

MARCELA MENDES SALAZAR

***“CARACTERIZAÇÃO DO TRANSCRIPTOMA E PAREDE
CELULAR DE TRÊS ESPÉCIES DE EUCALYPTUS COM
IMPORTÂNCIA INDUSTRIAL.”***

**CAMPINAS
2012**

UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

MARCELA MENDES SALAZAR



***“CARACTERIZAÇÃO DO TRANSCRIPTOMA E PAREDE
CELULAR DE TRÊS ESPÉCIES DE EUCALYPTUS COM
IMPORTÂNCIA INDUSTRIAL”***

Tese apresentada ao Instituto de
Biologia para obtenção do Título
de Doutora em Genética e
Biologia Molecular, na área de
Genética Vegetal e
Melhoramento.

| |
|--|
| Este exemplar corresponde à redação final da tese defendida pelo(a) candidato (a) <i>Marcela Mendes Salazar</i> e aprovada pela Comissão Julgadora. |
|--|

Orientador: Prof. Dr. Gonçalo Amarante Guimarães Pereira.

CAMPINAS

2012

FICHA CATALOGRÁFICA ELABORADA POR
ROBERTA CRISTINA DAL' EVEDOVE TARTAROTTI – CRB8/7430
BIBLIOTECA DO INSTITUTO DE BIOLOGIA - UNICAMP

Sa31c Salazar, Marcela Mendes, 1981-
Caracterização do transcriptoma e parede celular de
três espécies de *Eucalyptus* com importância industrial /
Marcela Mendes Salazar. – Campinas, SP: [s.n.], 2012.

Orientador: Gonçalo Amarante Guimarães Pereira.
Tese (doutorado) – Universidade Estadual de
Campinas, Instituto de Biologia.

1. *Eucalyptus*. 2. Madeira. 3. Transcriptoma. 4.
Parede celular vegetal. 5. Glicômica. I. Pereira,
Gonçalo Amarante Guimarães. II. Universidade
Estadual de Campinas. Instituto de Biologia. III. Título.

Informações para Biblioteca Digital

Título em Inglês: Characterization of the transcriptome and cell wall of three
Eucalyptus species with industrial importance

Palavras-chave em Inglês:

Eucalyptus

Wood

Transcriptome

Plant cell walls

Glycomics

Área de concentração: Genética Vegetal e Melhoramento

Titulação: Doutora em Genética e Biologia Molecular

Banca examinadora:

Gonçalo Amarante Guimarães Pereira [Orientador]

Marcos Antonio Machado

Fabio Marcio Squina

Marília Gaspar

Giancarlo Pasquali

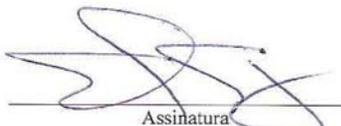
Data da defesa: 02-08-2012

Programa de Pós Graduação: Genética e Biologia Molecular

Campinas, 02 de agosto de 2012.

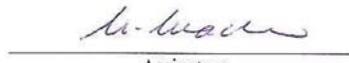
BANCA EXAMINADORA

Prof. Dr. **Gonçalo Amarante Guimarães Pereira**
(Orientador)



Assinatura

Prof. Dr. **Marcos Antonio Machado**



Assinatura

Prof. Dr. **Renato Vicentini dos Santos**

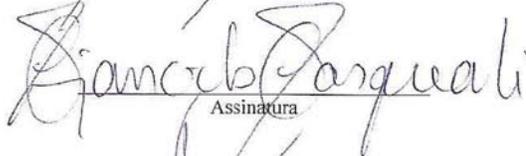
Assinatura

Prof(a). Dr(a). **Marilia Gaspar**



Assinatura

Prof. Dr. **Giancarlo Pasquali**



Assinatura

Prof. Dr. **Fabio Marcio Squina**



Assinatura

Prof. Dr. **Carlos Augusto Colombo**

Assinatura

Prof. Dr. **Marcelo Menossi Teixeira**

Assinatura

Aos meus pais, Sidnei Salazar e Regina
Salazar, e ao meu noivo Ivan, dedico.

À minha avó Olga, ofereço.

AGRADECIMENTOS

Ao professor Dr. Gonçalo A. G. Pereira por me conceder a oportunidade e orientação.
À FAPESP pelo apoio financeiro.

À International Paper do Brasil pelo apoio financeiro e disponibilização do material de estudo.

À UNICAMP e à pós-graduação do Instituto de Biologia por toda estrutura e pelo ótimo nível do curso de doutorado.

A todos do Laboratório de Genômica e Expressão (LGE) da UNICAMP, especialmente ao grupo de pesquisa do Eucalipto e aos bioinformatas Marcelo e Leandro pelas contribuições para esse trabalho. E aos meus amigos Alinne, Bruna, Silvia, Vanessa, Desis e Bruno pelo companherismo e por tornar tudo mais fácil.

À Eli, Sil, D.^a Ernê, Mari, Welbe e Bruno por toda a ajuda e prestatividade.

A todos do *Complex Carbohydrate Reserch Center* (CCRC), especialmente à Sivakumar Pattathil e Dr. Michael G. Hahn e também ao pessoal do LAFIECO (USP), em especial Adriana Grandis e Dr. Marcos Buckeridge pelas preciosas contribuições.

A todas as minhas amigas e amigos de Campinas e Goiânia que são tão especiais e fundamentais.

Agradeço à minha família, em especial meus pais pelo esforço, por todo carinho e dedicação nesses anos de doutorado e nos que precederam e que me fizeram chegar até aqui.

Ao meu noivo Ivan pela amizade, companherismo e amor que tornou essa conquista muito especial.

Agradeço, enfim, a Deus por permitir que todos esses “grandes ombros” estivessem sempre por perto para que eu pudesse me apoiar!

Sumário

| | |
|--|----|
| Lista de Tabelas | 3 |
| Lista de Figuras..... | 4 |
| Figures list..... | 5 |
| RESUMO..... | 6 |
| ABSTRACT..... | 6 |
| 1. Introdução geral | 8 |
| 2. Objetivos | 9 |
| 3. Revisão Bibliográfica..... | 10 |
| 3.1 O gênero <i>Eucalyptus</i> | 10 |
| 3.2 O Setor Florestal..... | 12 |
| 3.3 Genômica do Eucalipto..... | 14 |
| 3.4 Xilogênese e parede celular | 15 |
| 3.5 Biossíntese dos componentes da parede celular..... | 18 |
| 4. Referências | 29 |
| CAPÍTULO 1: Xylem transcription profiles indicate potential metabolic responses for economically relevant characteristics of <i>Eucalyptus</i> species | 35 |
| Abstract: | 35 |
| Background: | 37 |
| Results and Discussion:..... | 39 |
| Xylem RNA sequencing | 39 |
| Construction of a reference sequence database (<i>EUCANEXT</i>) | 40 |
| Expression analysis..... | 41 |
| Functional analysis..... | 46 |
| Conclusion:..... | 54 |
| Methodology:..... | 55 |
| List of abbreviations | 57 |
| Authors' contributions..... | 57 |
| Authors' information | 58 |
| Acknowledgements | 58 |
| References | 58 |
| CAPÍTULO 2: Caracterização da parede celular secundária de três espécies de <i>Eucalyptus</i> .. | 65 |
| 5. Introdução | 65 |

| | |
|--|-----|
| 6. Metodologia:..... | 67 |
| 7. RESULTADOS E DISCUSSÃO | 74 |
| 8. Conclusão | 87 |
| 9. Referências | 89 |
| CONCLUSÃO GERAL: | 92 |
| ANEXOS | 94 |
| CAPÍTULO 1.2: Avaliação do efeito da super-expressão de genes chave de <i>Eucalyptus</i> em planta modelo (<i>Arabidopsis thaliana</i>)..... | 107 |
| 10. Introdução | 107 |
| 11. Metodologia..... | 109 |
| 12. Resultados Preliminares | 114 |
| 13. Perspectivas..... | 117 |
| 14. Referências bibliográficas | 117 |

Lista de Tabelas

| | |
|---|-----|
| Tabela 1: Rendimento e rotação das florestas plantadas no setor de papel e celulose (adaptado de Bracelpa, 2011)..... | 13 |
| Tabela 2: Porcentagem e desvio padrão dos resíduos insolúveis em álcool tratados com DMSO | 76 |
| Tabela 3: Primers utilizados para avaliação da expressão por Real-Time PCR..... | 112 |
| Tabela 4: Resumo do andamento dos experimentos para obtenção de transformantes de <i>A. thaliana</i> .. | 116 |

Lista de Figuras

| | |
|---|----|
| Figura 1: Histologia para demonstrar as principais estruturas que compõem o córtex vascular | 17 |
| Figura 2: Estrutura da parede celular de plantas, evidenciando as diversas camadas que formam a parede celular secundária. | 18 |
| Figura 3: Representação esquemática da parede celular secundária lignificada. | 18 |
| Figura 4: Diferentes níveis de organização das cadeias de de celulose..... | 19 |
| Figura 5: Síntese de microfibras de celulose pelos complexos de celulose sintase presentes na membrana plasmática da célula. | 20 |
| Figura 6: Modelo hipotético da sacarose sintase associada ao complexo celulose sintase..... | 21 |
| Figura 7: Monossacarídeos constituintes de hemiceluloses e pectinas..... | 22 |
| Figura 8: A. Estrutura dos principais tipos de hemicelulose..... | 22 |
| Figura 9: Estrutura geral das pectinas: A. Homogalacturonano. B. Ramnogalacturonano I (sem ramificações). C. Ramnogalacturonano II.. | 26 |
| Figura 10: Fluxograma dos experimentos realizados..... | 67 |
| Figura 11: Cultivo de plântulas de eucalipto em casa de vegetação | 67 |
| Figura 12: Anticorpos monoclonais e respectivas epítopes correspondentes à polissacarídeos presentes na parede celular. | 72 |
| Figura 13: Gráfico “ <i>box plot</i> ” das médias das medidas de peso em mg (Y), com desvio padrão das 3 medidas para cada espécie; A- celulose total; B – lignina Klason. | 76 |
| Figura 14: Gráfico “ <i>box plot</i> ” das medidas das quantificações µg/mg de amostra total para as três espécies..... | 77 |
| Figura 15: Gráfico “ <i>box plot</i> ” das razões observadas para os diferentes monômeros de lignina (A – S/G, B – H/G) e pentoses e hexoses (C)..... | 78 |
| Figura 16: Peso em mg (eixo y) dos extratos obtidos por de <i>Glycome Profiling</i> a partir de AIR de folha de <i>E. grandis</i> e xilema de <i>E. urophylla</i> | 79 |
| Figura 17: <i>Heat map</i> das frações extraídas de xilema de três espécies de <i>Eucalyptus</i> e folha de <i>E. grandis</i> | 80 |
| Figura 18: <i>Heat maps</i> das frações extraídas de xilema de três espécies de <i>Eucalyptus</i> e folha de <i>E. grandis</i> . Média das amostras para as frações “Oxalato de amônio” e “Carbonato”. | 81 |
| Figura 19: Estrutura dos homogalacturonanos evidenciando interação entre as cadeiras através do íon cálcio..... | 82 |
| Figura 20: Interações entre os polissacarídeos pécticos..... | 82 |
| Figura 21: <i>Heat maps</i> das frações extraídas de xilema de três espécies de <i>Eucalyptus</i> e folha de <i>E. grandis</i> | 83 |
| Figura 22: <i>Heat maps</i> das frações extraídas de xilema de três espécies de <i>Eucalyptus</i> e folha de <i>E. grandis</i> . Média das amostras para as frações “Clorito de Sódio” e “4 M KOH PC”..... | 84 |

| | |
|---|-----|
| Figura 23: Modelo proposto de arquitetura da parede celular secundária de três espécies de <i>Eucalyptus</i> , com base nos resultados obtidos nesse capítulo..... | 88 |
| Figura 24: Visão geral de uma célula vegetal com as principais vias de biossíntese de celulose, hemicelulose e lignina. | 93 |
| Figura 25: Método de coleta de amostras de xilema de eucalipto.. | 109 |
| Figura 26: Cultivo de <i>Arabidopsis thaliana</i> em casa de vegetação..... | 110 |
| Figura 27: Vetores utilizados na clonagem: pFP101 e pCambia1302. | 114 |
| Figura 28: Expressão do gene UNK875 em bibliotecas de RNAseq (A) (Capítulo 1), Real-Time PCR (C) e PCR do cDNA (B): 1 e 2- PCR do cDNA de <i>E. globulus</i> , 3 – controle negativo, 4- Marcador de peso molecular 100bp (Invitrogen)..... | 115 |
| Figura 29: Imagem de sementes de <i>A. thaliana</i> borrifadas com <i>Agrobacterium</i> contendo gene glutamina sintetase obtida com microscópio com fluorescência. Filtro azul, aumento de 8X. | 116 |

Figures list

| | |
|--|----|
| Figure 1: Pipeline: Xylem transcriptome mapping and <i>de novo</i> assembly pipeline. | 40 |
| Figure 2: Venn diagrams showing the distribution of contigs between the <i>Eucalyptus</i> xylem libraries: Distributions of contigs between species..... | 41 |
| Figure 3: Cell wall-related genes: The expression profiles of selected genes related to cell wall construction..... | 47 |
| Figure 4: Simplified phenylpropanoid pathway and related genes..... | 49 |
| Figure 5: Important selected genes: The expression profiles of selected..... | 53 |
| Figure 6: Resistance-associated genes..... | 54 |

RESUMO

A celulose é o polímero mais abundante do planeta e está presente principalmente na parede secundária de células vegetais maduras. O conhecimento dos mecanismos moleculares envolvidos na sua biossíntese, bem como a regulação deste processo é recente e ainda elementar, apesar da sua grande importância. O *Eucalyptus* é o gênero florestal mais plantado no mundo, especialmente por ser matéria-prima para fabricação de celulose. Investimentos em pesquisas e desenvolvimento estão sendo realizados pelo setor florestal, no intuito de aumentar os ganhos de produtividade de celulose através da pesquisa na área da biologia molecular. Os dados analisados desmostram que a expressão gênica das espécies estudadas (*E. globulus*, *E. grandis* e *E. urophylla*) difere em genes essenciais para a formação dos compostos da parede celular e por fim da madeira. Os dados moleculares são condizentes com dados diferenciais aqui encontrados em relação à composição dos açúcares da parede. Os resultados gerados pela análise de composição da parede por diferentes técnicas, incluindo a recentemente desenvolvida “Perfil Glicômico” mostram uma grande diversidade e quantidade dos carboidratos da parede celular diferentemente distribuídos no xilema das espécies *Eucalyptus globulus*, *grandis* e *urophylla*, bem como em folha. Tais resultados apresentam um grande avanço para o entendimento da composição das paredes celulares de tais espécies. Assim, o objetivo do presente trabalho foi correlacionar os dados moleculares, com dados *highthroughput* do Perfil Glicômico para entender a composição e arquitetura da parede celular. Espera-se que esses dados possam contribuir para o entendimento da xilogênese e de como os genes trabalham para gerar árvores com características tão distintas, podendo direcionar o melhoramento das mesmas.

ABSTRACT

Cellulose is the most abundant polymer in the world and it is present mainly in secondary cell walls of plant mature cells. The knowledge of molecular mechanisms involved in biosynthesis and regulation of this process is recent despite its great importance. The *Eucalyptus* is the most planted forest genus in the world, especially for being raw material for pulp production. Investments in research and development, especially in the molecular biology field, are being carried out by the forestry sector in order to increase pulp productivity. The data analyzed showed that gene expression of the species the tree species studied here (*E. globulus*, *E. grandis* e *E. urophylla*) differs in essential genes of cell wall compounds formation. Molecular data are consistent with differential data found here in relation to the composition of cell wall sugars. The results generated by the wall composition analysis by various techniques, including the recently developed "Glycome Profiling", showed a wide diversity of carbohydrate of cell wall differently distributed in the xylem of the the *Eucalyptus* species. These results are a breakthrough in understanding the cell wall composition of these species. The objective of this study was to correlate molecular data with data highthroughput Glycome Profiling to understand the composition and architecture of the cell walls. It is hoped that these data will contribute to the understanding of wood formation and how genes work to generate trees with such different characteristics.

1. Introdução geral

O *Eucalyptus* é hoje o gênero florestal mais plantado no mundo, especialmente por suas características favoráveis para a produção de madeira, como rápido crescimento e grande plasticidade ambiental. A crescente demanda por produtos florestais nos últimos anos impulsionou o desenvolvimento do setor florestal. Além disso, as florestas plantadas de *Eucalyptus* são uma alternativa às florestas nativas, reduzindo os impactos ambientais. Assim, as florestas plantadas conquistaram cerca de 90 países, sendo que no Brasil, aproximadamente 4,7 milhões de hectares são ocupados por *Eucalyptus*. Com isso o Brasil conquistou a primeira posição no mercado florestal como o maior produtor e exportador de celulose de *Eucalyptus*, especialmente pelo fato do *Eucalyptus* plantado no país possuir um alto rendimento quando comparado aos de outros países. Esta alta produtividade também se deve à investimentos em pesquisa e desenvolvimento, incluindo programas de hibridação e clonagem e também às boas condições do clima e do solo brasileiro. Entretanto, ainda existem características inerentes à biologia das árvores que dificultam esse processo.

O sequenciamento do genoma e o estudo da expressão de genes de *Eucalyptus* surgem, então, como uma alternativa para se alcançar progressos de forma rápida, manipular e estudar caracteres até então desconhecidos ou que não podiam ser manipulados pelo do emprego das técnicas de melhoramento convencional. Estes dados servem de ponto de partida para diversos estudos funcionais de interesse. Especialmente os processos moleculares envolvidos na formação da madeira, fonte de biomassa para indústrias de papel e celulose, e genes particulares relacionados com tais processos.

O gênero *Eucalyptus* é composto por mais de 600 espécies, sendo três delas consideradas importantes por suas características contrastantes em relação à crescimento, adaptabilidade e qualidade da madeira: *E. globulus*, *E. grandis* e *E. urophylla*. Apesar da sua grande importância, ainda existe falta de informações no que diz respeito às principais diferenças moleculares do xilema destas. Tais diferenças na expressão gênica podem influenciar propriedades físicas da madeira, que são o resultado direto da estrutura da parede celular, da orientação celular, do tipo de células presentes, sua distribuição, disposição e as relativas proporções na qual estão presentes.

A parede celular é composta por, em média, 90% de açúcar. As propriedades físicas de cada parede dependem da combinação de polímeros nela existentes. Relacionar a expressão gênica com análises fenotípicas é de extrema importância para entender os mecanismos

envolvidos na formação da parede celular e que são responsáveis por gerar as diferenças observadas entre as espécies.

Assim, o objetivo deste trabalho foi (1) identificar genes/metabolismos diferencialmente expressos potencialmente relacionados com as características de interesse econômico nas três espécies de *Eucalyptus* estudadas, (2) estudar o efeito da super-expressão de genes candidatos em organismo modelo (*Arabidopsis thaliana*) e (3) caracterizar e identificar as diferenças na composição de açúcares da parede celular do xilema de *E. globulus*, *E. grandis* e *E. urophylla*.

No capítulo 1 descreve-se a geração e sequenciamento de bibliotecas de RNA mensageiro, geradas a partir de xilema em desenvolvimento de três espécies de *Eucalyptus*. Estas bibliotecas foram comparadas buscando a expressão diferencial de genes e de metabolismos relacionados com as características de cada uma destas espécies. Os resultados mostram que há a expressão diferencial de 5.909 genes e dentre estes, discutiu-se sobre os genes relacionados com a formação da parede celular (incluindo genes de lignificação), fatores de transcrição, genes do metabolismo de nitrogênio, ubiquitinas, *heat shock proteins* e genes relacionados à resposta a *stress*.

Genes selecionados a partir da análise do transcriptoma (capítulo 1) foram clonados e super-expressos em organismo modelo (*Arabidopsis thaliana*). Os dados dos experimentos ainda em andamento estão mostrados em anexo do capítulo 1.

No capítulo 2 avalia-se a composição, quantificação e estruturação das moléculas da parede celular do xilema das três espécies estudadas. Propoe-se também um modelo de arquitetura da parede celular das três espécies com base nos resultados encontrados neste capítulo. Tal modelo sugere que uma parede celular mais frouxa está presente em *E. globulus* e uma parede mais comprimida em *E. urophylla*.

2. Objetivos

- Identificar genes/metabolismos diferencialmente expressos potencialmente relacionados com as características de interesse econômico nas três espécies de *Eucalyptus* estudadas.
 - Estudar o efeito da super-expressão de genes candidatos em organismo modelo (*Arabidopsis thaliana*).

- Caracterizar e identificar as diferenças na composição de açúcares da parede celular do xilema de *E. globulus*, *E. grandis* e *E. urophylla*.

3. Revisão Bibliográfica

3.1 O gênero *Eucalyptus*

O gênero *Eucalyptus* pertence ao grande grupo das angiospermas, família *Myrtaceae*. É nativo da Austrália cobrindo 90% da área de todo o país, possui cerca de 700 espécies identificadas com numerosas subespécies, variedades botânicas e híbridos, sendo aproximadamente 200 delas descritas recentemente (por volta de 15 anos). Apenas duas espécies possuem o centro de origem fora da Austrália (*Eucalyptus urophylla* e *Eucalyptus deglupta*) (Dickinson, 1969; Brooker, 2000; Bertola, 2005; REMADE, 2008).

O *Eucalyptus* foi introduzido no Brasil por volta de 1868, a princípio como planta decorativa e somente a partir de 1950 foi utilizado como matéria prima para a produção de celulose (Lima, 1996; Bertola, 2005). Hoje, as espécies de *Eucalyptus* são aproveitadas em todas as cadeias produtivas: sua madeira é utilizada para a extração de celulose, produção de móveis e chapas e no setor de carvão vegetal; o óleo extraído de suas folhas é utilizado pela indústria alimentícia, farmacêutica e cosmética; o pólen de suas flores é utilizado na produção de mel, entre outros usos (Lima, 1996; Bertola, 2005). O *Eucalyptus* hoje é cultivado em diversos países de clima temperado e tropical, ocupando regiões de grande diversidade ambiental tanto em termos de precipitação quanto de temperaturas, ocorrendo ainda em regiões tropicais equatoriais até regiões semi-áridas e em elevações que vão desde o nível do mar até mais ou menos 4.000 metros de altitude. Tal versatilidade do gênero possibilitou a expansão de sua produção em diferentes regiões do território brasileiro, bem como em todo o mundo (Lima, 1996). Do ponto de vista genético, essa versatilidade se dá em um grande número de alelos para diferentes características, indicando o forte potencial para o melhoramento, principalmente pela da combinação de espécies, apesar do genoma relativamente pequeno (370-700 Mpb) (Foucart *et al.*, 2006).

Suas características favoráveis para a produção de madeira (facilidade de propagação vegetativa, crescimento rápido, alta produtividade, boa forma de fuste e boa resposta a tratamentos culturais de manejo e melhoramento genético) fizeram com que o *Eucalyptus* se consolidasse

como matéria-prima de base no setor florestal mundial e se tornasse gênero florestal mais plantado no mundo (Mora & Garcia, 2000; Silva, 2005).

Aproximadamente 20 espécies, a maioria pertencente ao subgênero *Symphyomyrtus*, são usadas em plantações comerciais em mais de 90 países (Potts and Pederick, 2000). Dentre estas, três espécies (*E. globulus*, *E. grandis* e *E. urophylla*) são largamente utilizadas em programas de melhoramento genético no Brasil por possuírem características importantes, como descritas abaixo:

- *Eucalyptus globulus* Labill.

É uma árvore de médio a grande porte, 15 a 25m, podendo alcançar 70m de altura (Santos, 2005). É nativa da Tasmânia e sudeste da Austrália. Foi a primeira espécie a se tornar mais largamente conhecida fora da Austrália, sendo sua madeira usada para fabricação de papel, lenha e construção civil (FAO, 2011). Para a indústria de papel e celulose, *E. globulus* se destaca por possuir melhor qualidade da madeira, possuindo alto rendimento de celulose, baixo teor de lignina e maior relação da razão de syringyl/guaiacyl (S/G) (explicado adiante) (Brito *et al.*, 1983, Rodrigues *et al.*, 1999). Entretanto, por não se adaptar bem ao clima tropical (Bison *et al.*, 2007) esta espécie não é largamente utilizada em plantações comerciais no Brasil.

- *Eucalyptus grandis* Hill ex-Mainden

As árvores desta espécie são muito altas (40 a 45m), podendo eventualmente atingir 75m de altura. Possuem tronco grosso, com casca de textura lisa (IPEF - chave de identificação, 2012). Ocorre naturalmente na Austrália, ao norte do Estado de New South Wales, ao sul de Queensland (próximo à região costeira e na parte central) e ao norte de Queensland em área de altitude (300 a 900m). A madeira de *E. grandis* é leve e fácil de ser trabalhada. É utilizada intensivamente, na Austrália e na República Sul Africana, como madeira de construção. Normalmente, a madeira oriunda de árvores com rápido crescimento apresenta problemas de empenamento, contrações e rachaduras. Plantações convenientemente manejadas, no entanto, podem produzir madeira excelente para serraria e laminação. É a principal matéria-prima para as indústrias de celulose, painéis aglomerados e chapas duras. É uma espécie presente em inúmeros países e, de regra geral, apresenta o seguinte comportamento:

- É considerada sensível a geadas severas, porém apresenta resistência a uma vasta gama de outros extremos ambientais;
- Apresenta relativa resistência a deficiência hídrica e doenças;

- A regeneração, por meio da brotação de cepas, é considerada boa;
 - O ritmo de crescimento e o rendimento volumétrico são, geralmente, superiores, quando comparados a outras espécies convencionais. (Potts & Pederick, 2000; REMADE, 2003)
- *Eucalyptus urophylla* S.T. Blake

É uma das duas espécies de ocorrência natural fora do território australiano, ocorrendo, naturalmente, na ilha de Timor e outras ilhas a leste do arquipélago indonésio. A madeira é considerada medianamente leve, onde as propriedades de resistência mecânica são moderadas. É uma madeira de relativa estabilidade e alta permeabilidade. No Brasil, é indicada para usos gerais e muito utilizada em laminação, componentes estruturais para construção, caixotaria, mourões, escoras, celulose e papel, chapas duras, painéis, lenha e carvão. Inúmeros esforços são realizados para a introdução da espécie fora das condições de sua zona natural. Os resultados são bastante satisfatórios, com a espécie apresentando alta plasticidade, adaptando-se a solos hidromórficos ou francamente arenosos, em diferentes altitudes. É considerada apta para regiões onde não ocorrem geadas e situações de déficits hídricos severos. No Brasil, a espécie tem sido plantada intensivamente em programas de melhoramento genético, principalmente de hibridação. É uma espécie que apresenta boa capacidade de regeneração por brotação. Apresenta maior potencial de crescimento em relação à resistência ao fungo causador do cancro do *Eucalyptus* (*Cryphonectria cubensis*) (Turnbull, 2000; Scanavaca Junior, 2002; REMADE, 2003).

3.2 O Setor Florestal

A crescente demanda por produtos florestais impulsionou o desenvolvimento do setor florestal nos últimos 20 anos. Tal desenvolvimento também é acompanhado pela crescente pressão de uma "indústria florestal sustentável", a qual engloba conceitos de melhorias na eficiência energética, conservação de recursos e a utilização de materiais seguros (FAO, 2011). Dentro deste contexto, as florestas plantadas de *Eucalyptus* e *Pinus* surgem como uma alternativa às florestas nativas. Tais florestas são plantadas em locais específicos e atendendo às normas de manejo sustentável, o que permite a redução dos impactos ambientais (Bracelapa, 2011).

As florestas plantadas estão em cerca de 90 países, sendo 6,3 milhões de hectares no Brasil, 74,8% correspondente a plantios de *Eucalyptus* e 25,2% de de *Pinus* (ABRAF, 2012).

Desse total, 2,2 milhões são cultivados pela indústria de celulose e papel (ABRAF, 2010), setor que utiliza 100% da madeira oriunda de florestas plantadas.

Em 2011, o Brasil conquistou a quarta posição como maior produtor de celulose do mundo, posição conquistada em 2008, passando à frente de países tradicionais produtores como a Finlândia e a Suécia. Dentre os primeiros produtores estão: Estados Unidos, Canadá e China. A produção brasileira de celulose apresentou crescimento em 2010 totalizando 14 milhões de toneladas e 9,4 milhões de toneladas de papel apesar da crise global que afetou o setor florestal entre 2008 e 2009 (ABRAF, 2010). O país também se destaca como o maior produtor e exportador de celulose branqueada de eucalipto do mundo. Dentre os principais destinos da celulose exportada estão países como Estados Unidos, China e países da Europa. Já o papel produzido no Brasil, tem como principal destino os países da América Latina (Bracelpa, 2011).

As florestas plantadas para o setor de papel e celulose no país estão entre as que apresentam maior produtividade mundial, devido ao alto rendimento (Tabela 1).

Tabela 1: Rendimento e rotação das florestas plantadas no setor de papel e celulose (adaptado de Bracelpa, 2011)

| Florestas de <i>Eucalyptus</i> | | | Florestas de <i>Pinus</i> | | |
|--------------------------------|-------------------------------------|---------|---------------------------|-------------------------------------|----------------|
| País | rendimento (m ³ /ha/ano) | rotação | País | rendimento (m ³ /ha/ano) | rotação (anos) |
| Brasil | 44 | 7 | Brasil | 38 | 15 |
| África do Sul | 20 | 8 - 10 | EUA | 10 | 25 |
| Portugal | 12 | 12 - 15 | Chile | 22 | 25 |

Além da produção e exportação de celulose, as plantações de *Eucalyptus* começam a se beneficiar das políticas de crédito de carbono, pelos quais países desenvolvidos devem comprar, normalmente de países em desenvolvimento, títulos (CER – certificado de emissões reduzidas) que correspondem a uma obrigação de captura de carbono na mesma proporção em que ele será emitido. As transações desses CERs fazem parte do chamado mercado de carbono (Rocha, 2003). O Brasil possui um lugar de destaque nesse mercado, estando entre os países mais procurados para investimentos em sequestro de carbono, bem como a Índia e a China, pelo fato de possuir grande extensão territorial localizada na zona tropical (Chang, 2002).

