Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”

Perfil transcricional da xilogênese em *Eucalyptus grandis*

Ana Paula Chiaverini Pinto

Tese apresentada para obtenção do título de Doutora em
Ciências. Área de concentração: Genética e
Melhoramento de Plantas

Piracicaba
2017

Ana Paula Chiaverini Pinto
Engenheira Agrônoma

Perfil transcricional da xilogênese em *Eucalyptus grandis*
versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:
Prof. Dr. **CARLOS ALBERTO LABATE**

Tese apresentada para obtenção do título Doutora em
Ciências. Área de concentração: Genética e
Melhoramento de Plantas

Piracicaba
2017

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Pinto, Ana Paula Chiaverini

Perfil transcricional da xilogênese em *Eucalyptus grandis*, / Ana Paula Chiaverini Pinto. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2017.

141 p.

Tese (Doutorado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. Elementos traqueais 2. Parede celular secundária 3. RNA-seq 4. Expressão gênica I. Título

Dedicatória

Aos meus queridos pais, Meire e Paulo, pela presença, suporte e incentivo para me tornar um ser humano melhor, sendo sempre exemplos de caráter e conduta.

Ao Raphael – meu companheiro da vida - pelo apoio incondicional e amor verdadeiro do dia-a-dia.

À Deus e a Seu universo ainda tão misterioso e fascinante para mim.

Agradecimentos

À Escola Superior de Agricultura “Luiz de Queiroz”;

Ao Prof. Dr. Carlos Alberto Labate pela oportunidade, orientação e profissionalismo;

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de doutorado e suporte financeiro;

Ao pesquisador Dr. Luiz Pedro Barreto Cid pelos conhecimentos e colaboração inicial no projeto na área de cultura de tecidos vegetais;

À Livia Maria Franceschini pela amizade, constante disponibilidade e ajuda essencial durante as análises de bioinformática;

Ao Prof. Gabriel Rodrigues Alves Margarido e ao Dr. Diego Mauricio Riano-Pachon pela atenção e auxílio na resolução de dúvidas sobre as análises de bioinformática;

À querida Thaís R. Cataldi por sua amizade e colaboração constante durante todo o desenvolvimento do meu trabalho, principalmente na revisão final da tese;

Ao Dr. Vinicius Augusto Carvalho de Abreu pela amizade, disponibilidade, auxílio nas análises dos dados e sugestões para a tese;

À Hana Karina Pereira da Silva pela ajuda com minhas infinitas dúvidas e pelas valiosas sugestões na redação da tese;

Aos meus queridos amigos e amigas do Laboratório Max Feffer de Genética de Plantas que tornaram esses anos de convívio tão especiais e inesquecíveis.

Gostaria de agradecer sinceramente a ajuda e amizade de todos, em especial a(o):

Juliana Guimarães Fonseca, Andressa Peres Bini, Flávia de Moraes Franco, Felipe Garbelini Marques, Ilara Gabriela Frasson Budzinski, Maria Leticia Bonatelli, Fabricio Edgar de Moraes, Vanessa Iurif, Ivan Miletovic Mozol, Mariana de Sousa e Silva, Lucas Souza Lopes e Mônica T. Veneziano Labate.

Enfim, a todos que contribuíram de alguma forma para a realização deste trabalho, meu muito obrigada!!!

Epígrafe

“A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que ninguém ainda pensou sobre aquilo que todo mundo vê.”

Arthur Schopenhauer

Sumário

RESUMO	7
ABSTRACT	8
1. INTRODUÇÃO E OBJETIVOS.....	9
2. REVISÃO BIBLIOGRÁFICA.....	13
2.1. O GÊNERO <i>EUCALYPTUS</i> E O SETOR FLORESTAL BRASILEIRO	13
2.2. PROCESSO <i>KRAFT</i> DE EXTRAÇÃO DE CELULOSE.....	15
2.3. PAREDE CELULAR VEGETAL: UMA VISÃO GERAL	17
2.4. PAREDE CELULAR SECUNDÁRIA: BIOSÍNTESE E REGULAÇÃO TRANSCRICIONAL.....	22
2.5. XILOGÊNESE E A DIFERENCIAÇÃO <i>IN VITRO</i> DE ELEMENTOS TRAQUEAIS	29
2.6. RNA-SEQ E TRANSCRIPTÔMICA.....	35
3. MATERIAIS E MÉTODOS	43
3.1. CULTIVO <i>IN VITRO</i> E COLETA DAS AMOSTRAS.....	43
3.2. EXTRAÇÃO DO RNA TOTAL, QUANTIFICAÇÃO E QUALIDADE	45
3.3. PREPARAÇÃO DA BIBLIOTECA E SEQUENCIAMENTO DO RNA	45
3.4. MAPEAMENTO DAS LEITURAS E ANÁLISE DOS DADOS.....	46
3.5. EXPRESSÃO DIFERENCIAL DE GENES, ANOTAÇÃO E CATEGORIZAÇÃO FUNCIONAL.....	46
4. RESULTADOS	49
4.1. OBTENÇÃO E CARACTERIZAÇÃO DOS ELEMENTOS TRAQUEAIS DE <i>EUCALYPTUS GRANDIS</i>	49
4.2. QUALIDADE DO RNA TOTAL EXTRAÍDO	51
4.3. SUMÁRIO DOS DADOS DE SEQUENCIAMENTO E MAPEAMENTO	51
4.4. IDENTIFICAÇÃO DOS GENES DIFERENCIALMENTE EXPRESSOS (GDEs)	52
4.5. ANOTAÇÃO FUNCIONAL E ANÁLISE DE ENRIQUECIMENTO DOS TERMOS GO	57
4.6. FATORES DE TRANSCRIÇÃO DE INTERESSE	66
4.7. GENES RELACIONADOS À BIOSÍNTESE DOS COMPONENTES DA PAREDE CELULAR SECUNDÁRIA	72
5. DISCUSSÃO	85
5.1. OBTENÇÃO <i>IN VITRO</i> E CARACTERIZAÇÃO MORFOLÓGICA DOS ELEMENTOS TRAQUEAIS DE <i>EUCALYPTUS GRANDIS</i>	85
5.2. PERFIL DA EXPRESSÃO GÊNICA E CLASSIFICAÇÃO FUNCIONAL	88
5.3. REGULAÇÃO TRANSCRICIONAL DA BIOSÍNTESE DA PAREDE CELULAR SECUNDÁRIA	92
5.4. GDEs ENVOLVIDOS NA XILOGÊNESE DE <i>EUCALYPTUS GRANDIS</i>	97
5.5. ENGENHEIRANDO A DEPOSIÇÃO DA PAREDE CELULAR SECUNDÁRIA.....	103
6. CONCLUSÕES.....	105
REFERÊNCIAS.....	107
ANEXOS	125

RESUMO

Perfil transcricional da xilogênese em *Eucalyptus grandis*

O eucalipto (*Eucalyptus* spp.) é a arbórea comercial mais importante do Brasil, tendo um papel marcante na indústria de papel e celulose. O aumento no rendimento da extração da polpa celulósica e a redução nos custos de produção são os principais objetivos dos programas de melhoramento genético de eucalipto. Como o processo de deslignificação é fundamental para a extração da celulose, é suma importante compreender a biossíntese e deposição da parede celular secundária durante a xilogênese. A diferenciação dos elementos traqueais é caracterizada pela deposição da parede celular secundária, logo estudos de expressão gênica envolvendo essas células xilemáticas é uma estratégia interessante. Foi realizado o sequenciamento do transcriptoma de células em suspensão em diferentes fases de desenvolvimento (células meristemáticas, células alongadas e elementos traqueais) gerando 348 milhões de leituras *paired-end* com 100 pb cada uma, com a maioria das leituras (80 %) alinhada e mapeada contra genoma de referência. A análise dos genes diferencialmente expressos (GDEs) identificou um total de 3426, 1044 e 4665 GDEs nas comparações M vs. A, A vs. T e M vs. T, respectivamente. Sendo 1316, 567, 2637 genes *up-regulated* e 2110, 476, 2028 genes *down-regulated* nas comparações M/A, T/A, T/M, respectivamente. Um total de 4229 genes foram anotados funcionalmente e mapeados em 310 rotas identificadas pelo KEGG, obtendo-se 40 sequências relacionadas à rota dos fenilpropanoides, a quinta rota mais representativa. Quando foram considerados os GDEs com $\log_2 \text{fold change} \leq -1,5$ e $\log_2 \text{fold change} \geq 1,5$, 34 rotas metabólicas foram identificadas, destas, a principal foi a rota dos fenilpropanoides, com 14 sequências relacionadas a enzima lactoperoxidase. Foi realizado o enriquecimento dos termos GO e nas redes a biossíntese de celulose e xiloglucanos, rotas de sinalização de etileno, rotas de biogênese e organização da parede celular vegetal e de morte celular, entre outras, foram destaque. Identificaram-se 31 fatores de transcrição, sendo 3 fatores de transcrição MYB, 2 fatores de transcrição NAC e 10 fatores de transcrição WRKY. Além disso, 21, 7 e 23 genes da rota da celulose e hemiceluloses e 24, 8 e 21 genes da rota de síntese dos monômeros da lignina foram constatados nas comparações M vs. A, A vs. T e M vs. T, respectivamente. Assim como, foram encontrados genes codificadores de lacases e peroxidases, participantes da polimerização dos monômeros da lignina. Os resultados possibilitaram uma compreensão mais ampla da regulação transcricional da biossíntese da parede celular secundária nas células xilemáticas de eucalipto e revelou novos genes participantes da xilogênese. Este conjunto de dados transcricionais ajudará na realização de manipulações genéticas visando à obtenção de plantas de eucalipto com características de extratibilidade aspiradas pela indústria.

Palavras-chave: Elementos traqueais; Parede celular secundária; RNA-seq; Expressão gênica

ABSTRACT

Transcriptional profile of xylogenesis in *Eucalyptus grandis*

Eucalyptus (*Eucalyptus* spp.) is the most important commercial tree in Brazil, playing a significant role in the pulp and paper industry. An increase in the yield of cellulosic pulp extraction and a reduction of production cost are the main objectives of the eucalyptus breeding programs. As the delignification process is fundamental for the extraction of cellulose, it is highly important to understand the biosynthesis and deposition of the secondary cell wall during xylogenesis. The differentiation of the tracheary elements is characterized by the deposition of the secondary cell wall, so studies of gene expression involving these xylem cells is an interesting strategy. Sequencing of the transcriptome of cells in suspension at different stages of development (meristematic cells, elongated cells and tracheal elements) was performed generating 348 million paired-end reads with 100 bp each, with most readings (80%) aligned and mapped against reference genome. The analysis of differentially expressed genes (DEG) identified a total of 3426, 1044 and 4665 DEGs in the comparisons M vs. A, A vs. T and M vs. T, respectively. Being 1316, 567, 2637 up-regulated genes and 2110, 476, 2028 down-regulated genes in the M/A, T/A, T/M, respectively. A total of 4229 genes were functionally annotated and mapped in 310 routes identified by the KEGG, yielding 40 sequences related to the phenylpropanoid route, the fifth most representative route. When DEGs with $\log_2 \text{fold change} \leq -1.5$ e $\log_2 \text{fold change} \geq 1.5$ were considered, 34 metabolic routes were identified. From these, the main route was the phenylpropanoids pathway, with 14 sequences correlated to the enzyme laccase. The GO terms enrichment was carried out and the networks of cellulose and xyloglucans biosynthesis, ethylene signaling routes, biogenesis routes and the plant cell wall organization and cell death, among others, were highlighted. Thirty-one transcription factors were identified, being 3 MYB transcription factors, 2 NAC transcription factors and 10 WRKY transcription factors. In addition, 21, 7 and 23 genes of the cellulose and hemicelluloses pathway and 24, 8 and 21 genes of lignin monomer synthesis route were found in the M vs. A, A vs. T and M vs. T comparisons, respectively. Besides this, encoding genes laccases and peroxidases were found, participating in the polymerization of lignin monomers. The results allowed a broader understanding of the transcriptional regulation of secondary cell wall biosynthesis in eucalyptus xylem cells and revealed new genes participating in xylogenesis. This set of transcriptional data will help in the accomplishment of genetic manipulations in order to obtain eucalyptus plants with extractability characteristics aspired by the industry.

Keywords: Tracheary elements; Secondary cell wall; RNA-seq; Gene expression

1. INTRODUÇÃO E OBJETIVOS

O setor florestal desempenha um importante papel no cenário socioeconômico do nosso país, contribuindo com a produção de bens e serviços, agregação de valores aos produtos florestais, assim como para a geração de empregos, divisas, tributos e rendas. As espécies florestais plantadas são destinadas aos mais diversos fins como para produção de painéis de madeira, pisos laminados, celulose, papel, produção energética e biomassa. Sendo esse setor responsável por 91% de toda a madeira produzida para fins industriais no País – os demais 9% vem de florestas nativas legalmente manejadas (IBÁ, 2015).

Seguramente, o Brasil detém um dos mais avançados programas de melhoramento florestal do mundo, sendo o eucalipto o seu principal componente. A cultura do eucalipto no Brasil encontrou bases sólidas para o seu desenvolvimento, inicialmente, motivada pelo potencial de uso da madeira de algumas espécies como fonte de biomassa para combustível (lenha, principalmente) e pelo sucesso de seu emprego apropriado como matéria-prima para a produção de celulose e papel.

A produção nacional de celulose em 2015, considerando-se fibra curta (eucalipto) e longa (pinus) e pasta de alto rendimento, foi de 17,3 milhões de toneladas, 5,2% maior do que em 2014. Com esse resultado, o Brasil manteve-se no quarto lugar no ranking dos países produtores de celulose de todos os tipos e como primeiro produtor mundial de celulose de eucalipto, com exportações na ordem de 11,5 milhões de toneladas.

Na cadeia produtiva, durante o processo de extração da celulose, o branqueamento é uma etapa fundamental que consiste no tratamento da fibra celulósica com produtos químicos para a remoção da lignina, responsável pela tonalidade escura da polpa celulósica. Esse processo de deslignificação é importante não só para a obtenção de uma celulose pura, mas também para proporcionar um aspecto de alvura elevado, característica fundamental de um produto final com alta qualidade.

De acordo com Rowell, (2013) dependendo do grau de cozimento efetuado, uma polpa tratada pelo processo *Kraft* pode conter cerca de 6% de lignina (baseada em sua massa seca), enquanto a não branqueada pode conter até 27% de lignina na sua composição. É importante salientar que a intensidade do branqueamento é determinada em função da finalidade a que se destina a celulose, bem como dos custos inerentes ao processo de branqueamento.

Considerando esse panorama do setor de papel e celulose, os estudos relacionados ao processo de lignificação em espécies arbóreas, como o eucalipto, tem sido cada vez mais importantes para compreender a dinâmica fisiológica e genética envolvida nesse processo, e assim, tentar otimizar o processo de deslignificação, e consequentemente, a extração da celulose.

Alguns estudos relacionados à genômica funcional da xilogênese e a identificação de genes envolvidos na rota biossintética da celulose e da lignina utilizam as suspensões celulares como metodologia, por constituírem uma fração celular homogênea, possibilitando o estudo bioquímico e molecular da formação do xilema de maneira mais simplificada, sem a complexidade apresentada pelos tecidos vegetais já formados.

Os elementos traqueais, mantidos em suspensões celulares, são os alvos desse estudo envolvendo a xilogênese, por constituírem as células do xilema. A formação desses elementos traqueais envolve uma série de mudanças citológicas, dentre elas, o rearranjo dos microtúbulos para o posicionamento e espessamento da parede celular secundária. Contudo, definitivamente, o processo de lignificação que é iniciado após o processo de espessamento, e a morte celular programada (MCP) são as mais importantes etapas, pois caracterizam a formação do elemento traqueal maduro (McCann, 1997).

Nesse contexto, estudos de expressão gênica com várias espécies vegetais, que utilizam as suspensões celulares, tem contribuído para fornecer dados a respeito de uma gama de genes candidatos e fatores de transcrição relacionados à diferenciação dos elementos traqueais e, por conseguinte, relacionados à via de biossíntese da celulose e da lignina (Demura, 2014; Lacayo et al., 2010; Ohashi-Ito et al., 2010; Pyo et al., 2007).

Conjuntamente com esses estudos, a utilização da transcriptômica possibilita um estudo mais refinado e preciso da expressão gênica em cada etapa desse processo. Usando a técnica de RNA-seq, estudos tem sido dirigidos com o objetivo de investigar o perfil de expressão durante estádios específicos do desenvolvimento do xilema em eucalipto identificando os processos biológicos mais relevantes para a formação da madeira e sua regulação transcricional (Paiva et al., 2011; Salazar et al., 2011; Wang et al., 2016).

Dessa forma, inicialmente foi realizada a germinação *in vitro* de sementes de *Eucalyptus grandis* para a indução de calos friáveis utilizados nas suspensões celulares. O estudo de expressão gênica diferencial englobou o desenvolvimento da metodologia de indução de células na etapa inicial de diferenciação (células meristemáticas iniciais - M), na

etapa intermediária (células alongadas - A) e os elementos traqueais (T) que apresentaram uma parede celular secundária bem definida e lignificada. A identificação dos genes que participam dos vários estádios de desenvolvimento do xilema foi conduzida por meio da análise do transcrito dessas células, utilizando a técnica de RNA-seq.

Os objetivos principais do trabalho englobaram a otimização da formação *in vitro* de calos friáveis de *Eucalyptus grandis* e consequentemente, a obtenção de suspensões celulares com elementos traqueais bem caracterizados. A realização da separação fracionada dos diferentes tipos ou estádios celulares (M, A e T) com a extração do RNA total dessas células para a realização do sequenciamento do RNA (RNA-seq) por meio da plataforma Illumina. Estudar o transcriptoma obtido, utilizando a bioinformática como ferramenta para identificar os genes e fatores de transcrição diferencialmente expressos e realizar a categorização funcional dos mesmos, e por fim, validar os principais genes diferencialmente expressos por PCR quantitativo em tempo real (RT-qPCR).

2. REVISÃO BIBLIOGRÁFICA

2.1. O gênero *Eucalyptus* e o setor florestal brasileiro

O gênero *Eucalyptus* pertence à família *Myrtaceae*, que apresenta uma ampla variedade de espécies, inclusive exemplares de plantas frutíferas como o jambeiro, a pitangueira, a goiabeira, a jabuticabeira, entre outros. O plantio sistemático do eucalipto começou nas primeiras décadas do século XIX e no século XX, já havia se tornado a espécie florestal mais plantada mundialmente. Os primeiros eucaliptos que chegaram ao Brasil foram plantados no Jardim Botânico do Rio de Janeiro como planta ornamental em 1825, e depois de 43 anos começaram a ser plantados no Rio Grande do Sul para a produção de lenha e formação de quebra-ventos (Suzano Papel e Celulose, 2008). No Brasil, os projetos de reflorestamento tiveram início com a introdução do eucalipto em 1904, como matéria-prima destinada à produção de lenha e dormentes no Estado de São Paulo (Mora and Garcia, 2000).

As principais espécies cultivadas atualmente no Brasil incluem o *Eucalyptus grandis*, *Eucalyptus camaldulensis*, *Eucalyptus saligna* e o *Eucalyptus urophylla*, entre outras. E além dessas espécies citadas, foram desenvolvidos também cruzamentos entre as espécies, resultando em híbridos, como é o caso do *Eucalyptus urograndis* (*E. grandis* X *E. urophylla*) (Biotecnologia, 2009).

É importante salientar que as árvores plantadas representam uma matéria-prima renovável e reciclável ao meio ambiente, à biodiversidade e à vida humana, contribuindo significativamente para a melhoria da qualidade de vida, pois além de gerar empregos diretos e indiretos, possibilitam benefícios econômicos, sociais e principalmente ambientais, pois protegem o solo e as nascentes de água, sequestram carbono e amenizam o clima, embelezam a paisagem e, ademais, fornecem madeira (IBÁ, 2015; Queiroz, L. R. de S.; Barrichelo, 2008).

Tendo como enfoque o nosso país, no ano de 2014, a área existente com plantio de essências florestais no Brasil foi de 9.364.427 hectares, dos quais 6.951.145 hectares foram plantados com *Eucalyptus* spp., 2.048.264 hectares com *Pinus*; e o restante com outras espécies. A região Sul detém a maior área plantada para silvicultura, com mais de 3,7

milhões de hectares, principalmente representados pelo plantio de *Pinus*, seguida pela região Sudeste, que apresenta 3 milhões de hectares plantados (2,9 milhões com eucalipto, 178.974 hectares com pinus e 4.107 hectares com outras espécies). Essa maior concentração de plantios florestais nas Regiões Sudeste e Centro-Oeste do País (cerca de 71%) é justificada pelo fato dessas regiões possuírem as principais unidades industriais do segmento florestal (IBGE, 2014).

A madeira oriunda das plantações florestais é utilizada para diversos fins, tais como: produção de chapas, lâminas, compensados, aglomerados, carvão vegetal, madeira serrada, móveis, e principalmente, para o segmento de papel e celulose, o qual de acordo com a ABRAF concentrou 72,5% da área plantada de eucalipto, seguido pelos segmentos de siderurgia e carvão vegetal (19,5%), painéis de madeira industrializada (7,3%) e produtores independentes (0,7%) (ABRAF, 2013).

A Indústria Brasileira de Árvores destacou que em 2014, o Brasil manteve mais uma vez sua liderança no ranking global de produtividade florestal. A produtividade média dos plantios brasileiros de eucalipto atingiu 36 m³/ha.ano e a produtividade dos plantios de pinus foi de 31 m³/ha.ano. Considerando a cadeia produtiva do setor florestal brasileiro em 2014, o número estimado de empregos mantidos de forma direta pelo setor foi de 610 mil, e o número de postos de trabalho diretos, indiretos e resultantes do efeito renda da atividade florestal foi da ordem de 4,23 milhões (IBÁ, 2015).

O principal produto do ponto de vista econômico obtido do eucalipto é a celulose, que é transformada em matéria-prima de primeira classe para produção de papel de escrita, impressão e *tissue* (fins sanitários). E este setor é tão importante que apresentou um faturamento de aproximadamente 34 bilhões de reais em 2010, um crescimento de 23% em relação ao observado em 2009 (BRACELPA, 2011).

A produção nacional de celulose em 2015, considerando-se fibra curta (eucalipto) e longa (pinus) e pasta de alto rendimento, foi de 17,3 milhões de toneladas, 5,2% maior do que em 2014. Com esse resultado, o Brasil manteve-se no quarto lugar no ranking dos países produtores de celulose de todos os tipos, superado somente pelos Estados Unidos, China e Canadá e como primeiro produtor mundial de celulose de eucalipto. Além disso, os embarques de celulose somaram 11,5 milhões de toneladas, ante 10,6 milhões de toneladas, em 2014, um aumento de 8,6%, tendo como principal mercado a Europa, que respondeu por 38,5% da receita, seguida pela China (33,2%) e América do Norte (17,6%) (IBÁ, 2015).

Outra vertente do setor florestal brasileiro se concentra na produção de papel para diversas finalidades (embalagens, imprimir e escrever, imprensa, fins sanitários, papel cartão) que totalizou 10,35 milhões de toneladas em 2015, mantendo-se praticamente estável em relação a 2014, colocando o Brasil em nono lugar no ranking mundial.

2.2. Processo *Kraft* de extração de celulose

Como já foi comentado, a celulose é a principal *commodity* do setor florestal brasileiro, fato confirmado pelo recorde positivo da balança comercial do setor em 2015, e esse aspecto fundamenta a importância do entendimento e do aperfeiçoamento do seu processo de extração (Shinya et al., 2016).

A celulose é um polímero natural presente nas plantas, sendo um dos principais componentes da célula vegetal, e é o biopolímero mais abundante no nosso planeta. Já a lignina é um polímero que atua como ligante permanente das fibras de celulose e hemiceluloses da parede celular. O processo de extração da celulose, objetiva dissolver a lignina da madeira com a finalidade de liberar tais fibras com o mínimo de degradação dos carboidratos presentes nas paredes celulares (Wiley, 1989).

Dentre os processos químicos disponíveis para a extração da celulose, o mais utilizado no Brasil é o Kraft, em que se obtém a celulose após tratamento químico. Esse processo de polpação Kraft tem como premissa a individualização das fibras da madeira pela remoção da lignina, e se mostrou adaptável tanto para *hardwood* (árvores folhosas) como para *softwood* (árvores coníferas). Com destaque para o eucalipto que tem se mostrado uma das espécies florestais mais adequadas para o processo de polpação Kraft (Santos et al., 2006).

Esse processo de polpação compreende basicamente as seguintes etapas: preparação da madeira, cozimento, pré-lavagem e depuração, deslignificação com O₂ e pós-lavagem, branqueamento, secagem e embalagem da polpa celulósica. Essa deslignificação requer o cozimento da madeira picotada em cavacos e mantidas em um digestor com a utilização de água, produtos químicos (hidróxido de sódio e sulfeto de sódio), pressão e temperaturas da ordem de 150 °C a 170 °C. Nesse processo produtivo, a deslignificação com O₂ é indispensável para a remoção da lignina residual que permanece na polpa mesmo após a etapa alcalina (Li et al., 2016a).

Como já mencionado os pontos críticos envolvidos no processo de polpação são a deslignificação e o branqueamento, sendo ambos problemáticos devido aos resíduos de lignina. Meng et al. (2015) ressaltam que os processos mais modernos de branqueamento demandam uma alta alvura e estabilidade nas polpas produzidas, com a preservação das propriedades de resistência física intrínsecas da mesma. Ao mesmo tempo em que se busca minimizar os compostos químicos, a energia e o consumo de água dispendidos na produção da celulose.

É importante salientar que mesmo quando a composição química da madeira é conhecida é difícil prever o seu comportamento durante a polpação uma vez que os componentes se encontram arrançados de forma complexa tornando o comportamento diferente de quando estão isolados. De qualquer forma, o conhecimento da composição química da madeira em seus componentes principais como celulose, lignina e extrativos é importante para interpretar o comportamento da madeira no processo de polpação, assim como para determinar a qualidade da polpa produzida (Dos Santos and Sansígolo, 2007).

É sabido que um baixo teor de lignina e um alto teor de carboidratos requererão condições menos severas de polpação e conduzirão a um alto rendimento gravimétrico (Baucher et al., 1998). Por essa razão, vários trabalhos tem buscado a caracterização de enzimas e genes envolvidos nessa rota metabólica visando a melhor compreensão dos mecanismos de regulação gênica envolvidos com a biossíntese da parede celular vegetal e seus componentes (Harakava, 2005; Mandrou et al., 2012; Wong et al., 2011; Yoon et al., 2015).

O entendimento e a manipulação dos elementos envolvidos na regulação desses genes podem possibilitar a obtenção de plantas de eucalipto com quantidades de lignina mais adequadas para fins industriais, podendo assim reduzir os custos do setor produtivo. Trabalhos com o objetivo de reduzir ou modificar o conteúdo de lignina através da sub-regulação de vias de biossíntese de lignina mostraram-se promissores (Baucher et al., 1998; Boerjan et al., 2003; Vanholme et al., 2008).

Um exemplo envolvendo a regulação na expressão de genes da rota de biossíntese de lignina foi realizado por Ziebell et al. (2016) que reduziram a expressão dos genes 4-Hidroxicinamato 3-Hidroxilase (*C3H*) e Cinamato 4-Hidroxilase (*C4H*) pela técnica de RNAi em *Eucalyptus urophylla* x *Eucalyptus grandis*. Ambos os genes estão envolvidos na produção dos monômeros da lignina: álcoois cumarílico, coniferílico e sinapílico. Os autores

observaram um aumento significativo na extratibilidade dos biopolímeros da parede celular (~50%), que estaria relacionado com a alteração na constituição da parede celular, pressupondo que a redução no conteúdo de lignina acarretou na formação de uma parede celular menos rígida e mais propícia aos processos extrativos.

2.3. Parede celular vegetal: uma visão geral

A parede celular é um componente fundamental das células vegetais sendo essencial para diversas atividades vitais à célula, incluindo a forma celular, que é necessária na formação dos tecidos e órgãos das plantas. Formando a interface entre células adjacentes, a parede celular vegetal muitas vezes desempenha um papel importante na comunicação intercelular, e devido a sua localização superficial a parede celular também tem um papel importante nas interações planta-patógeno (Keegstra, 2010).

As paredes celulares de plantas são predominantemente compostas por uma matriz complexa de polissacarídeos, dos quais a celulose é a mais abundante. As células vegetais controlam as propriedades mecânicas da parede organizando a síntese e a deposição de polímeros de parede e modificando a arquitetura da parede de acordo com as necessidades da célula. As paredes celulares primárias (PCPs) tem um arranjo multilaminado de microfibrilas de celulose para acomodar o seu crescimento enquanto mantem sua integridade estrutural. Certos tipos de células, tais como células de fibras e células vasculares, formam subsequentemente paredes celulares secundárias (PCs) espessas com alto teor de celulose e lignina para sustentar estruturalmente a planta e facilitar o transporte de água e solutos (Li et al., 2016b).

Albersheim et al., (2010) conceituam a parede primária (PCP) como uma estrutura mais fina, entre 50 a 200 nm, separada das células vizinhas pela lamela média e composta por uma proporção relativamente constante das quatro principais classes de macromoléculas estruturais: as proteínas e os polissacarídeos celulose, hemicelulose e pectina, apresentando na sua composição 20-30% de celulose, 30-35% de pectina, 25-30% de hemiceluloses e 1-10% de proteína (Tabela 1). Salientando que a parede primária é capaz de sustentar a expansão celular e o crescimento até que em um momento da diferenciação celular esse crescimento cessa e novas camadas são depositadas nessa parede, sempre no

lado próximo a membrana plasmática, criando o que foi denominado de parede celular secundária.

A parede celular secundária (PCS) contém uma estrutura diferente na composição de seus polímeros, a maioria desses polímeros é adicionada à parede primária durante sua transição para parede secundária. A parede secundária possui na sua composição uma quantidade maior de celulose e uma alta porcentagem de lignina, biopolímero geralmente ausente nas paredes primárias. Juntas, a celulose e lignina perfazem mais de 60% e em alguns casos, até 80% da parede celular secundária. As paredes secundárias ainda possuem um adicional de hemiceluloses e relativamente menos pectina que as primárias (Fagerstedt, 2015; Li et al., 2016c).

Tabela 1. Diferença na composição dos polímeros presentes nas paredes primárias e secundárias (adaptada de Albersheim et al., 2010).

Polímeros	Parede celular (%) ^a		
	Primária	Secundária	
	Dicotiledôneas	Angiospermas	Gimnospermas
Celulose	20-30	37-57	38-52
Lignina	0	17-30	26-36
Pectina	30-35	<10	<10
Hemicelulose	25-30	20-37	16-27

^a Os valores apresentados na tabela são uma média das porcentagens obtidas de vários tecidos de múltiplas espécies.

A complexidade da formação e deposição da parede celular, com todas as nuances da biossíntese dos seus componentes é muito estudada. Como já comentado, dentre os polissacarídeos existentes na parede celular a celulose é o mais importante e normalmente o mais abundante. Na parede celular vegetal as microfibrilas de celulose são incorporadas em uma matriz de polissacarídeos complexos, os quais são divididos em duas classes: as pectinas e as hemiceluloses (Lerouxel et al., 2006; Smith, 2001).

Na parede primária complexos enzimáticos de celulose sintase (CCSs) estão aderidos à membrana plasmática. Esses complexos são visualizados como rosetas hexaméricas de aproximadamente 25-30 nm, cada roseta é constituída por 6 subunidades menores, cada uma formada por seis proteínas CESA, totalizando 36 proteínas CESA formando essa

maquinaria de síntese de celulose. Essas proteínas CESA presentes na membrana plasmática são codificadas pela família gênica *CESA*. Quando a uridina difosfato D-glicose (UDP-glicose) é usada como substrato, cada subunidade da CESA pode sintetizar uma cadeia β -(1-4)D-glucano que constituirão arranjos paralelos com 36 cadeias de β -(1-4)D-glucanos para formarem as microfibrilas de celulose (Lerouxel et al., 2006; Taiz and Zeiger, 2010).

Estudos genéticos revelaram que diferentes isoformas das proteínas de celulose sintase (CESA) são responsáveis pela síntese de celulose durante a síntese da PCP e da PCS. As proteínas CESA1, -3 e -6 colaboram na síntese de celulose na PCP, enquanto as proteínas CESA4, -7 e -8 são necessárias para a síntese de celulose na PCS. É importante lembrar que a formação da celulose envolve não apenas a síntese de glucano, mas também a cristalização de cadeias múltiplas de glucano dentro da microfibrila. Estudos para compreender como esse processo é controlado mostraram que a direção da deposição das microfibrilas é orientada pelos microtúbulos adjacentes a membrana (Taylor et al., 2003; Watanabe et al., 2015).

As hemiceluloses são polissacarídeos de ligação que juntamente com a celulose forma uma rede forte, ainda que flexível (Cosgrove, 2005). Xiloglucano e arabinoxilana são duas das hemiceluloses mais abundantes. Como a celulose, o xiloglucano possui uma estrutura básica de resíduos D-glucano com ligações β -(1-4). Contudo, diferentemente da celulose, esse polissacarídeo tem cadeias laterais curtas que contêm xilose, galactose e muitas vezes, uma fucose terminal. Variando com o estado de desenvolvimento e com a espécie vegetal, a fração de hemiceluloses da parede pode conter também grandes quantidades de outros polissacarídeos importantes, como glucuronoarabinoxilanas e glucomananos (Taiz and Zeiger, 2010).

As pectinas são um grupo complexo e heterogêneo de polissacarídeos que se apresentam ligados covalentemente entre si, e contêm açúcares ácidos, como o ácido galacturônico, e açúcares neutros como a ramnose, galactose e arabinose. Uma das pectinas mais abundantes é o polissacarídeo complexo ramnogalacturonano I (RG I), que consiste em resíduos alternativos de ácido galacturônico e ramnose, e provavelmente tem ramos laterais que contêm outros domínios pectina. Os polissacarídeos pécticos podem ser ainda mais complexos, um exemplo disso é o complexo ramnogalacturonano II (RG II) que contém pelo menos 11 resíduos de açúcares e forma dímeros com ligações cruzadas de diéster de borato (Cosgrove, 2005; Vincken, 2003).

As pectinas neutras arabinano e arabinogalactanos também podem estar ligadas às pectinas ácidas promovendo a flexibilidade da parede e ligação à superfície da celulose. Além disso, as pectinas formam géis hidratados que empurram as microfibrilas facilitando seu deslizamento durante o crescimento celular, e as bloqueiam quando o crescimento cessa (Taiz and Zeiger, 2010; Zykwinska et al., 2005).

Resumindo, a celulose constitui o componente principal das paredes primária e secundária, mas a celulose nessas paredes apresenta diferenças estruturais. Na PCP, a celulose apresenta um baixo grau de polimerização e largura de microfibrila de 2-2,5nm. Em contraste, na PCS, a celulose tem um elevado grau de polimerização e largura de microfibrilas de 5-10 nm. Outros componentes da parede celular também diferem entre a PCP e a PCS. Por exemplo, a primária contém tipicamente xiloglucano como a principal hemicelulose (isto é, o componente polissacarídeos que é solúvel em álcali), mas a PCS contém xilana. Além disso, a PCP é rica em pectina semelhante a um gel, enquanto que a lignina e polímeros fenólicos específicos são mais abundantes na PCS (Figura 1). Estas diferenças na composição da parede celular conferem propriedades físicas diferentes à parede celular, tornando a PCP flexível e a PCS mecanicamente e biologicamente robusta (Nakano et al., 2015).

