

RUAN MALONI TEIXEIRA

**PANORAMA TRANSCRICIONAL SOB CONTROLE DA  
PROTEÍNA LIMYB, UM COMPONENTE DA VIA DE DEFESA  
ANTIVIRAL MEDIADA POR NIK1**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Aplicada, para obtenção do título de *Magister Scientiae*.

VIÇOSA  
MINAS GERAIS- BRASIL  
2017

**Ficha catalográfica preparada pela Biblioteca Central da Universidade  
Federal de Viçosa - Câmpus Viçosa**

T

O48p  
2017  
Teixeira, Ruan Maloni, 1990-  
Panorama transcricional sob controle da proteína LIMYB, um  
componente da via de defesa antiviral mediada por NIK1 / Ruan Maloni  
Teixeira. - Viçosa, MG, 2017.  
x, 88f. : il. (algumas color.) ; 29 cm.

Orientador: Elizabeth Pacheco Batista Fontes.  
Dissertação (mestrado) - Universidade Federal de Viçosa.  
Referências bibliográficas: f.81-88.

1. Biologia molecular. 2. Proteínas. 3. Plantas - Resistência a doenças e pragas. 4. Transcriptoma. I. Universidade Federal de Viçosa. Departamento de Bioquímica e Biologia Molecular. Programa de Pós-graduação em Bioquímica Aplicada. II. Título.

CDD 22 ed. 572.8

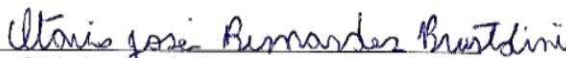
RUAN MALONI TEIXEIRA


**PANORAMA TRANSCRICIONAL SOB CONTROLE DA  
PROTEÍNA LIMYB, UM COMPONENTE DA VIA DE DEFESA  
ANTIVIRAL MEDIADA POR NIK1**

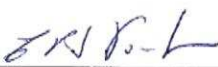
Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Aplicada, para obtenção do título de *Magister Scientiae*.

APROVADA: 20 de julho de 2017.

  
\_\_\_\_\_  
Humberto Josué de Oliveira Ramos

  
\_\_\_\_\_  
Otávio José Bernardes Brustolini  
(Coorientador)

  
\_\_\_\_\_  
Iara Pinheiro Calil

  
\_\_\_\_\_  
Elizabeth Pacheco Batista Fontes  
(Orientadora)

*Aos meus pais,*

*O Pai celeste e ao meu  
amado pai Sebastião, e as  
minhas mães, a Mãe fiel,  
mãe Maria e e minha amada  
mãe Eliane.*

*Dedico*

## **Agradecimentos**

Primeiramente agradeço a Deus, por me guiar até aqui e me dar forças para seguir em frente, apesar de todas as dificuldades que surgem pelo caminho, mostrando-me que sempre há mais luz do que trevas. Agradeço a Virgem Santíssima pela intercessão a mim sempre cedida.

Agradeço aos meus Pais, Sebastião e Eliane que sempre lutaram para que eu possa realizar meus sonhos, me dando educação, amor, apoio e preparo para os desafios no caminho. Sempre com muita humildade e fé.

Agradeço a minha orientadora Professora Elizabeth P.B. Fontes, pela contribuição para o meu crescimento profissional, me guiando por caminhos rumo a excelência. Agradeço também pela oportunidade de trabalho que me foi concedida e por toda compreensão nos momentos em que precisei.

Agradeço a Iara Calil por me guiar nos passos iniciais e por toda contribuição com este trabalho e pela amizade sincera que estabelecemos. Por se tornar um leão de profissionalidade em momentos que foram necessários para auxílio neste trabalho.

Ao Otávio e Cleysinho pela amizade e pelos papos de relatividade e espaços interdimensionais. E o meu muito obrigado pela condução das análises de bioinformática deste trabalho.

Agradeço ao professor Pedro Augusto e toda equipe do Laboratório de Biologia Molecular de Plantas, em especial aos companheiros Otto, Bruno, Luiz, Marco Aurélio, Larissa Miguel, Rafinha, Paola e Larissa Ávila.

Agradeço também aqueles que me mostraram que neste lugar estranho de Viçosa, onde cheguei, e pensei jamais ter amizade, acabei fazendo irmãos, tais como Virgílio, Danilo, Cleysinho, Dalton, Laura e Chris, fortes parceiros na vida.

Aos amigos de república George, Gustavo e Pablo que escutaram com toda paciência o relato dos dias de trabalho no laboratório e pela amizade.

Agradeço também a Analu Zanotti, pela parceria eterna construída em Alegre para o resto da vida, que veio me dar suporte em Viçosa.

A Universidade Federal de Viçosa pela oportunidade de crescimento em uma instituição de excelência.

Ao Núcleo de Análises de Biomoléculas (Nubiomol) pelo uso de equipamentos necessários a realização deste projeto.

A Fundação Arthur Bernardes (Funarbe) pelo apoio financeiro.

A coordenação de aperfeiçoamento de pessoal de nível superior (CAPES) pelo apoio financeiro.

## Sumário

<b>Resumo</b> .....	vii
<b>Abstract</b> .....	ix
<b>1. Introdução</b> .....	1
<b>2. Revisão bibliográfica</b> .....	5
2.1 Classificação e características estruturais de geminivirus.....	5
2.2 Tráfego do genoma viral e a interação com proteínas do hospedeiro...7	
2.3 A via sinalização de defesa contra begomovírus em plantas mediada pelo receptor imune NIK1.....	11
2.5 LIMYB associa-se a RPL10 no núcleo e suprime a tradução geral como mecanismo imune antiviral.....	13
2.6 LIMYB, uma proteína com domínios de ligação ao DNA Myb/SANT-like, pertencente à família de proteínas contendo domínios Myb.....	14
<b>3. Metodologia</b> .....	16
3.1 Obtenção plantas de <i>Arabidopsis thaliana</i> mutantes nulas para LIMYB complementadas com a proteína LIMYB fusionada a YFP.....	16
3.2 Caracterização de linhagens transformantes de <i>A. thaliana</i> para LIMYB quanto ao nível de acúmulo de transcritos e de proteínas.....	18
3.3 Sequenciamento de RNA total de linhagens transformantes de LIMYB e análise dos dados.....	18
3.4 Identificação de novos sítios de ligação ao genoma do fator de transcrição LIMYB em <i>A. thaliana</i> por meio de CHIP-seq.....	20
3.4.1 Coleta e fixação do material vegetal para Imunoprecipitação de Cromatina (ChIP).....	20
3.4.2 Lise nuclear e fragmentação da cromatina.....	21
3.4.3 Pré-clareamento e Imunoprecipitação.....	21

3.4.4 <i>Crosslink-reverso</i> e purificação do DNA imunoprecipitado.....	22
3.5 Análise dos novos sítios de ligação de LIMYB e identificação de cis- elementos e características conservadas.....	23
<b>4. Resultados e Discussão</b> .....	<b>25</b>
4.1 Caracterização de linhagens transformantes overexpressando LIMYB.....	25
4.2 Variação global da expressão gênica induzida pela expressão ectópica de LIMYB.....	27
4.3 Estudo da atividade transcricional de LIMYB em escala genômica....	33
4.4 Identificação de cis-elementos conservados e análise de colocalização da ligação do fator de transcrição por cobertura de RNA-seq e ChIP-se...	52
<b>5. Conclusão</b> .....	<b>58</b>
<b>6. Material Suplementar</b> .....	<b>60</b>
6.1 Figuras suplementares.....	60
6.2 Tabelas Suplementares.....	71
<b>7. Referências bibliográficas</b> .....	<b>81</b>

## Resumo

Teixeira, Ruan Maloni, M.Sc., Universidade Federal de Viçosa, julho de 2017. **Panorama transcricional sob controle da proteína LIMYB, um componente da via de defesa antiviral mediada por NIK1.** Orientadora: Elizabeth Pacheco Batista Fontes. Coorientadores: Pedro Augusto Braga dos Reis e Otávio José Bernardes Brustolini.

A proteína LIMYB (*L10-interacting MYB domain-containing protein*) foi recentemente identificada como um componente *downstream* da via de sinalização antiviral mediada pelo receptor NIK1 (*Nuclear shuttle protein-interacting kinase 1*). Diante da ativação de NIK1, LIMYB interage com RPL10 para reprimir a expressão de genes de proteínas ribossomais, levando a supressão global de tradução. Exceto pelos genes de proteínas ribossomais que são regulados por LIMYB, os demais alvos diretos de LIMYB em escala genômica não foram ainda determinados. Para determinar possíveis genes controlados por LIMYB e compreender sua função no desenvolvimento e nos processos coordenados pela regulação da tradução, foi avaliado o perfil transcricional de plantas overexpressando LIMYB, por RNA-seq e ChIP-seq. Nas análises por RNA-seq, foram identificados 2351 genes diferencialmente expressos, sendo que 1246 genes *down* regulados e 1105 genes *up* regulados, em plantas expressando ectopicamente LIMYB. Os estudos de ontologia gênica demonstraram que a categoria funcional mais enriquecida entre genes *down* regulados, de fato pertence aos constituintes ribossomais, enquanto que para genes *up* regulados foi notada uma categoria associada a resposta de níveis de nutrientes. Nas análises de ChIP-seq, o DNA associado à proteína foi fragmentado, imunoprecipitado e sequenciado, permitindo uma avaliação dos locais de ligação de LIMYB ao DNA nos respectivos genes sob controle da proteína. A análise integrativa de ChIP-seq e RNA-seq identificou 156 genes *down* regulados e destes, os constituintes ribossomais foram o grupo mais afetado, seguido por categorias relacionadas a metabolismo celular de aldeídos, fotossíntese e tradução. Nestas diversas categorias, foi possível identificar genes reprimidos por LIMYB e que também participam de processos de infecção viral, expandindo a atividade antiviral de LIMYB. Entre os genes reprimidos por LIMYB, foram identificados componentes chaves do aparato fotossintético

sugerindo que LIMYB pode funcionar como um molecular link que sincroniza supressão de tradução e inibição de fotossíntese. Já no conjunto de genes *up* regulados, a categoria relacionada a transporte de ânions foi a mais enriquecida, seguida por categorias relacionadas a escassez de determinados compostos como fosfato e na resposta a estímulos internos e externos. A partir dos genes diferencialmente expressos para compreender sua função no desenvolvimento e nos processos coordenados a regulação da tradução na intercessão dos experimentos de ChIP-seq e RNA-seq, foram encontrados em maior significância 3 sequencias consensos, CAAAC, AAGAAA e ATATATA. A interação de LIMYB com promotores, reprimindo e ativando genes alvo, foi elucidada neste trabalho, através de análise de cobertura (co-localização RNA-seq/ChIP-seq), representando a proximidade dos picos com os respectivos genes, e pelo mapeamento das sequencias consensos nos promotores escolhidos. Estes resultados permitem uma melhor compreensão da função de LIMYB como um regulador capaz de gerenciar processos celulares em função da inibição da tradução global, durante o desenvolvimento e na defesa antiviral.

## Abstract

Teixeira, Ruan Maloni, M.Sc., Universidade Federal de Viçosa, July, 2017. **Transcriptional landscape under control of LIMYB protein, a component of the antiviral defense pathway mediated by NIK1.** Adviser: Elizabeth Pacheco Batista Fontes. Co-advisers: Pedro Augusto Braga dos Reis and Otávio José Bernardes Brustolini.

LIMYB (L10-interacting MYB domain-containing protein) was recently identified as a downstream component of the NIK1 (Nuclear shuttle protein-interacting kinase 1)-mediated antiviral signaling. Upon NIK1 activation, LIMYB interacts with RPL10 to repress the expression of ribosomal protein genes, leading to the global translation suppression. Apart from the LIMYB-mediated regulation of RP genes, the genome-wide direct target genes of LIMYB are unknown. To identify genes controlled by LIMYB towards understanding LIMYB function in development and in processes coordinated by translation regulation, we carried out a transcriptional profiling of LIMYB-overexpressing plants by RNA-seq and ChIP-seq. RNA-seq analyses revealed 1246 down-regulated genes and 1105 up-regulated genes in LIMYB-overexpressing plants. Gene enrichment analysis, based on gene ontology, demonstrated that the ribosomal constituent category was the most-enriched down-regulated functional term, whereas the most-enriched up-regulated term was associated with response to nutrient levels. For the chip-seq analysis, the DNA associated with the protein was fragmented, immunoprecipitated and sequencing, allowing the location of LYMIB association with DNA on the respective genes under control of LIMYB. The integrative analysis between ChIP-seq and RNA-seq identified 156 down-regulated genes, from which the ribosomal constituents were the group more affected, followed by categories related to aldehyde metabolism, photosynthesis and translation. Within these categories, we identified LIMYB-repressed genes unrelated to translation that also participate in viral infections, expanding the antiviral activity of LIMYB. Among the LIMYB-repressed genes, we also identified key components of the photosynthetic apparatus, suggesting that LIMYB may function as a molecular link, which synchronizes translation suppression and photosynthesis inhibition. Within the up-regulated genes, the anion transport term was the most enriched category, followed by phoshate deprivation-related

category and response to internal and external stimuli. By analyzing the set of differentially expressed genes also found in ChIP-seq, the three consensus sequences CAAAC, AAGAAA and ATATATA were identified as the top-scoring motifs. The interaction of LIMYB with promoters, repressing and activating target genes, was elucidated by the coverage analysis (co-localization RNA-seq/ChIP-seq), mapping peaks from ChIP-seq and heats from RNA-seq within the gene locus. These results allowed a better understanding of the LIMYB function as a regulator of cell processes related to suppression of global translation, during development and antiviral defense.

## 1. Introdução

A busca de alternativas para solucionar complicações de estresses ambientais é considerada um grande desafio para o setor agrícola. A produtividade e o desenvolvimento de plantas têm interferência de inúmeros fatores naturais, bióticos e abióticos, como salinidade, seca e invasões de patógenos, que podem causar perdas consideráveis de safras em todo o mundo. Sob estresses, as plantas se tornam incapazes de atingir sua potencialidade genética, o que interfere na capacidade de culturas agrônomicas atingirem sua produtividade total (Cramer *et al.*, 2011). Para contornar estes fatores de estresse, evolutivamente as plantas têm desenvolvido diferentes estratégias de defesa contra ataques bióticos, visando combater o amplo espectro de patógenos ao qual estão submetidas, como vírus, bactérias, fungos e pequenos invertebrados que podem atuar como vetores na transmissão de agentes infectantes (Hull, 2014).

Na defesa contra agentes infecciosos, as plantas desenvolveram um sofisticado sistema imunológico em camadas. A primeira linha de defesa é denominada como PTI (*PAMP-triggered immunity*), ativada pela percepção de PAMPs (*pathogen-associated molecular pattern*), similar ao sistema imunológico de mamíferos, porém mais amplo e menos flexível (Jones & Dangl, 2006; Ausubel, 2005). Este reconhecimento ocorre por meio de PRRs (*pattern recognition receptors*) na membrana plasmática, que atuam na indução de sinalização de defesa (Ben Khaled *et al.* 2015; Boller & Felix, 2009; Monaghan & Zipfel, 2012; Dangl *et al.*, 2013). Muitos agentes virulentos se tornaram capazes de secretar efetores para interferir com a PTI. As plantas por sua vez evoluíram, produzindo proteínas de resistência, designadas NBS-LRR (*nucleotide binding site-leucine rich repeats*) para reconhecer efetores de virulência, ativando uma segunda linha de defesa, denominada ETI (*Effector triggered immunity*) acionando um tipo de resposta hipersensitiva e morte celular programada (Dodds & Rathejen, 2010; Stotz *et al* 2014).

No processo coevolutivo entre plantas e patógenos, os vírus da família *Geminiviridae* se destacam tanto na diversidade de hospedeiros, como na capacidade de superar as estratégias de defesa (Scholthof *et al.*, 2011; Hanley-Bowdoin *et al.*, 2013). O progresso evolutivo dos *geminivirus* abrange múltiplos

fatores cruciais e resultantes de novas adaptações que frequentemente conduzem a grandes prejuízos para economia (Mansoor *et al.*, 2006; Navas-Castillo *et al.*, 2011; Lefeuvre & Moriones, 2015). Neste contexto, melhores esclarecimentos dos mecanismos bioquímicos de virulência e de defesa relacionado à infecção por *geminivirus* são extremamente importantes para o desenvolvimento de cultivares geneticamente resistentes a este patógeno.

Um dos mecanismos primários de estratégia de defesa antiviral se dá pela via de silenciamento de RNA (Deleris *et al.*, 2006). Em plantas infectadas com vírus, proteínas *Dicer-like* (DCL) podem detectar e efetuar o processamento do RNA viral em RNAs curtos de interferência (siRNAs), seguido pelo auxílio de proteínas argonautas (AGO), formando complexos de silenciamento que reprimem genes virais (Blevins *et al.*, 2006; Carbonell *et al.*, 2012; Pooggin, 2013). No entanto, os vírus adquiriram fatores de virulência que suprimem componentes da maquinaria de silenciamento de RNA, inativando a resposta de defesa e aumentando a patogenicidade (Raja *et al.*, 2010).

Estudos realizados em *Arabidopsis thaliana* na busca de elementos que impedem a defesa contra vírus, deram início à caracterização de uma nova via de defesa antiviral, em que foi identificado a interação da proteína viral NSP (*Nuclear shuttle protein*) de *Begomovirus*, um gênero da família *Geminiviridae*, com a proteína NIK1 (*NSP-interacting kinase*), um receptor de membrana do tipo cinase (Mariano *et al.*, 2004; Fontes *et al.*, 2004). No lado citoplasmático da membrana celular, a ligação da proteína NSP ao *loop* de ativação de NIK, inibe sua atividade cinase, criando um ambiente mais favorável à replicação do vírus (Fontes *et al.*, 2004).

Foi evidenciado que NIK1 se autofosforila, no resíduo de treonina 474 para ativação do domínio cinase, e isto conduz a sua oligomerização, indicando que suas propriedades têm características comuns à função sinalizadora (Fontes *et al.*, 2004). Em adição, a inativação de NIK1 prejudica a resposta de sinalização antiviral imune. Quando NIK1 é ativada, ocorre a fosforilação de demais substratos e tem início a sinalização de defesa antiviral em que a proteína ribossomal (RP) RPL10 é fosforilada, e nesta forma passa a ser translocada para o núcleo (Rocha *et al.*, 2008; Carvalho, Santos, *et al.*, 2008).

RPL10 interage no núcleo com LIMYB (*L10-interacting MYB domain-containing protein*), uma proteína cuja a função não é completamente elucidada (Zorzatto *et al.*, 2015). No trabalho realizado por Zorzatto (2015) e colaboradores, por meio de experimento de imunoprecipitação de cromatina (ChIP), foi notado que LIMYB age como um fator de transcrição, capaz de se ligar a um promotor RP (RPL18) *in vivo*. Foi demonstrado que LIMYB se liga em uma região do promotor de RPL18 de 150 pares de base (pb), resultante da amplificação a partir de DNA precipitado de tecidos de plantas superexpressando LIMYB, mas não de plantas do tipo selvagem. Estes dados implicam que LIMYB pode funcionar como uma proteína de ligação a DNA que associa-se com os promotores de genes de proteínas ribossomais *in vivo*. A ativação da via antiviral mediada por NIK1, leva à repressão do processo de tradução global, devido à redução da expressão de genes ribossomais, mediado pela associação de LIMYB com DNA, um mecanismo de defesa ainda não esclarecido em plantas (Zorzatto *et al.*, 2015; Brustolini *et al.*, 2015).

A proteína LIMYB tem em sua composição dois domínios Myb/SANT-like, e se encaixa na família de fatores de transcrição Myb, mais especificamente no grupo 2R-Myb ou R2R3-Myb, um grupo majoritário de proteínas Myb em *A. thaliana*, composto por 126 membros. Esta família de proteínas está envolvida em processos celulares diversos, como metabolismo primário e secundário, diferenciação celular, no desenvolvimento e também em respostas a estresses bióticos e abióticos (Dubos *et al.*, 2010). Proteínas Myb/SANT podem interagir com o DNA, agindo como fatores de transcrição ou, no caso de domínios SANT, realizar controle do remodelamento da cromatina por associação com histonas (Bartholomew, 2014).

Dado a relevância de LIMYB na via de defesa antiviral mediada por NIK1, o esclarecimento de sua função, torna-se de extrema importância, visto que tem aumentado nos últimos anos a incidência e o agravamento de doenças por *Geminivirus*, (Mansoor *et al.*, 2006; Navas-Castillo *et al.*, 2011). Os complexos fatores epidemiológicos associados a surtos de doenças deste vírus, como dinâmica de migração de vetores, sua acelerada evolução e expansões imprevisíveis de hospedeiros, tornaram muito difícil desenvolver estratégias eficazes de gestão de doenças de longo e curto prazo (Loebenstein & Katis,

2014). Este projeto tem como finalidade identificar locais de interação no genoma ao qual esta proteína se liga, para compreensão do funcionamento da resposta de defesa, pela identificação dos genes que são afetados neste processo de regulação.

## 2. Revisão bibliográfica

### 2.1 Classificação e características estruturais de geminivirus

Os geminivirus (vírus da família *Geminiviridae*) infectam uma ampla gama de culturas agrônômicas relevantes por todo o mundo causando doenças como o mosaico amarelo ou mosqueado amarelo, enrugamento das folhas, baixa estatura, raias/listras e gerando grandes perdas, que variam de 20-100% da produtividade (e.g.: *Maize streak virus* - MSV, *Cassava mosaic virus* - CMV, *Cotton leaf curl virus* - CLCuV, *Tomato yellow leaf curl virus* - TYLCV) (Gibbs & Ohshima, 2010; Scholthof *et al.*, 2011). A incidência e a gravidade das doenças causadas por geminivirus têm aumentado nos últimos 25 anos (Mansoor *et al.*, 2006; Navas-Castillo *et al.*, 2011), devido a fatores como a capacidade de recombinação de diferentes geminivirus coinfectando uma planta, a expansão das fronteiras agrícolas, o transporte de material e de vetores virais infectados (Gray & Banerjee, 1999; Lefeuvre & Moriones, 2015).

A família *Geminiviridae* inclui 369 vírus de plantas divididos em 9 gêneros, tendo evoluído provavelmente de um plasmídeo de fitoplasma, ("ICTV": [http://www.ictvonline.org/virusTaxonomy.asp?taxnode\\_id=20152517](http://www.ictvonline.org/virusTaxonomy.asp?taxnode_id=20152517); acessado em maio de 2017), um tipo de bactéria parasita intracelular obrigatória de plantas (Krupovic *et al.*, 2009). Atualmente, o Comitê Internacional de Taxonomia de Vírus (ICTV) os classifica em nove gêneros (*Becurtovirus*, *Begomovirus*, *Curtovirus*, *Eragrovirus*, *Mastrevirus*, *Topocovirus*, *Turncurtovirus*, *Capulavirus* e *Grablovirus*), baseado na organização de seu genoma (monopartidos ou bipartidos), semelhança na sequência, tipos de insetos vetores e hospedeiros (Hanley-Bowdoin *et al.*, 2013 e "ICTV" - [www.ictvonline.gov](http://www.ictvonline.gov)).

Os geminivirus são encapsulados em capsídeos que medem aproximadamente 18x30 nm, sendo compostos de duas partículas icosaédricas geminadas imperfeitas (Zhang *et al.*, 2001). Os geminivirus possuem pequeno DNA fita simples (ssDNA) circular, que variam de 2500-3100 nucleotídeos no caso dos vírus de componente genômico único, e de 2600-2800 nucleotídeos em vírus com genoma bipartido. Devido à capacidade codificante limitada de seus genomas, os geminivirus são altamente dependentes da maquinaria celular do hospedeiro para completar seu ciclo de vida. Ainda que altamente

dependentes das funções do hospedeiro, são capazes de causar alterações que acarretam na reprogramação do ciclo celular das células infectadas para induzir a replicação, tanto do genoma viral como do genoma da planta hospedeira (Hanley-Bowdoin *et al.*, 2013). As diversas alterações incluem mudanças no padrão de expressão gênica, tráfego molecular e interferências em vias de sinalização, para redirecionamento ou bloqueio das defesas do hospedeiro. Além disso, os geminivirus codificam proteínas que atuam como supressores do aparato proteico de silenciamento, que têm como alvo as vias de produção de pequenos RNAs de interferência (siRNA), o que por conseguinte gera alterações na produção de micro RNAs (miRNA) e metilação do DNA, levando a anormalidades no desenvolvimento da planta infectada (Hanley-Bowdoin *et al.*, 2013).

Três gêneros da família *Geminiviridae* são bem caracterizados, *Mastrevirus*, *Curtovirus* e *Begomovirus*. Os mastrevirus (32 espécies) são transmitidos por espécies de cigarrinhas (*Homoptera: cicadellidae*), possuem um único componente genômico e infectam tanto monocotiledôneas quanto dicotiledôneas. Os curtovirus (3 espécies) também são transmitidos por cigarrinhas (*Hemiptera: cicadellidae*), têm apenas um componente genômico e infectam dicotiledôneas. Já os begomovirus (322 espécies) constituem o maior gênero, transmitidos por moscas brancas (*Bemisia tabaci*), também a dicotiledôneas ("ICTV": <http://www.ictvonline.org/virusTaxonomy/>; acessado em maio de 2017). Podem ter genomas monopartidos ou bipartidos, em que os componentes genômicos são denominados DNA-A e DNA-B. A maioria dos begomovirus monopartidos permanecem associados com alfasatélites ou betasatélites (Rojas *et al.*, 2005).

O genoma bissegmentado de begomovirus é composto no DNA-A por seis ORFs: duas no sentido viral (V) (AV1 e AV2) e quatro no sentido complementar (AC1, AC2, AC3 e AC4), enquanto o DNA-B contém duas ORFs: uma no sentido viral (BV1) e outra em sentido complementar (BC1) (Fondong, 2013). A ORF AV1 codifica a proteína CP (*coat protein*), constituinte do capsídeo viral (22 unidades da CP), que confere a especificidade do vírus. Por sua vez, a ORF AV2 codifica uma proteína capaz de suprimir o silenciamento de RNA (Chowda-Reddy *et al.*, 2008) e estabelecer certo controle da iniciação da replicação viral (Rouhibakhsh

*et al.*, 2011). As ORFs AC1, AC2, AC3 e AC4 codificam as proteínas Rep (*replication-associated protein*), TrAP (*trans-activating protein*), REn (*replication enhancer protein*) e AC4, respectivamente. Rep e REn são proteínas associadas ao processo de replicação do genoma viral, sendo a proteína Rep crucial para replicação (Elmer *et al.*, 1988; Fontes *et al.*, 1992; E., P., B., Fontes *et al.*, 1994; E., P., Fontes *et al.*, 1994) e a proteína REn um amplificador do mecanismo (Settlage *et al.*, 2001). A proteína TrAP codifica um fator de transcrição dos genes de sentido viral AV1 e BV1 (Rojas *et al.*, 2005). Adicionalmente, TrAP atua em conjunto com a proteína AC4 na supressão da defesa mediada por RNAi (Vanitharani *et al.*, 2005; Wang *et al.*, 2005).

O DNA-B codifica duas proteínas: NSP (*nuclear shuttle protein*) (BV1) e MP (*movement protein*) (BC1), respectivamente, exigidas no tráfego intracelular (núcleo-citoplasma) e intercelular de componentes do genoma do vírus (Hanley-Bowdoin *et al.*, 1999; Settlage *et al.*, 2001; Rojas *et al.*, 2005). A atuação das proteínas NSP e MP, deste modo, está conexa ao processo sistêmico de infecção e ao desenvolvimento da sintomatologia (Elmer *et al.*, 1988). Além disso, a proteína NSP é capaz de agir como fator de virulência, suprimindo defesas inatas da planta (Fontes *et al.*, 2004; Carvalho, Santos, *et al.*, 2008; Santos *et al.*, 2009). Os genomas de geminivirus possuem unidades transcricionais divergentes, sendo que a região 5' intergênica contém origens de replicação círculo rolante, e de dois promotores de RNA polimerase II (Hanley-Bowdoin *et al.*, 2013).

## 2.2 Tráfego do genoma viral e a interação com proteínas do hospedeiro

A infecção por begomovirus tem início quando a mosca branca, vetor do vírus, alimenta-se da seiva transportada pelo floema de folhas saudáveis, transmitindo os virions para as células associadas a esta região (Ghanim, 2014). O ssDNA (single stranded DNA) quando liberado pelo virion torna-se dsDNA (double stranded DNA) no núcleo, utilizando a maquinaria celular do hospedeiro para síntese do DNA complementar (Saunders *et al.*, 1992). O DNA do vírus é transcrito pela RNA polimerase do hospedeiro, levando posteriormente a produção da proteína viral Rep, que inicia a replicação viral através de dois

modos, mecanismo círculo-rolante e recombinação dependente (Jeske *et al.*, 2001). Uma vez replicado, o DNA viral organiza-se na forma de um monômero circular ou multimeros de ssDNA (Jeske *et al.*, 2001), servindo como molde para síntese de novas fitas de ssDNA viral na fase inicial da infecção ou sendo inserido em novos capsídeos e transportado para fora do núcleo em fases mais tardias, quando as proteínas CP e NSP estão presentes (Jeske *et al.*, 2001).

A proteína NSP interage com proteínas histonas H3, possibilitando que o DNA viral movimente-se como um minicromossomo (Zhou *et al.*, 2011). NSP pode também realizar ligação e inibição da proteína acetiltransferase (AtNSI) de *A. thaliana*, permitindo um maior nível de compactação do DNA viral em associação com histonas H3 (Carvalho & Lazarowitz, 2004; Carvalho *et al.*, 2006). Outra interação observada de NSP é com a proteína com NIG (*NSP-interacting GTPase*) (AT4G13350), uma GTPase presente em *A. thaliana* associada a porção externa do envelope nuclear, podendo facilitar o movimento do DNA viral para o citosol, possivelmente pelo poro nuclear (Carvalho *et al.*, 2008). O complexo NSP-DNA então se move para periferia celular através da interação com MP (Sanderfoot & Lazarowitz, 1995). O DNA viral pode ser transferido para a proteína MP por meio de um mecanismo envolvendo a hidrólise de GTP por NIG. Alternativamente, NIG poderia facilitar a interação de MP com o complexo NSP-DNA, que se move através de plasmodesmas (Carvalho *et al.*, 2008). Foi observado que a co-expressão transiente de NIG e NSP em folhas de tabaco resulta em um redirecionamento da proteína viral do núcleo para o citoplasma e que plantas transgênicas superexpressando NIG apresentam suscetibilidade aumentada à infecção por geminivírus, como esperado para um regulador positivo do ciclo de infecção (Carvalho *et al.*, 2008; Carvalho *et al.*, 2008). No estágio final da infecção, o vírus atinge o floema da planta e pode infectar um novo inseto vetor (Martin *et al.*, 2011).

### 2.3 A via sinalização de defesa contra begomovírus em plantas mediada pelo receptor imune NIK1

O sucesso de uma infecção viral em plantas depende em grande parte dos casos, da inibição da PTI por meio dos receptores imunes transmembrana do hospedeiro (Zipfel, 2009). Dentre os grupos de receptores imunes de plantas estão as proteínas RLKs (*Receptor-like kinase*), que além de participar em vias de sinalização de crescimento e desenvolvimento também atuam no reconhecimento de patógenos e na resposta de defesa (Osakabe *et al.*, 2013). As mais bem caracterizadas RLKs envolvidas na resposta à infecção por geminivírus, são RLKs com repetições LRR (*leucine rich repeat*), designadas NIK1 (*NSP-interacting Kinase*) (Fontes *et al.*, 2004). A primeira NIK a ser isolada foi identificada por sua capacidade de interagir com NSP de geminivírus *que infectam* tomateiros, sendo designada LeNIK1 (*Lycopersicum esculentum NIK*) (Mariano *et al.*, 2004). Posteriormente, essa interação foi também encontrada em soja (*Glicine max*) e *Arabidopsis thaliana*, sendo os receptores designados GmNIK1 e AtNIK1, respectivamente (Mariano *et al.*, 2004; Fontes *et al.*, 2004). Existem três homólogos de NIK no genoma de *A. thaliana*, designados NIK1 (AT5G16000), NIK2 (AT3G25560) e NIK3 (AT1G60800), sendo classificados na subfamília LRR-II-RLK, que possui 14 proteínas caracterizadas por um domínio extracelular com cinco repetições LRR em bloco contínuo, sendo quatro repetições completas (24 resíduos de aminoácido) e uma repetição incompleta (16 resíduos de aminoácido) (Santos *et al.*, 2010).

As proteínas NIK são receptores de membrana com atividade de serina/treonina cinase, as quais exibem capacidade de autofosforilação e de fosforilar substratos, atuando em vias de sinalização antiviral. Entretanto a função precisa de NIK1 na ativação da propagação do sinal antiviral, atuando como um co-receptor ou como PRR (*Pattern recognition receptor*), não foi ainda elucidada (Fontes *et al.*, 2004; Carvalho, Santos, *et al.*, 2008; Santos *et al.*, 2009).

A interação de NSP com o domínio cinase de NIK1 inibe a autofosforilação, que é essencial para atividade cinase dessa proteína (Fontes *et al.*, 2004). Foi observado que plantas de *A. thaliana* nocautes para NIK1 e NIK3 exibem

aumento na susceptibilidade à infecção por begomovirus (Fontes *et al.*, 2004), ao passo que a superexpressão de AtNIK1 em tabaco infectado com begomovirus é capaz de superar o efeito inibitório de NSP sobre a atividade de NIK1. Além disso, a superexpressão de AtNIK1 atenua os sintomas gerados pelo vírus em tomate (Carvalho *et al.*, 2008). Contudo, a efetividade da via de defesa antiviral mediada por NIK1 é limitada, uma vez que sua ativação depende do estabelecimento de infecção e a proteína viral NSP funciona como um supressor da via (Carvalho, Fontenelle, *et al.*, 2008; Santos *et al.*, 2010; Zorzatto *et al.*, 2015).