O alto nível de produtividade conquistado pelo país deve-se principalmente à investimentos em pesquisa e desenvolvimento, incluindo programas de hibridação e clonagem e também às boas condições do clima e do solo brasileiro (Bertola, 2005; Poke *et al.*, 2005;

Bracelpe, 2011). Entretanto, ainda existem características inerentes à biologia das árvores que dificultam esse processo. Entre elas estão: longo ciclo de vida, porte, maturação tardia, polinização cruzada, alta complexidade da análise dos descendentes após cruzamentos e retrocruzamentos, dificuldade de obtenção de ganhos significativos para caracteres complexos (concentração de lignina e celulose) e necessidade de grandes áreas para plantios experimentais (Campbell *et al.*, 2003; Valverde *et al.*, 2007).

Assim, o surgimento de projetos genoma de *Eucalyptus* proporcionam uma alternativa para se alcançar progressos de forma rápida, manipular e estudar caracteres até então desconhecidos ou que não podiam ser manipulados pelo emprego das técnicas de melhoramento convencional.

As empresas do setor de papel e celulose do país iniciaram investimentos que deverão atingir ao menos US\$ 5,5 bilhões nos próximos anos em pesquisas e desenvolvimento florestal a fim de aumentar a oferta para acompanhar o crescimento da demanda do setor. Os investimentos visam principalmente ganho em produtividade, aumentando também a qualidade da matéria prima (madeira e fibra), otimizando o teor de lignina, o rendimento em celulose, a densidade da madeira e outras características de interesse de cada setor de transformação da madeira (ABRAF, 2010).

Nesse contexto, em fevereiro de 2008, foi firmado um convênio de cooperação entre a UNICAMP e a empresa International Paper do Brasil (convênio Funcamp – International Paper do Brasil IP/IB/Gene Discovery: Convênio 3972) visando principalmente à utilização do banco de dados construídos para *Eucalyptus* para identificação de genes relacionados à formação e qualidade da madeira de *Eucalyptus* sp.

3.3 Genômica do Eucalipto

Apesar da grande importância do *Eucalyptus* para a indústria florestal, o número de genes caracterizados é pequeno e são principalmente aqueles relacionados com processos do desenvolvimento da planta. A maior parte deles está concentrada em genes envolvidos na formação da madeira, especialmente aqueles ligados à biossíntese de lignina (Grima-Pettenati & Goffner, 1999; Baucher *et al.*, 2003). O sequenciamento de ESTs proporcionou um grande avanço para o isolamento e identificação de determinados genes, além do estudo de sua expressão em um tecido particular (Hatey *et al.*, 1998).

Diversos projetos de sequenciamento de genomas expressos (*Expressed Sequence Tags* ou EST) de *Eucalyptus* já foram realizados por alguns grupos no mundo. No Brasil, os projetos FOREST (*Eucalyptus Genome Sequencing Project Consortium*)/FAPESP (Vicentini *et al.*, 2005) e GENOLYPTUS (Rede Brasileira de Pesquisa do Genoma do Eucalipto)/MCT) são os maiores.

Novas tecnologias de análises do perfil de transcriptoma vêm surgindo para melhorar o entendimento da extensão e complexidade dos transcriptomas de eucariotos. São baseadas na seleção do mRNAs pelas caudas poli-A e sua purificação de outros RNAs. Geralmente, elas podem ser desenvolvidas de duas maneiras: o mRNA é fragmentado e transcrito em bibliotecas de cDNA por primers randômicos, ou o RNA é transcrito utilizando *primers* oligo-dT e depois fragmentado, este último garantido a uniformidade de transcritos representados nas bibliotecas. Os cDNAs são então ligados a adaptadores e submetidos aos sequenciadores. Atualmente o RNA-seq vem sendo considerado o novo padrão para estudos em transcriptômica, com inúmeros trabalhos em plantas como *Arabidopsis*, milho e alfafa (Cheung *et al.*, 2006; Emrich *et al.* 2007; Yang *et al.*, 2011) e mais recentemente em *Eucalyptus* (Mizrachi *et al.*, 2010). A técnica fornece uma forma mais precisa de medir os níveis de transcritos e suas isoformas, quando comparado à outras técnicas de análise de transcriptoma, por exemplo microarranjo (Fu *et al.*, 2009). Além do sequenciamento de ESTs, o genoma de *Eucalyptus grandis* e *Eucalyptus globulus* foi recentemente finalizado pelo US Department of Energy (DOE) Joint Genome Institute (JGI) (<http://www.eucagen.org/>).

Todos esses dados servem de ponto de partida para diversos estudos funcionais de interesse. Especialmente os processos moleculares envolvidos na formação da madeira, fonte de biomassa para indústrias de papel e celulose, e genes particulares relacionados com tais processos.

3.4 Xilogênese e parede celular

O xilema das angiospermas é basicamente constituído pelas seguintes partes (Figura 1): (i) medula: porção interna, vestígio do meristema apical do ramo; (ii) xilema primário e secundário: tecido onde ocorre a circulação de água e de micronutrientes (Taiz & Zeiger, 2006). A madeira é uma denominação genérica utilizada para se referir ao tecido vascular de plantas lenhosas - **o xilema secundário**, cujas funções principais são: sustentação da planta e transporte de água e solutos. Em angiospermas, três tipos de células são responsáveis por

desempenhar as principais funções deste tecido: os elementos de vasos que transportam água e solutos; as fibras que fornecem sustentação (ambos possuem parede celular secundária); as células parenquimáticas estão envolvidas na transferência radial de assimilados entre floema e xilema e estoque temporário de amido e lipídeos (Fukuda, 1996; Plomion *et al.*,2001; Dejardin *et al.*, 2010).

A estrutura do xilema secundário é complexa, sendo resultado da diferenciação radial e longitudinal das células que compõem o tecido vascular. Essas células são formadas a partir da atividade de um meristema lateral: o câmbio vascular (Figura 1). Na zona cambial existem dois tipos de células: as iniciais radiais que dão origem aos raios xilemáticos, consistindo de células parenquimáticas, os quais são essenciais para a translocação de nutrientes entre xilema e floema; iniciais fusiformes que se dividem para formar os tecidos vasculares secundários: os elementos de vaso (condutores de água), fibras e células do parênquima axial (Dejardin *et al.*, 2010).

Os elementos de vasos e as fibras são as principais células no xilema secundário de dicotiledôneas com parede celular secundária. Nestas células há um acúmulo de substâncias polimerizadas (celulose, hemicelulose e lignina) e que fazem a madeira ser de grande interesse em diversas cadeias produtivas do setor florestal.

Após a divisão, a diferenciação das células do xilema com parede secundária envolve quatro passos principais até que as células adquiram as características distintivas do seu tipo celular: 1) Expansão; nesta etapa as células jovens crescem sem aumentar sua largura, são constituídas apenas de parede primária; 2) Deposição de parede celular secundária; diversas enzimas agem para modificação da parede primária (xiloglucanas endoglicosiltransferases, expansinas, pectina metil esterases e pectinases) para acomodar a nova parede em formação, a qual é coordenada pela expressão de diversos genes; 3) Lignificação; 4) Morte celular programada: ocorre para formar um tubo vazio com parede secundária (Hertzberg *et al.*, 2001; Plomion *et al.*,2001; Demura & Fukuda, 2007)

A parede celular das plantas está localizada na matriz extracelular vegetal, o apoplasto, o qual também é composto por água e compostos solúveis nela dissolvidos (açúcares, aminoácidos e compostos derivados do metabolismo secundário). As células estão conectadas entre si por uma camada denominada lamela média (formada por pectinas). As células vegetais possuem uma parede celular primária e podem ou não possuir uma parede celular secundária, dependendo do tipo celular. Além disto, sua composição é variável de acordo com o tipo celular e a idade da planta, principalmente no que se refere à presença ou ausência da parede celular

secundária. Células em crescimento, de parênquima, células fotossintetizantes presentes nas folhas e alguns outros tipos celulares possuem somente parede celular primária (Salisbury & Ross, 1991; Raven *et al.*, 2001; Buckeridge *et al.*, 2008).

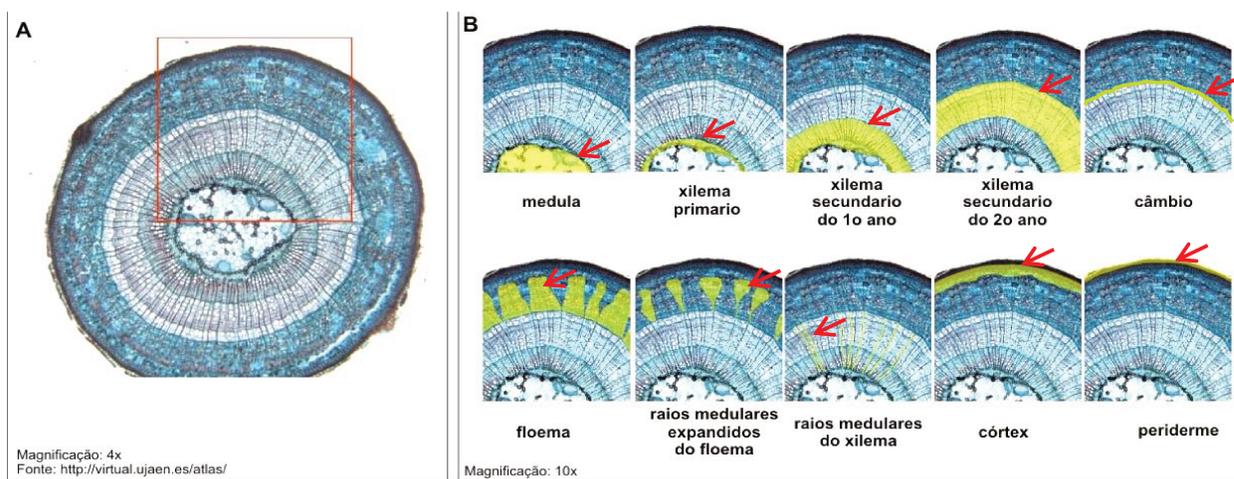


Figura 1: Histologia para demonstrar as principais estruturas que compõem o córtex vascular (gênero *Tília*, família Malvaceae). A. Corte transversal do caule. B. Detalhes do caule, evidenciando (em amarelo e apontado pelas setas vermelhas) cada uma das estruturas (identificadas sob os quadros). Imagem adaptada de <http://virtual.ujaen.es/atlas> (Pedrosa *et al.*)

Na maioria de células maduras, especialmente naquelas presentes em tecidos de xilema, existe uma parede celular secundária (Figura 2) que fornece o suporte para a planta. Este suporte mecânico se dá, principalmente, devido à presença de fibras de constituição celulósica compactadas e lignina (Taiz & Zeiger, 2006). As diferenças nas propriedades físicas de paredes primárias e secundárias devem-se, em grande parte, à quantidade de cadeias de celulose que estão presentes em cada microfibrila. Em paredes secundárias, cada microfibrila possui um número muito maior de cadeias de celulose quando comparadas com o número destas presentes nas microfibrilas constituintes da parede celular primária (Taylor *et al.*, 1999). A parede celular da maioria das células presentes no xilema é composta por diversas camadas de parede celular secundária consistindo de uma rede de celulose e hemicelulose saturada com lignina (Andersson-Gunneras *et al.*, 2006).

A parede celular é composta por microfibrilas de celulose ligadas às cadeias de hemicelulose inseridas em uma matriz de polissacarídeos pécticos e uma quantidade variável de proteínas estruturais (Figura 3). Estes elementos compõem cerca de 75% do peso seco da madeira, sendo o restante composto por lignina. No caso do *Eucalyptus*, a constituição química

varia entre espécies e, em menor escala, entre os exemplares da mesma espécie (Vital,1984; Buckeridge *et al.*, 2008).

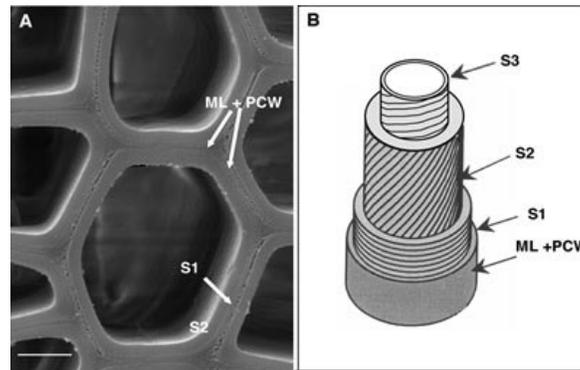


Figura 2: Estrutura da parede celular de plantas, evidenciando as diversas camadas que formam a parede celular secundária. Adaptado de Lodish *et al.* (2000) e Dejardin *et al.*, 2010. Microscopia eletrônica de *Populus sp.* (A) e esquema de interpretação das diferentes camadas da parede celular (B). ML: lamella média; PCW: parede celular primária; S1, S2, S3: camadas da parede celular secundária. A barra corresponde a 5mm.

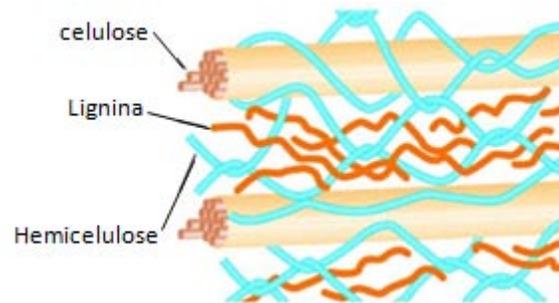


Figura 3: Representação esquemática da parede celular secundária lignificada. Imagem extraída de Boundet *et al.* (2003).

3.5 Biossíntese dos componentes da parede celular

A celulose é o principal constituinte da parede celular, um polissacarídeo hidrofóbico formado pela polimerização de resíduos de glicose, a partir de um substrato de uridina-difosfoglicose (UDP-glicose), formando um homopolímero linear com ligações β -1,4 glicano (Burton *et al.*, 2005). Nas paredes celulares primárias das plantas, a celulose contribui com cerca de 20 a 30% da massa seca (Mellerowicz, 2001). As microfibrilas de celulose das paredes primárias controlam o tamanho e a forma das células vegetais. Quando comparadas às paredes celulares primárias, paredes celulares secundárias contêm celulose com alto grau de polimerização fornecendo força mecânica e rigidez para toda a planta (Samuga & Joshi, 2004).

O processo de biossíntese de celulose é um dos mais importantes da natureza, já que este polissacarídeo é o principal constituinte das paredes da maioria das células maduras (Dhugga, 2001). Porém, em contraste com a grande abundância e importância da celulose, pouco se sabe sobre a regulação dos mecanismos moleculares envolvidos na sua biossíntese em plantas. Sabe-se, entretanto, que a biossíntese deste polímero é um processo complexo e que exige altos níveis de organização. A cadeia formada pelos resíduos de glicose é linear, pois durante a ligação um dos resíduos sofre uma rotação de 180° em relação ao resíduo anterior. Esta forma linear que a celulose assume permite que haja uma grande interação entre cadeias formando uma estrutura rígida. Pontes de hidrogênio formam-se dentro das cadeias e entre elas, de modo que uma estrutura cristalina pode ser observada. Toda esta interação entre as cadeias garante a esta estrutura propriedades notáveis, como: alta resistência à tensão, insolubilidade e a capacidade de ser relativamente imune à degradação enzimática e química (Taiz & Zeiger, 2006).

A estrutura que se forma pela interação íntima entre cadeias de celulose é chamada de microfibrila de celulose. Esta estrutura pode conter de 36 a 200 cadeias, uma variação que ocorre entre espécies. O comprimento das microfibrilas também é bastante variável. A organização das cadeias de celulose também se dá em níveis mais altos. Um conjunto de microfibrilas é chamado de macrofibrilas (Figura 4). A reunião deste último forma um feixe de celulose (Delmer & Amor, 1995; Taiz & Zeiger, 2006).

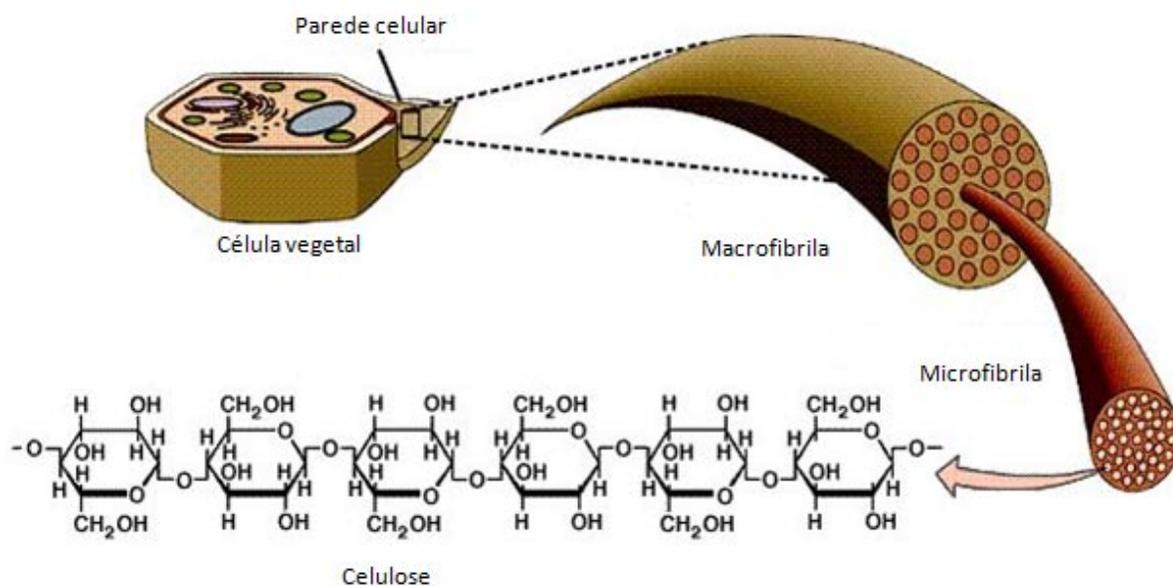


Figura 4: Diferentes níveis de organização das cadeias de de celulose. Imagens adaptada de: Moore *et al.*, 1998.

As microfibrilas de celulose são sintetizadas por um grande complexo proteico (>500 kDa) localizado na membrana celular, denominado Complexo Celulose Sintase. Este complexo, em forma de roseta, é composto por seis unidades globulares organizadas de maneira hexagonal, constituindo as unidades da roseta. Cada uma das unidades da roseta possui seis proteínas CESA (celulose sintase), de modo que, de cada roseta, serão geradas 36 cadeias de celulose (Figura 5). A proteína CESA atua como um membro catalítico do complexo e catalisa a transferência de glicose para a cadeia glicano crescente, utilizando como substrato UDP-glicose. A UDP-glicose liga-se ao sítio ativo no lado citoplasmático da membrana celular, enquanto a cadeia de celulose emerge da membrana para a parede celular (Richmond, 2000; Doblin *et al.*, 2002). Além da subunidade catalítica CESA, acredita-se que o complexo celulose sintase contenha pelo menos outros 15 polipeptídios envolvidos em fornecer o substrato, iniciar ou terminar o alongamento da cadeia de celulose, ou até mesmo na regulação do processo (Doblin *et al.*, 2002; Fujii *et al.*, 2010).

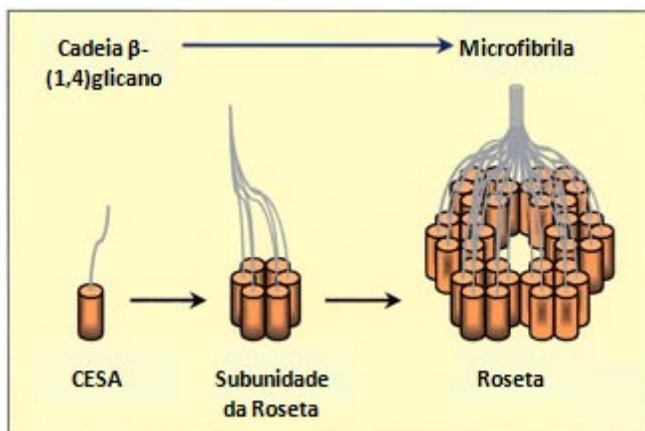


Figura 5: Síntese de microfibrilas de celulose pelos complexos de celulose sintase presentes na membrana plasmática da célula. Detalhe de um complexo, formado por 36 CESA. Figura adaptada de Doblin, *et al.*, 2002

A subunidade catalítica CESA é codificada por genes da família celulose sintase (*CesA*). Um grande número de genes *CesA* (celulose sintase) já foram clonados em diferentes espécies como *Pinus* (Nairn & Haselkorn, 2005), *Populus* (Suzuki, *et al.*, 2006) e *Eucalyptus* (Ranik & Myburg, 2006). A família de genes mais bem caracterizada é a de *A. thaliana*, a qual é formada por pelo menos 10 genes (Samuga & Joshi, 2002). A CESA está presente em todos os tecidos, entretanto, estudos indicam que os vários membros da família gênica são expressos diferencialmente entre tecidos que possuem parede secundária (Richmond, 2000).

Em *A. thaliana* pelo menos três genes da família são necessários para a formação de celulose na parede primária (*AtCesA1*, *AtCesA3* e *AtCesA6*) e três na parede secundária

(*AtCesA4*, *AtCesA7* e *AtCesA8*) (Tanaka *et al.*, 2003). Entretanto, os genes *AtCesA2* e *AtCesA5* parecem ter funções redundantes com o gene *AtCesA6* (Desprez *et al.*, 2007). Em *Eucalyptus* os genes *CesA1*, *CesA2* e *CesA3* são responsáveis pela formação da celulose na parede secundária, sendo o último mais expresso. E os genes *CesA4*, *CesA5* e *CesA6* atuam na formação da parede primária (Ranik & Myburg, 2006).

Apesar da celulose ser o polímero mais abundante do planeta, o funcionamento do complexo celulose sintase ainda não é totalmente entendido. Recentemente, Song *et al.* (2010) detectaram um segundo tipo de complexo celulose sintase em *Populus*, provavelmente associado com a formação de parede secundária. Além disso, outros modelos propostos sugerem que dois peptídeos unem-se para formar sítios catalíticos opostos de maneira tal que os resíduos de UDP-glicose são adicionados à extremidade não redutora da cadeia recém-formada sem que o resíduo tenha que sofrer uma rotação de 180° (Buckeridge *et al.*, 2001). Tais estudos abrem a possibilidade para um novo método de engenharia das microfibrilas de celulose.

A UDP-glicose é o substrato tanto para a síntese de celulose como para hemiceluloses e pectinas. Acredita-se que, no caso da formação da celulose, este substrato seja diretamente fornecido por uma forma da enzima sacarose sintase (SUSY) associada à membrana (Amor *et al.*, 1995; Fujii *et al.*, 2010) (Figura 6). Acredita-se que esta isoforma da SUSY seja um dos principais componentes do complexo celulose sintase (Fujii *et al.*, 2010).

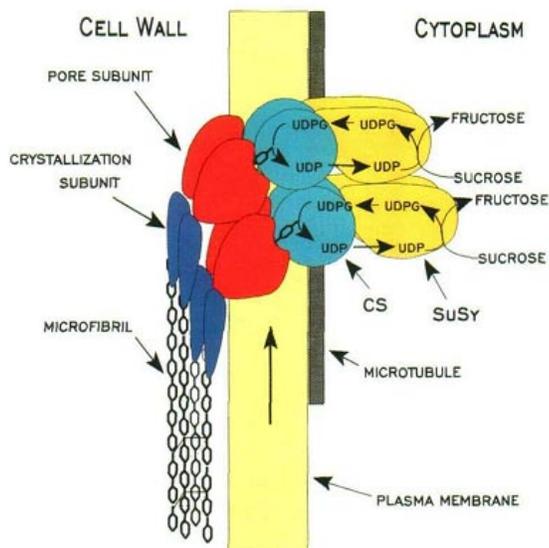


Figura 6: Modelo hipotético da sacarose sintase associada ao complexo celulose sintase. Extraído de Delmer & Amor (1995).

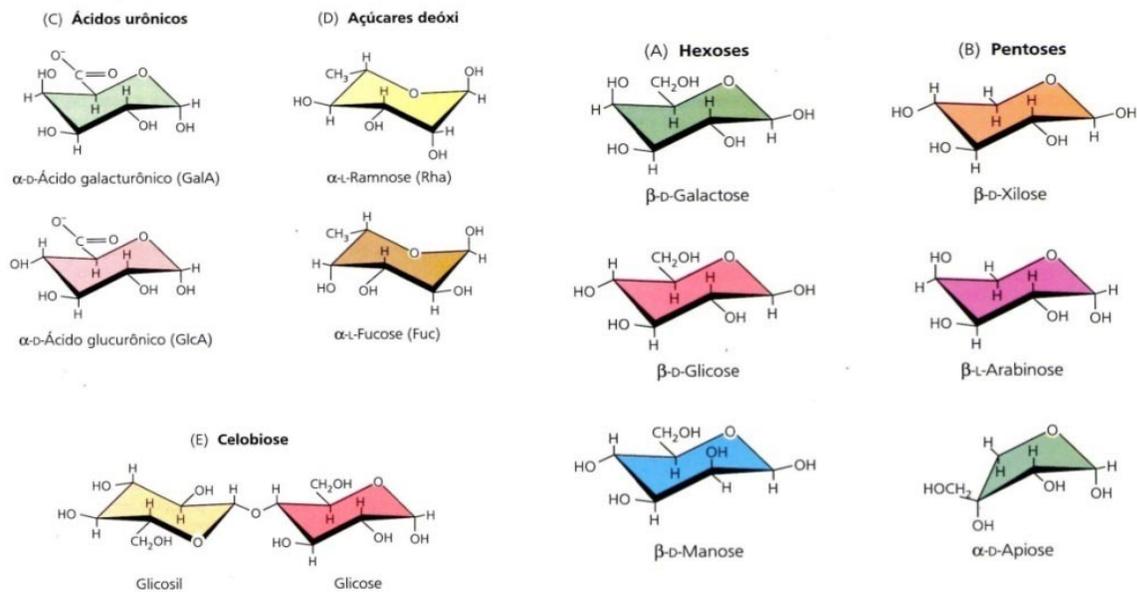


Figura 7: Monossacarídeos constituintes de hemiceluloses e pectinas.

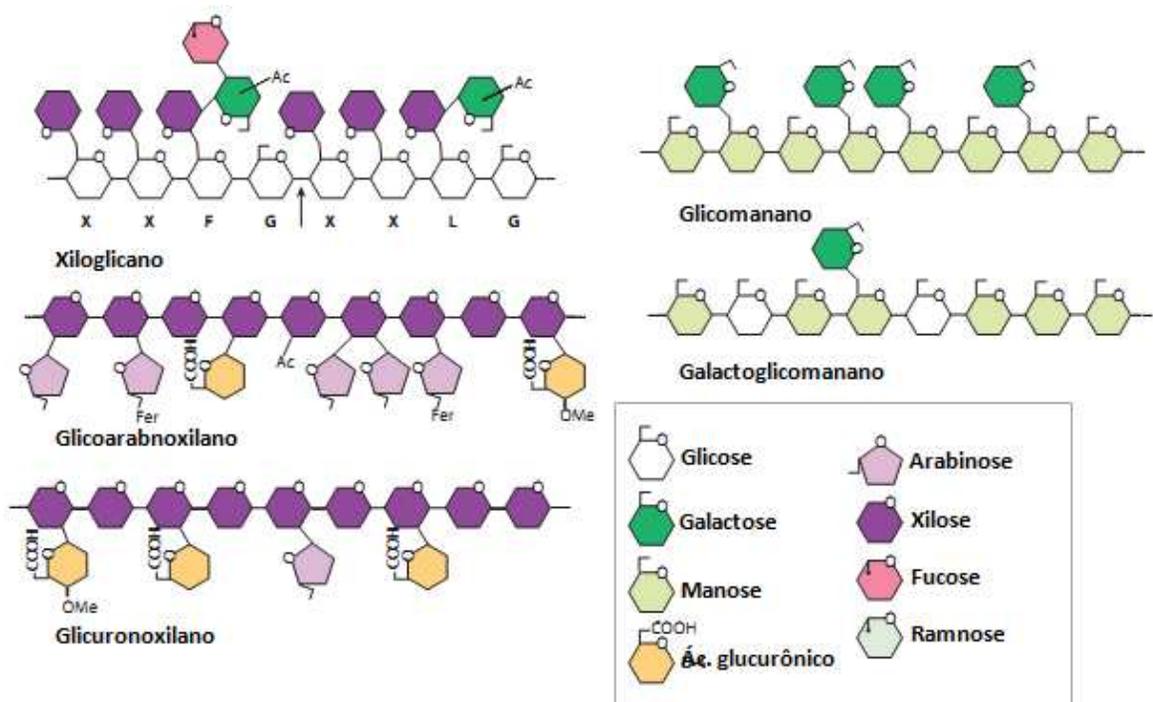


Figura 8: A. Estrutura dos principais tipos de hemicelulose. Adaptado de Scheller & Ulvskov (2010).

Os genes estudados da família *Susy* são divididos em três grupos de acordo com a predição dos aminoácidos e estrutura de íntron/éxon. Em *Arabidopsis*, a família da *Susy* é composta por seis membros, divididos em três grupos. A análise de transgênicos e mutantes mostram que as isoformas possuem funções específicas dentro da planta e estão expressas de maneira diferencial entre os tecidos (Bieniawska *et al.*, 2007). Os genes *AtSusy1* e *AtSusy4* pertencem ao mesmo grupo, assim como *AtSusy2* e *AtSusy3* e *AtSusy5* e *AtSusy6* formam mais dois grupos. Os genes *AtSusy1* e *AtSusy4* exibem uma identidade de aminoácidos elevada. Entretanto, a expressão de *AtSusy1* parece ser constante em todos os tecidos, enquanto *AtSusy4* parece ser mais expresso em raiz e caule (Baud, *et al.*, 2004). Uma forma de *Susy* associada com a membrana plasmática e com o complexo celulose sintase estudada em Fujii *et al.* (2010) (número de acesso AB495095) possui similaridade com as sequências de *AtSusy 1* e *4*, provavelmente pelo perfil de expressão, a forma associada à membrana seja *AtSusy4*. Em *Eucalyptus*, duas sequências de *Susy* foram isoladas de *E. grandis*: *EgSusy1* e *EgSusy3* (Zhou, H. & Myburg, A. A., 2005 – dados não publicados). Ambas pertencentes ao mesmo grupo, tendo *EgSusy1* identidade com *AtSusy1* e *EgSusy3* com *AtSusy4*. Ambas estão sendo estudadas neste trabalho.