Figura 1. Principais diferenças entre a parede celular primária e a secundária. (A) Modelo da parede celular primária. As microfibrilas de celulose na parede celular primária são relativamente curtas e finas, comparadas com aquelas na parede celular secundária, e a hemicelulose na parede celular primária é composta de xiloglucano. A parede celular primária é rica em pectina. (B) Modelo da parede celular secundária, que é depositada entre a parede celular primária e a membrana plasmática. A parede celular secundária contém principalmente microfibrilas de celulose relativamente longas e espessas, xilana hemicelulósica e lignina (Nakano et al., 2015).

A síntese da celulose e dos componentes da matriz polissacarídica envolve rotas distintas (Figura 2). A celulose é sintetizada por grandes complexos de membrana que produzem as microfibrilas que são depositadas diretamente na parede, enquanto os polissacarídeos são sintetizados no Complexo de Golgi e embalados em pequenas vesículas que se fundem com a membrana plasmática, sendo assim entregues a parede. A matriz de polissacarídeos se torna integrada à rede da parede por interações físicas e ligações enzimáticas, e diferentemente das microfibrilas de celulose, os polissacarídeos da matriz recém-sintetizados podem se difundir dentro da parede da célula também pelo auxílio da sua pressão de turgescência. (Cosgrove, 2005; Lerouxel et al., 2006; Li et al., 2016c; Wang et al., 2013).

Figura 2. Representação esquemática dos eventos-chave na biossíntese da parede celular. A biossíntese de celulose ocorre na membrana plasmática em grandes complexos denominados rosetas. A síntese de polissacarídeos e glicoproteínas de matriz ocorre no Complexo de Golgi onde os produtos se acumulam no lúmen antes do transporte para a parede celular via vesículas. Abreviaturas utilizadas na figura: CesaA: proteínas de celulose sintase que formam a roseta; açúcar NDP: açúcares de nucleotídeos que atuam como doadores para os açúcares que formam os polissacarídeos; Csl: proteínas semelhantes à celulose sintase que são conhecidas por estarem envolvidas na biossíntese de hemicelulose (Keegstra, 2010).

2.4. Parede celular secundária: biossíntese e regulação transcricional

Após finalizada a expansão das paredes primárias, as células podem continuar a sintetizar outra parede, conhecida como secundária. As paredes celulares secundárias (PCSs) são produzidas por tipos especializados de células vegetais e são particularmente importantes para o suporte mecânico e transporte de água. Mecanicamente, as PCSs podem ser comparadas com um concreto reforçado, com as microfibrilas de celulose análogas a varas de concreto que estão embebidas em uma matriz amorfa de lignina. Os componentes principais das paredes secundárias são celulose, hemiceluloses (xilana e glucomanano) e lignina. A composição de paredes secundárias, isto é, a proporção de celulose, hemiceluloses e lignina, pode variar entre as diferentes espécies de plantas e mesmo em diferentes tipos de células contendo parede secundária de uma mesma espécie. De maneira geral, as PCSs típicas são compostas por celulose (40-80%), hemiceluloses (10-40%), lignina (5-25%) e proteínas da parede celular (Kumar et al., 2016; Zhong et al., 2015).

Os tecidos vasculares das plantas são compostos pelo xilema, floema e células procambiais ou cambiais. A proliferação de células-tronco no meristema vascular produz células de progênie, que mantêm as suas propriedades de células-tronco ou se diferenciam

em xilema em direção ao centro e floema em direção à periferia das hastes das plantas (Plomion et al., 2001). O desenvolvimento do sistema vascular e a formação da parede celular secundária são processos biológicos altamente correlacionados e que podem ser regulados pela mesma rota de sinalização.

Como as paredes secundárias formam a madeira e as fibras que são matérias-primas importantes para nossos usos diários e uma fonte potencial para a produção de biocombustíveis, há estudos extensivos sobre os mecanismos moleculares que controlam a composição da parede secundária, a biossíntese e a sua montagem. Para compreender a síntese e a composição das PCSs alguns sistemas modelo foram estudados: o desenvolvimento dos feixes vasculares de *Arabidopsis* (Taylor-Teeple et al., 2015), a formação do xilema em madeiras de álamo (Ye and Zhong, 2015) e a indução da formação das PCSs em sistemas de cultura de tecidos (Lacayo et al., 2010; Oda et al., 2005).

Estudos mostraram que a construção da complexa estrutura da parede celular envolve a execução coordenada de uma série de rotas de biossíntese e de transporte. Além disso, vários membros de uma cascata reguladora participam na iniciação e regulação transcricional da formação da PCS (Didi et al., 2015). A iniciação e biossíntese da parede celular secundária inclui a síntese e interconversão dos precursores dos açúcares, polimerização dos componentes dos grandes polímeros, transporte dos iniciadores ou polímeros da parede celular e finalmente, a montagem dos polímeros pertencentes a estrutura da parede celular.

Conforme mostrado esquematicamente na Figura 3 os complexos celulose sintase são sintetizados pelos ribossomos ligados ao retículo endoplasmático rugoso, onde a montagem do complexo também ocorre, sendo depois transportados para a membrana plasmática pelo complexo de Golgi. Quando são inseridos na membrana plasmática, os CCSs movem-se no plano da membrana e direcionam as cadeias de glucano para dentro das paredes celulares que se ligam em microfibrilas de celulose. As hemiceluloses já são sintetizadas no Golgi pelas glicosiltransferases e as moléculas de hemicelulose pré-sintetizadas são transportadas para a parede celular, podendo ainda sofrer modificações estruturais. Os monolignóis de lignina são sintetizados na superfície do retículo endoplasmático e citoplasma sendo depois transportados para a parede celular onde são ativados pelas lacases e peroxidases, e então polimerizados para formação da molécula de lignina (Kumar et al., 2016).

Figura 3. Localização subcelular das vias de síntese e montagem da parede celular secundária. Legenda = PCS: parede celular secundária, MP: membrana plasmática, CCS: complexo celulose sintase, RE: retículo endoplasmático, GT: glicosiltransferases, LM: lamela média, PCP: parede celular primária (Kumar et al., 2016).

O genoma de *Arabidopsis* apresenta 10 genes que codificam as proteínas CESA, pertencentes à família glicosiltransferase GT2. Como já comentado, três proteínas CESA (CESA4, CES7, CESA8, também descritas como IRX5, IRX3 e IRX1, respectivamente) são os componentes básicos dos complexos celulose sintase. Além delas, uma proteína ligada à membrana com atividade de endo-1,4- β -d-glucanase extracelular chamada KORRIGEN1 (KOR1) interage com CESAs tanto na síntese da parede primária como da secundária (Mansoori et al., 2014). Propôs-se que o KOR1 funcione como uma celulase na liberação das microfibrilas crescentes a partir do complexo de celulose sintase e/ou no aumento da quantidade de microfibrilas de celulose não cristalizadas (Takahashi et al., 2009).

A xilana é outro importante constituinte da PCS sendo esta a hemicelulose mais importante tanto para mono como para dicotiledôneas. As xilanas são sintetizadas por produtos de genes *IRX*, incluindo *IRX7*/ FRAGILE FIBRE8 (FRA8), *IRX8*, *IRX9*, *IRX10*, *IRX14* e o IRREGULAR XYLEM-LIKE genes *IRX9-L*, *IRX10-L* e *IRX14-L*. A biossíntese da xilana requer um grande número de enzimas, incluindo xilosiltransferases, glucuroniltransferases, metiltransferases, acetiltransferases e arabinosiltransferases, que são responsáveis pelo alongamento e substituições na estrutura da xilana. A mutação do gene *IRX9* ou *IRX14* foi suficiente para causar uma redução drástica na atividade da xilosiltransferase, no

comprimento da cadeia de xilanas e no teor das xilanas (Brown et al., 2009; Zhong et al., 2015).

Finalmente, uma parte muito importante da PCS é a lignina, um polímero complexo de álcoois aromáticos. O processo de formação de lignina é iniciado com biossíntese de suas subunidades, os monolignóis, a partir da fenilalanina para produzir álcool cumarílico, álcool coniferílico e álcool sinapílico. A via começa com o passo fundamental, a desaminação da fenilalanina pela fenilalanina amônia-liase (PAL) a ácido trans-cinâmico. Ação sequencial de cinamato-4hidroxilase (C4H) e 4-cumarato:CoA ligase (4CL) produz p-cumaril-CoA. Esta parte da via envolvendo PAL, C4H e 4CL é comum a todos os fenilpropanoides e é por vezes referida como a via geral dos fenilpropanoides (C₆C₃). A ação da hidroxycinnamoil-CoA redutase (CCR) e cinamil álcool desidrogenase (CAD) na molécula de p-cumaril-CoA leva a álcool cumarílico e unidades H de lignina. Quatro enzimas, hidroxycinnamoil CoA:chiquimato/quinato hidroxycinnamoil transferase (HCT), p-cumarato 3-hidroxilase (C3H), CCR e CAD, convertem o p-cumaril-CoA em álcool coniferílico e unidades G de lignina. Um ramo lateral desta via conduz ao álcool sinapílico e unidades S pela ação da ferulato-5hidroxilase (F5H), ácido cafeico O-metiltransferase (COMT) e CAD (Kumar et al., 2016).

Em *Arabidopsis*, foram identificados 35 genes biossintéticos de monolignóis e a expressão de 16 deles, *PAL1*, *PAL2*, *PAL4*, *C4H*, *4CL1*, *4CL2*, *HCT*, *C3H1*, *CSE*, *CCoAOMT1*, *CCoAOMT7*, *CCR1*, *F5H*, *COMT*, *CAD3* e *CAD6*, está intimamente associado ao desenvolvimento de xilema lignificado e das fibras interfasciculares em caules, indicando que eles são os genes principais da biossíntese dos monômeros e responsáveis pela síntese de lignina durante a formação de paredes secundárias (Vanholme et al., 2010).

Os monolignóis são subsequentemente polimerizados no espaço apoplástico pelas peroxidases (PXR) e lacases (LACs) que possibilitam a formação da lignina pela polimerização oxidativa dos monolignóis (Yang and Wang, 2016). Nas células do protoxilema, as lacases se localizam apenas em regiões da parede celular submetidas a deposição da parede secundária, tornando provável que a reticulação pelas lacases determine os locais de deposição da PCS.

A deposição da lignina é temporalmente e espacialmente controlada, a lignificação ocorre em três estágios cada um caracterizado pela deposição de um carboidrato específico em locais específicos (Figura 4). O primeiro estágio da lignificação acompanha a expansão radial das paredes celulares, a deposição de pectina e a formação da primeira camada

celulósica na PCS (camada S1). A segunda fase é um estágio de lignificação mais lenta, e ocorre conjuntamente com a formação da maior camada da parede secundária, a camada S2. A maior parte da lignificação ocorre da terceira fase depois da formação da camada S3 (Albersheim et al., 2010).

Figura 4. Deposição ao longo do tempo dos polissacarídeos da parede primária e lignina durante a lignificação da parede celular secundária. A parte superior apresenta o tempo de formação das maiores camadas a parede secundária e componentes dos carboidratos. A parte inferior mostra a relação do tempo e localização da lignificação (Albersheim et al., 2010).

O padrão de deposição da parede celular secundária é determinado pelos microtúbulos corticais (MTs) que direcionam o movimento dos CCSs na membrana plasmática, conforme esquematizado na Figura 5. Os feixes de microtúbulos corticais não apenas controlam a trajetória do movimento das CESA na membrana plasmática, mas também orientam o transporte direcionado de vesículas transportando os complexos CESA, hemiceluloses e enzimas de polimerização dos monolignóis para domínios específicos da membrana plasmática e, deste modo, controlam a deposição padronizada das paredes secundárias. Foi demonstrado que os microtúbulos corticais dos elementos traqueais formam feixes sob domínios específicos da membrana plasmática onde ocorre espessamento da parede secundária (Oda et al., 2005). Recentemente, várias proteínas associadas a microtúbulos (MAPs) foram descobertas e atuam regulando a dinâmica global

dos microtúbulos corticais na célula e nos padrões de deposição da parede secundária (Oda and Fukuda, 2013).

Figura 5. Modelo para deposição da parede celular secundária. O citoesqueleto de actina dirige o fornecimento de vesículas contendo hemicelulose e o complexo de celulose sintase à membrana plasmática. Os feixes de microtúbulos (MT) marcam os locais de inserção da vesícula e mantêm as partições da membrana plasmática, assegurando que o complexo da celulose sintase se mova paralelamente aos MTs, ligando os MTs as microfibrilas de celulose. As células adjacentes fornecem os componentes que contribuem para a síntese de lignina e o espessamento da parede celular secundária (Turner et al., 2007).

Uma vez que as paredes secundárias são depositadas após a cessação do alongamento celular nas células formadoras de parede secundárias, os genes que codificam o conjunto de enzimas catalíticas necessárias para produzir polímeros de parede secundária têm de ser ativados de forma coordenada temporalmente e espacialmente. Um grande número de fatores de transcrição (FTs) tem sido identificados como controladores da expressão dos genes relacionados à biossíntese da PCS, a maioria pertencente a duas famílias principais: NAC e MYB (Figura 6).

Os FTs envolvidos na biossíntese da PCS podem ser amplamente classificados em três grupos. Os FTs do Nível 1 ligam-se diretamente aos genes estruturais responsáveis pela biossíntese dos componentes da parede celular. FTs Nível 2 podem regular FTs Nível 1, além dos genes estruturais, enquanto os genes Nível 3 podem regular os genes Nível 2 e os genes estruturais (Kumar et al., 2016; Zhong et al., 2015). Além disso, foi identificado um fator de transcrição denominado E2Fc - membro da família E2F - e que atua como um regulador chave acima dos reguladores Nível 3, bem como dos genes estruturais (Taylor-Teeple et al., 2015).

Conforme a Figura 6 existem 5 reguladores principais no Nível 3, o nível mais alto da rede hierárquica, denominados NST1, SND1/NST3, NST2, VND6 e VND7. Quando super

expressos esses genes são responsáveis pela deposição da PCS em vários tipos celulares. O NST1 e o SND1 são os reguladores principais da formação das fibras de xilema, e o VND6 e o VND7 agem nos vasos xilemáticos (metaxilema e protoxilema, respectivamente). VND6 e VND7 também ativam a expressão de uma vasta gama de genes envolvidos na morte celular programada, tal como a cisteína peptidase específica do xilema (XCP1) (Yang and Wang, 2016).

Todos esses genes ativam diretamente vários fatores de transcrição abaixo no Nível 2, incluindo a família MYB (MYB46, MYB83 e MYB55) e a família NAC (SND3 e XND1). MYB46 e MYB83, por sua vez, ativam uma rede de fatores de transcrição Nível 1 promovendo a biossíntese de lignina, celulose e/ou hemicelulose (Ruzicka et al., 2015). O BES1 mostrou-se relacionado ao controle transcricional dos genes CESA, enquanto os FTs MYB20, MYB69, MYB79, MYB85, MYB58 e MYB63 e BP controlam genes da biossíntese de lignina. Alguns destes fatores de transcrição, tais como MYB4, MYB7, MYB32 e KNAT7, atuam como reguladores negativos da formação de PCS e reprimem a expressão de SND1 por um circuito de retroalimentação negativa ou através da atividade de genes biossintéticos da PCS (Figura 6) (Liu et al., 2015).

Um estudo recente de Taylor-Teeple et al. (2015) lançou uma nova luz sobre a rede reguladora de genes que controlam a síntese de celulose, hemicelulose e lignina no xilema da raiz de *Arabidopsis*. Um total de 623 interações entre 45 promotores e 208 FTs foram identificados e demonstraram que a regulação transcricional da biossíntese da PCS é altamente combinatória, com múltiplos FTs se ligando a múltiplos promotores. Um dos achados mais significativos foi a identificação de um fator de transcrição mestre E2Fc, que se liga a aproximadamente 20 promotores, incluindo os dos genes que codificam VND6, VND7 e MYB46, bem como genes associados à biossíntese de celulose, hemicelulose e lignina (Taylor-Teeple et al., 2015).

Figura 6. Visão geral da rede regulatória transcricional controlando a formação de parede celular secundária em *Arabidopsis* (Ruzicka et al., 2015).

Com relação a deposição de lignina, as enzimas responsáveis pela oxidação dos monolignóis LACCASE4 (LAC4) e LAC17 mostraram-se importantes para a lignificação e foi recentemente demonstrado a localização específica dos LAC4 e LAC17 nos espessamentos da PCS em *Arabidopsis* (Zhao et al., 2013). Contrariamente à biossíntese, o transporte e polimerização dos monolignóis ainda são pouco conhecidos, como os fatores que determinam a localização específica de LAC4/LAC17 e, portanto, o padrão de lignificação específico da PCS em vasos xilemáticos, devem ser identificados (Ruzicka et al., 2015).

Foi demonstrado que o etileno está envolvido na formação de madeira em tensão. Diversos Fatores de Resposta ao Etileno (ERFs), que medeiam a sinalização de etileno, são induzidos em resposta à formação de madeira de tensão no álamo e, quando super expressos, alteram a formação de madeira em árvores transgênicas (Vahala et al., 2013).

2.5. Xilogênese e a diferenciação *in vitro* de elementos traqueais

O processo de xilogênese envolve a diferenciação celular e o processo de lignificação envolvidos na formação da madeira. A madeira é composta basicamente por dois tipos celulares com paredes secundárias: as fibras que mecanicamente suportam a estrutura das plantas e os elementos traqueais compostos pelos elementos de vasos e traqueídes, os quais transportam água e solutos e são responsáveis pela função primordial do xilema (Fukuda, 1996).

O estudo da formação dos elementos traqueais está intimamente relacionado ao entendimento do processo de xilogênese que é fundamental para melhorar a quantidade e qualidade da madeira produzida. Além disso, o uso da indução *in vitro* dos elementos traqueais possibilita um acesso privilegiado ao processo de formação da parede celular secundária que não seria possível nos tecidos vegetais complexos. Resumindo, a formação *in vitro* dos elementos traqueais proporciona um excelente sistema para o estudo da diferenciação das células xilemáticas (Demura and Fukuda, 2007; Devillard and Walter, 2014; Taiz and Zeiger, 2010).

Os elementos traqueais maduros são células mortas envoltas por uma rígida parede secundária, que pode ser depositada em vários padrões, como espiralado, reticulado, anular ou escalariforme (Figura 7). Essas células são formadas a partir de células do pró-câmbio ou células-filhas formadas das células procambiais. Essa deposição da parede secundária é determinada pela orientação dos microtúbulos que dirigem a movimentação dos complexos celulose-sintase na membrana plasmática. Recentemente algumas proteínas associadas aos microtúbulos (MAPs) foram identificadas e relacionadas à deposição e padronização da parede secundária encontrada nas fibras e nos elementos traqueais (Oda and Fukuda, 2012).

Figura 7. Modelo esquemático das células do xilema e floema, com destaque para as traqueídes e elementos de vaso. É possível visualizar os diferentes padrões de deposição das paredes secundárias nos traqueídes (Taiz and Zeiger, 2010)).

O processo de xilogênese envolve quatro passos para formação dos elementos traqueais: divisão celular, expansão celular (elongação e alargamento radial), espessamento da parede celular (envolvendo celulose, hemicelulose, proteínas da parede celular e

biossíntese e deposição de lignina), e a morte celular programada (McCann, 1997). Essas etapas envolvidas na diferenciação dos elementos traqueais foram ilustradas detalhadamente na Figura 8. A morte celular programada, o processo final da diferenciação dos elementos traqueais, ocorre independentemente do status do processo de lignificação em cultura de células de *Zinnia elegans* mostrando que a lignificação não é nem um indutor nem um pré-requisito para a morte celular programada nos elementos traqueais (Pesquet et al., 2013).

A regulação do processo de xilogênese está correlacionada à expressão de classes específicas de genes e aos processos relacionados ao crescimento e diferenciação celular. Recentes estudos moleculares desse processo tem revelado um alto controle gênico envolvido no processo de formação da madeira, principalmente ao nível transcricional (Ursache et al., 2013). A deposição da parede secundária merece destaque, pois caracteriza os elementos traqueais e está intimamente relacionada às vias de lignificação (Turner et al., 2007).

Figura 8. Principais etapas para formação *in vitro* um elemento traqueal em *Zinnia elegans* (Devillard and Walter, 2014).

Estudos pioneiros relacionados à indução *in vitro* de elementos traqueais foram realizados com *Zinnia elegans* utilizando-se um sistema de cultura de células do mesófilo foliar. Fukuda and Komamine (1980) conseguiram induzir um processo de transdiferenciação das células do mesófilo em elementos traqueais utilizando auxina e citocinina. Estudos

realizados com *Arabidopsis thaliana* possibilitaram a obtenção de elementos traqueais na presença de 1 μ M de brassinolide, e auxiliaram na compreensão dos mecanismos relacionados à regulação e controle da deposição e espessamento da parede secundária (Oda et al., 2005).

Trabalhos envolvendo a indução de elementos traqueais também foram realizados com outras espécies como *Pinus radiata* (Möller et al., 2006a), *Torreya nucifera* e *Cryptomeria japonica* (Yamagishi et al., 2012), mas nesses casos foram utilizadas suspensões celulares obtidas a partir de calos, com adição de 5 g/L carvão vegetal ativado ao meio de indução, na ausência de reguladores vegetais. Recentemente, foi realizada a indução da formação *in vitro* de elementos traqueais em suspensões celulares da conífera *Cryptomeria japonica* suplementando o meio de cultura com 1,1 mg/L de 2,4-D (ácido diclorofenoxiacético) (Yamagishi et al., 2015).

Estudos relacionados à indução *in vitro* de elementos traqueais em plantas modelo, demonstraram a influência de diferentes substâncias no aparecimento dessas células. Kákošová et al., (2013) constataram um efeito na viabilidade celular e na formação de elementos traqueais, em cultura de células de *Zinnia*, com a utilização do oligossacarídeo galactoglucomanana, sendo o efeito positivo máximo atingido pela adição simultânea do oligossacarídeo e fitohormônios de crescimento (auxina e citocinina) ao meio de cultura. Confirmando a importância do uso de brassinolide na indução de elementos traqueais, Yamagishi et al., (2013) demonstraram que mais de 70% dos elementos traqueais nos calos do híbrido *Populus sieboldii* x *P. grandidentata* foram induzidos em meio de cultura acrescido exclusivamente de brassinolide (1,0 μ M). A adição de enzimas ACC sintase (ácido 1-carboxílico-1-aminociclopropano), que participa da rota de biossíntese de etileno, estimulou a diferenciação dos elementos traqueais em suspensões celulares de *Zinnia* (Pesquet and Tuominen, 2011).

Outro fator importante na indução dos elementos traqueais foi a incidência de luz, que foi testada em *Pinus radiata*, em que os autores observaram que a porcentagem de diferenciação nos calos mantidos na presença de luz (fotoperíodo de 16h ou 24h) foi 25% maior em relação aos mantidos no escuro (Möller et al., 2006a).

Nos trabalhos realizados envolvendo os elementos traqueais, foi possível verificar uma maior frequência dessas células em agregados, demonstrando uma possível comunicação intercelular importante no processo de diferenciação das mesmas. Os

elementos traqueais foram obtidos por meio da transdiferenciação de células do mesófilo foliar ou da diferenciação de células meristemáticas em suspensões obtidas a partir dos calos. Em alguns casos, os elementos traqueais se desenvolveram e foram visualizados nos próprios calos mantidos nos meios de cultura sólidos (Möller et al., 2006b; Yamagishi et al., 2012, 2015).

O tempo para a indução e diferenciação das células varia muito entre as espécies e o tipo de sistema utilizado, sendo relatado períodos de 4 até 49 dias (Aloni, 1980; Fukuda and Komamine, 1980b; Pillai et al., 2011). Outro ponto interessante foi a sincronização observada na indução dos elementos traqueais, que se mostrou totalmente sincronizada (Fukuda and Komamine, 1980b), não-sincronizada (Möller et al., 2003) ou parcialmente sincronizada (Pesquet and Tuominen, 2011). A porcentagem de transdiferenciação em elementos traqueais foi bastante variável, com índices de 20% (Lacayo et al., 2010) a 60% (Fukuda and Komamine, 1980b) em suspensões de *Zinnia elegans*, e índices de 30% em *Arabidopsis thaliana* (Oda et al., 2005). Yamagishi et al. (2012) perceberam a ocorrência de elementos traqueais nos calos de *Torreya nucifera* e *Cryptomeria japonica* e um aumento na porcentagem dessas células quando os mesmos foram transferidos para meio líquido acrescido de carvão ativado, na ausência de fitohormônios.

Os mecanismos básicos da diferenciação dos elementos traqueais foram analisados por abordagens de biologia molecular, sendo possível isolar genes-chaves em *Zinnia* relacionados à expressão de enzimas da biossíntese de brassinosteróides (BRs) (ZeSTE1, ZeDIM, ZeDWF4, ZeCPD1 e ZeCPD2), sendo sugerido que a biossíntese dos BRs durante a diferenciação dos elementos traqueais pode ser controlada pela regulação coordenada da biossíntese de esteróis amplos e pela regulação específica da biossíntese de BRs, que ocorre em parte por níveis elevados de transcrição de genes que codificam enzimas, especificamente ZeDWF4 e ZeCPD1 (Yamamoto et al., 2007). Os genes *Tracheary Element Differentiation-Related6* (*TED6*) e *TED7* foram selecionados baseados na técnica de RNA interferente (RNAi). A diminuição da expressão dos genes *TED6* e *TED7* por RNAi atrasou a diferenciação dos elementos traqueais e a superexpressão desses genes aumentou a diferenciação em suspensões celulares de *Z. elegans*. Experimentos com *Zinnia* e *Arabidopsis* ainda demonstraram um papel ativo das proteínas TED6 e TED7 na formação da PCS e a possível participação dessas proteínas nos complexos celulose sintase CesaA (Endo et al., 2009).

Os padrões nos quais os espessamentos dos elementos traqueais são organizados variam de acordo com os feixes de microtúbulos que guiam a deposição da parede secundária. Derbyshire *et al.* (2015) nocauteou genes por RNAi mostrando que a formação de parede secundária depende da presença coordenada dos microtúbulos interagindo com proteínas que atuam na espessura da parede celular, na homogeneidade da parede celular e no padrão e localização cortical da parede.

Quanto ao processo de lignificação nos elementos traqueais, estudos foram realizados com a supressão de genes associados a rota de biossíntese dos monolignóis em *Pinus*. A supressão severa da 4CL (4-cumarato CoA ligase) em *Pinus radiata* afetou substancialmente o fenótipo da planta e resultou em plantas anãs com aparência de “bonsai”. Ocorreu a formação de traqueídes fracamente lignificadas e reduções do conteúdo de lignina de 36% a 50% nas plantas transgênicas mais severamente afetadas (Wagner *et al.*, 2009). O silenciamento mediado por RNAi do gene putativo de HCT (hidroxicinamoil shikimato transferase) teve um forte impacto no teor de lignina e na composição dos monolignóis. Sendo constatada uma redução de até 42% no teor de lignina nos elementos traqueais de *Pinus radiata* (Wagner *et al.*, 2007). A transformação por biobalística de elementos traqueais de *P. radiata* acarretou na supressão da expressão do PrCCoAOMT que codifica a enzima da biossíntese de lignina CCoAOMT (caffeoil-CoA 3-O-metiltransferase) afetando o teor e na composição da lignina, resultando num polímero de lignina contendo unidades p-hidroxifenil (H), catecol (C) e guaiacil (G). Os ensaios de ligninas solúveis em AcBr (brometo de acetila) revelaram reduções no teor de lignina de até 20% em linhagens transgênicas deficientes em PrCCoAOMT (Wagner *et al.*, 2011). A supressão do gene relacionado à rota biossintética de lignina que codifica a enzima cinamoil-CoA redutase (CCR) em elementos traqueais de *P. radiata* resultou em mais de 46% de redução no conteúdo de lignina em linhagens CCR-RNAi, além do impacto no perfil metabólico e na matriz da parede celular (Wagner *et al.*, 2013).

Quando usado como modelo, o cultivo *in vitro* dos elementos traqueais pode contribuir para a melhor compreensão da formação da madeira e tem vantagens sob o sistema *in planta*, pois a diferenciação sincrônica de um conjunto celular homogêneo possibilita o estudo da bioquímica e da biologia molecular da formação dessas células, livre da complexidade dos tecidos vegetais especializados (McCann, 1997).

Desta maneira, análises utilizando o sistema com elementos traqueais pode auxiliar na compreensão do complexo mecanismo envolvido na formação da parede celular secundária e da biossíntese da lignina durante a diferenciação do xilema, podendo esse conhecimento ser utilizado para modificar ou diminuir o conteúdo de lignina, reduzindo os custos no processamento da biomassa pelas indústrias de papel e celulose e etanol lignocelulósico (Devillard and Walter, 2014).

2.6. RNA-Seq e Transcriptômica

O RNA-seq descreve uma coleção de métodos experimentais e computacionais para determinar a identidade e abundância de sequências de RNA em amostras biológicas. Os métodos experimentais envolvem o isolamento do RNA a partir de amostras de células e/ou tecidos, a preparação de bibliotecas, o sequenciamento químico efetivo da biblioteca e a subsequente análise dos dados por bioinformática. Na prática, milhões de sequências curtas, chamadas de *reads* ou leituras, são sequenciadas a partir de posições aleatórias dos RNAs analisados. Estas leituras são mapeadas computacionalmente num genoma de referência para revelar um "mapa transcricional", onde o número de leituras alinhadas a cada gene dá uma medida do seu nível de expressão. No entanto, o esquema de análise descrito acima, que pode parecer simples à primeira vista, é de fato muito mais complexo e consiste em vários passos de processamento (Finotello and Di Camillo, 2015; Korpelainen et al., 2015) (Figuras 10 e 11).

Uma distinção crítica do RNA-seq de métodos anteriores, tais como os *microarrays*, é o alto rendimento das plataformas atuais de RNA-seq, a sensibilidade proporcionada por essas novas tecnologias, e a possibilidade de descobrir novos transcritos, modelos de gene, e pequenos RNA não-codificantes. Essas poderosas características do RNA-seq impulsionaram um progresso sem precedentes da pesquisa transcriptômica, produzindo uma impressionante quantidade de dados em todo o mundo. Para suportar este crescimento exponencial e lidar com as diferentes etapas da análise de dados, várias ferramentas computacionais foram desenvolvidas e são atualizadas em ritmo acelerado (Conesa et al., 2016).

Transcriptoma ou Transcritoma refere-se ao conjunto completo de transcritos (RNAs mensageiros, RNAs ribossômicos, RNAs transportadores e os microRNAs) de um dado organismo, órgão, tecido ou linhagem celular. Portanto, as abundâncias relativas dos transcritos refletirão o nível de expressão dos genes correspondentes, para um estágio de desenvolvimento específico ou condição fisiológica. Embora os RNAs não sejam os produtos finais do processo de transcrição-tradução, o estudo da expressão gênica e expressão gênica diferencial podem revelar aspectos importantes sobre os estádios celulares sob investigação (Rapaport et al., 2013).

O advento do sequenciamento de nova geração (NGS) revolucionou a transcriptômica e rapidamente estabeleceu o RNA-seq como a metodologia preferida para o estudo da expressão gênica. Os RNAs de uma amostra de interesse são inicialmente fragmentados e transcritos inversamente em DNAs complementares (cDNAs) e esses cDNAs obtidos que são então amplificados e submetidos ao NGS (Wang et al., 2009).

Atualmente a plataforma Illumina é a mais comumente utilizada, realizando o sequenciamento por síntese. Após as bibliotecas serem construídas (Figura 9a), o ds cDNA é passado através de uma célula de fluxo ou *flow cell* que irá hibridar as moléculas individuais com base na sua complementaridade com sequências adaptadoras (Figura 9b). As sequências hibridizadas, mantidas em ambas as extremidades do adaptador na célula de fluxo, serão amplificadas como uma ponte. Essas sequências recém-geradas irão hibridar-se com a célula de fluxo próximo e após muitos ciclos, uma região da célula conterá muitas cópias do ds cDNA original, esse processo é conhecido como *cluster generation* (Figura 9c). Depois que os *clusters* são gerados e uma cadeia é removida do cDNA, os reagentes são passados através da célula de fluxo para executar o sequenciamento por síntese. Nesse tipo de sequenciamento uma reação a cada ciclo de síntese adiciona de um único nucleotídeo, que pode ser A, C, G ou T, determinando um sinal fluorescente, que é fotografado, tal como sua localização e, assim, o nucleotídeo adicionado pode ser determinado, armazenado e analisado. A sequência de adições numa localização específica na célula de fluxo fornecerá a sequência de nucleotídeos precisa para uma parte original do cDNA (Figura 9d e 9e). Assim, as milhões de leituras curtas geradas podem ser, posteriormente, mapeadas num genoma de referência e o número de leituras alinhadas a cada gene, denominado de "contagens", dará uma medida digital dos níveis de expressão de genes na amostra sob investigação (Ozsolak and Milos, 2011; Verk et al., 2013; Wang et al., 2009).

Figura 9. Visão geral da tecnologia NGS usando Illumina. a) A biblioteca é composta por fragmentos de DNA de cadeia dupla que podem ser reconhecidos pelo sequenciador de nova geração. O DNA a ser sequenciado é flanqueado por seqüências adaptadoras. b) A extremidade 5' do adaptador em forma de Y contém um grupo fosfato e a extremidade 3' contém um dT. c) Geração de *clusters* na superfície da célula de fluxo ou *flow cell*. d) Sequenciamento por síntese com a adição dos nucleotídeos, fluorescência e detecção da coloração. e) Quando a reação de sequenciamento da primeira cadeia termina, a cadeia sintetizada é removida e o processo é repetido para a cadeia oposta (adaptada de Shin et al. (2014)).

A Figura 10 fornece uma visão geral do *pipeline* de análise de RNA-seq para detectar expressões diferenciais. Existem programas separados para cada etapa, embora algumas ferramentas possam combinar algumas delas. A primeira etapa é o controle de qualidade das leituras brutas geradas que envolve a análise da qualidade das seqüências, do conteúdo de GC e da presença de adaptadores (<http://support.illumina.com/downloads/illumina-customer-sequence-letter.html>), auxiliando na detecção de possíveis erros no sequenciamento, artefatos de PCR ou contaminação. O resultado desta etapa é um sumário

contendo o número de leituras e informações de qualidade que orientam as decisões de pré-processamento na próxima etapa (Conesa et al., 2016; Oshlack et al., 2010). Uma ferramenta popular para executar essas análises nas leituras geradas pelo Illumina é o *FastQC* (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).

Figura 10. O fluxo de trabalho de uma análise de expressão diferencial consiste em várias etapas inter-relacionadas. As numerações de 1. a 6. correspondem as etapas realizadas. Os formatos dos arquivos de saída típicos de cada etapa estão indicados entre parênteses. Figura adaptada de Korpelainen et al. (2015).

Idealmente, a maioria das leituras deve ter uma qualidade média de base de 25 ou mais, entretanto, a qualidade da leitura pode diminuir para a extremidade 3' e, se ficar demasiadamente baixa, as bases devem ser removidas para melhorar a capacidade de mapeamento (Korpelainen et al., 2015). Assim, a etapa 2. (Figura 10) de pré-processamento tem como principal objetivo remover as bases e leituras de baixa qualidade e artefatos, como os adaptadores. Programas como *Trimmomatic* pode ser utilizado para essa finalidade. Se as bases de baixa qualidade forem detectadas nas extremidades das leituras, a maneira mais simples de removê-las é cortando as leituras em um dado comprimento ou cortar um número específico de bases de qualquer das extremidades. Como algumas leituras podem se

tornar muito curtas, parâmetros utilizados no *Trimmomatic* podem filtrar leituras que são mais curtas do que um comprimento mínimo definido pelo usuário. Um fato importante desse programa é que têm suporte para leituras *paired-end*, podendo manter os arquivos de leitura sincronizados mesmo se uma leitura perder seu par no processo de corte (Bolger et al., 2014).