A interação entre NSP e o domínio cinase de NIK1 foi mapeada, demonstrando que NSP se liga a um fragmento de 80 aminoácidos que contém uma alça de ativação (*A-loop*) e o sítio ativo de NIK1 (Fontes *et al.*, 2004). Como já descrito na literatura, a fosforilação de resíduos de aminoácido no *A-loop* constitui um importante mecanismo regulatório da atividade de tirosina e serina/treonina cinases (Nolen *et al.*, 2004). O alinhamento das sequências dos domínios cinases da subfamília LRR-II-RLK com o receptor IRAK de humanos, revelou alto grau de conservação entre os resíduos de treonina 468, 469 e 474, contidos no *A-loop* de NIK1. O alinhamento da treonina 474 do *A-loop* de NIK1 com o resíduo de treonina 468 da proteína SERK de *A. thaliana*, essencial para atividade dessa enzima, destacou a importância específica desse resíduo para atividade de NIK1, o que foi confirmado pela perda de atividade da cinase com a substituição da Thr-474 por um resíduo de alanina (Shah *et al.*, 2001; Santos *et al.*, 2009). Um mutante de NIK1, em que a treonina 474 foi substituída por um resíduo de ácido aspártico, foi então obtido, a fim de introduzir uma carga negativa permanente nesta posição, de modo a mimetizar o estado ativo da enzima. O mutante AtNIK1 (NIK1-T474D) apresentou hiperatividade de cinase, quando comparado aos níveis normais de NIK1, e uma redução do efeito inibitório de NSP sobre a atividade da enzima (Santos *et al.*, 2009). Ao serem desafiadas com a infecção por *Begomovirus*, as plantas NIK1-T474D mostraram-se significativamente mais resistentes quando comparadas a plantas *wild type* (WT) (Zorzatto *et al.*, 2015; Brustolini *et al.*, 2015).

## 2.4 Proteína ribossomal RPL10A, um componente da via de sinalização antiviral mediada por NIK1

A proteína RPL10A (*QM-like*) (AT1G14320) foi identificada como efetor *downstream* da via mediada por NIK1, por meio de um *screening* de duplo-híbrido em leveduras utilizando NIK1 como isca (Rocha *et al.*, 2008). Ainda que a proteína rpl18 também tenha exibido associação com NIK1, somente RPL10A foi fosforilada por NIK1 em um ensaio *in vitro* (Rocha *et al.*, 2008). Adicionalmente, plantas nocautes de RPL10A exibiram aumento na susceptibilidade à infecção por *begomovirus*, recapitulando o fenótipo observado para plantas *nik1* e *nik3* (Carvalho, Santos, *et al.*, 2008). A fosforilação de RPL10A por NIK1 leva a uma mudança no padrão de localização da proteína ribossomal, sendo translocada do citoplasma para o núcleo, onde RPL10A deve gerar uma resposta de defesa antiviral (Carvalho, Santos, *et al.*, 2008). A inibição da atividade cinase de NIK1 pela proteína viral NSP inibe a fosforilação e translocação de RPL10A para o núcleo, fenótipo exibido por plantas mutantes *nik1*, que constitui uma situação mais favorável à infecção por *begomovirus* (Carvalho, Santos, *et al.*, 2008; Santos *et al.*, 2009; Santos *et al.*, 2010).

As proteínas ribossomais (PRs) são um grupo de proteínas altamente conservadas, envolvidas na montagem e funções do ribossomo (Korobeinikova *et al.*, 2012; Ban *et al.*, 2014). A proteína RPL10 atua especificamente na junção das subunidades 40S e 60S para formar o ribossomo funcional 80S (Eisinger *et al.*, 1997; Loftus *et al.*, 1997). Foi demonstrado em leveduras, que mutações em RPL10 são deletérias, já que a mesma organiza a união do sítio da aminoacil-tRNA e também é um pré-requisito a sua incorporação na subunidade 60S, para a união das subunidades e o início da tradução (Hofer *et al.*, 2007). Similarmente, a interrupção da expressão de RPL10A por inserção de T-DNA causa letalidade no gametófito feminino (Imai *et al.*, 2008). Em função de seu alto grau de conservação e de sua presença em todos supergrupos eucarióticos (Olvera & Wool, 1996; Ban *et al.*, 2014), a proteína RPL10A é comumente utilizada em construções de árvores filogenéticas e estudos de relações entre grupos distantes (e.g.: animais, plantas e fungos) (Matte-Tailliez *et al.*, 2002; Jain, 2011; Bachvaroff *et al.*, 2014).

No genoma de *A. thaliana* há três homólogos de RPL10, designados RPL10A (AT1G14320), RPL10B (AT1G26910) e RPL10C (AT1G66580). Experimentos de coimunoprecipitação indicam que as proteínas RPL10s em *Arabidopsis* são componentes do ribossomo 80S (Ferreyra *et al.*, 2010). A proteína RPL10A é relatada como supressor semi-dominante do mutante *acl5*, revertendo o fenótipo de nanismo observado no mutante (Imai *et al.*, 2008). Em adição, os genes RPL10s são diferencialmente regulados por radiação UV-B. Enquanto que o gene RPL10A não é regulado por UV-B, RPL10B é regulado negativamente e RPL10C é regulado positivamente, em todos os tecidos estudados. Há ainda evidências de funções não redundantes quanto a expressão em órgãos reprodutivos, em que RPL10A e RPL10B são expressos tanto nos órgãos reprodutivos masculino e feminino, ao passo que RPL10C é expresso somente no pólen (Ferreyra *et al.*, 2013). As três RPL10s exibem localização predominantemente nuclear, mas somente RPL10B tem sua concentração nuclear aumentada no tratamento com radiação UV-B (Ferreyra *et al.*, 2013). Dessa maneira, ainda que os três genes AtRPL10s sejam capazes de complementar um mutante nocaute *l10* de levedura, supõe-se que os genes RPL10 não sejam funcionalmente redundantes (Ferreyra *et al.*, 2010; Ferreyra *et al.*, 2013).

A proteína RPL10A é homóloga a proteína QM, primeiramente identificada em humanos (Dowdy *et al.*, 1991) e associada a diversas funções além de sua participação na montagem do ribossomo. Estas funções extraribossomal incluem regulação do crescimento, diferenciação celular e apoptose, muitas vezes atuando como fator de transcrição (Montecclaro & Vogt, 1993; Marty *et al.*, 1993; Imafuku *et al.*, 1999; Green *et al.*, 2000; Lillico *et al.*, 2002). A participação de QM na resposta imune é também extensivamente relatada na literatura (Wen *et al.*, 2005; Oh *et al.*, 2010; Chen *et al.*, 2012; Liu *et al.*, 2014). Adicionalmente, a proteína RPL10 é implicada no controle multifuncional da tradução em leveduras (Eisinger *et al.*, 1997; Gadai *et al.*, 2001). Nesse contexto, a demonstração de que a perda da função de RPL10 aumenta significativamente a susceptibilidade à infecção por *Begomovirus* (Rocha *et al.*, 2008) levou à conclusão de que RPL10A participa da via de resposta a begomovirus mediada por NIK1, atuando no núcleo na indução de uma resposta imune (Carvalho, Santos, *et al.*, 2008;

Zorzatto *et al.*, 2015). Recentemente, foi demonstrado que RPL10A interage no núcleo com a proteína LIMYB (*L10-interacting Myb domain-containing protein*) para suprimir a expressão de genes de proteínas ribossomais, levando à supressão de tradução global nas células vegetais, como um mecanismo de imunidade antiviral em plantas.

## 2.5 LIMYB associa-se a RPL10 no núcleo e suprime a tradução geral como mecanismo imune antiviral

A interação de RPL10A com uma proteína de função desconhecida, designada LIMYB (*L10 Interacting Myb-containing*), foi recentemente identificada por meio de um *screening* de uma biblioteca de cDNAs de *A. thaliana* pelo sistema dupla-híbrido de leveduras. A interação LIMYB-RPL10A também foi demonstrada *in vivo* por meio de coimunoprecipitação e *BiFC*, ocorrendo no núcleo e, com mais intensidade, no nucléolo das células (Zorzatto *et al.*, 2015). A proteína LIMYB exibe localização nuclear, com exclusão do nucléolo, exibindo contudo, distribuição por todo nucleoplasma quando coexpressa com RPL10A. A localização de RPL10A é também alterada de citoplasmática para nuclear na presença de LIMYB (Zorzatto *et al.*, 2015). LIMYB é expressa ubiquamente em todos os tecidos de plântulas de *A. thaliana*, sendo, no entanto, mais expressa no caule de plantas adultas (Zorzatto *et al.*, 2015). O crescimento de plantas nocautes *limyb* mostrou-se, aparentemente, retardado e reduzido, havendo porém a necessidade de ensaios de complementação de função para uma conclusão definitiva sobre o fenótipo (Zorzatto *et al.*, 2015). Recapitulando o fenótipo observado para plantas superexpressando NIK1-T474D e RPL10A, plantas transformadas com LIMYB sob o promotor 35S exibiram aumento na resistência à infecção por *Begomovirus*, ao passo que plantas *limyb* mostraram-se mais suscetíveis a infecção (Zorzatto *et al.*, 2015).

2.6 LIMYB, uma proteína com domínios de ligação ao DNA Myb/SANT-*like*, pertencente à família de proteínas contendo domínios Myb.

A proteína LIMYB é codificada pelo *locus* AT5G05800 e é constituída de 449 aminoácidos, com massa molecular de 52690 Da. Possui um variante de *splicing* com a mesma sequência de aminoácidos. Foram identificados em sua região N-terminal dois domínios Myb/SANT-*like* (Myb\_DNA-bind\_3: IPR024752 ou PF12776) (Finn *et al.*, 2014), constituídos de repetições em tandem de três alfa-hélices, organizadas em motivo hélice-volta-hélice (Zorzatto *et al.*, 2015). A família de fatores de transcrição Myb é classificada em 4 diferentes classes de acordo com o número de repetições adjacentes do motivo hélice-volta-hélice, que variam de 1 a 4, sendo designadas R1, R2, R3 e R4 (Dubos *et al.*, 2010). A comparação dos domínios presentes em LIMYB a outros domínios Myb de proteínas de *A. thaliana*, agrupou LIMYB juntamente a fatores transcricionais do grupo 2R-Myb ou R2R3-Myb (Zorzatto *et al.*, 2015), o maior grupo de proteínas Myb em *Arabidopsis thaliana*, com 126 membros (Dubos *et al.*, 2010).

O domínio SANT (acrônimo de '*switching-defective protein 3* (Swi3), *adaptor 2* (Ada2), *nuclear receptor co-repressor* (N-CoR), *transcription factor* (TF)III B')

) foi inicialmente identificado como um motivo ácido de 50 aminoácidos presente em co-repressores de receptores nucleares, sendo posteriormente encontrado em subunidades de diversos complexos remodeladores de cromatina (Aasland *et al.*, 1996). Análises de sequência e estruturais por meio de difração de raio-X revelaram a similaridade do domínio SANT com o domínio DBD (*DNA-binding domain*) de proteínas Myb-*like* (Aasland *et al.*, 1996), o que somado à dificuldade de diferenciar os domínios levou ao agrupamento dos domínios sob o nome de Myb/SANT (Boyer *et al.*, 2002). Contudo, o potencial de superfície ácido de alguns domínios SANT (pI = 4.6), em contraste com o potencial observado para domínios Myb DBD (pI = ~10), levou à proposição de divergência funcional entre os domínios (Boyer *et al.*, 2004). De fato, apesar de ser considerado que todas proteínas que contém domínios Myb/SANT ligam-se a caudas de histona e têm papel no remodelamento da cromatina, somente algumas proteínas SANT tem atividade enzimática e são capazes de atuar diretamente nesse processo (Bartholomew, 2014), características ausentes em

proteínas Myb (George & Ness, 2014). O próprio potencial de superfície é um parâmetro indicativo para diferenciação de domínios Myb/SANT capazes de ligação a caudas de histona ou de ligação a DNA, de maneira que potenciais positivos (domínios básicos) favorecem ligação ao DNA, ao passo que potenciais negativos (domínios ácidos) favorecem ligação às caudas positivamente carregadas das histonas (Boyer *et al.*, 2004). A análise dos domínios Myb/SANT de LIMYB quanto a seu pI revela que os domínios diferem, de modo que o domínio na posição 6-100 exibe caráter levemente ácido (pI = pH 6.34), ao passo que o domínio na posição 169-263 exibe caráter básico (pI = pH 8.83), mais próximo do observado para o domínio consenso Myb\_DNA-bind\_3, que exibe pI = pH 9.72 (Finn *et al.*, 2014).

Os domínios Myb/SANT exibem participação em diversas funções celulares, destacando-se sua presença em proteínas ligadas à progressão do ciclo celular, diferenciação e desenvolvimento. Como exemplo, a expressão ectópica do gene *Lefsm1* (*fruit SANT/Myb-like 1*) de tomate (*Solanum lycopersicum*), que contém um domínio Myb/SANT, resulta em alterações no desenvolvimento, como redução do crescimento e redução de dominância apical em tomate e *A. thaliana* (Barg *et al.*, 2005). Em camundongos, o gene UTF1 (*Undifferentiated embryonic cell transcription factor 1*) codifica uma proteína que contém um domínio Myb/SANT na porção N-terminal. A proteína UTF1 exibe localização nuclear, com exclusão do nucléolo, apresentando forte associação à cromatina. No núcleo, UTF1 age como repressor transcricional e está envolvida na diferenciação de células tronco embrionárias. O domínio Myb/SANT demonstrou-se essencial para a localização e função de UTF1 (Van Den Boom *et al.*, 2007). No fungo *Cochliobolus carbonum*, um gene designado BZcon1 (*Bipolaris zeicola conidiation*) contém dois domínios Myb/SANT, sendo essencial ao processo diferenciação de conídios e esporulação do fungo (Zhang *et al.*, 2014). Os genes FLP (*Four Lips*) e AtMYB88 de *Arabidopsis* codificam fatores transcricionais que contém, ambos, dois domínios Myb em suas sequências. FLP e AtMYB88 aparentemente agem no controle do ciclo celular e na diferenciação por meio da regulação da expressão de alvos como as CDKs (*cyclin-dependent kinases*) (Xie *et al.*, 2010).

Alternativamente, os domínios Myb/SANT podem ser encontrados em proteínas relacionadas às outras funções celulares. As proteínas TRF1 e TRF2 (*Telomeric repeat factors 1 e 2*), componentes do complexo *shelterin*, responsável por prevenir a falsa sinalização de dano nos telômeros de mamíferos, contém um Myb/SANT DBD que lhes confere capacidade de ligação ao DNA (Konishi & De Lange, 2008; Baker *et al.*, 2009; Choi *et al.*, 2011). A proteína Reb1 de *Schizosaccharomyces pombe*, associada ao processo de terminação de transcrição, contém duas cópias do domínio Myb/SANT em sua sequência. Os dois domínios Myb/SANT observados na proteína Reb1 são necessários e suficientes à sua capacidade de ligação ao DNA (Biswas & Bastia, 2008). Similarmente à proteína Reb1, a proteína Rtf1 de *S. pombe* possui dois domínios Myb/SANT em sua sequência e está envolvida na terminação do processo de transcrição (Eydmann *et al.*, 2008). Embora tenha sido comprovada a participação de LIMYB como supressor da expressão de genes de proteínas ribossomais, a sua atividade transcricional em escala genômica ainda não foi determinada, o que constitui o principal objetivo dessa investigação.

### 3. Metodologia

#### 3.1 Obtenção de plantas *Arabidopsis thaliana* mutantes nulas para LIMYB complementadas com a proteína LIMYB fusionada a YFP

Para que a caracterização da função de fator de transcrição da proteína LIMYB (AT5G05800) na via de defesa antiviral mediada por NIK1, identificada em estudo anterior (Zorzatto *et al.*, 2015), fosse aprofundada, foram executados de ensaios *in vivo* com plantas, para que LIMYB fosse facilmente identificada. Com este objetivo, foram obtidas do Salk Institute for Biological Studies as linhagens de *Arabidopsis thaliana* *limyb* Salk\_032054 e Salk\_082995. Nestas linhagens, expressão da proteína LIMYB foi interrompida através de inserção de T-DNA no 4º éxon do gene LIMYB e na região 5'UTR, para *limyb-32* (Salk\_032053) e *limyb-82* (Salk\_082995), respectivamente. Foi demonstrado (Zorzatto *et al.*, 2015) que o mutante *limyb-32* produz somente uma forma

truncada da proteína LIMYB e que o transcrito apresenta baixo nível de acúmulo, sendo assim, selecionado para utilização em ensaios de complementação.

A estirpe GV3101 de *Agrobacterium tumefaciens* foi transformada com o clone YFP-At5G05800, por eletroporação. O diagnóstico das colônias transformadas com os clones derivados de pYFP foi realizado por PCR, utilizando os primers para o promotor 35S de *CaMV*. Plantas de *A. thaliana* limyb-32 foram germinadas em meio MS ½ força e transferidas após 10 dias de desenvolvimento para vasos com substrato vegetal em câmara de crescimento, sob dias longos (16h/8h), temperatura de 22-24°C e 60% de umidade. Após o surgimento de botões florais, as plantas foram submetidas à transformação com as bactérias contendo a construção de interesse pelo método do mergulho floral. Posteriormente, as sementes das plantas submetidas à transformação foram coletadas e semeadas em meio MS ½ força contendo higromicina 30 µg/mL, a fim de selecionar plantas que continham o transgene de interesse. Adicionalmente, as plantas selecionadas foram submetidas à confirmação da inserção do transgene no genoma por PCR do DNA genômico e quantificação da expressão do transgene por PCR em tempo real.

**Tabela 1. Oligonucleotídeos utilizados em reação de PCR.**

Sequência alvo	Nome	Sequência (5'-->3')
limyb- Salk_032054	ATNOC32FWD (SALK032054LP)	GTC CCA AAC CTT GTT CTG CTC
limyb- Salk_032054	ATNOC32RVS (SALK032054RP)	AAG GAA GAG GCG AAG AAG TTG
LIMYB	RTAT5G05800BFWD	AAC CAG GGA CAC ATT TTT CGA
LIMYB	RTAT5G05800BRVS	ATC GTA CAT GGC GCC TGT TT
LIMYB	At5G05800BFWD	AAA AAG CAG GCT TCA CAA TGA GGC CAA AAG CAG T
LIMYB	At5G05800BRVS	AGA AAG CTG GGT CCT ATG TTG TAG TAG GAA TAG G
ACTINA	ARAACTIN2FWD	TAC AAC CAG CTT CGT GTT GC
ACTINA	ARAACTIN2RVS	GAT TGA TCC TCC GAT CCA GA
T-DNA	LBB1.3	ATT TGC CGA TTT CGG AAC
FWD Promotor 35 S	MC36	TCC TTC GCA AGA CCC TTC CTC
PROL18	PROL18POS908FWD	GTT GAA AGC CTA ATG CCA CAT
PROL18	PROL18POS755RVS	CTC TGT TTC TCC TTC AAT GAC

### 3.2 Caracterização de linhagens transformantes de *A. thaliana* para LIMYB quanto ao nível de acúmulo de transcritos e de proteínas

Na finalidade de confirmar e quantificar a expressão da proteína LIMYB nas linhagens transformantes obtidas a partir do mutante Salk\_032054, foram empregadas as metodologias de PCR em tempo real e *Western Blotting*.

Plantas de *Arabidopsis* das linhagens independentes obtidas, *limyb-32* pYFP-LIMYB, que tiverem a inserção da construção contendo LIMYB confirmada por PCR, foram coletadas para extração de RNA total e proteínas totais. O RNA foi extraído pelo método do Trizol® e, após verificação de sua qualidade, utilizado para síntese de cDNA por transcriptase reversa M-MLV, de acordo com as instruções do fabricante (Invitrogen™). A PCR em tempo real foi realizada no equipamento ABI 7500 (Applied Biosystems™), assim como informa o manual do fabricante, utilizando o corante SYBR® Green PCR Master Mix (Applied Biosystems™) e primers específicos para os genes avaliados (Tabela 1). Os valores relativos de expressão foram quantificados pelo método  $2^{-\Delta Ct}$ .

As proteínas totais foram extraídas por maceração do tecido foliar em N2 líquido e ressuspensão em tampão ( $\text{Na}_2\text{HPO}_4$  100 mM pH 8,0, 150 mM NaCl, 5mM EDTA, 5mM EGTA, 1%, PMSF 1mM e 0,1% triton x-100). Após a extração, as proteínas foram separadas por eletroforese, transferidas para membranas e sondadas com anticorpo anti-GFP policlonal, e o kit WestDura® foi utilizado na visualização do resultado.

### 3.3 Sequenciamento de RNA total de linhagens transformantes de LIMYB e análise dos dados

O procedimento de RNA-seq foi realizado para identificar quais genes têm seus transcritos *down* ou *up* regulados por superexpressão de LIMYB, e discutir as relações destes efeitos com os mecanismos de defesa antiviral contra o processo de infecção por *geminivírus*.

O RNA total foi isolado a partir de maceração do tecido vegetal em nitrogênio líquido com reagente Trizol® (Invitrogen), sendo incubado à

temperatura ambiente, por 5 min, e adicionados 200  $\mu\text{L}$  de clorofórmio. A mistura foi homogeneizada agitando levemente, por 15 s, incubando-se à temperatura ambiente, por 3 min, e posteriormente submetida à centrifugação a 12.000 x g por 10 min, a 4°C. O sobrenadante foi transferido para um recipiente (microtubo) livre de RNase, e foram adicionados 500  $\mu\text{L}$  de isopropanol. Em sequência, as amostras foram incubadas à temperatura ambiente, por 10 min, e centrifugadas a 12.000 x g, por 10 min. O RNA precipitado foi lavado com etanol 75% (V/V) e centrifugado a 9.500 x g, por 10 min tratado com DNase. Em seguida, o material foi seco à temperatura ambiente, por cerca de 10 min, e os pellets foram ressuspensos em 20  $\mu\text{L}$  de água livre de RNase. A qualidade do RNA total isolado foi avaliada por eletroforese em gel de agarose 1,2% em NBC ( $\text{H}_3\text{BO}_3$  0.05 mol.L<sup>-1</sup>,  $\text{Na}_3\text{C}_6\text{H}_5\text{O}_7$  1 mmol.L<sup>-1</sup> e NaOH 5mmol.L<sup>-1</sup> pH 7.5) corado com brometo de etídeo 0,1 g/mL.

O sequenciamento de RNA foi obtido utilizando a plataforma Illumina Hi-seq 2500. As bibliotecas para RNA-seq foram preparadas com o kit TrueSeq-RNA Sample Prep Kit Illumina. O kit foi empregado conforme o protocolo “*Low sample*”, realizando depleção do RNA ribossomal, usando modo paired-end, com os seguintes parâmetros de filtro de qualidade: máximo de uma base abaixo da qualidade de 30 nas primeiras 5 bases. A qualidade dos reads foi avaliada por meio das métricas estatísticas implementadas na ferramenta de análise de qualidade de dados FastQC (Andrews, 2017). Adaptadores illumina foram removidos utilizando software Trimmomatic (Bolger *et al.*, 2014) e as reads de baixa qualidade de sequenciamento foram filtradas e trimadas utilizando o pacote de software prinSEQ (Schmieder & Edwards, 2011). O procedimento de mapeamento dos reads foi executado usando o programa Bowtie 2 (Langmead *et al.*, 2009), e o conjunto de cDNAs foi obtido a partir do banco de dados *Arabidopsis* Information Resource (TAIR) (<http://www.arabidopsis.org>), décima versão. O processo de contagem de reads mapeadas foi realizado por meio do software HTseq count (Anders, 2010). Os dados totais foram normalizados usando fator de normalização TMM, e normalizado por DeSeq2 (*Trimmed Mean of M-values*; Robinson & Oshlack, 2010). A análise global da expressão gênica diferencial (DGE) foi realizada utilizando o pacote edgeR (Robinson *et al.*, 2010) e DEseq2 (Anders & Huber, 2010), ambos presentes no repositório

R/Bioconductor. A dispersão foi estimada pelo parâmetro *edgeR* tagwise. A expressão diferencial foi apurada utilizando o p-valor ajustado pelo método de “False discover rate” (FDR) com corte de 0,05. A classificação da ontologia gênica (Gene ontology) e análise de enriquecimento das vias metabólicas e de sinalização foram realizadas utilizando os pacotes *GoStats* (Falcon & Gentleman, 2007) e *PathView* (Luo & Brouwer, 2013) ambos presentes no repositório R/Bioconductor *GOstats*.

### 3.4 Identificação de novos sítios de ligação ao genoma do fator de transcrição LIMYB em *A. thaliana* por meio de ChIP-seq

Uma vez selecionadas linhagens do mutante *Salk\_032054* expressando LIMYB fusionado a YFP, as plantas foram submetidas a técnica de ChIP (Imunoprecipitação de Cromatina), com o objetivo de identificar outros possíveis alvos de ligação e regulação do fator de transcrição LIMYB.

O ensaio de ChIP-seq foi realizado de acordo com os parâmetros de qualidade estabelecidos pelos consórcios ENCODE e modENCODE (Landt & Marinov, 2012). Foram utilizadas plantas expressando LIMYB fusionada a proteína fluorescente YFP, sob o controle do promotor 35S de *CaMV* (*Cauliflower mosaic virus*). Como caracterização do anticorpo a ser utilizado no ensaio, foi realizado um imunoblot do extrato proteico total das plantas transformadas expressando LIMYB fusionado a YFP. O ensaio de ChIP seguiu o protocolo de otimizado especificamente para plantas de *A. thaliana*, estabelecido por Yamaguchi e colaboradores (2014), com algumas modificações. Após o ensaio de ChIP as amostras foram preparadas para sequenciamento em equipamento Illumina HiSeq 4000.

#### 3.4.1 Coleta e fixação do material vegetal para Imunoprecipitação de Cromatina (ChIP)

Plantas germinadas em meio MS ½ força foram coletadas após 20 dias de crescimento para o processo de fixação do tecido. Foram coletados 600 mg de tecido jovem em tubos falcon de 15 mL, seguido por lavagem com 10 mL PBS

1X, e após feito substituição deste por 10 mL de uma solução de fixação contendo 1% de Formaldeído em PBS 1X. Para infiltração da solução de formaldeído nas células vegetais foi aplicado uma sequência de vácuos de -500 Pa, de 1,5 min, 1 min, 1min respectivamente, com intervalo de 30 s entre cada vácuo. Após a solução de fixação foi removida e substituída por uma solução de glicina 0,125 M e novamente submetido a vácuo por 1,5 min. Em seguida a solução de glicina foi removida e a amostra novamente lavada com PBS 1X e seca brevemente com uso de papel toalha, podendo ser armazenada por até três meses a -80°C.

### 3.4.2 Lise nuclear e fragmentação da cromatina

As amostras fixadas e congeladas no processo anterior foram maceradas em nitrogênio líquido e após foi adicionado 3 mL de um tampão de extração nuclear (MOPS 100 mM pH 7,6, MgCl<sub>2</sub> 10 mM, sacarose 0,25 M, dextrano T-40 a 5%, ficoll 400 a 2,5%, 2-Mercaptoetanol 40 mM e 1X Inibidor de protease). Após, o material foi submetido a centrifugação por 5 min a 12000 x g, sendo o pellet ressuspendido em 75 µL de tampão de lise nuclear (50 mM Tris-HCl pH 8,0, 10 mM EDTA pH 8,0, 1% SDS), mantendo em gelo por 30 min, agitando em alguns momentos. Em seguida foram adicionados 700 µL de tampão de diluição ChIP 'sem triton' (Tris-HCl 16,7 mM pH 8,0, NaCl 167 mM, EDTA 1,2 mM, SDS 0,01%).

A fragmentação do DNA foi feita utilizando um sonicador, com amplitude constante de 80 µm, 6 pulsos, por aproximadamente 30 min, 30s on/off em gelo, para intervalos de corte de 100-1000pb. A solução sonicada foi acrescentada de tampão de diluição ChIP com triton (Tris-HCl 16,7 mM pH 8,0, NaCl 167 mM, EDTA 1,2 mM, Triton- X 1,1% e SDS 0,01%), seguido de centrifugação por 5 min a 4° C, 13200 rpm sendo o sobrenadante transferido para um novo microtubo para remoção de impurezas.

### 3.4.3 Pré-clareamento e Imunoprecipitação

Para eliminação de interações inespecíficas, 50 µL beads de proteína A-agarose foram incubadas com as amostras entre 1 a 2 h e após removidas por centrifugação a 900 x g por 3 min. Após, 18 µL de amostra para imput (Controle de DNA total fragmentado) foram armazenados a -80° C.

A imunoprecipitação foi feita adicionando a concentração adequada de anticorpo por tubo, e mantida a incubação overnight a 4° C com rotação. Em seguida, 50 µL de beads de proteína A-agarose foram incubadas por 4h a 4° C com rotação para captura dos complexos anticorpo-proteína-DNA.

Os pellets da imunoprecipitação contendo beads foram submetidos a uma sequência de lavagens após o passo de incubação. Inicialmente as amostras foram centrifugadas novamente a 900 x g por 3min, seguido da remoção do sobrenadante do microtubo e estocado a -80° C. O pellet de beads foi inicialmente lavado 2x com um tampão de baixa concentração de sal (SDS a 0,1%, Triton-1 a 1%, EDTA 2 mM, Tris-HCl 20 mM, pH 8,1, NaCl 150 mM) a 4° C por 5 min. As amostras foram novamente centrifugadas a 900 x g por 3 min para precipitação das beads e a solução de lavagem foi substituída por um tampão de alto teor de sal (SDS a 0,1%, Triton-X a 1%, EDTA 2 mM, Tris-HCl 20 mM, pH 8,1, NaCl 500 mM) e repetido o mesmo processo descrito acima. Em seguida, as beads foram lavadas por um tampão de LiCl 250 mM (LiCl 0,25 M, IGEPAL-CA630 a 1%, desoxicolato a 1%, EDTA 1 mM, Tris-HCl 10 mM pH 8,1) e por último uma lavagem com uma solução x 0,5 TE (Tris-HCl pH 8,0, 10 mM/EDTA pH 8,0 1M).

### 3.4.4 *Crosslink-reverso* e purificação do DNA imunoprecipitado

A interação proteína/DNA foi desfeita pelo processo de *crosslink-reverso*, para que fosse possível o isolamento dos ácidos nucleicos em coluna de purificação. Inicialmente foram pipetados no pellet de beads 50 µL de Tampão de lise nuclear e incubação durante 30 min a 65 ° C para separar a amostra das beads. O tampão de eluição foi removido e armazenado, e mais 50 µL de tampão

serão adicionados as beads e repetido o processo de aquecimento e suspensão. Após, os dois conjuntos de tampão de eluição foram combinados e descartadas as beads. Na sequência, foram adicionados 82 µL de Tampão de lise nuclear às amostras de input (item 3.4.3) para completar um total de 100 µL, e mais 6 µL de NaCl 5 M a as amostras de input e ChIP. As amostras foram incubadas a 65 °C overnight.

A purificação do DNA após o passo de crosslink foi feita utilizando kit QIAprep® Spin Miniprep, por adição de 550 µL de tampão PB a cada tubo, mantendo as amostras em aparelho agitação durante 30 min à temperatura ambiente. Em seguida, as colunas de purificação QIAprep foram carregadas com as amostras e centrifugadas durante 1 min a 13.200 rpm em todos passos descritos a seguir. Em sequência, carregada 750 µL de tampão PE à coluna, centrifugada, e repetido o passo para eliminação de resquícios do tampão. Ao final, as colunas foram colocadas em microtubos estéreis e as amostras foram eluídas em 150 µL de tampão 0,5x TE ao centro da coluna, deixando em repouso por 5 min, recarregando a coluna e centrifugando novamente.

A avaliação do tamanho dos fragmentos obtidos no procedimento de ChIP foi executada por eletroforese digital no equipamento Bioanalyzer 2100 (Agilent technologies), utilizando um chip High Sensitivity de DNA.

### 3.5 Análise dos novos sítios de ligação de LIMYB e identificação de cis-elementos e características conservadas

O sequenciamento dos fragmentos obtidos no procedimento de ChIP foi realizada utilizando o Kit TruSeq® ChIP Sample Preparation, na plataforma HiSeq 4000, Read Length Paired-end 100 bp, com Output de dados de 20 M paired-end reads por amostra. O protocolo utilizado, adiciona sequências de adaptadores nas extremidades das moléculas de ChIPed-DNA para gerar bibliotecas indexadas de leitura única ou de sequências pares. O DNA de ChIP de entrada (5-10 ng) foi fosforilado em extremidades sem cortes. Unicamente, nucleótido “A” foi adicionado às extremidades 3' dos fragmentos em preparação para ligação a um adaptador que tem base única "T". De modo que o

sequenciamento a partir das duas extremidades para fragmentos de DNA, selecionasse fragmentos com o tamanho entre 100-800 bp.

A qualidade das reads foi avaliada por meio das métricas estatísticas implementadas na ferramenta de análise de qualidade de dados FastQC (Andrews, 2017). Adaptadores illumina foram removidos utilizando software Trimmomatic (Bolger *et al.*, 2014) e as reads de baixa qualidade de sequenciamento foram filtradas utilizando o pacote de software printSEQ (Schmieder & Edwards, 2011). O procedimento de mapeamento dos reads foi executado usando o programa Bowtie 2 (Langmead *et al.*, 2009). Após esse processo, os reads obtidos foram mapeados contra o genoma de *Arabidopsis thaliana* disponível na base de dados online do TAIR (<ftp://ftp.arabidopsis.org/home/tair/Genes/>), com uma tolerância de 2 a 3 mismatches. O pacote de livre acesso disponível MACS2 (Zhang *et al.*, 2008) foi utilizado para análises dos dados de ChIP-seq. Os possíveis genes alvo e distribuição dos picos do ChIP-seq serão calculados com o software ChIPpeakAnno (Zhu *et al.*, 2010; Zhu L., 2013). A análise de enriquecimento de GO foi conduzida com o pacote GOstats (Falcon & Gentleman, 2007).