Na parede celular, a celulose está ligada às hemiceluloses por meio de pontes de hidrogênio, o que contribui para a força da parede celular, e ambas estão inseridas numa matriz de pectina. Em tecidos lignificados, a hemicelulose está ligada, também, à lignina através de pontes de hidrogênio. Os monossacarídeos constituintes das hemiceluloses e pectinas são D-manose, D-galactose, D-xilose, D-apiose, L-arabinose, L-ramnose, L-fucose, ácido D-manurônico, ácido D-galacturônico e ácido D-glicurônico (Figura 7) (Carpita & McCann, 2000).

Os polímeros de hemiceluloses são formados por uma mistura dos monômeros citados, além de diversas estruturas de ligação. Possuem um esqueleto principal constituído por uma cadeia com ligações do tipo β -1,4 formado por glicose, manose ou xilose e diversas ramificações laterais. Podem ser divididos nos seguintes grupos: xiloglicanos, xilanos, mananos, glicomananos e β -(1,3,1,4)-glicanos (Fengel & Wegener, 1985; Carpita & McCann, 2000; Scheller & Ulvskov, 2010).

Os xiloglicanos (XyG) estão presentes na maioria das espécies já estudadas e sua maior ocorrência é na parede primária, exceto para as gramíneas. São formados por uma cadeia principal do tipo β -1,4 glicano não-ramificado (**G**) e glicanos contendo ramificações de

resíduos α -1,6 xilosil (**X**) ao longo da cadeia (Scheller & Ulvskov, 2010; Hayashi & Kaida, 2011) (Figura 8).

Nas paredes primárias da maioria das dicotiledôneas, monocotiledôneas não-gramináceas e gimnospermas, a cadeia principal dos xiloglicanos é formada por unidades heptassacarídicas repetidas, sendo quatro resíduos consecutivos de β -D-glicosil e três ramificações de resíduos α -D-xilosil na cadeia de glicano. Cerca de metade destas unidades contém extensões de β -D-galactosil (L) (em algumas plantas, ocorre a ramificação por resíduos de L-arabinose (Carpita, 2011)). Um resíduo de α -L-furanosil é depois adicionado a um galactosil. A estrutura formada após a adição do resíduo de furanosil possui grande importância estrutural, pois este grupo trissacarídico permite que o xiloglicano assumira uma forma não torcida, o que facilita sua ligação à celulose. Este grupo também parece modular o crescimento induzido por auxina (Pena *et al.*, 2004).

Os xiloglicanos juntamente com a celulose, fornecem a força de tensão necessária à parede primária. Eles ligam-se às microfibrilas de celulose por de pontes de hidrogênio, limitando a agregação e contribuindo, assim, para as propriedades mecânicas da parede. Podem estar associados à celulose de três maneiras principais: vinculados à superfície, sem contato direto ou inseridas na microfibrila. Além disto, estão envolvidos na reestruturação da parede primária em células em crescimento. Nestas, os xiloglicanos são quebrados em determinadas zonas por enzimas chamadas xiloglicano endoglicosiltransferases e novamente religados em diferentes pontos de modo que a expansão celular ocorre sem que haja influência na sua resistência mecânica (Darley *et al.*, 2001).

A biossíntese dos xiloglicanos e outros polissacarídeos não celulósicos acontece, diferentemente da celulose, no complexo de Golgi. Entretanto, devido à semelhança nos domínios conservados entre a cadeia principal de ambos, acredita-se que membros da superfamília Celulose Sintase, os *Csl* (*celulose synthase like*) sejam responsáveis por codificar as sintases dos polissacarídeos não-celulósicos. Além disso, diversos trabalhos correlacionaram a ação de *Csl* com a síntese de polissacarídeos não-celulósicos (Liepman *et al.*, 2005; Cocuron *et al.*, 2007, Verherbruggen *et al.*, 2011). Genes do subgrupo *Cs/C* codificam glicano sintases que participam da formação da cadeia principal do xiloglicano. Em *A. thaliana*, a subfamília é formada por cinco membros, e apenas um deles (*Cs/C4*) teve ação comprovada na biossíntese de xiloglicano (Cocuron *et al.*, 2007). As ramificações são adicionadas pelas glicosil-transferases. Dentre as glicosiltransferases que participam da formação de xiloglicanos em dicotiledôneas estão: xilosiltransferases, galactosiltransferases e

fucosiltransferases, que adicionam resíduos de xilose (a partir de UDP-xilose), galactose (a partir de UDP-galactose) e fucose (a partir de GDP-fucose), respectivamente (Carpita, 2011). Sabe-se que tais glicano-sintases estão associadas à membrana do complexo de Golgi. Apesar dos muitos estudos já realizados, ainda não está claro o modo de ação de tais sintases (Carpita, 2011).

Os xilanos são um grupo diverso de polissacarídeos tendo em comum uma cadeia principal de resíduos de xilose (com ligações do tipo β -1,4). Uma ramificação comum nos xilanos ocorre com resíduos de glicuronosil (originado de UDP-glicuronato) e são denominados glicuronoxilano, predominante na parede celular secundária de dicotiledôneas. Em gramíneas, os xilanos contêm resíduos de arabinose e são chamados de arabinoxilanos (presentes também em gimnospermas) ou glicoarabinoxilano (GAXs), quando contêm os dois resíduos (Figura 8). A maior parte deles é acetilada em vários graus. Ao contrário dos xiloglicanos, os xilanos não possuem uma estrutura repetitiva (Scheller & Ulvskov, 2010).

Existem evidências de que *CsID5* estejam envolvidos na síntese de xilano (Bernal *et al.*, 2007), entretanto ainda é necessário a demonstração direta da ação deste gene (Carpita *et al.*, 2011). Genes de outras famílias também foram correlacionados com a síntese de xilano (Brown *et al.*, 2007; Lee *et al.*, 2011).

Os mananos e glicomananos também são um grupo de polissacarídeos com ligações do tipo β -1,4, contendo manose na cadeia principal (Figura 8). Quando a cadeia principal é composta apenas de resíduos de manose dá-se o nome de mananos ou galactomananos (quando ocorrem ramificações com resíduos de galactose). Quando composta por resíduos de manose e glicose distribuídos de forma não repetitiva, são chamados de glicomananos e galactoglicomananos. Estes polissacarídeos são os menos abundantes em espermatófitas. Entretanto parecem desempenhar papéis importantes. Membros da família *Cs/* do subgrupo A são responsáveis por codificar manano e glicomananos sintase, que participam da polimerização da cadeia principal. Estas enzimas parecem utilizar GDP-manose e GDP-glicose como precursores. O subgrupo D da família *Cs*, também parece estar envolvido na codificação destas proteínas. A glicosiltransferase responsável pela adição das cadeias laterais é a galactosiltransferase (Liepman *et al.*, 2005; Scheller & Ulvskov, 2010; Carpita, 2011).

Um outro tipo de hemicelulose é formada por cadeias de glicanos com ligações mistas do tipo β -1,3 e β -1,4 (Carpita, 1996). Entretanto, tal subclasse não será discutida nesse trabalho por se tratar de hemiceluloses não presentes em dicotiledôneas.

A pectina é constituída por um grupo de polissacarídeos ácidos, altamente hidratados, ricos em ácido galacturônico (versão ácida da galactose) (Figura 9). Os polissacarídeos pecticos estão divididos em três grupos principais: homogalacturonanos (HG), polímeros contendo uma cadeia com apenas ácido galacturônico, tendo alguns grupos carboxil metil-esterificados; ramnogalacturonanos-I (RG-I), o qual contém em sua cadeia principal também dissacarídeos formados por resíduos de ácido galacturônico e ramnose, a qual pode conter diversas ramificações, principalmente de arabinofuranosil ou galactopiranosil; ramnogalacturonanos-II (RG-II), grupo não estruturalmente ligado à RG-I e possuem estrutura altamente complexa (também contendo ácido galacturônico na cadeia principal) e ramificada (Willats *et al*, 2001; Lerouxel *et al.*, 2006).

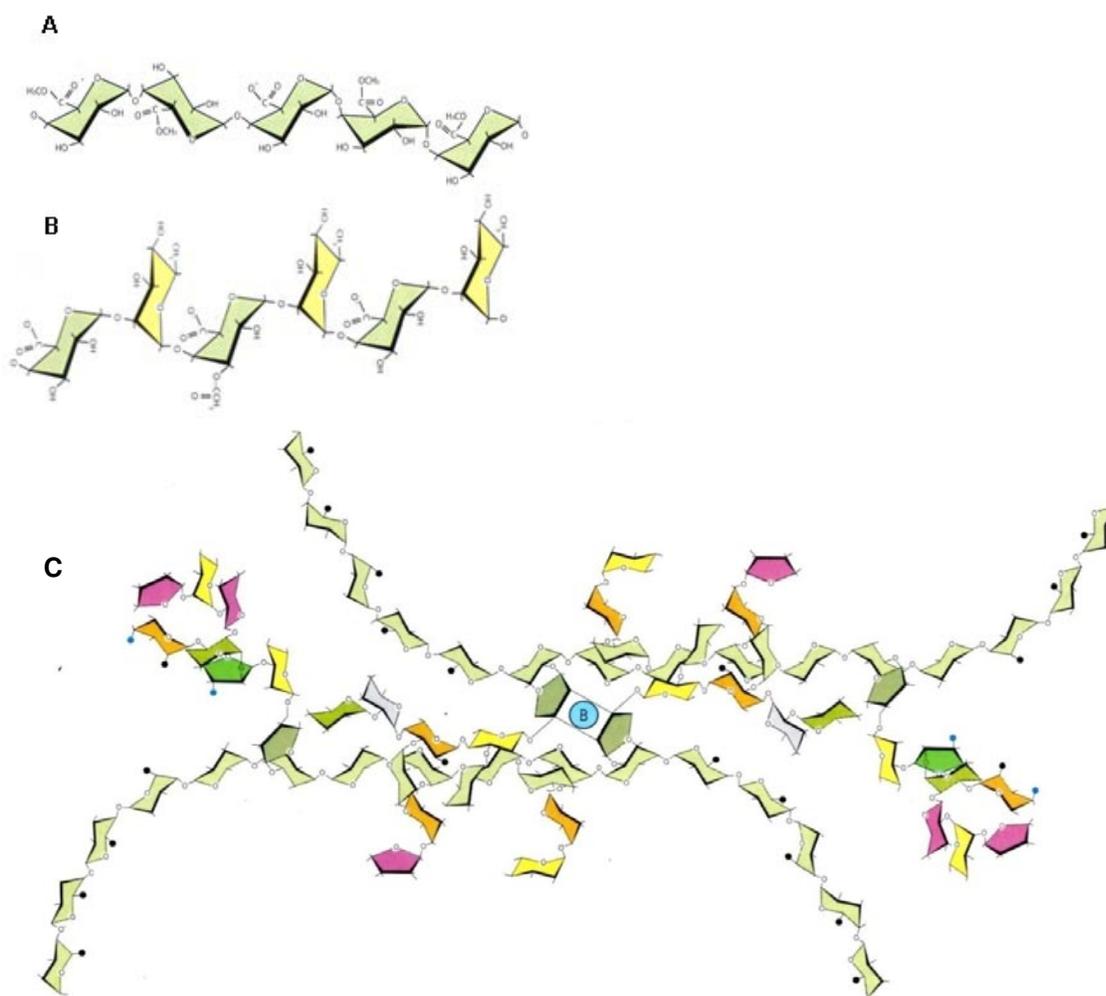


Figura 9: Estrutura geral das pectinas: A. Homogalacturonano. B. Ramnogalacturonano I (sem ramificações). C. Ramnogalacturonano II. As figuras, verde, amarela, laranja e rosa representam ácido galacturônico, ramnose, xilose e arabinose respectivamente.

A matriz de pectina fornece o ambiente ideal para a deposição, rearranjo e extensão da rede celulose – hemicelulose. Também está envolvida na porosidade da parede celular e na adesão entre células (Willats *et al.*, 2001; Lerouxel *et al.*, 2006).

As hemiceluloses e pectinas, ao contrário da celulose, são formadas no citoplasma e enviadas à parede celular pelas de vesículas secretoras. Os monômeros são formados, direcionados ao complexo de Golgi e lá, por meio da ação das glicano sintases e glicosiltransferases, são polimerizados (Perrin *et al.*, 2001).

A lignina é um polímero de constituição química difícil de ser estabelecida, principalmente devido à complexidade de sua estrutura, baseada em unidades fenilpropanóides interligadas por diferentes tipos de ligações, como também por sofrer modificações estruturais durante seu isolamento das paredes celulares. A lignina é depositada principalmente nas paredes celulares secundárias de elementos traqueais e fibras, principais células que formam a madeira, dentro da rede de celulose e hemicelulose (Raven *et al.*, 2001).

A lignina é responsável por fornecer uma superfície hidrofóbica que permite o transporte de água a uma altura de mais de 100 metros; contribui com a força mecânica que mantém as árvores em posição vertical; serve de barreira química e física à ação de insetos e patógenos (Franke *et al.*, 2002; Bhuiyan *et al.*, 2009).

Sabe-se que a lignina é formada a partir de três unidades básicas que diferem entre si pelo grau de metilação: álcool transconiferílico, álcool transsinapílico e álcool para-transcumarílico. Cada um destes tipos leva à formação de diferentes tipos de monômeros: guaicila (G), siringila (S) e cumarila (H), respectivamente (Zhong *et al.*, 1998).

Uma grande quantidade de genes da via de síntese de lignina já foi identificada: como os que codificam as enzimas fenilalanina amônia-liase (PAL), 4-cumarato-CoA ligase (4CL), cinamato 4-hidroxilase (C4H), ferulato 5-hidroxilase (F5H), cinamoil-CoA redutase (CCR), : álcool cinamílico desidrogenase (CAD), que participam das reações da etapa que antecede à polimerização dos monômeros constituintes da lignina (Boudet, 2000; Darley, 2001; Boudet, 2003).

Também fazem parte da parede celular as proteínas estruturais e as modificadoras da parede. As primeiras não possuem funções claras. São divididas em quatro classes principais: HRGPs – glicoproteínas ricas em hidroxiprolina; AGPs – proteínas ricas em arabinogalactanos; GRPs – proteínas ricas em glicina; e PRPs – proteínas ricas em prolina (Cassab, 1998). As

principais proteínas modificadoras da parede são as XTHs (xiloglucano endotransglicosilase/hidrolase), as extensinas e as várias classes de pectinases (Cosgrove, 1999, Rose *et al.*, 2002).

Além dos genes que codificam enzimas responsáveis pela síntese dos componentes estruturais, diversos outros genes identificados no contexto da formação da madeira codificam fatores de transcrição (Demura & Fukuda, 2007). Por exemplo, genes da família de fatores de transcrição NAC (NST1 e NST3) foram caracterizados em *A. thaliana* como essenciais ao processo de formação da parede secundária. Ortólogos destes genes foram também identificados em *Populus sp*, no qual parecem exercer as mesmas funções (Mitsuda, 2007). Fatores de transcrição da família Myb (*EgMYB1* e *EgMYB2*) que regulam a formação da parede celular secundária e a biossíntese de lignina já foram caracterizados em *Eucalyptus* (Goicoechea *et al.*, 2005; Legay *et al.*, 2007;).

Nossos resultados indicam que grande parte destes genes apresenta regulação diferencial no xilema das espécies estudadas (*E. grandis*, *E. globulus* e *E. urophylla*). Dada a diferença na qualidade da madeira e velocidade de crescimento destas espécies, tais genes constituem bons candidatos para a manipulação genética com vista à produção de *Eucalyptus* portadores de características de produtividade mais interessantes ao setor de papel e celulose.

4. Referências

ABRAF – Associação Brasileira de Produtores de Florestas Plantadas. Disponível em <http://www.abraflor.org.br/>. **Anuário Estatístico ABRAF 2010 – Ano Base 2009.**

ABRAF – Associação Brasileira de Produtores de Florestas Plantadas. Disponível em <http://www.abraflor.org.br/>. **Anuário Estatístico ABRAF 2012 – Ano Base 2011.**

AMOR, Y.; HAIGLER, C. H.; JOHNSON, S.; WAINSCOTT, M. & DELMER, D. P. A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. **Plant Biology**, v. 92, p. 9353 – 9357. 1995.

ANDERSSON-GUNNERAS, S.; MELLEROWICZ, E. J.; LOVE, J.; SEGERMAN, B.; OHMIYA, Y.; COUTINHO, P. M.; NILSSON, P.; HENRISSAT, B.; MORITZ, T. & SUNDBERG, B. Biosynthesis of cellulose-enriched tension wood in *Populus*: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. **The Plant Journal**, v. 45, p. 144-165. 2006.

BAUCHER, M.; HALPIN, C.; PETIT-CONIL, M.; BOERJAN, W. Lignin: genetic engineering and impact on pulping. **Critical Reviews in Biochemistry and Molecular Biology**, v. 38(4), p. 305–350. 2003.

BAUD, S.; VAULTIER, M.; ROCHAT, C. Structure and expression profile of the sucrose synthase multigene family in *Arabidopsis*. **Journal of Experimental Botany**, v. 55 (396), p. 397 – 409. 2004.

BERNAL, A. J.; KRUGER JENSEN, J.; HARHOLT, J.; SORENSEN, S.; MOLLER, I.; BLAUKOPF, C.; JOHANSEN, B.; DE LOTTO, R.; PAULY, M.; VIBE-SHELLER, H.; WILLATS, W. G. T. Disruption of *AtCslD5* results in reduced growth, reduced Xylan and Homogalacturonan synthase activity and altered Xylan occurrence in *Arabidopsis*. **The Plant Journal**, v. 52, p. 791-802. 2007.

BERTOLA, A. “**Eucalipto: Verdades e Mentiras**”. Disponível em: www.celuloseonline.com.br/imagembank/Docs/DocBank/dc/dc009.pdf. Acesso em 12 de março de 2011.

BHUIYAN, N. H.; SELVARAJ, G.; WEI, Y.; KING, J. Role of lignification in plant defense. **Plant Signal Behavior**, v. 4, p. 158–159. 2009.

BIENIAWSKA, Z; BARRATT, D. H. P.; GARLICK, A. P.; THOLE, V.; KRUGER, N. J.; MARTIN, C.; ZRENNER, R. & SMITH, A. M. Analysis of the sucrose synthase gene family in *Arabidopsis*. **The Plant Journal**, v. 49, p. 810 – 828. 2007.

BISON, O.; RAMALHO, M. A. P.; REZENDE, G. D. S. P.; AGUIAR, A. M.; DE RESENDE, M. D. V. Combining ability of elite clones of *Eucalyptus grandis* and *Eucalyptus urophylla* with *Eucalyptus globulus*. **Genetics and Molecular Biology**, v. 30(2), p. 417 – 422. 2007

BRITO, J. O.; BARRICHELO, L. E. G.; SEIXAS, F. Análise da Produção Energética e de Carvão Vegetal de Espécies de Eucalipto. **IPEF**, n.23, p.53-56. 1983.

BROWN, D. M.; GOUBET, F.; WONG, V. W.; GOODACRE, R.; STEPHENS, E.; DUPREE, P.; TURNER, S. R. Comparison on five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. **The Plant Journal**, v. 52 (6), p. 1154 – 1168. 2007.

BOUDET, A. M. Lignins and lignification: Selected issues. **Plant Physiology and Biochemistry**, v. 38 (1/2), p. 81-96. 2000.

BOUDET, A. M.; KAJITA, S.; GRIMA-PETTENATI, J.; GOFFNER, D. Lignins and lignocellulosics: a better control of synthesis for new and improved uses. **Trends in Plant Science**, v. 12 (8), p. 576 – 581. 2003.

BRACELPA – Associação Brasileira de Celulose e Papel. Disponível em: <http://www.bracelpa.org.br/bra2/>. Acesso em 15 de abril de 2011.

BROOKER, M. I. H. A new classification of the genus *Eucalyptus* L'Her. (Myrtaceae). **Australian Systematic Botany**, v.13, p. 79 – 148. 2000.

BUCKERIDGE, M. S.; VERGARA, C. E.; CARPITA, N. C. Insight into multi-site mechanisms of glycosyl transfer in (1-3), (1-4) β -D-glucan synthase. **Phytochemistry**, v. 57, p. 1045 – 1053. 2001.

BUCKERIDGE, M. S.; SILVA, G. B.; CAVALARI, A. A. **Parede celular**. Kerbauy, G.B. (ed.) Fisiologia Vegetal. Guanabara Koogan S.A . R.J. p. 165-181. 2008.

BURTON, R. A.; FARROKHI, N.; BACIC, A. & FINCHER, G. B. Plant cell wall polysaccharide biosynthesis: real progress in the identification of participating genes. **Planta**, v. 221, p. 309-312. 2005.

CAMPBELL, M. M.; BRUNNER, A. M.; JONES, H. M.; STRAUSS, S. H. Forestry's fertile crescent: the application of biotechnology to forest trees. **Plant Biotechnology Journal**, v. 1, p. 141-154, 2003.

CARPITA, N. C. Structure and biogenesis of the cell walls of grasses. **Annual review of plant physiology and plant molecular biology**, v. 47, p. 445 – 476. 1996.

CARPITA, N. C. Update on mechanisms of plant cell wall biosynthesis: How plants make cellulose and other (1 \rightarrow 4)- β -D-glycans. **Plant Physiology**, v. 155, p. 171 – 184. 2011.

CARPITA, N.C. & McCANN, M. The cell wall. **Biochemistry & Molecular Biology of Plants**. Rockville: American Society of Plant Physiologists, Chap.2, p. 52. 2000.

CASSAB, G. I. Plant cell wall proteins. **Annual Review on Plant Physiology and Plant Molecular Biology**, v.49, p. 281 – 309. 1998.

CHANG, M. Sequestro de carbono florestal: oportunidades e riscos para o Brasil. **Revista Paranaense de Desenvolvimento**, v.102, p. 85 - 101. 2002.

CHEUNG, F.; HAAS, B. J.; GOLDBERG, S. M. D.; MAY, G. D.; XIAO, Y.; TOWN, C. D. Sequencing *Medicago truncatula* expressed sequenced tags using 454 Life Sciences technology. **BMC Genomics**, v. 7 (272). 2006.

COCURON, J-C.; LEROUXEL, O.; DRAKAKAKI, G.; ALONSO, A. P.; LIEPMAN, A. H.; KEEGSTRA, K.; RAIKHEL, N.; WILKERSON, C. G. A gene from cellulose synthase-like C family encodes a β -1,4 glucan synthase. **Proceedings of the national academy of sciences of the United States of America**, v. 104 (20), p. 8550 - 8555. 2007.

COSGROVE, D. J. Enzymes and other agents that enhance cell wall extensibility. **Annual Review on Plant Physiology and Plant Molecular Biology**, v. 50, p. 391 – 417. 1999.

DARLEY, C. P.; FORRESTER, A. M.; MCQUEEN-MASON, S. J. The molecular basis of plant cell wall extension. **Plant molecular Biology**, v.47, p.179 - 195. 2001.

DÉJARDIN, A.; LAURANS, F.; ARNAUD, D.; BRETON, C.; PILATE, G.; LEPLÉ, J. Wood formation in Angiosperms. **Comptes Rendus Biologies**, v. 333, p. 325 – 334. 2010.

DELMER, D. P. & AMOR, Y. Cellulose Biosynthesis. **The Plant Cell**, v. 7, p. 987-1000.1995.

DEMURA, T. & FUKUDA, H. Transcriptional regulation in wood formation. **Trends in Plant Science**, v. 12 (2), p. 64 - 70. 2007.

DESPREZ, T.; JURANIEC, M.; CROWELL, E. F.; JOUY, H.; POCHYLOVA, Z.; PARCY, F.; HÖFTE, H.; GONNEAU, M.; VERNHETTES, S. Organization of cellulose synthase complexes involved in primary cell wall synthesis in *Arabidopsis thaliana*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 104, p.15572-15577. 2007.

DHUGGA, K. S. Building the wall: genes and enzyme complexes for polysaccharide synthases. **Current Opinion in Plant Biology**, v. 4, p. 488 - 493. 2001.

DICKINSON, J. C. The Eucalypt in the Sierra of Southern Peru. **Annals of the Association of American Geographers**, v. 59 (2), p. 294 - 306. 1969.

DOBLIN, M. S.; KUREK, I.; JACOB-WILK, D. & DELMER, D. P. Cellulose biosynthesis in Plants: from genes to rosettes. **Plant Cell Physiology**, v. 43 (12), p. 1407-1420. 2002.

EMRICH, S. J.; BARBAZUK, W. B.; LI, L.; SCHNABLE, P. S. Gene discovery and annotation using LCM-454 transcriptome sequencing. **Genome Research**, v. 17, p. 69 - 73. 2007.

FAO – FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. Disponível em: <http://www.fao.org/forestry>. Acesso em 23 de maio de 2011.

FENGEL, D. & WEGENER, G. Wood-Chemistry, Ultrastructure, Reactions. **Journal of Polymer Science: Polymer Letters Edition**, v. 23 (11), p. 601–602. 1985.

FOUCART, C.; PAUX, E.; LADOUCE, N.; SAN-CLEMENTE, H.; GRIMA-PETTENATI, J.; SIVADON, P. Transcript profiling of a xylem vs phloem cDNA subtractive library identifies new genes expressed during xylogenesis in *Eucalyptus*. **New Phytologist**, v. 170 (4), p. 739-752, 2006.

FRANKE, R.; HEMM, M. R.; DENAULT, J. W.; RUEGGER, M. O.; HUMPHREYS, J. M.; CHAPPLE, C. Changes in secondary metabolism and deposition of an unusual lignin in the ref8 mutant of *Arabidopsis*. **The Plant Journal**, v. 30, p. 47 – 59. 2002.

FU, N.; FU, X.; GUO, S.; YAN, Z.; XU, Y.; HU, H.; MENZEL, C.; CHEN, W.; LI, Y.; ZENG, R.; KHAITOVICH, P. Estimating accuracy of RNA-Seq and microarrays with proteomics. **BMC Genomics**, v. 10 (161). 2009

FUJII, S.; HAYASHI, T.; MIZUNO, K. Sucrose synthase is an integral component of the cellulose synthesis machinery. **Plant and Cell Physiology**, v. 5 (2), p. 294 – 301. 2010.

FUKUDA, H. Xylogenesis: initiation, progression and cell death. **Annual Review of Plant Physiology and Plant Molecular Biology**, v.47, p.299–325.1996.

GOICOECHEA, M.; LACOMBE, E.; LEGAY, S.; MIHALJEVIC, S.; RECH, P.; JAUNEAU, A.; LAPIERRE, C.; POLLET, B.; VERHAEGEN, D.; CHAUBET-GIGOT, N.; GRIMA-PETTENATI, J. EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. **The Plant Journal**, v. 43, p. 553 – 567. 2005.

GRIMA-PETTENATI, J. & GOFFNER, D. Lignin genetic engineering revisited. **Plant Science**, v. 145 (2), p. 51 – 65. 1999.

HATEY, F.; TOSSER-KLOPP, G.; CLOUSCARD-MARTINATO, C.; MULSANT, P.; GASSER, F. Expressed sequence tags for genes: a review. **Genetics Selection Evolution**, v.30 (6), p. 521–541.1998.

HAYASHI, T.; KAIDA, R. Functions of xyloglucan in plant cells. **Molecular Plant**, v.4(1), p. 17 – 24. 2011.

HERTZBERG, M.; ASPEBORG, H.; SCHRADER, J.; ANDERSSON, A.; ERLANDSSON, R.; BLOMQLVST, K.; BHALERAO, R.; UHLÉN, M.; TEERL, T. T.; LUNDEBERG, J.; SUNDBERG, B.;

NILSSON, P.; SANDBERG, G. A transcriptional roadmap to wood formation. **Proceedings of the National Academy of Sciences of the United States of America**, v. 98, p. 14732 - 14737. 2001.

IPEF – Instituto de Pesquisas e Estudos Florestais. **Chave de Identificação de Espécies Florestais**. Disponível em: <http://www.ipef.br/identificacao/cief/especies/grandis.asp>. Acesso em 4 de julho de 2012.

LEE, C.; TENG, Q.; ZHONG, R.; YE, Z-H. Molecular dissection of xylan biosynthesis during wood formation in poplar. **Molecular Plant**, v. 4 (4), p. 730 – 747. 2011.

LEGAY, S.; LACOMBE, E.; GOICOECHEA, M.; BRIERE, C.; SEGUIN, A.; MACKAY, J. & GRIMAPETTENATI, J. Molecular characterization of EgMyb1, a putative transcriptional repressor of the lignin biosynthetic pathway. **Plant Science**, v. 173 (5), p. 542 – 549. 2007.

LEROUXEL, O.; CAVALIER, D. M.; LIEPMAN, A. H.; KEEGSTRA, K. Biosynthesis of plant cell wall polysaccharides – a complex process. **Current Opinion in Plant Biology**, v. 9, p. 621 -630. 2006.

LIEPMAN, A. H.; WILKERSON, C. G.; KEEGSTRA, K. Expression of cellulose synthase-like (*Cs*) genes in insect cells reveals that *Cs/A* family members encode mannan synthases. **Proceedings of the national academy of sciences of the United States of America**, v. 102(6), p. 2221 – 2226. 2005.

LIMA, W. P. **Impacto Ambiental do Eucalipto**, São Paulo: EDUSP, 1993. 301p.

LODISH, H.; BERK, A.; ZIPURSKY, S. L.; MATSUDAIRA, P.; BALTIMORE, D. & DARNELL, J. E. **Molecular Cell Biology**, 4^o ed. New York. W. H. Freeman & Co., 2000. 1084p.

MELLEROWICZ, E.J.; BAUCHER, M.; SUNDBERG, B.; BOERJAN, W. Unravelling cell wall formation in the woody dicot stem. **Plant Molecular Biology**, v. 47, p. 239-274, 2001.

MITSUDA, N.; IWASE, A.; YAMAMOTO, H.; YOSHIDA, M.; SEKI, M.; SHINOZAKI, K.; OHMETAKAGI, M. NAC transcription factors, NST1 e NST2, are key regulators of the formation of secondary walls in woody tissues of Arabidopsis. **The Plant Cell**, v. 19, p. 270 – 280. 2007.

MIZRACHI, E.; HEFER, C. A.; RANIK, M.; JOUBERT, F.; MYBURG, A. A. De novo assembled expressed gene catalog of a fast-growing *Eucalyptus* tree produced by Illumina mRNA-Seq. **BMC Genomics**, v. 11 (681). 2010.

MORA, A. & GARCIA, C.H. **A cultura do eucalipto no Brasil**. São Paulo: Sociedade Brasileira de Silvicultura, 2000.112 p.

MOORE, R.; CLARK, D.; VODOPICH, D. Botany visual resource library. The Mc-Graw-Hill co. 1998.