A terceira etapa (Figura 10) abrange o alinhamento das leituras com o genoma de referência objetivando encontrar um ponto de origem para cada leitura. Se um genoma de referência não estiver disponível, as leituras podem ser mapeadas contra um transcriptoma (que pode ser criado *de novo* a partir das leituras, se necessário). Quando uma leitura é mapeada com a referência, uma sequência de alinhamento é criada (Trapnell et al., 2009).

Desta maneira, o mapeamento permite a identificação de novos genes ou transcritos requerendo o uso de um mapeador chamado *gapped* ou *spliced*. O desafio desta etapa é identificar corretamente as junções em encadeamento (*splicing*) de éxons, especialmente quando há erros no sequenciamento ou diferenças com a referência. Um dos mapeadores mais populares, o TopHat, realiza o mapeamento no genoma referência em 2 fases, primeiro todas as leituras contra o genoma de referência e depois realiza o mapeamento das *reads* inicialmente não mapeadas (IUM) nessas regiões de junção selecionadas. A saída desta etapa é um arquivo de alinhamento, que lista as leituras mapeadas e suas posições de mapeamento na referência (Conesa et al., 2016; Trapnell et al., 2012).

Um parâmetro importante de qualidade do mapeamento é a porcentagem de leituras mapeadas, que é um indicador global da precisão do sequenciamento e da presença de DNA contaminante. Por exemplo, é esperado que 70 a 90% das leituras regulares de RNA-seq se mapeiem no genoma humano (dependendo do mapeador de leitura usado), com uma fração significativa de leituras mapeando para um número limitado de regiões idênticas (*multi-mapping reads*) (Dobin et al., 2013).

As próximas etapas do RNA-seq estarão envolvidas em duas categorias: aquela relacionada à montagem dos transcritos e anotação do genoma (etapa 4.) E aquela envolvida com a quantificação de transcritos e genes (etapas 5 e 6.) (Figura 10). A ferramenta *Cufflinks* realiza a montagem dos transcritos estimando suas abundâncias e testando a expressão diferencial em experimentos de RNA-seq. Dois programas populares para análise de expressão diferencial são o *Cuffdiff* (que faz parte do pacote *Cufflinks*) e o *DESeq2* (Figura 11). A expressão diferencial realizada pelo *Cuffdiff* utiliza as leituras alinhadas

de duas ou mais condições e relata quais genes e transcritos são diferencialmente expressos utilizando uma análise estatística rigorosa e computando os níveis de expressão das isoformas (Trapnell et al., 2012). Em contrapartida, o *DESeq2* normalmente considera a contagem de transcritos realizada por programas de contagem, como o *featureCounts*, não considerando as isoformas dos genes para os cálculos de expressão e utiliza a distribuição binomial negativa como a base de sua modelagem (Love et al., 2014) (Figura 11).

Figura 11. Esquema detalhado com as etapas e programas utilizados, com os arquivos de entrada (*input*), para a realização de uma análise de expressão gênica diferencial (adaptado de Verk et al. (2013)).

Uma vantagem interessante do *Cuffdiff* é que sua saída pode ser alimentada em um pacote de visualização e análise muito útil em R, o *cummeRbund* (Trapnell et al., 2013). Mas o *Cuffdiff* não suporta desenhos experimentais muito complexos, somente comparando dois grupos, já o *DESeq2* suporta fatoriais (Soneson and Delorenzi, 2013).

As estimativas realizadas pelo *Cufflinks* são baseadas em quantas leituras suportam cada transcrito, levando em consideração os vieses como a distribuição da leitura de maneira não-uniforme ao longo do comprimento do gene e dos protocolos de preparação

das bibliotecas. As estimativas de abundância podem ser reportadas na contagem das leituras brutas ou em unidades normalizadas como RPKM (*reads per kilobase of exon model per million mapped read*) ou FPKM (*fragments per kilobase of exon per million fragments mapped*). A normalização é necessária por causa das possíveis diferenças no número de leituras e na composição do transcriptoma, e muitas ferramentas estatísticas possuem métodos de normalização embutidos (Korpelainen et al., 2015; Oshlack et al., 2010).

Em geral, uma anotação é qualquer comentário, explicação ou alguma marcação (coletivamente conhecida como metadados) anexada aos dados. No entanto, na bioinformática, o significado da anotação é mais complexo. As anotações típicas que estão associadas a um gene são a localização genômica, a estrutura exônica e intrônica, algumas noções funcionais na forma de vocabulários controlados, como o *Gene Ontology* (GO), ou as vias metabólicas em que as proteínas traduzidas atuam (por exemplo, KEGG, *Encyclopedia of Genes and Genomes de Kyoto*) (Korpelainen et al., 2015) (Figura 11).

A anotação mais comumente usada para agrupar genes é o *Gene Ontology* (Ashburner et al., 2000) que é um vocabulário controlado. É hierárquico, portanto, pode-se anotar um gene em diferentes níveis de detalhe, e um gene pode ter várias anotações. A análise de enriquecimento de conjuntos de genes fornece uma lista de funções moleculares, processos biológicos e localizações celulares mais representativas. O uso dessa informação funcional dos genes e dos transcritos, tais como as categorias de ontologias genéticas ou termos GO, permite análises mais refinadas das vias metabólicas *up-* ou *down-regulated* (Khatri and Drăghici, 2005) (Figura 11).

3. MATERIAIS E MÉTODOS

3.1. Cultivo *in vitro* e coleta das amostras

Sementes de *Eucalyptus grandis* (doadas pela empresa Suzano Papel e Celulose, identificação: 12CI/m33D1, monoprogênie) foram germinadas *in vitro* em frascos de vidro contendo o meio de cultura SP desenvolvido por (Barrueto Cid, 2005) que está detalhado no Anexo A. Os frascos foram mantidos à 26 °C na presença de luz (fotoperíodo de 16h). Após 45 dias, foi possível observar o desenvolvimento de plântulas com 5 cm de comprimento, nessa fase as folhas foram removidas e segmentos de 1 cm² foram utilizados como explantes para a indução *in vitro* de calos. A planta que apresentou maior porcentagem na indução de calos nos explantes foi clonada para serem utilizada nas etapas seguintes.

Para a indução dos calos *in vitro*, explantes provenientes da mesma planta clonada (planta 2) foram introduzidos no meio de cultura SP contendo 6 µM de Picloram (4-amino-3,5,6-tricloropicolínico) à 26 °C, em ausência de luz. Após 30 dias de incubação, os calos foram transferidos para meio de cultura novo suplementado com 3 µM de Picloram e 1 µM BAP (6-Benzylaminopurina) durante 60 dias, à 26 °C e em ausência de luz para estimular a multiplicação celular e o crescimento dos calos.

Após esse período, os calos friáveis foram utilizados nas suspensões celulares (Experimento 1). Para isso, cerca de 3 calos (0,8g a 1,0g) foram transferidos para frasco Erlenmeyer de 125 mL contendo 20 mL de meio líquido. Foram testados 4 meios líquidos distintos para a multiplicação celular e indução dos elementos traqueais, baseados em estudos realizados para outras espécies vegetais. Foram testados os meios líquidos utilizados por Fukuda and Komamine, (1980b), contendo 0,1 mg/L de NAA (ácido naftalenoacético) e 1mg/L de BAP, por Oda *et al.*, (2005) contendo 1µM de epibrassinolide, por Yamagishi *et al.*, (2012) contendo 5,0g /L de carvão vegetal ativado, e um controle que corresponde a um meio líquido SP desprovido de reguladores vegetais (Yamagishi et al., 2015).

Os calos foram mantidos nas respectivas suspensões celulares por um período de 30 dias, sendo realizadas avaliações periódicas, a cada 5 dias, com a contagem do número de elementos traqueais formados. As suspensões celulares foram mantidas a 26°C, na ausência de luz, sob agitação branda (80 rpm).

No experimento 2 um total de 10 calos de *E.grandis* foi transferido para 100 ml de meio líquido em Erlenmeyers de 250 ml contendo a mesma composição do meio de cultura sólido SP, mas sem a suplementação de reguladores de crescimento e solidificadores totalizando 4 frascos, sendo cada um considerado uma repetição biológica. As suspensões celulares foram mantidas à 26 °C, em ausência de luz e sob agitação de 120 rpm *overnight*.

Com o auxílio de peneiras com poros de tamanho de 50 µm, 100 µm e 500 µm, células em diferentes fases de desenvolvimento - células meristemáticas (M), células alongadas (A) e elementos traqueais (T) - foram retiradas do meio líquido. Cada amostra foi imediatamente congelada em nitrogênio líquido e armazenada à -80 °C, para posterior extração do RNA total. A metodologia adotada para a indução e isolamento das fases celulares foi esquematicamente demonstrada na Figura 12.

Para a visualização e caracterização das células em suspensão, uma alíquota de 100 µl da suspensão foi colocada sobre lâmina de vidro e coberta com lamínula. A observação das células foi realizada utilizando-se aumento de 100x do microscópio óptico Olympus BX53 (Olympus America, Center Valley, PA, USA) possibilitando a visualização com luz branca e fluorescência. A contagem do número de células presentes nas suspensões celulares foi realizada com a Câmara de Neubauer. Além disso, os calos foram submetidos à coloração com solução aquosa de safranina a 0,01% por 30 minutos.

Figura 12. Metodologia adotada para a indução dos elementos traqueais e isolamento das fases celulares.

3.2. Extração do RNA total, quantificação e qualidade

O RNA total foi extraído e purificado de 100 mg de células congeladas utilizando o *RNeasy Plant Mini Kit* e a *RNase-Free DNase Set* (Quiagen, Hilden, Germany), de acordo com o manual do fabricante. Foram separadas três fases celulares (M, A, T), cada uma com quatro repetições biológicas, totalizando 12 amostras. A quantificação, controle de qualidade e integridade do RNA total foram verificadas utilizando-se o instrumento *Agilent 2100 Bioanalyzer* (*Agilent Technologies*, Santa Clara, CA, USA) (Mueller and Schroeder, 2004) e a integridade das amostras foi observada pelo gel de agarose 1,5%.

3.3. Preparação da biblioteca e sequenciamento do RNA

As bibliotecas foram preparadas utilizando-se o *TruSeq Stranded mRNA Library Prep Kit* (Illumina, San Diego, CA, USA), a clusterização foi feita pelo *TruSeq PE Cluster Kit v3-cBot-HS* (Illumina, San Diego, CA, USA) e o sequenciamento das leituras com 100 pb *paired-end* foi

realizado pelo aparelho Illumina HiSeq 2500 com o *TruSeq SBS Kit v3-HS* (Illumina, San Diego, CA, USA), seguindo as instruções do fabricante.

3.4. Mapeamento das leituras e análise dos dados

As sequências brutas *paired-end* com 100 pb foram inicialmente tratadas para remoção dos adaptadores e bases de baixa qualidade utilizando o software *Trimmomatic* versão 0.35 (Bolger et al., 2014) (<http://www.usadellab.org/cms/?page=trimmomatic>) utilizando os parâmetros - phred33 SLIDINGWINDOW:4:22 MINLEN:60, o seguinte comando desliza uma janela de 4 bases partir da extremidade 5' e corta leituras quando a qualidade média cai abaixo de 22 (SLIDINGWINDOW:4:22). Ele também filtra as leituras que são mais curtas do que 60 bases (MINLEN: 60) após a limpeza.

A análise da qualidade das sequências foi realizada pelo programa *FastQC* (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), antes e após a limpeza. As leituras foram mapeadas contra o genoma do eucalipto (versão 2.0) (Myburg et al., 2014) disponibilizado no portal *Phytozome* (<https://phytozome.jgi.doe.gov/pz/portal.html>) pelo programa *TopHat* 2.1.1 (<http://tophat.cbcb.umd.edu/>) com os parâmetros -r 0 -l 5000 -i 20 --library-type fr-firststrand (Thumma et al., 2012; Trapnell et al., 2009). Os arquivos de saída do *TopHat* foram classificados e indexados utilizando a ferramenta *SAMtools* (Li et al., 2009) (<http://samtools.sourceforge.net/>) para cada uma das amostras. Posteriormente, foi realizada a contagem das leituras mapeadas pela função *featureCounts* do pacote *Rsubread* (Liao et al., 2013) do programa R (<https://www.r-project.org/>) para utilização no pacote *DESeq2*.

3.5. Expressão diferencial de genes, anotação e categorização funcional

Os genes diferencialmente expressos (GDEs) foram identificados pelo programa *Cufflinks* versão 2.2.1 (Trapnell et al., 2012) utilizando a ferramenta *Cuffdiff* (<http://cole-trapnell-lab.github.io/cufflinks/cuffdiff/index.html>) e pelo pacote *DESeq2* do programa R (versão 3.1.0) (<https://bioconductor.org/packages/release/bioc/html/DESeq2.html>) (Love et al., 2014). A quantificação e a normalização dos valores de expressão pelo FPKM (*Fragments Per Kilobase Of Exon Per Million Fragments Mapped*) foram calculadas pelo *Cuffdiff* e os

heatmaps foram construídos no ambiente R com o pacote *cummeRbund*. Para uma análise explanatória dos dados, uma PCA (*Principal Component Analysis*) foi construída pelo pacote *DESeq2* no programa R (<https://www.r-project.org/>).

Da totalidade dos GDEs identificados pelos programas foram considerados para o estudo os genes que se apresentaram diferencialmente expressos em ambos os programas, sendo considerado os parâmetros estatísticos de $p\text{-value} < 0,01$ e $q\text{-value} < 0,05$ para o programa *Cufflinks* e $padj < 0,05$ para o pacote *DESeq2*, e os genes presentes na intersecção do Diagrama de Venn (<http://bioinformatics.psb.ugent.be/webtools/Venn/>).

Para a etapa de anotação e categorização funcional foi utilizada a plataforma *Blast2GO* (versão 4.0), utilizando o banco de dados de proteínas não redundantes do NCBI, e resultando na identificação de enzimas, assinaturas metabólicas e termos GO. Também foi utilizado o arquivo de anotação do genoma de *E. grandis* (*Egrandis_297_v2.0.annotation_info.txt*) fornecido pelo website *Phytozome* (<https://phytozome.jgi.doe.gov/pz/portal.html>). A descrição geral das funções dos genes foi baseada nos três aspectos do *Gene Ontology*: função molecular, componente celular e processo biológico. O enriquecimento das categorias GO e identificação dos termos GO mais representativos foi realizado pelo aplicativo BINGO (versão 3.0.3), presente na plataforma *Cytoscape* 3.4.0 (Cline et al., 2007), utilizando a distribuição hipergeométrica como teste estatístico e um FDR de 0,05. Para a compreensão de funções biológicas complexas e identificação de vias metabólicas, os GDEs foram mapeados contra o banco de dados do *Kyoto Encyclopedia of Genes and Genomes* (KEGG).

4. RESULTADOS

4.1. Obtenção e caracterização dos elementos traqueais de *Eucalyptus grandis*

Calos de coloração esbranquiçada foram obtidos na superfície dos explantes de *E. grandis* após 30 dias de incubação. Esses calos foram removidos dos explantes e subcultivados em meio de cultura fresco a cada 30 dias. Após 60 dias, os calos aumentaram em tamanho, apresentando em média 1,5 cm de diâmetro (Figura 13A).

Foi possível observar ao microscópio ótico, a presença de elementos traqueais nas suspensões celulares após a instalação do experimento 1, em todos os meios de cultura líquidos testados. Aos 5 dias, as suspensões celulares do experimento 1 apresentaram o mesmo número médio de células (800.000 células/ml) com aproximadamente 5% de elementos traqueais. Entretanto após 10 dias de incubação os elementos traqueais não foram mais observados, ocorrendo somente a presença de células meristemáticas, fato que perdurou até o término das avaliações. Não foi obtida uma indução sincronizada dos elementos traqueais nos diferentes meios de cultura testados, e não houve diferença quantitativa quanto ao número de células e padrão celular exibido entre os diferentes meios de cultura líquidos.

No experimento 2, os calos friáveis foram transferidos para meio líquido e após 24h sob agitação estavam totalmente dissociados formando uma suspensão celular turva e sem contaminação microbiana, apresentando em média 1.000.000 de células/ml de suspensão. Foi possível constatar pelo microscópio ótico a presença de células em diferentes estádios de desenvolvimento, com diferentes tamanhos e formatos (Figura 13D). As células meristemáticas (M) exibiram um formato cilíndrico com diâmetros de aproximadamente 50 μm , as células alongadas (A) em formato longilíneo possuíam um comprimento ao redor de 100 μm e as traqueídes (T) apresentaram tamanhos muito maiores a partir de 400 μm atingindo até 800 μm de comprimento e perfizeram de 5 a 10% das células presentes (Figura 13C, 13E e 13F). Foi possível separar com sucesso os três tipos celulares (M, A e T) com base nos tamanhos e formatos.

Confirmou-se a presença das traqueídes nos calos por microscopia ótica e estes encontravam-se em conjuntos celulares ou agregados (Figura 13B). Os elementos traqueais observados nos calos e nas suspensões celulares exibiram diferentes graus de deposição da

parede celular secundária, que circundava a célula e formava no seu interior uma estrutura espiralada. Em células mais maduras e em processo de morte celular programada foi possível notar com nitidez uma parede celular secundária bem definida com deposição de lignina, constatada na coloração com safranina (Figura 13C), e a presença da lignina foi confirmada pela autofluorescência da molécula quando submetida à radiação ultravioleta (UV) (Figura 13E e 13F).

Figura 13. Etapas para indução de elementos traqueais *in vitro*. A) Formação de calos de *E. grandis* após 60 dias de incubação em meio de cultura com reguladores hormonais; B) Observação de elementos traqueais em agregados de células obtidos diretamente dos calos formados. C) Detalhe de um elemento traqueal formado pela diferenciação de uma célula meristemática, com parede secundária bem definida e corada com safranina; D) Visualização geral no microscópio óptico da suspensão celular após 24h sob agitação; E e F) Observação de elementos traqueais em microscópio de fluorescência, destacando a deposição da lignina autofluorescente. Barras: 100µm (B, C, E, F) e 200µm (D).

4.2. Qualidade do RNA total extraído

Foi possível extrair o RNA total das amostras e nenhuma amostra apresentou degradação indicando que todas possuíam alta qualidade. As amostras apresentaram um RIN (*RNA integrity number*) > 8.7, indicando que o RNA total extraído apresentava-se íntegro (Mueller and Schroeder, 2004) (Anexos B, C e D).

4.3. Sumário dos dados de sequenciamento e mapeamento

Para investigar os genes e fatores de transcrição envolvidos no desenvolvimento da parede celular secundária e no processo de lignificação em células xilemáticas de eucalipto, um total de 12 bibliotecas foram construídas. Foram gerados ao total de 420 milhões de leituras *paired-end* com 100 pb cada uma. A verificação do padrão de qualidade das sequências pelo programa FastQC reportou uma alta qualidade nas sequências com índices ao redor de 30 (Anexo E).

Após os processos de remoção dos adaptadores e filtragem das sequências por qualidade e tamanho pelo programa Trimmomatic constatou-se um aumento dos índices para 34 e aproximadamente 364 milhões de leituras (~86,7%) foram alinhados contra o genoma de eucalipto. Uma visão geral dos dados brutos obtidos no sequenciamento e após a limpeza é dada na Tabela 2. A maioria das leituras (~80%) foi alinhada contra genoma de referência. Os detalhes dos alinhamentos por amostra são fornecidos na Tabela 3.

Tabela 2. Visão geral dos dados do RNA-Seq. Número de leituras de 100pb obtidas em cada estágio celular sequenciado, antes e após a limpeza dos dados.

Estádio Celular	Pares de leituras brutas	Pares leituras após limpeza	Leituras remanescentes (%)
M1	42.065.576,00	37.229.270,00	88,50
M2	39.762.944,00	34.924.602,00	87,83
M3	38.768.906,00	34.319.172,00	88,52
M4	29.471.940,00	24.941.562,00	84,63
A1	39.394.456,00	34.810.554,00	88,36
A2	32.909.438,00	26.325.490,00	79,99
A3	26.551.876,00	23.453.458,00	88,33
A4	33.111.544,00	28.915.180,00	87,33
T1	43.677.536,00	38.281.904,00	87,65
T2	34.895.014,00	30.875.816,00	88,48
T3	29.847.090,00	24.045.856,00	80,56
T4	29.677.372,00	26.163.884,00	88,16
TOTAL	420.133.692,00	364.286.748,00	86,71

Tabela 3. Porcentagem (%) de leituras mapeadas em cada estágio celular.

Estádio Celular	Leituras mapeadas (%)	Pares de leituras alinhadas concordantemente (%)	Múltiplos alinhamentos por pares de leituras (%)
M1	87,7	80,6	9,9
M2	87,5	80,7	9,8
M3	87,9	81,1	10,9
M4	87,6	80,4	10,2
E1	87,6	80,4	10,2
E2	87,6	80,1	9,7
E3	87,8	81	11,9
E4	87,7	80,4	10,2
T1	87,8	80,9	11,8
T2	87,7	80,6	11,1
T3	87,9	80,9	11
T4	87,7	80,6	11
Média	87,70	80,64	10,64

4.4. Identificação dos genes diferencialmente expressos (GDEs)

A análise da expressão gênica foi realizada para identificar genes diferencialmente expressos (GDEs) durante o processo de diferenciação celular e deposição da parede celular

secundária. Considerando as três comparações M vs. A, A vs. T e M vs. T, um total de 12.387 genes e 13.653 genes foram considerados GDEs pelos programas *Cuffdiff* e pelo pacote *DESeq2*, respectivamente. Os genes diferencialmente expressos em dada comparação por programa está detalhado na Figura 14.

Figura 14. Detalhamento no número de genes diferencialmente expressos identificados em cada comparação, por cada programa ou pacote utilizado. As barras azuis referem-se aos resultados apresentados pelo programa *Cuffdiff* e as barras laranjas referem-se aos GDEs gerados pelo pacote *DESeq2*.

A construção do diagrama de Venn, contendo todos os resultados, mostrou o número de genes diferencialmente expressos exclusivamente para cada programa, bem com os GDEs igualmente encontrados nos dois programas *Cuffdiff* e *DESeq2*. Foi possível constatar que os mesmos 3426 genes foram diferencialmente expressos na comparação M vs. A, assim como, 1044 genes na comparação A vs. T e 4665 genes comparação M vs. T (Figura 15A). Além disso, por outro diagrama de Venn (Figura 15B) é possível observar como esses GDEs comuns estão distribuídos nas 3 comparações, sendo que 733, 175 e 1709 genes foram exclusivos das comparações M vs. A, A vs. T e M vs. T, respectivamente, e 272 genes foram comuns às 3 comparações.

Figura 15. Diagramas de Venn. A) Diagrama mostrando os genes comuns, representados pela intersecção, obtidos pelo Cuffdiff e DESeq2, em cada comparação testada. B) Diagrama mostrando o padrão de distribuição dos genes comuns entre todas as comparações.

Desses genes considerados no estudo, 1316, 567, 2637 GDEs foram significativamente mais expressos ou *up-regulated* e 2110, 476, 2028 foram menos expressos ou *down-regulated* nas comparações M/A, T/A, T/M, respectivamente. Como mostrado na Figura 16, o maior número de GDEs foi encontrado na comparação M vs. T, enquanto que o menor número de GDEs foi encontrado na comparação A vs. T. Delimitando os níveis de expressão para valores mais representativos ($\log_2 \text{fold change} \leq -1,5$ e $\log_2 \text{fold change} \geq 1,5$), os resultados quanto ao número de GDEs foi reduzido para 30, 46, 613 genes mais expressos e 445, 38, 58 genes menos expressos nas comparações M/A, T/A, T/M, respectivamente (Figura 16).

Figura 16. Gráficos com os GDEs comuns que foram *up-regulated* e *down-regulated* em cada comparação entre os tratamentos. O gráfico à esquerda abrange todos os genes e o gráfico à direita detalha somente genes com o $\log_2 \text{fold change}$ limitado.

Considerando os GDEs e os valores do $\log_2 \text{fold change}$ foi possível construir gráficos (*MA-plot*) correlacionando o $\log_2 \text{fold change}$ com as médias das contagens normalizadas comparando os níveis de expressão dos genes entre os tratamentos. Cada gene foi representado com um ponto, e os genes que tiveram um nível de expressão significativa para $p\text{-value} < 0,01$ foram representados em vermelho. O resultado visualizado nos gráficos revela que a comparação M vs. T apresentou um número maior de genes significativamente diferenciais (em vermelho), no intervalo do $\log_2 \text{fold change}$ de -2 a 2, quando comparado com o gráfico A vs. T que apresentou um número relativamente menor de GDEs (Anexo F).

Foi elaborado um gráfico de score de uma PCA em que a soma das duas primeiras componentes (PC1 e PC2) explica mais de 50% da variância dos dados. Para cada caso, as amostras biológicas foram agrupadas e diretamente conectadas. O gráfico separou perfeitamente pelos PC1 e PC2 os grupos, porém deixa acentuado a presença de um *outliner* no grupo das células meristemáticas (Figura 17). Esse comportamento de agrupamento das amostras também pode ser confirmado no dendograma (Anexo G).

Figura 17. Análise de componentes principais (PCA) realizada com os dados de expressão normalizados revela uma separação entre amostras pertencentes aos diferentes tratamentos. As amostras biológicas de cada tratamento foram representadas pelas cores vermelha (tratamento A), verde (tratamento M) e azul (tratamento T).

O *heatmap* com todos os genes estatisticamente significativos para o programa *DESeq2* (13.653 genes) mostrou uma associação entre as amostras de acordo com o perfil transcricional geral, sendo possível visualizar diferentes valores de expressão para os genes. Foi possível observar grupos de genes *up-regulated* (mais expressos) nos elementos traqueais e *down-regulated* (menos expressos) nas células meristemáticas, e vice-versa, fato verificado no *heatmap* pela intensidade das colorações azul e vermelha explicitando os genes *down* ou *up-regulated*, respectivamente. Também foi possível perceber que o padrão de expressão dos genes foi mais próximo entre as células alongadas e os elementos traqueais, enquanto as células meristemáticas apresentaram um perfil de expressão predominantemente diverso dos outros tratamentos (Figura 18).

Figura 18. Heatmap obtido pelo pacote *DESeq2* dos níveis relativos da expressão de todos os genes diferencialmente expressos, totalizando 13.653 genes. Os tons vermelhos representam genes com uma expressão mais elevada, quando os tons azuis representarem genes com expressão mais baixa.

4.5. Anotação funcional e análise de enriquecimento dos termos GO

Considerando todos os GDEs foi possível realizar a anotação funcional de 4229 genes pela plataforma Blast2GO. Observou-se com relação à função molecular que os genes foram classificados como relacionados a atividade catalítica e de ligação. As principais categorias de termos GO (ontologias genéticas) de componente celular pertencem à célula e membrana. E a maioria dos GDEs foram classificados em categorias GO envolvidas no processo metabólico, processo celular e processo de organismo-único. Esses resultados podem ser observados detalhadamente na Figura 19 que descreve as porcentagens de sequências relacionadas com os termos GO encontrados.

Figura 19. Anotações funcionais das sequências baseadas na categorização de ontologias genéticas (GO). Os dados estão resumidos em três categorias principais: função molecular, componente celular e processo biológico. Cada barra indica a porcentagem de uma categoria específica de genes dentro dessa categoria principal.

Estudando os genes que foram considerados diferencialmente expressos na comparação M vs. T, os mais expressos nos elementos traqueais apresentaram categorias GO de componente celular relacionados à membrana e partes da membrana (nível 2) e a componentes intrínsecos da membrana (nível 3). No aspecto de processo biológico, os processos metabólicos e celulares foram os destaques, com o processo metabólico de substâncias orgânicas, o processo de metabolismo primário e o processo metabólico celular como as categorias de nível 3 com mais termos GO representativos. A maioria das sequências apresentaram GOs de função molecular relacionada a atividade catalítica e de ligação, principalmente categorias relacionadas às ligações envolvendo compostos orgânicos cíclicos, compostos heterocíclicos e atividade de transferases. Em contrapartida, quando se observou os genes menos expressos entre os tratamentos M e T, a principal diferença foi na

categoria componente celular em que os genes apresentaram sequências pertencentes a célula e partes celulares, principalmente a intracelular.

A anotação do genoma pelo banco dos dados do KEGG possibilitou que os GDEs fossem identificados e relacionados a uma ou mais das 310 rotas identificadas. As rotas predominantes foram relacionadas ao metabolismo de tiamina, ao metabolismo de purina, a degradação de aminobenzoatos e a biossíntese de fenilpropanóides, como verificado na Figura 20. Nessa última rota, cerca de 40 sequências foram relacionadas às enzimas participantes da síntese dos monômeros da lignina (guaiacila, 5-hidroxi-guaiacil, siringila), principalmente a enzima lactoperoxidase (EC 1.11.1.7.). Outras enzimas como a cafeato O-metiltransferase (EC 2.1.1.68), aminobenzoato descarboxilase (EC 4.1.1.24) e tirosina descarboxilase (EC 4.1.1.25) também foram identificadas, sendo que todas participam em algum ponto da rota biossintética dos fenilpropanóides. Essas enzimas foram realçadas por cores distintas na rota dos fenilpropanóides fornecida pelo KEGG (Figura 21).

Figura 20. Rotas metabólicas identificadas no banco de dados do KEGG. Cada barra indica o número de sequências relacionadas a cada rota metabólica. Em destaque na cor laranja as sequências encontradas que foram relacionadas à rota bioquímica dos fenilpropanóides.

Figura 21. Rota da biossíntese dos fenilpropanoides fornecida pelo KEGG com as enzimas identificadas destacadas por cores distintas.

Quando se considerou os GDEs com $\log_2 \text{fold change} \leq -1,5$ e $\log_2 \text{fold change} \geq 1,5$, constatou-se que as rotas identificadas foram reduzidas, entretanto os resultados exibiram rotas mais específicas e interessantes para o estudo. Na comparação M vs. T os genes mais expressos nos elementos traqueais foram relacionados a 34 rotas metabólicas, destas a principal foi a via de biossíntese dos fenilpropanoides, com 14 sequências relacionadas a enzima lactoperoxidase (EC 1.11.1.7), que representou 18% do total das sequências ou genes.

Essa rota também foi destaque para os GDEs na comparação M vs. A em que a maioria dos genes mais expressos das células alongadas foram similarmente relacionados a biossíntese dos fenilpropanoides, com 16 sequências identificadas, que representaram 25% do total das sequências analisadas. Em A vs. T, devido ao número reduzido de genes *up-regulated* para os elementos traqueais o número de rotas foi claramente reduzido para 8 rotas identificadas, e dentre elas, o destaque continuou sendo para a biossíntese dos fenilpropanoides. As principais rotas metabólicas encontradas em cada comparação foram reunidas na Figura 22.

Figura 22. Números de seqüências correspondentes às rotas metabólicas do banco de dados do KEGG referentes aos genes com $\log_2 \text{fold change} \leq -1,5$ e $\log_2 \text{fold change} \geq 1,5$ para cada comparação testada.

Para calcular os termos GO mais representados pelos genes diferencialmente expressos foi utilizado o aplicativo BINGO que os exibiu como uma rede de termos GO significativos, nas categorias processo biológico, função molecular e componente celular. Foi realizado o enriquecimento dos termos GO para 2052 genes mais expressos nos elementos traqueais da comparação M vs. T que resultou em redes complexas com várias interconexões.

A rede obtida dos GOs na categoria processo biológico foi muito ampla com muitas conexões, considerando rotas relacionadas ao processo de deposição de parede celular secundária, observou-se GOs enriquecidos para o processo biológico de síntese de polissacarídeos celulares, como as vias para biossíntese de celulose e xiloglucanos, além da própria rota relacionada aos processos metabólicos envolvendo os fenilpropanóides, como o processo metabólico da lignina. Nessa rede também foram enriquecidos processos

biológicos relacionados à biossíntese de ácido jasmônico, e as rotas de sinalização mediadas pelo próprio ácido jasmônico, brassinosteróides e principalmente pelo etileno (Figuras 23A e 23B). Quanto a categoria componente celular, encontraram-se GOs relacionados à célula como um todo englobando tanto a parte intracelular como a parede celular, além de organelas como os cloroplastos e peroxissomos. E os termos GOs enriquecidos na categoria função molecular proporcionaram a obtenção de uma rede muito diversa com ênfase para a atuação de glicosiltransferases, peroxidases e celulose sintetases como enzimas representativas para os elementos traqueais (Figura 23C).

Figura 23. Ampliação das redes de termos GO enriquecidos e significativos para todos os genes mais expressos na comparação M vs. T. A intensidade da coloração mostra a significância estatística, com nível de significância de enriquecimento de $p < 0,05$ e a taxa de falsas descobertas de $FDR < 0,05$. O tamanho do nó é proporcional ao número de genes na categoria GO. As setas vermelhas destacam categorias GO relevantes no estudo da xilogênese. A e B) categoria GO: processo biológico e C) categoria GO: função molecular. As redes completas reportadas pelo aplicativo BINGO estão no Anexo H.

Quando foram considerados GOs correspondentes aos 423 genes anotados com $\log_2 \text{fold change} \geq 1,5$, na comparação M vs. T, a configuração da rede para os processos biológicos foi modificada, em que as rotas de biogênese e organização da parede celular vegetal e de morte celular foram proeminentes (Figura 24A). Os processos relacionados à resposta celular acarretadas pelo etileno e o processo metabólico do ácido jasmônico continuaram como termos enriquecidos (Figura 24B). Os GOs no aspecto de componente celular foram especialmente relacionados com a parede celular e parede celular vegetal e na categoria função molecular a atividade catalítica foi representativa com ênfase para a participação das transferases como peroxidases e UDP-glicosiltransferases (Figuras 24C e 24D).

Figura 24. Ampliação das redes de termos GO enriquecidos e significativos para os genes mais expressos na comparação M vs. T com $\log_2 \text{fold change} \geq 1,5$. A intensidade da coloração mostra a significância estatística, com nível de significância de enriquecimento de $p < 0,05$ e a taxa de falsas descobertas de $FDR < 0,05$. O tamanho do nó é proporcional ao número de genes na categoria GO. As setas vermelhas destacam categorias GO relevantes no estudo da xilogênese. A e B) categoria GO: processo biológico, C) categoria GO: componente celular e D) categoria GO: função molecular. As redes completas reportadas pelo aplicativo BINGO estão no Anexo I.

Em contrapartida, quando foram estudados os genes mais expressos nas células meristemáticas e os menos expressos nos elementos traqueais o padrão de correlações foi modificado. Para os processos biológicos os que apresentaram mais termos GO representativos estavam relacionados a organização celular e desenvolvimento, resposta a estímulos, morte celular programada e apoptose (Figura 25). Os termos enriquecidos para componente celular foram relacionados a vários locais da célula, havendo uma concentração de sequências na região intracelular e citoplasmática.

4.6. Fatores de transcrição de interesse

Dentre eles, os fatores responsivos ao etileno (ERFs) foram representativos, interagindo com genes estruturais das rotas de síntese de celulose, xilana e lignina. O ERF4 foi destaque pela correlação com genes estruturais das três vias biossintéticas estudadas (*COBL4*, *IRX9*, *C4H*). Esse fator de transcrição foi identificado nas comparações M vs. T e A vs. T, sendo mais expresso nos elementos traqueais. O ERF3 (*ethylene responsive element binding factor 3*) foi diferencialmente expresso nas 3 comparações, sendo que na M vs. A

apresentou-se mais expresso nas células meristemáticas, e nas comparações A vs. T e M vs. T, esse fator de transcrição foi mais expresso nos elementos traqueais. O ERF3 regula a expressão de genes importantes da rota de biossíntese da celulose (*COBL4*) e das heteroxilanas (*IRX9* e *GXM3*). Em contrapartida, o ERF2 somente foi constatado na comparação M vs. T com um nível de expressão maior nos elementos traqueais, atuando nos genes *COBL4* e *PAL4*, este último importante no início da rota de biossíntese da lignina.