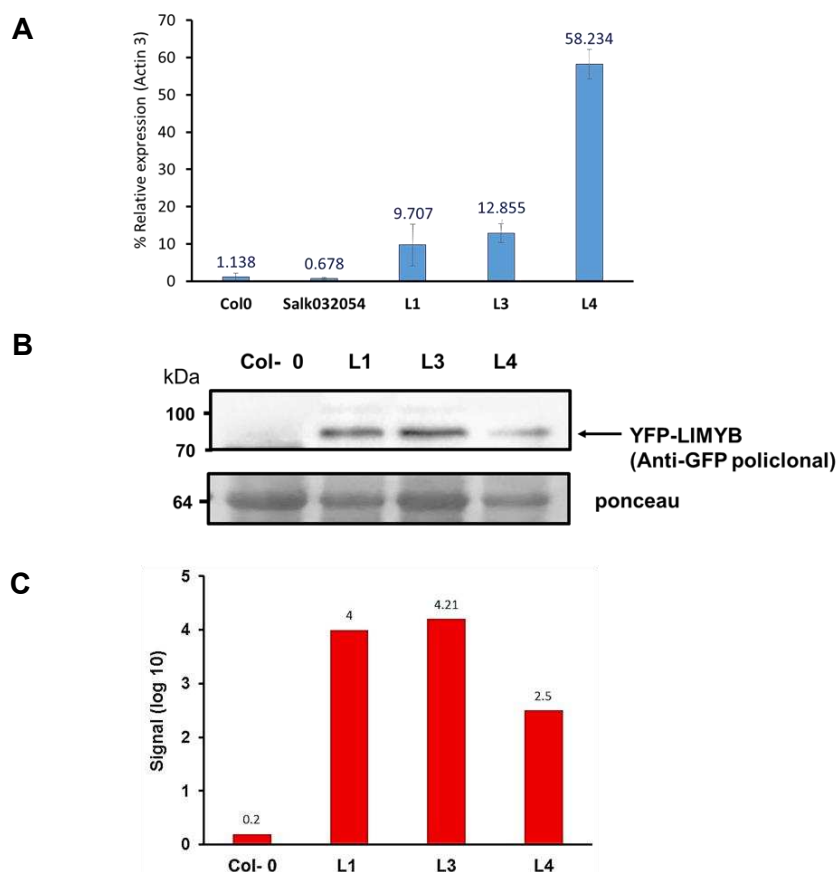
## 4. Resultados e Discussão

### 4.1 Caracterização de linhagens transformantes overexpressando LIMYB

Estudos apresentados em investigações anteriores mostraram que a proteína LIMYB, componente *downstream* na via de defesa antiviral mediada por NIK1, atua como um fator de transcrição, capaz de regular negativamente a expressão de genes ribossomais, levando a inibição global da tradução (Zorzatto *et al.*, 2015). LIMYB pertence ao grupo 2R-Myb ou R2R3-Myb de proteínas Myb em *Arabidopsis thaliana* (Dubos *et al.*, 2010). As proteínas do grupo R2R3-Myb estão envolvidas em diversos processos celulares de plantas, incluindo metabolismo primário e secundário, diferenciação celular, processos de desenvolvimento e respostas a estresses bióticos e abióticos (Dubos *et al.*, 2010). Sendo caracterizada como proteína da família MYB que possui ampla gama de funções, é possível que LIMYB opera em outros processos celulares básicos além do controle de tradução. Para melhor compreender a função desta proteína e seu papel na remodelação dos processos celulares, foi conduzido estudos sobre a abrangência genômica de LIMYB.

Linhagens de *Arabidopsis thaliana limyb* Salk\_032054, foram previamente transformadas com a construção pYFP-At5g05800ST (pUFV1887), na qual o cDNA de LIMYB está sob o controle do promotor 35S e contém YFP ligado à sua porção N-terminal, previamente descrita (Zorzatto *et al.*, 2014). Para as etapas subsequentes foram utilizadas três linhagens transformadas independentemente, denominadas LIMYB-32-L1, LIMYB-32-L3 e LIMYB-32-L4, que foram confirmadas por PCR (Figura S1).

O nível de expressão de LIMYB nas linhagens selecionadas, foi monitorado por qRT-PCR (Figura 1A), incluindo um mutante *limyb-32* e Col-0.



**Figura 1. Acúmulo de YFP-LIMYB nas linhagens transgênicas. (A)** Acúmulo de transcritos de YFP-LIMYB nas linhagens transgênicas L1, L3 e L4. A expressão gênica foi analisada por qRT-PCR, calculada no método de  $2^{-\Delta Ct}$ , foi utilizado actina, um gene endógeno como controle. As barras representam intervalo de confiança de 95% com base em replicatas de 3 amostras de experimentos independentes. **(B)** Acúmulo da proteína recombinante YFP-LIMYB em L1, L3 e L4. Proteínas totais foram extraídas de Col -0, L1, L3 e L4 e separadas por eletroforese, transferidas para membranas e sondadas com anticorpo anti-GFP policlonal. **(C)** Quantificação por intensidade de banda do nível de YFP-LIMYB em L1, L3L e L4. A quantificação das bandas foi feita utilizando o programa Quantity One BioRad.

Os transformantes apresentaram um nível de expressão superior a planta selvagem, sendo que LIMYB-32-L4 apresentou um alto nível de expressão em relação a LIMYB-32-L1 e LIMYB-32-L3, e em torno de 57 vezes mais que a planta selvagem, enquanto o índice de expressão no mutante *limyb-32* mostra-se abaixo de Col- 0, conforme já demonstrado por Zorzatto (2015).

O acúmulo da proteína LIMYB foi monitorado por *western blot* utilizando anti-GFP policlonal (Figura 1B). Nesta análise além das linhagens anteriores foi adicionado um novo transformante denominado LIMYB-32-L8 (Figura S2), que

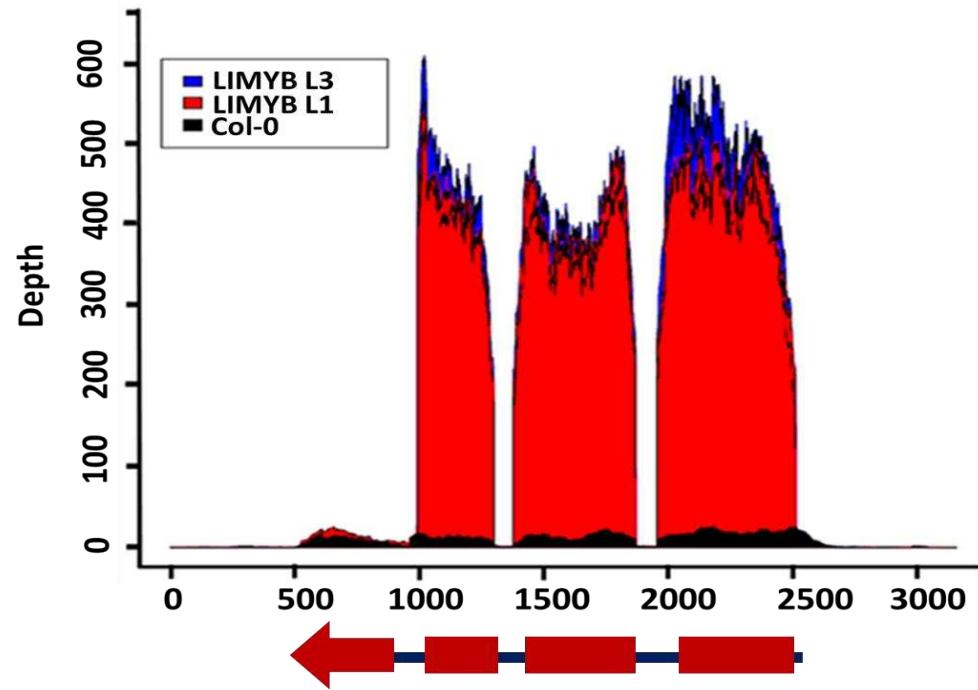
foi utilizado posteriormente utilizado como um dos controles em confirmações por PCR no experimento de CHIP. Apesar de maior nível de transcritos na linhagem LIMYB-L4, o acúmulo da proteína recombinante YFP-LIMYB foi superior nas linhagens LIMYB-32-L1 e LIMYB-32-L2 e não detectado em plantas selvagens (Figura 1B). Comparado com LIMYB-32-L4, LIMYB-32-L1 e LIMYB-32-L3 apresentaram acúmulo 2x superior de YFP-LIMYB (Figura 1C) e foram selecionadas para os estudos subsequentes.

#### **4.2 Variação global da expressão gênica induzida pela expressão ectópica de LIMYB**

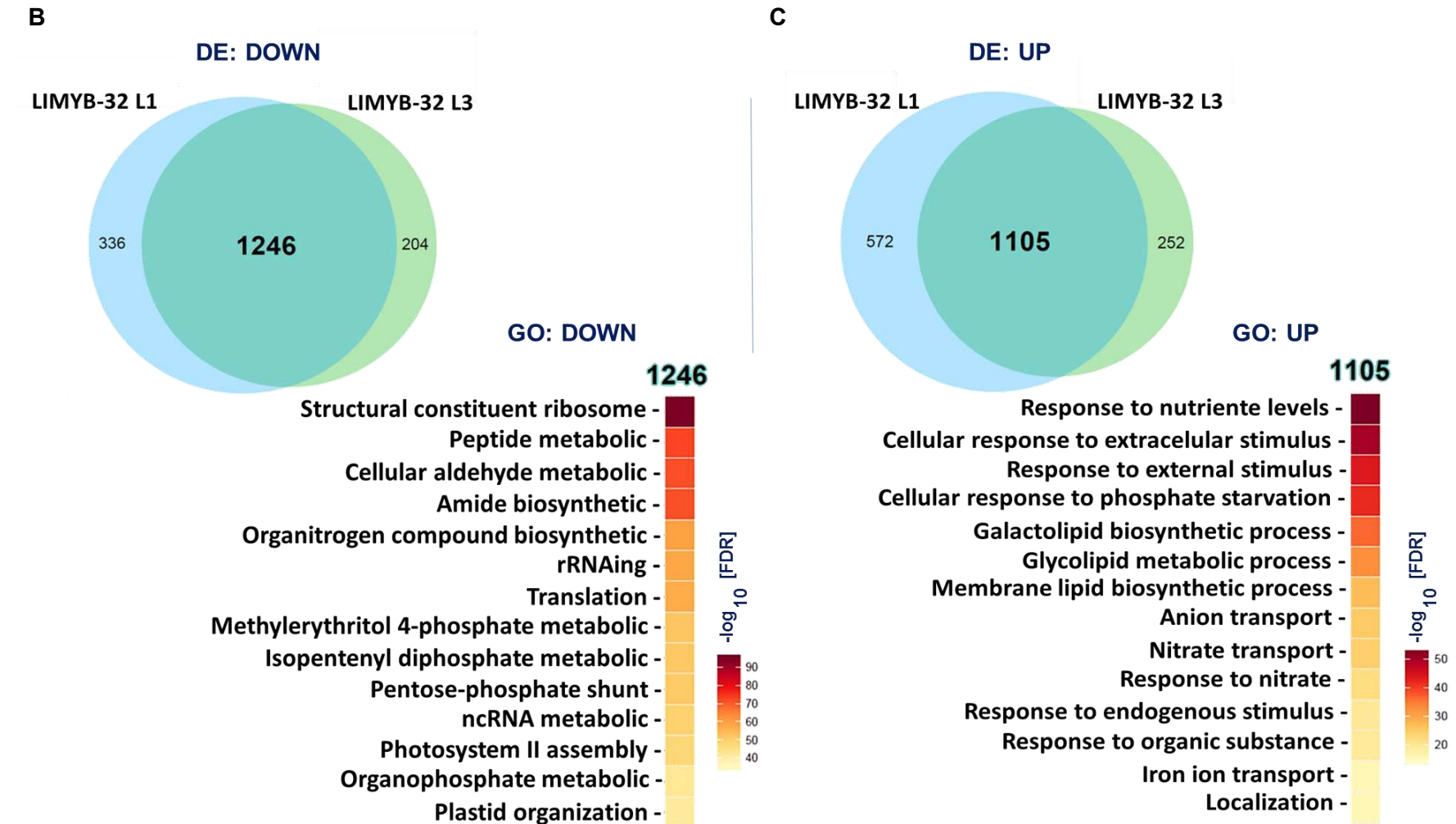
LIMYB interage com RPL10 para reprimir a expressão de genes de proteínas ribossomais, como resultado da ativação da via de sinalização antiviral mediada por NIK1 (Zorzatto *et al.*, 2015). Por pertencer ao maior grupo de proteínas MYB em *Arabidopsis*, envolvidas em uma ampla gama de redes transcricionais, foi levantada a hipótese de que LIMYB pudesse exercer funções adicionais em processos celulares básicos. Com a finalidade de testar esta ideia, inicialmente, foram realizadas, análises de sequenciamento de RNA com plantas expressando ectopicamente LIMYB (LIMYB-32-L1 e LIMYB-32-L3), e plantas selvagens, visando a identificação da variação global de expressão gênica induzida por LIMYB.

A avaliação da abundância relativa dos *hits* de RNA mapeados no *locus* de LIMYB (At5g05800), demonstrou que estes elementos de sequência apresentaram altos níveis de transcritos para LIMYB-32-L1 e LIMYB-32-L3 em relação a planta selvagem (Figura 2A). Os resultados de qRT-PCR (Figura 1A) são consistentes com os dados de RNA-seq, no qual Col-0 mantém um menor nível de expressão dos transcritos, quando comparado novamente com os níveis de LIMYB-32-L1 e LIMYB-32-L3. LIMYB-32-L1 e LIMYB-32-L3 também apresentaram perfil semelhante de intensidade exibido na análise de cobertura do RNA-seq. De fato, tomando como base estas duas análises de expressão, observa-se que o processo de complementação acarretou num quadro de superexpressão de LIMYB nestas linhagens.

A



ID	log2 FoldChange	Ajusted p-value	Description
AT5G05800 (L1)	3.50337190811801	1.34287359745286e-29	LIMYB-32
AT5G05800 (L3)	3.76386267848899	1.3861794344564e-34	LIMYB-32

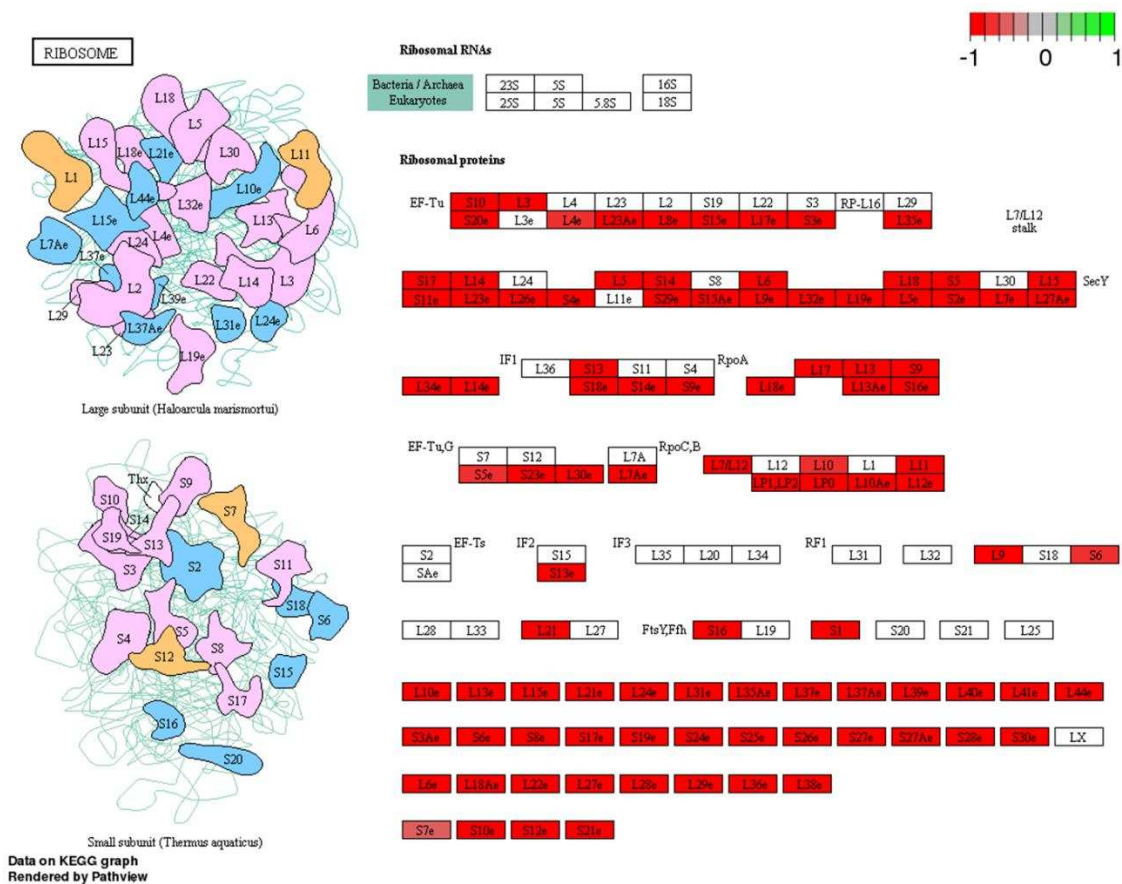


**Figura 2. Análise de RNA-seq da variação global da expressão gênica sob efeito da overexpressão de LIMYB.** (A) Representação esquemática do *Locus* de LIMYB (At5g05800) e dados de sequenciamento do RNA mapeados. O gene contém três introns e quatro exóons. A abundância relativa dos *hits* de RNA mapeadas no locus At5g05800 é mostrado em preto para Col -0, vermelho para L1 e azul em L3. (B) Intercessão de genes *down* regulados em diagrama de Venn na análise de DE; Representação por *Heat map* da GO de genes em comum down regulados. Para as análises de DE da GO foi usado um valor de FDR de  $-\log_{10}$ . (C) Diagrama de Vem de genes *up* regulados em expressão diferencial (DE) para L1 e L3; *Heat map* da ontologia gênica (GO), representando a distribuição de categorias funcionais da intercessão de genes *up* regulados; Para as análises de DE da GO foi usado um valor de FDR de  $-\log_{10}$ .

Análise de expressão diferencial (*Differential expression*, DE), utilizando os critérios  $p\text{-value} < 0.01$  e  $\log_2\text{-fold-change} > 1.0$  para genes *up*-regulados e  $\log_2\text{-fold-change} < 1.0$  para genes *down*-regulados, revelou que as linhagens LIMYB-32-L1 e LIMYB-32-L3 compartilham 1246 genes *down* regulados e 1105 genes *up* regulados por superexpressão de LIMYB (Figura 2B, C; Tabela S1, S2.2). Dentre estes transcritos *down* regulados (Figura 2B; Tabela S1, S1.2), o *heat map* com base na anotação das categorias de ontologia gênica (*Gene ontology*: GO; Tabela S3) identificou os constituintes estruturais de ribossomos como fortemente reprimidos, assumindo o topo da lista, dentre os grupos categorizados. Estes resultados são consistente com a função repressora de LIMYB sobre os genes de proteínas ribossomais (Zorzatto *et al.*, 2015).

Analisando as demais categorias diferencialmente expressas por LIMYB, são evidenciadas possíveis funções inéditas quanto sua capacidade de regulação. Dentre os genes *down* regulados, as categorias se enquadram entre processos de biossíntese de moléculas, desenvolvimento, tradução propriamente dita, e até mesmo fotossíntese, dentre outros, ampliando ainda mais o viés do mecanismo de atuação de LIMYB. As proteínas do grupo R2R3-Myb, ao qual LIMYB é catalogada, estão envolvidos em diversos processos específicos de plantas, como processos de desenvolvimento, diferenciação celular, metabolismo primário e secundário, e respostas a estresses bióticos e abióticos (Dubos *et al.*, 2010).

Recentemente, foi demonstrado que LIMYB atua como repressor transcricional e reduz drasticamente a síntese de transcritos correspondentes aos genes ribossomais analisados, incluindo *S13a*, *S25*, *L4/L1*, *L13*, *L28e* e *L18e* (Zorzatto *et al.*, 2013). Os dados de sequenciamento de RNA total de LIMYB-32-L1 e LIMYB-32-L3 reforçam a atividade repressora de LIMYB sob estes genes ribossomais e ampliam o elenco de genes ribossomais *down*-regulados por superexpressão de LIMYB, representados via KEGG (*Kyoto Encyclopedia of Genes and Genomes*) (Figura 3; Tabela suplementar 3).



**Figura 3. Overexpressão de LIMYB *down* regulando genes de proteínas ribossomais.** A análise de expressão diferencial representada via KEGG, demonstrando que o enriquecimento das redes de genes ribossomais foi inferior a Col-0, em plantas sob a atuação de LIMYB.

Os fatores de transcrição são amplamente envolvidos em vários processos biológicos e desempenham papéis importantes nas respostas das plantas ao estresse biótico. As proteínas MYB compreendem uma grande família de fatores de transcrição de plantas que desempenham papéis regulatórios em diferentes processos biológicos como respostas de defesa (Dubos *et al.*, 2010; Katiyar *et al.*, 2012). Em relação a *down* regulação dos genes ribossomais apresentados acima (Figura 3), estudos comprovaram a relação desta repressão com o mecanismo de defesa da via antiviral mediada por NIK1 (Zorzatto *et al.*, 2015).

A disposição de genes *up* regulados na análise de RNA-seq (Figura 2C, Tabela S2), revela que o enriquecimento de categorias funcionais (GO), ficou mais intenso entre processos de resposta aos níveis de nutrientes, seguido por processos relacionados a estímulos externos e internos e transporte de íons (Tabela 2). Este conjunto de genes *up* regulados podem estar relacionados tanto

à escassez de moléculas como fosfato, quanto ao acúmulo flutuante de produtos, gerados pelo bloqueio ou retardo de diversos processos de biossíntese de macromoléculas induzidos pela overexpressão de LIMYB, conforme notado entre genes *down* regulados (Figura 2B).

**Tabela 2. Genes agrupados na categoria de resposta ao nível de nutrientes (GO:0031669).**

ID	Log2 fold change	padj	Description
AT1G73010	8.372574	1.7E-261	phosphate starvation-induced gene 2 Encodes PPsPase1
AT3G09922	7.421661	5.8E-76	induced by phosphate starvation1 AF236376 Arabidopsis thaliana IPS1
AT2G46880	7.370094	1.56E-73	purple acid phosphatase 14 purple acid phosphatase 14 (PAP1 4);
AT2G11810	6.628731	4E-148	monogalactosyldiacylglycerol synthase type C MGD3
AT1G73220	6.614551	1.28E-69	organic cation/carnitine transporter1
AT2G45130	6.310646	4.7E-126	SPX domain gene 3 Expression is upregulated in the shoot of cax1/cax3
AT5G43370	6.05199	1.97E-27	phosphate transporter 2 Encodes a phosphate transporter Pht1;2.
AT3G01420	5.598822	1.5E-57	Peroxidase superfamily protein Encodes an alpha-dioxygenase
AT3G52820	5.461233	1.77E-58	purple acid phosphatase 22 purple acid phosphatase 22 (PAP22);
AT3G44510	5.114266	1.06E-73	alpha/beta-Hydrolases superfamily protein
AT5G48400	5.0017	9.42E-22	Glutamate receptor protein member ion channel subunit familyATGLR1.2;
AT1G18970	4.97005	7.78E-26	germin-like protein 4 Encodes a germin-like protein
AT1G17710	4.917454	9.3E-168	Pyridoxal phosphate phosphatase-related protein
AT4G24890	4.890727	3.36E-65	purple acid phosphatase 24 purple acid phosphatase 24 (PAP24)
AT2G45570	4.751457	9.96E-16	cytochrome P450, family 76, subfamily C, member of CYP76C
AT1G22990	4.611067	1.46E-30	Heavy metal transport/detoxification superfamily protein
AT5G03545	4.523645	2.2E-105	Expressed in response to phosphate starvation
AT1G08310	4.454611	1.55E-82	alpha/beta-Hydrolases superfamily protein
AT1G69930	4.285187	9.23E-10	glutathione S-transferase TAU 11 Encodes glutathione transferase
AT2G34500	4.177359	7.3E-21	cytochrome P450, family 710, subfamily A,
AT1G18980	4.071488	5.68E-20	RmlC-like cupins superfamily protein RmlC-like cupins superfamily protein;
AT4G04490	4.027803	5.89E-36	cysteine-rich RLK (RECEPTOR-like protein kinase)
AT1G02310	3.977802	2.66E-31	Glycosyl hydrolase superfamily protein;
AT3G09220	3.888482	1.08E-39	laccase 7 putative laccase, a member of laccase family of genes
AT1G69480	3.847622	3.57E-26	EXS (ERD1/XPR1/SYG1) family protein EXS (ERD1/XPR1/SYG1)
AT5G44610	3.811521	3.96E-27	microtubule-associated protein 18 with seven repeated VEEKK motifs
AT2G47190	3.788077	1.53E-36	myb domain protein 2 Encodes a MYB transcription factor
AT2G48130	3.785032	1.86E-12	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S
AT5G48850	3.741878	3.94E-26	Tetratricopeptide repeat (TPR)-like superfamily
AT3G04530	3.677546	4.71E-22	phosphoenolpyruvate carboxylase kinase 2

A expressão ectópica de proteínas R3-Myb revela que muitos MYBs são capazes de regular processos de desenvolvimento estrutural como formação de tricomas e capilares da raiz, dentre outros, e são responsivos a sinais hormonais

e de estresses in *Arabidopsis* (Chen *et al.*, 2006; Zheng *et al.*, 2016; Tian *et al.*, 2017). Mudanças sejam elas de crescimento ou desenvolvimento, requerem a integração de múltiplos sinais ambientais, como luz e temperatura, na expressão de genes endógenos (Fitter *et al.*, 2002; Willis *et al.*, 2008). Genes de fatores de transcrição MYB, como por exemplo LUX ARRHYTHMO (LUX), atuando juntamente com outros genes, auxiliam na regulação do crescimento de alongamento em estudos realizados com *Arabidopsis*, reprimindo diretamente a expressão de fatores como, bHLH transcription factor Phytochrome Interacting Factor 4 (PIF4), necessários à sensibilidade de temperatura e ao florescimento (Nusinow *et al.*, 2011; Huang *et al.*, 2016).

A atuação molecular de LIMYB, a partir destas análises, se mostra além do mecanismo de repressão de genes ribossomais, indicando que, de fato, a regulação para inibição global da tradução é gerenciada conjuntamente à coordenação de outros processos celulares.

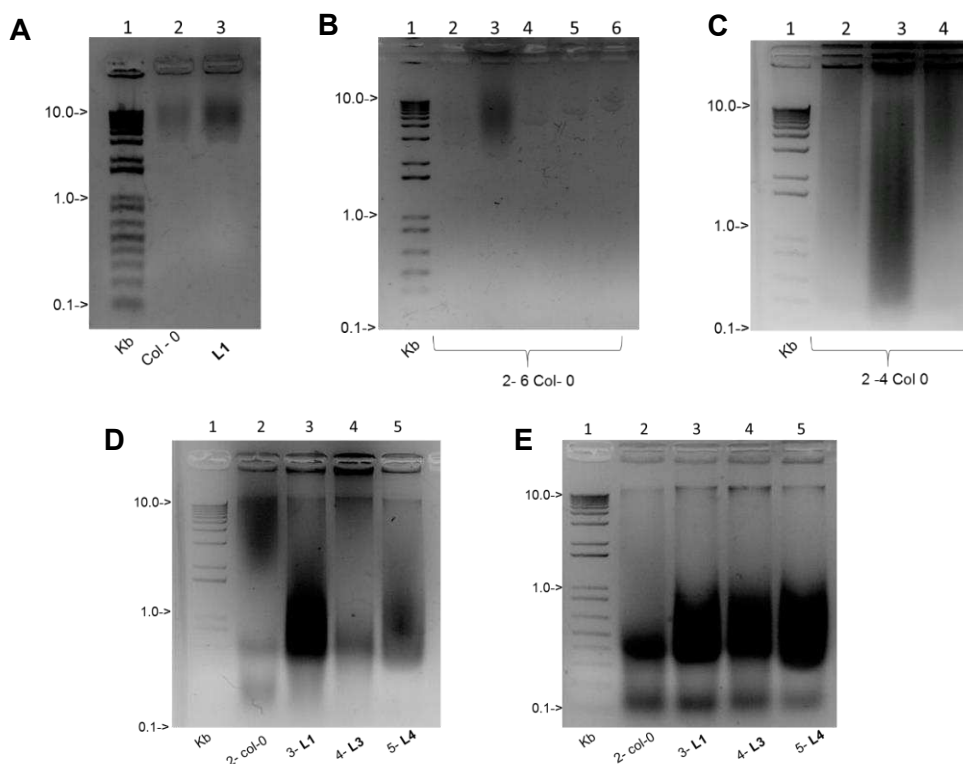
#### **4.3 Estudo da atividade transcricional de LIMYB em escala genômica**

O experimento de RNA-seq revelou um conjunto de genes diferencialmente expressos por superexpressão de LIMYB, enriquecidos em diferentes categorias, além dos genes constituintes dos ribossomos. Entretanto, os alvos diretos de LIMYB em escala genômica não foram ainda determinados. Nessa investigação, a técnica de *ChIP-sequence* foi utilizada para investigar a associação específica de LIMYB com o DNA genômico, caracterizar e confirmar a dinâmica de ligação ao genoma e identificar o conjunto de genes alvo, principalmente aqueles afetados drasticamente pela expressão ectópica de LIMYB. A imunoprecipitação de cromatina é uma valiosa alternativa para sondar interações *in vivo* em diferentes condições fisiológicas, estimar sítios específicos e por seguinte mensurar a força interação da proteína ao DNA (Kuo & Allis, 1999; Yamaguchi *et al.*, 2014).

Inicialmente, as linhagens Col-0, LIMYB-32-L1, LIMYB-32-L3 e LIMYB-32-L4 germinadas em meio MS ½ força por 20 dias, foram submetidas a fixação por

formaldeído, utilizando infiltração a vácuo, para que o formaldeído penetrasse nas células, possibilitando o *cross-link* DNA-Proteína.

Condições ótimas de sonicação/fragmentação do DNA foram estabelecidas empiricamente para obter uma resolução desejada da fragmentação da amostra para o procedimento de ChIP. O padrão de sonicação foi determinado para o tecido e modelo de sonicador disponível, utilizando diferentes tempos de sonicação, mantendo fixo amplitude de 80  $\mu$ m e 6 pulsos, em busca de tamanho médio dos fragmentos de DNA entre 100-1000 pb. Alíquotas de amostras extraídas e fragmentadas, foram mantidas a 65°C *overnight* e utilizadas para o teste de eficiência da sonicação, em gel de agarose 1,5% (Figura 4).

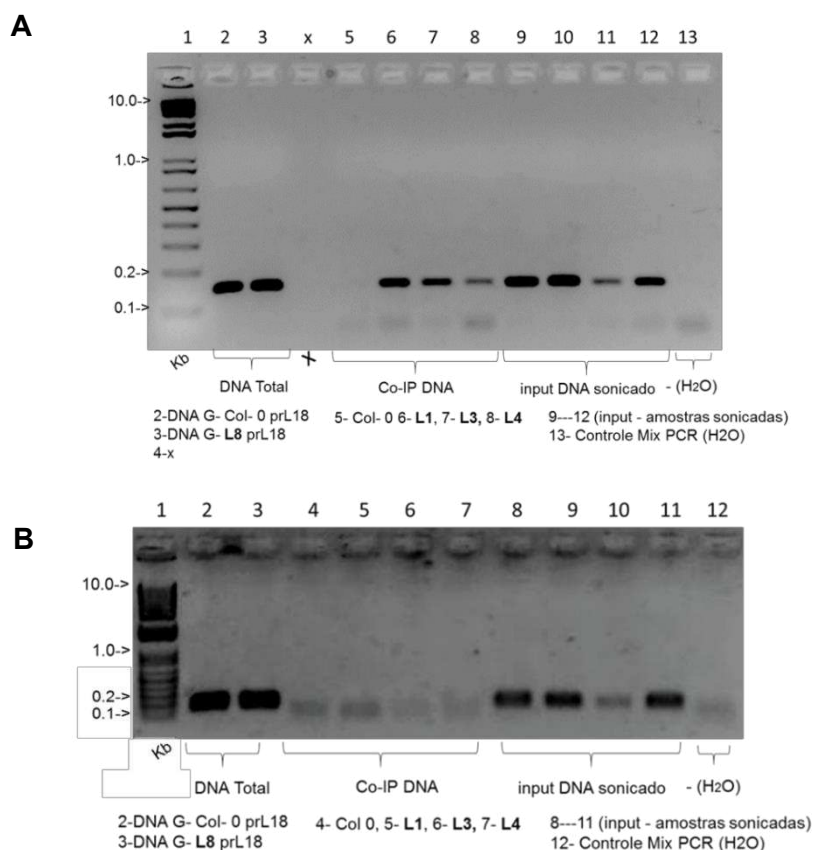


**Figura 4. Padronização das condições sonicação e fragmentação do DNA total para procedimento de ChIP. (A)** Teste de extração/Sonicação, utilizando intervalos de 30 s on- 30 s off 5 – 10 min, a 4°C. **(B)** Teste de extração/Sonicação, com intervalos de 50 s on- 50 s off, 5 a 5 – 25 min, 4°C **(C)** Teste de extração/Sonicação 30 s on- 30 s off, 10 a 35 min. **(D)** Padronização do tamanho de fragmentos entre aproximadamente 100 -1000 pb, utilizando fragmentação por 35 min, 30 s on- 30 s off, 4°C.