NAIRN, J. C. & HASELKORN, T. Three loblolly pine *CesA* genes expressed in developing xylem are orthologous to secondary cell wall *CesA* genes of Angiosperms. **New Phytologist**, v. 166, p. 907 – 915. 2005.

PENA, M. J.; RYDEN, P.; MADSON, M.; SMITH, A. C.; CARPITA, N. C. The galactose residues of xyloglucan are essential to maintain mechanical strength of primary cell walls in *Arabidopsis* during growth. **Plant Physiology**, v. 134, p. 443 – 541. 2004.

PERRIN, R.; WILKERSON, C.; KEEGSTRA, K. Golgi enzymes that synthesize plant cell wall polysaccharides: finding and evaluating candidates in the genomic era. **Plant Molecular Biology**, v. 47, p. 115 – 130. 2001.

PLOMION, C.; LEPROVOST, G.; STOKES, A. Wood formation in trees. **Plant Physiology**, v. 127, p. 1513 – 1523. 2001.

POKE, F. S., VAILLANCOURT, R. E., POTTS, B. M.; REID, J. B. Genomic research in Eucalipto. **Genetica**. v. 125, p. 79 - 101. 2005.

POTTS, B. M. & PEDERICK, L. A: Morphology, phylogeny, origin, distribution and genetic diversity of *Eucalyptus*. In **Diseases and pathogens of *Eucalyptus***. 1st edition. Edited by Keane PJ, Kile GA, Podger FD, Brown BN. Collingwood: CSIRO Publishing: 2000. 11- 34.

RANIK, M & MYBURG, A. A. Six new cellulose synthase genes from *Eucalyptus* are associated with primary and secondary cell wall biosynthesis. **Tree Physiology**, v. 26, p. 545 – 556. 2006.

RAVEN, P. H.; EVERT, R. F.; EICHHORN, S. E. **Biologia Vegetal**, 6. ed. Rio de Janeiro: Guanabara Koogan, 2001. 906p.

REMADE - REVISTA DA MADEIRA - EDIÇÃO Nº116 - SETEMBRO DE 2008. **Por que uma espécie exótica como eucalipto?**

REMADE - REVISTA DA MADEIRA - EDIÇÃO Nº75 - AGOSTO DE 2003. **Descrição de algumas espécies de eucalipto.**

RICHMOND, T. Higher plant cellulose synthases. **Genome Biology**, v. 1 (4), p. 3000.1 – 3000-6. 2000.

ROCHA, M.T. Aquecimento global e o mercado de carbono: uma aplicação do modelo CERT. 2003. 196p. **Tese (doutorado)** – Escola de Agronomia Luiz de Queiroz, Piracicaba, 2003.

RODRIGUES, J.; MEIER, D.; FAIX, O. & PEREIRA, H. Determination of tree to tree variation in syringyl/guaiacyl ratio of *Eucalyptus globulus* wood lignin by analytical pyrolysis. **Journal of Analytical and Applied Pyrolysis**, v. 48, p. 121 – 128. 1999.

ROSE, J. K. C.; BRAAM, J.; FRY, S. C.; NISHITANI, K. The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature. **Plant and Cell Physiology**, v. 43 (12), p. 1421 – 1435. 2002.

SALISBURY, F. B. & ROSS, C. W. **Plant Physiology**. 4. ed. Belmont: Wassworth Publishing Company, 1991. 682p.

SAMUGA, A. & JOSHI, C. P. A new cellulose synthase gene (*PtrCesA2*) from aspen xylem is orthologous to *Arabidopsis AtCesA7* (*irx3*) gene associated with secondary cell wall synthesis. **Gene**, v. 296, p. 37-44. 2002.

SAMUGA, A. & JOSHI, C. P. Differential expression patterns of two new primary cell wall-related cellulose synthase cDNAs, *PtrCesA6* and *PtrCesA7* from aspen trees. **Gene**, v. 334, p. 73-82. 2004.

SANTOS, S. N. Genes de Lignificação em *Eucalyptus*: estrutura e diversidade genética dos genes *4cl* e *coaomt*. 2005. 209p. **Dissertação (mestrado)** – Universidade Católica de Brasília, Brasília, 2005.

SCANAVACA JUNIOR, L. Caracterização silvicultural, botânica e tecnológica do *Eucalyptus urophylla* S.T. Blake e de seu potencial para utilização em serraria. 2001. 108p. **Dissertação (mestrado)** – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2002.

SHELLER, H. V. & ULVSKOV, P. Hemicelluloses. **Annual Review in Plant Biology**, v. 61, p. 263 – 289. 2010.

SILVA, J.C. Cresce a presença do eucalipto no Brasil. **Revista da Madeira**, v. 92, p.61-66, 2005.

SONG, D.; SHEN, J. & LI, L. Characterization of cellulose synthase complexes in *Populus* xylem differentiation. **New Phytologist**, v. 187, p. 777 – 790. 2010.

SUZUKI, S.; LI, L.; SUN, Y.; CHIANG, V. L. The cellulose synthase gene superfamily and biochemical functions of xylem-specific cellulose synthase-like genes in *Populus trichocarpa*. **Plant Physiology**, v. 142, p. 1233 – 1245. 2006.

TAIZ, L. & ZEIGER, E. **Fisiologia Vegetal**. 3 ed. Artmed Porto Alegre.2006

TANAKA, K.; MURATA, K.; YAMAZAKI, O.; MIYAO, A.; HIROCHIKA, H. Three distinct rice cellulose synthase catalytic subunit genes required for cellulose synthesis in the secondary wall. **Plant Physiology**, v. 133. p. 73 – 83. 2003.

TAYLOR, N. G.; SCHEIBLE, W. CUTLER, S.; SOMERVILLE, C. R.; TURNER, S. R. The *irregular xylem3* locus of *Arabidopsis* encodes a cellulose synthase required for secondary cell wall synthesis. **The Plant Cell**, v.11, p. 769-779. 1999.

TURNBULL, J. W. Economic and social importance of *Eucalyptus*. In **Diseases and pathogens of Eucalyptus**. 1st edition. Edited by Keane PJ, Kile GA, Podger FD, Brown BN. Collingwood: CSIRO Publishing, 2000:1- 9.

VALVERDE, R. S.; NEIVA, S. A.; SOARES, N. S. Biotecnologia e competitividade das plantações florestais. In: BORÉM, A. (Ed.). **Biotecnologia Florestal**. Viçosa: Universidade Federal de Viçosa, 2007, cap.17, p. 363-374.

VERHERTBRUGGEN, Y.; YIN, L.; OIKAWA, A.; SCHELLER, H. V. Mannan synthase activity in the CSLD family. **Plant Signaling & Behavior**, v. 6(10), p. 1620 – 1623. 2011.

VICENTINI, R.; SASSAKI, F.T.; GIMENES, M. A.; MAIA, I. G.; MENOSSI, M.; In silico evaluation of the *Eucalyptus* transcriptome. **Genetics and Molecular Biology**, v.28, p. 487–495. 2005.

VITAL, B. R. **Método de determinação da densidade da madeira**. Viçosa, SIF: 21p, 1984.

YANG, S. S.; TU, Z. J.; CHEUNG, F.; WENZHONG X.; LAMB, J. F. S.; JUNG, H-J. G.; VANCE, C. P.; GRONWALD, J. W. Using RNA-Seq for gene identification, polymorphism detection and transcript profiling in two alfafa genotypes with divergent cell wall composition in stems. **BMC Genomics**, v. 12 (199). 2011.

WILLATS, W. G. T.; McCARTNEY, L.; MACKIE, W.; KNOX, P. Pectin: cell biology and prospects for functional analysis. **Plant Molecular Biology**, v. 47, p. 9 – 27. 2001.

ZHONG, R.; MORRISON III, W.H.; NEGREL, J. & YE, Z. Dual methylation lignin biosynthesis. **The Plant Cell**, v. 10, p. 2033-2045. 1998.

CAPÍTULO 1: Xylem transcription profiles indicate potential metabolic responses for economically relevant characteristics of *Eucalyptus* species

Xylem transcription profiles indicate potential metabolic responses for economically relevant characteristics of *Eucalyptus* species

Submetido BMC GENOMICS

Marcela Mendes Salazar¹; Leandro Costa Nascimento¹; Eduardo Leal Oliveira Camargo¹; Danieli Cristina Gonçalves¹; Jorge Lepikson Neto¹; Wesley Leoricy Marques¹; Paulo José Pereira Lima Teixeira¹; Piotr Mieczkowski²; Jorge Maurício Costa Mondego³; Marcelo Falsarella Carazzolle¹; Ana Carolina Deckmann¹.; Gonçalo Amarante Guimarães Pereira^{1*}

¹Laboratório de Genômica e Expressão, Departamento de Genética Evolução e Bioagentes, Instituto de Biologia, Universidade Estadual de Campinas, CEP: 13083-970, Campinas, São Paulo, Brasil.

²Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill (UNC), Chapel Hill, NC, USA.

³Centro de Pesquisa e Desenvolvimento em Recursos Genéticos vegetais, Instituto Agrônomo de Campinas, CEP: 13001-970, Campinas, São Paulo, Brasil.

*Correspondence: goncalo@unicamp.br

Abstract:

Background: *Eucalyptus* is one of the most important sources of industrial cellulose. Three species of this botanical group are intensively used in breeding programs: *E. globulus*, *E. grandis* and *E. urophylla*. *E. globulus* is adapted to subtropical/temperate areas and is considered a source of high-quality cellulose; *E. grandis* grows rapidly and is adapted to tropical/subtropical climates; and *E. urophylla*, though less productive, is considered a source of

genes for robustness. Wood, or secondary xylem, results from cambium vascular differentiation and is mostly composed of cellulose, lignin and hemicelluloses. In this study, the xylem transcriptomes of the three *Eucalyptus* species were investigated with the objective of finding clues to understanding the biochemical basis of the differences among the properties of these species.

Results: Data analyses showed the following. (i) The majority of the *Eucalyptus* genes were expressed in xylem. (ii) Most of the genes that were expressed in a species-specific manner were unknown and should be further investigated. (iii) Relevant differences were observed in the phenylpropanoid: the xylem of *E. grandis* tended to have a higher expression of genes for lignin formation, whereas *E. urophylla* exhibited a deviated pathway for flavonoids formation. (iv) The expression of stress-related genes was considerably stronger in *E. urophylla*, which suggests that these genes may contribute to the molecular basis of robustness. (v) Nitrogen metabolism genes showed higher levels of expression in *E. globulus*, which could explain why this species has the highest cellulose content of the three species.

Conclusions: The comparison among the transcriptomes reveals the molecular responses that underlie some of *Eucalyptus* wood characteristics were distinct in these three species. This information may contribute to the understanding of xylogenesis, thereby increasing the potential of genetic engineering approaches for the improvement of *Eucalyptus* forest plantation productivity.

Keywords: ***Eucalyptus*, RNAseq, transcriptome analysis, secondary xylem**

Background:

The eucalypt is a ubiquitous tree in tropical and subtropical regions and is well known for its fast growth and ability to adapt. The *Eucalyptus* genus is composed of more than 600 species with an origin center based in Oceania [1]. Approximately 20 species, most belonging to the subgenus *Symphyomyrtus* [2], are used in commercial plantations in more than 90 countries. *Eucalyptus* wood is utilized in many ways, including for pulp, paper, civil construction and furniture and energy production [3]. In the near future, *Eucalyptus* may be an important source of second-generation biofuels and renewable chemicals. For all of these reasons, *Eucalyptus* has become the most planted forest genus in the world [4], which is a remarkable fact as these species are still in the early stages of domestication [2].

Three *Eucalyptus* species, *E. globulus*, *E. grandis* and *E. urophylla*, are generally considered to be the genetic sources of the most desired breeding features [5]. *E. globulus* trees are well known for their superior wood properties for the paper and pulp industries and its lignin presents a high syringyl/guaiacyl ratio (S/G ratio), which is an important indicator of low recalcitrance during pretreatments [6]. In addition, the cellulose productivity of *E. globulus* is significantly higher than the average of other species; e.g., approximately 25% less wood is required to produce a ton of cellulose [7]. However, this species is not well adapted to tropical conditions [8] and is primarily cultivated in countries such as Portugal, Spain and Uruguay.

In contrast, the wood of *E. urophylla* presents a high lignin content with a low S/G ratio and this species is very well adapted to high temperatures, humidity and drought, which are typical conditions of tropical areas. This species is widely used in breeding programs due to its superior resistance to extreme conditions, mainly to drought stress and diseases [9]. Amongst these three species, *E. grandis* exhibits the fastest growth and moderate resistance to disease and environmental extremes [10].

The lignin and cellulose contents of *E. grandis* are believed to be intermediate between *E. globulus* and *E. urophylla* despite the fact that the precise wood composition of *E. grandis* has not been reported in published studies. Furthermore, there is lack of knowledge regarding the major molecular differences of the xylems from these three species despite previous efforts to understand *Eucalyptus* wood formation. The comprehension of cell wall metabolism will lead to the improvement of *Eucalyptus* breeding programs and allow for the adoption of genetic strategies to enhance tree productivity, such as transgenic technology.

Wood forms as a result of secondary cell wall accumulation in xylem tissues, which occurs in four steps: (i) cell expansion, (ii) the deposition of the secondary cell wall, (iii) lignification and (iv) programmed cell death [11, 12]. Knowledge of the molecular mechanisms

involved in cell wall biosynthesis and its regulation have recently become available, and this area of research has been significantly enhanced by new “omics” technologies [13, 14].

Genomics research is especially relevant for the study of *Eucalyptus*. Since 2000, several research consortiums have produced EST databases, such as Forest [15] and Genolyptus [7], Microarrays [16, 17, 18], SAGE [19] and DArT [20]. These databases are currently used to gain insights on the genetics of wood formation and various genes have been selected for the production of transgenic trees [21, 22].

However, despite these efforts, the mechanism of wood formation has not been sufficiently elucidated. Several genes are involved in secondary xylem development; this process has a complex metabolism, which may explain why previous studies have not provided a comprehensive genetic model of the wood formation processes. Thus, a broad analysis of genes expressed in the xylems of species with different wood characteristics will help to elucidate this intricate molecular network.

In recent years, new sequencing technologies have greatly improved the capacity to generate and assemble complex genomes as well as the ability to map and quantify virtually all genes present in any given transcriptome with extreme precision [23, 24]. The use of these technologies has allowed for the sequencing and assembly of the *E. grandis* genome, and the *de novo* assembly of mRNAseq reads has generated a highly informative panel of *Eucalyptus* coding regions [25].

The objective of the present study was the generation and sequencing of RNA libraries from the xylems of *E. grandis*, *E. urophylla* and *E. globulus* to allow for a comparison of gene expression among the three species. The data analysis was performed by searching for differentially expressed genes to elucidate the metabolic pathways that are potentially involved in the distinct wood properties observed among these species.

Our results revealed that only a small number of genes (5,909), or approximately 20% of the data set, are differentially expressed in the xylems of these species and most of these genes have an unknown function. However, we identified the differential expression of a number of genes important for cell wall synthesis, which could account for phenotypic diversity in xylem and maybe related to the wood characteristics of *E. globulus*, *E. grandis* and *E. urophylla*.

Results and Discussion:

The objective of the present work was to explore the transcriptional programs underlying the xylems of three *Eucalyptus* species, *E. grandis*, *E. globulus* and *E. urophylla*, that are widely employed as genetic sources in breeding programs.

Our results demonstrated the following. (i) The majority of genes present in *Eucalyptus* were expressed in the xylem, which suggests that differentiation is a result of the general transcriptional regulation of genes rather than the expression of specific genes. (ii) Most of the genes that are expressed in a species-specific manner are unknown and should be further investigated. (iii) Among the known biochemical pathways, significant differences were observed in the phenylpropanoid pathway. For example, *E. grandis* xylem exhibited a higher expression of genes for lignin formation, whereas *E. urophylla* deviated from the lignin pathway toward flavonoids formation. (iv) Stress-related genes show higher levels of expression in *E. globulus* and *E. urophylla* than in *E. grandis*, which may indicate that the sampled *E. globulus* trees were growing under adverse conditions and in *E. urophylla*, the superior robustness has a molecular basis. (v) Nitrogen metabolism genes show higher levels of expression in *E. globulus*, which could explain the highest relative content of cellulose in this species, as the addition of nitrogen fertilizer has been correlated with increased cellulose in species such as *Populus* [26]. These and other results are described in detail in next sections.

Xylem RNA sequencing

Xylem tissues were collected from 3-year-old individuals of *E. globulus*, *E. grandis* and *E. urophylla*. Each sample was used for the construction of RNAseq libraries, which were sequenced using an Illumina Genome Analyzer II_x. Approximately 78 million single-end reads with lengths of 36 bp (approximately 2.75 Gbp) were produced. Of this total, 28,838,976, 21,234,801 and 21,047,013 reads were derived from *E. globulus*, *E. grandis* and *E. urophylla*, respectively (Figure 1).

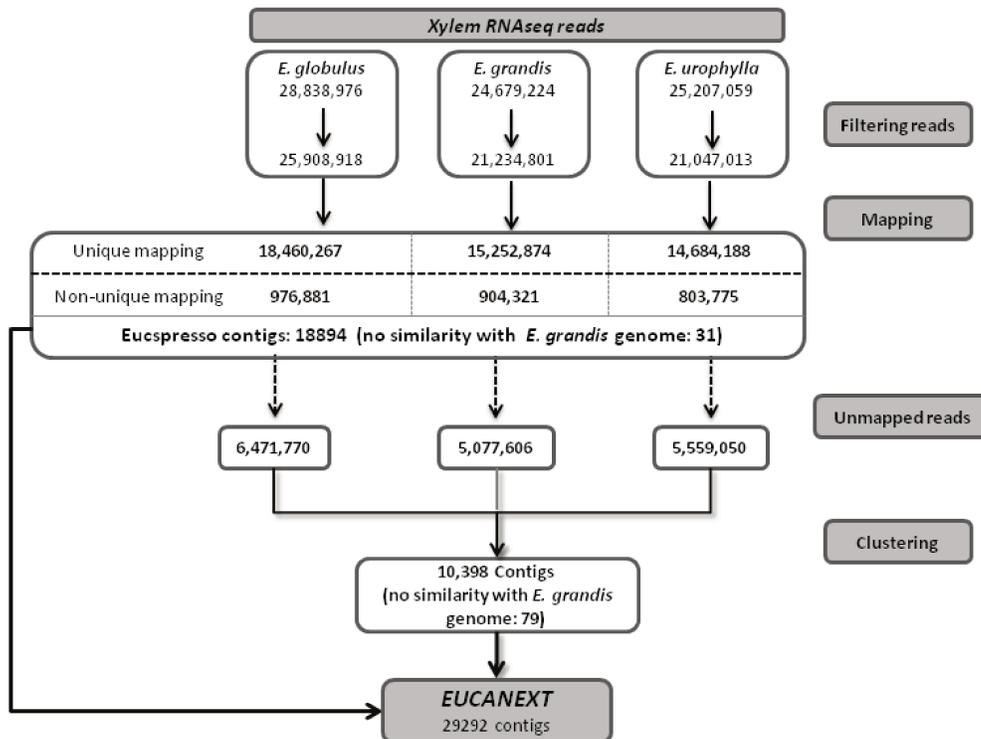


Figure 1: Pipeline: Xylem transcriptome mapping and *de novo* assembly pipeline.

Construction of a reference sequence database (*EUCANEXT*)

Xylem RNAseq reads of all three species were filtered to exclude ribosomal and low-quality sequences. The filtered reads were then mapped against the “*Eucspresso*” database (<http://eucspresso.bi.up.ac.za>), which contains the complete gene catalog generated by Mizrachi and colleagues during the *de novo* assembly of the hybrid *E. grandis* X *E. urophylla* (*urograndis*) ESTs obtained from several tissues [25].

Although this mapping was performed against a database of a hybrid between *E. grandis* and *E. urophylla*, the *E. globulus* reads were mapped with a precision comparable to that of the other species (Figure 1) and a high level of sequence conservation among the three genomes was observed.

Reads from the three species that did not map against the *Eucspresso* dataset (17,108,426 reads of the three species) were clustered together to produce a *de novo* assembly of 10,398 contigs (mean size: 407.75 bp). Together, the *Eucspresso* contigs and the *de novo* assembly resulted in a database containing 29,292 contigs, which was named ***EUCANEXT*** (<http://www.lge.ibi.unicamp.br/Eucalyptusdb>) (Nascimento *et al.*, unpublished data). Figure 1

shows the pipeline used for the construction of the *EUCANEXT* database (the distribution of the contigs lengths are shown in Additional File 1, Figure S1).

Expression analysis

For gene expression analysis, the 29,292 contigs present in the *EUCANEXT* database were compared against the xylem RNAseq reads, and the FPKM (fragments per kilobase of exons per million of fragments mapped) values were calculated for all of the contigs from each of the three species (Additional File 1, Figure S2).

To detect genes that were differentially expressed among the libraries, three pairwise comparisons were performed: *E. grandis* X *E. globulus*, *E. grandis* X *E. urophylla* and *E. globulus* X *E. urophylla*. For these comparisons, contigs with FPKM values equal to 0 were considered to be non-expressed genes. Thus, contigs were distributed into groups according to their expression profile comparisons, the results were represented by a Venn diagram (Figure 2). The selected genes discussed here are shown in Tables S1 and S2 of Additional File 2.

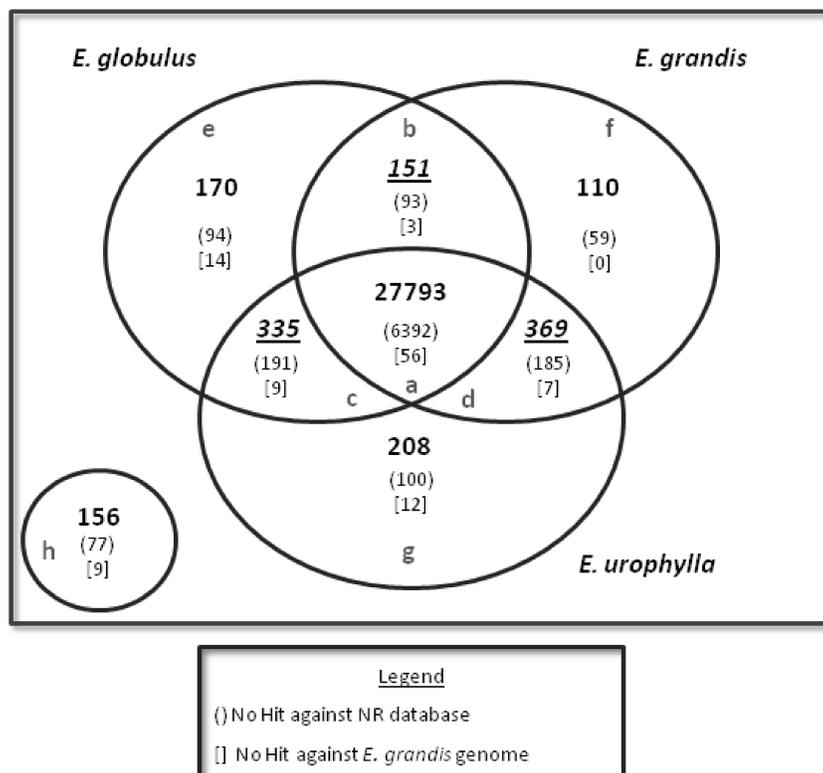


Figure 2: Venn diagrams showing the distribution of contigs between the *Eucalyptus* xylem libraries: Distributions of contigs between species. Bold numbers indicate contigs present in all species (a), contigs present in only two species (b, c and d), contigs present in only one species (e, f and g), and contigs not present in any species (h). The black box shows the legend for other numbers. (Presence of contigs defined as FPKM>0).

In the Figure 2, each subgroup present in the Venn diagram was identified by a Latin letter (*a, b, c...*). Group 'a' was composed of contigs with FPKM>0 in all three *Eucalyptus* species; this group represented genes that were expressed in the xylem tissues of all species. Groups 'b', 'c' and 'd' were composed of contigs with FPKM>0 in only two of the studied species: 'b' - *E. globulus* and *E. grandis*; 'c' - *E. globulus* and *E. urophylla*; and 'd' - *E. grandis* and *E. urophylla*. Groups 'e', 'f' and 'g' were composed of contigs with FPKM>0 in only one species: 'e' - *E. globulus*; 'f' - *E. grandis*; and 'g' - *E. urophylla*. Finally, group 'h' was composed of contigs with FPKM=0 in all species, which represents genes that were not expressed in any of the libraries. The roles of these genes are discussed in the next section (**Functional Analysis**). Each group was analyzed separately as follows.

The analyses of the contigs also included a sequence comparison with the *E. grandis* genome database (<http://Eucalyptusdb.bi.up.ac.za/>) to verify the quality of the dataset. *EUCANEXT* contigs were mapped against the *E. grandis* genome using BLASTn, and positive identifications of 29,106 contigs (99.36%) (e-value < 1e-10) were obtained (Figure 1). As expected, all 110 contigs that were presented only in *E. grandis* (Figure 2, group 'f') displayed a similarity with the genome. Other contigs in groups with FPKM>0 in *E. grandis* (Figure 2; groups 'a', 'b' and 'd') did not map in the genome, which may have occurred because the *E. grandis* genome contains sequence gaps. Interestingly, the proportion of contigs that did not map in the genome for groups of contigs present in only one species (Figure 2; groups 'e' and 'g') was high and included 14 contigs from *E. globulus* and 12 contigs from *E. urophylla*. Therefore, one may speculate that some of these contigs may constitute species-specific genes and are probably contributing to the distinct wood properties observed in these species.

Contigs with FPKM>0 in all species (Figure2, group 'a')

The contigs comprising this category (27,793) represented 94.88% of the *EUCANEXT* contigs. The large number of genes common to the three transcriptomes that were studied here suggests that the distinctive characteristics of each xylem may be derived from differences in the level of the expression of the same group of genes rather than the expression of a distinct set of genes.

A total of 5,175 genes were identified as being differentially expressed in at least one species (Additional File 3, Figure S3). These genes encode ubiquitins, heat shock proteins and transcription factors (MYB, AP2, NAC and WRKY family) in addition to *no hits* (30%), i.e., genes

displaying no similarity with any other sequence deposited in the public non-redundant (NR)nucleotide database (NCBI, National Center of Biotechnology Information) and hypothetical proteins (9%). In addition, many of these differentially expressed genes were related to cell wall metabolism (cellulose synthases, UDP-transferases and XTHs, for example), nitrogen metabolism (nitrate and nitrite reductases), phenylpropanoid metabolism and resistance to biotic and abiotic stress.

Based on the pairwise comparisons, 22,618 genes were not differentially expressed among the species (p -value >0.01 , fold change <2 ; Additional File 3, Figure S3). This result was expected because the majority of transcripts are expected to be involved in the basal metabolism of cells and are essentially present at constant levels in equivalent tissue samples collected from similar species.

For example, of the 200 genes that were most abundantly expressed (FPKM >504) in all three of the *Eucalyptus* xylems studied here, the majority (72%) constituted functionally characterized genes (GO categories are presented in Additional File 3, Figure S4 A), 17% encoded hypothetical proteins and 11% were *no hits* genes.

The genes with known functions included many of the traditional housekeeping genes used in RT-qPCR analysis, such as those encoding histone and actin [27], glyceraldehyde-3-phosphate dehydrogenase [28], calmodulins and tubulin [29]. Several other genes were identified as having roles in cell wall construction and plant development, such as members of the cellulose synthase superfamily, sucrose synthase, lignin production-related genes, UDP-glucose dehydrogenase, ubiquitins, chitinases and heat shock proteins. These 200 genes can be considered as putative housekeeping genes and can be used as reference genes in qPCR studies and other expression experiments involving the xylems of these three species.

From the analysis of the overall functional profile of this group, it is possible to suggest that the cells of all three *Eucalyptus* species follow a common transcriptional program, which can be considered “xylem programming” and includes many genes involved in the production of precursors for secondary growth. Even in *E. globulus*, which is the species that is most susceptible to alterations in climatic conditions, the basic “xylem programming” presented essentially the same gene set as the other species.

Contigs with FPKM >0 in only two species (Figure 2, groups ‘b’, ‘c’ and ‘d’)

The group of contigs with FPKM >0 in only two species was composed of 855 contigs. As showed in Figure 2, these three groups had high numbers of *no hits* genes. Among the contigs

that represented functionally characterized genes, we expected that each group of genes would play different functions in the xylem. Interestingly, the GO categories for level 3 biological processes were similar for all three groups (Additional File 3, Figure S4 B).

Based on the genes that were expressed only in *E. globulus* and *E. grandis* xylems (Figure 2, group 'b'), 63% (96 contigs) were not differentially expressed between them. Of the remaining genes, 17 showed higher levels of expression in *E. grandis*, whereas 38 showed higher levels of expression in *E. globulus* (Additional File 3, Figure S5 A). The majority (61%) constituted *no hits* genes, 29% were functionally characterized and 10% encoded proteins of unknown function (10%). Among the known genes, several were identified as protein kinases and proteins related to stress and pathogen responses.

Among the 335 genes expressed only in the xylems of *E. globulus* and *E. urophylla* (Figure 2, group 'c'), 104 were considered differentially expressed, 57% constituted *no hits* genes and 5% presented an unknown function. Among the 62 genes that showed higher levels of expression in *E. globulus*, several encoded resistance-related proteins and xyloglucan endotransglucosylase/hydrolase (XTH). Additionally, in *E. globulus* xylem, the differential expression of genes encoding a sarcosine oxidase and a nitrate reductase was identified. Many resistance-related genes and ATP phosphoribosyltransferases were included among the 42 genes that showed higher levels of expression in *E. urophylla* (Additional File 3, Figure S5 B).

Of the 369 contigs present only in *E. grandis* and *E. urophylla* (Figure 2, group 'd'), 50% constituted *no hits* genes and approximately 6% were genes with unknown functions. Only 87 genes were considered differentially expressed: 40 showed higher levels of expression in *E. grandis* and 47 showed higher levels of expression in *E. urophylla* (Additional File 3, Figure S5 C). Several of these genes were identified as being related to disease resistance and heat shock proteins.

From the analysis of the functional profiles of these groups, it is possible to suggest that stress, disease and pathogen response-related genes presented highly variable levels of expression in the xylem programming of the studied *Eucalyptus* species.