Outro fator de transcrição foi o DREB26 (*ERF family protein*), pertencente da família de genes APETALA2 e que interage em mais de uma via, podendo regular a expressão do *CESA7*, *COBL4* da rota de biossíntese de celulose e *GUX1* e *IRX10* da rota de biossíntese de xilana. O DREB26 foi significativo nas comparações M vs. T e M vs. A, sendo *up-regulated* nas células alongadas e nos elementos traqueais, quando comparado às células meristemáticas.

O fator de transcrição CBF4/DREB1 (*C-repeat-binding family 4*) foi observado em todas as comparações realizadas e interage com os genes participantes das rotas das xilanas (*IRX9*) e da lignina (*CAD4*, *C4H* e *PAL4*). Esse fator de transcrição apresentou níveis de expressão maiores nos elementos traqueais, ou seja, um valor de $\log_2$ *fold change* maior na comparação com as células meristemáticas, do que com as células alongadas. Já na comparação M vs. A, o CBF4 foi mais expresso nas células alongadas. O GRF1 (*growth-regulating family 1*) somente foi identificado na comparação M vs. A sendo mais expresso nas células meristemáticas e aparentemente, está relacionado a genes que codificam enzimas que atuam no final das rotas de biossíntese de xilanas (*IRX9*) e lignina (*LAC17*).

Outros fatores de transcrição foram relacionados com genes estruturais participantes de uma única via biossintética. O AIL6 (*AINTEGUMENTA-like 6*) apareceu interagindo unicamente com o gene *COBL4* da rota de biossíntese de celulose e foi identificado nas comparações M vs. T e M vs. A, e em ambas, o AIL6 foi mais expresso nas células meristemáticas. Considerando a rota de biossíntese de xilana os fatores de transcrição JAM 2 (*bHLH family protein*) e IRX11 (*KNOTTED-like homeobox of Arabidopsis thaliana 7*) apareceram regulando os genes *GUX1* e *GUX2*, respectivamente. O fator de transcrição JAM2 foi verificado nas comparações M vs. T e M vs. A, sendo *up-regulated* nos elementos traqueais e nas células alongadas, quando comparado o nível de expressão desse fator com as células meristemáticas.

No caso especial da rota de biossíntese de lignina, foi possível identificar um maior número de fatores de transcrição, entre eles, o CAMTA2 (*Calmodulin-binding transcription*

activator protein with CG-1 and Ankyrin domains), o AUX2-11 (*AUXIN INDUCIBLE 2-11*) e o RAP2.12 (*related to AP2 12*), estes fatores de transcrição foram relacionados a genes únicos que codificam enzimas da rota de lignificação, *PAL4*, *OMT* e *HCT*, respectivamente. Esses três fatores de transcrição foram identificados somente na comparação M vs. T, o CAMTA2 foi expresso tanto nos elementos traqueais como nas células meristemáticas, o AUX2-11 foi mais expresso nas células meristemáticas enquanto que o RAP2.12 foi mais expresso nos elementos traqueais.

O fator de transcrição PAT1 (*GRAS family protein*) regula a expressão de três genes estruturais importantes (*PAL4*, *4CL1* e *HCT*) e codificam enzimas participantes da formação dos monômeros da lignina. Foi verificado nas comparações M vs. A e M vs. T, sendo mais expresso nas células alongadas e elementos traqueais em comparação às meristemáticas, contudo, o nível de expressão observado desse fator de transcrição foi maior nas células alongadas.

Os últimos fatores de transcrição identificados são pertencentes às famílias NAC e MYB, e são os mais descritos na literatura. O fator de transcrição MYB63 regula a expressão do gene *CESA8* e de quatro genes da rota biossintética da lignina (*C4H*, *4CL1*, *HCT* e *CCoAOMT1*), o MYB20 regula os genes *C4H*, *4CL1* e *HCT* e o fator MYB56 o gene *C4H*. Todos esses fatores de transcrição foram observados exclusivamente na comparação M vs. T, o MYB63 apresentou maior expressão nos elementos traqueais, assim como, o MYB20. Por outro lado, o fator de transcrição MYB56 apresentou-se *up-regulated* nas células meristemáticas.

Com relação aos fatores de transcrição da família NAC, foi possível identificar os fatores de transcrição NAC102, NAC073 (*SND2*) e NAC010 (*SND3*). O NAC102 regula a expressão de quatro genes da rota biossintética da lignina: *4CL1*, *HCT*, *CCoAOMT1* e *CAD4*, englobando os principais passos dessa rota, enquanto o NAC073 interage com o *CCoAOMT1*. Esses fatores de transcrição foram reconhecidos nas comparações M vs. A e M vs. T, sendo mais expressos ou *up-regulated* nas células alongadas e elementos traqueais, respectivamente. Sendo o fator de transcrição NAC102 que apresentou o maior o nível de expressão.

Foi possível identificar a ocorrência de 10 fatores da transcrição da família WRKY, a maioria deles apresentou-se mais expresso nas células alongadas e principalmente nos elementos traqueais. Com destaque para os fatores de transcrição WRKY53 (Eucgr.E00850) e

WRKY41 (Eucgr.D02359, Eucgr.K00331) que se apresentaram consideravelmente mais expressos nos elementos traqueais do que nas células meristemáticas (Anexos J, K e L).

O *heatmap* (Figura 26) construído contendo a abundância relativa dos fatores de transcrição em cada tratamento mostrou visualmente que houve diferença na abundância desses fatores, alguns apresentando um aumento crescente no valor da abundância ($\log_{10}$ FPKM+1) entre as células meristemáticas, células alongadas e elementos traqueais. Fato verificado pela intensificação da cor entre os tratamentos, as células meristemáticas apresentaram alguns fatores de transcrição representados por colorações mais claras e as células alongadas e elementos traqueais já apresentaram coloração mais similar e mais escura.

Esse padrão pode ser observado na primeira metade do *heatmap*, com destaque para os fatores de transcrição NAC010 (*LocusName*: Eucgr.D00591), MYB20 (Eucgr.F02864), NAC102 (Eucgr.I00057), DREB26 (Eucgr.F04365, Eucgr.F03772), AUX2-11 (Eucgr.K03314) e CBF4 (Eucgr.A02818, Eucgr.A02820), em que a diferença na coloração foi bem nítida. As células meristemáticas apresentaram uma coloração mais clara enquanto que as células alongadas e elementos traqueais apresentaram um padrão similar, com uma coloração mais intensa, indicando maior abundância dos transcritos. Outros comportamentos também foram observados, tais como a mesma intensidade de coloração em todos os tratamentos indicando uma regularidade na abundância dos transcritos (RAP2.12) e o aumento ou diminuição gradativa da intensidade da cor, como no caso do ERF3 (Eucgr.F00659) e do AUX2-11 (Eucgr.H04336), respectivamente (Figura 26).

Tabela 4. Identificação das interações entre FTs e promotores envolvidos na biossíntese de celulose (amarelo), xilana (verde) e lignina (azul). Cada caixa cinza indica a interação, enquanto a caixa branca indica ausência de interação.

[illegible]

Figura 26. Heatmap dos valores de $\log_{10} \text{FPKM} + 1$ referentes aos fatores de transcrição identificados em cada comparação. A intensidade da expressão (FPKM) é indicada pela coloração, variando baixa (amarelo claro) a intensa (marrom) com valores maiores de FPKM. Cada fator de transcrição foi identificado pelo seu *Gene Class Symbol* e *LocusName* correspondente.

Para confirmar os resultados observados no *heatmap* acima foi construído um gráfico de barras (Figura 27) com os valores do FPKM +1, sendo possível observar melhor o padrão de abundância relativa dos transcritos, em que o gene RAP2.12 (*LocusName*: Eucgr. I02155), CAMTA2 (Eucgr. I02662) e ERF3 (Eucgr. F03499) apresentaram os maiores valores. A maioria dos genes apresentou as barras correspondentes aos tratamentos A e T (cores azul e verde, respectivamente) maiores que a barra cor de rosa (tratamento M).

Os transcritos dos FTs CBF4 (Eucgr. A02818, Eucgr.A02820, Eucgr.A02822), AUX2-11 (Eucgr.K03314), DREB 26 (Eucgr.F04365, Eucgr.F03772) e NAC102 (Eucgr.I00057) apresentaram-se mais abundantes nos tratamentos A e T, com uma grande disparidade nos valores do FPKM quando comparado às células meristemáticas. Por outro lado, os FTs ATERF (Eucgr.K00126), MYB (Eucgr.E01658), ERF3 (Eucgr.F03499), AIL6 (Eucgr.I00564) e principalmente, ATAUX2-11 (Eucgr.H04336) foram mais expressos nas células meristemáticas (Figura 27).

Além disso, os transcritos dos FTs CBF4 (Eucgr.A02820, Eucgr.A02820), ERF3 (Eucgr.K00734) ERF2 (Eucgr.L02041), ERF3 (Eucgr.F00659), RAP2.12 (Eucgr.I02155), NAC102

Figura 27. Gráfico de barras contendo os FTs e seus respectivos valores de abundâncias relativas (FPKM+1) e *locusName*, para cada estágio celular analisado, esses representados pelas cores azul (células alongadas), rosa (células meristemáticas) e verde (elementos traqueais).

4.7. Genes relacionados à biossíntese dos componentes da parede celular secundária

Como mostrado anteriormente, foi possível identificar genes diferencialmente expressos (GDEs) em cada comparação realizada. Analisando amplamente esses GDEs foi possível selecionar 170 genes na comparação M vs. T, 153 genes na comparação M vs. A e 65 genes na comparação A vs. T que estiveram relacionados em algum momento à biossíntese e deposição da parede celular secundária nas células (Anexos J, K e L). Para compreender de forma ampla a regulação gênica e as diferenças no transcriptoma de cada fase celular alguns genes-chaves foram destacados. Essa ênfase foi baseada em estudos prévios realizados com eucalipto (Hefer et al., 2015; Kumar et al., 2016; Myburg et al., 2014).

Os GDEs selecionados foram relacionados à enzima que codificam e ao sítio de atuação na rota de biossíntese dos respectivos componentes estruturais da parede celular. No caso da rota da celulose e hemiceluloses a partir da sacarose foram especificados 21, 7 e 23 genes (provenientes das comparações M vs. A, A vs. T e M vs. T, respectivamente). Enquanto na rota dos fenilpropanóides, o número de genes identificados codificadores de enzimas chaves da rota de síntese dos monômeros da lignina foi diferente correspondendo a 24, 8 e 21 genes provenientes das comparações M vs. A, A vs. T e M vs. T, respectivamente).

Analisando os genes envolvidos na rota de síntese da celulose e das hemicelulose foi possível visualizar nas comparações M vs. A e M vs. T que ocorreram genes codificadores de proteínas e enzimas que participaram de praticamente todos os passos da rota, sendo que alguns desses genes foram observados nas duas comparações e outros foram exclusivos para cada uma (Figura 28A e 28C). Já na comparação A vs. T, o número de GDEs foi inferior ao envolvimento de genes nas vias finais dessas rotas (Figura 28B).

Na comparação M vs. A, foi possível constatar que a maioria dos genes foram mais expressos nas células alongadas, como os genes *CINV2* (*cytosolic invertase 2*), *H XK* e *HKL* (*hexokinase*) e *SUSY* (*sucrose synthase*) que codificam enzimas do início da rota de síntese da celulose e xilanas. Do mesmo modo foi observado nos genes intermediários e finais dessa rota como o *UXS6* (*UDP-XYL synthase 6*), *UGD* (*UDP-glucose 6-dehydrogenase family protein*), *IRX* (*Exostosin family protein*), *GATL* (*galacturonosyltransferase*), *TBL* (*trichome birefringence*) e *CESA9* (*cellulose synthase 9*) os quais codificam enzimas e proteínas fundamentais para síntese e montagem das moléculas dos polímeros e que foram mais expressos nas células alongadas. Em contrapartida, alguns genes apresentaram-se mais expressos nas células meristemáticas como o *UXS2* (*NAD(P)-binding superfamily protein*) e a *CESA1* (*cellulose synthase 1*), como indicado na Figura 28A.

No caso da comparação A vs. T os genes identificados foram relacionados a enzimas atuantes nas vias finais das rotas, ocorrendo genes mais expressos nas células alongadas como o gene *GATL1* (*galacturonosyltransferase-like 1*) e genes *CESA9* (*Cellulose synthase A9*, Eucgr.I00286 e Eucgr.F03635) sendo que esses últimos dois genes *CESA* também foram identificados como mais expressos nas células alongadas na comparação anterior, como já comentado, e genes mais expressos nos elementos traqueais como o *GATL10* (*galacturonosyltransferase-like 10*), *CESA1* (*cellulose synthase 1*) e o *IRX15* (Figura 28B). Tanto o *GATL* quanto o *CESA* são genes que já haviam sido identificados, no caso do *GATL10* este gene mostrou-se mais expresso nas células alongadas na comparação M vs. A, e diferentemente, na comparação A vs. T apresentou-se mais expresso nos elementos traqueais. Enquanto o *CESA1* foi *up-regulated* nas células meristemáticas comparação M vs. A. No caso específico do gene *IRX15* este foi identificado como mais expresso nos elementos traqueais e não foi observado na comparação M vs. A.

Na comparação M vs. T em que foram identificados mais genes participantes, observou-se que a maioria deles também foram expressos na comparação M vs. T com

Figura 28. Genes preferencialmente expressos nas vias de biossíntese de celulose e hemiceluloses. Cada gene possui um círculo ao lado (vermelho ou azul) indicando em qual tratamento ou fase celular foi *up-regulated*. A descrição completa dos genes e seus valores de expressão são fornecidos nas tabelas dos Anexos J, K e L.

Foi possível construir um *heatmap* e um gráfico de barras representando o comportamento dos transcritos de cada gene identificado (Figuras 29 e 30). De maneira geral, foi possível observar o predomínio de genes com valores maiores de FPKM nos tratamentos A e T. Esse comportamento foi constatado na intensidade da coloração no *heatmap* e na amplitude das barras de coloração azul e verde.

Os genes *SUS3* (Eucgr.H01094), *UGD3* (Eucgr.J01372), *UXS2* (Eucgr.H01112) e *CINV2* (Eucgr.F02588) apresentaram as maiores abundâncias relativas de transcritos juntamente com a coloração mais intensa no *heatmap* (Figura 29). Pode-se observar que alguns genes contiveram uma abundância relativa dos transcritos superior para um determinado tipo celular, como no caso dos genes *TBR* (Eucgr.A02463), *TBL29* (Eucgr.J00196), *GATL2* (Eucgr.F01531, Eucgr.H01923), *IRX15-L* (Eucgr.I01114) e principalmente, *GATL10* (Eucgr.I01882) que se apresentaram mais representados nos elementos traqueais. Enquanto outros genes encontraram-se mais expressos nas células alongadas, como os genes *GATL1* (Eucgr.A00485), *HLK1* (Eucgr.B03711), *CESA9* (Eucgr.F03635, Eucgr.I00286), *TBL33* (Eucgr.G03092) e *TBL22* (Eucgr.I01832), e diferentemente, com transcritos mais abundantes nas células meristemáticas, como os genes *CESA1* (Eucgr.J01639, Eucgr.H00939) (Figuras 29 e 30).

Figura 29. Heatmap dos valores de abundâncias relativas dos transcritos (log₁₀FPKM + 1) dos genes diferencialmente expressos da rota de biossíntese da celulose e hemiceluloses. O nível de expressão dos genes é indicado pela intensidade da coloração apresentada por cada tratamento. Cada gene foi identificado pelo seu *Gene Class Symbol* e *LocusName* correspondente.

Figura 30. Gráfico de barras contendo os GDEs participantes da rota de biossíntese da celulose e hemiceluloses e seus respectivos valores de abundâncias relativas (FPKM+1) e *locusName*, para cada estágio celular analisado, esses representados pelas cores azul (células alongadas), rosa (células meristemáticas) e verde (elementos traqueais).

Analisando os GDEs relacionados à rota de síntese dos monômeros da lignina: álcool cumarílico, álcool coniferílico e álcool sinapílico foi possível identificar genes codificadores de enzimas importantes dessa rota. Nas comparações realizadas foi possível constatar que

os mesmos genes foram expressos nas três comparações, principalmente na M vs. A e M vs. T. Na comparação M vs. A os genes *PAL* (*phenylalanine ammonia-lyase*), *C4H1* (*cinnamate-4-hydroxylase*), *4CL9*, *HCT* (*quinic acid hydroxycinnamoyl transferase*), *CCoAOMT*, *COMT1*, *CCR7* e *CCR9*, e a maioria dos genes *CAD* (*cinnamyl-alcohol dehydrogenase*) foram mais expressos nas células alongadas (Figura 31A). Notavelmente, alguns genes dessa rota de biossíntese de lignina foram observados como mais expressos nas meristemáticas, sendo verificados tanto no início da via como o gene *4CL7* e nas vias finais representadas pelos genes *CCR5*, *COMT1* e genes *CAD* (*CAD21*, *CAD34* e *CAD37*).

Esse fato também foi observado na comparação M vs. T (Figura 31C) onde diversos genes foram mais expressos nos elementos traqueais como os genes *PAL*, *C4H1*, *4CL9*, *HCT*, *CCR7*, e genes *CAD*, enquanto outros foram mais expressos nas células meristemáticas, tal como os genes *4CL8* e *4CL9*, *C3H4*, *CCR5* e *CCR9*, além dos *CAD* (*CAD15* e *CAD37*). É importante salientar que muitos genes que foram mais expressos nas células alongadas (comparação M vs. A), continuaram mais expressos nos elementos traqueais, como os genes *PAL7* e *PAL9*, *HCT4* e *HCT5*, *CCR7* e *CAD4*, *CAD5* e *CAD9*. Assim como, genes que foram mais expressos nas células meristemáticas, continuaram como destaque para esse tipo celular, como no caso dos genes *COMT37*, *CCR5* e *CAD37*.

Um número inferior de genes foi identificado na comparação A vs. T (Figura 31B), sendo que somente o gene *F5H1* (*ferulic acid 5-hydroxylase 1*) foi exclusivo dessa comparação estando mais expresso nas células alongadas. Os genes *4CL9*, *CCR5* e *COMT37* foram mais expressos nas células alongadas, sendo também encontrados nas duas outras comparações. Os genes *CCoAOMT4*, *COMT14* também ocorreram na comparação M vs. A e o gene *C3H4* foi identificado na comparação M vs. T.

Tratando-se dos genes exclusivos, somente 3 genes apareceram unicamente em uma comparação entre as fases celulares. O *F5H1* já comentado e que foi identificado somente na comparação A vs. T e os genes *COMT1* e *PAL3* sendo que o primeiro somente ocorreu na comparação M vs. A e o segundo na comparação M vs. T.

Figura 31. Genes preferencialmente expressos na via biossintética dos fenilpropanoides que conduz à produção de monômeros da lignina. Cada gene possui um círculo ao lado (vermelho ou azul) indicando em qual tratamento ou fase celular foi *up-regulated*. A descrição completa dos genes e seus valores de expressão são fornecidos nas tabelas dos Anexos J, K e L.

Considerando a abundância relativa dos transcritos de genes participantes da biossíntese de lignina foi possível constatar que esses genes apresentaram de modo geral valores maiores de $\log_{10}$ FPKM +1 que os genes das rotas da celulose e xilanas. Ao que tudo indica, os genes apresentaram maior abundância relativa de transcritos nas células alongadas e nos elementos traqueais do que nas células meristemáticas, com exceção para os genes *CCR9* (LocusName: Eucgr.I01783), *CAD17* (Eucgr.H02414), *4CL9* (Eucgr.F03543), *COMT37* (Eucgr.F02625), *COMT35* (Eucgr.F02623) com mais transcritos nas células meristemáticas, principalmente o gene *4CL8* (Eucgr.F02624), o qual apresentou o maior valor de FPKM obtido (Figuras 32 e 33).

Pode-se verificar pela intensidade da coloração do *heatmap* que muitos genes tiveram riqueza de transcritos (valores maiores de $\log_{10}$ FPKM+1) nas células alongadas, fato confirmado e melhor visualizado no gráfico de barras (Figura 33), como os genes *COMT1* (Eucgr. A01397), *HCT5* (Eucgr. J03126), *PAL9* (Eucgr.J01079), *CAD4* (Eucgr.D00468), *CAD5* (Eucgr.D00471), *C3H4* (Eucgr.G03199), *F5H1* (Eucgr.I02371). Por outro lado, alguns genes apresentaram maior abundância relativa de transcritos nos elementos traqueais tal como os

genes *C4H1* (Eucgr.J01844), *4CL7* (Eucgr.B03943) e *CAD9* (Eucgr.D01087). Por fim, o gene *CAD10* (Eucgr. D00471) somente apresentou transcritos nos tratamentos A e T (Figuras 32 e 33).

Figura 32. Heatmap dos valores de abundâncias relativas dos transcritos ($\log_{10}\text{FPKM} + 1$) dos genes diferencialmente expressos na rota de biossíntese da lignina. O nível de expressão dos genes é indicado pela intensidade da coloração apresentada por cada tratamento. Cada gene foi identificado pelo seu *Gene Class Symbol* e *LocusName* correspondente.

Figura 33. Gráfico de barras contendo os GDEs participantes da rota de biossíntese da lignina e seus respectivos valores de abundâncias relativas ($\text{FPKM} + 1$) e *locusName*, para cada estágio celular analisado, esses representados pelas cores azul (células alongadas), rosa (células meristemáticas) e verde (elementos traqueais).

Para compreender as etapas finais da lignificação foi verificada a expressão diferencial das lacases e peroxidases entre os tratamentos M, A e T. Os genes que codificam as lacases foram o *LAC7* e o *LAC14*, sendo que todos os *LAC7* (Eucgr.B02752, Eucgr.K03009, Eucgr.K03008) apresentaram-se mais expressos nos elementos traqueais com maior abundância relativa dos transcritos, como mostrado no gráfico em que as barras de coloração verde são destacadas com valores maiores de FPKM +1 (Figura 34).

Em contrapartida, o gene *LAC14* apresentou exemplares que foram mais expressos nos elementos traqueais como os genes Eucgr.E04220, Eucgr.F02650, Eucgr.J00889, Eucgr.F02642 e outros mais expressos nas células meristemáticas tais como os genes Eucgr.F04160, Eucgr.F04163 e principalmente o Eucgr.L01345. Todavia dois genes de lacases (Eucgr.B02752 e Eucgr.J00889) foram fracamente expressos nas células meristemáticas quando comparados os valores do FPKM com as células alongadas e elementos traqueais. De maneira geral, foi possível observar um padrão de expressão diverso da família de genes codificadores das lacases (Figura 34).

Figura 34. Gráfico de barras contendo os GDEs codificadores de lacases e seus respectivos valores de abundâncias relativas (FPKM+1) e *locusName*, para cada estágio celular analisado, esses representados pelas cores azul (células alongadas), rosa (células meristemáticas) e verde (elementos traqueais).

Quando foi analisada a regulação da expressão da família de genes codificadores das peroxidases foram identificados 24 genes diferencialmente expressos nos tipos celulares estudados, sendo que a maioria deles foi mais expresso nas células alongadas. Esse fato pode ser confirmado no *heatmap* construído (Figura 35), em que é nítido uma coloração

mais intensa na coluna central que representa os valores de $\log_{10}\text{FPKM} + 1$ dos genes do tratamento A (células alongadas).

Além disso, no gráfico de barras é possível confirmar essa constatação, pois a grande maioria dos genes apresentaram as respectivas barras azuis proeminentes, como no caso dos genes Eucgr. J02352, Eucgr. E03715, Eucgr.E04120, Eucgr.H03910 que se mostraram mais expressos nas células alongadas. Poucos genes apresentaram-se mais expressos nos elementos traqueais, dentre eles, Eucgr.J02173, Eucgr.F04198, Eucgr.A00682 e os genes que apresentaram maior abundância de transcritos, os genes Eucgr.A01385 e Eucgr.A00966, mostraram um padrão de expressão similar entre os tratamentos com valores de FPKM bem próximos (Figura 36).

Figura 35. Heatmap dos valores de abundâncias relativas dos transcritos ($\log_{10}\text{FPKM} + 1$) dos genes diferencialmente expressos codificadores de peroxidases. O nível de expressão dos genes é indicado pela intensidade da coloração apresentada por cada tratamento. Cada gene foi identificado pelo seu *LocusName*.

Figura 36. Gráfico de barras contendo os GDEs codificadores de peroxidases e seus respectivos valores de abundâncias relativas (FPKM+1) e *locusName*, para cada estágio celular analisado, esses representados pelas cores azul (células alongadas), rosa (células meristemáticas) e verde (elementos traqueais).

5. DISCUSSÃO

5.1. Obtenção *in vitro* e caracterização morfológica dos elementos traqueais de *Eucalyptus grandis*.

A área de cultura de tecidos vegetais para a indução *in vitro* de calos em *Eucalyptus grandis* vem sendo estudada visando principalmente à otimização de protocolos para a regeneração e propagação de plantas e como ferramenta na área de transformação genética (Fernando et al., 2016; Hajari et al., 2006; Li et al., 2015a; Nakhooda and Mandiri, 2016; Oberschelp et al., 2015). Entretanto, a metodologia para a indução da formação *in vitro* de células xilemáticas de eucalipto, como os elementos traqueais, não havia sido realizada anteriormente e, portanto, os resultados obtidos nesse trabalho foram inovadores.

Verificou-se que os calos foram induzidos na totalidade dos explantes incubados, ocorrendo a multiplicação celular e posterior aumento no volume dos calos, demonstrando que o protocolo utilizado foi satisfatório para a calogênese, e ademais, para a obtenção de calos friáveis (Figura 13A), os quais formaram aglomerados celulares não-oxidados em que as células se apresentavam fracamente aderidas e, por essa razão, esse tipo de calo foi apropriado para formação das suspensões celulares testadas.

Essa alta porcentagem na formação de calos também foi observada por Fernando et al. (2016); Hajari et al. (2006); Li et al. (2015) que obtiveram 98,8%, 86% e 95% dos explantes com a presença de calos em *Eucalyptus polybractea*, *Eucalyptus grandis* e *Eucalyptus urophylla*, respectivamente. O tempo para a indução dos calos nos explantes de *E. grandis* foi o mesmo observado por Li et al. (2015) de 30 dias, mas em outros experimentos, o tempo decorrido foi, entre 6 a 8 semanas para que os calos se desenvolvessem nos explantes (Fernando et al., 2016; Hajari et al., 2006; Nakhooda and Mandiri, 2016; Oberschelp et al., 2015). De maneira geral, os resultados obtidos indicaram que o meio de cultura utilizado se mostrou apropriado para o objetivo proposto nesse trabalho. Porém, constatou-se que o período de quatro meses despendido para obtenção, constituição adequada e quantidade suficiente de calos foi um fato que precisou ser considerado e planejado, pela necessidade de subcultivos e cuidados para evitar contaminações microbianas.

Como descrito, os meios líquidos testados (baseados em artigos de referência na área) não induziram a formação *in vitro* dos elementos traqueais ao final do período de incubação, demonstrando que no caso do eucalipto a obtenção dessas células deveria utilizar outra metodologia. Quando os calos friáveis de eucalipto foram diretamente observados ao microscópio óptico descobriu-se a formação de elementos traqueais em agrupamentos celulares, demonstrando que o próprio meio de cultura sólido havia induzido parte das células a diferenciarem-se em elementos traqueais. Esse comportamento também foi reportado em coníferas, *Pinus radiata*, álamo híbrido (*Populus sieboldii* x *P. grandidentata*) e soja (Möller et al., 2006a; Yamagishi et al., 2012, 2013, 2015) em calos com a coloração esbranquiçada e recém-formados. Nesses casos, foi possível observar a presença dos elementos traqueais pela birrefringência exibida por essas células. Esse fenômeno confirmou a importância do contato célula a célula e das interações químicas celulares para que ocorra a formação das células xilemáticas; esse papel da vizinhança na diferenciação celular foi descrito por Ménard and Pesquet (2015).

No caso do eucalipto a indução dos elementos traqueais foi tarefa laboriosa pela necessidade da diferenciação dessas células nos próprios calos durante o período total de 4 meses de incubação. Fato este não observado nas plantas-modelo *Zinnia elegans* e *Arabidopsis thaliana* as quais encabeçaram as pesquisas de indução *in vitro* de elementos traqueais (Fukuda and Komamine, 1980a, 1980b; Lacayo et al., 2010; Oda et al., 2005). Essas espécies necessitaram da utilização de suspensões celulares, contendo células isoladas do mesófilo foliar, por um período de incubação significativamente menor de 4 a 5 dias para a ocorrência da transdiferenciação em elementos traqueais.

Outro fato importante foi com relação a porcentagem de elementos traqueais obtidos nos experimentos com eucalipto frente a outras espécies. Como comentado, a porcentagem observada de elementos traqueais foi de 5 a 10% nas células dissociadas dos calos, sendo esse valor foi baixo quando comparado a outros estudos em que as porcentagens oscilaram de 30% (Fukuda and Komamine, 1980b) a 76% em *Z. elegans* (Twumasi et al., 2009), de 40% em *A. thaliana* (Pesquet et al., 2010) e de 2 a 45% em *P. radiata* (Möller et al., 2006a). Essa baixa porcentagem de elementos traqueais nos calos de eucalipto acarretou na necessidade da obtenção de um número maior de calos para a caracterização microscópica dessas células e obtenção do RNA total utilizado nos experimentos de RNA-seq.

Quanto às observações microscópicas dos elementos traqueais de *E. grandis* foi possível verificar que estes já estavam com uma parede celular secundária bem definida e lignificada. Essa constatação foi confirmada pela intensa coloração avermelhada dessas células pela safranina (Figura 13C). Como a substância apresenta alta afinidade por lignina, esse corante também foi utilizado em outros trabalhos como recurso para a localização de lignina depositada nos elementos traqueais formados em outras espécies (McCann, 1997; Oda et al., 2005; Pillai et al., 2011).

A visualização do padrão de lignificação dos elementos traqueais de eucalipto foi possível pela autofluorescência da molécula pelo raio de excitação verde (filtro de excitação 546 nm; filtro de emissão 590 nm), essa autofluorescência possibilitou verificar que a deposição da lignina se deu na parede da célula e internamente, em formato espiralado (Figura 13E e 13F). Essa técnica para visualizar o padrão de deposição da lignina por autofluorescência também foi utilizada com sucesso por Pesquet and Tuominen (2011) com *Z. elegans* em que a deposição da lignina foi verificada ao longo do tempo de indução *in vitro* dos elementos traqueais a partir de células do mesófilo foliar.

Com relação às características morfológicas baseadas principalmente no tamanho e padrão de espessamento e lignificação dos elementos traqueais, observou-se, como já discutido, somente o padrão espiralado ou helicoidal nos elementos traqueais de eucalipto, sendo que na literatura foram descritos também outros padrões como o reticulado, o anular, o escalariforme e o perfurado. E em muitos casos, os elementos traqueais de uma mesma espécie apresentaram vários padrões de deposição da parede celular secundária, como em *Arabidopsis* (Oda et al., 2005) e no híbrido *Populus sieboldii* x *P. grandidentata* (Yamagishi et al., 2013) que foram visualizados três padrões (espiralado, escalariforme e reticulado), em *Cryptomeria japonica* (Yamagishi et al., 2015) além desses três padrões também constatarem-se elementos traqueais perfurados e em *Torreya nucifera* (Yamagishi et al., 2012) houve a predominância de células com perfurações com formato de pontos.

Para finalizar a caracterização desses elementos podemos comentar sobre o tamanho e formato dos mesmos. Os elementos traqueais de eucalipto visualizados apresentaram tamanhos variando de 400 a 800 µm de comprimento, aspecto que os caracterizou como os maiores elementos traqueais descritos na literatura. Para se ter uma ideia de como os elementos traqueais de eucalipto são extensos podemos considerar os já descritos na literatura. A conífera *Cryptomeria japônica* apresentou elementos traqueais *in*

vitro com tamanhos variando de 100 a 200 μm e o híbrido *Populus sieboldii* x *P. grandidentata* exibiram elementos traqueais com 100 μm de comprimento. Em contrapartida, os menores elementos traqueais descritos foram verificados em *Arabidopsis*, *Pinus* e *Zinnia*, com respectivamente, uma média de 20 μm , 40 μm e 60 μm de comprimento.

É importante salientar que as características gerais apresentadas pelas células *in vitro* de eucalipto, com sua correspondente correlação com os estudos previamente realizados com outras espécies, puderam confirmá-los como sendo de fato elementos traqueais. Portanto, o aspecto visual do elemento traqueal resultado do padrão de deposição da lignina e conseqüentemente, da parede celular secundária interna somado ao tamanho apresentado por desse tipo celular tornam-se um elemento poderoso para caracterização inédita dessas células xilemáticas obtidas *in vitro*.

5.2. Perfil da expressão gênica e classificação funcional

Considerando os aspectos técnicos, o RNA total extraído das diferentes frações celulares apresentou uma alta integridade e qualidade confirmando que o protocolo utilizado obteve sucesso mesmo utilizando uma quantidade inicial de material celular pequena e que não havia sido utilizada anteriormente para essa finalidade. Sendo, portanto, o primeiro trabalho realizado utilizando a técnica de RNA-seq e o transcriptoma de calos em suspensão de eucalipto.

Na literatura há poucos exemplos de trabalhos que utilizaram calos para estudos de expressão gênica e RNA-seq, como exemplo recentes temos os trabalhos de Siah et al. (2016) que utilizaram calos de uma espécie indiana chamada *Aquilaria malaccensis* para estudar genes relacionados a resposta a estresses e biossíntese de terpenos. E outro caso, foi a publicação de Ma et al. (2015) que usou calos friáveis de mandioca (*Manihot esculenta* Crantz) para realizar uma ampla investigação da expressão gênica diferencial de embriões somáticos, calos novos e velhos, sob diferentes fontes de açúcar, buscando compreender e aperfeiçoar o processo de calogênese nessa espécie.

Quanto aos resultados apresentados pelo sequenciamento do RNA, estes foram promissores com elevados valores totais de *reads* (leituras) e além disso, apresentaram altos

índices de qualidade de bases acarretando na manutenção de 88% das leituras obtidas mesmo após o processo de limpeza (Tabela 2). Tudo isso comprovou que o sequenciamento em geral obteve êxito e que as sequências obtidas podem ser depositadas nos bancos de dados *online* de nucleotídeos e serem utilizadas em outros trabalhos, com novos objetivos. Ademais, as leituras obtidas com alta qualidade geraram uma porcentagem elevada de alinhamentos contra o genoma de referência, permitindo assim, o mapeamento e a determinação assertiva de onde as leituras se originaram (Tabela 3).

Foi interessante observar nas primeiras análises gráficas (PCA e dendograma) que a disposição dos dados apresentados por cada replicata já seguia um nível de similaridade e agrupamento esperado em que as amostras A e T formaram grupos distintos, porém mais próximos e as amostras M ficam unidas, porém à parte. Essa similaridade ou disparidade entre os tratamentos foi melhor compreendida quando os padrões de expressão dos genes foram analisados.

Analizando o número total de genes identificados, pelas duas ferramentas computacionais, os valores de 12.387 e 13.653 citados foram maiores que os encontrados em outros trabalhos com *Eucalyptus spp.*, considerando obviamente as peculiaridades intrínsecas de cada publicação (Figura 14). Por exemplo, Hefer et al. (2015) conseguiram identificar um total 10.890 genes partir de hastes e folhas de *E. grandis*, enquanto Thumma et al. (2012) encontraram 7.420 transcritos provenientes do câmbio de *E. nitens*.