O DNA fragmentado nas condições acima (Figura 4E), foi utilizado em seguida para coimunoprecipitação (CoIP) utilizando anticorpo específico anti-GFP policlonal. Posteriormente o complexo DNA/proteína foi submetido ao processo de cross-link reverso, e o ChIPed-DNA (DNA imunoprecipitado por ChIP) submetido à confirmação por PCR, utilizando oligonucleotídeos específicos para 150 pb do promotor do gene *RPL18* (Figura 5A; Zorzatto *et al.*, 2015). A região de 150 bp foi amplificada a partir do DNA co-imunoprecipitado (Co-IP) dos transformantes (colunas 6-8), mas não a partir de DNA Co-IP de Col-0 (coluna 5) e do controle negativo para contaminações da reação de PCR (coluna 13), comprovando a eficiência do procedimento de ChIP (Figura 5A). As colunas 9-12 correspondem à amplificação do segmento do promotor RPL18 a partir de inputs Col-0 e de linhagens transgênicas. Em seguida, para confirmar a especificidade do anticorpo utilizado, foi efetuado CoIP com anticorpo não correlacionado, anti-IgG, no qual nenhuma das amostras apresentou amplificação para o material co-imunoprecipitado (Figura 5B, colunas 4-7). Coletivamente, estes resultados demonstraram a eficiência e especificidade das reações de co-imunoprecipitação de DNA, utilizando anticorpos policlonais anti-GFP.

O DNA co-imunoprecipitado foi enriquecido utilizando kit WGA polimerase, com duas repetições por amostra, para análises de qualidade e concentração, por nanodrop e eletroforese digital por Bioanalyser agilent 2100 (Figura S3). Após, o ChIPed-DNA foi sequenciado em plataforma *illumina HiSeq 4000*. As qualidades de sequências foram avaliadas via FastQC (Figura S4), e após os *reads* de sequência (individualmente para cada amostra) foram mapeados contra o genoma de *Arabidopsis thaliana*. As sequências mapeadas foram submetidas a análises de enriquecimento de picos (*peak calling*), para elicitação de sequências enriquecidas via MACS2. Posteriormente, visando à elucidação das funcionalidades de LIMYB, frente as regiões identificadas, e comuns a todas as linhagens transformantes, análises com base na anotação de ontologia gênica (GO) e KEGG foram utilizadas para agrupamentos de genes em cargo de suas funções. Determinando, um ponto de corte ótimo de log<sub>2</sub>FC (Fold Change) >0.01 e log<sub>10</sub>FDR (False Rate Discovery) > 1.0 para definir os genes alvos de

LIMYB dentro do elenco de genes diferencialmente expressos pela overexpressão de LIMYB.



**Figura 5. Diagnóstico por PCR do ChIPed-DNA por Anti-GFP Policlonal e Anti-IgG.** (A) Amplificação de 150 pb do promotor RPL18, mostrando a ligação *in vivo* para transformantes YFP-LIMYB, detectado por PCR; DNA G, DNA gênomico; prL18, promotor ribossomal L18; L8, Linhagem transformante YFP-LIMYB. (B) Controle do ChIP para anticorpo IgG não específico a ligação com a tag de YFP;

As leituras de RNA-Seq e ChIP-Seq foram alinhadas separadamente ao genoma de referência (*Arabidopsis thaliana*), e em seguida, comparadas para delimitar o conjunto de genes diferencialmente expressos e identificados por ChIP-seq. Foram considerados possíveis alvos diretos de LIMYB apenas os genes identificados por ChIP-seq que foram *down*-regulados ou *up* regulados pela expressão de LIMYB (Figura 6 A, B; Tabela S4, S4.2, S5, S5.2). YFP-LIMYB foi enriquecido na região 5' flanqueadora de 156 genes *down* regulados e 204 genes *up* regulados pela superexpressão do fator transcricional. A maior intensidade de repressão identificada pela anotação com base na análise de *heat*

*map*, a partir da ontologia gênica, se manteve entre genes pertencentes aos constituintes ribossomais (Figura 6A), que também foram encontrados em comum na análise integrativa de ChIP-seq e RNA-seq, demonstrados em representação via KEGG (Figura 6C; Tabela S6). Estes resultados validam os resultados do ChIP-seq, uma vez que genes fortemente *down* regulados, e submetidos à confirmação da interação via ChIP-seq, comumente representam promotores reconhecidos pelo fator transcricional correspondente por alta afinidade (Song *et al.*, 2016).

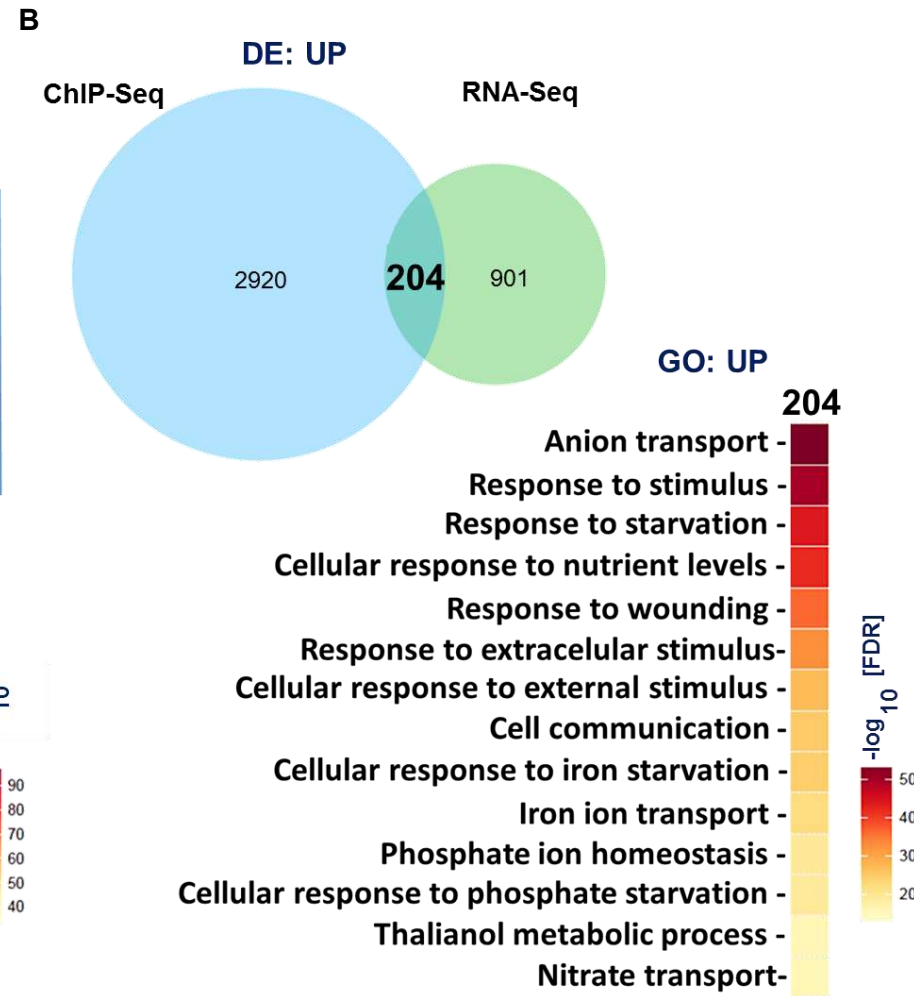
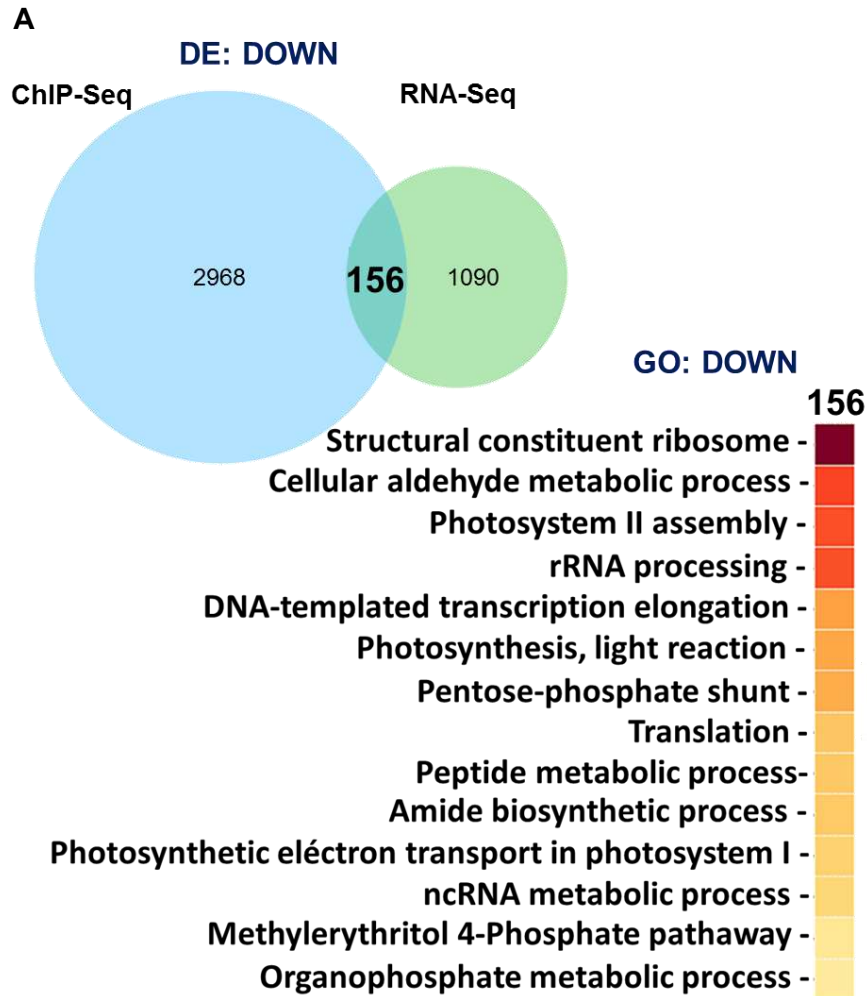
A análise integrativa da atividade funcional de LIMYB comprova inequivocamente sua capacidade de ligação a promotores, não só de uma rede genes ribossomais, mas também a outras classes de genes. Estas incluem classes down-reguladas, conforme detectado simultaneamente nas análises de DE por RNA-seq e ChIP-seq, evidenciando que LIMYB atua também na repressão de genes associados a processos biológicos como metabolismo de aldeído, fotossíntese, transcrição e tradução, dentre outros.

A repressão de processos metabólicos que envolvem proteínas relacionadas ao uso e síntese de componentes aldeídos e derivados, incluiu genes codificantes de DnaJ/HSP40, gliceraldeído 3- fosfato desidrogenase, subunidade delta de ATP-sintase, frutose-bifosfatase aldolase 2, dentre outros (Tabela S7). Destaca-se uma analogia com mecanismos que utilizam essas moléculas em fins de metabolismo primário, nos passos de respiração celular, e de caráter regulatório, relatado em estudos que descrevem a importância destes componentes na resposta a ABA e no tempo de florescimento (Linde *et al.*, 2011; Park & Kim, 2014). Dentre estes, o gene que codifica para a proteína DnaJ/HSP40 apresentou o maior valor para enriquecimento de pico (correspondente a ligação de LIMYB) nas análises de ChIP-seq, e para RNA-seq o valor mais negativo de Fold-Change (Tabela S7). As proteínas DnaJ são co-chaperonas moleculares e funcionam sozinhas ou em associação com parceiros Hsp70, que contribuem no dobramento, redobramento, degradação, e principalmente para a homeostase do nível proteico celular, em condições de estresse (Wang *et al.*, 2004; Hennessy *et al.*, 2005; Craig *et al.*, 2006). Durante parte do processo evolutivo das plantas, a ação conjunta destas proteínas tem sido importante para auxiliar na recuperação efetiva do PSII (Fotosistema II),

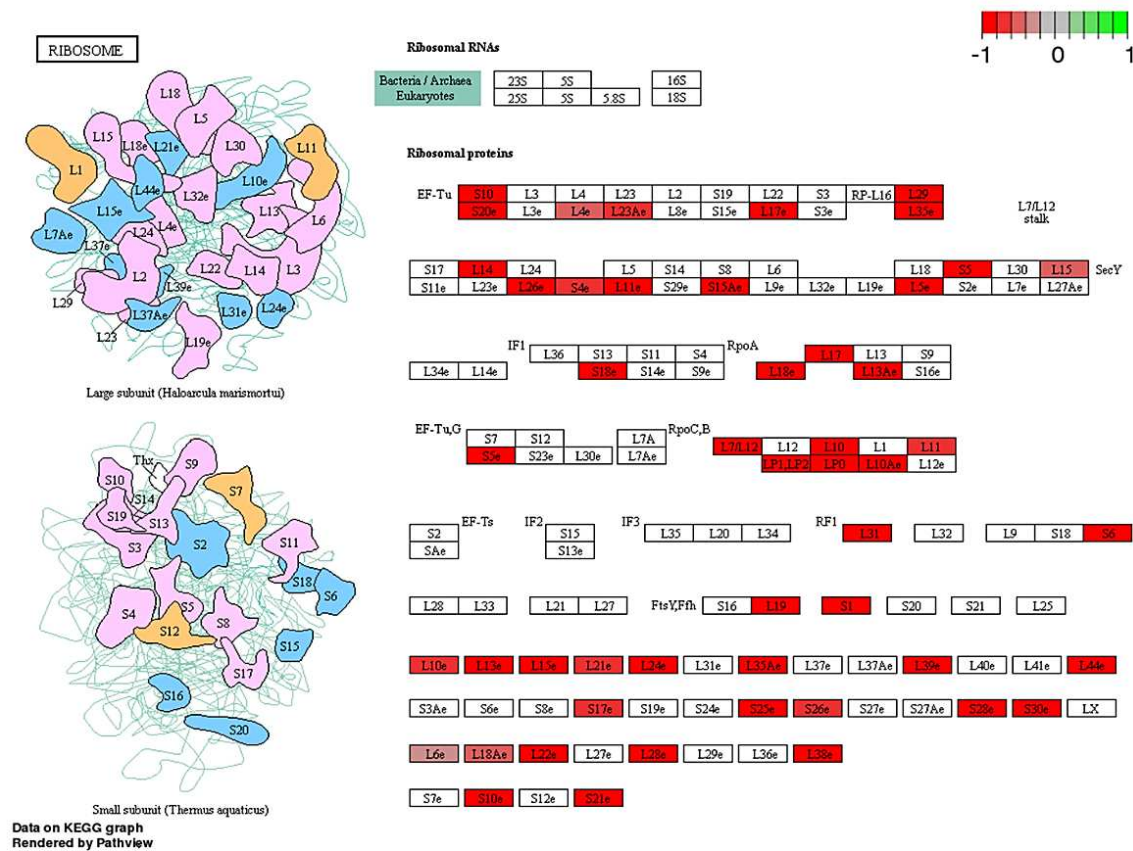
atuando em um dos mecanismos cruciais de proteção e reparo rápido deste complexo (Aro *et al.*, 2005, Takahashi e Badger, 2011). Hsp70 executa dois mecanismos para proteção do PSII contra danos, um que promove a síntese de nova proteína D1 do centro de reação (*PsbA domain*) e outro capaz de redobrar, caso não haja dano intenso (Schroda *et al.*, 2001). DNAj estimula a atividade ATPase de HSP70 e as demais funções, e comumente é encontrada em altas concentrações nos compartimentos celulares ao qual opera (Nakatsukasa *et al.*, 2008). Além de HSP40, a chaperona HSP70 também é notada entre um dos alvos de regulação por LIMYB, contido em componentes de processamento de retículo (Tabela S8). Atuando na regulação da síntese e dobramento de D1 de PSII, DNAJ torna-se um dos prováveis pontos de regulação. Outra razão que reforça esta premissa é o fato de que DNAj/HSP40 é um dos alvos de infecção viral por *TMV* (*Tobacco mosaic virus*), recrutado pelas proteínas de montagem viral (*Coat protein*), e pela proteína MP (*Movement protein*), facilitando o estabelecimento da infecção viral (Shimizu *et al.*, 2009; Hofius *et al.*, 2007). Outros componentes alvos de *TMV* são os elementos dos complexos da cadeia de transporte de elétrons, que foram relatados em altos níveis de expressão, direcionados pela remodelação de processos metabólicos, por ação de proteínas virais (Liao, *et al.*, 2015). Como componente de uma via de sinalização antiviral, não é supressa que alvos diretos de LIMYB estejam envolvidos em infecção viral.

Na análise integrativa realizada neste trabalho, também foi detectado enriquecimento de YFP-LIMYB aos promotores de genes que codificam proteínas que compõem a cadeia de transporte de elétrons e que foram regulados negativamente por LIMYB (Figura S6, Tabela S9). Esta categoria está envolvida a funções de biossíntese, visto que a maior parte da energia em forma de ATP, direcionada a processos de síntese proteica no citosol, é fornecida pela fosforilação oxidativa no complexo ATPsintase (Nelson & Cox, 2014). Foi demonstrado um modelo de defesa antiviral no qual, o efeito do ácido salicílico, afeta a fosforilação mitocondrial e componentes da cadeia de transporte de elétrons, funcionando como um mecanismo de defesa basal contra o vírus *TMV* (Liao, *et al.*, 2015). Similarmente, é sugerido aqui que a inibição por LIMYB compõe um quadro relacionado a regulação de processos biológicos de

modulação traducional e inclui uma possível estratégia de defesa antiviral, conforme já descrito na defesa contra *begomovirus* (Zorzatto *et al.*, 2015).



C



**Figura 6. Análise integrativa da atividade transcricional de LIMYB em escala genômica por ChIP-seq e RNA-seq. (A)** Diagrama de Venn de genes *down* regulados em DE nas análises comparativas ChIP-Seq/RNA-Seq; *Heat map* da ontologia gênica (GO), representando a composição diferencial de categorias funcionais da intercessão de genes *down* regulados. Utilizando corte de  $-\log_{10}$  do p-valor. **(B)** Intercessão de genes *up* regulados em DE em análise integrativa; *Heat map* das categorias funcionais da GO para componentes em *up* regulação. Utilizando corte de  $-\log_{10}$  do p-valor. **(C)** Representação de componentes ribossomais detectados por ChIP-seq e RNA-seq, remetendo em estado de down regulação pela forte associação de LIMYB a seus respectivos promotores. Utilizando corte de  $-\log_{10}$  do p-valor, os dados de componentes ribossomais detectados por ChIP-seq, podem ser vistos separadamente na figura S5, tabela S6.

Consistente com sua função como regulador negativo de tradução global, os resultados de CHIP-seq/RNA-seq revelam que LIMYB interage com promotores de genes de fatores de iniciação da tradução, como eIF2, EIF2C, eIF3, EF1B e eIF2 $\alpha$ . (Tabela 3). Consequentemente a regulação de tradução mediada por LIMYB se estende além do controle negativo da expressão de genes de proteínas ribossomais para acomodar outros componentes da maquinaria de tradução. A regulação negativa da expressão de fatores traducionais regulatórios pode explicar a velocidade com que ocorre a supressão geral de tradução diante a ativação da via de sinalização antiviral mediada por NIK1 e efetivada por LIMYB (Zorzatto *et al.*, 2015). Foi demonstrado que 8 horas após a ativação de NIK1 são suficientes para se observar supressão parcial de tradução global mediada por atuação de LIMYB. Entretanto, a vida média de ribossomos é superior a 24 h e, portanto, a simples repressão de genes de proteínas ribossomais não explicaria a resposta rápida de supressão de tradução resultante da atuação de LIMYB. Provavelmente, a repressão de expressão de outros componentes da maquinaria de tradução de baixa estabilidade pode reconciliar a velocidade de supressão de tradução com o mecanismo de atuação de LIMYB. A regulação destes fatores, faz parte da necessidade de controle traducional pelas plantas, devido a suas habilidades exclusivas entre os seres vivos, como a fotossíntese, e outras necessidades para o seu desenvolvimento e capacidade de resposta aos estresses em vida sésil (Browning *et al.*, 2015).

**Tabela 3.** Componentes de Fatores de tradução detectados em down regulação sob a expressão ectópica de LIMYB.

Gene_id	width	Fold enrich.	log10 q-value	Distance to feature	log2 FC	padj	Description
AT5G12110	185	6.32	24.02	-363	-1.16	0	Translation elongation factor EF1B;
AT5G05470	357	9.76858	52.78	-1750	-0.98	0	Eukaryotic translation initiation factor 2 $\alpha$ -sub.
AT3G11400	244	7.17	30.07	-863	-0.9	0	eukaryotic translation initiation factor 3G1
AT2G36010	252	8.27491	40.5	-829	-0.82	0	E2F transcription factor 3
AT1G70782	198	6.57	25.71	-688	-0.63	0	28 Upstream open reading frames (uORFs)
AT2G40290	242	5.71	19.94	-78	-0.58	0	Eukaryotic translation initiation factor 2 sub. 1
AT1G57720	311	8.71348	49.17	-400	-0.57	0	Translation elongation factor EF1B $\gamma$ -chain
AT1G04170	180	4.62	13.22	-573	-0.45	0.01	eukaryotic translation initiation factor 2 sub. $\gamma$

Detectado nesta análise, eIF2 é um fator de iniciação da tradução bem estudado, responsável pela ligação do met-tRNA subunidade ribossômica menor na presença de GTP (Schmitt *et al.*, 2010). eIF2 é um heterotrímero composto por três subunidades, designadas eIF2 $\alpha$ , eIF2 $\beta$  e eIF2- $\gamma$  (Schmitt *et al.*, 2010). Em mamíferos existem quatro cinases de eIF2, HRI (*heme-regulated inhibitor*), PKR (*double-stranded RNA-dependent protein kinase*), PERK (*PKR-like endoplasmic reticulum kinase*), GCN2 (*general control nonderepressible 2*). Estas cinases regulam a atividade deste fator fosforilando um resíduo conservado de serina na posição 51. Esta fosforilação inibe o início da tradução, em resposta a vários estresses. Quando isso ocorre, eIF2B torna-se incapaz de dissociar-se do eIF2 $\alpha$  fosforilado e, deste modo, a troca de GTP é inibida pelo esgotamento de eIF2 disponível para a formação do complexo ternário (Donnelly *et al.*, 2013). A regulação de GCN2 na formação do complexo ternário em plantas não é detalhada na tradução de mRNAs com uORFs. Dentre estes fatores listados, muito pouco ainda é conhecido sobre eEF1B a nível regulatório, sabe-se é responsável pela troca de GDP de eEF1A inativo por GTP para torna-lo ativo e restaurar sua competência de ligação ao aminoacil-tRNA, para dar continuidade ao processo de alongamento da cadeia polipeptídica (Rodnina e Wintermeyer, 2009). No entanto, também apresenta-se como alvo na replicação viral (Sasvari *et al.*, 2011; Hwang *et al.*, 2013). Foi demonstrado que eEF1B interage diretamente com o domínio metiltransferase de RdRp do vírus TMV (Hwang *et al.*, 2013).

A atuação dos fatores de tradução na biossíntese de proteínas nos últimos anos também tem sido estudada na sua correlação com a atividade fotossintética, mas os mecanismos que regem essa comunicação entre os processos permanecem pouco elucidados (Boex-Fontvieille *et al.*, 2013). A análise de expressão diferencial, realizada nesta investigação, demonstra que LIMYB também regula a expressão de proteínas relacionadas a montagem do complexo PSII (Figura 7A, Tabela S10). Estudos de proteômica descrevem mudanças no padrão de fosforilação de proteínas ribossomais e fatores de iniciação de eucarióticos sob efeito do processo fotossintético (Piques *et al.*, 2009; Reinbothe *et al.*, 2010). Estas observações fornecem uma conexão direta para a atuação de LIMYB na regulação de genes relacionados à fotossíntese e

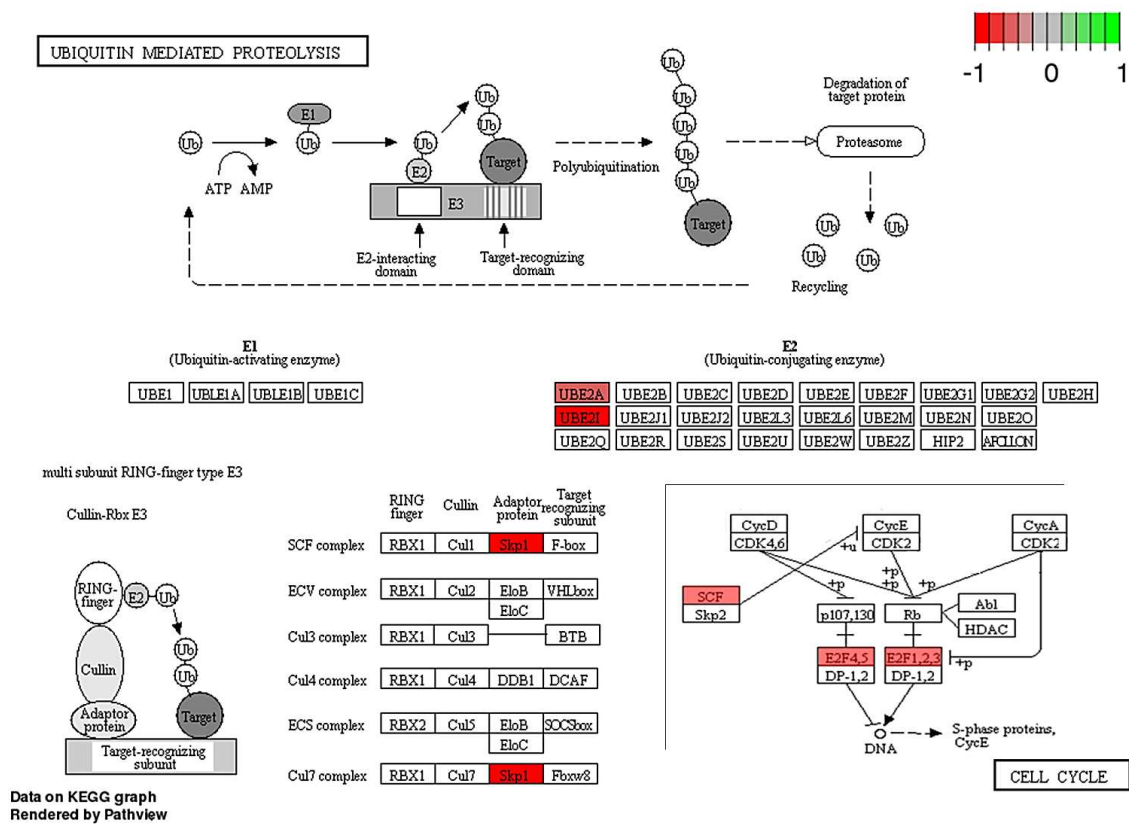
à tradução. A distinção recente de padrões de fosforilação em componentes ribossômicos, influenciados pela luz, mostrou um aumento significativo do nível de fosforilação do RPS6, RPL29, RPS14A e RPL13D, sendo estas últimas as mais responsivas (Boex-Fontvieille *et al.*, 2013). Um ponto crucial deste controle, parece estar baseado na fosforilação da proteína ribossomal S6 (RPS6), visto que o grau de fosforilação deste constituinte afeta a formação e o ritmo da maquinaria de tradução (Mehuyas, 2008). RPS6 é fosforilado pela cinase S6 (S6K), que depende da fosforilação de outras cinases, como PDK1 e o complexo TOR (Chang, 2005). Em cada uma dessas etapas de montagem da maquinaria de tradução, a fosforilação dos eIFs é crucial para o controle da iniciação e existem múltiplos sítios de fosforilação nessas proteínas, que podem ativar ou reprimir a tradução (Mayberry *et al.*, 2011). Dentre os fatores de iniciação da tradução, destacados neste trabalho, eIF3 e constituintes ribossomais, tais como S6 e RPL13, foram encontrados anteriormente em correlação a influência pela luz, sendo ambos fosforilados, em resposta significativa à fotossíntese (Boex-Fontvieille *et al.*, 2013). Outro ponto comum que liga a regulação de LIMYB aos componentes da fotossíntese, designadamente ao PSII, é que proteína PsbO, componente do CEO (complexo de evolução de oxigênio, envolvida na formação deste complexo, sendo conhecida como uma das principais proteínas fotossintéticas capaz de influenciar o crescimento e outros processos de desenvolvimento nas plantas (Gururani *et al.*, 2012; Gururani *et al.*, 2015).



Dados sobre o controle da fotossíntese têm relevante importância sobre a compreensão do desenvolvimento, para melhorias na assimilação de biomassa nas culturas agrícolas. Muito importante neste quesito são as moléculas componentes do complexo coletor de luz I e II (LHC I e LHC II), que na análise de expressão diferencial foram identificados como alvos de repressão de LIMYB (Figura 7B, Tabela S11). Estes complexos são organizados em estruturas de anéis com espaçamento e orientação molecular de modo que estas variações originem diferentes escalas de transferência energia, sendo alvos da proteína de revestimento de *TMV*, para uso desta energia no sucesso do vírus (Ying-Zhong *et al.*, 2008). Estudos de expressão gênica realizados em *Phyllostachys edulis* na resposta à alta luz identificaram mudança nos padrões globais de expressão gênica, a partir de observações em LHC e PsbS (subunidade do PSII), com destaque para genes e fatores de transcrição, como MYB e bZIP, que atuam na regulação do metabolismo fotossintético, controlando a expressão proteica em função da alta e ou baixa iluminosidade (Zhao *et al.*, 2016). Essas observações, associadas aos resultados desta investigação, expandem o eixo da atuação de LIMYB, em torno de processos de desenvolvimento, e também como alvos de infecção viral por geminivirus.

A repressão de genes relacionados a genes da montagem de complexos de ubiquitinação (Figura 8; Tabela S12), envolvidos numa série de ações celulares, relata mais um elemento com função no desenvolvimento, cujo a repressão por LIMYB, pode também estar associada na defesa contra geminivirus, relacionado ao fato de que estas maquinarias são importantes alvos para que a infecção viral ocorra de forma completa (Lozano-Durán *et al.*, 2011). As modificações pós-tradicionais de proteínas por ligação de ubiquitina alteram funções proteicas e regulam uma série de processos, incluindo, desenvolvimento, ciclo celular e respostas a estresse biótico e abiótico, vias de repressão de sinalização hormonal, degradação de fatores de transcrição específicos e regulação da biossíntese de hormônios (Aaron & Estelle, 2017; Castro *et al.*, 2012; Marino *et al.*, 2006). Um importante alvo durante a invasão de geminivírus, é o complexo SCF (*SKP1*, *CUL1/CDC53*, *F box proteins*) E3 ligase, uma vez que proteínas virais podem interferir em sua função para

favorecer a infecção (Lozano -Durán *et al.*, 2011; Zhou *et al.*, 2017). Estudos de silenciamento gênico de ASK2 (*S-phase kinase - associated protein 1*), que codifica SKP1-like, componente de diferentes complexos SCFs, mostraram que a infecção por TYLCSV (*Tomato yellow leaf curl Sardinia vírus*) é acometida (Lozano-Durán *et al.*, 2011). Em *Arabidopsis*, SKP1-like, desempenham importantes papéis em um grande número de processos celulares (Liu *et al.*, 2004; Umezawa *et al.*, 2004).



**Figura 8. Overexpressão de LIMYB *down* regulando genes de complexos de ubiquitinação, e fatores envolvidos na regulação do ciclo celular.** A proteína demonstrou reprimir a expressão de genes necessários a montagem do complexo de ubiquitinação, um alvo da infecção viral, visto a participação em processos de modulação pós-traducional que auxiliam no sucesso do contágio. No canto inferior esquerdo da imagem são representados fatores de transcrição E2F4 e E2F1, que atuam na regulação do ciclo celular, também alvos de infecção.

O complexo SCF é regulado também por derrubilação na subunidade culina (*CUL1*), que está presente em demais complexos de ubiquitinação. Este

processo é realizado por outro complexo, denominado CSN (COP9 signalosome; Hotton & Callis, 2008). Sendo assim, a repressão de genes associados ao complexo SCF por LIMYB se destaca pelo fato de que a proteína TrAP de geminivírus intervém com a atividade de derrubação do complexo CSN sobre Culina 1, por meio da sua interação com CSN5A (COP9 signalosome 5A; Lozano-Durán *et al.*, 2011). A interação TrAP-CSN altera processos celulares regulados pelos complexos SCF, incluindo a sinalização ao jasmonato (Lozano-Durán *et al.*, 2011), padrão de expressão gênica, inibem as vias de morte celular, alteram o fluxo de macromoléculas, interferindo com a sinalização celular e turnover de proteínas para desviar ou dificultar as defesas do hospedeiro e a sinalização hormonal (Hanley-Bowdoin *et al.*, 2013). Considerando que os complexos SCF ligases têm papel chave na regulação de muitos mecanismos celulares, a competência de geminivírus em controlar estes complexos resulta em uma poderosa tática para modulação dos processos do hospedeiro (Hanley-Bowdoin *et al.*, 2013).