Contigs with FPKM>0 in only one species (Figure 2, groups 'e', 'f' and 'g')

The groups of genes expressed in only one of the species studied here included a high number of *no hits* genes (52%). The contigs encoding known proteins (48%) were represented

by genes displaying similar functions, as demonstrated by the Gene Ontology analysis (Additional File 3, Figure S4C).

Several stress resistance-related genes of these groups were found in all of the species. We identified the presence of sarcosine oxidase and expansin genes in *E. globulus*, pyrophosphatase and exportin in *E. grandis*, and ATP phosphoribosyltransferase, inorganic pyrophosphatase and a protein related to cell death in *E. urophylla*.

These results reinforce the concept that the expression of genes responding to environmental conditions, such as stress resistance, was one of the major differences between the *Eucalyptus* xylem transcriptomes analyzed here. Another interesting result was the observed up-regulation of a sarcosine oxidase gene in *E. globulus* xylem, as shown for group 'c' above.

Contigs with FPKM=0 in all species (Figure 2, group 'h')

Of the 18,894 *Eucspresso* contigs present in the EUCANEXT database, only 156 (less than 1%) were not detected in any of the three xylem RNAseq libraries (FPKM=0). Because *Eucspresso* represents genes from several tissues, this finding indicates that the xylem tissues analyzed in this work presented a wide range of transcriptional activity, expressing >99% of all genes identified in the *Eucalyptus* transcriptome [25].

However, of the 156 undetected genes, 48% were identified to be expressed in the xylem of the hybrid "urograndis", as described by Mizrachi *et al.* [25]. This apparent contradiction may reflect a specific biologic event, such as the presence of genes expressed as a result of the hybridization between *E. grandis* and *E. urophylla* or simply a difference in the growth environment or plant age.

Among the genes expressed in "urograndis", but not in the other tree species we found 5 representatives of AP2/ERF transcription factors. These transcription factors appear to regulate the developmental, physiological and biochemical responses of plants under varied environmental conditions, such as drought, high salinity and extreme temperatures (see review in [30]). For the remaining set of 156 genes, most were classified as *no hits* or hypothetical genes. One XTH was also identified, although a specific function could not be assigned.

No Hits

The analysis of *no hits* present in groups 'a' to 'g' within the Venn diagram (Figure 2) demonstrated that they corresponded to nearly 11% of the non-differentially expressed genes present in all three transcriptomes, which further increases to approximately 60% among the genes present in only one or two transcriptomes. This result may suggest that the constitutive genes, which were likely those responsible for primary metabolism, were previously identified in many species and thus are represented by diverse sequences in public databases. However, genes that are species-specific or even tissue-specific are remain to be identified. It is expected that the number of *Eucalyptus* genes with known functions will increase following the release of the complete genome annotation of *E. grandis*,

Functional analysis

Certain metabolisms involved in secondary wall formation or preservation were analyzed in detail and compared among the three species.

Secondary cell wall formation genes

The expression of cell wall-related genes, as well as genes that directly or indirectly influence wood formation, was different between species. Among these, genes from the cellulose synthase superfamily, the sucrose synthase family, xyloglucan endotransglucosylase/hydrolase (XTH) and expansions were detected.

Many genes of the cellulose synthase superfamily (*CesA* – *cellulose synthase* and *Csl* – *cellulose synthase like*) were observed in our data, most were differentially expressed among the species. Despite the detailed investigations that are necessary to evaluate the relative contribution of each member, in general, the cellulose synthase superfamily genes showed a lower level of expression in *E. urophylla* and/or *E. globulus*, as shown in Figure 3 (Additional File 2, Tables S1).

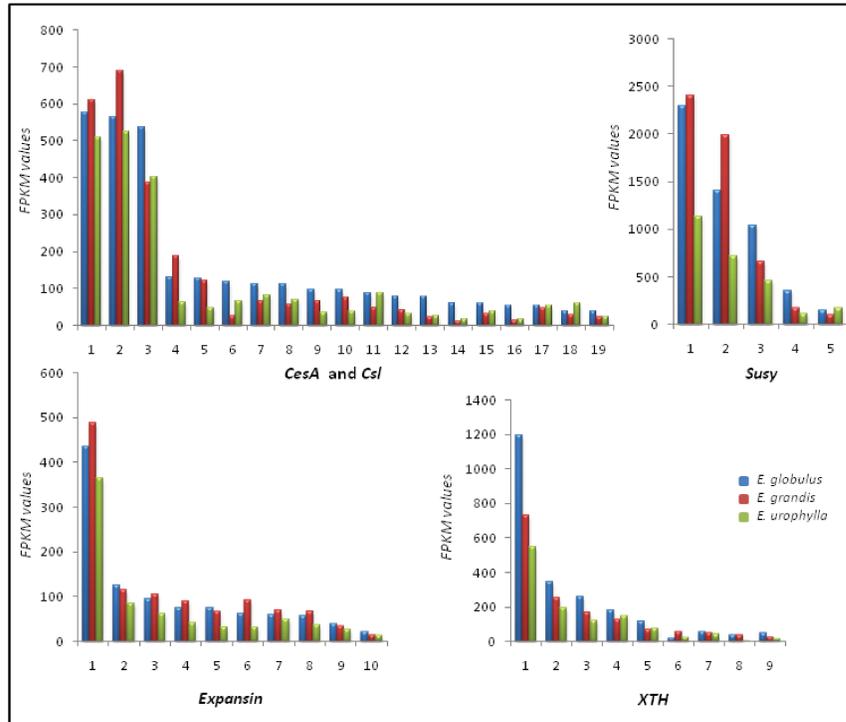


Figure 3: Cell wall-related genes: The expression profiles of selected genes related to cell wall construction. Numbers on the x-axis represent contigs listed in Table S1. The y-axis represents the FPKM values.

Members of the sucrose synthase (*Susy*), *XTH* and expansin gene families were found to be differentially expressed among the three species. The results indicated that the *Susy* and expansin gene families had a lower level of expression in *E. urophylla* compared to the other two species. Several contigs were identified as *XTH* genes, and although not all members of this family were differentially expressed, the majority were expressed at higher and lower levels in *E. globulus* and *E. urophylla*, respectively. These results are shown in Figure 3 and Additional File 2 (Table S1).

The importance of cell wall genes is highlighted by previously published studies. For example, the cellulose synthase superfamily is responsible for cellulose (*CesA*) and hemicellulose (*Csl*) biosynthesis and deposition [31, 32]. The sucrose synthase enzyme catalyzes the reversible conversion of sucrose and UDP to UDP-glucose and fructose [33]. XTHs belong to a large family named xyloglucan endotransglucosylase/hydrolase [34] and are involved in restructuring primary cell walls to accommodate secondary cell walls [35]. Expansins are cell wall proteins involved in many cell wall modification processes including relaxation and extension in particular, which are essential for cell wall enlargement [36]. The action of these proteins results in an increase in cell wall extensibility and growth [37].

Figure 4: Simplified phenylpropanoid pathway and related genes: A, The expression profiles of selected genes related to the phenylpropanoid pathway. Numbers on the x-axis represent contigs listed in Table S1. The y-axis represents the FPKM values. B, Common compounds shared by the lignin branch (right) and flavonoids branch (left) are depicted in white. Pathway genes are shown in white boxes (*PAL* - phenylalanine ammonium lyase; *C4H* - cinnamate-4-hydroxylase; *4CL* - 4-coumarate-CoA ligase; *HCT* - hydroxycinnamoyl-CoA:quinic acid shikimate p-hydroxycinnamoyltransferase; *C3H*-Coumaroyl-CoA 3-hydroxylase; *CCoAOMT* - caffeoyl-CoA 3-O-methyltransferase; *CCR* - cinnamoyl-CoA reductase; *F5H* - ferulate 5-hydroxylase; *COMT* - caffeic acid:5-hydroxyferulic acid O-methyltransferase; *CAD* - cinnamyl alcohol dehydrogenase; *CHS* - chalcone synthase; *CHI* - chalcone isomerase; *F3H* - flavonone 3-hydroxylase; *DFR* - dihydroflavonol 4-reductase; *ANS* - anthocyanidin synthase). The colored rectangles indicate genes that showed higher levels of expression in *E. grandis* (red) or in *E. urophylla* (green).

The first three enzymes of the phenylpropanoid pathway, phenylalanine ammonium lyase (*PAL*), 4-coumarate-CoA ligase (*4CL*) and cinnamate-4-hydroxylase (*C4H*) are responsible for converting phenylalanine to p-coumaroyl-CoA [39]. Many contigs related to these three genes were found in our data (Table 1), which indicates that large gene families encode each enzyme. In our discussion, only the genes with high expression in xylem were considered. In this case, when all gene members were taken into account, the number of transcripts in *E. grandis* was greater than that of the other two species, whereas this value was lower in *E. globulus*.

A number of genes of the lignin branch were also identified. The gene families encoding enzymes coumaroyl-CoA 3-hydroxylase (*C3H*), cinnamoyl-CoA reductase (*CCR*) and cinnamyl alcohol dehydrogenase (*CAD*) presented distinct expression profiles among the members; however, the genes with the highest levels of expression were observed in *E. grandis*. Conversely, the gene families encoding the enzymes hydroxycinnamoyl-CoA: quinic acid p-hydroxycinnamoyltransferase (*HCT*), caffeoyl-CoA 3-O-methyltransferase (*CCoAOMT*) and caffeic acid: 5-hydroxyferulic acid O-methyltransferase (*COMT*) presented a similar expression profile in the xylems of the three species and showed a higher level of expression in *E. grandis*.

In contrast, various genes of the flavonoid branch of the phenylpropanoid pathway, namely those encoding the enzymes chalcone synthase (*CHS*), chalcone isomerase (*CHI*), flavonone 3-hydroxylase (*F3H*) and dihydroflavonol 4-reductase (*DFR*), were expressed at significantly higher levels in *E. urophylla* compared to the two other species.

In general, our results indicated that the lignin genes showed higher levels of expression in *E. grandis* compared to *E. globulus* and *E. urophylla*, whereas genes present in the flavonoid branch showed higher levels of expression in *E. urophylla*. These results illustrate the complexity of the molecular events underlying lignification, as the species that apparently presents a superior lignin content, *E. urophylla*, appears to deviate from the phenylpropanoid pathway towards flavonoid formation.

The flavonoid branch of the phenylpropanoid pathway is known to be up-regulated by various environmental stressors [40]. In addition, flavonoids were reported to have protective functions in response to a broad variety of stressors [41]. The up-regulation of this branch in *E. urophylla* suggests that this species has defense mechanisms that enable it to survive in a wide range of environments.

Transcription factors

Secondary cell wall formation is regulated by several transcription factors, such as WRKY [42, 43], NAC [44, 45], MYB [46 - 48] and SHN [49]. Many MYB and WRKY transcription factors were identified as being differentially expressed among the xylems of the three species studied here, and in most cases, they presented lower levels of expression in *E. globulus* (Figure 5A and Additional File 2, Table S2).

Another interesting class of transcription factors present in *E. grandis*, *E. globulus* and *E. urophylla* xylem transcriptomes is the AP2/ERF class. This is a large family of transcription factors that is implicated in the abiotic stress response [30, 50]. Many members of this family were expressed in the xylems of the three *Eucalyptus* species, and as above, these genes were down-regulated in *E. globulus* (Figure 5A and Additional File 2, Table S2). Interestingly, as mentioned above, certain members of the AP2/ERF class were absent in all xylems studied here, and these genes were previously discussed in the section “*Contigs with FPKM=0 in all species*”.

The expression of *Myb* and *WRKY* transcription factors, which are involved in many different developmental processes and defense responses [42, 47], together with the expression of AP2/ERF indicates a possible molecular explanation of the differences in the adaptability of *E. grandis* and *E. urophylla* compared to *E. globulus*, a species that is poorly adapted to tropical conditions.

Members of the *Myb* and *WRKY* families are also involved in phenylpropanoid pathway regulation [51, 52]. Importantly, one recent work showed that mutations in *WRKY* transcription factors are associated with increased stem biomass [53] and the over-expression of these factors led to improved stress resistance [54]. In this context, it is interesting to observe that the expression of transcription factors followed the same profile of genes present in the phenylpropanoid pathway (Figure 5); i.e., these factors were more strongly expressed in *E. grandis* and *E. urophylla* relative to *E. globulus*.

Nitrogen metabolism genes

In *E. globulus*, genes for nitrate and nitrite reductases and sarcosine oxidase were expressed at higher levels compared to the other two species (Figure 5B and Additional File 2, Table S2).

Nitrate reductase is an essential enzyme for the metabolism of nitrogen and is responsible for the reduction of nitrate into nitrite which is in turn reduced to ammonia by nitrite reductase [55]. Studies of nitrogen supply performed in *Populus* [26, 56] and *Eucalyptus* (Camargo *et al.*, unpublished data) have suggested that nitrogen availability may be associated with an increase in cellulose and hemicelluloses contents as well as a decrease in the total amount of lignin in cell walls. Thus, the results indicating an increased expression of nitrate and nitrite reductase genes in the xylem of *E. globulus* suggest a possible mechanism by which the wood of this species presents superior cellulose productivity relative to the other species evaluated here.

Several enzymes are involved in nitrogen assimilation in plants. Ammonia is converted to glutamine and glutamate, and the nitrogen present in these compounds is then incorporated in other amino acids by enzymes known as aminotransferases [57].

Goyer *et al.* [58] showed that *A. thaliana* can assimilate nitrogen in glycine using sarcosine present in the environment through the action of a sarcosine oxidase; however, the precise mechanism of action of this enzyme remains elusive. The sarcosine oxidase gene was up-regulated in *E. globulus*; therefore, if the function of this gene is conserved with that of the same gene in *A. thaliana*, this species may use sarcosine as a nitrogen source, thereby conveying a superior advantage over the other studied species. Glycine metabolism is important for the biosynthesis of cell wall structural proteins (glycine-rich proteins) [59]. The potential of this protein to modify cell wall structures was previously studied in *Pinus* using a transcriptional analysis [60]. Based on this evidence, our results indicate that the presence of glycine may contribute to a more elaborate cell wall architecture in *E. globulus*.

Ubiquitins and heat shock genes

Other results relevant to wood production by the studied species include the differential expression of ubiquitins and heat shock proteins (Figure 5C and Additional File 2, Table S2).

Ubiquitins are involved in developmental plasticity and environmental adaptation and can play major roles in almost all aspects of plant growth and development [61, 62]. In addition, these proteins were found to be involved in the regulation of xylem differentiation, thereby allowing cells to initiate and progress through the stages of xylogenesis [16]. Many contigs were found to be related to ubiquitin genes, and among the most strongly expressed genes, there was a tendency of lower expression in *E. globulus*. However, several members were up-regulated in *E. urophylla*, which suggests that these species present different physiological responses during growth, development and environmental adaptation.

Heat shock proteins (HSP) are important in maintaining protein homeostasis inside cells and promoting the proper folding, stability and degradation of polypeptides [63]. In addition, they act as molecular chaperones, regulating protein maturation and the transition between the inactive and active states of signaling molecules, such as receptors, protein kinases and transcriptional regulators [64]. Furthermore, many classes of heat shock proteins are involved in abiotic stress responses [65]. Moreover, previous studies have suggested a role of molecular chaperones in wood formation in *Eucalyptus* and *Pinus*, thereby emphasizing the importance of HSP genes as candidate genes for the genetic improvement of wood quality [13, 66, 67].

Our results indicated the differential expression of several classes of HSPs among the three *Eucalyptus* species. However, we observed a clear expression pattern only for small heat shock proteins (smHSP): smHSP expression was induced in *E. urophylla* and repressed in *E. globulus* xylems (Additional File 2, Table S2 and Additional File 3, Figure S4 E). Because small HSPs appear to be involved in stress tolerance [68], we speculated that *E. urophylla* has a superior ability to overcome environmental challenges, such as biotic and abiotic stresses. In addition, these results support the reduced adaptability to tropical areas that has been observed for *E. globulus*.

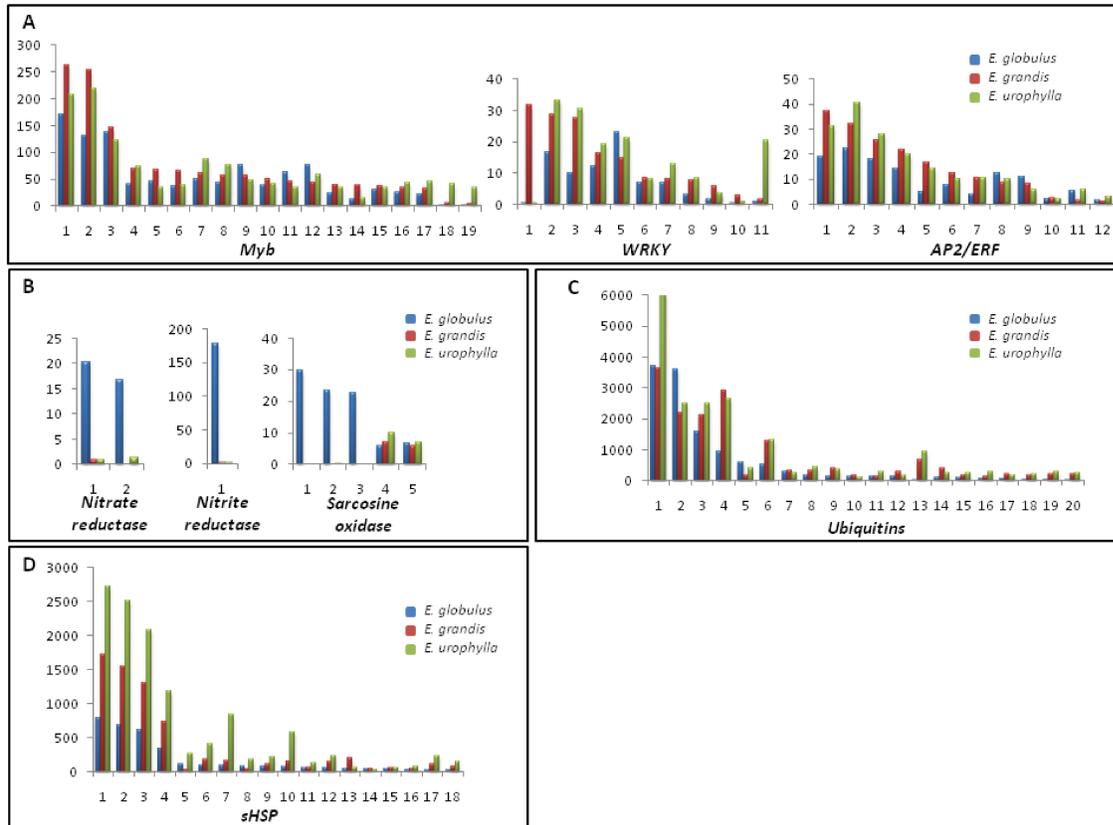


Figure 5: Important selected genes: The expression profiles of selected genes related to: A, transcription factors; B, nitrogen metabolism; C, ubiquitins; and D, heat shock proteins. The numbers on the x-axis represent the contigs listed in Table S2. The y-axis represents the FPKM values.

Disease and stress-resistance genes

With respect to disease- and stress-resistance genes, interestingly, these groups showed a high level of expression in the xylems of *E. globulus* and *E. urophylla* (Figure 6 and Additional File 2, Table S2).

Among these genes, we identified a large class of disease-resistance R genes, such as CC-NBS-LRR (CC) genes that are key plant disease-resistance genes [69] and many members of the multidrug resistance (MDR) protein family. Although the large MDR gene family has been found in other plants, their roles have not been well studied. The association of MDR genes with the transport of compounds from the environment, such as herbicides and antibiotics as well as metabolites, such as flavonoids, which allows plants to co-exist with toxins, has been previously reported [70].

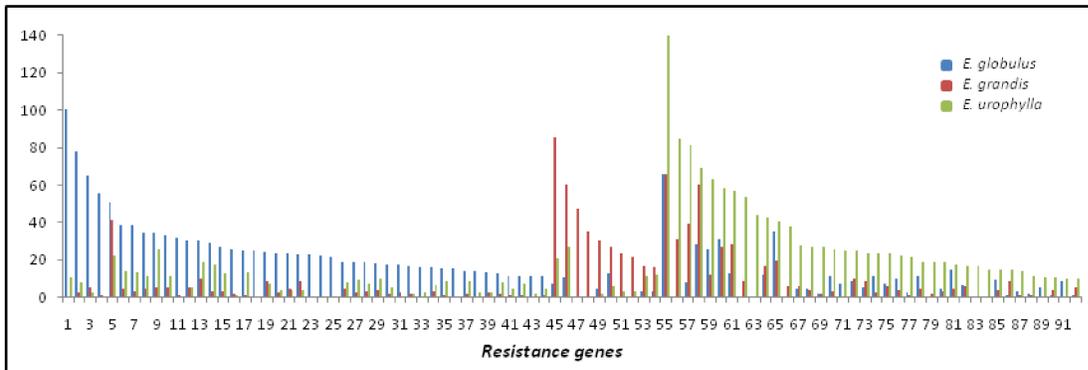


Figure 6: Resistance-associated genes: The expression profiles of selected genes related to resistance. Numbers on the x-axis represent the contigs listed in Table S2. The y-axis represents the FPKM values.

Although greater numbers of CC genes were expressed in the xylem of *E. globulus*, MDR genes showed a higher level of expression in the xylem of *E. urophylla*. Because these genes (CC and MDR) are associated with a variety of biotic and abiotic stresses, it is possible that their expression is related to the cultivation of these *Eucalyptus* species in areas outside of their natural habitat. However, a thorough investigation should be performed to understand the nature of these differences.

Taken together, the results demonstrating a high number of stress tolerance genes and the deviation of the phenylpropanoid pathway toward flavonoid formation in *E. urophylla* suggest that this species is naturally genetically programmed to tolerate several types of environmental stresses. In contrast, *E. globulus* grows poorly in tropical areas despite the expression of a large number of disease-resistance genes. These findings highlight the importance of stress tolerance genes in the adaptability of trees from temperate locations to tropical areas.

Conclusion:

The three species studied here are of economic importance in the forest industry mainly for use as pulp, paper and, more recently, biofuels. These species are common sources of genes for breeding programs due to their phenotypic and agronomic characteristics. The RNAseq performed here provides a transcriptional overview of multigene family participation in *Eucalyptus* wood formation. Thus, our results may contribute to an understanding of the mechanisms that determine the quality of premium wood found in *E. globulus*. The following results are of particular interest to the pulp and paper industry. The coordinated expression of genes leads to secondary cell wall formation and the down regulation of lignin genes. The expression of lignin genes in *E. grandis*, which was more strongly induced relative to *E. globulus*

and *E. urophylla*, illustrates the complexity of the wood formation process. The production of flavonoids coupled with the large number of stress response genes indicate that the genetic programming of *E. urophylla* allows this species to grow in diverse environmental conditions. The large number of *no hits* genes show that many of the transcriptional differences among the species are unknown. We believe that an understanding of xylogenesis may help in the elucidation of the molecular basis of differences in wood quality and will allow for modifications in breeding programs toward better gene combinations.

Methodology:

Plant materials

Samples of xylem tissues of three species (*E. globulus*, *E. grandis* and *E. urophylla*) were collected from three-year-old plants from the International Paper fields in Mogi Guaçu, SP, Brazil (latitude (S): 22°21'; longitude (W): 46°58'). The trees were grown in fields having an experimental area of approximately 100 m². Xylem tissue samples were harvested, immediately frozen in liquid nitrogen and stored at -80°C.

RNA extraction

The RNA material was extracted according to the protocol described by Zeng and Yang [71] with modifications proposed by Provost *et al* [72].

mRNA sequencing

The mRNA sequencing was performed at the High Throughput Sequencing Facility at the Carolina Center for Genome Sciences (University of North Carolina, USA). For each xylem sample, 10 µg of total RNA were used to prepare the mRNAseq library according to the protocol provided by Illumina. The gel extraction step was modified by dissolving excised gel slices at room temperature to avoid the underrepresentation of AT-rich sequences [73]. Library quality control and quantification were performed using a Bioanalyzer Chip DNA 1000 series II (Agilent). For each library, 36 bp single-end sequences were generated in one lane of an Illumina Genome Analyzer II_x.

Public data sets

A publically available gene catalog of a commercially grown *E. grandis* X *E. urophylla* hybrid (“urograndis”) was downloaded from *Eucspresso* (<http://eucspresso.bi.up.ac.za>). The data set consisted of 18,994 contigs greater than 200 bp,

and 99.83% of contigs had similarities with the *E. grandis* genome assembly [25]. The *Eucalyptus grandis* genome assembly was downloaded from Phytozome (<http://www.phytozome.net/Eucalyptus>).

Read alignment

The Illumina reads were aligned against the *Eucspresso* contigs using the SOAP2 aligner [74]. To prepare the data for Genebrowser analysis, reads were aligned against the *E. grandis* genome using TopHat [75] to allow for spliced alignments. Both programs were configured to allow up to two mismatches (SNPs can generate mismatches in the alignment, especially in these cases because the sequences are from different species), discard sequences with ambiguities (Ns) and return all optimal alignments.

De novo assembly

The *de novo* assembly (without a reference) was performed using reads that did not map against the public data set (previously described) using the Trinity assembler [76]; contigs of at least 200 bp were allowed, and the parameter “--run_butterfly” was used.

Contig annotation

The Autofact program [77] was used to perform an automatic annotation of all EUCANEXT contigs. The main feature of Autofact is its capacity to resume the annotation based on sequence similarity searches in several databases. BLASTx [78] (e-value cutoff of $1e^{-5}$) was used to align the contigs against the following public databases: the non-redundant (NR) database of *NCBI*; the Uniref90 and Uniref100 database, which containing clustered sets of proteins from Uniprot [79]; the Pfam database of proteins families [80]; the KEGG database of metabolic pathways [81]; and TAIR (version 10), which is a database of *Arabidopsis* proteins. Functional annotation (GO) was performed using BLAST2GO [82] with the default parameters.

Determination of gene expression levels

The expression levels for each gene were estimated using the FPKM (fragments per kilobase of exon per million fragments) mapped value [83]. Three pairwise comparisons (*E. grandis* X *E. globulus*, *E. grandis* X *E. urophylla* and *E. globulus* X *E. urophylla*) between the FPKM values of each species were performed. A differentially expressed gene was defined as having a greater than two-fold difference in FPKM values between species based on a T-test with a 99% confidence rate (cut-off of 0.01).

List of abbreviations

4CL - 4-coumarate-CoA ligase
AP2/ERF - Apetala2/ethylene response factor
BLAST - Basic local alignment search tool
C3H - Coumaroyl-CoA 3-hydroxylase
C4H - Cinnamate-4-hydroxylase
CAD - Cinnamyl alcohol dehydrogenase
CC-NBS-LRR - Coiled-coil –nucleotide-binding site-leucine-rich repeat protein
CCoAOMT - Caffeoyl-CoA 3-O-methyltransferase
CCR - Cinnamoyl-CoA reductase
CesA - Cellulose synthase
Csl - Cellulose synthase-like
CHI - Chalcone isomerase
CHS - Chalcone synthase
COMT - Caffeic acid:5-hydroxyferulic acid O-methyltransferase
DArT - [Diversity arrays technology](#)
DRF - Dihydroflavonol 4-reductase
EST - Expressed sequence tag
F3H - Flavonone 3-hydroxylase
F5H - Ferulate 5-hydroxylase
FPKM - Fragments per kilobase of exons per million
GO - Gene ontology
HCT - Hydroxycinnamoyl-CoA: quinate shikimate p–hydroxycinnamoyltransferase
HSP - Heat shock proteins
MDR - Multi drug resistance
MyB - (myeloblastosis) family of [transcription factors](#)
Nac - (NAM, ATAF1,2 and CUC2) family of [transcription factors](#)
NCBI - National Center for Biotechnology Information
NR - Non-redundant GenBank database
PAL - Phenylalanine ammonium lyase
RT-qPCR - Real time quantitative polymerase chain reaction
SAGE - [Serial analysis of gene expression](#)
SHN - SHINE transcription factor
Susy - Sucrose synthase
WRKY - WRKY aminoacid domain of transcription factor
XTH - xyloglucan endotransglucosylase/hydrolase

Authors' contributions

MMS: preparation of the manuscript, sampled the material, prepared the libraries and participated in data analysis; LCN: bioinformatics analysis, construction of EUCANEXT database; ELOC, DCG, JLN, WLM: collaboration in sampling the material, preparation of libraries and data analysis; PJPLT, PM: sequencing of libraries by RNAseq, collaboration in data analysis; JMCM, ACD: coordination of data analysis and revision of manuscript. MFC: coordination of bioinformatics and data analysis; GAGP: coordination of the molecular and bioinformatics analysis and organization of the manuscript. All authors have read and approved the final version of the manuscript.

Authors' information

¹Laboratório de Genômica e Expressão, Departamento de Genética Evolução e Bioagentes, Instituto de Biologia, Universidade Estadual de Campinas, CEP: 13083-970, Campinas, São Paulo, Brasil.

²Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill (UNC), Chapel Hill, NC, USA. ³Centro de Pesquisa e Desenvolvimento em Recursos Genéticos vegetais, Instituto Agrônomo de Campinas, CEP: 13001-970, Campinas, São Paulo, Brasil.

Acknowledgements

The authors would like to acknowledge to International Paper for kindly provided plant material and EUCAGEN team, especially Dr. Alexander Myburg from University of Pretoria. These sequence data were produced by the US Department of Energy Joint Genome Institute <http://www.jgi.doe.gov/> in collaboration with the user community. This work was supported by research funding of Fapesp (process number 2007/54877-0) and grants from International Paper (IP/IB/Gene Discovery : 3972).

References

1. Ladiges PY, Udovicic F, Nelson G: **Australian biogeographical connections and the phylogeny of large genera in the plant family *Myrtaceae***. *J Biogeogr* 2003, **30**:989 – 998.
2. Grattapaglia D, Kirst M: ***Eucalyptus* applied genomics: from gene sequences to breeding tools**. *New Phytol* 2008, **179**:911 – 929.
3. Turnbull JW: ***Eucalypt* plantations**. *New Forest* 1999, **17**:37 – 52.
4. Paiva JAP, Prat E, Vautrin S, Santos, MD, San-Clemente H, Brommonschenkel S, Fonseca, PGS, Grattapaglia D, Song X, Ammiraju JSS, Kudrna D, Wing RA, Freitas AT, Bergès H, Grima-Pettenati J: **Advancing *Eucalyptus* genomics: identification and sequencing of lignin biosynthesis genes from deep-coverage BAC libraries**. *BMC Genomics* 2011, **12**:137.
5. Grattapaglia D: **Genomics of *Eucalyptus*, a global tree for energy, paper, and wood**. In *Genomic of tropical crop plants. Volume 1*. Edited by Moore PH, Ming R. New York: Springer; 2008: 259 – 297.
6. Rodrigues J, Meier D, Faix O, Pereira H: **Determination of tree to tree variation in syringyl/guaiacyl ratio of *Eucalyptus globulus* wood lignin by analytical pyrolysis**. *J Anal Appl Pyrol* 1999, **48**:121 – 128.
7. Grattapaglia D: **Integrating genomics into *Eucalyptus* breeding**. *Genet Mol Res* 2004, **3**:369-379.