Os resultados iniciais mostraram que o maior número de genes diferencialmente expressos (GDEs) totais e exclusivos foram constatados na comparação entre as células meristemáticas e os elementos traqueais, fato que pode ser compreendido pelas grandes diferenças morfológicas e fisiológicas apresentadas por esses dois tipos celulares, sendo essa dissemelhança altamente correlacionada à peculiar regulação gênica constatada em cada tipo celular. Em contrapartida, a comparação entre as células alongadas e os elementos traqueais apresentou um número mais reduzido de GDEs, comprovando que o padrão de expressão dos genes nesses tipos celulares foi um tanto similar, porém não idêntico.

Esse tipo de estratégia de realizar estudos comparativos entre diferentes idades, partes ou estádios em arbóreas para compreender a regulação gênica envolvida no desenvolvimento do xilema é muito explorada. A exemplo disso, Ouyang et al. (2016) compararam segmentos apical, médio e basal do caule da árvore *Neolamarckia cadamba* para compreender seu processo de xilogênese e o trabalho de Zhang et al. (2016) comparou

árvores com diferentes idades (7, 15 e 21 anos) de *Cunninghamia lanceolata* (Lamb.) Hook para estudar o desenvolvimento do seu xilema imaturo. E como destaque, há o trabalho de Hefer et al. (2015) que estudou o perfil transcricional durante a formação da parede celular secundária de *Populus trichocarpa* e *Eucalyptus grandis* comparando amostras de folha e caule.

A singularidade no padrão de expressão de cada tratamento foi apresentada pelo *heatmap* construído (Figura 18) em que se observou a variação do nível de expressão dos genes pela intensidade da coloração apresentada, de azul claro a vermelho intenso, simbolizando genes menos expressos (*down-regulated*) a mais expressos (*up-regulated*), nessa ordem. E desta maneira, foi possível obter uma visão global do perfil de expressão apresentada pelos tratamentos, constatando-se um antagonismo na expressão de grupos de genes nos tratamentos M e T, e além disso, grupos de gene que nas células meristemáticas manifestaram-se menos expressos, mas que aumentaram ligeiramente seus níveis de expressão nas células alongadas e mais intensamente nas traqueídes, e notavelmente, esse comportamento foi invertido em outros grupos de genes.

Quando foi realizada a categorização funcional inicial da totalidade dos GDEs foi possível delinear que a maioria dos genes anotados participou predominantemente de processos metabólicos na célula e na membrana celular, relacionados com funções catalíticas e de ligação. Esses primeiros resultados foram mais genéricos e elementares, pois as células precisavam expressar genes relacionados à manutenção das suas atividades vitais, logo a maioria dos resultados estão fortemente relacionados a essas categorias funcionais. Essa abrangência nos resultados da categorização funcional inicial também foi constatada em outros trabalhos que utilizaram essa metodologia (Domingos et al., 2016; Ouyang et al., 2016; Xu et al., 2016).

Quando foi realizada anotação dos GDEs pelo banco de dados do KEGG e posteriormente, foram considerados somente os genes mais significativos ($\log_2 \text{fold change} \leq -1,5$ e $\log_2 \text{fold change} \geq 1,5$) para o estudo, os resultados foram bem mais relevantes. Um resultado significativo foi a rota dos fenilpropanoides (Figuras 20 e 22). Tanto quando foram considerados a totalidade dos genes, em que a rota biossintética dos fenilpropanoides ficou em quinta posição, de um total de 310 rotas identificadas pelo banco de dados; como quando foram considerados os genes mais expressos nos tratamentos T e A, e no caso

particular dessas fases celulares, a importância dessa rota de biossíntese dos monômeros da lignina ficou realçada.

A percepção da existência de sequências codificadoras de enzimas-chaves relacionadas à biossíntese dos precursores da lignina demonstrou como essa rota está intimamente relacionada à deposição da parede celular secundária nas células alongadas, na subsequente formação dos elementos traqueais de eucalipto e no desenvolvimento do xilema. Além disso, a rota dos fenilpropanoides foi demonstrada como fundamental na deposição dos monolignóis antes e logo após a morte celular programada nos elementos traqueais em *A. thaliana* (Smith et al., 2013), sendo essa relevância estendida aos elementos traqueais de eucalipto.

Continuando o estudo funcional dos genes identificados foi realizado o enriquecimento dos termos GO obtidos. O uso dessa ferramenta para o estudo dos GOs representativos se deu pela abrangência dos resultados iniciais. Desta maneira, foi possível refinar as vias metabólicas e interpretar os resultados com mais acurácia. Os GOs representativos para a comparação M vs. T, principalmente dos genes mais expressos no tratamento T ($\log_2 \text{fold change} \geq 1,5$), confirmaram o enriquecimento das vias de componentes da parede celular secundária, bem como, de processos biológicos relacionados a biossíntese de ácido jasmônico, e as rotas de sinalização mediadas pelo próprio ácido jasmônico, pelos brassinosteróides e principalmente pelo etileno (Figura 24).

Yamamoto et al. (2007) já havia comentado sobre a co-regulação de genes relacionados à biossíntese de brassinosteróides durante a diferenciação *in vitro* de células xilemáticas de *Z. elegans*. Assim como, Sehr et al. (2010) mostraram que a adição de ácido jasmônico estimulou o crescimento secundário cambial em brotos de *Arabidopsis* reafirmando a participação dessas moléculas no desenvolvimento das células xilemáticas. Assim a ocorrência de termos GOs relacionados a essas rotas bioquímicas é coerente e valida a importância dessas moléculas para o desenvolvimento vascular em eucalipto.

Além dessa constatação, pode-se considerar que o ponto alto foi a participação do etileno como estimulador na diferenciação dos elementos traqueais de eucalipto. Essa dependência e efeito pode estar relacionada a atuação da ACC oxidase (*1-aminocyclopropane-1-carboxylic acid*), a qual sintetizou o etileno no apoplasto das células xilemáticas *in vitro* de *Z. elegans* (Pesquet and Tuominen, 2011) e, por conseguinte, essa atuação pode ser estendida aos elementos traqueais de eucalipto. Adicionalmente, os

fatores responsivos ao etileno (ERFs) também mereceram um destaque atuando como fatores de transcrição e ativando a expressão de genes codificadores de enzimas fundamentais nas rotas de biossíntese dos componentes da parede celular secundária (celulose, hemiceluloses e lignina), logo os ERFs conectam a ação do etileno com o desenvolvimento da madeira.

A relevância de termos GOs relacionados à morte celular e apoptose nos elementos traqueais, foi um fato condizente, pois esse tipo celular realmente está sofrendo modificações bioquímicas e estruturais para se tornar uma célula oca e de condutância hídrica. Entretanto, esses GOs também foram relacionados inesperadamente às células meristemáticas, mostrando que mesmo nas fases iniciais do desenvolvimento celular há uma regulação gênica fina que redireciona as células totipotentes a se diferenciarem em células alongadas, que em seguida, como explicado por Pesquet et al. (2013), sofrerão uma deposição da parede secundária celulósica e lignificação dos elementos traqueais.

Resumindo, os resultados das redes de termos GOs significativos expuseram a participação de personagens importantes para o enredo do desenvolvimento secundário vegetal. As rotas apresentadas fizeram sentido no contexto considerado, com a proeminência para a biogênese e organização da parede celular. Além disso, foi reafirmada a ativação de genes relacionados às rotas de biossíntese da celulose, do ácido jasmônico, das xilanas e dos fenilpropanoides, assim como, a sinalização mediada pelo etileno e a atividade catalítica das peroxidases, tudo isso trabalhando em conjunto para possibilitar a completa formação e funcionalidade dos elementos traqueais.

5.3. Regulação transcricional da biossíntese da parede celular secundária

A constatação da participação ativa de 31 fatores de transcrição (FTs) esquematizados na Tabela 4. mostrou a importância dos mesmos para a ativação de vários genes codificadores de enzimas essenciais para a biossíntese dos componentes da parede celular secundária. Alguns deles, como as famílias MYB e NAC, já foram bem descritos na literatura com várias espécies, inclusive com eucalipto (Hussey et al., 2011, 2015; Liu et al., 2015; Nakano et al., 2015; Soler et al., 2015; Zhong et al., 2008; Zhou et al., 2009), enquanto outros FTs foram identificados e correlacionados a parede celular secundária mais

recentemente, como no trabalho de Taylor-Teeple et al. (2015) que verificou a ocorrência de 623 interações entre 45 promotores e 208 fatores de transcrição construindo uma rede regulatória da síntese da parede celular secundária em *Arabidopsis* e que foi utilizada como um norteador para o estudo com o eucalipto.

Os fatores de transcrição MYB63 e MYB20 foram identificados e mais expressos nos elementos traqueais de eucalipto do que nas suas células meristemáticas, fato coerente pelos relatos bibliográficos da participação da família MYB na biossíntese de lignina e a formação de parede celular secundária (Zhong et al., 2008; Zhou et al., 2009). Sendo essa função altamente conservada como foi demonstrado quando se estudou subgrupos R2R3-MYB de *Eucalyptus grandis* comparando-os filogeneticamente com *Arabidopsis thaliana*, *Oryza sativa*, *Populus trichocarpa* e *Vitis vinífera* e observou-se que os subgrupos que continham genes-chave que regulavam a biossíntese de lignina e formação de parede celular secundária eram conservados em todas as espécies analisadas (Soler et al., 2015).

O fator de transcrição MYB63 foi relatado como especificamente expresso em fibras e vasos submetidos ao espessamento da parede secundária atuando como um importante regulador da biossíntese de lignina em várias espécies, inclusive em eucalipto (Hefer et al., 2015; Ouyang et al., 2016; Zhang et al., 2016; Zhou et al., 2009), além de acreditar-se que os genes *LAC4* e *LAC17* são regulados por esse fator de transcrição (Zhou et al., 2009). Taylor-Teeple et al. (2015) ainda demonstrou que o MYB63 regula a expressão do gene *CESA8* e de quatro genes da rota biossintética da lignina (*C4H*, *4CL1*, *HCT* e *CCoAOMT1*), comprovando a relevância da sua participação na formação dos elementos traqueais.

Da mesma maneira, o MYB20 foi identificado como um dos fatores de transcrição envolvidos na regulação da biossíntese de parede celular secundária em *Arabidopsis* por Zhong et al. (2008). E além disso, esse fator de transcrição MYB20, juntamente com o MYB63 foram mais expressos durante a diferenciação *in vitro* de elementos de vaso do xilema de *Arabidopsis* e mais ativos nas células vasculares já formadas (Nakano et al., 2010), o que pode explicar o fato desses MYBs terem sido *up-regulated* exclusivamente nos elementos traqueais, na comparação M vs. T, mas não terem sido constatados GDEs quando foi considerado as comparações M vs. A. e A vs. T.

Diferentemente, o MYB56 que segundo Taylor-Teeple et al. (2015) regula a expressão o gene *C4H* não foi relatado na literatura com essa função biológica, mas sim correlacionado ao controle do tamanho de sementes de *A. thaliana* (Zhang et al., 2013),

sendo que os autores concluíram na publicação que esse fator de transcrição poderia ter sua função estendida, participando também no desenvolvimento da parede celular das sementes. E em outro trabalho mais recente, foi descrito como participante no controle do processo de florescimento em *A. thaliana* (Chen et al., 2015), os autores comentaram que a degradação do MYB56 pela via da ubiquitina-proteassoma ativaria genes que iniciariam o florescimento, logo esse fator de transcrição atuaria como um regulador negativo do florescimento.

Portanto, o MYB56 não foi descrito na literatura correlacionado diretamente às células xilemáticas ou ao desenvolvimento da parede celular secundária, entretanto foi ligado a processos importante do desenvolvimento da planta, o que pode explicar a sua ativação e expressão no caso das células meristemáticas, sendo que não há relatos desse fator de transcrição específico no caso de eucalipto.

Com relação a família NAC, esta engloba FTs que são bem relatados na literatura, inclusive com publicações que estudaram a história evolutiva dessa família NAC em plantas (Zhu et al., 2012), e especificamente em eucalipto (Hussey et al., 2015). No estudo presente foi possível identificar a participação dos FTs NAC102, NAC073 (SND2) e NAC010 (SND3), sendo que os dois primeiros se mostraram sutilmente mais expressos nos elementos traqueais e o último mais expresso nas células alongadas, demonstrando que nessas células a maquinaria transcricional da deposição da parede celular secundária estava ativa. Essa participação da família NAC em estudo com células também foi observado por Demura et al. (2002) que foram os primeiros a relatar que a expressão do fator de transcrição Z567 do domínio NAC havia sido aumentada durante transdiferenciação *in vitro* de elementos traqueais de *Zinnia elegans*.

O fator NAC102 que de acordo com Taylor-Teeples et al. (2015) estaria relacionado a expressão dos genes da rota da lignina, não foi descrito em outros trabalhos, em compensação, o NAC073 (SND2) e o NAC010 (SND3) foram observados em vários estudos transcricionais envolvendo o desenvolvimento vascular de várias espécies, como *Arabidopsis* e *Populus* (Grant et al., 2010), *Medicago truncatula* (Wang et al., 2016), *Raphanus sativus* (Yu et al., 2016), e em *E. grandis* (Hefer et al., 2015; Hussey et al., 2011). Inclusive, Hussey et al. (2011) já postulavam que a super-expressão do SND2 teria o potencial de aumentar o desenvolvimento de fibras em árvores de eucalipto.

Não foram encontrados nos resultados obtidos dois fatores de transcrição da família NAC, o VND6 e o VND7 que são descritos como reguladores diretos de genes que governam a morte celular programada e a formação de parede secundária durante a diferenciação do xilema (Ohashi-Ito et al., 2010; Yamaguchi et al., 2011), atuando através da ligação à região promotora do chamado TERE (Elemento cis Regulador do Elemento Traqueal), responsável pela expressão coordenada de genes relacionados à diferenciação de elementos traqueais (Pyo et al., 2007).

É importante salientar que os FTs MYB e NAC são caracterizados por serem codificados por muitos genes, com relatos segundo a revisão de Nakano et al. (2015) de mais de 126 genes em *arabidopsis*, 109 em arroz, 141 em eucalipto e 192 em álamo no caso da família MYB, e com mais de 100 membros envolvidos em *arabidopsis*, arroz, soja, álamo e eucalipto, considerando a família NAC. Por essa razão, no presente trabalho esperava-se encontrar mais FTs pertencentes a essas famílias do que o observado, fato que pode estar relacionado ao tipo de material vegetal que foi utilizado e a metodologia utilizada nas análises de bioinformática.

Uma família de fatores de transcrição que foi comentada nos resultados, e que apresentou membros que foram diferencialmente expressos nas comparações realizadas no trabalho foi a WRKY. Há registros na literatura que essa família é participante da formação da parede secundária (Guillaumie et al., 2010; Hefer et al., 2015; Salazar et al., 2011; Wang et al., 2016), recentemente foi demonstrado que um novo fator WRKY13 regula positivamente a biossíntese de lignina em caules de *Arabidopsis* se ligando diretamente ao promotor *NST2* (Li et al., 2015b).

Considerando espécies arbóreas, foi realizado um estudo com álamo indicando que PtrWRKY19 exercia uma função ortóloga ao AtWRKY12 regulando negativamente o desenvolvimento secundário em células da medula de *Populus trichocarpa* (Wang et al., 2010; Yang et al., 2016). Com híbridos de eucalipto (*Eucalyptus urophylla* × *Eucalyptus grandis*) detectou-se a presença de 10 FTs WRKY, e a maioria não apresentou diferença de nível de expressão entre genótipos com teores de lignina diferentes (Shinya et al., 2016). Logo pela literatura existente, ainda não está evidente se os fatores de transcrição WRKY desempenham papéis igualmente críticos no desenvolvimento da PCS em espécies arbóreas.

Com relação a outras classes de fatores de transcrição, os resultados foram interessantes e novos FTs foram encontrados. Uma classe em evidência foi a dos elementos

responsivos ao etileno (ERFs), que na literatura estão associados a complexos sistemas de sinalização relacionados ao desenvolvimento vegetal e a resposta a estresses bióticos e abióticos (Alonso and Stepanova, 2004; Nakano et al., 2006; Stepanova and Alonso, 2009). No caso em estudo, os ERF2, ERF3, ERF4 foram identificados como diferencialmente expressos majoritariamente nos elementos traqueais, mesmo sem adição externa de precursores de etileno no meio de cultura utilizado.

Para compreender a ativação desses FTs, deve-se considerar que o etileno foi descrito como um estimulador endógeno do crescimento cambial durante a formação da madeira em *Populus* (Love et al., 2009), além disso, a expressão heteróloga do SHINE/WAX INDUCER (SHN/WIN), pertencente à família ERF, em arroz alterou drasticamente a composição da celulose, da hemicelulose e da lignina sem afetar o crescimento global da planta (Ambavaram et al., 2011). Esses trabalhos conjuntamente com de Vahala et al. (2013) que identificaram ERFs no híbrido *Populus tremula x tremuloides* que quando superexpressados modificaram a formação da madeira, suportam a ideia que alguns ERFs tem motivos de ligação que regulam direta ou indiretamente a biossíntese da parede celular secundária.

Com isso, a descoberta desses FTs nas células estudadas é um dado interessante, pois esses fatores de transcrição AP2/ERF conjuntamente com outros fatores já comentados, MYB e NAC foram identificados em tecidos do xilema de híbridos de eucalipto com diferentes conteúdos de lignina, e apresentaram um maior nível de expressão no híbrido que continha maior teor de lignina (Shinya et al., 2016).

Os outros fatores de transcrição que foram identificados e correlacionados aos genes de *E. grandis*, foram descritos em artigos que consideraram a rede de regulação da parede celular secundária de *Arabidopsis* (Kumar et al., 2016; Taylor-Teeple et al., 2015). Como por exemplo, o IRX11 pertencente à família dos IRX (*irregular xylem*), a qual foi relacionada na literatura à via de biossíntese das xilanas (Chiniquy et al., 2013; Shinya et al., 2016). Essa relação do IRX com as xilanas foi observada no caso das células de eucalipto, o IRX11 foi mais expresso nas células alongadas, interagindo com o promotor do gene GUX2 que atua no final da rota de biossíntese das xilanas.

Outro fator de transcrição o JAM2 foi detectado como similarmente mais expresso nas células alongadas e elementos traqueais, mas na literatura a expressão desse gene está ligada a uma regulação negativa de receptores de jasmonato em *Arabidopsis* (Sasaki-

Sekimoto et al., 2013), contudo, a aplicação de ácido jasmônico (JA) atuou positivamente no crescimento secundário de brotos de *Arabidopsis* (Sehr et al., 2010), com isso, o resultado apresentado foi contraditório no sentido da atuação desse fator de transcrição na xilogênese.

Os outros FTs DRED26, AIL6, GRF1, CBF4, PAT1, CAMTA2, AUX2-11 e RAP2.12 foram descritos na literatura com outras funções biológicas, como exemplo, o DREB26 participante da família APETALA2 (AP2) foi relacionado a papéis importantes no crescimento e desenvolvimento das plantas e em respostas de defesa, e particularmente a superexpressão do DREB26 resultou em plantas deformadas de *Arabidopsis* (Krishnaswamy et al., 2011).

Compilando os resultados observados nas diferentes fases celulares estudadas, foi possível compreender a expressão e a participação da maioria dos FTs encontrados na regulação transcricional da formação da parede celular secundária, principalmente os NAC, MYB, WRKY e ERF. Essa atuação como elementos-chaves foi observada pela expressão diferencial desses FTs, que se apresentaram mais expressos nas células em que a rota de biossíntese dos componentes da parede celular secundária faria sentido biológico, como nas alongadas e principalmente nos elementos traqueais de eucalipto. Outros FTs não foram relacionados ao desenvolvimento xilemático e à formação dos elementos traqueais, mas a identificação desses fatores de transcrição pode possibilitar novas vertentes de estudo e além disso, análises mais aprofundadas nos dados obtidos para a busca de novos fatores de transcrição que não foram encontrados até o momento.

5.4. GDEs envolvidos na xilogênese de *Eucalyptus grandis*

O estudo do desenvolvimento do xilema, e da formação da parede celular secundária é um assunto bem debatido na literatura, proporcionando cada vez mais sua compreensão e manipulação visando a seleção e melhoramento na produção de madeira. Com isso, o estudo dos genes mais expressos nas células diferenciadas de eucalipto transpassa o estudo básico, para atingir objetivos aplicados.

Sendo assim, observando a rota de biossíntese da celulose muitos genes apresentaram diferença nos seus níveis de expressão, sendo *up-regulated* nas células em processo de diferenciação como nas células alongadas e nos elementos traqueais, em

oposição ao observado nas meristemáticas (Figura 28). Por exemplo, os genes do início da rota *CINV2* (Eucgr. F02588), *HXK3* (Eucgr.F01647) e principalmente o *SUSY* ou *SUS4* (Eucgr.C00769) foram mais expressos nas células alongadas e nos elementos traqueais, e além disso, os transcritos dos genes *CINV2* e o *SUS4* foram abundantes. Ainda nessa etapa da rota, é interessante observar que o *SUS3* (Eucgr.H01094) foi um dos genes com maior FPKM, sendo altamente abundante nas três fases celulares, principalmente as células meristemáticas, demonstrando a importância do *SUSY* na síntese da parede celular, sendo esse gene considerado a principal fonte de UDP-glicose para a síntese de celulose na diferenciação do xilema de *E. grandis* (Gallo de Carvalho et al., 2008).

De maneira similar, o grupo de genes *CESA* também foram destaque com os genes homólogos ao *CESA1* de *Arabidopsis* (Eucgr. H00939, Eucgr.J01639) sendo mais expressos nas células meristemáticas, isso pode estar relacionado ao fato desse gene participar da formação da parede celular primária em *Arabidopsis* (Ranik and Myburg, 2006), e similarmente também seriam expressos nas células meristemáticas de eucalipto que apresentam somente esse tipo de parede celular. Enquanto os genes *CESA9* (*Cellulose synthase A9*, Eucgr.I00286, Eucgr.F03635) foram mais expressos nas células alongadas que estão em processo de diferenciação com alta síntese de celulose, sendo essa classe ou grupo de genes *CESA* descrita em vários estudos como relacionada à formação da parede celular secundária (Ambavaram et al., 2011; Mckinley et al., 2016; Mendu et al., 2011), além serem funcionalmente conservados em diversas espécies vegetais, inclusive em *Eucalyptus* (Creux et al., 2013).

Partindo para a rota de biossíntese das hemiceluloses, uma grande gama de genes foi identificada e com alta abundância de transcritos, principalmente os *UDP-xylose synthase* *UXS2* e *UXS6*, que foram ligeiramente mais expressos nas células meristemáticas e nos elementos traqueais, respectivamente. Seguindo na via, o gene *irregular xylem* (*IRX*) gene *IRX10* (Eucgr.G01977) foi identificado, e assim como os genes *GATL2*, *GATL3* e *GATL10* foram mais expressos nas fases celulares A e T, em que a parede celular secundária estava em formação e deposição. A expressão do *IRX10* é fundamental para a elongação do esqueleto de xilanas (Brown et al., 2009), e esses três genes *UXS*, *IRX10* e *GATL* apresentaram níveis de expressão alterados em híbridos de eucalipto com teores maiores de lignina (Shinya et al., 2016). Desta maneira, a presença e expressão diferencial desses genes corrobora a regulação gênica existente para a diferenciação dos elementos traqueais.

Alguns genes TBL foram identificados como GDEs, resultado pertinente pois esses genes codificam proteínas da família DUF231, que estão envolvidos na acetilação das xilanas, fato confirmado pela adição de grupos acetil às xilanas sintetizadas *in vitro* pela expressão do gene *TBL29* (Urbanowicz et al., 2014). Os genes *TBL33* (Eucgr.G03092), *TBL29* (Eucgr.J00196), *TBL16* (Eucgr.C03303), *TBL21* (Eucgr.B03360) foram expressos nas células de eucalipto em processo de diferenciação, e também foram identificados por Hefer et al. (2015) em *Populus* e *Eucalyptus*, entretanto somente o *TBL29* foi descrito com mais detalhes na literatura.

O TBR foi sugerido por Bischoff et al. (2010) como relacionado com o estado de esterificação das pectinas, contribuindo assim para a síntese e deposição da celulose na parede secundária, entretanto faltam estudos com estes grupos de genes TBR e TBL sugerindo que as funções bioquímicas exatas desses genes devem ser melhor investigadas (Yuan et al., 2016).

O gene da família DUF579 o *IRX15L* (Eucgr.I01114) foi mais expresso nos elementos traqueais, com maior abundância de transcritos nessa fase celular. A expressão desse gene é essencial para a biossíntese normal das xilanas, pois a mutação desse gene em *Arabidopsis* acarretou na redução do teor de xilanas, com a presença de cadeias mais curtas (Brown et al., 2011). Assim como comentado no caso dos genes da família DUF231, os genes com o domínio conservado DUF579 também merecem atenção e devem ser investigados, pois sua função bioquímica também é pouco clara.

Estudando os genes participantes da rota de biossíntese da lignina, foram identificados genes codificadores das dez enzimas envolvidas na catalisação da biossíntese dos monolignóis (PAL, C4H, 4CL, HCT, C3H, CCoAOMT, F5H, COMT, CCR e CAD) demonstrando que toda a rota de biossíntese foi ativada nas diferentes fases celulares estudadas (Figura 31). É sabido que a biossíntese da lignina via fenilpropanoides é uma rota molecular bem estudada, e como mostrado, com no mínimo onze famílias de genes conhecidas envolvidas no processo. Esse fato, reforçado a disponibilidade do genoma de *E. grandis* (Myburg et al., 2014) possibilitou a realização de uma análise genômica mais abrangente dos genes envolvidos nessa rota nessa espécie.

Comparando os resultados obtidos nesse estudo, com dois trabalhos recentes envolvendo a identificação de genes participantes no desenvolvimento e lignificação do xilema em *Eucalyptus*, foi possível constatar semelhanças e genes inéditos importantes na

rota de biossíntese da lignina. No trabalho de Hefer et al. (2015) utilizando a camada externa do xilema jovens de *E. grandis* foram identificados 24 genes atuantes na rota dos fenilpropanoides, enquanto Carocha et al. (2015) comparou 6 tecidos diferentes, incluindo xilema de *E. grandis*, encontrando 38 genes expressos em tecidos lignificados. Sendo que os genes encontrados nesses dois artigos foram na sua maioria similares.

Com esse conhecimento prévio, nas comparações M vs. A, A vs. T e M vs. T foi possível distinguir 28 genes codificadores das enzimas participantes da rota de biossíntese dos monômeros da lignina, sendo que 8 deles também foram citados por Hefer et al. (2015) e 9 deles foram comuns aos encontrados por Carocha et al. (2015), dentre eles, os genes *PAL3*, *PAL7* e *PAL9*, *C4H1*, *HCT4* e *HCT5*, *C3H4*, *F5H1* e *COMT1*.

Assim, é promissor considerar que 19 genes encontrados não haviam sido mencionados anteriormente na via de biossíntese da lignina em eucalipto, sendo pertencentes as famílias 4CL (*4CL7*, *4CL8*, *4CL9*), COMT (*COMT14*, *COMT35*, *COMT37*), CCR (*CCR5*, *CCR7*, *CCR9*) e CAD (*CAD4*, *CAD5*, *CAD9*, *CAD10*, *CAD12*, *CAD15*, *CAD20*, *CAD21*, *CAD34*, *CAD37*) que foram homólogos a genes de *Arabidopsis*. Sendo visualizado um padrão de expressão interessante, em que maioria dos genes que se apresentaram mais expressos nas células meristemáticas não continuaram sua expressão diferencial nos outros tipos celulares, enquanto aqueles que foram mais expressos nas alongadas continuaram sua expressão nos elementos traqueais.

Outro fato interessante, é que os genes *CCR* e *CAD*, majoritariamente expressos nas células em alongamento e nos elementos traqueais, e em outros estudos foram expressos em tecidos altamente lignificados, como caule secundário e xilema em desenvolvimento, foram também observados nas células meristemáticas. Essa expressão ampla da família CAD foi comentada por Carocha et al. (2015) que constataram que o gene *EgrCAD2* havia sido expresso tanto no xilema como outros tipos de tecidos de eucalipto, como nas folhas jovens e nas flores.

Na rota de biossíntese da lignina novas vias alternativas ainda estão sendo descobertas, como aquela envolvendo a enzima *caffeoyl shikimate esterase* (CSE) que atua sobre o produto do C3H promovendo uma rota alternativa para o cafeoil-CoA (Ruben Vanholme et al., 2013). O gene *EgrCSE* foi encontrado preferencialmente em tecidos xilemáticos e lignificados de *E. grandis* (Carocha et al., 2015; Shinya et al., 2016), entretanto esse gene não foi encontrado neste estudo.

Em contraste com sua biossíntese, o transporte e a polimerização dos monolignóis permanecem pouco compreendidos, há um consenso de que algumas enzimas como C4H, C3H e F5H se localizam no retículo endoplasmático (ER), enquanto todas as outras enzimas biossintéticas de monômeros de lignina, como CCoAOMT, CCR e COMT, são encontradas no citoplasma. Algumas evidências sugerem que pode ocorrer a formação de complexo na superfície externa da ER (Barros et al., 2015), mas claramente quando as unidades de monolignol são sintetizadas, elas precisam ser transportadas do citosol para a parede celular.

Tanto a difusão passiva como o tráfico de vesículas tem sido propostos como um mecanismo, mas a atenção recente tem se concentrado no papel de transportadores específicos, como os transportadores ABC que foram identificados como bons candidatos com base nos dados de expressão com *Arabidopsis* e interações celulares envolvendo a formação dos elementos traqueais (Ehlting et al., 2005; Ménard and Pesquet, 2015; Pesquet et al., 2013). Esse tipo de transportadores foi encontrado nas comparações M vs. T, M vs. A e A vs. T, sendo que os genes Eucgr.C01000 e o Eucgr. J02160 foram descritos como *ABC transporter family protein*.

O gene Eucgr.C01000 foi ligeiramente mais expresso nas células meristemáticas, na comparação M vs. T e o gene Eucgr.J02160 foi intensamente mais expresso nas células alongadas, no caso M vs. A e A vs. T. Este resultado sugere que nas células alongadas, em processo de diferenciação para elementos traqueais, os monolignóis estariam sendo transportados ativamente através da membrana plasmática por esses transportadores ABC.

A identificação desses transportadores ABC nas células de eucalipto é uma descoberta importante, pois há poucos relatos e estudos caracterizando o mecanismo de transporte desses precursores de lignina em arbóreas. E além disso, há relatos divergentes pois no xilema de álamo e cipreste não foi encontrada nenhuma evidência do envolvimento de transportadores ABC, mas identificou-se um tipo de transportador vacuolar (*proton-dependent coniferin transport*) (Tsuyama et al., 2013). Logo, dualidade apresentada nos estudos nesse tema demonstra que pesquisas adicionais são necessárias para fornecer uma evidência definitiva do papel central desses transportadores na biossíntese da lignina.

Quanto à polimerização, lacases e peroxidases têm sido sugeridas como participantes da ativação dos monolignóis. As lacases são codificadas por uma grande família de genes, chegando a 17 genes em *Arabidopsis*, sendo demonstrado que algumas possuem um papel

bem claro na polimerização oxidativa do monolignóis durante o desenvolvimento vascular da planta, em particular para os *LAC4* e *LAC17* (Ruzicka et al., 2015; Zhao et al., 2013).

Especificamente os genes *LAC7* e *LAC14* que foram encontrados nos resultados não são descritos na literatura como relacionados diretamente à ativação, polimerização ou formação da lignina, no trabalho de Turlapati et al. (2011) foi realizada uma caracterização da família LAC, e o *LAC7* foi expresso nos pelos radiculares e o *LAC14* em órgãos reprodutivos de *Arabidopsis*. Desta maneira, é aceito que as lacases participam da formação da lignina, mas existe uma lacuna na identificação dos genes LAC envolvidos na polimerização dos monolignóis nas células xilemáticas de eucalipto, por essa razão, a identificação dos dois genes nesse trabalho é um dado intrigante.

No caso das peroxidases, a existência de uma família numerosa de genes, chegando a 73 genes identificados, pode ser um entrave no estudo e na compreensão da atuação das peroxidases na polimerização da lignina. Esse fato é tão verdadeiro, que não há consenso quanto a dinâmica das lacases e peroxidases no processo, há autores que acreditam que as lacases e peroxidases trabalhem juntas na lignificação dos ETs (Schuetz et al., 2014), enquanto outros autores suportam um modelo em que a polimerização da lignina nos tecidos vasculares é conduzida principalmente por lacases, e não pela ação combinatória de lacases e peroxidases, e que essas últimas atuariam no final da via após ação inicial das lacases (Zhao et al., 2013).

Quando se trata da formação dos elementos traqueais, trabalhos pioneiros já descreviam um pico de atividade de peroxidases, imediatamente antes da lignificação, e durante a atividade de biossíntese da lignina no decorrer da transdiferenciação em *Zinnia* (Fukuda and Komamine, 1982). Os resultados observados em eucalipto mostraram que os genes codificadores das peroxidases apresentaram-se mais expressos nas células alongadas, e não nos elementos traqueais, indicando a possível síntese das peroxidases na fase anterior a formação dos elementos traqueais propriamente ditos, sugerindo que algumas lacases e peroxidases são sintetizadas previamente para serem utilizadas depois durante a lignificação da parede celular. Entretanto, é relevante considerar que outras lacases foram produzidas pelos elementos traqueais, indicando a complexidade desse processo, que ainda é agravada pelo fato que as células vizinhas parenquimáticas podem fornecer monômeros de lignina e espécies reativas de oxigênio (ROS) aos elementos traqueais, possibilitando ainda uma lignificação *post-mortem* (Ménard and Pesquet, 2015; Pesquet et al., 2013). Com isso, pode-

se reconhecer que a trajetória biossintética da lignina se mostrou mais complexa e reticulada do que se pensava inicialmente.

5.5. Engenheirando a deposição da parede celular secundária

A enorme importância econômica da indústria da celulose tem sido uma força motriz para decifrar o processo de deposição da parede celular secundária, em especial a lignificação. E como comentado, a resistência à degradação da lignina é um grande obstáculo no processamento da madeira, durante a fabricação de celulose e papel, pela necessidade de tratamentos químicos agressivos e onerosos. Essa problemática levou a estudos genéticos buscando novos genes candidatos para a modificação da lignina, modificação dos seus monômeros e a regulação da biossíntese das lignoceluloses, através da alteração na expressão dos genes dessa rota. (Baucher et al., 1998; Boerjan et al., 2003; Vanholme et al., 2008; Yang et al., 2013).

Com espécies arbóreas, esses estudos já mostraram resultados interessantes, como Min et al. (2012) que relataram que linhagens de *Populus trichocarpa* foram obtidas pela diminuição da expressão da enzima 4CL, e se mostraram mais passíveis a hidrólise enzimática. A regulação negativa mediada por RNAi da enzima C3H em álamo resultou em plantas com menor teor de lignina e maior liberação da glicose por hidrólise enzimática (Mansfield et al., 2012). E mais recentemente, o sistema CRISPR/Cas9 foi utilizado com sucesso para realizar mutação dos genes *4CL1* e *4CL2* em híbridos de *Populus*, sendo observado a redução do conteúdo de lignina e da razão dos monolignóis S/G nos mutantes *4CL1* (Zhou et al., 2015).