A tentativa de domínio dos processos de ubiquitinação por invasores em plantas tem sido relatada em diversos estudos, estando relacionada a susceptibilidade da invasão de organismos capazes de modular complexos que ativam ubiquitina (Ntoukakis *et al.*, 2009). Os dados analisados neste trabalho também revelam que genes, envolvidos na codificação de proteínas componentes do complexo E2 (*Ubiquitin-conjugating Enzyme*), também foram *down* regulados (Figura 6). O processo de ubiquitinação envolve uma cascata enzimática consecutiva de três passos, catalisada por três tipos de enzimas: enzimas ativadoras de ubiquitina (E1), Enzimas conjugadoras (E2) e ubiquitina ligase (E3). Inicialmente a ubiquitina, uma proteína com sequência conservada de 76 aminoácidos, é ativa, de forma dependente de ATP, associando-se a ubiquitina a uma enzima E1. A ubiquitina ativada é transferida de E1 para o resíduo cisteína no sítio ativo de E2. Uma ligase de E3 recruta então uma proteína como substrato para que E2 realize a transferência da ubiquitina de E2 para o substrato, geralmente em um resíduo de lisina (Hershko & Ciechanover, 1998; Zhou *et al.*, 2017). Enzimas do tipo E3 têm tido o foco principal até agora em estudos relacionados a contornos da PTI ou ETI por invasores (Marino *et al.*, 2012; Li *et al.*, 2014). Por outro lado, as enzimas E2 foram por muitas vezes

subestimadas equivocadamente no seu papel assistencial no processo de ubiquitinação, e poucas foram investigadas em plantas. No entanto, as enzimas E2 foram descobertas no agenciamento da processabilidade e topologia da formação de cadeia de poliubiquitina e, de tal modo, determinando o destino de proteínas modificadas (Ye & Rape, 2009; Jiang & Chen, 2011; Zhou *et al.*, 2017). A inibição de componentes deste complexo torna-se uma potencial estratégia durante a invasão de patógenos que os têm como alvo.

O complexo SCF está envolvido na regulação do conjunto proteico ciclina E/cdk2. Ciclina E, assim como outras ciclinas, esta é ubiquitinada pelas proteínas do complexo SCF (Moroy & Geisen, 2003). Durante o ciclo celular, E/cdk2 regula RBR (retinoblastoma-related protein) por fosforilação, desestabilizando sua ligação com o fator de transcrição E2F, o que permite a transcrição de genes alvo deste fator durante a fase G1 (Figura 6). Portanto, pode-se acreditar que é parte da estratégia de defesa que E2F seja incluído no processo de repressão, visto que a inibição de RBR para dar início à fase S do ciclo celular representa uma particularidade de vírus de DNA que contagiam plantas e animais (Hanley-Bowdoin *et al.*, 2013). A proteína Rep de geminivirus interage com RBR e interfere com a interação entre RBR e E2F, liberando o fator de transcrição para ativar a expressão de genes necessários para a replicação do DNA, e conseqüentemente do genoma viral (Argüello-Astorga *et al.*, 2004; Egelkrout *et al.*, 2001; Hanley-Bowdoin *et al.*, 2013).

Os genes *up* regulados, comumente identificados por ChIP-seq (Figura 6B, Tabela S6), foram predominantemente representados em categorias relacionadas ao transporte de ânions, respostas a níveis de nutrientes, bem como escassez de algumas moléculas, como fosfato. Estes resultados podem estar em alinhamento com o quadro de LIMYB atuando na repressão de genes de subunidades de ATPases, tanto da cadeia de transporte de elétrons (Figura S6), como da ATPase que compõe o fotossistema (Figura 7A). Os genes relacionados à movimentação de ânions na célula (Tabela 3), foram representados na categoria mais enriquecida pelas análises de DE. O gene mais enriquecido foi o Transportador ABC (*ATP-binding cassette*), envolvido na resposta ao hormônio ácido abscísico, e outros genes codificantes de transportadores. Que pode estar relacionado a atividade de ABA frente a escassez de compostos como moléculas de fosfato

inorgânico e íons de  $\text{NO}_3^-$  que podem interferir na osmolaridade celular (Roelfsema *et al*, 2012). ABA desencadeia cascatas de sinalização que regulam canais iônicos de membrana por fosforilação, aumentos na concentração de  $\text{Ca}^{2+}$  citosólico livre, e outros mensageiros secundários. Estudos recentes mostraram que a importação de ABA em vacúolos isoladas de células do mesófila é mediada por transportadores do tipo ABC, com movimentação de gradiente de prótons e ATP (Munemasa *et al*, 2015; Jezek & Blatt, 2017).

Tabela 4. Genes da categoria anion transport, GO:0006820. Categoria mais enriquecida entre genes *up* regulados na análise integrativa.

gene_id	width	Fold Enrich.	x_log10 _qvalue	Distance To feature	log2 FC	padj	description
AT1G15520	210	4.85	15.25	-1412	4.79	0.00	ABC transporter
AT2G29330	250	11.67	67.36	-878	2.91	0.00	tropinone reductase
AT5G43350	181	6.20	23.18	-1188	2.82	0.00	phosphate transporter 1
AT4G02270	228	7.17	30.07	-352	2.75	0.00	root hair specific 13 root hair specific
AT4G26200	178	5.23	16.85	-2912	2.24	0.00	1-amino-cyclopropane-1-carboxylate
AT3G18830	174	4.86	14.64	-16	2.08	0.00	polyol/monosaccharide transporter 5
AT2G29750	173	4.62	13.22	-1470	2.07	0.00	UDP-glucosyl transferase 71C1
AT5G49520	202	6.08	22.36	-3873	2.03	0.00	WRKY DNA-binding protein
AT3G23430	183	5.71	19.94	-186	2.02	0.00	N-terminal domain. phosphate 1
AT3G59700	284	4.54	13.57	-284	1.57	0.00	lectin-receptor kinase
AT2G39350	268	3.77	8.63	-179	1.49	0.00	ABC-2 type transporter family ABC
AT3G13380	170	4.99	15.37	-631	1.40	0.00	BRI1-like 3 Similar to BRI,
AT1G25530	333	15.99	138.62	-2303	1.38	0.00	Transmembrane a. acid transporter
AT5G44240	177	4.74	13.92	-650	1.27	0.00	aminophospholipid ATPase 2
AT1G72700	170	5.11	16.10	-1273	1.21	0.00	ATPase E1-E2 dehalogenase-like
AT2G26660	293	11.43	65.15	-510	1.10	0.00	SPX domain gene 2 SPX domain
AT1G58030	250	10.90	72.00	-2360	1.00	0.00	cationic amino acid transporter 2
AT4G37640	196	6.69	26.56	-1185	0.98	0.00	calcium ATPase 2 Encodes
AT1G80760	173	5.23	16.85	-1344	0.98	0.00	NOD26-like intrinsic protein 6
AT3G02880	185	4.86	14.64	-714	0.82	0.00	Leucine-rich repeat protein kinase
AT4G11670	215	7.06	31.05	-1103	0.72	0.00	Protein of unknown function DUF810
AT3G02360	271	6.20	23.18	-356	0.71	0.00	6-phosphogluconate dehydrogenase

Quando é necessário contornar problemas relacionados à limitação de nutrientes, tais como fosfato que são essenciais a diversos processos (Hinsinger, 2001), as plantas utilizam respostas adaptativas que visam restaurar o nível interno de fosfato inorgânico (Pi). Após, ativam mecanismos que aumentam a acessibilidade e a absorção de Pi externo, resultando em respostas que alteram

a arquitetura da raiz, aumento da absorção de Pi, secreção de ácidos orgânicos e secreção de fosfatases (Poirier e Bucher, 2002; Yuan e Liu, 2008; Péret *et al.*, 2011). A sincronização da disponibilidade de Pi com o crescimento e o desenvolvimento da planta é orquestrada por vários fitohormônios, incluindo ABA e etileno (Devaiah *et al.*, 2009; Lei *et al.*, 2011).

Os efeitos regulatórios de ABA dependem principalmente da sua concentração no tecido vegetal, que é determinada pelo equilíbrio entre biossíntese e catabolismo (Nambara e Marion-Poll, 2005). Dentre os genes da categoria de maior intensidade de componentes *up* regulados, é registrado um gene que codifica para UDP-glicosiltransferase (Tabela 4), que em situações de desequilíbrio tem papel importante na homeostase de ABA, para evitar atrasos de desenvolvimento e principalmente fechamento estomático excessivo (Burla *et al.*, 2015; Zeevaart & Yang, 2005). A glicosilação de ABA é um processo reversível catalisado por glucosidases e hidrolisases, que resulta num papel de controle dos níveis de ABA nos tecidos vegetais (Lee *et al.*, 2006; Xu *et al.*, 2012). Também foi constatado que as moléculas glicosiladas de ABA pela atividade de UDP-glicosiltransferase, são transportadas via transportadores ABC (Burla *et al.*, 2015). Em uma das respostas de condições deficientes de fósforo, o Pi pode ser liberado a partir da parede celular, geralmente da fração de pectina, acompanhada por uma redução do nível de ABA e um aumento na produção de etileno, e que o restabelecimento de ABA inibe a remobilização da Pi de parede celular. Outros elementos que podem auxiliar nestas situações são cinases tais como cinases dependentes de  $Ca^{2+}$  (CDPKs), também registradas nesta análise de DE (Figura S7; Tabela S13). As CDPKs ajudam na regulação de canais de íons na membrana plasmática, respostas de defesa a patógenos (Kollist *et al.*, 2014; Munemasa *et al.*, 2015; Jezek & Blatt de 2017), e recentemente tornou-se claro que essas cinases também têm parte na manutenção do tonoplasto (Latz *et al.*, 2013; Wege *et al.*, 2014).

A análise integrativa também constatou um grupo de genes correspondentes a enzimas envolvidas na biossíntese de fenilpropanoides (Figura S8, Tabela S14), que são compostos gerados a partir de esqueletos carbônicos de fenilalanina, e são direcionados tanto a produção de metabólitos secundários, que vão gerar produtos de interesse ao desenvolvimento ou

respostas de defesa a estresses (Vogt, 2010). Esta via de biossíntese, também cataboliza precursores de monolignóis, que são componentes presentes nas estruturas de lignina, conferindo suporte estrutural, e barreira numa estratégia de enrijecimento da parede celular em resposta a patógenos (Boerjan *et al.*, 2003).

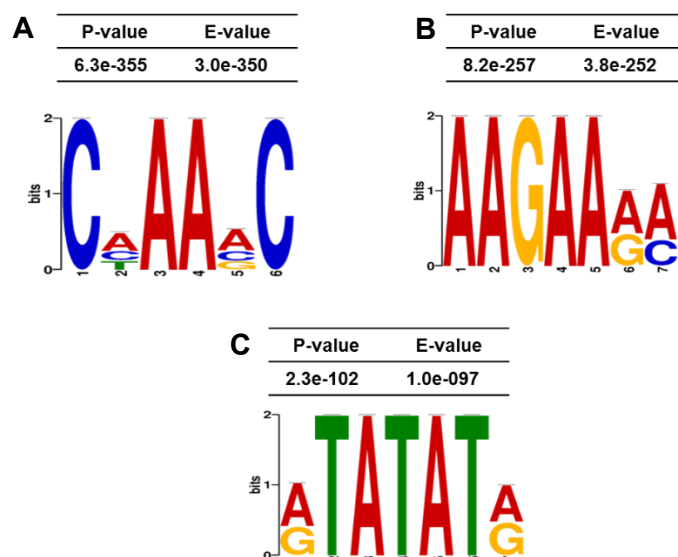
O *begomovirus Tomate yellow leaf curl China (TYLCCNV)* é capaz de beneficiar seu vetor, a mosca branca (*Bemisia tabaci*), através da supressão das defesas de plantas hospedeiras. Esta relação possui caráter mutualístico, e atenua os custos de energia, aumentando assim o desempenho das moscas brancas em plantas infectadas com vírus (Gu *et al.*, 2017). A supressão da produção de ROS, cujo CDPKs são intermediárias na resposta, diminui tanto a resistência ao vírus, como também alivia a inibição do reparo de D1 do PSII, a nível de tradução do mRNA (Nishiyama *et al.*, 2006), que vão acarretar no beneficiamento do vírus com a ativação da tradução. A observação de todos os mecanismos na regulação de LIMYB, estabelece um caráter de repressor a processos de estímulos com fins direcionados a da tradução global, sendo destacado que grande maioria de genes sob repressão inédita desta proteína, tem sido registrados como alvos de infecção viral, o que faz sentido, visto que os vírus tentam dominar processos de desenvolvimento celular para o seu progresso.

#### **4.4 Identificação de cis-elementos conservados e análise de colocalização da ligação do fator de transcrição por cobertura de RNA-seq e CHIP-seq**

A detecção de múltiplos eventos de ligação, resultando em expressão diferencial, despertou o interesse na identificação da dinâmica vinculativa de LIMYB. Inicialmente, realizou-se análise de enriquecimento de motivos dentro dos picos de LIMYB para a descoberta de cis-elementos de ligação de LIMYB, utilizando a plataforma MEME-ChIP (meme 4.9.1). Os motivos foram agrupados pela aplicação de agrupamento hierárquico usando distâncias de motivos, calculadas pelos coeficientes correlacionados Pearson como métrica de

comparação de bases e método de alinhamento Ungapped Smith-Waterman. A intensidade das ligações foi medida a partir do log2 (contagens normalizadas de reads/peaks) e avaliadas no ponto de corte E-valor de  $10^{-5}$  para identificação das sequências.

Os motivos com maior significância e enriquecimento para análise conjunta do total de elementos de sequência detectados por CHIP-seq foram CAAAAC (E-value =  $3,0 \times 10^{-350}$ ) e AAGAAA (E-value =  $3,8 \times 10^{-252}$ ) (Figura 9A, 9B e Figura S10), respectivamente, que são elementos conservados presentes no grupo de reconhecimento por domínios MYB (Gupta *et al.*, 2007; Tom *et al.*, 2011; Katiar *et al.*, 2012).



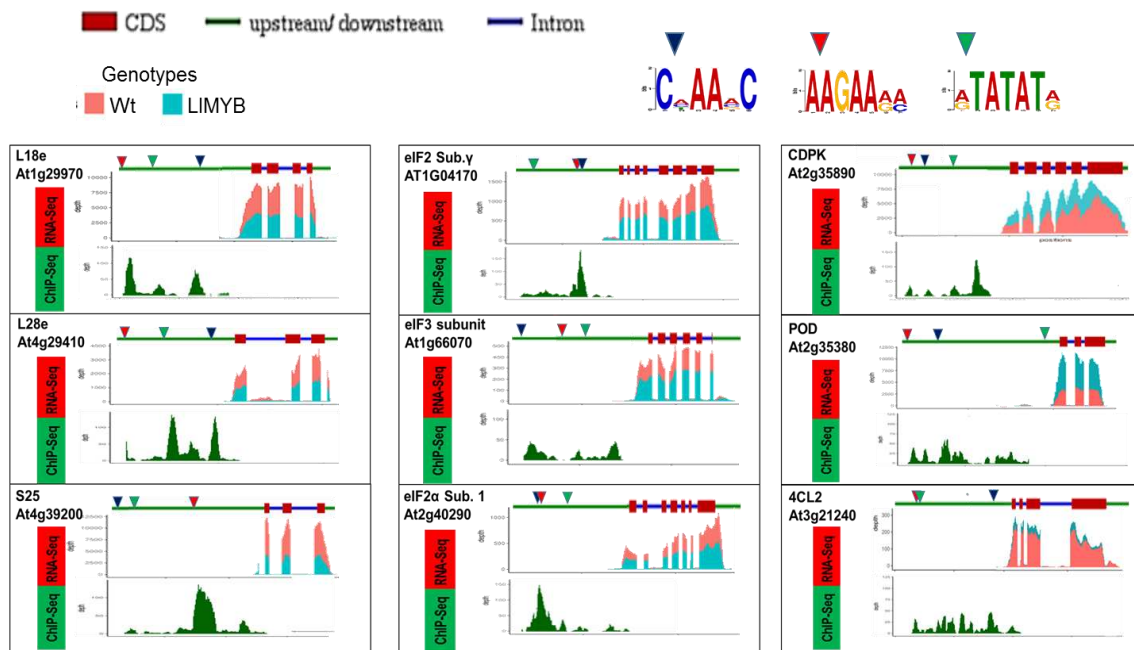
**Figura 9. Análise de *motif enrichment* nos promotores alvos da atuação de LIMYB na regulação de processos biológicos.** Os dados de CHIP-seq utilizados na análise integrativa foram submetidos a análise de cis-elementos regulatórios, para identificar e caracterizar possíveis sítios de atuação de LIMYB e validar sua função como fator de transcrição. (A) e (B) elementos conservados característicos de domínios MYB, e (C) elemento conservado ATATATA, detectados pela ligação de LIMYB a regiões promotoras via CHIP-seq.

A descoberta de elementos cis-regulatórios em regiões promotoras é essencial para entender o padrão de expressão de genes controlados por fatores MYB, como o caso da proteína em estudo. Esses resultados sugerem que um grande conjunto de genes de fato respondem a função do fator em estudo. Outro

motif em alto nível significância e enriquecimento, foi a sequência ATATATA (E-value =  $1.0 \times 10^{-097}$ ) (Figura 9C), com alta semelhança ao elemento regulatório TATA-box, conhecido por constituir um motivo conservado em 30 a 40% dos principais promotores da maioria das espécies (Katti *et al.*, 2000), já bem caracterizado por ser sítio da ligação de fatores basais de transcrição. Prontamente, pode-se sugerir que LIMYB possa exercer competição com estes fatores em seus próprios sítios de atuação, resultando em processos que podem acarretar em repressão total da tradução.

Para esclarecer se de fato LIMYB apresentava especificidade de ligação quanto ao seu pressuposto papel de repressor, separadamente, foram realizadas análises de motivos conservados entre genes *up* e *down* regulados constatados pela DE (Figura S10). Foram encontrados motivos apenas para genes *down* regulados, a partir do ponto corte de significância utilizado. Em seguida foram realizados diagnósticos do tipo *deth* de cobertura entre os dois procedimentos utilizados na análise integrativa, representados por mapeamento de gene, e demonstrando a localização destes elementos cis em seus promotores, para elucidar a atuação de LIMYB frente aos respectivos *locus* afetados. Dentre os genes *down* regulados, foram escolhidos 3 genes ribossomais pelas características já evidenciadas da proteína em estudo (Zorzatto *et al.*, 2015), 3 genes relacionados a fatores de iniciação da tradução, e outros 3 dentre os genes *up* pertencentes ao processo de biossíntese de Fenilpropanoides, que foram encontrados em comum na análise integrativa. Foi utilizado o pacote *Deeptools* contido em software R, com normalização para as relações de intensidade de picos e *hits* de expressão partir dos valores de *Fold enrichment*, para elucidar a relação de *peaks* de ChIP-seq com reads de RNA-seq e a proximidade da ligação da proteína aos elementos detectados, como os mais relacionados com a atividade deste fator (Figura 10).

A anotação dos picos de ligação detectados por ChIP-seq frente aos respectivos genes, submetidos a controle do nível da expressão de transcritos de RNA, fornece uma estrutura para interpretar características que refinam a compreensão da atividade de LIMYB pela capacidade de ligação aos sítios correspondentes a elementos conservados de ligação de fatores de transcrição (Figura S11).



**Figura 10. Análise de cobertura da relação de picos de ligação detectados por ChIP-seq frente aos respectivos promotores de genes diferencialmente expressos nas análises de RNA-seq.** A esquerda, são representados no 1ºquadro, picos de ligação de LIMYB *upstream* a genes codificantes de proteínas ribossomais, e a relação proximidade com os motivos enriquecidos; demonstrando o efeito de *down* regulação em contraste com a planta selvagem. No centro da imagem (2ºquadro) há a representação da atuação de LIMYB sobre promotores de genes codificantes de fatores de iniciação da tradução. A direita, são apresentadas no 3ºquadro, as relações de picos e *hits* de genes *up* regulados em plantas overexpressando LIMYB em contraste com a planta selvagem. Acima dos quadros são apresentadas as cores correspondentes a cada componente das representações.

Uma importante constatação neste trabalho é que a capacidade regulatória de LIMYB inclui um raro e pouco descrito evento celular, visto que durante o desenvolvimento, são pouco conhecidos mecanismos que controlam direta e efetivamente a expressão de constituintes ribossomais (Lio & Qian, 2013). Evidências vem sendo levantadas sobre a influência da heterogeneidade do ribossomo prevalente entre espécies, em diferentes estágios de desenvolvimento e em diferentes tecidos (Vesper *et al.*, 2011). A composição variante dos ribossomos, tanto em rRNA, quanto em proteínas ribossômicas, fornece um desafio aos mecanismos regulatórios no controle da maquinaria de tradução (Lio & Qian, 2013). Em observação do alinhamento dos picos detectados por ChIP-seq com os motivos encontrados nota-se uma afinidade correspondente ao p-valor encontrado para cada motivo, que representa também nesta análise uma relação de afinidade. Reforçando esta análise, estes

motivos encontrados (Figura 9A, B), foram alinhados a um motif específico dentro da categoria funcional de constituintes ribossomais (GO:0003735) detectados em DE (Figura S10, Tabela S4).

A detecção por CHIP-seq, também constatou a ligação de LIMYB a sítios de proteínas sem qualquer notação de sua função ou classificação em bancos de dados (Tabela S15). A partir desta observação foi utilizado um estudo em modelo de aprendizado de máquina (*machine learning*), com seleção para proteínas com caráter de constituintes ribossomais, no âmbito da predição de novos candidatos a proteínas ribossomais pela plataforma RAMA (*A machine learning approach for ribosomal protein prediction in plants*). O software usa para predição um grupo de características fundamentais da cadeia lateral de aminoácidos: aromático, negativamente, carregado, não polar, substituto, polar, sem carga, hidrofóbico, além de massa molecular, volume e comprimento. A predição é feita em 2 etapas, com 6 passos utilizando 6 *machine learnings* em cada etapa, aplicando um procedimento de duas etapas para classificar um conjunto de proteínas com função desconhecida como RP: Primeiro, os modelos criados a partir dos conjuntos de dados de RPs / RP não são aplicados; segundo, as proteínas classificadas como RPs são dadas como entrada para os modelos criados a partir dos conjuntos de dados de RPs / histonas. Somente as proteínas consideradas como positivas nesses dois filtros de classificação são relatadas como RPs. Feita a análise, evidenciou-se que estes elementos desconhecidos possuem forte caráter de proteínas ribossomais (Tabela S15), sendo preditos da mesma forma em diferentes espécies de plantas, além de *Arabidopsis*. Esta constatação reforça a afinidade de LIMYB frente a repressão de proteínas ribossomais, e sugerindo que de fato os sítios de ligação de LIMYB não são aleatórios, conforme também demonstrado na análise de enriquecimento de motivos.

A detecção da *down* regulação de eIFs também reforça o potencial repressor de LIMYB no quadro de coordenador de processos de defesa e ou desenvolvimentos direcionados a tradução global. São encontrados poucos registros em literatura da repressão de eIFs. Estudos em mamíferos descreveram estruturas de RNA G-quadruplex (GQ) atuam como reguladores de uma variedade diversa de processos celulares, e relataram a capacidade desta

estrutura na repressão de fatores iniciadores da tradução, em estudo direcionado para a inibição de eIF4E (Yang *et al.*, 2016). Nas análises de *deth* de cobertura realizadas neste trabalho para os eIFs selecionados (Figura 10A), observa-se que os motivos conservados (Figura 9A, B) estão mais próximos a localização espacial dos picos ligação. Estes motivos também foram apresentaram alinhamento com *motif* específico a categoria de GO:0006412, que corresponde a processos de tradução (Figura S17), corroborando com a classificação de LIMYB como repressor em uma rede processos de desenvolvimentos que envolvem mecanismos de tradução.

O estudo aqui realizado ganha forte significado frente a importância na agricultura da compreensão de mecanismos que regem o crescimento e desenvolvimento vegetal, para facilitar o desenvolvimento de estratégias no campo da engenharia genética para produção de novos transgenes resistentes a doenças que geram grandes perdas econômicas, e possivelmente diminuir a necessidade de componentes químicos contra pragas, cada vez mais agressivos ao meio ambiente e a saúde humana. O uso de agrotóxicos e fertilizantes sintéticos, embora ajudem no aumentar da produção agrícola, são causadores de danos ao meio ambiente e a saúde das pessoas que o ingerem ou ficam sob exposição, seja por acúmulo de resíduos tóxicos em alimentos, contaminação da água e do solo, intoxicação de rebanhos e produtores rurais, aparecimento de pragas resistentes, a interrupção do sistema de controle biológico por inimigos naturais, ocasionando surtos de insetos-praga (Correa & Salgado *et al.*, 2011).

## 5. Conclusão

A partir dos dados obtidos neste trabalho foi possível assinalar possíveis funções de LIMYB, baseado em estudos de expressão diferencial induzida por LIMYB correlacionados à análise integrativa de RNA-seq/ChIP-Seq. Os resultados dessa investigação permitiram a identificação de cis-elementos conservados em promotores de genes alvos de LIMYB, envolvidos tanto em processos de desenvolvimento como em estratégias de defesa antiviral. Os processos de desenvolvimento sob controle de LIMYB parece estar sincronizado com o controle direto de LIMYB na tradução global. É possível sugerir que evolutivamente na defesa contra a infecção viral, as células vegetais desenvolveram um mecanismo de resposta que ativasse um regulador da tradução, capaz de reprimir conjuntamente a síntese do máximo possível de elementos tidos como alvos virais para suscetibilidade. Em contrapartida, durante a evolução, os geminivirus desenvolveram estratégias de virulência produzindo a proteína NSP, que além de suas funções no transporte do vírus, é capaz de inibir a via desencadeada pelo receptor NIK1 (*NSP-interacting Kinase*), cuja cascata de fosforilação desencadeada tem como alvo *downstream* LIMYB (Mariano *et al.*, 2004; Fontes *et al.*, 2004; Zorzatto *et al.*, 2015). A não superação dessa barreira por geminivirus seria certamente um fator culminante na manutenção da existência do vírus.

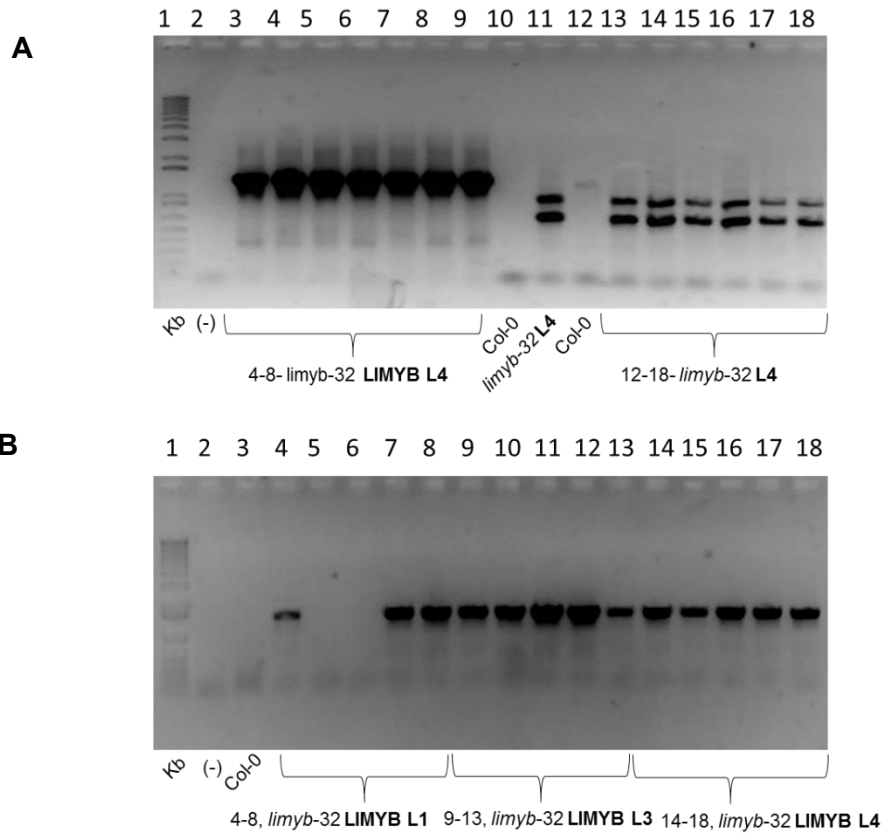
Além da repressão de genes envolvidos na maquinaria de tradução, também foi demonstrado que LIMYB reprime a expressão de componentes chave do aparato fotossintético. Conseqüentemente, é possível que LIMYB seja um link molecular que coordena em sincronia a supressão global de tradução com a inibição de fotossíntese.

A identificação de novos sítios de ligação de LIMYB a uma série de promotores de genes de proteínas ribossomais, fatores de iniciação da tradução e alongação, componentes de montagem dos complexos fotossintéticos e cadeia de transporte de elétrons, complexos de ubiquitinação demonstra o forte potencial desta proteína na remodelação das funções celulares. Levando em consideração estes aspectos é possível classificar LIMYB como um regulador chave no processo de tradução em estágios do desenvolvimento. Em busca de

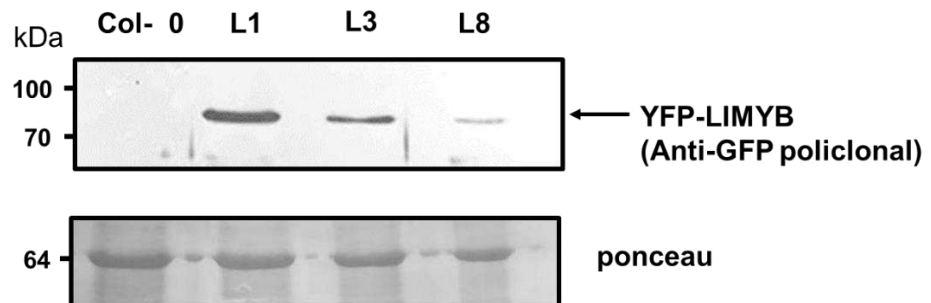
reforçar este pressuposto, torna-se viável a realização de análises que possam validar ainda mais os dados obtidos na análise integrativa quanto à intensidade de ligação de LIMYB aos seus respectivos sítios.

## 6. Material Suplementar

### 6.1 Figuras suplementares



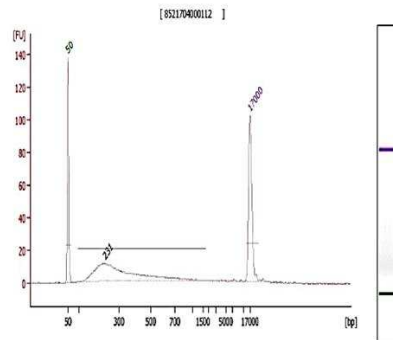
**Figura S1. Diagnóstico por PCR da incorporação do transgene na linhagem nocaute.** (a) Amplificação da região codificante de transgene LIMYB no transformante L4 e confirmação da condição nocaute de *limyb-32* L4. (b) Confirmação da incorporação do transgene 35S::LIMYB em três transformantes independentes.



**Figura S2. Detecção de YFP-LIMYB nas linhagens transgênicas, detectada por western blotting.** Acúmulo de YFP-LIMYB nas linhagens transgênicas L1, L3 e L8. Proteínas totais foram extraídas de Col-0, L1, L3 e L8 (transformante independente adicional), separadas por eletroforese, transferidas para membranas e sondadas com anticorpo anti-GFP policlonal.