8. Bison O, Ramalho MAP, Rezende GDSP, Aguiar AM, De Resende MDV: **Combining ability of elite clones of *Eucalyptus grandis* and *Eucalyptus urophylla* with *Eucalyptus globulus*.** Genet Mol Biol 2007, **30(2)**: 417 – 422.
9. Turnbull JW: **Economic and social importance of *Eucalyptus*.** In *Diseases and pathogens of Eucalyptus*. 1st edition. Edited by Keane PJ, Kile GA, Podger FD, Brown BN. Collingwood: CSIRO Publishing, 2000:1- 9.
10. Potts BM, Pederick LA: **Morphology, phylogeny, origin, distribution and genetic diversity of *Eucalyptus*.** In *Diseases and pathogens of Eucalyptus*. 1st edition. Edited by Keane PJ, Kile GA, Podger FD, Brown BN. Collingwood: CSIRO Publishing, 2000:11- 34.
11. Fukuda H: **Xylogenesis: initiation, progression, and cell death.** Annu Rev Plant Phys 1996, **47**:299 – 325.
12. Plomion C, Leprovost G, Stokes A: **Wood formation in trees.** Plant Physiol 2001, **127**:1513 – 1523.
13. Wang M, Qi X, Zhao S, Zhang S, Lu M: **Dynamic changes in transcripts during regeneration of the secondary vascular system in *Populus tomentosa* Carr. revealed by cDNA microarrays.** BMC Genomics 2009, **10**:215.
14. Yang SS, Tu ZJ, Cheung F, Xu WW, Lamb JFS, Jung HG, Vance CP, Gronwald JW: **Using RNA-seq for gene identification, polymorphism detection and transcript profiling in two alfalfa genotypes with divergent cell wall composition in stems.** BMC Genomics 2011, **12**:199.
15. Vicentini R, Sasaki FT, Gimenes MA, Maia IG, Menossi M: **In silico evaluation of the *Eucalyptus* transcriptome.** Genet Mol Biol 2005, **28**:487–495.
16. Paux E, Tamasloukht M, Ladouce N, Sivadon P, Grima-Pettenati J: **Identification of genes preferentially expressed during Wood formation in *Eucalyptus*.** Plant Molecular Biology. Plant Mol Biol 2004, **55**:263 – 280.
17. Kirst M, Basten CJ, Myburg AA, Zeng Z, Sederoff RR: **Genetic architecture of transcript-level variation in differentiating xylem of a *Eucalyptus* hybrid.** Genetics 2005, **169**:2295 – 2303.
18. Barros E, Staden C, Lezar S. **A microarray-based method for the parallel analysis of genotypes and expression profiles of wood-forming tissues in *Eucalyptus grandis*.** BMC Biotechnol 2009, **9(51)**.
19. Carvalho MCGG, Caldas DGG, Carneiro RT, Moon DH, Salvatierra GR, Franceschini LM, Andrade A, Celedon PAF, Oda S, LABATE CA: **SAGE transcript profiling of the juvenile cambial region of *Eucalyptus grandis*.** Tree physiol 2008, **25**:905 – 919.
20. Sansaloni CP, Petroli CD, Carling J, Hudson CJ, Steane DA, Myburg AA, Grattapaglia D, Vaillancourt RE, Kilian A: **A high-density Diversity Arrays Technology (DArT) microarray for genome-wide genotyping in *Eucalyptus*.** Plant Methods, **6(16)**.

21. Tournier V, Grat S, Marque C, Kayal WE, Penchel R, Andrade G, Boudet AM, Teulières C: **An efficient procedure to stably introduce genes into a economically important pulp tree (*Eucalyptus grandis* X *Eucalyptus urophylla*).** Transgenic Res 2003, **12(4)**:403 – 411.
22. Kawaoka A, Nanto K, Ishii K, Ebinuma H. **Reduction of lignin content by suppression of expression of the LIM domain transcription factor in *Eucalyptus camaldulensis*.** Silvae Genet 2006, **55(6)**:269 – 277.
23. Wang B, Guo G, Wang C, Lin Y, Wang X, Zhao M, Guo Y, He M, Zhang Y, Pan L: **Survey of the transcriptome of *Aspergillus oryzae* via massive parallel mRNA sequencing.** Nucleic Acids Res 2010, **38(15)**:5075–5087.
24. Liu S, Lin L, Jiang P, Wang D, Xing Y. **A comparison of RNA-Seq and high-density exon array for detecting differential gene expression between closely related species.** Nucleic Acids Res 2011, **39(2)**:578 – 588.
25. Mizrachi E, Hefer CA, Ranik M, Joubert F, Myburg AA: **De novo assembled expressed gene catalogue f a fast-growing *Eucalyptus* tree produced by Illumina mRNA-Seq.** BMC Genomics 2010, **11**:681.
26. Pitre FE, Pollet b, Lafarguette F, Cooke JEK, MacKay JJ, Lapierre C. **Effects of increased nitrogen supply on the lignifications of poplar wood.** J Agric Food Chem 2007, **55(25)**:10306 – 10314.
27. Oliveira LA, Breaton MC, Bastolla FM, Camargo SS, Margis R, Frazzon J, Pasquali G: **Reference genes for the normalization of gene expression in *Eucalyptus* species.** Plant Cell Physiol 2012, **53(2)**:405-422.
28. Huggett J, Dheda K, Bustin S, Zulma A: **Real-time RT-PCR normalization; strategies and considerations.** Genes Immun 2005, **6**:279 – 284.
29. Radonic A, Thulke S, MacKay IM, Landt O, Siegert W, Nitsche A: **Guideline to reference gene selection for quantitative real-time PCR.** Biochem Bioph Res Co 2005, **313(4)**:856 – 862.
30. Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K: **AP2/ERF family transcription factors is plant abiotic stress responses.** Biochem Biophys Acta 2012, **1819(2)**:86 – 96.
31. Taylor NG, Howells RM, Huttly AK, Vickers K, Turner SR: **Interactions among three distinct *CesA* proteins essential for cellulose synthesis.** P Natl Acad Sci USA 2003, **100(3)**:1450–1455.
32. Carpita NC: **Update on mechanisms of plant cell wall biosynthesis: how plants make cellulose and other (1 →4)-β-D-glycans.** Plant Physiol 2011, **155**:171 – 184.
33. Baud S, Vaultier M, Rochat C: **Structure and expression profile of the sucrose synthase multigene family in *Arabidopsis*.** J Exp Bot 2004, **55 (396)**:397 – 409.
34. Rose JKC, Braam J, Fry SC, Nishitani K: **The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature.** Plant Cell Physiol 2202, **43 (12)**:1421-1435.

35. Bourquin V, Nishikubo N, Abe H, Brumer H, Denman S, Eklund M, Christiernin M, Teeri TT, Sundberg B, Mellerowicz EJ: **Xyloglucan endotransglycosylases have a function during the formation of secondary cell walls of vascular tissues.** *Plant Cell* 2002, **14**:3073 – 3088.
36. Sampedro J, Cosgrove DJ: **The expansin superfamily.** *Genome Biol* 2005, **6**(242).
37. Cosgrove, DJ: **Growth of the plant cell wall.** *Nat Rev Mol Cell Bio* 2005, **6**: 850 – 861.
38. Ferrer JL, Austin MB, Stewart JR. C, Noel JP: **Structure and function of enzymes involved in the biosynthesis of phenylpropanoids.** *Plant Physiol Biochemistry.* *Plant Physiol Biochem* 2008, **46**(3):356-370.
39. Boerjan W, Ralph J, Baucher M: **Lignin biosynthesis.** *Annu Rev Plant Biol* 2003, **54**:519 – 546.
40. Olsen KM, Sliemstad R, Lea US, Brede C, Lovdal T, Ruoff P, Verheul M, Lillo C: **Temperature and nitrogen effects on regulators and products of the flavonoid pathway: experimental and kinetic model studies.** *Plant Cell Environ* 2009, **32**:286 – 299.
41. Treutter D: **Significance of flavonoids in plant resistance: a review.** *Environ chem. Lett* 2006, **4**:147 – 157.
42. Ulker B, Somssich IE: **WRKY transcription factors: from DNA binding towards biological function.** *Curr Opin Plant Biol* 2004, **7**:491 – 498.
43. Wang Q, Wang M, Zhang X, Hao B, Kaushik SK, Pan Y: **WRKY gene family evolution in *Arabidopsis thaliana*.** *Genetica* 2011, **139**(8):973 - 983.
44. Mitsuda N, Iwase A, Yamamoto H, Yoshida M, Seki M, Shinozaki K, Ohme-Takagi M: **NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in Woody tissues in *Arabidopsis*.** *Plant Cell* 2007, **19**:270 – 280.
45. Ohashi-Ito k, Oda Y, Fukuda H: ***Arabidopsis* vascular related NAC-domain6 directly regulates the genes that govern programmed cell death and secondary wall formation during xylem differentiation.** *Plant Cell* 2010, **22**(10):3461 – 3473.
46. Goicoechea M, Lacombe E, Legay S, Mihaljevic S, Rech P, Jauneau A, Lapierre C, Pollet B, Verhaegen D, Chaubet-Gigot N, Gruma-Pettenati J: **EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis.** *Plant J* 2005, **43**:553 – 567.
47. Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G, Zhiqiang L, Yunfei Z, Xiaoming Q, Yunping S, Li Z, Xiaohui D, Jingchu L, Xing-Wang D, Zhangliang C, Hongya G, Li-Jia Q: **The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family.** *Plant Mol Biol* 2006, **60**:107 – 124.
48. Legay S, Sivadon P, Blervacq A, Pavy N, Baghdady A, Tremblay L, Lévassieur C, Ladouce N, Lapierre C, Seguin a, Hawkins S, MacKay J, Grima-Pettenati J: **EgMYB1, an R2R3 MYB transcription factor from *Eucalyptus* negatively regulates secondary cell wall formation in *Arabidopsis* and poplar.** *New Phytol* 2010, **188**(13):774 – 786.

49. Ambavaram MMR, Krishnan A, Trijatmiko KR, Pereira A: **Coordinated Activation of Cellulose and Repression of Lignin Biosynthesis Pathways in Rice.** *Plant Physiol* 2010, **155**:916-931.
50. Park HY, Seok HY, Woo DH, Lee SY, Tarte VN, Lee EH, Lee CH, Moon YH: **AtERF71/HRE2 transcription factor mediates osmotic stress response as well as hypoxia response in *Arabidopsis*.** *Biochem Bioph Res Co* 2011, **414(1)**:135 – 141.
51. Endt DV, Kijne JW, Memelink J: **Transcription factors controlling plant secondary metabolism: what regulates the regulators?** *Phytochemistry* 2002, **61(2)**:107 - 114.
52. Hichri I, Barrieu F, Bogs J, Kappel C, Delrot S, Lauvergeat V: **Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway.** *J Exp Bot* 2011, **62(8)**:2465 - 2483.
53. Wang H, Avci U, Nakashima J, Hahn MG, Chen F, Dixon RA: **Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants.** *P NATL ACAD SCI USA* 2010, **107(51)**.
54. Guillaumie S, Mzid R, Mechin V, Leon C, Hichri I, Destrac-Irvine A, Trossat-Magnin C, Delrot S, Lauvergeat V: **The grapevine transcription factor WRKY influences the lignin pathway and xylem development in tobacco.** *Plant Mol Biol* 2010, **72(1-2)**:215 - 234.
55. Hoff T, Stummann BM, Henningsen KW: **Structure, function and regulation of nitrate reductase in higher plants.** *Physiol Plantarum* 1992, **84(4)**:616 - 624.
56. Pitre FE, Lafarguette F, Boyle B, Pavy N, Caron S, Dallaire N, Poulin PL, Ouellet M, Morency MJ, Wiebe N, Lim EL, Urbain A, Mouille G, Cooke JEK, MacKay JJ, Peter G: **High nitrogen fertilization and stem leaning have overlapping effects on wood formation in poplar but invoke distinct molecular pathways.** *Tree Physiol* 2010, **30(10)**:1273 - 1289.
57. Taiz L, Zaiger E: *Plant Physiology.* Sunderland: Sinauer Associates.2003
58. Goyer A, Johnson TL, Olsen LJ, Collakova E, Shachar-Hill Y, Rhodes D, Hanson AD: **Characterization and metabolic function of a peroxisomal sarcosine and pipercolate oxidase from *Arabidopsis*.** *J Biol Chem* 2004, **279 (17)**:16947 – 16953.
59. Ringli C, Keller B, Ryser U: **Glycine-rich proteins as structural components of plant cell walls.** *Cell Mol Life Sci* 2001, **58**:1430 - 1441.
60. Zhang Y, Sederoff RR, Allona I: **Differential expression of genes encoding cell wall proteins in vascular tissues from vertical and bent loblolly pine trees.** *Tree Physiol* 2000, **20(7)**:457 - 466.
61. Stone SL, Callis J: **Ubiquitin ligases mediate growth and development by promoting protein death.** *Curr Opin Plant Biol* 2007, **10(7)**:624 - 632.
62. Miura K, Lee J, Miura T, Hasegawa PM: **SIZ1 controls cell growth and plant development in *Arabidopsis* through salicylic acid.** *Plant Cell Physiol* 2010, **51(1)**:103 – 113.
63. Tirolì-Cepeda AO, Ramos CHI: **An Overview of the Role of Molecular Chaperones in Protein Homeostasis.** *Protein Peptide Lett* 2011, **18(2)**:101-9.

64. Gaestel M: **Molecular Chaperones in Signal Transduction in Molecular Chaperones in Health and Disease**. *Handb Exp Pharmacol* 2006, **172**:93 – 109.
65. Vinocur B, Altman A: **Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations**. *Curr Opin Biotechnol* 2005, **16(2)**:123 – 132.
66. Plomion C, Pionneau C, Brach J, Costa P, Baille`res H: **Compression Wood-Responsive Proteins in Developing Xylem of Maritime Pine (*Pinus pinaster* Ait.)**. *Plant Physiol* 2000, **123**:959–969.
67. Gion JM, Lalanne C, Le Provost G, Ferry-Dumazet H, Paiva J, Chaumeil P, Frigerio JM, Brach J, Barre A, Daruvar A, Claverol S, Bonneu M, Sommerer N, Negrone L, Plomion C: **The proteome of maritime pine wood forming tissue**. *Proteomics* 2005, **5**:3731–3751.
68. Wang W, Vinocur B, Shoseyov O, Altman A: **Role of plant heat-shock proteins and molecular chaperons in the abiotic stress response**. *Trends Plant Sci* 2004, **9(5)**:244 – 252.
69. Pan Q, Wendel J, Fluhr R: **Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes**. *J Mol Evol* 2000, **50(3)**:203 – 213.
70. Conte SS, Lloyd AM: **Exploring multiple drug and herbicide resistance in plants-spotlight on transporter proteins**. *Plant Sci* 2011, **180**:196 – 203.
71. Zeng Y, Yang T: **RNA isolation from highly viscous samples rich in polyphenols and polysaccharides**. *Plant Mol Biol Rep* 2002, **20**:417 – 417.
72. Provost GL, Herrera R, Paiva JA, Chaumeil P, Salin F, Plomion C: **A micromethod for high throughput RNA extraction in forest trees**. *Biol Res* 2007, **40**:291 – 297.
73. Quail MA, Kozarewa I, Smith F, Scally A, Stephens PJ, Durbin R, Swerdlow H, Turner DJ: **A large genome center's improvements to the Illumina sequencing system**. *Nat Methods* 2008, **5**:1005–1010.
74. Li R, Yu C, Li Y, Lam T-W, Yiu S-M, Kristiansen K, Wang J: **SOAP2: an improved ultrafast tool for short read alignment**. *Bioinformatics* 2009, **25(15)**: 1966-1967.
75. Trapnell C, Pachter L, Salzberg, SL: **TopHat: discovering splice junctions with RNA-Seq**. *Bioinformatics* 2009, **25(9)**:1105-1111.
76. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Haconen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A: **Full-length transcriptome assembly from RNA-Seq data without a reference genome**. *Nat Biotech* 2011, **29**:644 – 652.
77. Koski LB, Gray LW, Lang BF, Burger G: **AutoFACT: An Automatic Functional Annotation and Classification Tool**. *BMC Bioinformatics* 2005, **6**:151.
78. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs**. *Nucl. Acids Res* 1997, **25(17)**:3389-3402.

79. Suzek BE, Huang H, McGarvey P, Mazumber R, Wu: **Uniref: comprehensive and non-redundant UniProt reference clusters**. *Bioinformatics* 2007, **23(10)**:1282-1288.
80. Bateman A, Birney E, Cerruti L, Durbin R, Ewlinger L, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer ELL: **The Pfam Protein Families Database**. *Nucl Acids Res* 2002, **30(1)**:276-280.
81. Kanehisa M, Goto S: **KEGG: Kyoto Encyclopedia of Genes and Genomes**. *Nucl Acids Res* 2000, **28(1)**:27-30.
82. Conesa A, Gotz S, García-Gómez JM, Terol J, Talón M, Robles M: **Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research**. *Bioinformatics* 2005, **21(18)**:3674-3676.
83. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq**. *Nature Methods* 2008, **5**:621-628.

CAPÍTULO 2: Caracterização da parede celular secundária de três espécies de *Eucalyptus*

5. Introdução

Lignina e carboidratos compreendem os principais componentes químicos da madeira. Em menor quantidade estão os extrativos e materiais inorgânicos. Em resumo, a madeira tem uma composição de aproximadamente 50% de carbono, 6% hidrogênio, 44% de oxigênio e pouca quantidade de íons metálicos (Pettersen, 1984). Tal composição se dá principalmente pela presença de parede celular secundária na maioria das células que formam a madeira (Salisbury & Ross, 1991).

A composição química da madeira, entretanto, não pode ser precisamente definida para uma determinada espécie ou até mesmo uma árvore. Esta composição varia de acordo com a porção da árvore analisada (tronco, galho), tipo de madeira (tensão, reação), localização geográfica, clima e solo (Pettersen, 1984). As diferenças fenotípicas encontradas nas três espécies aqui estudadas (*E. globulus*, *E. grandis* e *E. urophylla*), assim como outras propriedades físicas da madeira, são resultado direto da estrutura da parede celular, da orientação celular, do tipo de células presentes, sua distribuição, disposição e as relativas proporções nas quais estas estão presentes (Burger & Richter, 1991).

As informações dos parâmetros físicos, químicos e anatômicos da madeira influenciam diretamente a qualidade da madeira para a produção de papel e celulose (Mimms, 1993). Por esse motivo é fundamental o conhecimento de tais propriedades para obtenção de um produto final de qualidade.

O conteúdo de lignina, por exemplo, e sua composição em termos de unidades fenilpropanóides são parâmetros importantes para a produção da polpa de celulose, pois durante esse processo, o cozimento da madeira é interrompido sem que ocorra a remoção total da lignina, para evitar a degradação da celulose. Assim, essa lignina residual aumentará o uso de reagentes químicos usados durante o processo de branqueamento da polpa (Rodrigues *et al.*, 1999). A lignina é um polímero constituído de diferentes monômeros, principalmente aqueles derivados dos álcoois coniferílico (guaiacila – G), álcoois sinapílico (siringila – S) e *p*-cumarílico (*p*-hidroxifenila - H). Em *Eucalyptus*, a lignina é formada principalmente pelos monômeros S e G, contendo pouca quantidade de H. Esses monômeros diferem entre si, pelo

seu grau de metilação, quanto mais metilado, mais facilmente extraído (Barbosa *et al.*, 2008, Shi *et al.*, 2010). Assim, S é o monômero mais metilado e H o menos metilado.

Além da quantidade de cada componente químico, as propriedades físicas de cada parede dependem da combinação de polímeros nela existentes que são restritas, considerando o grande número de possibilidades (Buckeridge *et al.*, 2008).

Na parede celular, os polímeros estão ligados entre si por inúmeras ligações, covalentes e não-covalentes (Albersheim *et al.*, 2011). As moléculas de celulose adjacentes estão ligadas entre si por um grande número de pontes de hidrogênio (Buckeridge *et al.*, 2008). A lignina interage com os polissacarídeos da parede por vários tipos de ligação que são chamadas de “complexos lignina-carboidrato” (LCC). Os três polissacarídeos pécticos (homogalacturano, ramnogalacturano I e II) estão covalentemente ligados entre si por ligações do tipo alfa-1,4 através dos resíduos de ácido galacturônico (Albersheim *et al.*, 2011). Pectina e hemicelulose estão provavelmente ligadas por ligações com ácido diferúlicos e também transesterificação, que provavelmente ocorre entre os resíduos metilesterificados dos homogalacturonanos e outros polissacarídeos (Albersheim *et al.*, 2011). Entretanto, um grande número de interações entre as moléculas da parede ainda está por ser descoberto, uma vez que associações covalentes são encontradas entre vários polímeros extraídos da parede aos quais não se atribui nenhuma interação conhecida (Brett & Waldron, 1996).

Relacionar a expressão gênica com análises fenotípicas é de extrema importância para entender os mecanismos envolvidos na formação da parede celular e que são responsáveis por gerar as diferenças observadas entre as espécies. Novas metodologias surgem como alternativa para caracterizar a estrutura da parede celular. Dentre essas tecnologias podemos destacar o uso de anticorpos monoclonais que reconhecem diferentes tipos de polímeros da parede celular (Pattathil *et al.*, 2010).

Essa nova técnica de determinação o perfil glicômico de um material utiliza uma grande coleção de anticorpos monoclonais que reconhecem uma variedade de polissacarídeos da parede, desse modo, possibilitando o monitoramento da estrutura/estrutibilidade de seus componentes (DeMartini *et al.*, 2011). A técnica aplicada nesse trabalho permitiu observar mudanças significativas no perfil glicômico do xilema das três espécies analisadas. Este perfil está diretamente relacionado com o quão firmemente os polímeros estão ligados dentro da parede.

Assim, o objetivo do presente capítulo foi investigar, por diferentes metodologias, a composição, quantificação e estruturação das moléculas da parede celular do xilema das três

espécies estudadas, conforme fluxograma da figura 10. As análises permitiram a criação de modelo de estruturação de parede celular diferencial para as três espécies.

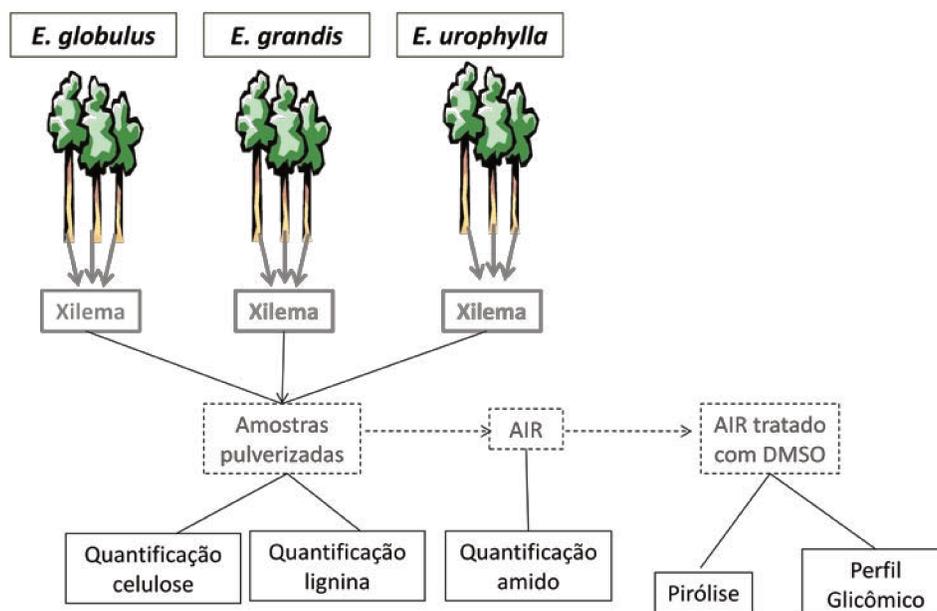


Figura 10: Fluxograma dos experimentos realizados.

6. Metodologia:

6.1 Material Vegetal



Figura 11: Cultivo de plântulas de eucalipto em casa de vegetação

Plantas de três espécies de *Eucalyptus* (*E. globulus*, *E. grandis* e *E. urophylla*) foram cultivadas em casa de vegetação no IB/Unicamp (Figura 11). Utilizou-se no plantio substrato

PlantMax HA + vermiculita e as plantas foram irrigadas 3 vezes por semana com água. Três plantas de cada uma das espécies, com 1,5 anos, foram utilizadas nas análises seguintes. Destas plantas coletou-se o caule após a retirada da casca. Esse material foi imediatamente congelado em nitrogênio líquido e depois liofilizado e pulverizado em moinho de bola (Tecnal). Amostras de folha de duas plantas de *E. grandis* também foram coletadas, congeladas, liofilizadas e pulverizadas. Estas foram utilizadas apenas nas análises do perfil glicômico, como mostrado adiante.

6.2 Quantificação de celulose total:

Para a quantificação do conteúdo total de celulose nas amostras de xilema, 50 mg do material pulverizado de cada amostra foi pesada e colocada em tubos cônicos de vidro, seguindo o protocolo descrito por Brendel *et al.* (2000), com modificações. Adicionou-se 2 mL de ácido acético 80% e 200 uL de ácido nítrico 69%. O tudo foi fechado e misturado cuidadosamente (com bolas de gude). As amostras foram autoclavadas por 25 minutos (121°C) e após esse tempo esperou-se que estas atingissem a temperatura ambiente.

Adicionou-se então, 2,5 mL de etanol 99% nas amostras, as quais foram homogenizadas no vortex, centrifugadas a 3500 rpm por 5 minutos e descartou-se o sobrenadante. Para a limpeza das amostras, esta sequência de passos foi realizada com os seguintes reagentes: 5 mL de etanol 99% (2,5 mL para misturar e depois mais 2,5 mL); 5 mL de água deionizada; 5 mL de NaOH 17% (p/v) (deixadas por 10 minutos na temperatura ambiente); 5 mL de água deionizada; 2,2 mL de água deionizada + 600 uL de ácido acético (misturadas em vortex) 2,2 mL de água deionizada; 5 mL de água deionizada.

O precipitado foi liofilizado e pesado para determinação da quantidade total de celulose na amostra.

6.3 Quantificação de lignina de Klason:

Para a quantificação do conteúdo de lignina de Klason na amostras de xilema, 10mg de cada amostra pulverizada foi pesada e colocada em tubos cônicos de vidro, seguindo o protocolo descrito por Theander *et al.* (1986). Adicionou-se 100 uL de ácido sulfúrico 72%, as amostras foram incubadas a 30°C por 45 minutos. O ácido foi então diluído para 4% com água destilada (1,7 mL). O tudo foi fechado e misturado cuidadosamente (com bolas de gude). As

amostras foram autoclavadas por 1 hora (121°C). O sobrenadante foi retirado e o hidrolisado foi lavado exaustivamente com água quente para remover o ácido.

O precipitado foi liofilizado e pesado para determinação do conteúdo de lignina Klason nas amostras de xilema.

6.4 Preparo dos Resíduos Insolúveis em Álcool (AIR)

O preparo dos AIRs visa retirar os açúcares solúveis presentes nos tecidos. Essa etapa, bem como a determinação do conteúdo de amido foram realizadas no Laboratório de Fisiologia e Bioquímica de Plantas - Lafieco (Departamento de Botânica – IB) USP, coordenado pelo Dr. Marcos Buckridge. O protocolo descrito a seguir e o descrito no Perfil Glicômico são baseados em Gorshokova *et al.* (1996), Carpita (1983) e Moraes *et al.* (2001).

Para a preparação dos resíduos insolúveis em álcool 1 conjunto de 500 mg de cada amostra de xilema das três espécies e folha de *E. grandis* (conjunto 1) e outro conjunto de 500 mg formado apenas por amostras de xilema (conjunto 2) foram pesadas. Em seguida, cada amostra foi ressuspensa em 30 mL de etanol 80% e incubadas a 80° C por 20 minutos. As amostras foram depois centrifugadas a 3000 g por 15 minutos e o sobrenadante descartado.

Esse passo foi repetido mais quatro vezes e a seguir, as amostras foram lavadas duas vezes com água destilada para retirar o excesso de etanol.

O conjunto 2 de amostras lavadas foi secado para posterior determinação do conteúdo de amido. Com o conjunto 1 seguiu-se os protocolos descritos a seguir:

O precipitado resultante da última lavagem foi ressuspensa em 30 mL de DMSO 90% e incubado à temperatura ambiente por 12 horas com agitação constante. Esse passo é realizado a fim de remover o amido das amostras. Após 12 horas, as amostras foram centrifugadas, o sobrenadante descartado e houve nova extração com DMSO 90% por mais 2 horas.

As amostras foram novamente centrifugadas e o sobrenadante descartado. A lavagem com água destilada foi feita 6 vezes para total remoção do DMSO.

Após todo esse processo, a parede celular resultante das amostras foi liofilizada por 24 horas e pesada com o intuito de mensurar a quantidade de resíduos retirados.

6.5 Determinação do conteúdo de amido

A extração de amido foi realizada por método enzimático (Amaral *et al.*, 2007) descrito a seguir: Ao material moído foram adicionados 0,5 mL (120 U mL^{-1}) de α -amilase, diluída em tampão MOPS 10 mM pH 6,5. As amostras foram incubadas a 75°C por 30 minutos. Este procedimento foi repetido, totalizando 120 unidades de enzima. As amostras foram resfriadas até 50°C , e então foi adicionado 0,5 mL de uma solução contendo 30 U mL^{-1} de amiloglicosidase em tampão acetato de sódio 100 mM pH 4,5. As amostras foram incubadas a 50°C por 30 minutos. Este procedimento foi repetido. Após estas incubações, foram acrescentados 100 mL de ácido perclórico 0,8 M para parar a reação enzimática e precipitar proteínas. Após uma rápida centrifugação (2 minutos a 9.300 g), procedeu-se a dosagem de amido nos extratos, através de quantificação da glicose liberada no processo de hidrólise do amido. Para tal foram retiradas alíquotas do extrato às quais foram adicionados 300 mL do reagente Glicose PAP Liquiform (CENTERLAB). Após incubação por 15 minutos a 37°C , o teor de glicose foi determinado em espectrofotômetro acoplado a leitor de ELISA em comprimento de onda de 490 nm. Para a confecção da curva padrão foi utilizada solução de glicose (SIGMA), nas concentrações de 0; 1; 2,5; 5,0; 7,5; 10 e $12,5 \text{ mg/mL}$.