Considerando esse panorama e os resultados apresentados nessa tese com *E. grandis*, as possibilidades de manipulações nos genes participantes na rota da biossíntese da lignina são inúmeras, por exemplo, se o gene *C3H* fosse perturbado, o fluxo para os monômeros G e S é bloqueado e os mutantes depositariam mais lignina H na parede celular. Se o gene *F5H* fosse silenciado, o ramo S seria bloqueado e os mutantes poderiam possuir exclusivamente lignina G. Entretanto, se o gene *F5H* fosse superexpressado, o fluxo G poderia ser redirecionado intensificando a rota para síntese dos monômeros S, de forma que

os transgênicos acumulariam predominantemente lignina S. Desta maneira, tanto o arcabouço como a quantidade da lignina poderia ser alterada.

Essa estratégia foi utilizada com sucesso por Ziebell et al. (2016) que reduziram a expressão dos genes *C3H* e *C4H* via a técnica de RNAi em híbridos de *Eucalyptus urophylla* x *Eucalyptus grandis* obtendo uma maior extração da lignina por NaOH e um conteúdo menor de lignina nas plantas. Nesse estudo os autores buscavam plantas com parede celulares menos rígidas para a utilização nas indústrias de etanol lignocelulósico, mas com certeza essa metodologia poderia ser expandida para reduzir os custos no processamento da biomassa nas indústrias de papel e celulose.

Considerando os resultados de expressão das lacases e peroxidases nas fases celulares estudadas, as lacases (LACs) poderiam ser alvos potenciais de modificações genéticas visando uma maior eficiência no processo de polpação. Esse conceito foi realizado com o mutante *lac5* de *Brachypodium* que apresentou um conteúdo menor de lignina e exibiu uma sacarificação melhorada, nesse trabalho o objetivo era voltado para obtenção de plantas para a produção de biocombustíveis (Wang et al., 2015). E assim como as lacases, as peroxidases também podem ser manipuladas, nesse caso em particular, é necessário um estudo mais aprofundado pelos entraves já discutidos com essa família gênica.

Claramente essas enzimas envolvidas na síntese e polimerização dos monômeros da lignina, como as lacases e as peroxidases identificadas, serão uma ferramenta interessante para a obtenção de plantas de eucalipto com características agronômicas desejadas, além da manipulação genética dos genes participantes da biossíntese de lignina mencionados, com destaque para os genes *CCR* e *CAD* que apresentaram muitos exemplares novos identificados nesse estudo com eucalipto.

6. CONCLUSÕES

A premissa proposta nesse trabalho de investigar os fatores de transcrição e genes envolvidos na xilogênese, delineando um perfil transcricional dessa rota foi atendida. A obtenção *in vitro* dos elementos traqueais possibilitou pela primeira vez descrever essas células em eucalipto. Além disso, a estratégia de utilizar células em diferentes estádios de desenvolvimento para estudar os perfis de expressão gênica de seus transcritomas foi válida e obteve resultados interessantes.

As análises de bioinformática possibilitaram uma visão geral da atividade gênica celular, com isso, fatores de transcrição e genes já descritos ou não na literatura, como relacionados às rotas de biossíntese e deposição parede celular secundária, foram identificados. Com destaque para os genes codificadores de enzimas participantes no final da rota de síntese dos monômeros da lignina, como genes *CCR* e *CAD*, transportadores e genes codificadores de lacases e peroxidases, complementando o conhecimento sobre a formação da madeira em eucalipto. Com isso, novos estudos podem ser conduzidos utilizando os resultados desse trabalho, ou, utilizando os dados transcriptômicos gerados com novos questionamentos.

Assim, holisticamente esse trabalho demonstrou como os avanços mais recentes da biotecnologia incluindo o sequenciamento, manipulação genética e engenharia metabólica dos componentes da parede celular secundária podem beneficiar os pesquisadores na obtenção de uma madeira que atenda a seus objetivos. Como uma madeira mais facilmente processada, economizando energia e substâncias químicas poluentes, e que produz uma polpa com as propriedades aspiradas pela indústria.

REFERÊNCIAS

- ABRAF (2013). *Anuário estatístico ABRAF 2013 ano base 2012*.
- Albersheim, P., Darvill, A., Roberts, K., Sederoff, R., and Staehelin, A. (2010). *Plant Cell Walls*. Garland Science.
- Aloni, R. (1980). Role of auxin and sucrose in the differentiation of sieve and tracheary elements in plant tissue cultures. *Planta* 150, 255–263. doi:10.1007/BF00390835.
- Alonso, J. M., and Stepanova, A. N. (2004). The Ethylene Signaling Pathway. *Science* (80-.). 306, 1513–1515. doi:10.1126/science.1104812.
- Ambavaram, M. M. R., Krishnan, A., Trijatmiko, K. R., and Pereira, A. (2011). Coordinated activation of cellulose and repression of lignin biosynthesis pathways in rice. *Plant Physiol.* 155, 916–31. doi:10.1104/pp.110.168641.
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene Ontology: Tool for The Unification of Biology. *Nat. Genet.* 25, 25–29. doi:10.1038/75556.
- Barros, J., Serk, H., Granlund, I., and Pesquet, E. (2015). The cell biology of lignification in higher plants. *Ann. Bot.* 115, 1053–1074. doi:10.1093/aob/mcv046.
- Barrueto Cid, L. P. (2005). “El cultivo de tejidos,” in *Biotecnología Vegetal*, ed. D. J. Prieto, H., Jordan, M., Barrueto Cid, L.P., Cordeiro, M.C.R., Durzan (Santiago-Chile), 35.
- Baucher, M., Monties, B., Montagu, M. Van, and Boerjan, W. (1998). Biosynthesis and Genetic Engineering of Lignin. *CRC. Crit. Rev. Plant Sci.* 17, 125–197. doi:http://dx.doi.org/10.1080/07352689891304203.
- Biotecnologia, C. de I. sobre (2009). GUIA DO EUCALIPTO: oportunidades para um desenvolvimento sustentável. *Rev. Int. Segur. Soc.* 62, 127–129. doi:10.1111/j.1752-1734.2009.01350.x.
- Bischoff, V., Nita, S., Neumetzler, L., Schindelasch, D., Urbain, A., Eshed, R., et al. (2010). TRICHOME BIREFRINGENCE and its homolog AT5G01360 encode plant-specific DUF231 proteins required for cellulose biosynthesis in Arabidopsis. *Plant Physiol.* 153, 590–602. doi:10.1104/pp.110.153320.
- Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin Biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–546. doi:10.1146/annurev.arplant.54.031902.134938.

- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170.
- BRACELPA (2011). Dados do Setor. *Rev. Abinee* Março 2011, 38–42.
- Brown, D. M., Zhang, Z., Stephens, E., Dupree, P., and Turner, S. R. (2009). Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in *Arabidopsis*. *Plant J.* 57, 732–746. doi:10.1111/j.1365-313X.2008.03729.x.
- Brown, D., Wightman, R., Zhang, Z., Gomez, L. D., Atanassov, I., Bukowski, J. P., et al. (2011). *Arabidopsis* genes IRREGULAR XYLEM (IRX15) and IRX15L encode DUF579-containing proteins that are essential for normal xylan deposition in the secondary cell wall. *Plant J.* 66, 401–413. doi:10.1111/j.1365-313X.2011.04501.x.
- Carocha, V., Hefer, C., Cassan-wang, H., Fevereiro, P., Alexander, A., Paiva, J. A. P., et al. (2015). Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. *New Phytol.*, 1297–1313. doi:10.1111/nph.13313.
- Chen, L., Bernhardt, A., Lee, J., and Hellmann, H. (2015). Identification of *Arabidopsis* MYB56 as a novel substrate for CRL3^{BPM} E3 ligases. *Mol. Plant* 8, 242–250. doi:10.1016/j.molp.2014.10.004.
- Chiniquy, D., Varanasi, P., Oh, T., Harholt, J., Katnelson, J., Singh, S., et al. (2013). Three novel rice genes closely related to the *Arabidopsis* IRX9, IRX9L, and IRX14 genes and their roles in xylan biosynthesis. *Front. Plant Sci.* 4, 83. doi:10.3389/fpls.2013.00083.
- Cline, M., Smoot, M., and Cerami, E. (2007). Integration of biological networks and gene expression data using Cytoscape. *Nat. Protoc.* 2, 2366–2382. doi:10.1038/nprot.2007.324.
- Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., et al. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biol.* 17, 13. doi:10.1186/s13059-016-0881-8.
- Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* 6, 850–861. doi:10.1038/nrm1746.
- Creux, N. M., De Castro, M. H., Ranik, M., Maleka, M. F., and Myburg, A. A. (2013). Diversity and cis-element architecture of the promoter regions of cellulose synthase genes in *Eucalyptus*. *Tree Genet. Genomes* 9, 989–1004. doi:10.1007/s11295-013-0611-2.
- Demura, T. (2014). Tracheary element differentiation. *Plant Biotechnol. Rep.* 8, 17–21. doi:10.1007/s11816-013-0293-0.

- Demura, T., and Fukuda, H. (2007). Transcriptional regulation in wood formation. *Trends Plant Sci.* 12, 64–70. doi:10.1016/j.tplants.2006.12.006.
- Demura, T., Tashiro, G., Horiguchi, G., Kishimoto, N., Kubo, M., Matsuoka, N., et al. (2002). Visualization by comprehensive microarray analysis of gene expression programs during transdifferentiation of mesophyll cells into xylem cells. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15794–15799. doi:10.1073/pnas.232590499.
- Derbyshire, P., Ménard, D., Green, P., Saalbach, G., Buschmann, H., Lloyd, C. W., et al. (2015). Proteomic analysis of microtubule interacting proteins over the course of xylem tracheary element formation in *Arabidopsis*. *Plant Cell* 27, 2709–2726. doi:10.1105/tpc.15.00314.
- Devillard, C., and Walter, C. (2014). Formation of plant tracheary elements in vitro – A review. *New Zeal. J. For. Sci.* 44, 1–14. doi:10.1186/s40490-014-0022-7.
- Didi, V., Jackson, P., and Hejátko, J. (2015). Hormonal regulation of secondary cell wall formation. *J. Exp. Bot.* 66, 5015–5027. doi:10.1093/jxb/erv222.
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi:10.1093/bioinformatics/bts635.
- Domingos, S., Fino, J., Paulo, O. S., Oliveira, C. M., and Goulao, L. F. (2016). Molecular candidates for early-stage flower-to-fruit transition in stenospermocarpic table grape (*Vitis vinifera* L.) inflorescences ascribed by differential transcriptome and metabolome profiles. *Plant Sci.* 244, 40–56. doi:10.1016/j.plantsci.2015.12.009.
- Ehlting, J., Mattheus, N., Aeschliman, D. S., Li, E., Hamberger, B., Cullis, I. F., et al. (2005). Global transcript profiling of primary stems from *Arabidopsis thaliana* identifies candidate genes for missing links in lignin biosynthesis and transcriptional regulators of fiber differentiation. *Plant J.* 42, 618–640. doi:10.1111/j.1365-313X.2005.02403.x.
- Endo, S., Pesquet, E., Yamaguchi, M., Tashiro, G., Sato, M., Toyooka, K., et al. (2009). Identifying new components participating in the secondary cell wall formation of vessel elements in *zinnia* and *Arabidopsis*. *Plant Cell* 21, 1155–1165. doi:10.1105/tpc.108.059154.
- Fagerstedt, K. (2015). The plant cell wall. *J. Integr. Plant Biol.* 57, 328–329. doi:10.1111/jipb.12351.

- Fernando, S. C., Goodger, J. Q. D., Gutierrez, S. S., Johnson, A. A. T., and Woodrow, I. E. (2016). Plant regeneration through indirect organogenesis and genetic transformation of *Eucalyptus polybractea* R.T. Baker. *Ind. Crops Prod.* 86, 73–78. doi:10.1016/j.indcrop.2016.03.025.
- Finotello, F., and Di Camillo, B. (2015). Measuring differential gene expression with RNA-seq: Challenges and strategies for data analysis. *Brief. Funct. Genomics* 14, 130–142. doi:10.1093/bfpg/elu035.
- Fukuda, H. (1996). Xylogenesis: initiation, progression, and cell death. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47, 299–325.
- Fukuda, H., and Komamine, A. (1980a). Direct evidence for cytodifferentiation to tracheary elements without intervening mitosis in a culture of single cells isolated from the mesophyll of *Zinnia elegans*. *Plant Physiol.* 65, 61–64. doi:10.1104/pp.65.1.61.
- Fukuda, H., and Komamine, A. (1980b). Establishment of an experimental system for the study of tracheary element differentiation from single cells isolated from the mesophyll of *Zinnia elegans*. *Plant Physiol.* 65, 57–60. doi:10.1104/pp.65.1.57.
- Fukuda, H., and Komamine, A. (1982). Lignin synthesis and its related enzymes as markers of tracheary-element differentiation in single cells isolated from the mesophyll of *Zinnia elegans*. *Planta* 155, 423–430.
- Gallo de Carvalho, M. C. D. C., Caldas, D. G. G., Carneiro, R. T., Moon, D. H., Salvatierra, G. R., Franceschini, L. M., et al. (2008). SAGE transcript profiling of the juvenile cambial region of *Eucalyptus grandis*. *Tree Physiol.* 28, 905–919.
- Grant, E. H., Fujino, T., Beers, E. P., and Brunner, A. M. (2010). Characterization of NAC domain transcription factors implicated in control of vascular cell differentiation in *Arabidopsis* and *Populus*. *Planta* 232, 337–352. doi:10.1007/s00425-010-1181-2.
- Guillaumie, S., Mzid, R., Méchin, V., Léon, C., Hichri, I., Destrac-Irvine, A., et al. (2010). The grapevine transcription factor WRKY2 influences the lignin pathway and xylem development in tobacco. *Plant Mol. Biol.* 72, 215–234. doi:10.1007/s11103-009-9563-1.
- Hajari, E., Watt, M. P., Mycock, D. J., and McAlister, B. (2006). Plant regeneration from induced callus of improved *Eucalyptus* clones. *South African J. Bot.* 72, 195–201. doi:10.1016/j.sajb.2005.07.003.
- Harakava, R. (2005). Genes encoding enzymes of the lignin biosynthesis pathway in *Eucalyptus*. *Genet. Mol. Biol.* 28, 601–607. doi:10.1590/S1415-47572005000400015.

- Hefer, C. A., Mizrachi, E., Myburg, A. A., Douglas, C. J., and Mansfield, S. D. (2015). Comparative interrogation of the developing xylem transcriptomes of two wood-forming species: *Populus trichocarpa* and *Eucalyptus grandis*. *New Phytol.* 206, 1391–1405. doi:10.1111/nph.13277.
- Hussey, S. G., Mizrachi, E., Spokevicius, A. V., Bossinger, G., Berger, D. K., and Myburg, A. A. (2011). SND2, a NAC transcription factor gene, regulates genes involved in secondary cell wall development in *Arabidopsis* fibres and increases fibre cell area in *Eucalyptus*. *BMC Plant Biol.* 11, 173. doi:10.1186/1471-2229-11-173.
- Hussey, S. G., Sa??di, M. N., Hefer, C. A., Myburg, A. A., and Grima-Pettenati, J. (2015). Structural, evolutionary and functional analysis of the NAC domain protein family in *Eucalyptus*. *New Phytol.* 206, 1337–1350. doi:10.1111/nph.13139.
- IBÁ (2015). - Indústria brasileira de Árvores. *Anu. estatístico da IBA ano base 2014. 2015*, 100. doi:10.1007/s13398-014-0173-7.2.
- IBGE (2014). Produção da Extração Vegetal e da Silvicultura. 29, 1–56. doi:ISSN 0101-4234.
- Kákošová, A., Digonnet, C., Goffner, D., and Lišková, D. (2013). Galactoglucomannan oligosaccharides are assumed to affect tracheary element formation via interaction with auxin in *Zinnia* xylogenic cell culture. *Plant Cell Rep.* 32, 479–487. doi:10.1007/s00299-012-1379-9.
- Keegstra, K. (2010). Plant Cell Walls. *Am. Soc. Plant Biol.* 154, 483–486. doi:10.1104/pp.110.161240.
- Khatrri, P., and Drăghici, S. (2005). Ontological analysis of gene expression data: Current tools, limitations, and open problems. *Bioinformatics* 21, 3587–3595. doi:10.1093/bioinformatics/bti565.
- Korpelainen, E., Tuimala, J., Somervuo, P., Huss, M., and Wong, G. (2015). *RNA-seq Data Analysis*. Chapman and Hall/CRC.
- Krishnaswamy, S., Verma, S., Rahman, M. H., and Kav, N. N. V (2011). Functional characterization of four APETALA2-family genes (RAP2.6, RAP2.6L, DREB19 and DREB26) in *Arabidopsis*. *Plant Mol. Biol.* 75, 107–127. doi:10.1007/s11103-010-9711-7.
- Kumar, M., Campbell, L., and Turner, S. (2016). Secondary cell walls: Biosynthesis and manipulation. *J. Exp. Bot.* 67, 515–531. doi:10.1093/jxb/erv533.

- Lacayo, C. I., Malkin, A. J., Holman, H.-Y. N., Chen, L., Ding, S.-Y., Hwang, M. S., et al. (2010). Imaging cell wall architecture in single *Zinnia elegans* tracheary elements. *Plant Physiol.* 154, 121–133. doi:10.1104/pp.110.155242.
- Lerouxel, O., Cavalier, D. M., Liepman, A. H., and Keegstra, K. (2006). Biosynthesis of plant cell wall polysaccharides - a complex process. *Curr. Opin. Plant Biol.* 9, 621–630. doi:10.1016/j.pbi.2006.09.009.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi:10.1093/bioinformatics/btp352.
- Li, J., Zhang, C., Hu, H., and Chai, X. S. (2016a). Effect of shortening kraft pulping integrated with extended oxygen delignification on biorefinery process performance of eucalyptus. *Bioresour. Technol.* 202, 119–124. doi:10.1016/j.biortech.2015.11.074.
- Li, L. M., Ouyang, L. J., and Gan, S. M. (2015a). Towards an efficient regeneration protocol for eucalyptus urophylla. *J. Trop. For. Sci.* 27, 289–297.
- Li, S., Bashline, L., Zheng, Y., Xin, X., Huang, S., Kong, Z., et al. (2016b). Cellulose synthase complexes act in a concerted fashion to synthesize highly aggregated cellulose in secondary cell walls of plants. *Proc. Natl. Acad. Sci.* 113, 11348–11353. doi:10.1073/pnas.1613273113.
- Li, W., Tian, Z., and Yu, D. (2015b). WRKY13 acts in stem development in *Arabidopsis thaliana*. *Plant Sci.* 236, 205–213. doi:10.1016/j.plantsci.2015.04.004.
- Li, Z., Fernie, A. R., and Persson, S. (2016c). Transition of primary to secondary cell wall synthesis. *Sci. Bull.* 61, 838–846. doi:10.1007/s11434-016-1061-7.
- Liao, Y., Smyth, G. K., and Shi, W. (2013). The Subread aligner: Fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* 41. doi:10.1093/nar/gkt214.
- Liu, J., Osbourn, A., and Ma, P. (2015). MYB Transcription Factors as Regulators of Phenylpropanoid Metabolism in Plants. *Mol. Plant* 8, 689–708. doi:10.1016/j.molp.2015.03.012.
- Love, J., Björklund, S., Vahala, J., Hertzberg, M., Kangasjärvi, J., and Sundberg, B. (2009). Ethylene is an endogenous stimulator of cell division in the cambial meristem of *Populus*. *Proc. Natl. Acad. Sci. U. S. A.* 106, 5984–5989. doi:10.1073/pnas.0811660106.

- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi:10.1186/s13059-014-0550-8.
- Ma, Q., Zhou, W., and Zhang, P. (2015). Transition from somatic embryo to friable embryogenic callus in cassava: dynamic changes in cellular structure, physiological status, and gene expression profiles. *Front Plant Sci* 6, 1–14. doi:10.3389/fpls.2015.00824.
- Mandrou, E., Hein, P. R. G., Villar, E., Vigneron, P., Plomion, C., and Gion, J. M. (2012). A candidate gene for lignin composition in Eucalyptus: Cinnamoyl-CoA reductase (CCR). *Tree Genet. Genomes* 8, 353–364. doi:10.1007/s11295-011-0446-7.
- Mansfield, S. D., Kang, K. Y., and Chapple, C. (2012). Designed for deconstruction - poplar trees altered in cell wall lignification improve the efficacy of bioethanol production. *New Phytol.* 194, 91–101. doi:10.1111/j.1469-8137.2011.04031.x.
- Mansoori, N., Timmers, J., Desprez, T., Kamei, C. L. A., Dees, D. C. T., Vincken, J. P., et al. (2014). KORRIGAN1 interacts specifically with integral components of the Cellulose synthase machinery. *PLoS One* 9, 4–11. doi:10.1371/journal.pone.0112387.
- McCann, M. C. (1997). Tracheary element formation: Building up to a dead end. *Trends Plant Sci.* 2, 333–338. doi:10.1016/S1360-1385(97)84620-1.
- Mckinley, B., Rooney, W., Wilkerson, C., and Mullet, J. (2016). Dynamics of biomass partitioning, stem gene expression, cell wall biosynthesis, and sucrose accumulation during development of *Sorghum bicolor*. *Plant J.* 662–680. doi:10.1111/tpj.13269.
- Ménard, D., and Pesquet, E. (2015). Cellular interactions during tracheary elements formation and function. *Curr. Opin. Plant Biol.* 23, 109–115. doi:10.1016/j.pbi.2014.12.001.
- Mendu, V., Stork, J., Harris, D., and DeBolt, S. (2011). Cellulose synthesis in two secondary cell wall processes in a single cell type. *Plant Signal. Behav* 6, 1638–43. doi:10.4161/psb.6.11.17709.
- Meng, Q., Fu, S., Li, B., and Lucia, L. A. (2015). The impact of xylanase and hot acid pretreatment on HexAs in Eucalyptus kraft pulp bleaching. *J. Wood Chem. Technol.* 35, 239–250. doi:10.1080/02773813.2014.902967.

- Min, D., Li, Q., Jameel, H., Chiang, V., and Chang, H. M. (2012). The cellulase-mediated saccharification on wood derived from transgenic low-lignin lines of black cottonwood (*Populus trichocarpa*). *Appl. Biochem. Biotechnol.* 168, 947–955. doi:10.1007/s12010-012-9833-2.
- Möller, R., Ball, R. D., Henderson, A. R., Modzel, G., and Find, J. (2006a). Effect of light and activated charcoal on tracheary element differentiation in callus cultures of *Pinus radiata* D. Don. *Plant Cell. Tissue Organ Cult.* 85, 161–171. doi:10.1007/s11240-005-9065-z.
- Möller, R., Koch, G., Nanayakkara, B., and Schmitt, U. (2006b). Lignification in cell cultures of *Pinus radiata*: activities of enzymes and lignin topochemistry. *Tree Physiol.* 26, 201–210. doi:10.1093/treephys/26.2.201.
- Möller, R., McDonald, A. G., Walter, C., and Harris, P. J. (2003). Cell differentiation, secondary cell-wall formation and transformation of callus tissue of *Pinus radiata* D. Don. *Planta* 217, 736–747. doi:10.1007/s00425-003-1053-0.
- Mora, A. L., and Garcia, C. H. (2000). *A cultura do Eucalipto no Brasil*. São Paulo: Sociedade Brasileira de Silvicultura,.
- Mueller, O., and Schroeder, A. (2004). RNA Integrity Number (RIN) – Standardization of RNA Quality Control Application. *Nano*, 1–8.
- Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., et al. (2014). The genome of *Eucalyptus grandis*. *Nature* 510, 356–362. doi:10.1038/nature13308.
- Nakano, T., Suzuki, K., Fujimura, T., and Shinshi, H. (2006). Genome-Wide Analysis of the ERF Gene Family. *Plant Physiol.* 140, 411–432. doi:10.1104/pp.105.073783.
- Nakano, Y., Nishikubo, N., Goué, N., Ohtani, M., Yamaguchi, M., Katayama, Y., et al. (2010). MYB transcription factors orchestrating the developmental program of xylem vessels in *Arabidopsis* roots. *Plant Biotechnol.* 27, 267–272. doi:10.5511/plantbiotechnology.27.267.
- Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N. A., and Ohtani, M. (2015). NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front. Plant Sci.* 6, 288. doi:10.3389/fpls.2015.00288.

- Nakhooda, M., and Mandiri, E. (2016). Using synergistic exogenous phytohormones to enhance somatic embryogenesis from leaf explants of a *Eucalyptus grandis* clone. *South. For. a J. For. Sci.* 78, 73–80. doi:10.2989/20702620.2015.1108620.
- Oberschelp, G. P. J., Gonçalves, A. N., Meneghetti, E. C., Graner, E. M., and de Almeida, M. (2015). *Eucalyptus dunnii* Maiden plant regeneration via shoot organogenesis on a new basal medium based on the mineral composition of young stump shoots. *Vitr. Cell. Dev. Biol. - Plant* 51, 626–636. doi:10.1007/s11627-015-9715-8.
- Oda, Y., and Fukuda, H. (2012). Secondary cell wall patterning during xylem differentiation. *Curr. Opin. Plant Biol.* 15, 38–44. doi:10.1016/j.pbi.2011.10.005.
- Oda, Y., and Fukuda, H. (2013). The dynamic interplay of plasma membrane domains and cortical microtubules in secondary cell wall patterning. *Front. Plant Sci.* 4, 511. doi:10.3389/fpls.2013.00511.
- Oda, Y., Mimura, T., and Hasezawa, S. (2005). Regulation of secondary cell wall development by cortical microtubules during tracheary element differentiation in Arabidopsis cell suspensions. *Plant Physiol.* 137, 1027–1036. doi:10.1104/pp.104.052613.
- Ohashi-Ito, K., Oda, Y., and Fukuda, H. (2010). Arabidopsis VASCULAR-RELATED NAC-DOMAIN6 directly regulates the genes that govern programmed cell death and secondary wall formation during xylem differentiation. *Plant Cell* 22, 3461–73. doi:10.1105/tpc.110.075036.
- Oshlack, A., Robinson, M. D., and Young, M. D. (2010). From RNA-seq reads to differential expression results. *Genome Biol.* 11, 220. doi:10.1186/gb-2010-11-12-220.
- Ouyang, K., Li, J., Zhao, X., Que, Q., Li, P., Huang, H., et al. (2016). Transcriptomic analysis of multipurpose timber yielding tree *Neolamarckia cadamba* during xylogenesis using RNA-Seq. *PLoS One* 11, 1–22. doi:10.1371/journal.pone.0159407.
- Ozsolak, F., and Milos, P. M. (2011). RNA sequencing: advances, challenges and opportunities. *Nat. Rev. Genet.* 12, 87–98. doi:10.1038/nrg2934.
- Paiva, J. a P., Prat, E., Vautrin, S., Santos, M. D., San-Clemente, H., Brommonschenkel, S., et al. (2011). Advancing Eucalyptus genomics: identification and sequencing of lignin biosynthesis genes from deep-coverage BAC libraries. *BMC Genomics* 12, 137. doi:10.1186/1471-2164-12-137.

- Pesquet, E., Korolev, A. V., Calder, G., and Lloyd, C. W. (2010). The microtubule-associated protein AtMAP70-5 regulates secondary wall patterning in Arabidopsis wood cells. *Curr. Biol.* 20, 744–749. doi:10.1016/j.cub.2010.02.057.
- Pesquet, E., and Tuominen, H. (2011). Ethylene stimulates tracheary element differentiation in *Zinnia elegans* cell cultures. *New Phytol.* 190, 138–149. doi:10.1111/j.1469-8137.2010.03600.x.
- Pesquet, E., Zhang, B., Gorzsas, A., Puhakainen, T., Serk, H., Escamez, S., et al. (2013). Non-cell-autonomous postmortem lignification of tracheary elements in *Zinnia elegans*. *Plant Cell* 25, 1314–28. doi:10.1105/tpc.113.110593.
- Pillai, K. V, McDonald, A. G., and Wagner, F. G. (2011). Developing a model system in vitro to understand tracheary element development in Douglas-fir (*Pseudotsuga menziesii*). *Maderas. Cienc. y Tecnol.* 13, 3–18. doi:10.4067/S0718-221X2011000100001.
- Plomion, C., Leprovost, G., and Stokes, A. (2001). Wood Formation in Trees. *Plant Physiol.* 127, 1513–1523. doi:10.1104/pp.010816.1.
- Pyo, H., Demura, T., and Fukuda, H. (2007). TERE; a novel cis-element responsible for a coordinated expression of genes related to programmed cell death and secondary wall formation during differentiation of tracheary elements. *Plant J.* 51, 955–965. doi:10.1111/j.1365-313X.2007.03180.x.
- Queiroz, L. R. de S.; Barrichelo, L. E. G. (2008). *A celulose de eucalipto: uma oportunidade brasileira*. 1st ed. Vinhedo: Avis Brasilis Editora.
- Ranik, M., and Myburg, A. A. (2006). Six new cellulose synthase genes from *Eucalyptus* are associated with primary and secondary cell wall biosynthesis. *Tree Physiol.* 26, 545–556.
- Rapaport, F., Khanin, R., Liang, Y., Pirun, M., Krek, A., Zumbo, P., et al. (2013). Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biol* 14, R95. doi:10.1186/gb-2013-14-9-r95.
- Rowell, R. M. (2013). “Chemical modification of wood,” in *Handbook of Wood Chemistry and Wood Composites* (Madison: CRC Press), 537–589.
- Ruben Vanholme, Cesarino, I., Rataj, K., Xiao, Y., Sundin, L., Goeminne, G., et al. (2013). Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in Arabidopsis. *Science* (80-.). 341, 1103–1106. doi:10.1126/science.1241602.
- Ruzicka, K., Ursache, R., Hejátko, J., and Helariutta, Y. (2015). Xylem development - from the cradle to the grave. *New Phytol.* 207, 519–535. doi:10.1111/nph.13383.

- Salazar, M. M., Nascimento, L. C., Camargo, E. L. O., Vidal, R. O., Goncalves, D. C., and Marques, W. L. (2011). Comparative transcriptome analysis of tree Eucalyptus species using RNAseq technology : analysis of genes interfering in wood quality aspects. *BMC Proc.* 5, 175. doi:10.1186/1753-6561-5-S7-P175.
- Santos, R. B., Gomide, J. L., and Hart, P. W. (2006). Kraft Pulping of Reduced Metal Content Eucalyptus Wood : Process Impacts. 10, 6538–6547.
- Dos Santos, S. R., and Sansígolo, C. A. (2007). Influência da densidade básica da madeira de clones de Eucalyptus grandis x Eucalyptus urophylla na qualidade da polpa branqueada. *Cienc. Florest.* 17, 53–63.
- Sasaki-Sekimoto, Y., Jikumaru, Y., Obayashi, T., Saito, H., Masuda, S., Kamiya, Y., et al. (2013). Basic Helix-Loop-Helix transcription factors JASMONATE-ASSOCIATED MYC2-LIKE1 (JAM1), JAM2, and JAM3 are negative regulators of jasmonate responses in Arabidopsis. *Plant Physiol* 163, 291–304. doi:10.1104/pp.113.220129.
- Schuetz, M., Benske, A., Smith, R. A., Watanabe, Y., Tobimatsu, Y., Ralph, J., et al. (2014). Laccases direct lignification in the discrete secondary cell wall domains of protoxylem. *Plant Physiol.* 166, 798–807. doi:10.1104/pp.114.245597.
- Sehr, E. M., Agusti, J., Lehner, R., Farmer, E. E., Schwarz, M., and Greb, T. (2010). Analysis of secondary growth in the Arabidopsis shoot reveals a positive role of jasmonate signalling in cambium formation. *Plant J.* 63, 811–822. doi:10.1111/j.1365-313X.2010.04283.x.
- Shin, J., Ming, G., and Song, H. (2014). Decoding neural transcriptomes and epigenomes via high-throughput sequencing. *Nat Neurosci* 17, 1463–1475. doi:10.1038/nn.3814.
- Shinya, T., Iwata, E., Nakahama, K., Fukuda, Y., Nanto, K., Rosa, A. C., et al. (2016). Transcriptional profiles of hybrid Eucalyptus genotypes with contrasting lignin content reveal that monolignol biosynthesis-related genes regulate wood composition. *Front. Plant Sci.* 7, 1–12. doi:10.3389/fpls.2016.00443.
- Siah, C. H., Namasivayam, P., and Mohamed, R. (2016). Transcriptome reveals senescing callus tissue of Aquilaria malaccensis, an endangered tropical tree, triggers similar response as wounding with respect to terpenoid biosynthesis. *Tree Genet. Genomes* 12, 33. doi:10.1007/s11295-016-0993-z.
- Smith, L. G. (2001). Plant cell division: building walls in the right places. *Nat. Rev. Mol. Cell Biol.* 2, 33–39. doi:10.1038/35048050.

- Smith, R. A., Schuetz, M., Roach, M., Ellis, B., and Samuels, L. (2013). Neighboring parenchyma cells contribute to Arabidopsis xylem lignification, while lignification of interfascicular fibers is cell autonomous. *Plant Cell* 25, 3988–3999. doi:10.1105/tpc.113.117176.
- Soler, M., Camargo, E. L. O., Carocha, V., Cassan-Wang, H., San Clemente, H., Savelli, B., et al. (2015). The Eucalyptus grandis R2R3-MYB transcription factor family: Evidence for woody growth-related evolution and function. *New Phytol.* 206, 1364–1377. doi:10.1111/nph.13039.
- Soneson, C., and Delorenzi, M. (2013). A comparison of methods for differential expression analysis of RNA-seq data. *BMC Bioinformatics* 14, 91. doi:10.1186/1471-2105-14-91.
- Stepanova, A. N., and Alonso, J. M. (2009). Ethylene signaling and response: where different regulatory modules meet. *Curr. Opin. Plant Biol.* 12, 548–555. doi:10.1016/j.pbi.2009.07.009.
- Suzano Papel e Celulose (2008). Eucaliptocultura e desenvolvimento socioambiental. Available at: file:///C:/Users/AnaPaula/Documents/Tese_Arquivos/TESE/Eucalipto/Guia_do_Eucalipto_junho_2008.pdf.
- Taiz, L., and Zeiger, E. (2010). *Plant Physiology*. doi:10.1104/pp.900074.
- Takahashi, J., Rudsander, U. J., Hedenström, M., Banasiak, A., Harholt, J., Amelot, N., et al. (2009). KORRIGAN1 and its aspen homolog PttCel9A1 decrease cellulose crystallinity in arabidopsis stems. *Plant Cell Physiol.* 50, 1099–1115. doi:10.1093/pcp/pcp062.
- Taylor-Teeple, M., Lin, L., Lucas, M. De, Turco, G., Toal, T. W., Gaudinier, A., et al. (2015). An Arabidopsis gene regulatory network for secondary cell wall synthesis. *Nature* 517, 571–575. doi:10.1038/nature14099.
- Taylor, N. G., Howells, R. M., Huttly, A. K., Vickers, K., and Turner, S. R. (2003). Interactions among three distinct Cesa proteins essential for cellulose synthesis. *Proc. Natl. Acad. Sci.* 100, 1450–1455. doi:10.1073/pnas.0337628100.
- Thumma, B. R., Sharma, N., and Southerton, S. G. (2012). Transcriptome sequencing of Eucalyptus camaldulensis seedlings subjected to water stress reveals functional single nucleotide polymorphisms and genes under selection. *BMC Genomics* 13, 364. doi:10.1186/1471-2164-13-364.