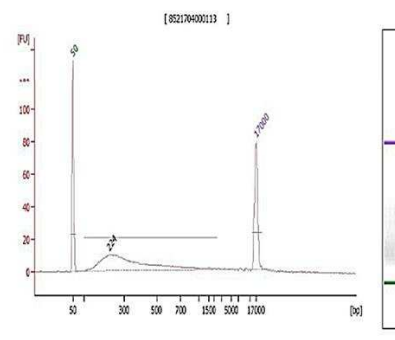
**A**

(1) Sample name: L1-2



Overall Results for sample 11 :  
Number of peaks found: 1

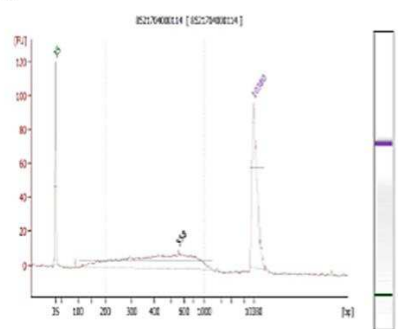
(2) Sample name: L1.D1.2



Overall Results for sample 12 :  
Number of peaks found: 1

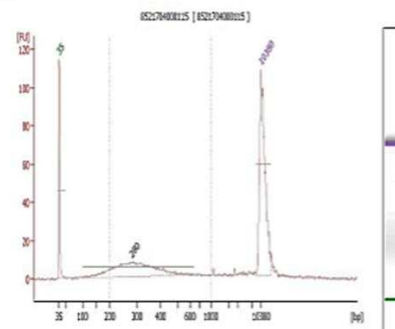
**B**

(3) Sample name: L3



Overall Results for sample 9 :  
Number of peaks found: 1  
Noise: 0.5

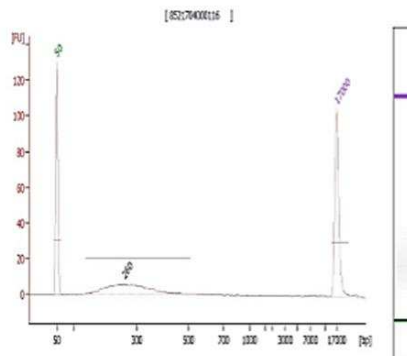
(4) Sample name: L3-2



Overall Results for sample 10 :  
Number of peaks found: 1  
Noise: 0.4

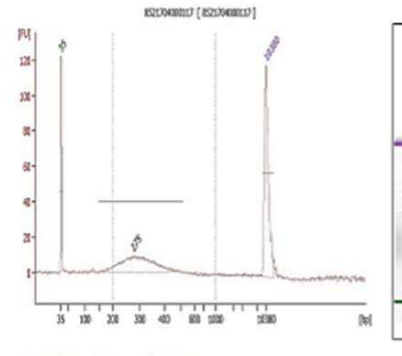
**C**

(5) Sample name: L4



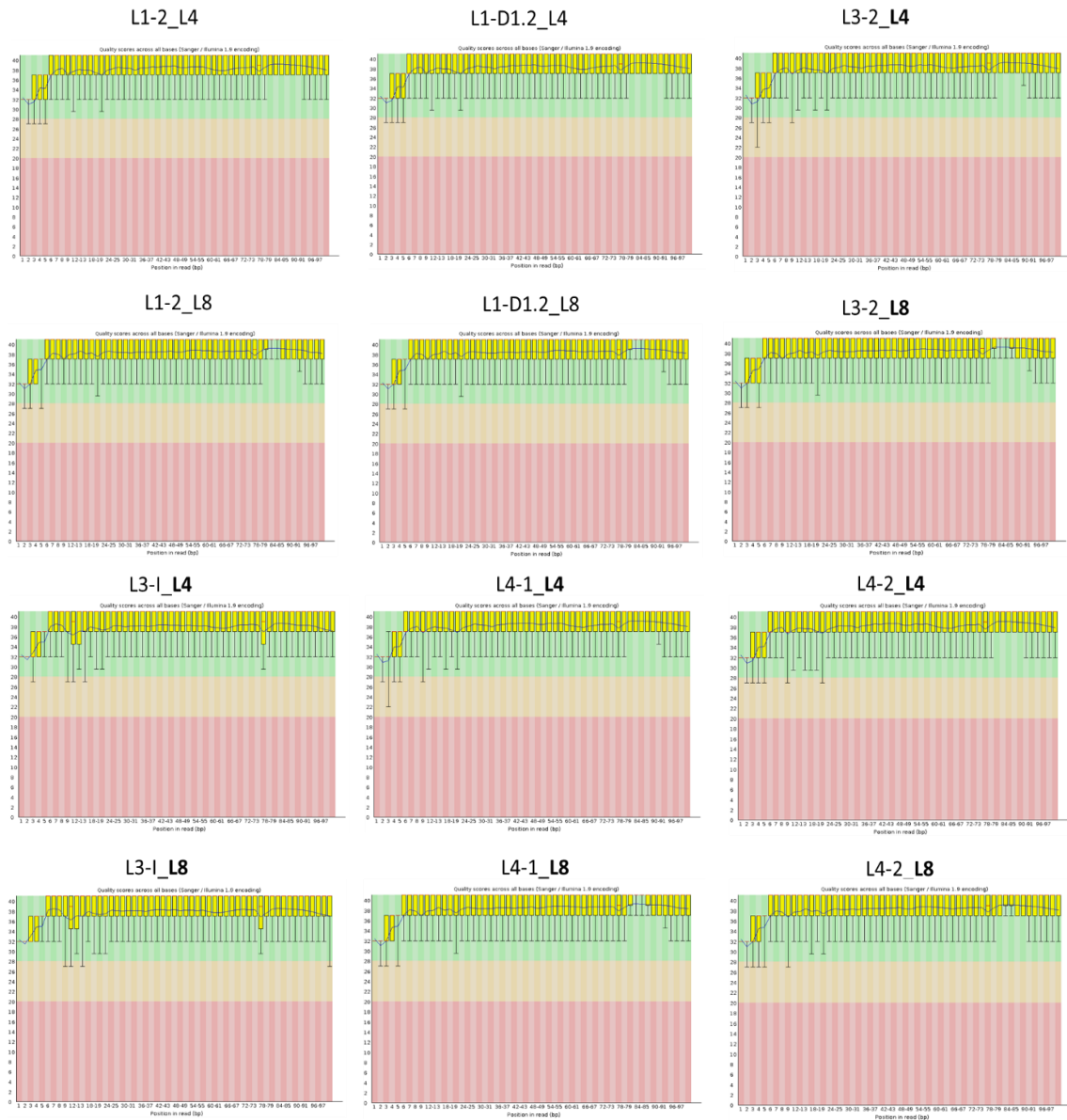
Overall Results for sample 1 :  
Number of peaks found: 1

(6) Sample name: L4 - 2



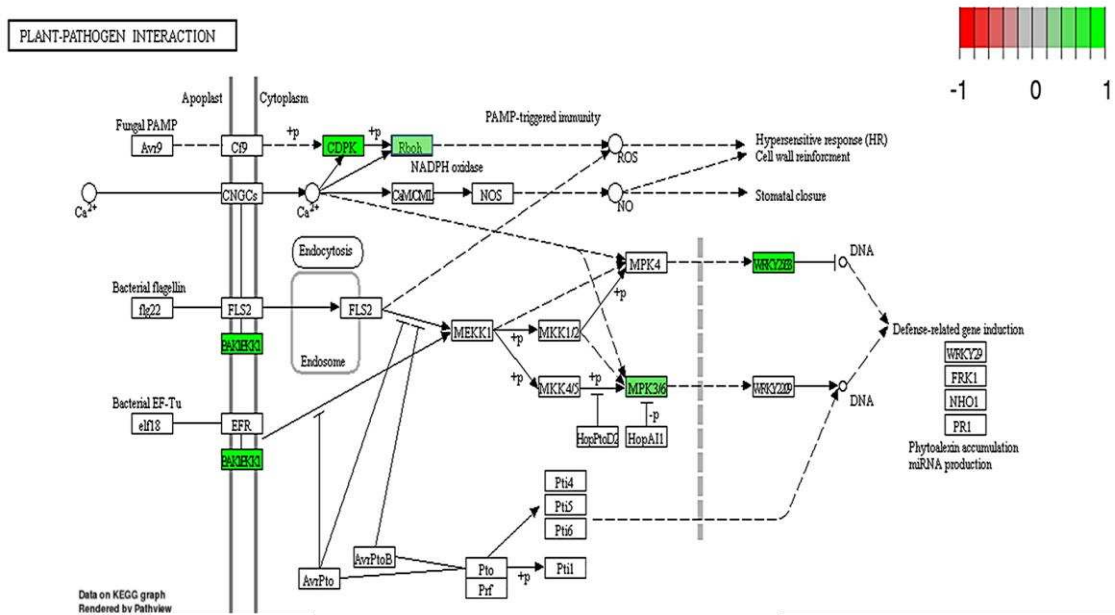
Overall Results for sample 11 :  
Number of peaks found: 1  
Noise: 0.8

**Figura S3. Análise por eletroforese digital para diagnóstico do ChIPed-DNA enriquecido.** Foi utilizado na análise chip eletrônico de alta sensibilidade em plataforma Bioanalyzer Agilent 2100. **(A)** São apresentados os padrões de fragmentação das amostras L1-2 (L1-1), e L1.D1 (L1-2), apresentando fragmentos de 100 a 1000 pb com pico maior de concentração em torno de 300 pb. **(B)** Padrões de fragmentação das amostras L3-1 e L3-2, com fragmentos de 100 a 1000 pb, apresentando picos de maior concentração para fragmentos em torno de 500pb para L3-1 e 300 pb para L3-2. **(C)** Padrão de fragmentação das amostras de L4-1 e L4-2, apresentando picos de maior concentração em torno de 250 a 300 pb, respectivamente.

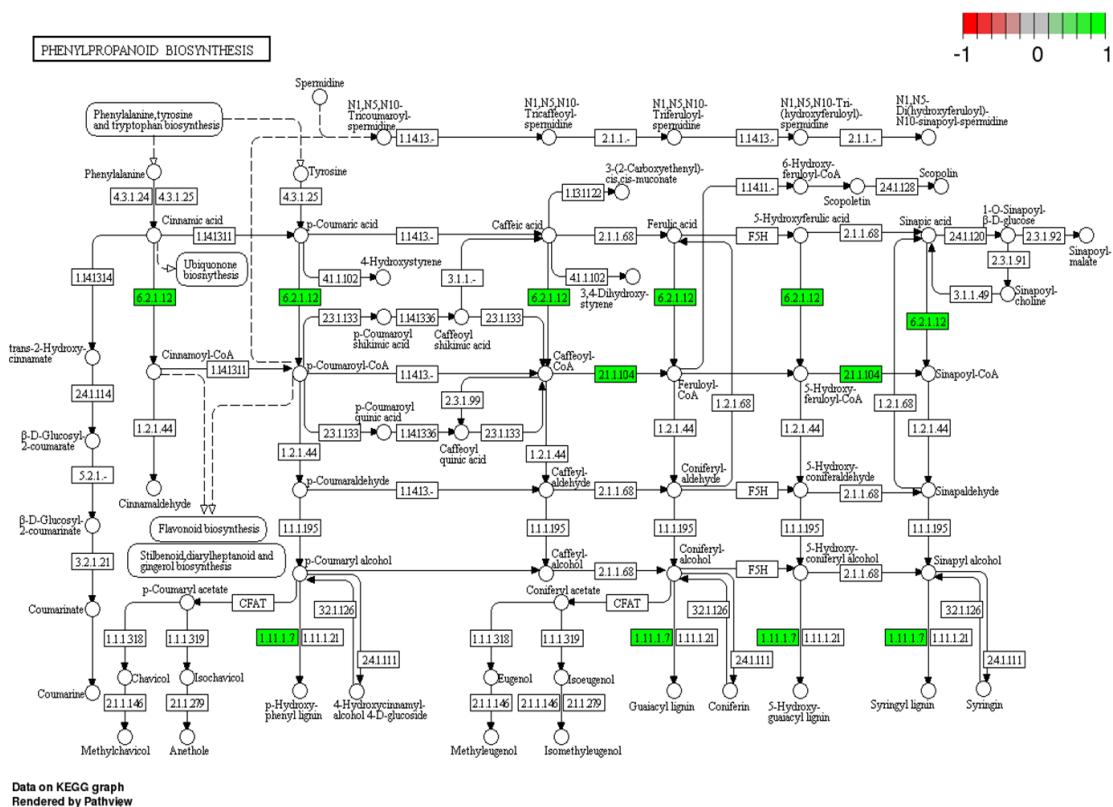


**Figura S4. Análise via FastQC da qualidade de bases (Per base sequence quality; Dados já filtrados) das amostras sequenciados por CHIP-seq.** São apresentadas as qualidades dos pares de base das amostras sequencias em pares de Lanes (Representado pelas letras em negrito, no final de cada identificação acima da figura). Os dados foram considerados confiáveis, acima da “razão de 30”, correspondente a área verde do gráfico.

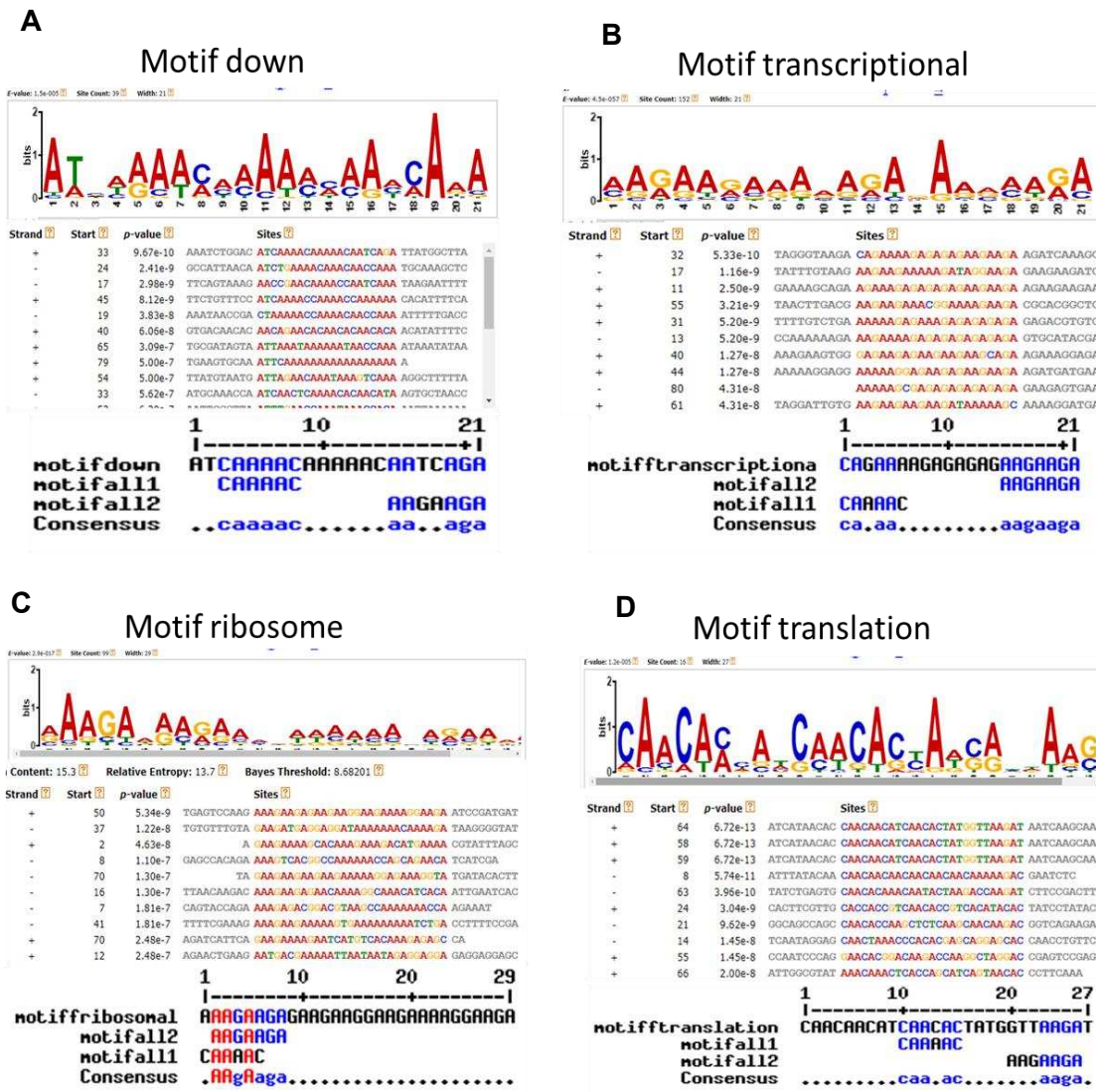




**Figura S7. Representação via KEGG, de genes *up* regulados em linhagens overexpressando em LIMYB, presentes na categoria de interação planta patógeno. Dados de significância e intensidade de ativação podem ser vistos na tabela suplementar 13)**

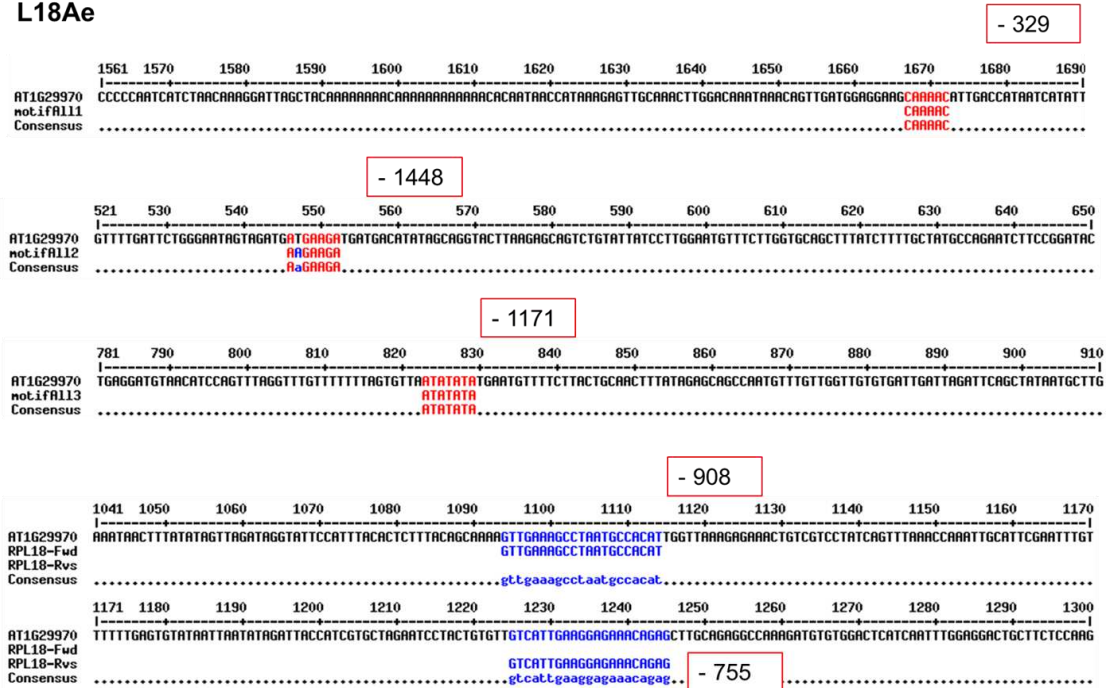


**Figura S9. Representação via KEGG, de genes *up* regulados em linhagens overexpressando em LIMYB, presentes na categoria de biossíntese de Fenilpropanoides. Dados de significância e intensidade de ativação podem ser vistos na tabela suplementar 13)**



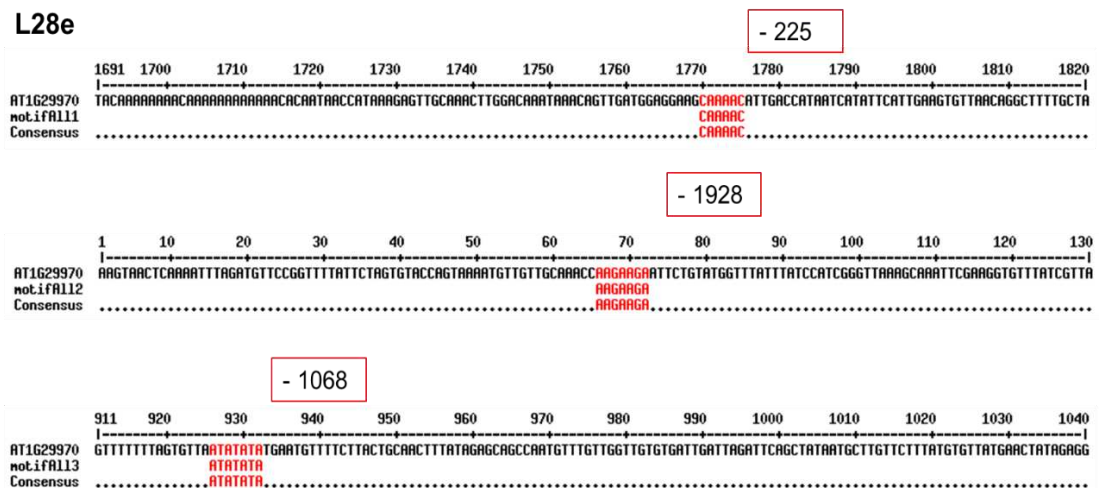
**Figura S10. Análise de motivos enriquecidos por MEME-ChIP, para detecção de motivos relacionados as funções categorizadas.** Foi realizado pareamento de bases dos motivos relacionados a cada categoria com os motivos cis conservados encontrados para a atividade de LIMYB entre todos o genes estudados, utilizando a plataforma Multialin (*Multiple sequence alignment with hierarchical clusterin*). Na parte superior de cada imagem é mostrado o motivo e suas possíveis variações de base. A mais significativa foi utilizada para o passo de pareamento. **(A)** Pareamento de bases dos motivos encontrados com um motivo encontrado exclusivamente dentro do grupo de genes *down* regulados. **(B)** Pareamento de bases dos motivos conservados encontrados para LIMYB, presentes na composição de motivo específico a regiões de fatores de transcrição (GO:0006355) detectados via CHIP-seq. **(C)** Pareamento de base dos motivos mais significantes cis-conservados no GO:0003735, correspondente a constituintes ribossomais. **(D)** Pareamento de bases para rastrear a presença dos elementos cis-conservados encontrados para LIMYB, dentro de um motivo específico a processos de regulação da tradução.

## L18Ae



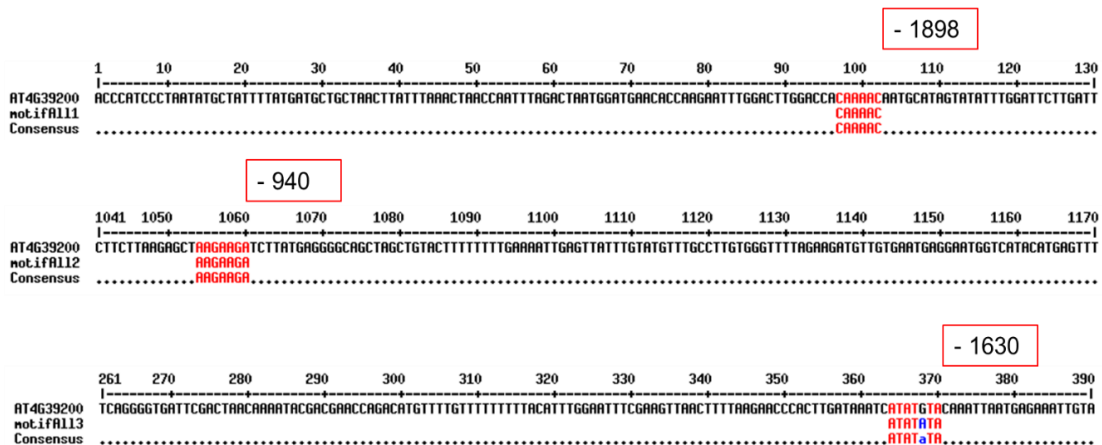
**Figura S11. Pareamento de bases dos elementos cis-conservados detectados por motif enrichment aos promotores de estudo de cobertura CHIP-seq/RNA-seq.** Os motivos são localizados e comparados quanto a proximidade dos picos detectados elucidados na figura 10. Em vermelho são destacados a posição upstream ao início do sítio de transcrição do gene. Na marcação Azul abaixo são demonstrados o alinhamento dos primers utilizados para amplificação dos 150 pb utilizados na constatação do procedimento de CHIP.

## L28e



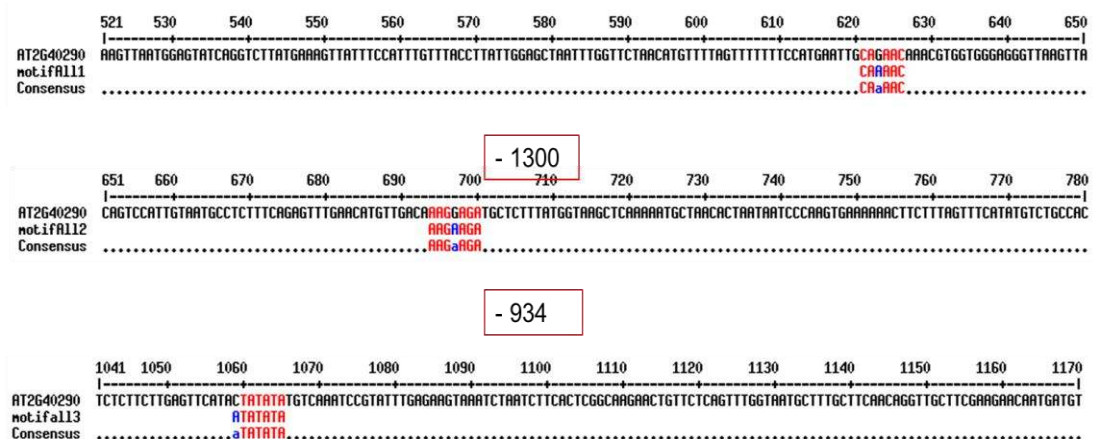
**Figura S12. Pareamento de bases dos elementos cis-conservados detectados por motif enrichment aos promotores de estudo de cobertura CHIP-seq/RNA-seq.** Os motivos são localizados e comparados quanto a proximidade dos picos detectados elucidados na figura 10. Em vermelho são destacados a posição upstream ao início do sítio de transcrição do gene.

## S25



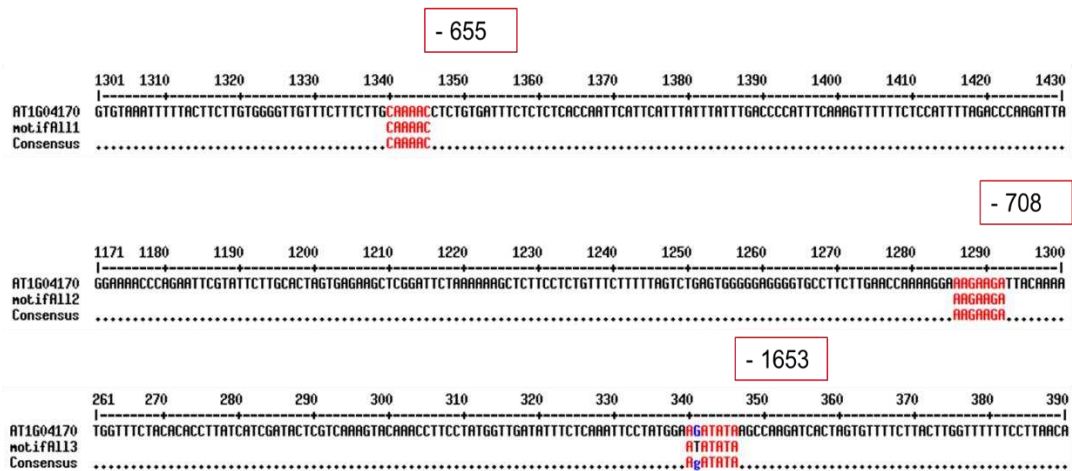
**Figura S13. Pareamento de bases dos elementos cis-conservados detectados por motif enrichment aos promotores de estudo de cobertura ChIP-seq/RNA-seq.** Os motivos são localizados e comparados quanto a proximidade dos picos detectados elucidados na figura 10. Em vermelho são destacados a posição upstream ao início do sítio de transcrição do gene.

## eIF2a sub.1



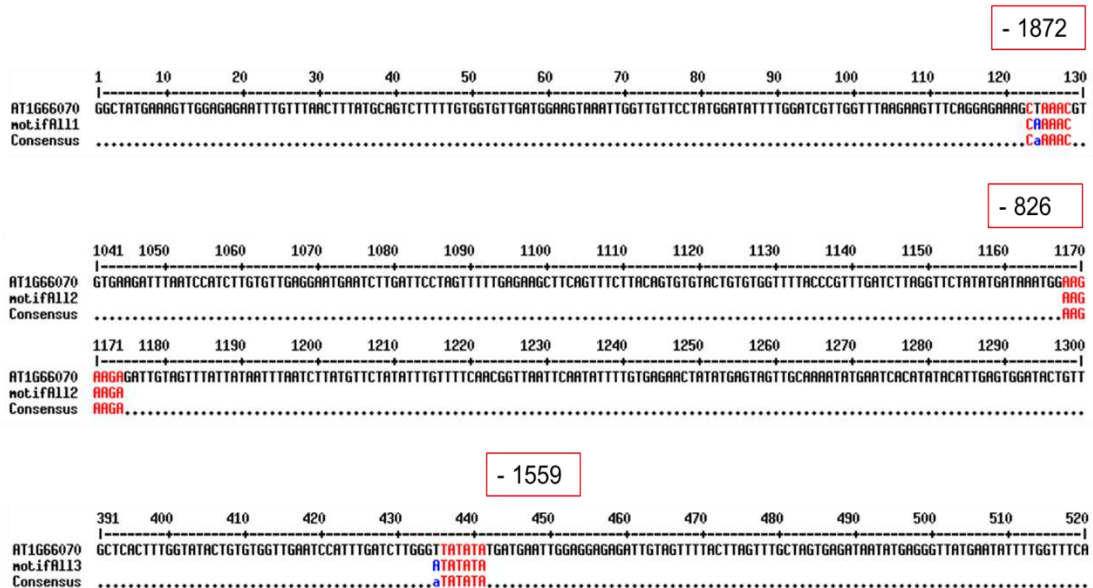
**Figura S14. Pareamento de bases dos elementos cis-conservados detectados por motif enrichment aos promotores de estudo de cobertura ChIP-seq/RNA-seq.** Os motivos são localizados e comparados quanto a proximidade dos picos detectados elucidados na figura 10. Em vermelho são destacados a posição upstream ao início do sítio de transcrição do gene.

### eIF2 sub. gamma



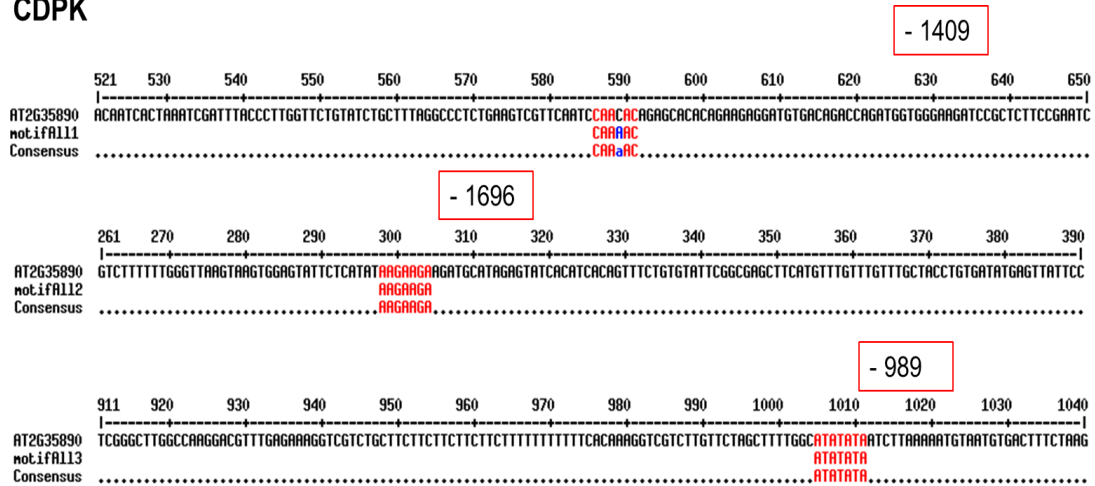
**Figura S15. Pareamento de bases dos elementos cis-conservados detectados por motif enrichent aos promotores de estudo de cobertura ChIP-seq/RNA-seq.** Os motivos são localizados e comparados quanto a proximidade dos picos detectados elucidados na figura 10. Em vermelho são destacados a posição upstream ao início do sítio de transcrição do gene.

### eIF3 sub



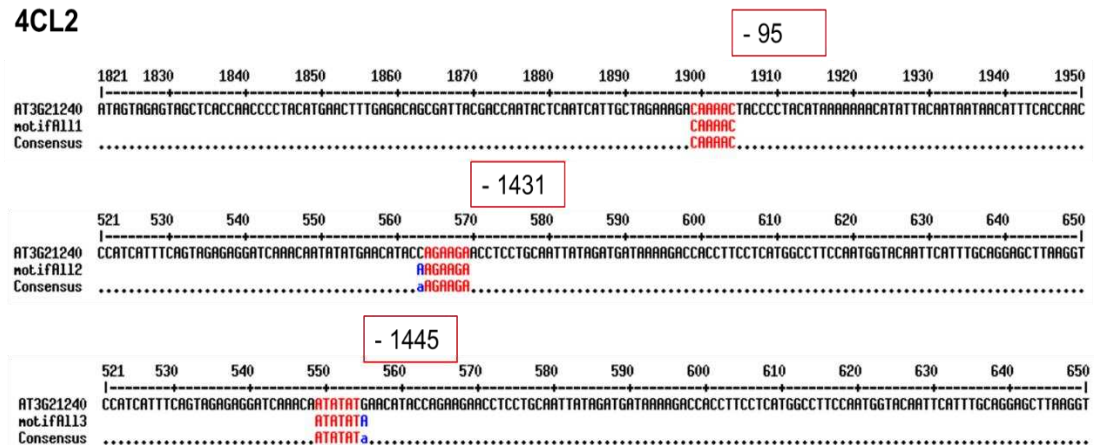
**Figura S16. Pareamento de bases dos elementos cis-conservados detectados por motif enrichent aos promotores de estudo de cobertura ChIP-seq/RNA-seq.** Os motivos são localizados e comparados quanto a proximidade dos picos detectados elucidados na figura 10. Em vermelho são destacados a posição upstream ao início do sítio de transcrição do gene.

## CDPK



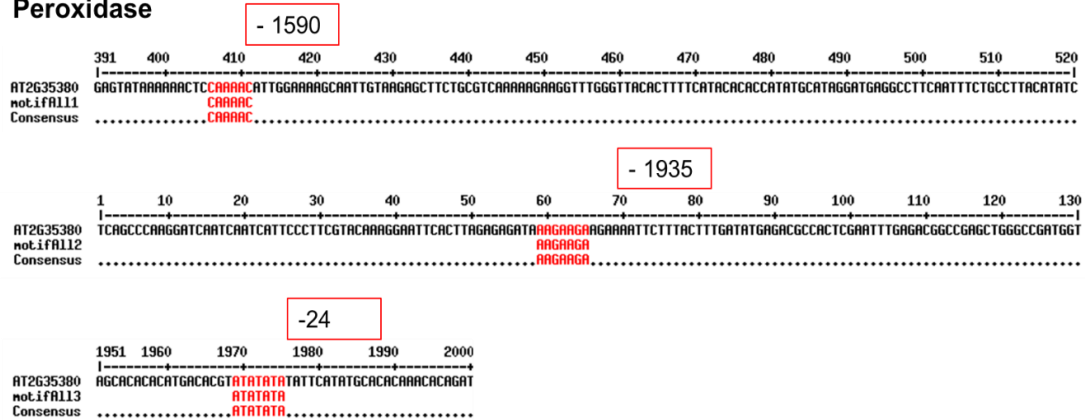
**Figura S17. Pareamento de bases dos elementos cis-conservados detectados por motif enrichent aos promotores de estudo de cobertura ChIP-seq/RNA-seq.** Os motivos são localizados e comparados quanto a proximidade dos picos detectados elucidados na figura 10. Em vermelho são destacados a posição upstream ao início do sítio de transcrição do gene.