6.6 Pirólise

A pirólise analítica foi realizada em colaboração com o Instituto de Investigação Científica Tropical – IICT- Portugal, sob orientação do Dr. José Carlos Rodrigues. Para tal utilizou-se um Pyroprobe CBS 1000 com uma sonda de filamento ligado a uma bobina de GC (Agilent 6890), com detector de ionização por de uma interface aquecida (270°C). A pirólise foi realizada a 600°C para 5s, usando 75-77 μg das amostras. Coluna capilar: DB1701 (60 m x 0,25 mm, 0,25 μm filme, J & W Scientific). Condições GC: injetor 270°C , C detector 270°C , temperatura, programa de 45°C , 4 minutos isotérmica, em seguida, taxa de aquecimento 4- 250°C minutos, 6- 270°C para manter 8 minutos A identificação de produtos de pirólise foi realizada usando amostras seleccionadas anteriormente por Py-GC/MS (CDS Pyroprobe 1000 ligado a um HP 6890 com um detector de massa HP 5973 selectivos). Os produtos foram identificados pelo seu espectro de massa e tempo de retenção.

6.7 Perfil Glicômico da parede celular de eucalipto

O perfil glicômico das amostras de xilema e folha foi realizado no Complex Carbohydrate Research Center – University of Georgia (Athens – GA - USA), sob a supervisão

do Dr. Michael G. Hahn. Para obtenção do perfil glicômico realizou-se o fracionamento de resíduos de parede celular, a estimativa do açúcar total e os ensaios de ELISA.

Fracionamento dos Resíduos de Parede Celular Insolúveis em Álcool

Este método envolve extração sequencial para obter frações enriquecidas com diversos polisacarídeos da parede, incluindo pectina, arabinogalactanos, xilanos, xiloglicanos. Tais etapas foram feitas da seguinte maneira:

Os resíduos insolúveis em álcool (AIR) tratados com DMSO foram ressuspensos em 20 mL de oxalato de amônio 50 mM (pH 5.0) e incubados por 24 horas sob agitação constante. As amostras foram depois centrifugadas a 4000 g por 15 minutos e o sobrenadante transferido para novo tubo marcado como "fração oxalato de amônio" e armazenado a 4°C.

O precipitado resultante foi lavado com o mesmo volume de água destilada e depois ressuspendido em 20 mL de carbonato de sódio 50 mM, contendo 0,5% de borohidreto de sódio e incubado por 24 horas sob agitação constante. Após incubação, um novo precipitado foi obtido a partir da centrifugação a 4000 g por 15 minutos. O sobrenadante foi armazenado a 4°C em um novo tubo marcado como "fração de carbonato".

O precipitado foi ressuspendido em 20 mL de hidróxido de potássio a 1M contendo 0,5% de borohidreto de sódio e depois incubado por 24 horas sob agitação constante. As amostras foram novamente centrifugadas (4000 g, 15 minutos) e o sobrenadante armazenado a 4°C em um novo tubo marcado "fração 1 M KOH". Antes de armazenar, os sobrenadantes foram neutralizados como descrito no final desta seção.

O precipitado resultante foi diretamente ressuspendido em 20 mL de hidróxido de potássio a 4M contendo 0,5% de borohidreto de sódio e depois incubado por 24 horas sob agitação constante. Após esta etapa, o precipitado foi obtido por centrifugação (4000 g, 15 minutos). O sobrenadante foi armazenado a 4°C em um novo tubo marcado "fração 4 M KOH". Antes de armazenar, os sobrenadantes também foram neutralizados como descrito no final desta seção.

O precipitado foi lavado 3 vezes com água destilada a fim de remover todo KOH antes da próxima extração e depois ressuspendido em água destilada (20 mL).

Com o objetivo de quebrar os polímeros de lignina, as amostras foram tratadas com clorito de sódio e ácido acético glacial da seguinte maneira:

As amostras ressuspendidas em água foram colocadas em banho-maria a 70°C sob agitação. À elas foi adicionado 0.125 g de clorito de sódio e 50 µL de ácido acético glacial. Esse passo foi feito 3 vezes, sendo cada um com 1 hora de incubação a 70°C. O produto final contém gás clorino, que foi removido por borbulhamento de ar nas amostras até que elas passassem de uma coloração amarela pra branco-translúcido (este passo foi feito na capela).

As amostras foram centrifugadas (4000 g por 15 minutos) e o sobrenadante armazenado a 4°C em um novo tubo marcado “fração clorito”.

Finalmente, o pellet resultante foi lavado com água destilada e depois ressuspendido em 20 mL de KOH 4 M contendo 1% de borohidreto de sódio e incubado por 24 horas sob agitação constante. Depois, as amostras foram centrifugadas (4000 g por 15 minutos) e o sobrenadante transferido para novo tudo marcado “fração KOHPC 4 M”. Antes de armazenar a 4°C, os sobranadantes também foram neutralizados como descrito no final desta seção.

O precipitado residual foi lavado em água destilada e armazenado a 4°C.

- Neutralização das frações de KOH 1 M, KOH 4 M e KOHPC 4 M:

Os tubos foram colocados no gelo (verticalmente) e vedados com “parafilm” (com pequenos furos). Adicionou-se de 3 a 5 gotas de 2-Octanol e depois, vagorosamente, ácido acético glacial até que o pH7 fosse atingido.

Figura 12: Anticorpos monoclonais e respectivas epítopes correspondentes à polissacarídeos presentes na parede celular.



Todas as frações de parede celular foram dializadas contra 4 trocas de água destilada (amostra:água ~ 1:60) à temperatura ambiente sob leve agitação constante por um total de 60 horas. Cada troca de água destilada obedeceu a um intervalo mínimo de 6 horas. As amostras foram depois transferidas para novos tubos de 50 mL e depois liofilizadas.

Estimativa do açúcar total

Todos os extratos foram pesados. De cada extrato, 2 mg foram dissolvidas em 10 mL de água destilada (volume final) em um tubo de 15 mL. O cálculo do açúcar total nas amostras é feito por ensaio colorimétrico com fenol-ácido sulfúrico em microplaca, adaptado de Dubois *et al.* (1956) e Masuko *et al.* (2005).

A estimativa foi realizada em duplicata para cada amostra em tubos de ensaio. 100 µL da amostra diluída foi adicionado a cada tubo. A eles adicionou-se 100µL de fenol a 5% e 500 µL de ácido sulfúrico (18 M). As amostras foram levemente homogenizadas em vortex e incubadas à temperatura ambiente por 20 minutos.

Após esse período, 250 µL de cada amostra foi transferida para uma placa de ELISA e as amostras foram lidas (ELISA *plate reader*) em comprimento de onda 490nm.

ELISA – Enzyme-linked Immunosorbent Assay

Após estimativa do teor açúcar total, as amostras foram diluídas a uma concentração de açúcar de 60 µg mL⁻¹. Para cada amostra, 50 µL da diluição foram distribuídos em placas de ELISA e estas incubadas *overnight* a 37°C para secagem. Para controle também aplicou-se água deionizada em poços da placa.

O ensaio foi realizado com 156 diferentes anticorpos monoclonais específicos para a maioria dos polissacarídeos presentes da parede celular (Figura 12), seguindo o protocolo descrito por Pattathil *et al.* (2010), como descrito a seguir:

As placas foram bloqueadas com 200 mL de leite em pó desnatado a 1% (p/v) em TBS (50 mM, Tris-HCl, p H 7,6, contendo 100 mM de cloreto de sódio) por 1 h. Os passos de aspiração e lavagem foram realizados em ELx405 *microplate washer* (Bio-Tek Instruments). O agente bloqueador foi removido por aspiração e 50 mL dos 156 anticorpos monoclonais foram distribuídos nas placas e incubados por 1h à temperatura ambiente. O sobrenadante foi então removido e os poços foram lavados 3 vezes com 300 mL de leite em pó desnatado a 1% (p/v) em TBS (tampão de lavagem). Adicionou-se então 50 mL o anticorpo secundário (anti-rato

ou anti-camundongo dependendo do anticorpo primário) diluído 1: 5000 em tampão de lavagem e incubou-se por 1h. Em seguida, as placas foram lavadas 5 vezes com 300 mL de tampão de lavagem. Uma solução de 3,3', 5,5'-*tetramethylbenzidine* (Vector Laboratories) foi preparada de acordo com instruções do fabricante, e 50 mL foi adicionado em cada poço. Após 20 minutos, a reação foi terminada com a adição de 50 mL de 0,5 N de ácido sulfúrico. Em seguida, as cores de cada amostra foram lidas como a diferença em A450 and A 655 usando um leitor de microplaca (modelo 680, Bio-Rad). Subtraiu-se da leitura de cada amostra a leitura realizada para os poços contendo água.

7. RESULTADOS E DISCUSSÃO

7.1 Quantificação de celulose total e lignina Klason.

Na parede celular, as moléculas de celulose adjacentes estão ligadas entre si por um grande número de pontes de hidrogênio, as quais competem com a água de hidratação das hidroxilas e, por isto, a microfibrila é uma estrutura cristalina, compacta e desidratada. Esta desidratação dificulta o ataque de enzimas, pois para a quebra de uma ligação glicosídica (hidrólise) é necessário a presença de água. As ligações β exigem alta energia para quebra. Por este motivo, há uma severa limitação ao acesso à celulose tanto enzimaticamente como por hidratação (Buckeridge *et al.*, 2008).

Diversos métodos são aplicados para estudo dos componentes da parede celular. A hidrólise ácida é comumente utilizada para separação dos polissacarídeos da parede. Tais polissacarídeos da parede diferem entre si pela sua suscetibilidade à hidrólise ácida. Além disso, alguns monossacarídeos, como a xilose, são mais suscetíveis à hidrólise que outros (Albersheim *et al.*, 2011). O método aqui utilizado para quantificação de celulose envolve um único passo de aplicação de ácidos para promoverem tanto a deslignificação e remoção de polissacarídeos não celulósicos, evitando a perda de celulose (Brendel *et al.*, 2000). O conteúdo estimado de celulose é em geral 40 a 50% aproximadamente da massa seca do caule (Pettersen, 1984). Para *E. globulus* de 1,5 anos Rencoret e colaboradores (2011) encontraram aproximadamente 40%. Entretanto os resultados aqui encontrados para árvores também de 1,5 são de em torno de 14% de celulose (Figura 13, A), mostrando que provavelmente houve perda de celulose durante o processo. Além disso, não houve diferenças significativas entre as médias do conteúdo de celulose para as três espécies, porém observou-se uma variação intraespecífica considerável.

Com base nos resultados encontrados para o conteúdo de celulose nas amostras de xilema das três espécies, podemos concluir que o método utilizado não foi preciso para determinar tais medidas, uma vez que o conteúdo observado não deve estar representando o conteúdo real.

A lignina total de uma amostra é definida como a soma de lignina solúvel em ácido e a lignina insolúvel em ácido (definida como Lignina Klason) (Raiskila *et al.*, 2007). Diversos métodos para isolamento de lignina geralmente refletem um valor superestimado devido a diversos compostos “contaminantes”, como por exemplo, os carboidratos ligados à lignina (LCC) (Guerra *et al.*, 2006; Balakshin *et al.*, 2008). Não foi observada diferença significativa para a média do conteúdo de lignina Klason extraída para cada espécie (Figura 13, B). Entretanto, uma variação interespecífica grande foi observada. Além disso, as quantidades de lignina extraídas aqui foram bem inferiores aos conteúdos total de lignina geralmente encontrados em madeiras de *Eucalyptus* (geralmente de 20 a 30%) (Ferraz *et al.*, 2000; Fardiml & Durán, 2004; Rencoret *et al.*, 2011). Em Rencoret *et al.* (2011) árvores de *E. globulus* também de 1,5 anos foram avaliadas e observou-se um conteúdo total de lignina de aproximadamente 20% da massa seca do caule, sendo aproximadamente 17% de lignina Klason. Entretanto os resultados aqui encontrados para árvores também de 1,5 variam de 16% a 40%, aproximadamente. Essa grande variação mostra que o método foi impreciso e que para algumas amostras provavelmente deve haver uma contaminação com outros carboidratos.

Desse modo, acredita-se que para a confiabilidade do resultado seja necessária a repetição do experimento com maior número de amostras e repetições entre as mesmas.

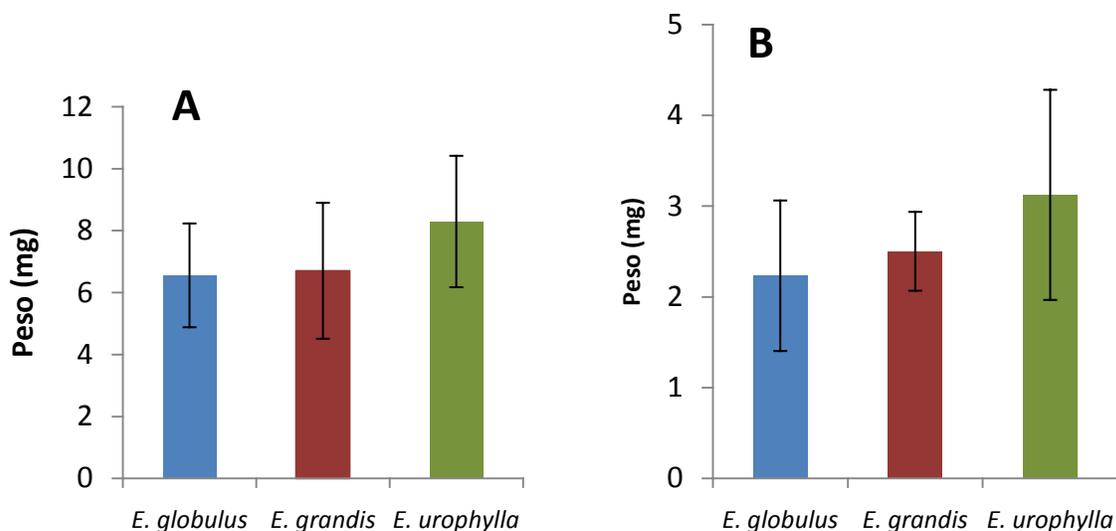


Figura 13: Gráfico “box plot” das médias das medidas de peso em mg (Y), com desvio padrão das 3 medidas para cada espécie; A- celulose total; B – lignina Klason.

7.2 Preparo dos Resíduos Insolúveis em Álcool (AIR)

Após a retirada dos resíduos solúveis em Álcool e amido, as amostras foram liofilizadas e pesadas novamente (Tabela 2). Esse primeiro passo de preparo das amostras é realizado com o intuito de isolar os resíduos de parede celular, como é possível observar na tabela 5, mais de 50% do material de folha era solúvel em álcool, mostrando que, como esperado, a quantidade de “parede celular” presente nas folhas é bem menor que em xilema. O qual, ao contrário, possui a maior parte de suas células contendo parede celular secundária. Também não houve diferença na porcentagem de resíduos insolúveis em álcool nas amostras de xilema.

Tabela 2: Porcentagem e desvio padrão dos resíduos insolúveis em álcool tratados com DMSO (Fgr – amostras de folha de *E. grandis*; E. gl - amostras de xilema de *E. globulus*; E. gr - amostras de xilema de *E. grandis*; E. ur - amostras de xilema de *E. urophylla*).

| Amostras | % Media de AIR tratados com DMSO | Desvio Padrão |
|--------------|----------------------------------|---------------|
| <i>Fgr</i> | 40.13 | 0.16 |
| <i>E. gl</i> | 86.90 | 2.84 |
| <i>E. gr</i> | 88.27 | 0.97 |
| <i>E.ur</i> | 88.12 | 1.68 |

7.3 Determinação do conteúdo de amido

O conteúdo total de amido nas amostras variou significativamente na espécie *E. globulus* em comparação com as outras espécies (Figura 14).

Pode-se especular aqui que, como *E. globulus* é a espécie com crescimento mais lento, a glicose produzida durante a fotossíntese seja, principalmente, convertida em celulose e hemicelulose na parede celular, pois já observamos no Capítulo 1, genes responsáveis por essa via estão mais expressos nessa espécie.

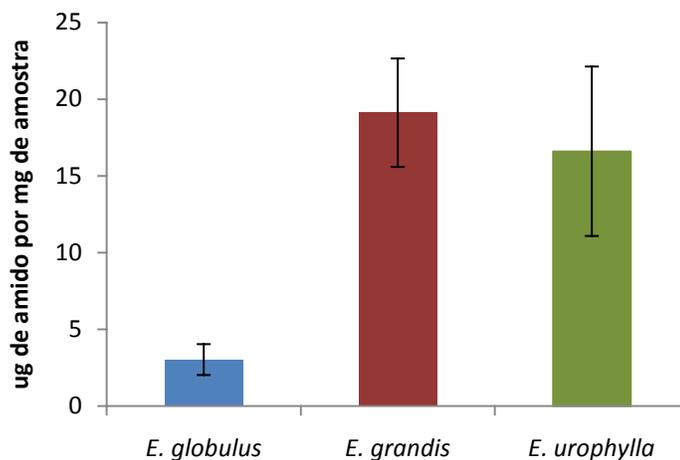


Figura 14: Gráfico “*box plot*” das medidas das quantificações µg/mg de amostra total para as três espécies. Para cada espécie, 3 amostras foram mensuradas (com três réplicas para cada).

7.4 Pirólise

Os produtos que resultam de pirólise foram identificados por espectrometria de massa e as razões aqui obtidas foram com base nas áreas dos picos correspondentes aos diferentes compostos. A lignina é um polímero constituído de diferentes monômeros, principalmente aqueles derivados dos álcoois coniferílico (guaiacila – G), álcoois sinapílico (siringila – S) e *p*-cumarílico (*p*-hidroxifenila - H). Em *Eucalyptus*, a lignina é formada principalmente pelos monômeros S e G, contendo pouca quantidade de H. Esses monômeros diferem entre si, pelo seu grau de metilação, quanto mais metilado, mais facilmente extraído (Barbosa *et al.*, 2008, Shi *et al.*, 2010). Assim, S é o monômero mais metilado e H o menos metilado. Na figura 15 (A) podemos observar uma diferença significativa apenas na razão S/G de *E. globulus* em comparação com as outras espécies. A maior quantidade de monômeros S indica que a lignina desse material é mais facilmente extraída em procesos de produção de polpa de celulose. Um

material no qual a lignina é mais facilmente extraída gera vantagens para as indústrias tanto em relação à maior produção, quanto em relação ao custo de produção. Também foi determinada a razão H/G para as amostras (Figura 15 B) e não foi observada uma diferença significativa.

A relação de pentoses e hexoses presentes nas amostras também foi medida (Figura 15 C). Os açúcares da parede celular são principalmente formados por monossacarídeos do tipo pentoses (ex. xilose) e hexoses (ex. glicose) (Buckeridge *et al.*, 2008). Para as amostras avaliadas não foram observadas diferenças significativas desses monossacarídeos.

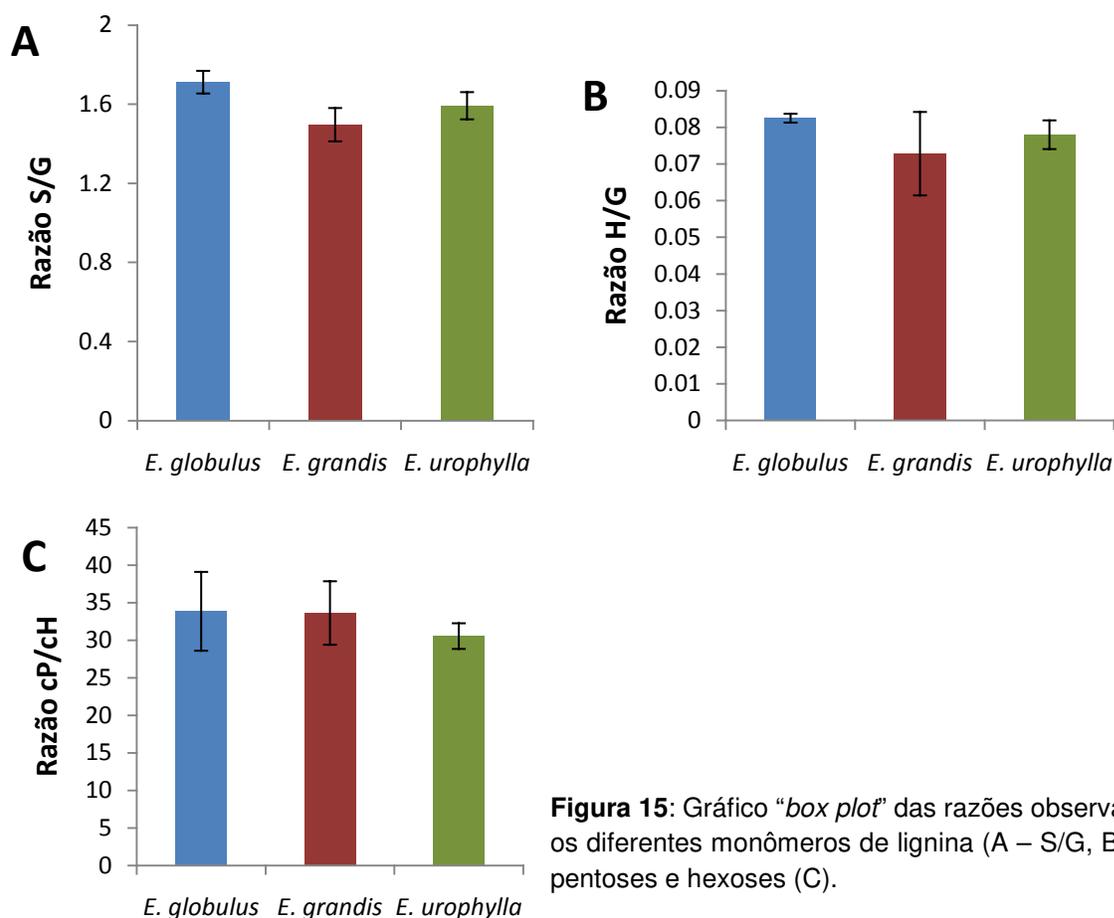


Figura 15: Gráfico “box plot” das razões observadas para os diferentes monômeros de lignina (A – S/G, B – H/G) e pentoses e hexoses (C).

7.5 Perfil Glicômico da parede celular de eucalipto

Os resíduos insolúveis em álcool foram submetidos a uma série sucessiva de extrações, que têm como objetivo extrair os polissacarídeos presentes na parede celular. Em grande parte

das plantas já submetidas a essa série de extrações, a pectina é o polissacarídeo principal presente na “fração oxalato de amônio”, uma vez que esse polissacarídeo está frouxamente associado com a celulose e hemicelulose. Como é possível observar na figura 16, para esta fração obteve-se uma maior quantidade de material em folhas (A) que em xilema (B). Tal fato é condizente com o atual conhecimento da presença deste polissacarídeo nas paredes celulares, ele é o principal constituinte da parede primária, além de ser quebrado pela ação das poligalacturonases para acomodação de uma parede celular secundária. Também pode ser observado que as frações clorito e 4M KOH PC, possuem mais material em xilema. Após a quebra da lignina, as hemicelulose e celuloses estão mais facilmente expostas.

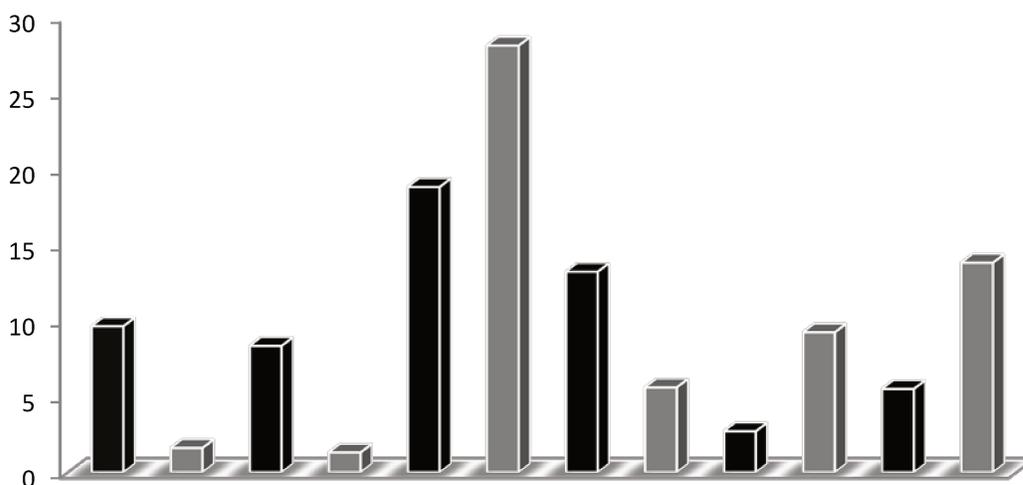


Figura 16: Peso em mg (eixo y) dos extratos obtidos por de *Glycome Profiling* a partir de AIR de folha de *E. grandis* e xilema de *E. urophylla*. A distribuição das frações e tecidos seguem a mesma ordem da figura 33. Colunas pretas: folha, colunas cinza: xilema. As frações seguem a seguinte ordem: fração oxalato de amônio,carbonato, 1M KOH, 4M KOH, Clorito e 4M KOH PC.

Após a estimativa do açúcar total em cada um dos extratos, as amostras foram submetidas à ELISA utilizando um conjunto de 156 anticorpos monoclonais que reconhecem diferentes carboidratos presente na parede celular.

As medidas de absorvância foram feitas para todas as 66 amostras contra 156 anticorpos, totalizando 10.296 medidas. Entretanto, para a construção dos *heat maps*, obteve-se a média para cada uma das amostras.

Na figura 17 é possível observar que o padrão de marcação dos 156 anticorpos para as amostras de folha é claramente diferente do padrão observado em xilema para as primeiras quatro frações. Tal resultado já era esperado uma vez que a parede celular das plantas possui

diferente composição para cada tecido. Além disso, a presença de uma parede celular secundária em células de xilema contribui fortemente para a diferença observada. Nas últimas 2 frações, o padrão de marcação para xilema *E. urophylla* se diferencia do padrão de xilema para as outras espécies, indicando uma aquitetura celular diferente, detalhada a seguir.

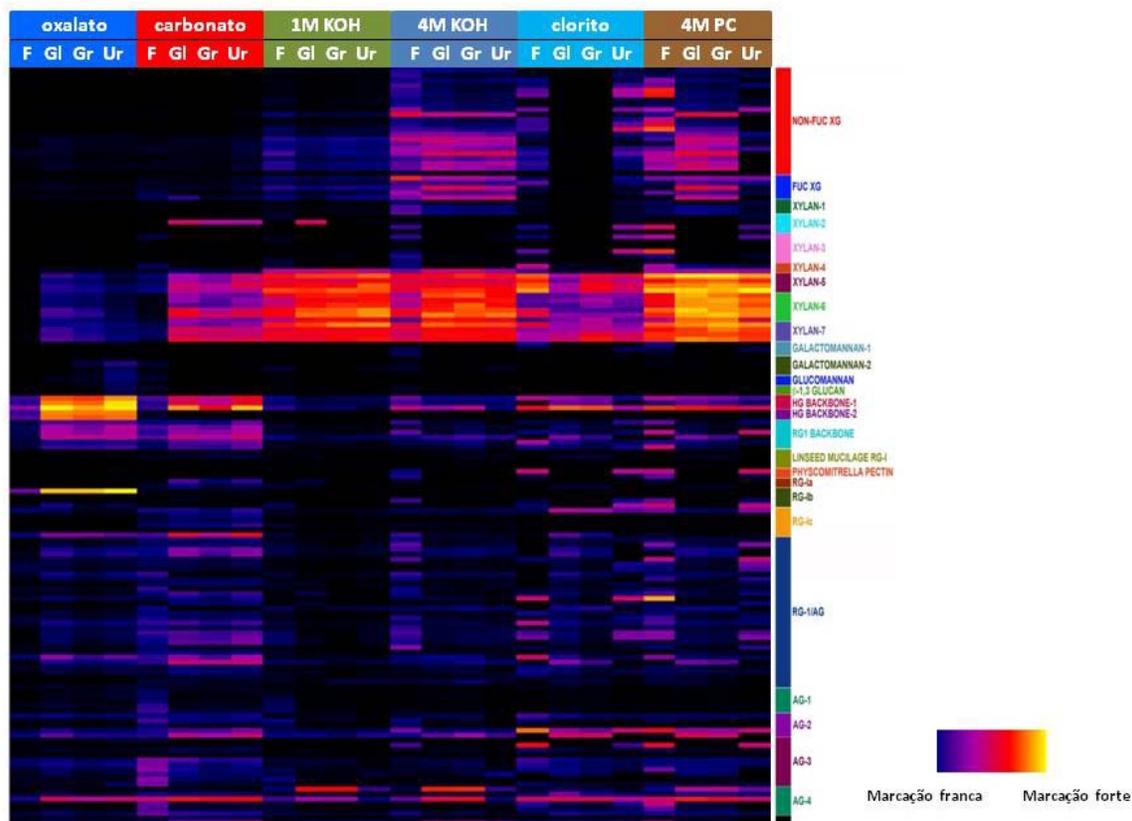


Figura 17: Heat map das frações extraídas de xilema de três espécies de *Eucalyptus* e folha de *E. grandis*. Visualização do perfil geral de todos os tratamentos realizados.

Os polímeros da parede estão ligados entre si por inúmeras ligações, covalentes e não-covalentes. Individualmente, a maior parte dessas moléculas individuais é solúvel em solução aquosa, entretanto a parede em si é insolúvel, pois as moléculas interagem entre si e essas ligações são a base da insolubilidade (Albersheim *et al.*, 2011).

Como esperado, o principal polissacarídeo extraído na fração “oxalato de amônio” foi a pectina (Figura 17 e Figura 18, mais detalhada) sendo a marcação mais forte em epitopos de homogalacturonano. Os homogalacturonanos não esterificados na presença de cátions divalentes (Ca^{2+}) podem formar interações entre os grupos carboxil adjacentes (Figura 19) e esse tipo está principalmente presente na lamela média. Dentro da parede celular, a maior parte dos homogalacturonanos é esterificada (Albersheim *et al.*, 2011).