- Trapnell, C., Hendrickson, D. G., Sauvageau, M., Goff, L., Rinn, J. L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* 31, 46–53. doi:10.1038/nbt.2450.
- Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat : discovering splice junctions with RNA-Seq. 25, 1105–1111. doi:10.1093/bioinformatics/btp120.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578. doi:10.1038/nprot.2012.016.
- Tsuyama, T., Kawai, R., Shitan, N., Matoh, T., Sugiyama, J., Yoshinaga, A., et al. (2013). Proton-dependent coniferin transport, a common major transport event in differentiating xylem tissue of woody plants. *Plant Physiol.* 162, 918–926. doi:10.1104/pp.113.214957.
- Turlapati, P. V., Kim, K. W., Davin, L. B., and Lewis, N. G. (2011). The laccase multigene family in *Arabidopsis thaliana*: Towards addressing the mystery of their gene function(s). *Planta* 233, 439–470. doi:10.1007/s00425-010-1298-3.
- Turner, S., Gallois, P., and Brown, D. (2007). Tracheary Element Differentiation. *Annu. Rev. Plant Biol.* 58, 407–433. doi:10.1146/annurev.arplant.57.032905.105236.
- Twumasi, P., Schel, J. H. N., van Ieperen, W., Woltering, E., Van Kooten, O., and Emons, A. M. C. (2009). Establishing in vitro *Zinnia elegans* cell suspension culture with high tracheary element differentiation. *Cell Biol. Int.* 33, 524–533. doi:10.1016/j.cellbi.2009.01.019.
- Urbanowicz, B. R., Pena, M. J., Moniz, H. A., Moremen, K. W., and York, W. S. (2014). Two *Arabidopsis* proteins synthesize acetylated xylan in vitro. *Plant J.* 80, 197–206. doi:10.1111/tpj.12643.
- Ursache, R., Nieminen, K., and Helariutta, Y. (2013). Genetic and hormonal regulation of cambial development. *Physiol. Plant.* 147, 36–45. doi:10.1111/j.1399-3054.2012.01627.x.
- Vahala, J., Felten, J., Love, J., Gorzsás, A., Gerber, L., Lamminmäki, A., et al. (2013). A genome-wide screen for ethylene-induced Ethylene Response Factors (ERFs) in hybrid aspen stem identifies ERF genes that modify stem growth and wood properties. *New Phytol.* 200, 511–522. doi:10.1111/nph.12386.
- Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant Physiol.* 153, 895–905. doi:10.1104/pp.110.155119.

- Vanholme, R., Morreel, K., Ralph, J., and Boerjan, W. (2008). Lignin engineering. *Curr. Opin. Plant Biol.* 11, 278–285. doi:10.1016/j.pbi.2008.03.005.
- Verk, M. C. Van, Hickman, R., Pieterse, C. M. J., and Wees, S. C. M. Van (2013). RNA-Seq : revelation of the messengers. *Trends Plant Sci.* 18, 1–23. doi:10.1016/j.tplants.2013.02.001.
- Vincken, J.-P. (2003). If Homogalacturonan Were a Side Chain of Rhamnogalacturonan I. Implications for Cell Wall Architecture. *Plant Physiol.* 132, 1781–1789. doi:10.1104/pp.103.022350.
- Wagner, A., Donaldson, L., Kim, H., Phillips, L., Flint, H., Steward, D., et al. (2009). Suppression of 4-Coumarate-CoA Ligase in the Coniferous Gymnosperm *Pinus radiata*. *Plant Physiol.* 149, 370–383. doi:10.1104/pp.108.125765.
- Wagner, A., Ralph, J., Akiyama, T., Flint, H., Phillips, L., Torr, K., et al. (2007). Exploring lignification in conifers by silencing hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase in *Pinus radiata*. *Proc. Natl. Acad. Sci. U. S. A.* 104, 11856–11861. doi:10.1073/pnas.0701428104.
- Wagner, A., Tobimatsu, Y., Goeminne, G., Phillips, L., Flint, H., Steward, D., et al. (2013). Suppression of CCR impacts metabolite profile and cell wall composition in *Pinus radiata* tracheary elements. *Plant Mol. Biol.* 81, 105–117. doi:10.1007/s11103-012-9985-z.
- Wagner, A., Tobimatsu, Y., Phillips, L., Flint, H., Torr, K., Donaldson, L., et al. (2011). CCoAOMT suppression modifies lignin composition in *Pinus radiata*. *Plant J.* 67, 119–129. doi:10.1111/j.1365-313X.2011.04580.x.
- Wang, H., Avci, U., Nakashima, J., Hahn, M. G., Chen, F., and Dixon, R. A. (2010). Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. *Proc. Natl. Acad. Sci. U. S. A.* 107, 22338–43. doi:10.1073/pnas.1016436107.
- Wang, H., Yang, J. H., Chen, F., Torres-Jerez, I., Tang, Y., Wang, M., et al. (2016). Transcriptome analysis of secondary cell wall development in *Medicago truncatula*. *BMC Genomics* 17, 23. doi:10.1186/s12864-015-2330-6.
- Wang, Y., Bouchabke-coussa, O., Lebris, P., Antelme, S., Soulhat, C., Gineau, E., et al. (2015). LACCASE5 Is Required for Lignification of the *Brachypodium distachyon* Culm. *Plant Physiol.* 168, 192–204. doi:10.1104/pp.114.255489.

- Wang, Y., Chantreau, M., Sibout, R., and Hawkins, S. (2013). Plant cell wall lignification and monolignol metabolism. *Front. Plant Sci.* 4, 220. doi:10.3389/fpls.2013.00220.
- Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63. doi:10.1038/nrg2484.
- Watanabe, Y., Meents, M. J., McDonnell, L. M., Barkwill, S., Sampathkumar, A., Cartwright, H. N., et al. (2015). Visualization of cellulose synthases in Arabidopsis secondary cell walls. *Science (80-.).* 350, 198–203.
- Wiley, J. & S. (1989). “The pulp and paper making processes,” in *Technologies for Reducing Dioxin in the Manufacture of Bleached Wood Pulp*, 17–26.
- Wong, M. M. L., Cannon, C. H., and Wickneswari, R. (2011). Identification of lignin genes and regulatory sequences involved in secondary cell wall formation in *Acacia auriculiformis* and *Acacia mangium* via de novo transcriptome sequencing. *BMC Genomics* 12, 13. doi:10.1186/1471-2164-12-342.
- Xu, H., Cao, D., Feng, J., Wu, H., Lin, J., and Wang, Y. (2016). Transcriptional regulation of vascular cambium activity during the transition from juvenile to mature stages in *Cunninghamia lanceolata*. *J. Plant Physiol.* 200, 7–17. doi:10.1016/j.jplph.2016.06.003.
- Yamagishi, Y., Sato, T., Uchiyama, H., Yoshimoto, J., Nakagawa, R., Nakaba, S., et al. (2012). Tracheary elements that resemble secondary xylem in calli derived from the conifers, *Torreya nucifera* and *Cryptomeria japonica*. *J. Wood Sci.* 58, 557–562. doi:10.1007/s10086-012-1288-0.
- Yamagishi, Y., Uchiyama, H., Sato, T., Kitamura, K., Yoshimoto, J., Nakaba, S., et al. (2015). In vitro induction of the formation of tracheary elements from suspension-cultured cells of the conifer *Cryptomeria japonica*. *Trees - Struct. Funct.*, 1283–1289. doi:10.1007/s00468-014-1139-2.
- Yamagishi, Y., Yoshimoto, J., Uchiyama, H., Nabeshima, E., Nakaba, S., Watanabe, U., et al. (2013). In vitro induction of secondary xylem-like tracheary elements in calli of hybrid poplar (*Populus sieboldii* x *P. grandidentata*). *Planta* 237, 1179–1185. doi:10.1007/s00425-013-1839-7.
- Yamaguchi, M., Mitsuda, N., Ohtani, M., Ohme-Takagi, M., Kato, K., and Demura, T. (2011). VASCULAR-RELATED NAC-DOMAIN 7 directly regulates the expression of a broad range of genes for xylem vessel formation. *Plant J.* 66, 579–590. doi:10.1111/j.1365-313X.2011.04514.x.

- Yamamoto, R., Fujioka, S., Iwamoto, K., Demura, T., Takatsuto, S., Yoshida, S., et al. (2007). Co-regulation of brassinosteroid biosynthesis-related genes during xylem cell differentiation. *Plant Cell Physiol.* 48, 74–83. doi:10.1093/pcp/pcl039.
- Yang, F., Mitra, P., Zhang, L., Prak, L., Verhertbruggen, Y., Kim, J. S., et al. (2013). Engineering secondary cell wall deposition in plants. *Plant Biotechnol. J.* 11, 325–335. doi:10.1111/pbi.12016.
- Yang, J. H., and Wang, H. (2016). Molecular Mechanisms for Vascular Development and Secondary Cell Wall Formation. *Front. Plant Sci.* 7, 1–8. doi:10.3389/fpls.2016.00356.
- Yang, L., Zhao, X., Yang, F., Fan, D., Jiang, Y., and Luo, K. (2016). PtrWRKY19, a novel WRKY transcription factor, contributes to the regulation of pith secondary wall formation in *Populus trichocarpa*. *Nat. Publ. Gr.*, 1–12. doi:10.1038/srep18643.
- Ye, Z. H., and Zhong, R. (2015). Molecular control of wood formation in trees. *J. Exp. Bot.* 66, 4119–4131. doi:10.1093/jxb/erv081.
- Yoon, J., Choi, H., and An, G. (2015). Roles of lignin biosynthesis and regulatory genes in plant development. *J. Integr. Plant Biol.* 57, 902–912. doi:10.1111/jipb.12422.
- Yu, R., Wang, J., Xu, L., Wang, Y., Wang, R., Zhu, X., et al. (2016). Transcriptome Profiling of Taproot Reveals Complex Regulatory Networks during Taproot Thickening in Radish (*Raphanus sativus* L.). *Front. Plant Sci.* 7, 1–17. doi:10.3389/fpls.2016.01210.
- Yuan, Y., Teng, Q., Zhong, R., and Ye, Z. H. (2016). TBL3 and TBL31, Two Arabidopsis DUF231 domain proteins, are required for 3-O-monoacetylation of xylan. *Plant Cell Physiol.* 57, 35–45. doi:10.1093/pcp/pcv172.
- Zhang, Y., Han, X., Sang, J., He, X., Liu, M., Qiao, G., et al. (2016). Transcriptome analysis of immature xylem in the Chinese fir at different developmental phases. *PeerJ* 4, 21. doi:10.7717/peerj.2097.
- Zhang, Y., Liang, W., Shi, J., Xu, J., and Zhang, D. (2013). MYB56 Encoding a R2R3 MYB Transcription Factor Regulates Seed Size in *Arabidopsis thaliana*. *J. Integr. Plant Biol.* 55, 1166–1178. doi:10.1111/jipb.12094.
- Zhao, Q., Nakashima, J., Chen, F., Yin, Y., Fu, C., Yun, J., et al. (2013). LACCASE Is Necessary and Nonredundant with PEROXIDASE for Lignin Polymerization during Vascular Development in *Arabidopsis*. 25, 3976–3987. doi:10.1105/tpc.113.117770.

- Zhong, R., Lee, C., Zhou, J., McCarthy, R. L., and Ye, Z.-H. (2008). A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell* 20, 2763–2782. doi:10.1105/tpc.108.061325.
- Zhong, R., Ye, Z., and E-mail, C. (2015). Secondary Cell Walls : Biosynthesis , Patterned Deposition and Transcriptional Regulation Special Focus Issue – Review. 56, 195–214. doi:10.1093/pcp/pcu140.
- Zhou, J., Lee, C., Zhong, R., and Ye, Z.-H. (2009). MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *Plant Cell* 21, 248–66. doi:10.1105/tpc.108.063321.
- Zhou, X., Jacobs, T. B., Xue, L.-J., Harding, S. A., and Tsai, C.-J. (2015). Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate : CoA ligase specificity and redundancy. *New Phytol.* 208, 298–301. doi:10.1111/nph.13470.
- Zhu, T., Nevo, E., Sun, D., and Peng, J. (2012). Phylogenetic analyses unravel the evolutionary history of nac proteins in plants. *Evolution (N. Y.)*. 66, 1833–1848. doi:10.1111/j.1558-5646.2011.01553.x.
- Ziebell, A., Gjersing, E., Hinchey, M., Katahira, R., Sykes, R. W., Johnson, D. K., et al. (2016). Downregulation of p-Coumaroyl Quinate/Shikimate 3'-Hydroxylase (C3'H) or Cinnamate-4-hydroxylase (C4H) in *Eucalyptus urophylla* × *Eucalyptus grandis* Leads to Increased Extractability. *Bioenergy Res.*, 1–9. doi:10.1007/s12155-016-9713-7.
- Zykwinska, A. W., Ralet, M. J., Garnier, C. D., and Thibault, J.-F. J. (2005). Evidence for in vitro binding of pectin side chains to cellulose. *Plant Physiol.* 139, 397–407. doi:10.1104/pp.105.065912.1.

ANEXOS

ANEXO A. Tabela com a descrição do meio de cultura SP utilizado para indução *in vitro* dos calos.

Componentes	Concentração
Macronutrientes	50 ml/l
Micronutrientes	1 ml/l
Ferro	1 ml/l
Inositol	50 mg/l
Sacarose	20 g/l
PVP (polivinilpirrolidone)	1 g/l
Vitaminas	3 mg/l
Tiamina	10 mg/l
L-Glutamina	146 mg/l
Fitagel	3 g/l
pH	5,8

ANEXO B. Eletroferogramas do Agilent 2100 Bioanalyzer indicando integridade do RNA total extraído das amostras biológicas.

ANEXO C. Tabela contendo as concentrações (ng/ul) do RNA total extraído pelo *RNeasy Mini Kit* de cada amostra, a relação 260/280 obtidos pelo *NanoDrop* e o RIN reportado pelo *Agilent 2100 Bioanalyzer*.

Amostras	Tratamento	Concentração (ng/ul)	Relação 260/280	Volume (ul)	RNA Integrity Number (RIN)
M1	1	46,4	1,917	50	9.4
M2	1	39,2	2,026	50	9.7
M3	1	30,6	2,096	50	9.2
M4	1	27,6	2,029	50	9.3
A1	2	29,9	1,954	50	8.7
A2	2	33,2	1,834	50	8.9
A3	2	37,0	2,044	50	9.0
A4	2	35,0	2,096	50	8.9
T1	3	128,0	2,159	50	8.7
T2	3	216,8	1,90	50	8.7
T3	3	122,3	2,192	50	8.7
T4	3	74,05	2.113	50	9.1

ANEXO D. Gel de agarose 1,5% confirmando a integridade das amostras de RNA total extraídas das frações celulares.

ANEXO E. Exemplo de resultados de qualidade (amostra T2/R1) apresentados pelo programa FastQC, antes (dados brutos) e após a limpeza pelo programa Trimmomatic.

ANEXO F. MA-plot das comparações A vs. T, M vs. A e M vs. T. obtidas pelo pacote DESeq. O $\log_2$ fold change para uma comparação particular é plotada no eixo y e a média das contagens normalizadas é mostrada no eixo x. Cada gene é representado com um ponto. Genes com um valor de p ajustado abaixo de um limite $padj < 0,05$ são mostrados em vermelho.

ANEXO G. Dendrograma mostrando a relação em termos de similaridade/proximidade entre as diferentes amostras biológicas de cada tratamento.

ANEXO H. Exemplos das redes completas de termos GO enriquecidos e significativos para todos os genes mais expressos na comparação M vs. T, reportadas pelo aplicativo BINGO.

ANEXO I. Exemplos das redes completas de termos GO enriquecidos e significativos para os genes mais expressos na comparação M vs. T, com $\log_2 \text{fold change} \geq 1,5$.

ANEXO J. Tabela contendo as informações estatísticas e descrição funcional dos genes diferencialmente expressos identificados na comparação M vs. T relacionados às rotas de biossíntese dos componentes da parede celular secundária descritos na literatura. Os genes discutidos nesse trabalho estão realçados, com a cor laranja os fatores de transcrição, com a cor azul os genes participantes da rota da celulose e hemicelulose, e com a cor amarelo os genes da rota de biossíntese da lignina.

LocusName	sample_1	sample_2	value_1	value_2	log2 fold_change	p_value	Símbolo (arabi-symbol)	Descrição
Eucgr.A02818	M	T	1,50306	38,3715	4,67406	5,00E-05	CBF4,DREB1D	C-repeat-binding factor 4
Eucgr.A02820	M	T	0,488659	6,95565	3,83129	5,00E-05	CBF4,DREB1D	C-repeat-binding factor 4
Eucgr.F04198	M	T	4,80005	65,7248	3,77532	5,00E-05		Peroxidase superfamily protein
Eucgr.D01088	M	T	0,08702	0,834739	3,26191	0,00455	ATCAD7, CAD7, CAD10 ELI3,ELI3	Elicitor-activated gene 3-1
Eucgr.E00850	M	T	7,66316	73,2906	3,25762	5,00E-05	ATWRKY53,WRKY53	WRKY family transcription factor
Eucgr.B02752	M	T	0,490152	3,36831	2,78072	5,00E-05	LAC7	Laccase 7
Eucgr.A02822	M	T	0,167587	1,09041	2,70188	0,0074	CBF4,DREB1D	C-repeat-binding factor 4
Eucgr.E02950	M	T	7,51109	48,3275	2,68575	5,00E-05	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.A00682	M	T	0,295051	1,79951	2,60857	0,00115		Peroxidase superfamily protein
Eucgr.H03521	M	T	0,188158	1,04349	2,4714	5,00E-05		Copper amine oxidase family protein
Eucgr.J01617	M	T	0,58914	3,14176	2,41489	5,00E-05		Peroxidase superfamily protein
Eucgr.D02359	M	T	16,8816	85,2246	2,33582	5,00E-05	AtWRKY41,WRKY41	WRKY family transcription factor
Eucgr.I01120	M	T	6,27461	30,6615	2,28883	5,00E-05	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.L02416	M	T	3,66506	17,3519	2,24319	5,00E-05		Peroxidase superfamily protein
Eucgr.H01923	M	T	0,543478	2,38337	2,13271	5,00E-05	GATL2	Galacturonosyltransferase-like 2
Eucgr.E04120	M	T	0,569105	2,45362	2,10814	5,00E-05		Peroxidase superfamily protein
Eucgr.J01616	M	T	4,62425	18,1799	1,97505	5,00E-05		Peroxidase superfamily protein
Eucgr.K00331	M	T	1,97064	7,33286	1,89571	5,00E-05	AtWRKY41,WRKY41	WRKY family transcription factor
Eucgr.F03595	M	T	2,64791	9,78218	1,8853	5,00E-05	ATUGT85A1,UGT85A1	UDP-Glycosyltransferase superfamily protein
Eucgr.I00057	M	T	2,53124	9,10892	1,84744	5,00E-05	ANAC102,NAC102	NAC domain containing protein 102
Eucgr.L01657	M	T	4,32282	15,5328	1,84527	5,00E-05	UGT72E1	UDP-glucosyl transferase 72E1
Eucgr.J01618	M	T	7,55338	27,05	1,84043	5,00E-05		Peroxidase superfamily protein
Eucgr.F03414	M	T	3,98751	14,2575	1,83816	5,00E-05	PAT1	GRAS family transcription factor
Eucgr.H03667	M	T	4,31486	14,9746	1,79514	5,00E-05		Peroxidase superfamily protein
Eucgr.E03715	M	T	0,246959	0,851761	1,78618	0,0024		Peroxidase superfamily protein
Eucgr.A02463	M	T	0,541709	1,80274	1,7346	5,00E-05	TBR	Plant protein of unknown function (DUF828)
Eucgr.H02855	M	T	36,1905	120,258	1,73245	5,00E-05	BGLU46	Beta glucosidase 46
Eucgr.K03009	M	T	19,0348	63,1254	1,72959	5,00E-05	LAC7	Laccase 7
Eucgr.D01087	M	T	23,0663	74,8918	1,69902	5,00E-05	ATCAD7,CAD7, CAD9, ELI3,ELI3-1	Elicitor-activated gene 3-1
Eucgr.K03008	M	T	9,43292	29,9522	1,66689	5,00E-05	LAC7	Laccase 7
Eucgr.H00071	M	T	7,75833	23,8351	1,61927	5,00E-05	BGLU46	Beta glucosidase 46
Eucgr.H03362	M	T	5,77224	17,5441	1,60378	5,00E-05	ANAC083,NAC083,VNI2	NAC domain containing protein 83
Eucgr.J01619	M	T	8,7109	26,1775	1,58743	5,00E-05		Peroxidase superfamily protein
Eucgr.F03596	M	T	8,36921	24,9415	1,57538	5,00E-05	AtUGT85A3,UGT85A3	UDP-glucosyl transferase 85A3
Eucgr.I01882	M	T	18,2938	53,8955	1,55881	5,00E-05	GATL10	Galacturonosyltransferase-like 10
Eucgr.L02740	M	T	8,54271	24,9619	1,54696	5,00E-05		Peroxidase superfamily protein

Eucgr.F04195	M	T	9,59157	27,6555	1,52773	5,00E-05		Peroxidase superfamily protein
Eucgr.D00471	M	T	0,832332	2,29585	1,4638	5,00E-05	ATCAD1,CAD1, CAD	Cinnamyl-alcohol dehydrogenase
Eucgr.F01531	M	T	3,0764	8,47509	1,46198	5,00E-05	GATL2	Galacturonosyltransferase-like 2
Eucgr.H00890	M	T	13,3674	36,1237	1,43423	5,00E-05	RCI3,RCI3A	Peroxidase superfamily protein
Eucgr.L02041	M	T	0,617401	1,64524	1,41402	0,0015	ATERF-2,ATERF2,ERF2	Ethylene responsive element binding factor 2
Eucgr.B03906	M	T	8,31939	22,0023	1,40311	5,00E-05	HCT	Hydroxycinnamoyl-coa shikimate/quinat hydroxycinnamoyl transferase
Eucgr.I01116	M	T	0,951261	2,51394	1,40204	5,00E-05	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.A00940	M	T	7,90799	20,6141	1,38225	5,00E-05	HYR1	UDP-Glycosyltransferase superfamily protein
Eucgr.I02632	M	T	8,99855	23,353	1,37584	5,00E-05	CAMTA3,SR1	Signal responsive 1
Eucgr.L03642	M	T	8,61608	22,2909	1,37135	5,00E-05	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.J02173	M	T	2,21947	5,49357	1,30753	5,00E-05		Peroxidase superfamily protein
Eucgr.J00196	M	T	12,6761	31,0755	1,29367	5,00E-05	ESK1,TBL29, DUF231	Plant protein of unknown function (DUF828)
Eucgr.A02436	M	T	1,64761	3,99346	1,27727	5,00E-05	bZIP65,TGA10	Bzip transcription factor family protein
Eucgr.J01468	M	T	16,471	39,7971	1,27273	5,00E-05	GT72B1,UGT72B1	UDP-Glycosyltransferase superfamily protein
Eucgr.G02223	M	T	10,2953	24,5881	1,25597	5,00E-05	ATCAD9,CAD9, CAD21	Cinnamyl alcohol dehydrogenase
Eucgr.E04220	M	T	198,287	462,214	1,22097	5,00E-05	LAC14	Laccase 14
Eucgr.D00468	M	T	5,36939	12,482	1,21702	5,00E-05	ATCAD1,CAD1, CAD4	Cinnamyl-alcohol dehydrogenase
Eucgr.J01601	M	T	0,801081	1,84681	1,20502	0,00015	AtMYB43,MYB43	MYB domain protein 43
Eucgr.J02484	M	T	1,04242	2,38965	1,19686	5,00E-05		Peroxidase superfamily protein
Eucgr.B03943	M	T	2,15722	4,87891	1,17738	5,00E-05		AMP-dependent synthetase and ligase family protein 4CL7
Eucgr.G02921	M	T	69,0422	155,978	1,17579	5,00E-05	UXS6	UDP-XYL synthase 6
Eucgr.L00520	M	T	1,02847	2,32044	1,1739	5,00E-05	CAMTA1,EICBP.B	Ethylene induced calmodulin binding protein
Eucgr.J01844	M	T	114,539	258,154	1,1724	5,00E-05	ATC4H, C4H, C4H1	Cinnamate-4-hydroxylase
Eucgr.F04449	M	T	0,883675	1,97356	1,15921	5,00E-05	ATMYB63,MYB63	MYB domain protein 63
Eucgr.G02325	M	T	14,1725	30,9696	1,12776	5,00E-05	CCR7	NAD(P)-binding Rossmann-fold superfamily protein
Eucgr.F01164	M	T	23,9944	52,288	1,12378	5,00E-05	ATERF-4,ATERF4,ERF4,RAP2.5	Ethylene responsive element binding factor 4
Eucgr.I00608	M	T	4,09797	8,79345	1,10152	5,00E-05	bHLH093	Beta HLH protein 93
Eucgr.I01117	M	T	7,3709	15,805	1,10047	5,00E-05	UGT73B3	UDP-glucosyl transferase 73B3
Eucgr.H01249	M	T	27,8365	57,896	1,05648	5,00E-05	UGT72E1	UDP-glucosyl transferase 72E1
Eucgr.G01977	M	T	0,759318	1,50242	0,984513	0,00075	ATGUT1,GUT2,IRX10	Exostosin family protein
Eucgr.B03636	M	T	1,09687	2,16752	0,982651	5,00E-05	TBR	Plant protein of unknown function (DUF828)
Eucgr.F01003	M	T	0,699069	1,38116	0,982375	0,0001	PAT1	GRAS family transcription factor
Eucgr.G02458	M	T	2,50587	4,92443	0,974644	5,00E-05	BGLU40	Beta glucosidase 40
Eucgr.A02702	M	T	6,57472	12,7511	0,955627	5,00E-05	UGT88A1	UDP-glucosyl transferase 88A1
Eucgr.F00659	M	T	32,5251	62,3211	0,938166	5,00E-05	ATERF3,ERF3	Ethylene responsive element binding factor 3
Eucgr.A00936	M	T	7,14194	13,6343	0,932857	5,00E-05	HYR1	UDP-Glycosyltransferase superfamily protein
Eucgr.I01114	M	T	6,50577	12,4118	0,931924	5,00E-05	IRX15-L	Protein of unknown function (DUF579)
Eucgr.D02324	M	T	117,604	218,265	0,892144	5,00E-05	AtWRKY22,WRKY22	WRKY family transcription factor
Eucgr.B02744	M	T	3,41054	6,27359	0,879292	5,00E-05	AtUGT85A2,UGT85A2	UDP-glucosyl transferase 85A2
Eucgr.G02852	M	T	2,77837	5,10223	0,876891	5,00E-05	ATPAL2,PAL2, PAL7	Phenylalanine ammonia-lyase 2

Eucgr.K01666	M	T	5,15265	9,46189	0,876815	5,00E-05	TUB6	Beta-6 tubulin
Eucgr.J01079	M	T	16,231	29,7208	0,872724	5,00E-05	ATPAL2,PAL2, PAL9	Phenylalanine ammonia-lyase 2
Eucgr.J01372	M	T	138,126	251,283	0,863327	5,00E-05	UGD3	UDP-glucose 6-dehydrogenase family protein
Eucgr.A00942	M	T	12,2216	22,0839	0,853565	5,00E-05	HYR1	UDP-Glycosyltransferase superfamily protein
Eucgr.C02389	M	T	7,17342	12,9111	0,847881	5,00E-05		Peroxidase superfamily protein
Eucgr.B03653	M	T	4,18179	7,50159	0,843076	0,00795	BGLU17	Beta glucosidase 17
Eucgr.F02864	M	T	3,41426	6,12379	0,842853	5,00E-05	AtMYB20,MYB20	Myb domain protein 20
Eucgr.F02650	M	T	25,6324	45,3551	0,823297	5,00E-05	LAC14	Laccase 14
Eucgr.B03360	M	T	1,52908	2,67641	0,807632	0,0001	TBL27	Trichome birefringence-like 27
Eucgr.D02223	M	T	2,30493	4,01824	0,801842	5,00E-05	HB-7,HDG5	Homeobox-7
Eucgr.J02352	M	T	5,37387	9,26286	0,785497	5,00E-05		Peroxidase superfamily protein
Eucgr.K00786	M	T	27,6453	47,5444	0,782243	5,00E-05	ATWRKY6,WRKY6	Wrky family transcription factor
Eucgr.B03654	M	T	718,235	1234,81	0,781763	5,00E-05	BGLU17	Beta glucosidase 17
Eucgr.F02171	M	T	1,37305	2,33482	0,765924	0,0018	bHLH093	Beta hlh protein 93
Eucgr.E00037	M	T	5,88053	9,99126	0,76472	5,00E-05	DAG1	Dof-type zinc finger dna-binding family protein
Eucgr.J03126	M	T	27,6666	46,927	0,76227	5,00E-05	HCT5	Hydroxycinnamoyl-coa shikimate/quinat hydroxycinnamoyl transferase
Eucgr.D00591	M	T	1,80322	2,97644	0,723015	0,0028	ANAC010,NAC010,SND3	NAC domain containing protein 10
Eucgr.K02875	M	T	0,984789	1,62216	0,720031	0,01075	UGT74F1	UDP-glycosyltransferase 74 F1
Eucgr.B00859	M	T	17,0258	27,9355	0,714377	5,00E-05	BGLU44	B-S glucosidase 44
Eucgr.J00851	M	T	29,3685	46,9146	0,67577	5,00E-05	CBB1, DIM, DIM1, DWF1, EVE1	Cell elongation protein / DWARF1 / DIMINUTO (DIM)
Eucgr.F03978	M	T	8,42998	13,4578	0,674842	5,00E-05	HCT4	Hydroxycinnamoyl-coa shikimate/quinat hydroxycinnamoyl transferase
Eucgr.E04011	M	T	114,664	181,776	0,664745	5,00E-05	ATWRKY6,WRKY6	WRKY family transcription factor
Eucgr.C01769	M	T	2,47484	3,90978	0,659756	5,00E-05	AtCESA1,CESA1,RSW1	Cellulose synthase 1
Eucgr.F01543	M	T	5,73613	8,86325	0,627758	5,00E-05	UGT72B3	UDP-glucosyl transferase 72B3
Eucgr.K02506	M	T	27,1355	41,4492	0,611162	5,00E-05	UGD3	UDP-glucose 6-dehydrogenase family protein
Eucgr.D00592	M	T	5,11214	7,69054	0,589158	0,00075	ANAC073,NAC073,SND2	NAC domain containing protein 73
Eucgr.L00599	M	T	7,38292	11,011	0,576678	0,0003	ATUGT85A1,UGT85A1	UDP-Glycosyltransferase superfamily protein
Eucgr.B03054	M	T	3,55697	5,28674	0,57173	0,0025	GATL3	Galacturonosyltransferase-like 3
Eucgr.F02090	M	T	10,5229	15,6124	0,569152	5,00E-05	ARF5,IAA24,MP	Transcriptional factor B3 family protein / auxin-responsive factor AUX/IAA-related
Eucgr.C03303	M	T	12,5832	18,5516	0,560044	0,0001	TBL16	Trichome birefringence-like 16
Eucgr.G02848	M	T	15,7575	23,2124	0,558851	5,00E-05	ATPAL1,PAL1, PAL3	PHE ammonia lyase 1
Eucgr.F02588	M	T	141,894	207,38	0,547461	5,00E-05	CINV2	Cytosolic invertase 2
Eucgr.K03547	M	T	4,06642	5,88527	0,533349	0,0014	TBL6	Trichome birefringence-like 6
Eucgr.L01603	M	T	8,18149	11,7377	0,520712	0,00885		Peroxidase superfamily protein
Eucgr.A01385	M	T	211,31	297,647	0,494241	5,00E-05	RCI3,RCI3A	Peroxidase superfamily protein
Eucgr.A00935	M	T	56,5039	79,0821	0,485001	0,00015	HYR1	UDP-Glycosyltransferase superfamily protein
Eucgr.F01647	M	T	6,29677	8,79971	0,482844	0,00225	HXK3	Hexokinase 3
Eucgr.G01186	M	T	69,0804	95,4384	0,466293	0,0003	TUA2	Tubulin alpha-2 chain

Eucgr.K00264	M	T	105,655	145,804	0,46466	0,00025	TUB8	Tubulin beta 8
Eucgr.J00923	M	T	15,4724	21,3355	0,46356	0,0007	ARF10	Auxin response factor 10
Eucgr.A00966	M	T	368,018	506,247	0,460065	0,00055	RCI3,RCI3A	Peroxidase superfamily protein
Eucgr.E04217	M	T	34,9951	47,1433	0,429902	0,0007	LAC14	Laccase 14
Eucgr.A00947	M	T	55,4377	73,0112	0,397249	0,0022	UGT71B6	UDP-glucosyl transferase 71B6
Eucgr.E04216	M	T	20,6298	26,9519	0,385655	0,0075	LAC14	Laccase 14
Eucgr.D01612	M	T	47,2657	61,1945	0,372607	0,0029	TUB6	Beta-6 tubulin
Eucgr.F04472	M	T	39,7504	50,6622	0,34994	0,00535	AtMC4,MC4	Metacaspase
Eucgr.C02178	M	T	19,5551	24,5	0,325232	0,01045	ARF7,BIP,IAA21,IAA23,IAA25	Transcriptional factor B3 family protein / auxin-responsive factor AUX/IAA-related
Eucgr.K01043	M	T	26,9516	21,0084	-0,35941	0,00525		RING/U-box superfamily protein
Eucgr.D01117	M	T	20,7055	15,9998	-0,37196	0,00705	AtUGP1,UGP,UGP1	Udp-glucose pyrophosphorylase 1
Eucgr.A01019	M	T	70,4225	52,8691	-0,41361	0,00085	SPL12	Squamosa promoter-binding protein-like 12
Eucgr.B03537	M	T	235,593	176,538	-0,41632	0,0014	ANAC083,NAC083,VNI2	NAC domain containing protein 83
Eucgr.H01487	M	T	23,3756	17,4026	-0,4257	0,0014	BHLH42,TT8	Basic helix-loop-helix (bhlh) DNA-binding superfamily protein
Eucgr.F00740	M	T	59,3422	44,1185	-0,42767	0,00065	ATWRKY6,WRKY6	WRKY family transcription factor
Eucgr.J01998	M	T	26,5441	19,3394	-0,45685	0,00085	WRKY4	WRKY DNA-binding protein 4
Eucgr.K00043	M	T	6,02502	4,30934	-0,4835	0,0105	NACA2	Nascent polypeptide-associated complex subunit alpha-like protein 2
Eucgr.G03217	M	T	22,7851	16,2811	-0,48489	0,0004	BZIP45,TGA6	TGACG motif-binding factor 6
Eucgr.J00491	M	T	4,70883	3,29751	-0,51399	0,0033	bZIP21,TGA9	Bzip transcription factor family protein
Eucgr.B03529	M	T	5,61913	3,90118	-0,52644	0,0044	ATTCP20,PCF1,TCP20	TEOSINTE BRANCHED 1, cycloidea, PCF (TCP)-domain family protein 20
Eucgr.H04783	M	T	24,512	16,6807	-0,55531	5,00E-05	CAMTA3,SR1	Signal responsive 1
Eucgr.F03543	M	T	24,4902	16,0543	-0,60924	0,0001	4CL9	
Eucgr.F03605	M	T	82,0404	52,2293	-0,65147	5,00E-05	CCR5	NAD(P)-binding Rossmann-fold superfamily protein
Eucgr.H03910	M	T	96,4575	61,383	-0,65205	5,00E-05		Peroxidase superfamily protein
Eucgr.L03550	M	T	3,78833	2,40742	-0,65407	0,0005	bHLH121	Basic helix-loop-helix (bhlh) DNA-binding superfamily protein
Eucgr.A02694	M	T	6,62383	4,17322	-0,66651	0,00035	UGT88A1	UDP-glucosyl transferase 88A1
Eucgr.F02624	M	T	604,425	378,306	-0,67601	5,00E-05	ATOMT1,OMT1, 4CL8	O-methyltransferase 1
Eucgr.B03858	M	T	112,601	70,1195	-0,68334	5,00E-05	UGT88A1	UDP-glucosyl transferase 88A1
Eucgr.H02414	M	T	13,2767	8,17608	-0,69942	5,00E-05	ATCAD9,CAD9, CAD37	Cinnamyl alcohol dehydrogenase 9
Eucgr.I00564	M	T	43,5095	26,7748	-0,70045	5,00E-05	AIL6,PLT3	AINTEGUMENTA-like 6
Eucgr.H01112	M	T	235,033	143,42	-0,71261	5,00E-05	ATUXS2,AUD1,UXS2	NAD(P)-binding Rossmann-fold superfamily protein
Eucgr.F04160	M	T	34,4695	20,8319	-0,72652	5,00E-05	LAC14	Laccase 14
Eucgr.J01963	M	T	16,6683	9,91276	-0,74975	5,00E-05	UGT76E2	UDP-glucosyl transferase 76E2
Eucgr.E01104	M	T	13,6901	7,84611	-0,80308	5,00E-05	ATCAD1,CAD1, CAD15	Cinnamyl-alcohol dehydrogenase
Eucgr.F04044	M	T	15,226	8,70275	-0,80699	5,00E-05	bHLH121	Basic helix-loop-helix (bhlh) DNA-binding superfamily protein
Eucgr.K01643	M	T	28,8859	16,4039	-0,81633	5,00E-05	BZIP45,TGA6	TGACG motif-binding factor 6
Eucgr.F04163	M	T	11,3525	6,28031	-0,8541	5,00E-05	LAC14	Laccase 14
Eucgr.C01000	M	T	0,972918	0,534842	-0,86321	0,00035		ABC transporter family protein
Eucgr.G03199	M	T	86,8635	46,2011	-0,91082	5,00E-05	CYP98A3	C3H4 cytochrome P450, family 98, subfamily A,

								polypeptide 3
Eucgr.I00493	M	T	24,8714	13,1705	-0,91718	5,00E-05	UGT73B3	UDP-glucosyl transferase 73B3
Eucgr.H01094	M	T	704,646	372,671	-0,91899	5,00E-05	ATSUS3,SUS3	Sucrose synthase 3
Eucgr.I00492	M	T	1,08325	0,570939	-0,92396	0,0069	UGT73B3	UDP-glucosyl transferase 73B3
Eucgr.A02320	M	T	11,4927	5,99406	-0,93911	5,00E-05	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.A02709	M	T	6,24868	3,2425	-0,94645	5,00E-05	UGT88A1	UDP-glucosyl transferase 88A1
Eucgr.H00939	M	T	7,18183	3,63643	-0,98183	5,00E-05	AtCESA1,CESA1,RSW1	Cellulose synthase 1
Eucgr.I01783	M	T	4,69986	2,32162	-1,01748	5,00E-05	CCR9	CCR9 NAD(P)-binding Rossmann-fold superfamily protein
Eucgr.A00929	M	T	8,93921	4,23117	-1,07909	5,00E-05	HYR1	UDP-Glycosyltransferase superfamily protein
Eucgr.J01639	M	T	21,4122	9,55869	-1,16355	5,00E-05	AtCESA1,CESA1,RSW1	Cellulose synthase 1
Eucgr.F02625	M	T	5,69853	2,50776	-1,18419	5,00E-05	ATOMT1,OMT1, COMT37	O-methyltransferase 1
Eucgr.A00926	M	T	3,20964	1,3685	-1,22982	0,0001		UDP-Glycosyltransferase superfamily protein
Eucgr.A00927	M	T	83,9029	34,3504	-1,28839	5,00E-05	HYR1	UDP-Glycosyltransferase superfamily protein
Eucgr.I00409	M	T	39,4245	16,0406	-1,29737	5,00E-05	UGT73B3	UDP-glucosyl transferase 73B3
Eucgr.A02710	M	T	3,37997	1,32044	-1,35599	5,00E-05	UGT88A1	UDP-glucosyl transferase 88A1
Eucgr.D02351	M	T	4,68312	1,75002	-1,4201	5,00E-05	GT72B1,UGT72B1	UDP-Glycosyltransferase superfamily protein
Eucgr.L01345	M	T	26,7676	8,72917	-1,61657	5,00E-05	LAC14	Laccase 14
Eucgr.J01015	M	T	2,88787	0,793886	-1,863	5,00E-05		UDP-Glycosyltransferase superfamily protein

ANEXO K. Tabela contendo as informações estatísticas e descrição funcional dos genes diferencialmente expressos identificados na comparação A vs. M relacionados às rotas de biossíntese dos componentes da parede celular secundária descritos na literatura. Os genes discutidos nesse trabalho estão realçados, com a cor laranja os fatores de transcrição, com a cor azul os genes participantes da rota da celulose e hemicelulose, e com a cor amarelo os genes da rota de biossíntese da lignina.