## 4CL2

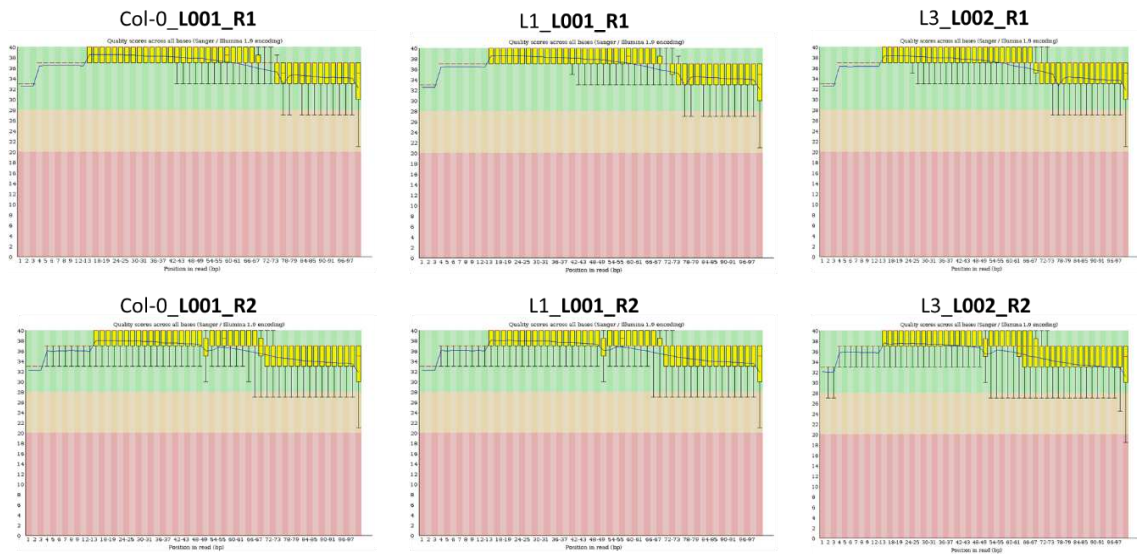


**Figura S18. Pareamento de bases dos elementos cis-conservados detectados por motif enrichent aos promotores de estudo de cobertura ChIP-seq/RNA-seq.** Os motivos são localizados e comparados quanto a proximidade dos picos detectados elucidados na figura 10. Em vermelho são destacados a posição upstream ao início do sítio de transcrição do gene.

## Peroxidase



**Figura S19. Pareamento de bases dos elementos cis-conservados detectados por motif enrichent aos promotores de estudo de cobertura CHIP-seq/RNA-seq. Os motivos são localizados e comparados quanto a proximidade dos picos detectados elucidados na figura 10. Em vermelho são destacados a posição upstream ao início do sítio de transcrição do gene.**



**Figura S20. Análise via FastQC da qualidade de bases (Per base sequence quality; Dados já filtrados) das amostras sequenciados por RNA-seq. São apresentadas as qualidades dos pares de base das amostras sequencias em pares de Lanes (Representado pelas letras em negrito, no final de cada identificação acima da figura). Os dados foram considerados confiáveis, acima da "razão de 30", correspondente a área verde do gráfico.**

## 6.2 Tabelas Suplementares

**Tabela S1. Genes down regulados sob a overexpressão de LIMYB via RNA-seq**

id	log2foldchang	padj	description
AT1G01080	-1.154484369	3.34923E-17	RNA-binding (RRM/RBD/RNP motifs) family protein
AT1G01100	-1.833444827	1.19032E-60	60S acidic ribosomal protein family
AT1G01170	-1.2643295	4.9661E-18	Protein of unknown function (DUF1138)
AT1G01620	-1.373376942	4.48864E-40	plasma membrane intrinsic protein 1C
AT1G02280	-1.053074758	4.66938E-25	translocon at the outer envelope membrane of chloroplast
AT1G02610	-1.302742452	7.35333E-11	RING/FYVE/PHD zinc finger superfamily protein
AT1G02640	-1.568946312	2.01005E-15	beta-xylosidase 2
AT1G02690	-1.008934958	4.58144E-07	importin alpha isoform 6
AT1G02730	-1.020611075	1.77906E-10	cellulose synthase-like D5
AT1G02820	-1.155694601	1.90042E-06	Late embryogenesis abundant 3 (LEA3) family protein
AT1G03055	-1.47821755	0.039514328	Encodes the ortholog of rice D27. It is plastid-localized
AT1G03600	-1.313989747	3.32831E-21	photosystem II family protein
AT1G03650	-1.106234195	2.51883E-05	Acyl-CoA N-acyltransferases (NAT) superfamily protein
AT1G04110	-1.223147664	1.80983E-05	Subtilase family protein
AT1G04270	-1.389689144	6.0171E-49	cytosolic ribosomal protein S15
AT1G04290	-1.457160722	3.44088E-31	Thioesterase superfamily protein
AT1G04410	-1.015422806	7.28119E-42	Lactate/malate dehydrogenase family protein
AT1G04480	-1.209537058	1.43479E-29	Ribosomal protein L14p/L23e family protein
AT1G04635	-1.473638708	4.37937E-10	ribonuclease P family protein / Rpp14 family protein
AT1G04640	-1.497074374	3.38349E-14	lipoyltransferase 2
AT1G05190	-1.352584767	7.33645E-14	Ribosomal protein L6 family
AT1G05210	-1.182024355	4.94902E-09	Transmembrane protein 97, predicted
AT1G05810	-1.177501032	1.25804E-14	RAB GTPase homolog A5E
AT1G06360	-1.752952196	1.1649E-24	Fatty acid desaturase family protein
AT1G06515	-1.057586857	2.01612E-08	Protein of unknown function (DUF3317)
AT1G06680	-1.114232779	2.65172E-14	photosystem II subunit P-1
AT1G07070	-1.415221348	2.91455E-26	Ribosomal protein L35Ae family protein

Link para ver a tabela completa: <http://inctipp.bioagro.ufv.br/maloni/>

**Tabela S1.2. Categorias em down regulação sob a overexpressão de LIMYB.**

DOWN REGULATED						
GO_ID	Pvalue	OddsRatio	ExpCount	Count	Size	Term
GO:0003735	0.000	17 270	22	187	398	structural constituent of ribosome
GO:0042586	0.003	Inf	0	2	2	peptide deformylase activity
GO:0006081	0.000	6 955	19	95	339	cellular aldehyde metabolic process
GO:0019252	0.000	5 042	11	43	190	starch biosynthetic process
GO:0006807	0.000	1 813	197	305	3878	nitrogen compound metabolic process
GO:0006364	0.000	11 669	14	99	252	rRNA processing
GO:0006412	0.000	8 645	32	178	578	translation
GO:0019288	0.000	10 925	13	87	229	isopentenyl diphosphate biosynthetic process,
GO:0046490	0.000	10 624	13	87	233	isopentenyl diphosphate metabolic process
GO:0006098	0.000	12 093	11	81	200	pentose-phosphate shunt
GO:0034660	0.000	6 546	23	110	413	ncRNA metabolic process
GO:0010207	0.000	14 313	10	79	177	photosystem II assembly
GO:0090407	0.003	1 684	24	38	446	organophosphate biosynthetic process
GO:0009657	0.000	13 235	8	61	145	plastid organization

**Tabela S2. Genes up regulados sob a overexpressão de LIMYB via RNA-seq**

id	log2foldchang	padj	description
AT1G01420	1.029845428	0.00055952	UDP-glucosyl transferase 72B30
AT1G01560	1.515530129	0.0253124	MAP kinase 11
AT1G02310	3.203692724	2.9008E-19	Glycosyl hydrolase superfamily protein
AT1G02340	1.461940627	0.00014882	basic helix-loop-helix (bHLH) DNA-binding superfamily
AT1G02390	4.887531643	2.3624E-15	glycerol-3-phosphate acyltransferase 2
AT1G02400	1.396018471	4.649E-11	gibberellin 2-oxidase 6
AT1G02460	1.631202943	5.806E-09	Pectin lyase-like superfamily protein
AT1G02470	2.791083755	6.1714E-34	Polyketide cyclase/dehydrase and lipid transport superfamily
AT1G02520	1.907061576	5.8776E-08	P-glycoprotein 11
AT1G03020	3.717518592	2.7442E-16	Thioredoxin superfamily protein
AT1G03080	1.051760975	3.2125E-14	kinase interacting (KIP1-like) family protein
AT1G03550	1.275059762	6.979E-13	Secretory carrier membrane protein (SCAMP) family
AT1G03790	4.690446553	1.2341E-20	Zinc finger C-x8-C-x5-C-x3-H type family protein
AT1G04220	2.499766854	1.1467E-19	3-ketoacyl-CoA synthase 2
AT1G04310	1.35844265	6.6171E-15	ethylene response sensor 2
AT1G04610	2.771008128	2.2477E-05	YUCCA 3
AT1G04700	1.505398463	0.00090022	PB1 domain-containing protein tyrosine kinase
AT1G04770	1.759569259	1.2284E-25	Tetratricopeptide repeat (TPR)-like superfamily protein
AT1G05000	1.952532327	6.1892E-34	Phosphotyrosine protein phosphatases superfamily
AT1G05100	3.119458469	1.6242E-13	mitogen-activated protein kinase kinase kinase 18
AT1G05260	2.039831049	1.9829E-07	Peroxidase superfamily protein
AT1G05610	3.134108772	3.7309E-15	ADP-glucose pyrophosphorylase small subunit 2
AT1G06090	1.507920605	0.01089317	Fatty acid desaturase family protein
AT1G07000	1.073536041	0.00036338	exocyst subunit exo70 family protein B2
AT1G07430	3.666828363	3.0469E-12	highly ABA-induced PP2C gene 2
AT1G07620	1.656909528	5.2447E-05	GTP-binding protein Obg/CgtA
AT1G07710	1.461199438	0.00482421	Ankyrin repeat family protein
AT1G10040	1.502374596	0.01128544	alpha/beta-Hydrolases superfamily protein
AT1G07900	2.603626974	7.9966E-05	LOB domain-containing protein 1
AT1G08165	2.724635707	2.3367E-14	unknown protein; FUNCTIONS IN: molecular_function
AT1G08230	1.528420928	1.2684E-26	Transmembrane amino acid transporter family protein
AT1G08250	1.178417423	3.2889E-09	arogenate dehydratase 6
AT1G08310	4.32611041	8.07E-78	alpha/beta-Hydrolases superfamily protein
AT1G08650	2.644992428	7.181E-65	phosphoenolpyruvate carboxylase kinase 1
AT1G08830	1.73473203	7.341E-37	copper/zinc superoxide dismutase 1
AT1G08910	1.049551195	8.2653E-18	zinc ion binding; zinc ion binding
AT1G08920	1.188881852	1.3056E-06	ERD (early response to dehydration) six-like 1
AT1G09080	1.806782307	0.00094758	Heat shock protein 70 (Hsp 70) family protein
AT1G09180	1.917001648	2.692E-06	secretion-associated RAS super family 1

Link para ver a tabela completa: <http://inctipp.bioagro.ufv.br/maloni/>

**Tabela S2.2. Categorias *up* reguladas sob a overexpressão de LIMYB.**

UP REGULATED						
GO_ID	Pvalue	OddsRatio	ExpCount	Count	Size	Term
GO:0031669	0.000	9 878	16	103	310	cellular response to nutrient levels
GO:0071496	0.000	8 795	18	107	349	cellular response to external stimulus
GO:0009991	0.000	8 089	19	105	363	response to extracellular stimulus
GO:0016036	0.000	15 508	7	58	130	cellular response to phosphate starvation
GO:0019375	0.000	17 956	5	48	99	galactolipid biosynthetic process
GO:0006664	0.000	11 256	7	49	132	glycolipid metabolic process
GO:0046467	0.000	8 331	8	49	161	membrane lipid biosynthetic process
GO:0006820	0.000	4 271	32	110	619	anion transport
GO:0015706	0.000	6 564	11	53	207	nitrate transport
GO:0010167	0.000	6 295	10	49	197	response to nitrate
GO:0009719	0.000	2 390	94	193	1824	response to endogenous stimulus
GO:0010033	0.000	2 163	135	249	2651	response to organic substance
GO:0006826	0.000	6 467	6	32	125	iron ion transport
GO:0051179	0.000	1 756	165	255	3295	localization

**Tabela S3. Constituintes ribossomais (GO:0003735), na intercessão entre Linhagens L1 e L3 por RNA-seq.**

Gene_ID	log2foldch	padj	description
AT1G07070	-1.42	0	Ribosomal protein L35Ae family protein
AT1G14320	-1.08	0	Ribosomal protein L16p/L10e family protein
AT1G17560	-1.42	0	Ribosomal protein L14p/L23e family protein
AT1G18540	-1.02	0	Ribosomal protein L6 family protein
AT1G22780	-1.72	0	Ribosomal protein S13/S18 family
AT1G23290	-1.35	0	Ribosomal protein L18e/L15 superfamily protein
AT1G23410	-1.54	0	Ribosomal protein S27a / Ubiquitin family protein
AT1G26880	-1.23	0	Ribosomal protein L34e superfamily protein
AT1G32990	-1.1	0	plastid ribosomal protein l11
AT1G35680	-1.62	0	Ribosomal protein L21
AT1G48350	-1.75	0	Ribosomal L18p/L5e family protein
AT1G52300	-1.21	0	Zinc-binding ribosomal protein family protein
AT1G56045	-1.38	0	Ribosomal protein L41 family
AT1G57660	-1.43	0	Translation protein SH3-like family protein
AT1G74970	-1.04	0	ribosomal protein S9
AT1G77750	-1.53	0	Ribosomal protein S13/S18 family
AT1G78630	-1.4	0	Ribosomal protein L13 family protein
AT1G79850	-1.07	0	ribosomal protein S17
AT2G05220	-1.08	0	Ribosomal S17 family protein
AT2G09990	-1.46	0	Ribosomal protein S5 domain 2-like superfamily protein
AT2G18020	-1.04	0	Ribosomal protein L2 family
AT2G18400	-1.57	0	ribosomal protein L6 family protein
AT2G19720	-1.47	0	ribosomal protein S15A B
AT2G19730	-1.36	0	Ribosomal L28e protein family
AT2G20450	-1.43	0	Ribosomal protein L14
AT2G25210	-1.4	0	Ribosomal protein L39 family protein
AT2G27720	-1.65	0	60S acidic ribosomal protein family
AT2G32060	-1.51	0	Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protei
AT2G33800	-1.49	0	Ribosomal protein S5 family protein
AT2G36160	-1.16	0	Ribosomal protein S11 family protein
AT2G36170	-1.57	0	Ubiquitin supergroup;Ribosomal protein L40e
AT2G36620	-1.42	0	ribosomal protein L24

**Tabela S4. Genes em intercessão RNA-seq/ChIP-seq, down regulados sob a overexpressão de LIMYB.**

gene_id	width	fold_enrichmen	x_log10_qvalu	distancet	log2foldchang	padj2	description
AT4G15670	187	5.72712	21.54313	-290	-4.1522414	4E-56	Thioredoxin superfamily proteinThioredo
AT5G50915	232	8.02524	36.49087	-3018	-3.26060339	4E-25	basic helix-loop-helix (bHLH) DNA-bindi
AT1G19610	247	5.71494	19.94048	-805	-3.20404266	1E-41	Arabidopsis defensin-like proteinPredict
AT3G19040	241	6.20132	23.18301	-40	-2.44444346	0.002	histone acetyltransferase of the TAFII25
AT1G31690	209	8.14683	37.43564	-2335	-2.43270933	5E-24	Copper amine oxidase family proteinCc
AT1G22500	171	5.83654	20.73765	-2942	-2.41067464	3E-32	RING/U-box superfamily proteinGene er
AT1G77760	170	4.86378	14.63918	-524	-2.36595639	2E-42	nitrate reductase 1Encodes the cytosol
AT5G25130	263	8.39002	39.34327	-2788	-2.36383527	2E-18	cytochrome P450, family 71, subfamily
AT5G25130	175	5.35016	17.60725	-1680	-2.36383527	2E-18	cytochrome P450, family 71, subfamily
AT3G28270	215	6.93089	28.30316	-781	-2.35760916	7E-06	Protein of unknown function (DUF677)E
AT4G25780	198	6.07972	22.35884	-370	-2.25063084	3E-16	CAP (Cysteine-rich secretory proteins,
AT5G51390	317	6.44451	24.85628	-406	-2.20697647	7E-19	unknown protein; Has 30201 Blast hits
AT4G16880	184	5.10697	16.10255	-456	-2.20409308	3E-25	Leucine-rich repeat (LRR) family protei
AT4G34760	208	4.499	12.5266	-78	-2.18381079	5E-39	SAUR-like auxin-responsive protein farr
AT1G66230	170	4.98537	15.36551	-583	-2.15020315	4E-12	myb domain protein 20Encodes a putat
AT3G23805	229	7.78205	34.6234	-686	-2.06843866	2E-31	ralf-like 24Member of a diversely expres
AT3G54880	346	5.58131	24.03821	-1009	-2.05853685	1E-40	unknown protein; BEST Arabidopsis the
AT1G72430	285	8.14683	37.43564	-2734	-2.05669988	2E-41	SAUR-like auxin-responsive protein farr
AT1G54740	179	5.95813	21.54378	-3237	-2.01177242	1E-51	Protein of unknown function (DUF3049)
AT1G23100	301	8.51161	40.3068	-2009	-1.9722981	8E-20	GroES-like family proteinGroES-like fan
AT2G28630	257	5.35016	17.60725	-2781	-1.91048044	2E-58	3-ketoacyl-CoA synthase 12Encodes K
AT1G29920	250	5.81333	44.38213	-14	-1.89883201	4E-69	chlorophyll A/B-binding protein 2Encode
AT4G02130	240	8.69157	43.11537	-1818	-1.89049212	3E-61	galacturonosyltransferase 6Encodes a
AT3G47650	325	13.57951	86.90302	-948	-1.8896719	4E-31	DnaJ/Hsp40 cysteine-rich domain supe
AT4G34730	172	4.01262	9.8797	-92	-1.88190208	1E-22	ribosome-binding factor A family protein
AT4G34730	222	6.20132	23.18301	-758	-1.88190208	1E-22	ribosome-binding factor A family protein
AT5G57785	227	7.29567	30.96981	-636	-1.79457663	2E-24	unknown protein; Has 30201 Blast hits
AT1G10960	283	9.94904	52.2709	-398	-1.78941146	6E-21	ferredoxin 1ferredoxin 1 (FD1); FUNCTI
AT3G53890	201	5.71494	19.94048	-340	-1.76255559	2E-56	Ribosomal protein S21e Ribosomal prc
AT3G16240	328	4.499	12.5266	-873	-1.75549859	1E-28	delta tonoplast integral proteinDelta ton
AT5G05580	242	6.93089	28.30316	-2794	-1.75510265	5E-41	fatty acid desaturase 8Encodes a temp
AT2G37470	179	5.1534	16.61361	-1519	-1.73193619	2E-37	Histone superfamily proteinHistone sup
AT1G76240	170	5.10697	16.10255	-74	-1.72343358	4E-10	Arabidopsis protein of unknown functior
AT1G22780	173	4.74219	13.92323	-867	-1.72247563	1E-59	Ribosomal protein S13/S18 familyS18 r
AT3G44590	279	9.69697	54.84788	-1091	-1.70907477	4E-36	60S acidic ribosomal protein family60S
AT4G30250	229	11.67307	67.35648	-1658	-1.70019825	7E-07	P-loop containing nucleoside triphosph
AT1G31290	192	6.5661	25.7057	-2235	-1.69940607	6E-31	ARGONAUTE 3ARGONAUTE 3 (AGO3)
AT3G62030	207	7.17408	30.07376	-332	-1.68212573	5E-30	rotamase CYP 4nuclear-encoded chlor
AT2G28605	170	4.91371	15.14424	-1574	-1.67378438	8E-19	Photosystem II reaction center PsbP fa

Link para ver a tabela completa: <http://inctipp.bioagro.ufr.br/maloni/>

**Tabela S4.2. Categorias em *down* regulação sob a overexpressão de LIMYB, via análise integrativa RNA-seq/ChIP-seq.**

GO_ID	Pvalue	OddsRatio	ExpCount	Count	Size	Term
GO:0003735	0.000	17 270	22	187	398	structural constituent of ribosome
GO:0006081	0.000	6 955	19	95	339	cellular aldehyde metabolic process
GO:0010207	0.000	6 512	2	12	177	photosystem II assembly
GO:0006364	0.000	6 907	3	18	256	rRNA processing
GO:0006354	0.000	4 981	7	30	133	DNA-templated transcription, elongation
GO:0019684	0.002	4 293	2	7	156	photosynthesis, light reaction
GO:0006098	0.000	4 673	2	10	200	pentose-phosphate shunt
GO:0006412	0.000	4 761	6	25	526	translation
GO:0006518	0.000	5.298	7	31	566	peptide metabolic process
GO:0043604	0.000	5	7	30	578	amide biosynthetic process
GO:0009773	0.000	22 606	3	29	51	photosynthetic electron transport in photosystem I
GO:0034660	0.000	6 546	23	110	413	ncRNA metabolic process
GO:0019288	0.005	3 186	3	8	229	methylethylthritol 4-phosphate pathway
GO:0046490	0.006	3 129	3	8	233	isopentenyl diphosphate metabolic process
GO:0090407	0.003	1 684	24	38	446	organophosphate biosynthetic process

**Tabela S5. Genes em intercessão RNA-seq/ChIP-seq, *up* regulados sob a overexpressão de LIMYB.**

gene_id	width	fold_enric	x_log10_q	distancet	log2foldch	padj2	description
AT1G0230	191	6.05	24.7	-1207	0.44	0	Cysteine proteinases superfamily protein
AT1G0239	172	4.26	11.18	-1118	4.89	0	glycerol-3-phosphate acyltransferase 2
AT1G0247	200	4.52	13.21	-790	2.79	0	Polyketide cyclase/dehydrase and lipid transport superfamily
AT1G0329	195	5.35	17.61	-368	0.7	0	unknown protein; FUNCTIONS IN: molecular_function unknow
AT1G0338	213	3.53	7.45	-1136	0.72	0	homolog of yeast autophagy 18 (ATG18) G
AT1G0470	202	4.5	12.53	-559	1.51	0	PB1 domain-containing protein tyrosine kinase
AT1G0591	292	5.39	18.48	-1119	0.62	0	cell division cycle protein 48-related / CDC48-related
AT1G0752	219	5.93	21.47	-1664	0.68	0	GRAS family transcription factor
AT1G0790	182	6.47	28.83	-2222	2.6	0	LOB domain-containing protein 1
AT1G0816	237	6.39	25.11	-354	2.72	0	unknown protein; FUNCTIONS IN: molecular_function unknow
AT1G0892	227	9.02	47.64	-2559	1.19	0	ERD (early response to dehydration) six-like 1 Encodes ESL1, a
AT1G0954	170	5.33	17.55	-3429	2.82	0	myb domain protein 61 Encodes putative transcription factor.
AT1G1089	240	6.69	26.56	-339	0.95	0	unknown protein; FUNCTIONS IN: molecular_function unknow
AT1G1148	170	4.8	17.03	-714	0.41	0	eukaryotic translation initiation factor-related
AT1G1695	198	7.42	31.87	-953	4.31	0	unknown protein; FUNCTIONS IN: molecular_function unknow
AT1G1783	238	6.32	24.02	-990	2.34	0	Protein of unknown function (DUF789)
AT1G2069	214	6.23	29.57	-2418	1.18	0	SWI-SNF-related chromatin binding protein
AT1G2088	203	5.35	17.61	-500	1	0	RNA-binding (RRM/RBD/RNP motifs) family protein
AT1G2311	170	5.59	19.15	-60	5.48	0	unknown protein; FUNCTIONS IN: molecular_function unknow
AT1G2627	170	5.47	18.37	-999	0.72	0	Phosphatidylinositol 3- and 4-kinase family protein Phosphati
AT1G2819	200	5.84	20.74	-1276	1.79	0	unknown protein; BEST Arabidopsis thaliana protein match is:
AT1G3075	209	5.79	21.77	-590	2.4	0.01	unknown protein; FUNCTIONS IN: molecular_function unknow
AT1G3075	170	5.46	20.19	-1444	2.4	0.01	unknown protein; FUNCTIONS IN: molecular_function unknow
AT1G3370	248	10.28	60.75	-1169	1.33	0	Beta-glucosidase, GBA2 type family protein
AT1G3572	170	4.62	13.22	-651	1.2	0	annexin 1
AT1G4475	199	3.89	9.25	-1158	0.67	0	purine permease 11
AT1G4712	229	7.54	32.78	-4459	0.33	0.01	Granulin repeat cysteine protease family protein cysteine pro
AT1G4733	185	6.2	23.18	-982	0.61	0	CBS domain-containing protein with a domain of unknown fur
AT1G4837	211	5.62	23.07	-372	1.36	0	YELLOW STRIPE like 8 Arabidopsis thaliana metal-nicotianamir
AT1G4923	174	5.71	19.94	-781	1.14	0	RING/U-box superfamily protein
AT1G5347	235	10.81	71.6	-376	0.94	0	mechanosensitive channel of small conductance-like 4 mecha
AT1G5349	222	4.86	14.64	-2348	2.92	0	RING/U-box superfamily protein
AT1G5502	197	3.65	8.03	-992	1.83	0.03	lipoxigenase 1
AT1G5602	207	5.96	21.54	-400	0.99	0	unknown protein; BEST Arabidopsis thaliana protein match is:
AT1G5803	250	10.9	72	-2360	0.91	0	cationic amino acid transporter 2
AT1G5975	210	8.75	42.25	-5043	0.61	0	auxin response factor 1 Encodes a member of the auxin respo
AT1G6136	227	6.2	23.18	-315	1.17	0	S-locus lectin protein kinase family protein
AT1G6142	229	7.38	37.03	-800	1.01	0.01	S-locus lectin protein kinase family protein
AT1G6237	296	10.21	54.43	-4242	1.15	0	RING/U-box superfamily protein

Link para ver a tabela completa: <http://inctipp.bioagro.ufv.br/maloni/>

**Tabela S5.2. Categorias em down regulação sob a overexpressão de LIMYB, via análise integrative RNA-seq/ChIP-seq.**

GO_ID	Pvalue	OddsRatio	ExpCount	Count	Size	Term
GO:0006820	0.000	4 271	32	110	619	anion transport
GO:0050896	0.000	1 895	74	126	1814	response to stimulus
GO:0042594	0.000	9 613	16	100	306	response to starvation
GO:0031669	0.000	9 878	16	103	310	cellular response to nutrient levels
GO:0009611	0.000	2 845	17	44	335	response to wounding
GO:0009991	0.000	8 089	19	105	363	response to extracellular stimulus
GO:0071496	0.000	8 795	18	107	349	cellular response to external stimulus
GO:0007154	0.000	6 558	21	102	439	cell communication
GO:0010106	0.000	6 851	6	31	116	cellular response to iron ion starvation
GO:0006826	0.000	5 413	1	6	99	iron ion transport
GO:0030643	0.000	3 412	5	7	362	phosphate ion homeostasis
GO:0016036	0.000	2 657	3	7	124	cellular response to phosphatte starvation
GO:0080003	0.000	5 897	2	3	3	thalianol metabolic process
GO:0010167	0.000	6 295	10	49	197	response to nitrate

**Tabela S6. Constituintes ribossomais GO:0003735), na intercessão entre RNA-seq e ChIP-seq.**

gene_id	width	fold_enri	x_log10_c	distance	log2foldc	padj2	description
AT1G09690	185	4.13	10.52	-152	-0.84	0.00	Translation protein SH3-like family protein
AT1G22780	173	4.74	13.92	-867	-1.72	0.00	Ribosomal protein S13/S18 family
AT1G74060	371	6.04	23.03	-79	-0.38	0.00	Ribosomal protein L6 family protein
AT2G27530	184	5.23	16.85	-144	-1.49	0.00	Ribosomal protein L1p/L10e family
AT2G34480	222	6.93	28.30	-1334	-0.62	0.00	Ribosomal protein L18ae/LX family protein
AT2G36620	184	4.62	13.22	-1250	-1.42	0.00	ribosomal protein L24e/PL24A encodes ribosomal pr
AT2G37270	271	7.79	36.16	-75	-0.75	0.00	ribosomal protein 58S
AT3G07110	172	5.11	16.10	-350	-1.26	0.00	Ribosomal protein L13 family protein
AT3G09200	188	4.27	15.26	-75	-0.77	0.00	Ribosomal protein L10 family protein
AT3G25520	173	5.84	20.74	-455	-1.26	0.00	ribosomal protein L5e
AT3G27830	212	6.69	26.56	-220	-1.12	0.00	ribosomal protein L12-A60S ribosomal protein L12-/
AT3G44590	279	9.70	54.85	-1091	-1.71	0.00	60S acidic ribosomal protein family
AT3G53890	201	5.71	19.94	-340	-1.76	0.00	Ribosomal protein S21e
AT3G55280	184	5.18	19.86	-1065	-0.83	0.00	ribosomal protein L23AB60S ribosomal protein L23A
AT4G13170	237	7.05	29.18	-595	-1.62	0.00	Ribosomal protein L13 family protein
AT4G31985	193	6.80	29.21	-2389	-1.48	0.00	Ribosomal protein L39 family protein
AT4G35490	180	5.96	21.54	-302	-0.74	0.00	mitochondrial ribosomal protein L11
AT4G39200	230	5.59	19.15	-793	-1.26	0.00	Ribosomal protein S25 family protein
AT5G02870	264	9.24	46.22	-402	-0.60	0.00	Ribosomal protein L4/L1 family
AT5G13510	230	8.66	54.93	-630	-1.02	0.00	Ribosomal protein L10 family protein
AT5G22440	235	10.45	60.19	-693	-1.42	0.00	Ribosomal protein L1p/L10e family
AT5G27770	303	4.28	11.84	-364	-1.54	0.00	Ribosomal L22e protein family
AT5G46160	205	8.51	40.31	-972	-1.10	0.00	Ribosomal protein L14p/L23e family protein
AT5G59850	266	11.19	62.97	-624	-1.57	0.00	Ribosomal protein S8 family protein
AT5G64140	170	5.35	17.61	-222	-1.10	0.00	ribosomal protein S28e
ATMG00560	193	3.04	37.87	-1156	-0.59	0.00	Nucleic acid-binding, OB-fold-like protein

**Tabela S7. Processos metabólicos associados a derivados de aldeídos (GO:0006081), na intercessão entre RNA-seq e ChIP-seq.**

gene_id	width	fold_enri	x_log10_c	log2foldc	padj2	description
AT3G47650	325	13.58	86.9	-1.89	0	DnaJ/Hsp40 cysteine-rich domain su
AT3G62030	207	7.17	30.07	-1.68	0	rotamase CYP 4H nuclear-encoded chl
AT3G12930	234	6.32	24.02	-1.63	0	Lojap-related proteinLojap-related
AT5G18660	170	4.13	10.52	-1.5	0	NAD(P)-binding Rossmann-fold sup
AT5G45930	218	9	44.22	-1.5	0	magnesium chelatase i2Encodes a s
AT5G45930	247	6.32	24.02	-1.5	0	magnesium chelatase i2Encodes a s
AT4G24750	235	6.85	28.49	-1.34	0	Rhodanese/Cell cycle control phosph
AT2G22990	212	6.29	24.32	-1.52	0	sinapoylglucose 1Sinapoylglucose:n
AT4G09650	400	11.33	66.7	-1.33	0	ATP synthase delta-subunit geneEn
AT3G26650	215	5.3	19.24	-1.36	0	glyceraldehyde 3-phosphate dehyd
AT5G09660	184	5.59	19.15	-1.24	0	peroxisomal NAD-malate dehydrog
AT4G38970	221	6.1	29.43	-1.15	0	fructose-bisphosphate aldolase 2Fr
AT1G74070	219	8.08	40.79	-1.06	0	Cyclophilin-like peptidyl-prolyl cis-t
AT3G51820	299	10.97	61.55	-0.9	0	UbiA prenyltransferase family prote
AT5G45390	335	5.59	19.15	-1.11	0	CLP protease P4One of several nucl
AT4G13770	192	6.16	23.45	-1.02	0	cytochrome P450, family 83, subfar
AT4G14890	285	5.81	20.67	-0.84	0	2Fe-2S ferredoxin-like superfamily
AT5G66470	216	7.32	42.53	-0.87	0	RNA binding;GTP bindingRNA bindi
AT1G11750	223	5.06	18.69	-0.87	0	CLP protease proteolytic subunit 6
AT3G52380	253	10.94	60.8	-0.66	0	chloroplast RNA-binding protein 33
AT4G34240	241	5.71	19.94	-0.6	0	aldehyde dehydrogenase 311Encod
AT4G24620	198	5.59	19.15	-0.36	0.01	phosphoglucose isomerase 1The PC

**Tabela S8. Genes relacionados a proteínas de processamento no retículo endoplasmático, detectados na análise integrativa.**

Gene_ID	width	fold_enri	x_log10_c	log2foldc	padj2	description
AT1G10230	203	6.57	25.71	-1.03	0	SKP1-like 18
AT1G54050	405	5.98	22.82	-1.48	0	HSP20-like chaperones superfamily
AT2G04030	406	6.32	24.02	-1.62	0	Chaperone protein htpG family pro
AT2G04030	173	5.4	18.16	-1.62	0	Chaperone protein htpG family pro
AT2G40290	242	5.71	19.94	-0.64	0	Eukaryotic translation initiation fac
AT3G62560	259	8.39	39.34	-0.54	0	Ras-related small GTP-binding fam
AT5G02150	269	4.36	14.73	-0.78	0	Fes1C
AT5G02490	221	8	40.5	-1.25	0	Heat shock protein 70 (Hsp 70) fam
AT5G02490	221	8	40.5	-1.25	0	Heat shock protein 70 (Hsp 70) fam
AT5G56010	201	6.32	24.02	-0.47	0	heat shock protein 81-3