LocusName	sample_1	sample_2	value_1	value_2	log2 fold_change	p_value	Símbolo	Descrição
Eucgr.E04120	A	M	26,4875	0,569105	-5,54047	5,00E-05		Peroxidase superfamily protein
Eucgr.E03715	A	M	8,76454	0,246959	-5,14933	5,00E-05		Peroxidase superfamily protein
Eucgr.A02818	A	M	35,9262	1,50306	-4,57906	5,00E-05	CBF4,DREB1D	C-repeat-binding factor 4
Eucgr.J02160	A	M	9,45006	0,702275	-3,75021	5,00E-05		ABC transporter family protein
Eucgr.E00850	A	M	88,3068	7,66316	-3,52651	5,00E-05	ATWRKY53,WRKY53	WRKY family transcription factor
Eucgr.A02820	A	M	3,62124	0,488659	-2,88958	5,00E-05	CBF4,DREB1D	C-repeat-binding factor 4
Eucgr.J01614	A	M	2,01339	0,326986	-2,62232	0,005		Peroxidase superfamily protein
Eucgr.B02752	A	M	2,86302	0,490152	-2,54624	5,00E-05	LAC7	Laccase 7
Eucgr.H00890	A	M	76,4054	13,3674	-2,51495	5,00E-05	RCI3,RCI3A	Peroxidase superfamily protein
Eucgr.L02416	A	M	20,918	3,66506	-2,51284	5,00E-05		Peroxidase superfamily protein
Eucgr.J01617	A	M	3,22538	0,58914	-2,45279	5,00E-05		Peroxidase superfamily protein
Eucgr.J01616	A	M	23,9984	4,62425	-2,37565	5,00E-05		Peroxidase superfamily protein
Eucgr.F04198	A	M	22,9923	4,80005	-2,26003	5,00E-05		Peroxidase superfamily protein
Eucgr.J01563	A	M	1,41301	0,298137	-2,24472	0,00025	BHLH92	Basic helix-loop-helix (bhlh) DNA-binding superfamily protein
Eucgr.H03362	A	M	27,0393	5,77224	-2,22785	5,00E-05	ANAC083,NAC083,VNI2	NAC domain containing protein 83
Eucgr.H01245	A	M	1,97579	0,434088	-2,18637	5,00E-05		UDP-Glycosyltransferase superfamily protein
Eucgr.J01618	A	M	33,7321	7,55338	-2,15893	5,00E-05		Peroxidase superfamily protein
Eucgr.C02389	A	M	31,5944	7,17342	-2,13893	5,00E-05		Peroxidase superfamily protein
Eucgr.H00196	A	M	5,49477	1,27066	-2,11248	5,00E-05	BHLH92	Basic helix-loop-helix (bhlh) DNA-binding superfamily protein
Eucgr.D02359	A	M	72,6272	16,8816	-2,10506	5,00E-05	AtWRKY41,WRKY41	WRKY family transcription factor
Eucgr.F03394	A	M	1,10252	0,258241	-2,09402	5,00E-05	BGLU41	Beta glucosidase 41
Eucgr.K00331	A	M	8,17941	1,97064	-2,05334	5,00E-05	AtWRKY41,WRKY41	WRKY family transcription factor
Eucgr.J02352	A	M	22,1916	5,37387	-2,04598	5,00E-05		Peroxidase superfamily protein
Eucgr.A00682	A	M	1,2173	0,295051	-2,04465	0,0039		Peroxidase superfamily protein
Eucgr.C00924,Eucgr.C00925	A	M	75,3667	18,6388	-2,01562	5,00E-05	CCoAOMT4	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
Eucgr.J01619	A	M	33,9043	8,7109	-1,96057	5,00E-05		Peroxidase superfamily protein
Eucgr.F03414	A	M	15,1994	3,98751	-1,93046	5,00E-05	PAT1	GRAS family transcription factor
Eucgr.J01601	A	M	2,8382	0,801081	-1,82496	5,00E-05	AtMYB43,MYB43	MYB domain protein 43
Eucgr.D00471	A	M	2,84452	0,832332	-1,77295	5,00E-05	ATCAD1,CAD1,CAD5	Cinnamyl-alcohol dehydrogenase
Eucgr.I01116	A	M	3,15342	0,951261	-1,729	5,00E-05	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.I02632	A	M	29,2956	8,99855	-1,70292	5,00E-05	CAMTA3,SR1	Signal responsive 1
Eucgr.D00468	A	M	17,4693	5,36939	-1,70199	5,00E-05	ATCAD1,CAD1,CAD4	Cinnamyl-alcohol dehydrogenase
Eucgr.K02355	A	M	2,74017	0,842472	-1,70157	5,00E-05	UGT74F1	UDP-glycosyltransferase 74 F1
Eucgr.L03642	A	M	27,4322	8,61608	-1,67077	5,00E-05	UGT73B4	UDP-glycosyltransferase 73B4

Eucgr.F04195	A	M	29,0825	9,59157	-1,60031	5,00E-05		Peroxidase superfamily protein
Eucgr.L00520	A	M	3,03772	1,02847	-1,56249	5,00E-05	CAMTA1,EICBP.B	Ethylene induced calmodulin binding protein
Eucgr.D01090 Eucgr.D01091	A	M	5,84579	2,03992	-1,51889	5,00E-05	ATCAD7,CAD7,ELI3,ELI3-1,CAD12	Cinnamyl alcohol dehydrogenase 9
Eucgr.A02463	A	M	1,53681	0,541709	-1,50435	5,00E-05	TBR	Plant protein of unknown function (DUF828)
Eucgr.H00071	A	M	21,4319	7,75833	-1,46594	5,00E-05	BGLU46	Beta glucosidase 46
Eucgr.L02740	A	M	23,2582	8,54271	-1,44497	5,00E-05		Peroxidase superfamily protein
Eucgr.H01923	A	M	1,42972	0,543478	-1,39544	0,00255	GATL2	Galacturonosyltransferase-like 2
Eucgr.I00057	A	M	6,64758	2,53124	-1,39298	5,00E-05	ANAC102,NAC102	NAC domain containing protein 102
Eucgr.A02436	A	M	4,283	1,64761	-1,37825	5,00E-05	bZIP65,TGA10	Bzip transcription factor family protein
Eucgr.H03915	A	M	84,8734	33,9251	-1,32296	5,00E-05	RCI3,RCI3A	Peroxidase superfamily protein
Eucgr.E02950	A	M	18,2087	7,51109	-1,27753	5,00E-05	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.D02223	A	M	5,51504	2,30493	-1,25865	5,00E-05	HB-7,HDG5	Homeobox-7
Eucgr.J01079	A	M	38,7855	16,231	-1,25677	5,00E-05	ATPAL2,PAL2,PAL9	Phenylalanine ammonia-lyase 2
Eucgr.K00786	A	M	65,9893	27,6453	-1,2552	5,00E-05	ATWRKY6,WRKY6	WRKY family transcription factor
Eucgr.I00020	A	M	1,76455	0,740912	-1,25192	0,00035		Peroxidase superfamily protein
Eucgr.H03910	A	M	226,417	96,4575	-1,23101	5,00E-05		Peroxidase superfamily protein
Eucgr.I01120	A	M	14,7063	6,27461	-1,22883	5,00E-05	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.K03008	A	M	22,1053	9,43292	-1,22861	5,00E-05	LAC7	Laccase 7
Eucgr.G02223	A	M	24,0815	10,2953	-1,22594	5,00E-05	ATCAD9,CAD9,CAD34	Cinnamyl alcohol dehydrogenase 9
Eucgr.J03126	A	M	63,9797	27,6666	-1,20947	5,00E-05	HCT5	Hydroxycinnamoyl-coa shikimate/quinic acid hydroxycinnamoyl transferase
Eucgr.K03009	A	M	43,5061	19,0348	-1,19258	5,00E-05	LAC7	Laccase 7
Eucgr.G02325	A	M	31,6288	14,1725	-1,15815	5,00E-05	CCR7	NAD(P)-binding Rossmann-fold superfamily protein
Eucgr.D01087	A	M	51,0121	23,0663	-1,14506	5,00E-05	ATCAD7,CAD9,ELI3,ELI3-1	Cinnamyl-alcohol dehydrogenase
Eucgr.J01844	A	M	246,333	114,539	-1,10477	5,00E-05	ATC4H,C4H1,CYP73A5	Cinnamate-4-hydroxylase
Eucgr.H02855	A	M	75,8304	36,1905	-1,06717	5,00E-05	BGLU46	Beta glucosidase 46
Eucgr.E04315	A	M	2,51411	1,21197	-1,05269	0,00015	UGT71C4	UDP-glucosyl transferase 71C4
Eucgr.F03978	A	M	17,2767	8,42998	-1,03522	5,00E-05	HCT4	Hydroxycinnamoyl-coa shikimate/quinic acid hydroxycinnamoyl transferase
Eucgr.J01832	A	M	10,7115	5,26147	-1,02563	5,00E-05	TBL22	Trichome birefringence-like 22
Eucgr.F03595	A	M	5,37418	2,64791	-1,02119	5,00E-05	ATUGT85A1,UGT85A1	UDP-Glycosyltransferase superfamily protein
Eucgr.H03667	A	M	8,72184	4,31486	-1,01532	5,00E-05		Peroxidase superfamily protein
Eucgr.A00485	A	M	7,88942	3,90662	-1,014	5,00E-05	ATGATL1,GATL1,GLZ1,PARVUS	Nucleotide-diphospho-sugar transferases superfamily protein
Eucgr.D02324	A	M	237,14	117,604	-1,0118	5,00E-05	AtWRKY22,WRKY22	WRKY family transcription factor
Eucgr.D01935	A	M	1,98353	0,995015	-0,99528	5,00E-05	IXR11,KNAT7	KNOTTED-like homeobox of Arabidopsis thaliana 7
Eucgr.L00991	A	M	21,0431	10,7211	-0,9729	5,00E-05	ATUGT85A2,UGT85A2	UDP-glucosyl transferase 85A2
Eucgr.F01531	A	M	5,99018	3,0764	-0,96135	5,00E-05	GATL2	Galacturonosyltransferase-like 2
Eucgr.J01372	A	M	268,621	138,126	-0,95958	5,00E-05	UGD3	UDP-glucose 6-dehydrogenase family protein
Eucgr.J01882	A	M	34,8223	18,2938	-0,92866	5,00E-05	GATL10	Galacturonosyltransferase-like 10
Eucgr.J00196	A	M	24,0745	12,6761	-0,9254	5,00E-05	ESK1,TBL29	Plant protein of unknown function (DUF828)
Eucgr.L01657	A	M	8,12345	4,32282	-0,91012	5,00E-05	UGT72E1	UDP-glucosyl transferase 72E1
Eucgr.B03711	A	M	12,8916	6,87137	-0,90776	5,00E-05	ATHKL1,HKL1	Hexokinase-like 1

Eucgr.D00591	A	M	3,38109	1,80322	-0,90692	0,00035	ANAC010,NAC010,SND3	NAC domain containing protein 10
Eucgr.G02921	A	M	128,229	69,0422	-0,89317	5,00E-05	UXS6	UDP-XYL synthase 6
Eucgr.F03635	A	M	8,39425	4,54452	-0,88527	5,00E-05	CESA09,CESA9	Cellulose synthase A9
Eucgr.E04220	A	M	363,005	198,287	-0,8724	5,00E-05	LAC14	Laccase 14
Eucgr.G01977	A	M	1,38671	0,759318	-0,86889	0,0024	ATGUT1,GUT2,IRX10	Exostosin family protein
Eucgr.G02458	A	M	4,39133	2,50587	-0,80935	5,00E-05	BGLU40	Beta glucosidase 40
Eucgr.L01603	A	M	13,511	8,18149	-0,7237	0,0004		Peroxidase superfamily protein
Eucgr.A00936	A	M	11,6836	7,14194	-0,7101	5,00E-05	HYR1	UDP-Glycosyltransferase superfamily protein
Eucgr.F03596	A	M	13,6438	8,36921	-0,70508	5,00E-05	AtUGT85A3,UGT85A3	UDP-glucosyl transferase 85A3
Eucgr.B03943	A	M	3,48788	2,15722	-0,69318	0,00045	4CL7	AMP-dependent synthetase and ligase family protein
Eucgr.H01243	A	M	13,2588	8,20684	-0,69205	5,00E-05	UGT72E1	UDP-glucosyl transferase 72E1
Eucgr.E01115	A	M	10,2935	6,45213	-0,67389	0,00105	ATCAD1,CAD1,CAD20	Cinnamyl-alcohol dehydrogenase
Eucgr.I00608	A	M	6,35167	4,09797	-0,63223	0,0007	bHLH093	Beta HLH protein 93
Eucgr.B03654	A	M	1112,86	718,235	-0,63174	5,00E-05	BGLU17	Beta glucosidase 17
Eucgr.G02852	A	M	4,25164	2,77837	-0,61378	0,001	ATPAL2,PAL2,PAL7	Phenylalanine ammonia-lyase 2
Eucgr.G03092	A	M	7,39342	4,83732	-0,61203	0,0004	TBL33	Trichome birefringence-like 33
Eucgr.F01647	A	M	9,61063	6,29677	-0,61002	0,0001	HXK3	Hexokinase 3
Eucgr.E04217	A	M	53,3616	34,9951	-0,60865	5,00E-05	LAC14	Laccase 14
Eucgr.I00286	A	M	64,6004	44,194	-0,54769	5,00E-05	CESA09,CESA9	Cellulose synthase A9
Eucgr.C00769	A	M	84,1691	58,6699	-0,52067	5,00E-05	ATSUS4,SUS4	Sucrose synthase 4
Eucgr.D00592	A	M	7,25913	5,11214	-0,50587	0,00385	ANAC073,NAC073,SND2	NAC domain containing protein 73
Eucgr.A01397	A	M	90,3664	64,2564	-0,49195	0,0001	ATOMT1,OMT1,COMT1	O-methyltransferase 1
Eucgr.A01385	A	M	291,767	211,31	-0,46546	0,0003	RCI3,RCI3A	Peroxidase superfamily protein
Eucgr.E03226	A	M	19,195	14,3956	-0,4151	0,0018	ANAC073,NAC073,SND2	NAC domain containing protein 73
Eucgr.E04216	A	M	27,3725	20,6298	-0,408	0,0043	LAC14	Laccase 14
Eucgr.J00923	A	M	20,495	15,4724	-0,40558	0,0026	ARF10	Auxin response factor 10
Eucgr.G02469	A	M	18,2494	13,801	-0,40308	0,00395	WRKY3	WRKY DNA-binding protein 3
Eucgr.J00851	A	M	38,4244	29,3685	-0,38776	0,0029	CBB1,DIM1,DWF1,EVE1	Cell elongation protein / DWARF1 / DIMINUTO (DIM)
Eucgr.F02588	A	M	182,795	141,894	-0,36541	0,0039	CINV2	Cytosolic invertase 2
Eucgr.A00966	A	M	470,363	368,018	-0,354	0,00775	RCI3,RCI3A	Peroxidase superfamily protein
Eucgr.C03303	A	M	15,9703	12,5832	-0,34389	0,0102	TBL16	Trichome birefringence-like 16
Eucgr.B00859	A	M	21,5442	17,0258	-0,33958	0,01055	BGLU44	B-S glucosidase 44
Eucgr.H01487	A	M	18,0843	23,3756	0,370267	0,00405	BHLH42,TT8	Basic helix-loop-helix (bhlh) DNA-binding superfamily protein
Eucgr.E01117	A	M	59,9295	77,9469	0,379225	0,00265	ATCAD1,CAD1,CAD21	Cinnamyl-alcohol dehydrogenase
Eucgr.G01757	A	M	13,7009	17,9357	0,388571	0,00775	bHLH121	Basic helix-loop-helix (bhlh) DNA-binding superfamily protein
Eucgr.H04135	A	M	23,3904	31,3415	0,422156	0,0082	AtWRKY42,WRKY42	WRKY family transcription factor
Eucgr.J01963	A	M	12,2063	16,6683	0,449481	0,0017	UGT76E2	UDP-glucosyl transferase 76E2
Eucgr.E01658	A	M	23,1507	32,1006	0,471545	0,00055		Alcohol dehydrogenase transcription factor MYB/SANT-like family protein
Eucgr.B03537	A	M	169,82	235,593	0,47229	0,0003	ANAC083,NAC083,VNI2	NAC domain containing protein 83
Eucgr.B03529	A	M	3,98046	5,61913	0,497413	0,0065	ATTCP20,PCF1,TCP20	TEOSINTE BRANCHED 1, cycloidea, PCF (TCP)-domain family protein 20

Eucgr.H01112	A	M	164,086	235,033	0,518408	0,0001	ATUXS2,AUD1,UXS2	NAD(P)-binding Rossmann-fold superfamily protein
Eucgr.K00043	A	M	4,16208	6,02502	0,533662	0,00425	NACA2	Nascent polypeptide-associated complex subunit alpha-like protein 2
Eucgr.K00405	A	M	4,73433	6,95923	0,555767	0,0021	DAG1	DOF-type zinc finger DNA-binding family protein
Eucgr.F02648	A	M	4,84157	7,14799	0,562061	0,00035	LAC14	Laccase 14
Eucgr.F03499	A	M	52,7276	80,5564	0,611442	5,00E-05	ATERF3,ERF3	Ethylene responsive element binding factor 3
Eucgr.D00264	A	M	3,42285	5,24946	0,616969	5,00E-05	ARF8,ATARF8	Auxin response factor 8
Eucgr.H00939	A	M	4,58717	7,18183	0,646749	0,0022	AtCESA1,CESA1,RSW1	Cellulose synthase 1
Eucgr.B02222 Eucgr.B02223	A	M	45,8538	72,1424	0,653806	0,0001	CCR3	Thioredoxin superfamily protein
Eucgr.A00926	A	M	2,01944	3,20964	0,668457	0,00375		UDP-Glycosyltransferase superfamily protein
Eucgr.D02603	A	M	2,8784	4,65354	0,693059	0,0079	AtGRF1,GRF1	Growth-regulating factor 1
Eucgr.A01867	A	M	48,4136	79,1329	0,708864	5,00E-05	ATOMT1,OMT1, COMT14	O-methyltransferase 1
Eucgr.H02414	A	M	8,07979	13,2767	0,716507	5,00E-05	ATCAD9,CAD9, CAD37	Cinnamyl alcohol dehydrogenase 9
Eucgr.G01342	A	M	69,8563	118,988	0,768351	5,00E-05	UGT84A1	UDP-Glycosyltransferase superfamily protein
Eucgr.B03858	A	M	65,0178	112,601	0,792315	5,00E-05	UGT88A1	UDP-glucosyl transferase 88A1
Eucgr.A00927	A	M	46,0287	83,9029	0,866186	5,00E-05	HYR1	UDP-Glycosyltransferase superfamily protein
Eucgr.F04163	A	M	6,19206	11,3525	0,874516	5,00E-05	LAC14	Laccase 14
Eucgr.I01783	A	M	2,48196	4,69986	0,921137	5,00E-05	CCR9	NAD(P)-binding Rossmann-fold superfamily protein
Eucgr.F04160	A	M	18,091	34,4695	0,930045	5,00E-05	LAC14	Laccase 14
Eucgr.J01639	A	M	11,037	21,4122	0,956092	5,00E-05	AtCESA1,CESA1,RSW1	Cellulose synthase 1
Eucgr.I00492	A	M	0,553201	1,08325	0,969496	0,00445	UGT73B3	UDP-glucosyl transferase 73B3
Eucgr.D02351	A	M	2,37407	4,68312	0,980107	5,00E-05	GT72B1,UGT72B1	UDP-Glycosyltransferase superfamily protein
Eucgr.I00564	A	M	21,6135	43,5095	1,0094	5,00E-05	AIL6,PLT3	AINTEGUMENTA-like 6
Eucgr.F03599	A	M	34,3571	69,2342	1,01088	5,00E-05	AtUGT85A3,UGT85A3	UDP-glucosyl transferase 85A3
Eucgr.K03266	A	M	15,0841	30,56	1,01862	5,00E-05	ATERF1,ERF1	Ethylene response factor 1
Eucgr.F02624	A	M	296,414	604,425	1,02795	5,00E-05	ATOMT1,OMT1	O-methyltransferase 1
Eucgr.I00493	A	M	12,145	24,8714	1,03412	5,00E-05	UGT73B3	UDP-glucosyl transferase 73B3
Eucgr.A02320	A	M	5,5159	11,4927	1,05905	5,00E-05	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.A02314	A	M	0,648908	1,36459	1,07238	0,0024	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.I00409	A	M	18,2941	39,4245	1,10772	5,00E-05	UGT73B3	UDP-glucosyl transferase 73B3
Eucgr.A02709	A	M	2,67907	6,24868	1,22182	5,00E-05	UGT88A1	UDP-glucosyl transferase 88A1
Eucgr.F02623	A	M	0,891274	2,28714	1,3596	5,00E-05	ATOMT1,OMT1, COMT35	O-methyltransferase 1
Eucgr.L01345	A	M	10,0528	26,7676	1,41289	5,00E-05	LAC14	Laccase 14
Eucgr.A02710	A	M	1,25733	3,37997	1,42665	5,00E-05	UGT88A1	UDP-glucosyl transferase 88A1
Eucgr.A00929	A	M	3,28164	8,93921	1,44573	5,00E-05	HYR1	UDP-Glycosyltransferase superfamily protein
Eucgr.F03543	A	M	8,85407	24,4902	1,46779	5,00E-05	OPCL1, 4CL9	OPC-8:0 coa ligase1
Eucgr.F02625	A	M	1,80085	5,69853	1,66191	5,00E-05	ATOMT1,OMT1, COMT37	O-methyltransferase 1

ANEXO L. Tabela contendo as informações estatísticas e descrição funcional dos genes diferencialmente expressos identificados na comparação A vs. T relacionados às rotas de biossíntese dos componentes da parede celular secundária descritos na literatura. Os genes discutidos nesse trabalho estão realçados, com a cor laranja os fatores de transcrição, com a cor azul os genes participantes da rota da celulose e hemicelulose, e com a cor amarelo os genes da rota de biossíntese da lignina.

LocusName	sample_1	sample_2	value_1	value_2	log2 fold_change	p_value	Símbolo	Descrição
Eucgr.A02318	A	T	0,179219	0,897483	2,32416	0,00975	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.J02484	A	T	0,754298	2,38965	1,66359	5,00E-05		Peroxidase superfamily protein
Eucgr.F04198	A	T	22,9923	65,7248	1,51529	5,00E-05		Peroxidase superfamily protein
Eucgr.E02950	A	T	18,2087	48,3275	1,40821	5,00E-05	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.H00692	A	T	0,475288	1,25095	1,39615	0,00185	ATSAGT1,GT,SAGT1,SGT1,UGT74F2	UDP-glucosyltransferase 74F2
Eucgr.A02314	A	T	0,648908	1,70674	1,39516	0,00015	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.J00889	A	T	0,680041	1,6631	1,29018	5,00E-05	ATLAC15,LAC15,TT10	Laccase/Diphenol oxidase family protein
Eucgr.I01117	A	T	6,6478	15,805	1,24943	5,00E-05	UGT73B3	UDP-glucosyl transferase 73B3
Eucgr.A00940	A	T	9,51505	20,6141	1,11535	5,00E-05	HYR1	UDP-Glycosyltransferase superfamily protein
Eucgr.I01120	A	T	14,7063	30,6615	1,06	5,00E-05	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.B03906	A	T	10,561	22,0023	1,05891	5,00E-05	HCT	Hydroxycinnamoyl-coa shikimate/quinic acid hydroxycinnamoyl transferase
Eucgr.J01468	A	T	19,9883	39,7971	0,993509	5,00E-05	GT72B1,UGT72B1	UDP-Glycosyltransferase superfamily protein
Eucgr.A02820	A	T	3,62124	6,95565	0,941703	5,00E-05	CBF4,DREB1D	C-repeat-binding factor 4
Eucgr.L01657	A	T	8,12345	15,5328	0,93515	5,00E-05	UGT72E1	UDP-glucosyl transferase 72E1
Eucgr.E00037	A	T	5,40111	9,99126	0,887412	5,00E-05	DAG1	Dof-type zinc finger DNA-binding family protein
Eucgr.K03266	A	T	15,0841	27,6826	0,875952	5,00E-05	ATERF1,ERF1	Ethylene response factor 1
Eucgr.F03596	A	T	13,6438	24,9415	0,870303	5,00E-05	AtUGT85A3,UGT85A3	UDP-glucosyl transferase 85A3
Eucgr.F03595	A	T	5,37418	9,78218	0,864112	5,00E-05	ATUGT85A1,UGT85A1	UDP-Glycosyltransferase superfamily protein
Eucgr.F03543	A	T	8,85407	16,0543	0,858547	5,00E-05	4CL9	
Eucgr.F01164	A	T	29,5521	52,288	0,82322	5,00E-05	ATERF-4,ATERF4,ERF4,RAP2.5	Ethylene responsive element binding factor 4
Eucgr.H01249	A	T	32,7951	57,896	0,819982	5,00E-05	UGT72E1	UDP-glucosyl transferase 72E1
Eucgr.A02308	A	T	2,93316	5,05238	0,784509	0,00015	UGT73B4	UDP-glycosyltransferase 73B4
Eucgr.F02642	A	T	0,91994	1,58118	0,781394	0,00115	LAC14	Laccase 14
Eucgr.H03667	A	T	8,72184	14,9746	0,779817	5,00E-05		Peroxidase superfamily protein
Eucgr.B02744	A	T	3,78366	6,27359	0,729508	5,00E-05	AtUGT85A2,UGT85A2	UDP-glucosyl transferase 85A2
Eucgr.A02702	A	T	7,80537	12,7511	0,708089	5,00E-05	UGT88A1	UDP-glucosyl transferase 88A1
Eucgr.K00734	A	T	2,65262	4,29874	0,696498	0,011	ATERF3,ERF3	Ethylene responsive element binding factor 3
Eucgr.H02855	A	T	75,8304	120,258	0,665282	5,00E-05	BGLU46	Beta glucosidase 46
Eucgr.F02650	A	T	28,6945	45,3551	0,66049	5,00E-05	LAC14	Laccase 14
Eucgr.I01882	A	T	34,8223	53,8955	0,630152	5,00E-05	GATL10	Galacturonosyltransferase-like 10
Eucgr.F00659	A	T	41,51	62,3211	0,586261	5,00E-05	ATERF3,ERF3	Ethylene responsive element binding factor 3
Eucgr.C01769	A	T	2,6095	3,90978	0,583315	0,0002	AtCESA1,CESA1,RSW1	Cellulose synthase 1
Eucgr.I01114	A	T	8,41935	12,4118	0,559935	0,0006	IRX15-L	Protein of unknown function (DUF579)
Eucgr.A01867	A	T	48,4136	70,0129	0,532209	0,00015	ATOMT1,OMT1	O-methyltransferase 1 COMT14
Eucgr.B00859	A	T	21,5442	27,9355	0,374799	0,004	BGLU44	B-S glucosidase 44
Eucgr.F02090	A	T	12,2813	15,6124	0,346224	0,0075	ARF5,IAA24,MP	Transcriptional factor B3 auxin-responsive factor AUX/IAA-related

Eucgr.A01397	A	T	90,3664	70,9169	-0,34966	0,0047	ATOMT1,OMT1,COMT1	O-methyltransferase 1
Eucgr.G02469	A	T	18,2494	13,9248	-0,39019	0,00505	WRKY3	WRKY DNA-binding protein 3
Eucgr.I00286	A	T	64,6004	47,6514	-0,43902	0,00035	CESA09,CESA9	Cellulose synthase 9
Eucgr.F03605	A	T	71,1841	52,2293	-0,44669	0,0005	CCR5	NAD(P)-binding Rossmann-fold superfamily protein
Eucgr.H01243	A	T	13,2588	9,65196	-0,45805	0,0064	UGT72E1	UDP-glucosyl transferase 72E1
Eucgr.I02371	A	T	106,392	77,3441	-0,46003	0,0004	CYP84A1,FSH1	Ferulic acid 5-hydroxylase 1
Eucgr.K01643	A	T	24,539	16,4039	-0,58104	5,00E-05	BZIP45,TGA6	TGACG motif-binding factor 6
Eucgr.F03635	A	T	8,39425	5,16059	-0,70187	5,00E-05	CESA09,CESA9	Cellulose synthase 9
Eucgr.J00491	A	T	6,17311	3,29751	-0,90462	5,00E-05	bZIP21,TGA9	Bzip transcription factor family protein
Eucgr.E04315	A	T	2,51411	1,27158	-0,98343	0,0005	UGT71C4	UDP-glucosyl transferase 71C4
Eucgr.H00890	A	T	76,4054	36,1237	-1,08073	5,00E-05	RCI3,RCI3A	Peroxidase superfamily protein
Eucgr.G03199	A	T	103,775	46,2011	-1,16746	5,00E-05	CYP98A3, C3H4	Cytochrome P450, family 98, subfamily A, polypeptide 3
Eucgr.J02352	A	T	22,1916	9,26286	-1,26048	5,00E-05		Peroxidase superfamily protein
Eucgr.C02389	A	T	31,5944	12,9111	-1,29105	5,00E-05		Peroxidase superfamily protein
Eucgr.F03394	A	T	1,10252	0,446455	-1,30422	0,0002	BGLU41	Beta glucosidase 41
Eucgr.F00459	A	T	2,17344	0,879449	-1,30531	0,0002	bHLH093	Beta HLH protein 93
Eucgr.A00485	A	T	7,88942	3,069	-1,36215	5,00E-05	ATGATL1,GATL1,GLZ1,PARVUS	Nucleotide-diphospho-sugar transferases superfamily protein
Eucgr.F00351	A	T	2,28398	0,878197	-1,37893	5,00E-05	ATUGT85A1,UGT85A1	UDP-Glycosyltransferase superfamily protein
Eucgr.H00196	A	T	5,49477	2,04102	-1,42877	5,00E-05	BHLH92	Basic helix-loop-helix (bhlh) DNA-binding superfamily protein
Eucgr.C00924,Eucgr.C00925	A	T	75,3667	26,0172	-1,53446	5,00E-05	CCoAOMT4	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
Eucgr.H01245	A	T	1,97579	0,571416	-1,78981	5,00E-05		UDP-Glycosyltransferase superfamily protein
Eucgr.H03915	A	T	84,8734	24,3591	-1,80085	5,00E-05	RCI3,RCI3A	Peroxidase superfamily protein
Eucgr.H03910	A	T	226,417	61,383	-1,88307	5,00E-05		Peroxidase superfamily protein
Eucgr.J01563	A	T	1,41301	0,269229	-2,39186	0,00015	BHLH92	Basic helix-loop-helix (bhlh) DNA-binding superfamily protein
Eucgr.K02355	A	T	2,74017	0,505585	-2,43824	5,00E-05	UGT74F1	UDP-glycosyltransferase 74 F1
Eucgr.E03715	A	T	8,76454	0,851761	-3,36316	5,00E-05		Peroxidase superfamily protein
Eucgr.E04120	A	T	26,4875	2,45362	-3,43233	5,00E-05		Peroxidase superfamily protein
Eucgr.J02160	A	T	9,45006	0,802429	-3,55788	5,00E-05		ABC transporter family protein