**Tabela S9. Genes relacionados a fosforilação oxidativa, detectados na análise integrativa.**

Gene_ID	width	fold_enric	x_log10_q	distancet	log2foldch	padj2	description
AT1G15120	173	4.65	16.49	-305	-0.72	0	Ubiquinol-cytochrome C reductase h
AT1G15120	173	4.65	16.49	-305	-0.72	0	Ubiquinol-cytochrome C reductase h
AT1G16700	194	8.51	40.31	-957	-0.72	0	Alpha-helical ferredoxin
AT1G16700	194	8.51	40.31	-957	-0.72	0	Alpha-helical ferredoxin
AT1G16700	194	8.51	40.31	-957	-0.72	0	Alpha-helical ferredoxin
AT1G75630	186	5.47	18.37	-408	-0.96	0	vacuolar H+-pumping ATPase 16 kDa
AT1G75630	186	5.47	18.37	-408	-0.96	0	vacuolar H+-pumping ATPase 16 kDa
AT4G04640	237	7.52	35.64	-2103	-0.76	0	ATPase, F1 complex, gamma subunit
AT4G09650	400	11.33	66.7	-1043	-1.33	0	ATP synthase delta-subunit gene Enc
AT4G11150	189	6.2	23.18	-1200	-0.63	0	vacuolar ATP synthase subunit E1 Enr
AT4G32260	331	11.01	76.31	-236	-0.98	0	ATPase, F0 complex, subunit B/B., ba
AT5G40810	173	5.59	19.15	-2033	-0.7	0	Cytochrome C1 family
AT5G40810	170	4.38	11.85	-273	-0.7	0	Cytochrome C1 family
AT5G47890	234	5.8	21.81	-1038	-1.04	0	NADH-ubiquinone oxidoreductase B
AT5G47890	200	5.84	22.34	-676	-1.04	0	NADH-ubiquinone oxidoreductase B

**Tabela S10. Genes relacionados a proteínas de montagem do fotossistema II (PSII), detectados na análise integrativa.**

gene_id	width	fold_enric	x_log10_qv	distancet	log2foldch	padj2	description
AT5G66570	368	8.02	43.63	-1109	-0.7	0	PS II oxygen-evolving complex 1 Encodes
AT1G06680	205	8.22	38.21	-208	-1.11	0	photosystem II subunit P-1
AT4G05180	300	6.81	28.34	-878	-0.59	0	photosystem II subunit Q-2
AT1G79040	170	4.86	14.64	-586	-1.34	0	photosystem II subunit R
AT1G44575	229	6.51	25.96	-1501	-1.08	0	Chlorophyll A-B binding family protein
AT2G30570	170	4.86	14.64	-88	-1.34	0	photosystem II reaction center W Encode
AT1G67740	257	11.55	66.25	-1158	-0.67	0	photosystem II BY
AT1G52230	264	11.53	68.15	-2151	-0.97	0	photosystem I subunit H2 Phosphorylatio
AT1G10960	283	9.95	52.27	-398	-1.79	0	ferredoxin 1
AT1G20020	243	5.88	21.69	-1150	-1.16	0	ferredoxin-NADP(+)-oxidoreductase 2
AT4G04640	237	7.52	35.64	-2103	-0.76	0	ATPase, F1 complex, gamma subunit prot
AT4G09650	400	11.33	66.7	-1043	-1.33	0	ATP synthase delta-subunit gene Encode:
AT4G32260	331	11.01	76.31	-236	-0.98	0	ATPase, F0 complex, subunit B/B., bacteri

**Tabela S11. Genes relacionados a proteínas de montagem do complexo coletor de luz da fotossíntese, detectados na análise integrativa.**

gene_id	width	fold_enric	x_log10_q	distancet	log2foldcl	padj2	description
AT1G29910	372	10.72	104.41	-889	-0.92	0	chlorophyll A/B binding protein 3 mer
AT1G45474	206	6.2	23.18	-964	-1.3	0	photosystem I light harvesting comple
AT3G27690	473	7.37	31.72	-482	-0.93	0	photosystem II light harvesting comple
AT3G47470	699	5.67	19.79	-1654	-1.03	0	light-harvesting chlorophyll-protein co
AT4G10340	336	8.54	40.96	-291	-0.88	0	light harvesting complex of photosyste
AT4G10340	180	4.26	11.18	-1234	-0.88	0	light harvesting complex of photosyste
AT5G54270	462	5.71	21.1	-672	-0.85	0	light-harvesting chlorophyll B-binding
AT5G54270	217	4.92	17.51	-2445	-0.85	0	light-harvesting chlorophyll B-binding

**Tabela S12. Genes relacionados a proteínas de montagem dos complexos ubiquitinação, detectados na análise integrativa.**

gene_id	width	fold_enric	x_log10_q	distancet	log2foldcl	padj2	description
AT1G14400	192	6.09	24.02	-1584	-0.43	0	ubiquitin carrier protein 1
AT3G57870	219	7.54	32.78	-546	-0.98	0	sumo conjugation enzyme 1 Encodes a SUMO ligase that dire
AT1G75950	187	5.96	21.54	-1411	-0.101	0	S phase kinase-associated protein 1 SKP1 is core component
AT1G75950	187	5.96	21.54	-1411	-0.102	0	S phase kinase-associated protein 1 SKP1 is core component
AT3G57870	219	7.54	32.78	-546	-0.103	0	sumo conjugation enzyme 1
AT5G08320	182	5.47	21.54	-831	-0.83	0	E2F-associated phosphoprotein, C-terminal (InterPro:IPR015

**Tabela S13. Genes relacionados a interação planta-patógeno, detectados na análise integrativa.**

gene_id	width	fold_enric	x_log10_q	distancet	log2foldcl	padj2	description
AT2G13790	211	4.99	15.37	-127	1.14	0	Encodes calcium dependent protein k
AT2G13790	211	4.99	15.37	-127	1.14	0	somatic embryogenesis receptor-like
AT2G17290	296	14.21	93.68	-448	0.56	0	Calcium-dependent protein kinase fa
AT2G38470	317	9.13	61.57	-440	0.9	0	WRKY DNA-binding protein 33
AT2G43790	180	5.23	16.85	-1017	0.53	0	MAP kinase 6 Encodes a MAP kinase ir
AT2G46440	181	6.41	26.03	-1124	2.26	0	cyclic nucleotide-gated channels Men
AT3G48080	294	5.11	16.1	-776	2.57	0	alpha/beta-Hydrolases superfamily pr

**Tabela S14. Genes relacionados a biossíntese de fenilpropanóides, detectados na análise integrativa.**

gene_id	width	fold_enric	x_log10_q	distancetofe	log2foldcl	padj	description
AT2G35380	181	4	11.46	-622	2.42	0	Peroxidase superfamily protein Peroxidase superfamily protein
AT3G21240	208	5.59	19.15	-164	1.64	0	4-coumarate:CoA ligase 2
AT4G26220	170	5.3	17.79	-604	2.17	0	S-adenosyl-L-methionine-dependent methyltransferases

**Tabela S15. Predições de novos candidatos à proteínas ribossomais pelo método de *machine learning* (Utilizando o software RAMA).**

Accession	<i>A.thaliana</i>		<i>O.lucimarinus</i>		<i>O.satíva</i>		<i>S.lycopersicum</i>		<i>G.max</i>		<i>Z.mays</i>		Classification		Description
AT1G11655.1	0.98	0.89	1.0	0.8	0.98	0.99	0.99	0.97	0.97	0.98	0.95	1.0	0.98	0.94	<a href="#">Unknow function</a>
AT2G34800.1	0.93	0.88	1.0	1.0	0.97	1.0	0.95	1.0	0.97	0.91	0.8	1.0	0.94	0.96	<a href="#">Unknow function</a>
AT3G43682.1	0.92	0.86	1.0	0.86	0.94	0.86	0.99	0.94	0.85	0.95	0.95	0.92	0.94	0.9	<a href="#">Unknow function</a>
AT3G43829.1	0.94	1.0	0.99	1.0	0.99	1.0	0.88	0.92	0.85	0.99	0.91	1.0	0.93	0.98	<a href="#">Unknow function</a>
AT3G18282.1	0.9	0.92	0.99	0.92	0.94	0.96	0.95	0.91	0.95	0.97	0.85	0.94	0.93	0.94	<a href="#">Unknow function</a>
AT5G51451.1	0.98	0.93	1.0	0.79	0.88	0.96	0.96	0.94	0.85	0.97	0.94	0.95	0.93	0.92	<a href="#">Unknow function</a>
AT1G32260.1	0.93	1.0	0.99	0.97	0.88	1.0	0.97	0.95	0.81	0.91	0.9	1.0	0.91	0.97	<a href="#">Unknow function</a>
AT3G32120.1	0.91	0.97	0.98	0.94	0.94	0.89	0.94	0.92	0.93	0.99	0.78	0.93	0.91	0.94	<a href="#">Unknow function</a>
AT3G05936.1	0.87	1.0	1.0	0.98	0.96	0.99	0.96	0.92	0.84	1.0	0.75	1.0	0.9	0.98	<a href="#">Unknow function</a>
AT3G25099.1	0.98	0.94	1.0	0.96	0.89	1.0	0.82	0.91	0.78	0.96	0.95	0.88	0.9	0.94	<a href="#">Unknow function</a>
AT5G57785.1	0.97	0.95	0.99	0.88	0.87	0.82	0.88	0.93	0.78	0.97	0.92	0.88	0.9	0.9	<a href="#">Unknow function</a>
AT2G33509.1	0.95	0.9	1.0	0.71	0.88	0.94	0.89	0.93	0.87	0.96	0.8	0.97	0.9	0.9	<a href="#">Unknow function</a>
AT5G28465.1	0.93	0.91	1.0	1.0	0.92	0.95	0.79	0.88	0.92	0.99	0.76	0.91	0.89	0.94	<a href="#">Unknow function</a>
AT5G51812.1	0.86	0.99	0.98	0.99	0.96	1.0	0.87	0.99	0.77	0.96	0.86	1.0	0.88	0.99	<a href="#">Unknow function</a>
AT2G35480.1	0.91	1.0	0.93	0.92	0.71	1.0	0.88	0.97	0.91	0.91	0.91	1.0	0.88	0.97	<a href="#">Unknow function</a>
AT5G57760.1	0.88	0.91	1.0	0.91	0.75	0.96	0.77	0.97	0.89	0.98	0.96	1.0	0.88	0.96	<a href="#">Unknow function</a>
AT5G29044.1	0.98	1.0	1.0	0.96	0.87	0.81	0.74	0.8	0.89	0.89	0.79	0.94	0.88	0.9	<a href="#">Unknow function</a>
AT2G11015.1	0.98	0.99	0.93	0.97	0.92	0.99	0.81	0.97	0.82	0.93	0.75	0.93	0.87	0.96	<a href="#">Unknow function</a>
AT5G64401.1	0.89	0.97	1.0	0.98	0.9	0.99	0.78	0.91	0.86	0.93	0.78	0.98	0.87	0.96	<a href="#">Unknow function</a>
AT1G72141.1	0.81	0.95	1.0	0.85	0.95	0.92	0.81	0.91	0.72	0.98	0.9	0.99	0.86	0.93	<a href="#">Unknow function</a>
AT3G52070.1	0.74	0.98	1.0	0.76	0.88	0.98	0.87	0.89	0.71	0.87	0.94	0.98	0.86	0.91	<a href="#">Unknow function</a>
AT3G24929.1	0.78	1.0	1.0	0.99	0.91	1.0	0.85	0.94	0.73	0.99	0.85	1.0	0.85	0.99	<a href="#">Unknow function</a>
AT4G02910.1	0.92	0.92	0.99	0.96	0.77	0.92	0.81	0.91	0.84	0.93	0.77	0.95	0.85	0.93	<a href="#">Unknow function</a>

## 7. Referências bibliográficas

**Aaron S. and Estelle, M.** (2010). "The Ubiquitin-Proteasome System Regulates Plant Hormone Signaling." *The Plant journal: for cell and molecular biology* 61, 1029–1040.

**Altman A. and Arengada P.** (2009). Efeitos of N, P, K and S on metabolismo: new knowledge gained from multi-level analysis. *Curr. Opin. Plant Biol.* 12, 275–283.

**Anders S. and Huber W.** (2010). Differential expression analysis for sequence count data. *Genome Biology.* 11, 106.

**Argüello-Astorga G., Lopez-Ochoa L., Kong L. J., Orozco, B. M., Settlage S. B., Hanley-Bowdoin L.** (2004). A novel motif in geminiviral replication proteins interacts with the plant retinoblastoma related protein. *J Virol.* 78, 4817-4826.

**Aro E. M., Suorsa M., Rokka A., Allahverdiyeva Y., Paakkarinen V., Saleem, A., Battchikova N., Rintamäki E.** (2005). Dynamics of photosystem II: a proteomic approach to thylakoid protein complexes. *Journal of Experimental Botany.* 56, 347–356.

**Ausubel F. M.** (2005). Are innate immune signaling pathways in plants and animals conserved?. *Nature immunology.* 6, 973–979.

**Bartholomew B.** (2014). ISWI Chromatin Remodeling: One primary actor or coordinated effort? *Current Opinion in Structural Biology.* 24, 150–155.

**Ben Khaled, S., Postma, J., Robatzek, S.** (2015). A Moving View: Subcellular Trafficking Processes in Pattern Recognition Receptor-Triggered Plant Immunity. *Annu. Rev. Phytopathol.* 53, 379-402.

**Blevins T., Rajeswaran R., Shivaprasad P.V., Beknazariants D., Si-Ammour A., Park H.S., Vazquez F., Robertson D., Meins F. J., Hohn, T.** (2006). Four plant dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Research.* 34, 6233–6246.

**Boex-Fontvieille E., Daventure M., Jossier M., Zivy M., Hodges M., & Tcherkez, G.** (2013). Photosynthetic Control of Arabidopsis Leaf Cytoplasmic Translation Initiation by Protein Phosphorylation. *PLoS ONE.* 8, 70692.

**Bolger A. M., Lohse M., Usadel B.** (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 30,15, 2114-2120.

**Boller T. and Felix, G.** (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review of Plant Biology.* 60, 379–406.

**Browning K. S., Bailey-Serres J.** (2015). Mechanism of Cytoplasmic mRNA Translation. *The Arabidopsis Book / American Society of Plant Biologists.* 13, 0176.

**Brustolini O. J. B., Machado J. P. B., Condori-Apfata J. a., et al.** (2015). Sustained NIK-mediated antiviral signalling confers broad-spectrum tolerance to begomoviruses in cultivated plants. *Plant Biotechnology Journal*. 1–12.

**Carbonell A., Fahlgren N., Garcia-Ruiz H., Gilbert K. B., Montgomery T. A., Nguyen T., Cuperus J. T., Carrington J. C.** (2012). Functional analysis of three Arabidopsis ARGONAUTES using slicer-defective mutants. *Plant Cell*. 24, 3613–3629.

**Carvalho C. M., Fontenelle M. R., Florentino L. H., Santos A. a, Zerbini F. M. and Fontes E. P. B.** (2008). A novel nucleocytoplasmic traffic GTPase identified as a functional target of the bipartite *geminivirus* nuclear shuttle protein. *The Plant journal: for cell and molecular biology*. 55, 869–80.

**Castro P. H., Tavares R. M., Bejarano E. R., Azevedo H.** (2012). SUMO, a heavyweight player in plant abiotic stress responses. *Cell Mol Life Sci*. 69, 3269-3283.

**Chang I. F., Szick-Miranda K., Pan S., Bailey-Serres, J.** (2005) Proteomic characterization of evolutionarily conserved and variable proteins of Arabidopsis cytosolic ribosomes. *Plant Physiol*. 137, 848–862

**Chen K. M., Holmström M., Raksajit W., Suorsa M., Piippo M., Aro E. M.** (2010). Small chloroplast-targeted DnaJ proteins are involved in optimization of photosynthetic reactions in Arabidopsis thaliana. *BMC Plant Biology*. 43, 1471–2229.

**Chen Y., Yang, X., He K.** (2006). The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Molecular Biology*. 60, 107–124.

**Corrêa J. C. R. and Salgado H. R. N.** (2011). Atividade inseticida das plantas e aplicações: revisão. *Revista Brasileira Plantas Mediciniais*. v.13, p.500-506,

**Cramer G. R., Urano K., Delrot S., Pezzotti M., Shinozaki, K.** (2011). Effects of abiotic stress on plants: a systems biology perspective. *BMCPlantBiol*. 11, 1471-2229.

**Dangl J. L., Horvath D. M., Staskawicz B. J.** (2013). Pivoting the plant immune system from dissection to deployment. *Science*. 341, 746–751.

**Deleris A., Gallego-Bartolome, J., Bao, J., Kasschau, K. D., Carrington, J. C., Voinnet, O.** (2006). Hierarchical action and inhibition of plant dicer-like proteins in antiviral defense. *Science*. 313, 68–71.

**Dubos C., Stracke R., Grotewold E., Weisshaar B., Martin C., Lepiniec L.** (2010). MYB transcription factors in Arabidopsis. *Trends in Plant Science*. 15, 573–581.

**Egelkrout E. M., Robertson D., Hanley-Bowdoin, L.** (2001). Proliferating cell nuclear antigen transcription is repressed through an E2F consensus element and activated by geminivirus infection in mature leaves. *Plant Cell*. 13, 1437-1452.

**Falcon, S. and Gentleman, R.** (2007). Using GOstats to test gene lists for GO term association. *Bioinformatics*. 23, 257–258.

**Fitter, A. H. and Fitter, R. S.** (2002). Rapid changes in flowering time in British plants. *Science*. 296, 1689–1691.

**Fontes E. P. B., Santos A. a., Luz D. F., Waclawovsky A. J. and Chory J.** (2004). The *geminivirus* nuclear shuttle protein is a virulence factor that suppresses transmembrane receptor kinase activity. *Genes and Development*. 18, 2545–2556.

**Ghanim, M.** (2014). Plant Virus Epidemiology Symposium Ecology, evolution and control of plant viruses and their vectors, *Science direct*. 186, 47–54.

**Gojon A., Nacry P., Davidian J.-C.** (2009). Root uptake regulation: a central process for NPS homeostasis in plants. *Curr. Opin. Plant Biol.* 12, 328–338.

**Gu J., Zhou Z., Li Z., et al.** (2017). Photosynthetic Properties and Potentials for Improvement of Photosynthesis in Pale Green Leaf Rice under High Light Conditions. *Frontiers in Plant Science*. 8,1082.

**Gururani M. A., Venkatesh J., Ganesan M., Strasser R. J., Han Y., Kim J. I., Lee H. Y., Song P. S.** (2015). In vivo assessment of cold tolerance through chlorophyll-a fluorescence in transgenic zoysiagrass expressing mutant phytochrome A. *PLOS ONE*.10.

**Gururani M.A., Upadhyaya C.P., Strasser R.J., Woong Y.J., Park S.W.** (2012). Physiological and biochemical responses of transgenic potato plants with altered expression of PSII manganese stabilizing protein. *Plant Physiol. Biochem.* 58,182–194.

**Hanley-Bowdoin L., Bejarano E. R., Robertson D., Mansoor S.** (2013). Geminiviruses: masters at redirecting and reprogramming plant processes. *Nature reviews. Microbiology*. 11, 777–88

**Hanley-Bowdoin, L., Settlage, S.B., Orozco, B.M., Nagar, S. and Robertson, D.** (1999). Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. *Critical reviews in biochemistry and molecular biology*. 18, 71–106.

**Hennessy F., Nicoll W. S., Zimmermann R., Cheetham M. E., Blatch G. L.** (2005). Not all J domains are created equal: implications for the specificity of Hsp40–Hsp70 interactions. *Protein Science*. 14, 1697–1709.

**Hershko A. and Ciechanover A.** (1998). The ubiquitin system. *Annu Rev Biochem* 67: 425–479

**Hofius D., Maier A. T., Dietrich C., Jungkunz I., Bornke F., Maiss, E. Sonnewald U.** (2007). Capsid protein-mediated recruitment of host DnaJ-Like proteins is required for Potato virus Y infection in tobacco plants. *J. Virol.* 81, 11870–11880.

**Hotton S. K and Callis J.** (2008). Regulation of cullin RING ligases. *Annu Rev Plant Biol.* 59, 467-489.

**Huang H. and Nusinow D. A.** (2016). Into the evening: complex interactions in the Arabidopsis circadian clock. *Trends Genet.* <https://dx.doi.org/10.1016/j.tig.08.002>.

**Hwang J., Oh C., Kang B.** (2013). Translation elongation factor 1B (eEF1B) is an essential host factor for infection in plants, *Virology*. 439,105-114.

**Jiang X, Chen Z. J.** (2011). The role of ubiquitylation in immune defence and pathogen evasion. *Nat Rev Immunol* ,12, 35–48.

**Katiyar A., Smita, S., Lenka S. K., Rajwanshi R., Chinnusamy V., Bansal K. C.** (2012). Genome-wide classification and expression analysis of MYB transcription factor families in rice and *Arabidopsis*. *BMC Genomics*. 13, 544

**Katiyar, A., Smita, S., Lenka, S. K., Rajwanshi, R., Chinnusamy, V., Bansal, K. C.** (2012). Genome-wide classification and expression analysis of MYB transcription factor families in rice and *Arabidopsis*. *BMC Genom.* 13, 544

**Katti M. V., Sakharkar M. K., Ranjekar P. K., Gupta V. S.** (2000). TRES: comparative promoter sequence analysis. *Bioinformatics*. Aug; 16(8):739-40. Keywords: Eukaryotic translation elongation factor 1A; Eukaryotic translation elongation factor 1B; Host factor; Tobacco mosaic virus; Virus resistance

**Langmead B., Trapnell C., Pop M., Salzberg S. L.** (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10, 3-25.

**Lefeuvre P. and Moriones E.** (2015). Recombination as a motor of host switches and virus emergence : geminiviruses as case studies. *Current Opinion in Virology*. 10, 14–19

**Li B., Lu D., Shan L.** (2014). Ubiquitination of pattern recognition receptors in plant innate immunity. *Mol Plant Pathol*. 15, 737–746.

**Liao Y., Tian M., Zhang H., Li X., Wang Y., Xia X., Zhou J., Zhou Y., Yu J., Shi K. and Klessig, D. F.** (2015). Salicylic acid binding of mitochondrial alpha-ketoglutarate dehydrogenase E2 affects mitochondrial oxidative phosphorylation and electron transport chain components and plays a role in basal defense against tobacco mosaic virus in tomato. *New Phytol*. 205, 1296–1307.

**Lin, Q. and Aoyama, T.** (2012). Pathways for epidermal cell differentiation via the homeobox gene *GLABRA2*: update on the roles of the classic regulator. *J. Integr. Plant Biol*. 54, 729–737.

**Liu B., Qian S-B.** (2014). Translational reprogramming in stress response. *Wiley interdisciplinary reviews RNA*. 5,301-305.

**Liu F., Ni W., Griffith M. E., Huang Z., Chang C., Peng W., Ma H., Xie D.** (2004). The ASK1 and ASK2 genes are essential for *Arabidopsis* early development. *Plant Cell*. 16, 5-20.

**Loebenstein G. and Katis N.** (2014). “Control of plant virus diseases seed-propagated crops,” in *Advance Virus Reserch*, eds L. Gad and K. Nikolaos (Cambridge, MA: Academic Press), 11.

**Lozano-Durán R., Rosas-Díaz T., Gusmaroli G., Luna A.P., Taconnat L., Deng X. W., Bejarano E. R.** (2011). Geminiviruses subvert ubiquitination by altering CSN-mediated derubylation of SCF E3 ligase complexes and inhibit jasmonate signaling in *Arabidopsis thaliana*. *Plant Cell*. 23, 1014-1032.

**Luo W. and Brouwer C.** (2013). Pathview: an R/bioconductor package for pathway-based data integration and visualization. *Bioinformatics*. 29, 14, 1830–1831.

**Mansoor S., Zafar Y., Briddon R. W.** (2006). *Geminivirus* disease complexes: the threat is spreading. *Trends in Plant Science*. 11, 209–212.

**Mariano A. C., Andrade M. O., Santos A. A., Carolino S. M. B., Oliveira M. L., Baracat-Pereira M. C., Brommonshenkel S. H., Fontes, E. P. B.** (2004). Identification of a novel receptor-like protein kinase that interacts with a geminivirus nuclear shuttle protein. *Virology*. 318, 24–31.

**Marino D, Peeters N, Rivas S.** (2012). Ubiquitination during plant immune signaling. *Plant Physiol*. 160, 15–27.

**Matus J. T., Cavallini, E., Loyola R., Höll J., Finezzo L., Dal Santo S., Violet S., Commisso M., Roman F., Schubert A., Alcalde J. A., Bogs J., Ageorges A., Tornielli G. B. and Arce-Johnson, P.** (2017). A group of grapevine MYBA transcription factors located in chromosome 14 control anthocyanin synthesis in vegetative organs with different specificities compared with the berry color locus. *Plant J*. 91, 220–236

**Mayberry L. K., Allen M. L., Nitka, K. R., Campbell L., Murphy P. A., et al.** (2011). Plant cap-binding complexes eukaryotic initiation factors eIF4F and eIFISO4F: molecular specificity of subunit binding. *J Biol Chem*. 286, 42566–42574.

**Mehuyas O.** (2008). Physiological roles of ribosomal protein S6: one of its kind. *Intl Rev Cell Mol Biol*. 268, 1–37.

**Monaghan J. and Zipfel C.** (2012). Plant pattern recognition receptor complexes at the plasma membrane. *Current Opinion in Plant Biology*. 15, 349–357.

**Moroy T. and Geisen C.** (2004). Cyclin E. *Int J Bioch Cell Biolog*. 36,1424-39.

**Navas-Castillo J., Fiallo-Olivé, E., Sánchez-Campos, S.** (2011). Emerging virus diseases transmitted by whiteflies. *Annual review of phytopathology*. 49, 219–248.

**Nelson, D. L. and Cox, M. M.** (2014). Principios de bioquímica de lehniger. 6. ed *Porto Alegre: Artmed*.

**Ntoukakis V., Mucyn T. S., Gimenez-Ibanez S., et al.** (2009). Host inhibition of a bacterial virulence effector triggers immunity to infection. *Science*. 324, 784–787.

**Nusinow, D. A. et al.** (2011) The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature* .475, 398–402.

**Park M. Y. and Kim S. Y.** (2014). The Arabidopsis J Protein AtJ1 is Essential for Seedling Growth, Flowering Time Control and ABA Response. *Plant Cell Physiol*. 55 2152-2163.

**Petroni, K. and Tonelli, C.** (2011). Recent advances on the regulation of anthocyanin synthesis in reproductive Organs. *Plant Sci*. 181, 219–229.

**Piques M., Schulze W. X., Hohne M., Usadel B., Gibon Y., et al.** (2009). Ribosome and transcript copy numbers, polysome occupancy and enzyme dynamics in Arabidopsis. *Mol Syst Biol*. 5, 314.

- Pooggin M. M.** (2013). How can plant DNA viruses evade siRNA-directed DNA methylation and silencing? *International Journal of Molecular Sciences*. 14, 15233–15259.
- Raja P., Wolf J. N., Bisaro, D. M.** (2010). RNA silencing directed against *geminiviruses*: Post-transcriptional and epigenetic components. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 1799, 337-351.
- Reinbothe C., Pollmann S., Reinbothe S** (2010). Singlet oxygen signaling links photosynthesis to translation and plant growth. *Trends Plant Sci* 15: 499–506.
- Robinson M. D. and Oshlack A.** (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*. 11, 3-25.
- Rocha C.S., Santos, A. A., Machado J. P. B., Fontes, E. P. B.** (2008). The ribosomal protein L10/QM-like protein is a component of the NIK-mediated antiviral signaling. *Virology*. 380, 165–169.
- Rodnina M. V. and W. Wintermeyer.** (2009). 'Recent Mechanistic Insights into Eukaryotic Ribosomes', *Curr Opin Cell Biol*. 21, 435-43.
- Santos A. A., Lopes, K. V. G., Apfata, J. C., Fontes, E. P. B.** (2010). NSP-interacting kinase, NIK: A transducer of plant defence signalling. *Journal of Experimental Botany*. 61, 3839–3845.
- Santos, A. A., Carvalho, C. M., Florentino, L. H., Ramos, H. J. O., Fontes, E. P. B.** (2009). Conserved threonine residues within the A-loop of the receptor NIK differentially regulate the kinase function required for antiviral signaling. *Plos One*. 4, 1–15.
- Schmieder R. and Edwards R.** (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics* (Oxford, England). 27, 863–864.
- Scholthof K. B. G., Adkins S., Czosnek, H., et al.** (2011). Top 10 plant viruses in molecular plant pathology. *Molecular Plant Pathology*. 12, 938–954.
- Schroda M., Vallon O., Wollman F. A., Beck C. F.** (1999). A chloroplast-targeted heat shock protein 70 (HSP70) contributes to the photoprotection and repair of photosystem II during and after photoinhibition. *Plant Cell*. 11, 1165–1178.
- Shimizu T., Yoshii A., Sakurai K., Hamada K., Yamaji Y., Suzuki M., Namba S. Hibi T.** (2009). Identification of a novel tobacco DnaJ-like protein that interacts with the movement protein of Tobacco mosaic virus. *Arch. Virol*. 154, 959–967.
- Shobhit G., Stamatoyannopolous J. A., Timothy B., William S. N.** (2007). "Quantifying similarity between motifs", *Genome Biology*. 8, R24.
- Simon, A.** A quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed 17 May 2017.
- Song L., Huang S. C., Wise A., et al.** (2016). A transcription factor hierarchy defines an environmental stress response network. *Science*. 354, (6312).
- Stotz H. U., Mitrousis G. K., De Wit P. J. G. M., Fitt B. D. L.** (2014). Effector-triggered defence against apoplastic fungal pathogens. *Trends Plant Science*. 19, 491-500.

- Takahashi S., Murata N.** (2006). Glycerate-3-phosphate, produced by CO<sub>2</sub> fixation in the Calvin cycle, is critical for the synthesis of the D1 protein of photosystem II. *Biochimica et Biophysica Acta – Bioenergetics*. 1757, 198–205.
- Tian H., Wang X., Hongyan G., Cheng Y., Hou C., Chen Jin-Gui., Wang S.** (2017). NTL8 regulates trichome formation in Arabidopsis by directly activating R3 MYB genes 9 TRY and TCL1. *Plant Physiology*. 174.
- Umezawa, T., Yoshida, R., Maruyama, K., Yamaguchi-Shinozaki, K., Shinozaki, K.** (2004). SRK2C, a SNF1-related protein kinase 2, improves drought tolerance by controlling stress-responsive gene expression in Arabidopsis thaliana. *Proc Natl Acad Sci*. 101, 17306-17311.
- Van der Linde K., Gutsche N., Leffers H. M., Lindermayr C., Müller B., Holtgreffe S., Scheibe R.** (2011). *Plant Physiol Biochem*. 49, 946-57.
- Wang W., Vinocur B., Shoseyov O., Altman A.** (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in Plant Science* 9, 244–252.
- Willis C. G., Ruhfel B., Primack R. B., Miller-Rushing A. J., Davis C. C.** (2008). Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change. *Proc. Natl Acad. Sci*. 105, 17029–17033.
- Yamaguchi, N., Winter, C. M., Wu, M.-F., Kwon, C. S., William, D. A., & Wagner, D.** (2014). PROTOCOLS: Chromatin Immunoprecipitation from Arabidopsis Tissues. The Arabidopsis Book / American Society of Plant Biologists. 12, 170.
- Yang W1., Zhou X1., Zimmermann.** (2016). HR1, Cavener DR2, Klann E3, Ma T4. Repression of the eIF2 $\alpha$  kinase PERK alleviates mGluR-LTD impairments in a mouse model of Alzheimer's disease. *Neurobiol Aging*. 41, 19-24.
- Ye Y., Rape M.** (2009). Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol*. 10, 755–764.
- Ying-Zhong Ma., Rebekah A., Miller G. R., Fleming., Matthew B. F.** (2008). Energy Transfer Dynamics in Light-Harvesting Assemblies Templated by the Tobacco Mosaic Virus Coat Protein. *The Journal of Physical Chemistry*. 22, 6887-6892.
- Zhao H., Lou Y., Sun H., et al.** (2016). Transcriptome and comparative gene expression analysis of Phyllostachys edulis in response to high light. *BMC Plant Biology*. 16, 34.
- Zhen L., Jin-Ping Y., De-Kuan Li., Qin L., Qiujie Y., Zhi-Bin L., Li-Ming Y., Jian-Mei W., Xu-Feng L., Yi Y.** (2015). UDP-Glucosyltransferase71C5, a Major Glucosyltransferase, Mediates Abscisic Acid Homeostasis in Arabidopsis. *Plant Physiology* Apr, 167 (4) 1659-1670. but not rice, affects trichome and root hair formation. *Scientific Reports*. 6, 19254.
- Zheng K., Tian H., Hu Q., Guo H., Yang L., Cai L., Wang X., Liu B., Wang S.** (2016). Ectopic expression of R3 MYB transcription factor gene OsTCL1 in Arabidopsis, but not rice, affects trichome and root hair formation. *Scientific Reports*. PMC4725938

**Zheng Y., Anderson S., Zhang Y., Garavito R.M.** (2011). The structure of sucrose synthase-1 from *arabidopsis thaliana* and its functional implications. *J. Biol. Chem.* 286 (41), 36108-36118.

**Zhou B., Mural R. V., Chen X., Oates M. E., Connor R. A., Martin G. B., Zeng L.** (2017). A Subset of Ubiquitin-Conjugating Enzymes Is Essential for Plant Immunity. *Plant Physiology.* 173, 1371–1390.

**Zorzatto C., Machado J.P.B., Lopes K.V.G., et al.** (2015). NIK1-mediated translation suppression functions as a plant antiviral immunity mechanism. *Nature.* 520, 679–682.