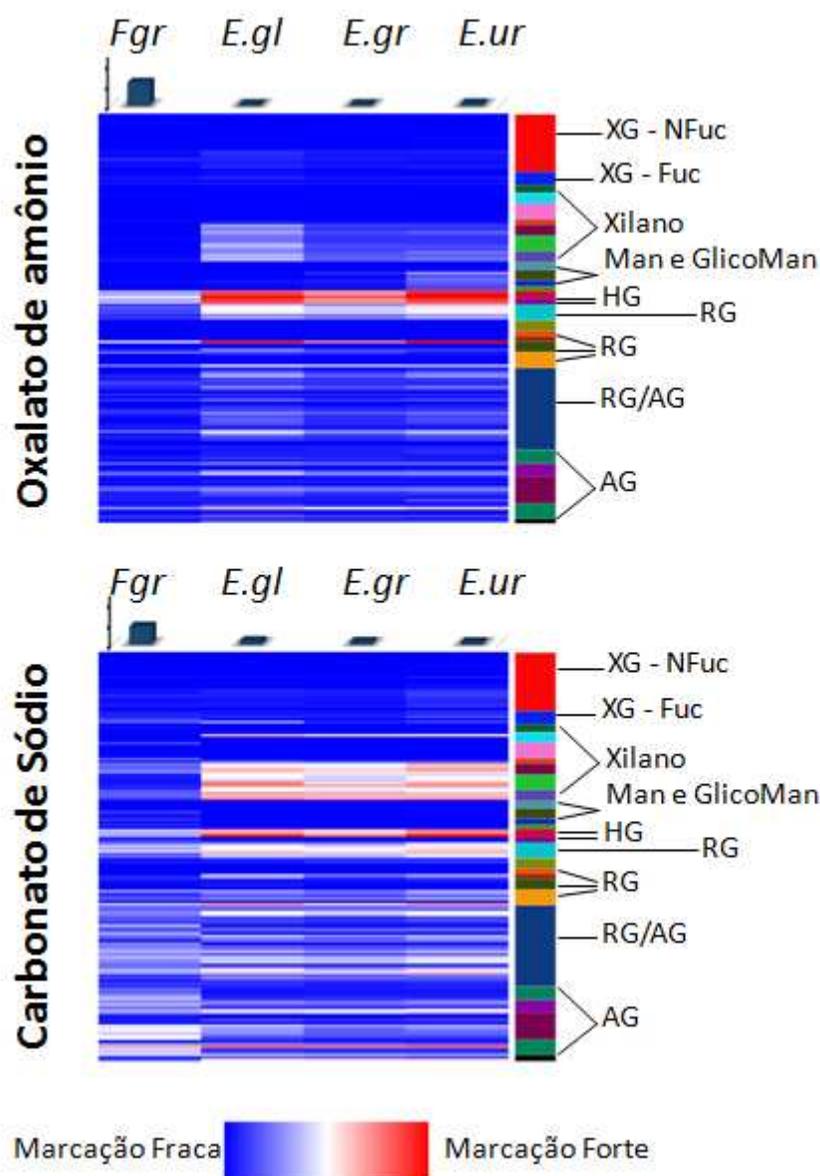


Figura 18: *Heat maps* das frações extraídas de xilema de três espécies de *Eucalyptus* e folha de *E. grandis*. Média das amostras para as frações “Oxalato de amônio” e “Carbonato”. A coluna colorida da direita mostra os anticorpos utilizados (de acordo com a figura 12 e os principais carboitratos reconhecidos identificados por XG-NFuc: xiloglucano não fucosilado; XG-FUC: xiloglucano fucosilado; Xilano; Man – manano; Glicoman-glicomanano; HG – cadeia principal de homogalacturano; RG – cadeia principal de ramnogalacturano I; AG: arabinogalactano) . O gráfico de barras sobre cada *heat map* representa a quantidade em mg de cada fração extraída por g de tecido.

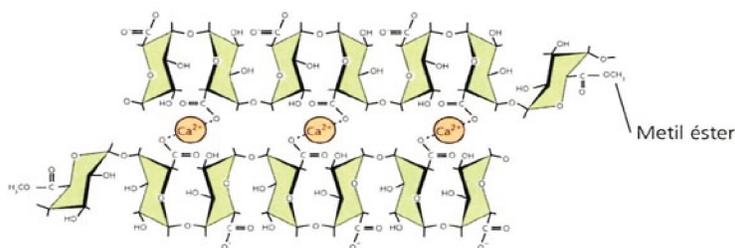


Figura 19: Estrutura dos homogalacturonanos evidenciando interação entre as cadeiras através do íon cálcio.

O oxalato de amônio é um sal que pode formar precipitados com os íons dos metais alcalinos terrosos (ex: Ca^{2+}). Assim, quando este sal é adicionado nas amostras, as pontes de cálcio dos homogalacturonanos são quebradas, desestabilizando esses polímeros, que ficam mais solúveis e então diluídos no sobrenadante que deu origem a essa fração. Assim, na figura 18, observa-se principalmente a pectina liberada que estava presente na lamela média.

Sugere-se que os três polissacarídeos pécticos estão covalentemente ligados entre si por ligações do tipo alfa-1,4 através dos resíduos de ácido galacturônico (Figura 20). Desse modo, os RG ligados aos HG serão diluídos na fração oxalato de amônio. Por esse motivo, a marcação de RGI observada nesse extrato, se deve provavelmente ao RG ligado ao homogalacturonano que se tornou solúvel após ruptura das pontes de hidrogênio.

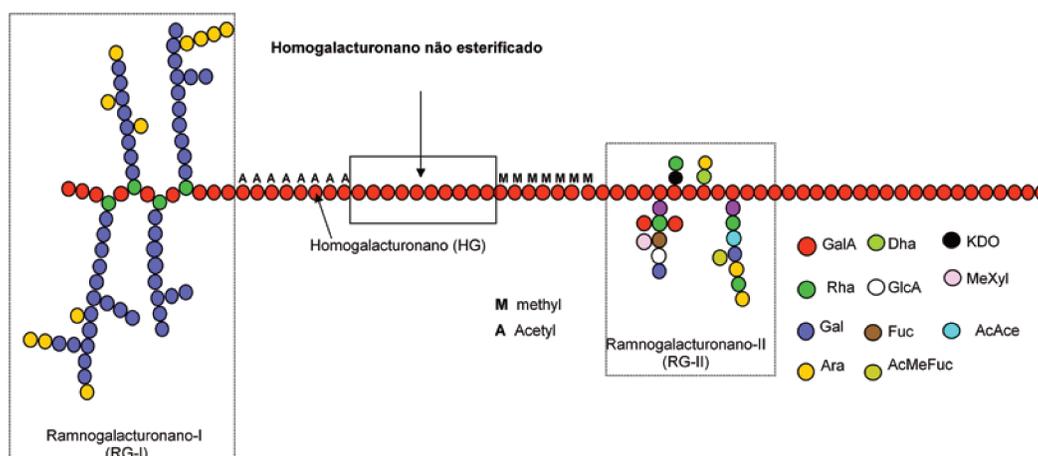


Figura 20: Interações entre os polissacarídeos pécticos.

Também pode-se observar a marcação, apesar de mais fraca, de epitopos de xilano nessa fração (Figura 17 e Figura 18, mais detalhado). Os xilanos são as hemiceluloses principais na parede celular secundária enquanto que as pectinas estão principalmente

presentes na lamela média e parede primária. Entretanto, existe pectina presente na parede secundária, provavelmente remanescente daquela presente na parede primária.

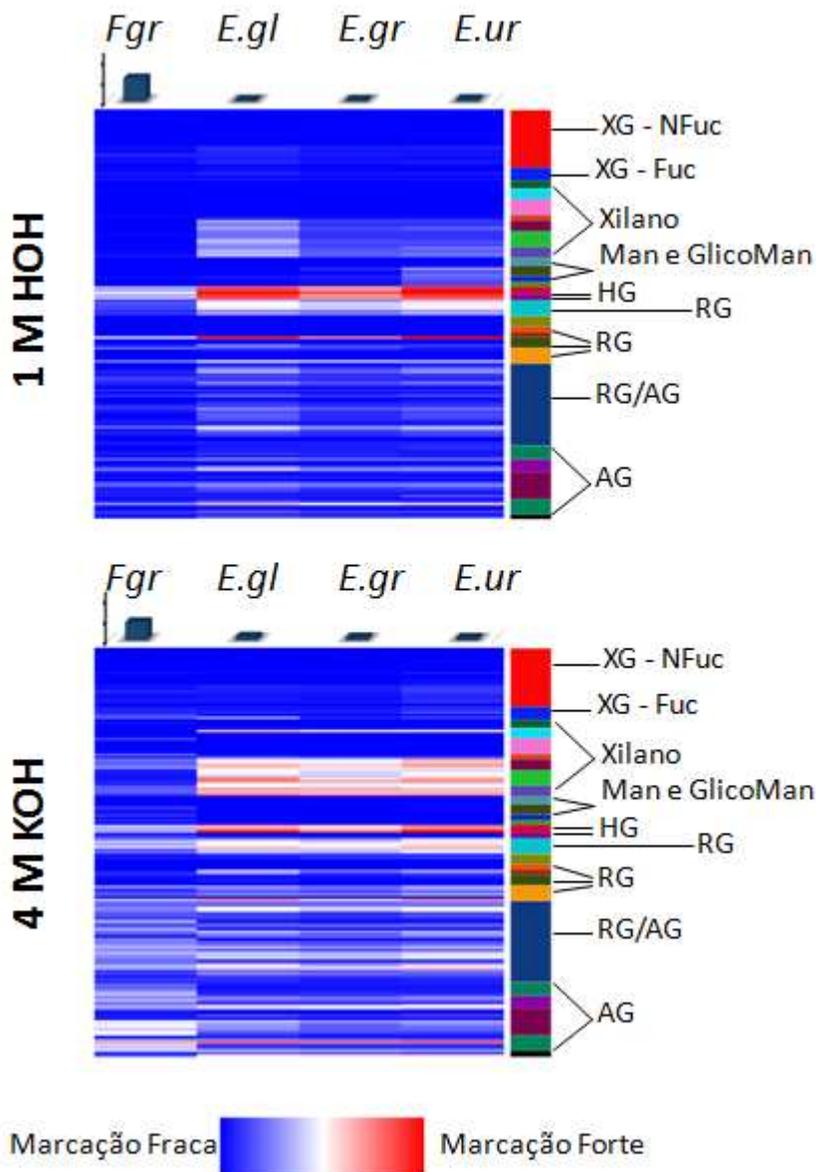


Figura 21: *Heat maps* das frações extraídas de xilema de três espécies de *Eucalyptus* e folha de *E. grandis*. Média das amostras para as frações “1 M KOH” e “4 M KOH”. A coluna colorida da direita mostra os anticorpos utilizados (de acordo com a figura 12 e os principais carboidratos reconhecidos identificados por XG-NFuc: xiloglucano não fucosilado; XG-FUC: xiloglucano fucosilado; Xilano; Man – manano; Glicoman-glicomanano; HG – cadeia principal de homogalacturano; RG – cadeia principal de ramnogalacturano I; AG: arabinogalactano). O gráfico de barras sobre cada *heat map* representa a quantidade em mg de cada fração extraída por g de tecido.

Um grande número de interações entre as moléculas da parede ainda está por ser descoberto, uma vez que associações covalentes são encontradas entre vários polímeros extraídos da parede aos quais não se atribui nenhuma interação conhecida. Estes incluem proteína-pectina, pectina-xilano, xilano, e outras associações mais complicadas como xilano-pectina-proteína e etc. (Brett & Waldron, 1996).

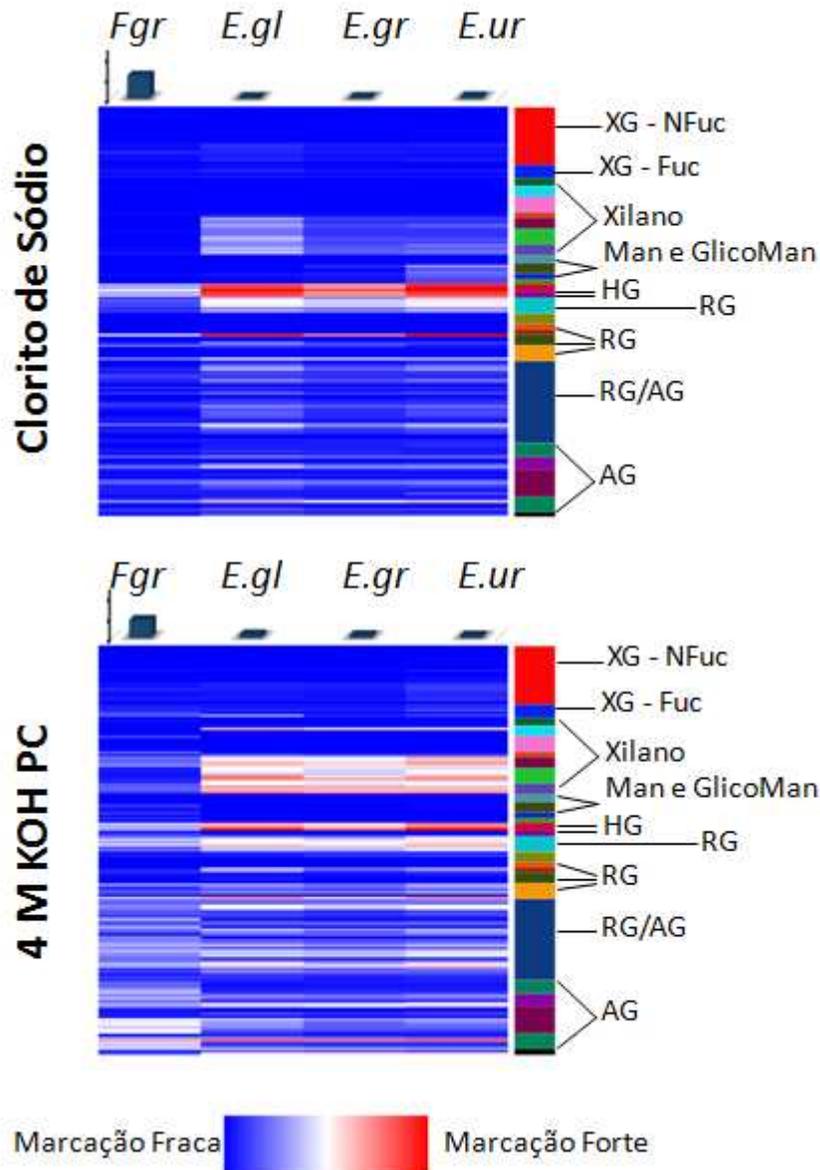


Figura 22: Heat maps das frações extraídas de xilema de três espécies de *Eucalyptus* e folha de *E. grandis*. Média das amostras para as frações “Clorito de Sódio” e “4 M KOH PC”. A coluna colorida da direita mostra os anticorpos utilizados (de acordo com a figura 12 e os principais carboítratos reconhecidos identificados por XG-NFuc: xiloglucano não fucosilado; XG-FUC: xiloglucano fucosilado; Xilano; Man – manano; Glicoman-glicomanano; HG – cadeia principal de homogalacturano; RG – cadeia

principal de ramnogalacturano I; AG: arabinogalactano) . O gráfico de barras sobre cada *heat map* representa a quantidade em mg de cada fração extraída por g de tecido.

De fato, já foram observados dois tipos de interações que podem explicar a ligação pectina-hemiceluloses. As ligações com ácido diferúlicos (como aqueles encontrados entre arabinoxilanos – principal hemicelulose de monocotiledôneas) e também a transesterificação, que pode provavelmente ocorrer entre os resíduos metilesterificados dos homogalacturonano e outros polissacarídeos (Albersheim *et al.*, 2011).

Assim, acredita-se que o xilano liberado nesse extrato deve provavelmente estar ligado com a pectina remanescente na parede secundária. Na fração “Carbonato de Sódio” ainda observa-se marcação de epitopos de pectina e marcação mais forte dos xilanos e uma marcação, ainda que bem fraca, de xiloglicanos. A solução de carbonato de sódio alcalina tem a propriedade de extrair pectina e também algumas hemiceluloses em pequenas quantidades (Carpita, 1984).

Polissacarídeos pécticos com alto teor de ácido galacturónico (homogalacturonano) são geralmente encontrados na primeira fração. Enquanto polissacarídeos pécticos com elevados níveis de açúcar neutro são extraídos pelo carbonato de sódio e soluções de KOH. As hemiceluloses são encontradas em todos os extractos alcalinos, incluindo o carbonato de sódio. O borohidreto de sódio (NaBH₄) é adicionado às extracções alcalinas para prevenir a degradação dos polissacarídeos. Pois, em meio alcalino, o açúcar nas extremidades reduzidas se rearranja até que é clivado, expondo assim outra extremidade reduzida.

Xiloglucanos perfazem 30% da parede primária de dicotiledôneas, se aderem fortemente à celulose por pontes de hidrogênio e impedem que as microfibrilas colapsem. Tratamentos com 1M e 4M KOH, resultam na remoção das interações entre hemicelulose e celulose, colapsando as microfibrilas. Após tratamento com KOH 4M, uma quantidade significativa de xiloglicano é solubilizada (Melton & Smith, 2001). A marcação para xiloglucano torna-se mais forte a partir da extração com KOH 1M (Figura 17 e Figura 21, mais detalhada).

A lignina é depositada principalmente nas paredes celulares secundárias e também pode ser encontrada na lamela média. Esta não pode ser extraída por solventes aquosos, por isso a aplicação de clorito de sódio é feita para oxidar o composto e gerar moléculas de baixo peso molecular e solúveis em água. O clorito de sódio (NaClO₂) é uma solução alcalina e não

reativa, entretanto quando acidificado (adição ácido acético) forma um gás reativo (Sarkanen, 1962).

A lignina interage com os polissacarídeos da parede por vários tipos de ligação que são chamadas de “complexos lignina-carboidrato” (*LCC*). Dentre estas estão as ligações benzil-éster (ligações entre ácidos urônicos dos polissacarídeos e grupos hidroxila da lignina); e os diferúlicos (ligações dos ácidos ferúlicos esterificados ou ligados através de ligações éter aos polissacarídeos) (Albersheim *et al.*, 2011).

Existe uma forte evidência que há uma maior afinidade de formação de ligações covalentes da lignina com a hemicelulose, do que com celulose. A lignina se liga aos resíduos galactosil no xiloglucano e arabinosil na xilose. Podendo, desse modo, estarem ligados aos arabinogalactanos. A ligação da lignina com a pectina também pode ocorrer (Prof. Marcos Buckeridge, comunicação pessoal).

Algumas dessas ligações são mais estáveis em soluções alcalinas, o que pode contribuir para o lento processo de deslignificação da madeira em processos alcalinos de separação de polpa de celulose (Albersheim *et al.*, 2011).

A maior diferença entre os extratos foi para a fração “Clorito” (Figura 17 e Figura 23, mais detalhada). Nesta fração, existe uma quantidade considerável de xilano (porção superior) e xiloglucanos presente nas amostras de folha e *E. urophylla*, enquanto não há marcação para os mesmos nas amostras de *E. globulus* e *E. grandis*. A adição de clorito de sódio e ácido acético glacial às amostras tem o objetivo de quebrar os polímeros de lignina e seus monômeros são lavados. Assim, com a retirada da lignina, a qual está fortemente ligada às hemiceluloses, tais polissacarídeos ficam expostos e são assim mais facilmente extraídos. Além disso, existe uma grande diferença nos xilanos da porção inferior (Figura 22) para as diferentes espécies e tecidos observados. Uma grande quantidade de pectinas também foi marcada nas frações seguintes ao tratamento com clorito de sódio. Observa-se que a marcação de epítomos de xilano em *E. globulus* é menor, isso pode indicar que durante os processos de deslignificação a perda de material associado ao processo nessa espécie seja menor, aumentando, assim, seu rendimento.

Em média, 10% do peso seco da parede celular se deve à proteínas glicoproteínas e proteoglicanos, que podem ser divididos em 2 grupos principais: 1) as proteínas estruturais insolúveis (glicoproteínas ricas em hidroxiprolina, proteínas ricas em glicina e prolina) e as proteínas solúveis (enzimas, proteínas de transporte, proteínas arabinogalactanos e outras). Arabinogalactanos são proteoglicanos constituídos 90% de por carboidratos. Muitos são solúveis, mas alguns possuem caudas lipídicas que servem para ancoramento na membrana.

Parecem ter um papel estrutural na interface entre a membrana plasmática e a parede celular, mantendo a rede celulose-hemicelulose numa conformação aberta e regulando o grau de hidratação e porosidade da parede e da lamela (Albersheim *et al.*, 2011). Acredita-se que os arabinogalactanos solúveis são aqueles liberados nas duas primeiras frações.

As paredes celulares das três espécies possuem arquiteturas diferentes, mostrado pelas diferentes interações entre as moléculas assumem nas diferentes espécies.

8. Conclusão

Os polímeros da parede estão ligados entre si por inúmeras ligações, covalentes e não-covalentes. A quantidade e o tipo de ligação influenciam a arquitetura da parede celular, que por sua vez têm forte influência nos processos de extração de celulose das madeiras dessas espécies. Com base nos resultados encontrados nesse capítulo pôde-se propor um modelo de arquitetura para as paredes secundárias das três espécies (Figura 23).

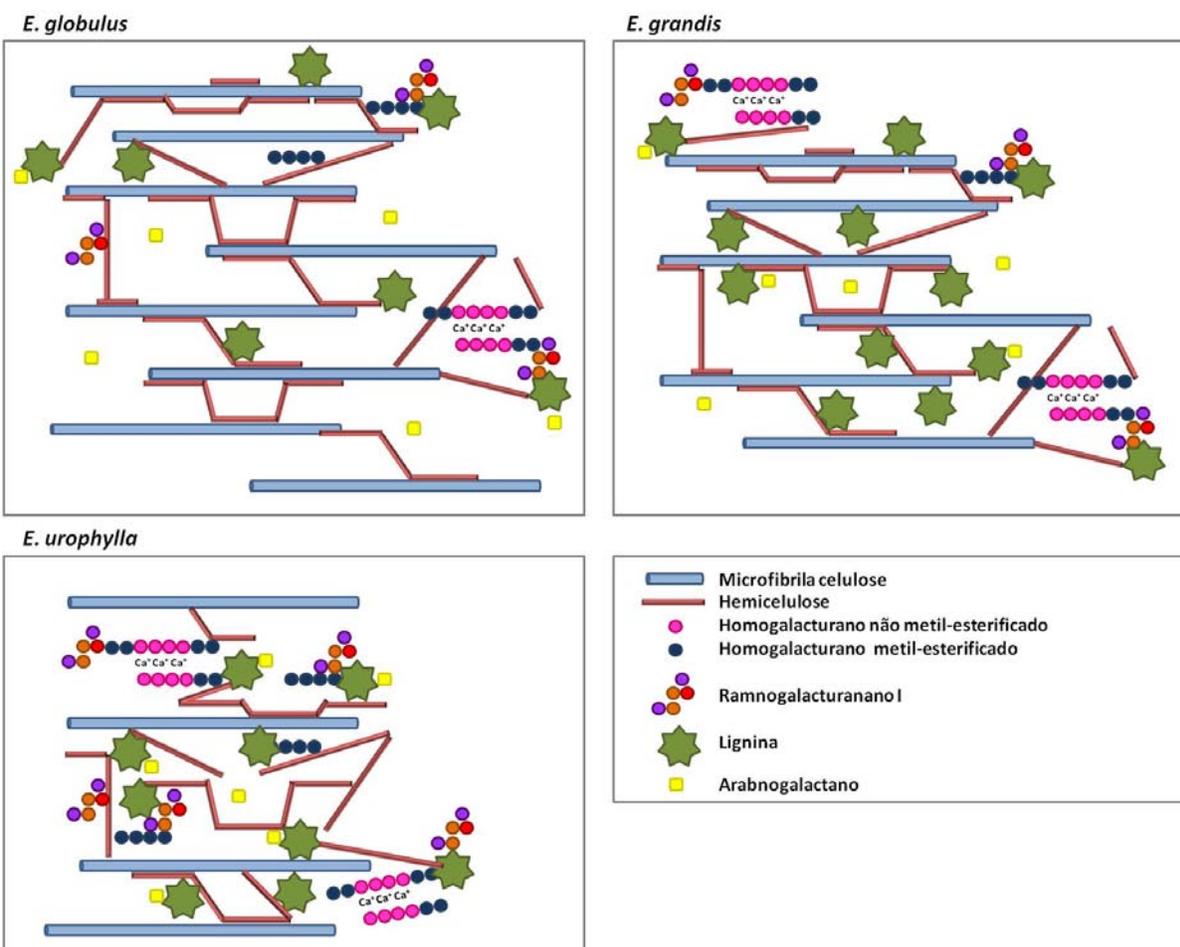


Figura 23: Modelo proposto de arquitetura da parede celular secundária de três espécies de *Eucalyptus*, com base nos resultados obtidos nesse capítulo.

O modelo de arquitetura proposto aqui sugere uma parede celular mais frouxa em *E. globulus* e uma parede mais “comprimida” para *E. urophylla*. A hemicelulose presente em *E. globulus* está principalmente ligada à celulose e pectina, enquanto que em *E. urophylla*, parece estar mais ligada à lignina. Esta ligação com a lignina parece também ocorrer com outras moléculas da parede em *E. urophylla*. Não só o tipo de monômeros de lignina, mas também o tipo de ligação que essas moléculas fazem parecem aumentar a recalcitrância desta espécie à tratamentos para obtenção de polpa de celulose. Assim, podemos observar que não há diferença entre as razões de monômeros do tipo S e G entre *E. grandis* e *E. urophylla*, mas a maior recalcitrância dessa última provavelmente se deve pois a lignina ali presente se liga mais fortemente às outras moléculas, de forma a “travar” essa parede e impedir o seu desmembramento. Além disso, parece haver uma maior quantidade de celulose em *E. globulus* (Cap. 1).

9. Referências

ALBERSHEIM, P.; DARVILL, A.; ROBERTS, K.; SEDEROFF, R.; STAEHELIN, A. **Plant Cell Walls – from chemistry to biology**. Garland Science, New York. 430pp. 2011.

AMARAL, L. I. V.; GASPAR, M.; COSTA, P. M. F.; AIDAR, M. P. M.; BUCKERIDGE, M. S. Novo método enzimático rápido e sensível de extração e dosagem de amido em materiais vegetais. **Hoehnea**, v. 34 (4), p. 425 – 431. 2007.

BALAKSHIN, M. Y.; CAPANEMA, E. A.; CHANG, H. M. **Recent advances in the isolation and analysis of lignins and lignin-carbohydrate complexes**. *In*: Characterization of Lignocellulosic Materials. Blackwell Publishing, Oxford, pp 148–170. 2008.

BARBOSA, L. C. A.; MALTHA, C. R. A.; SILVA, V. L. Determinação da relação siringila/guaiacila da lignina em madeiras de eucalipto por pirólise acoplada à cromatografia gasosa e espectrometria de massas (PI-CG/EM). **Química Nova**, v. 31(8), p. 2035 – 2041. 2008.

BRENDEL, O; IANETTA, P.P.M. & STEWART, D. A rapid and simple method to isolate pure alpha-cellulose. **Phytochemical. Analysis**. 11: 7-10. 2000.

BRETT, C. & WALDRON, K. N. **Physiology and biochemistry of plant cell walls**. 1ed. Champ and Hall, London. 259pp. 1996.

BUCKERIDGE, M.S., SILVA, G.B. & CAVALARI, A. A. **Parede Celular**. *In*: Kerbauy, G.B. (ed). 2ed. Fisiologia Vegetal, Guanabara Koogan: Rio de Janeiro. P. 165-181. 2008.

BURGER, L.M.; RICHTER, H.G. Anatomia da madeira. São Paulo. Livraria Nobel S.A. 154p. 1991.

CARPITA, N. C. Hemicellulosic polymers of cell-walls of *Zea* coleoptiles. **Plant Physiology**. v. 72, p. 512 – 521. 1983.

CARPITA, N. C. Fractionation of hemicelluloses from maize cell walls with increasing concentrations of alkali. **Phytochemistry**, v. 23(5), p. 1089 – 1093. 1984.

DeMARTINI, J. D.; PATTATHIL, S.; AVCI, U.; SZEKALSKI, K.; MAZUMDER, K.; HANH, M. G. WYMAN, C. E. Application of monoclonal antibodies to investigate plant cell wall deconstruction for biofuels production. **Energy & environmental science**, v. 4 (4332). 2011.

DUBOIS, M.; GILLES, K. A.; HAMILTON, J. K.; REBERS, P. A.; SMITH, F. Calorimetric method determination of sugars and related substances. **Analytical chemistry**. v. 28 (3), p. 350 – 356. 1956.

FARDIM, P. & DURÁN, N. Retention of cellulose, xylan and lignin in kraft pulping of *Eucalyptus* studied by multivariate data analysis: influences of physicochemical and mechanical properties of pulp. **Journal of the Brazilian Chemical Society**, v.15 (4), p. 514-522. 2004.

FERRAZ, A.; BAEZA, J.; RODRIGUEZ, J.; FREER, J. Estimating the chemical composition of biodegraded pine and eucalyptus wood by DRIFT spectroscopy and multivariate analysis. **Bioresource Technology**, v. 74, p. 201-212 2000.

GORSHKOVA, T. A.; WYATT, S. E.; SALNIKOV, V. V.; GIBEAUT, D. M.; IBRAGIMOV, M. R.; LOZOVAYA, V. V.; CARPITA, N. C. Cell-wall polysaccharides of developing flax plants. **Plant Physiology**. v. 110, p. 721 – 729. 1996.

GUERRA, A.; FILPPONEN, I.; LUCIA, L. A.; SAQUING, C.; BAUMBERGER, S.; ARGYROPOULOS, D. S. Toward a better understanding of the lignin isolation process from wood. **Journal of agriculture and food chemistry**, v. 54, p. 5939 – 5947. 2006.

MASUKO, T.; MINAMI, A.; IWASAKI, N.; MAJIMA, T.; NISHIMURA, S-I.; LEE, Y, C. Carbohydrate analysis by phenol-sulfuric acid method in microplate format. **Analytical Biochemistry**. v. 339 (1), p. 69 -72. 2005.

MIMMS, A. **Kraft pulping, a compilation of notes**. Atlanta. TAPPI PRESS. ed.2, 181p. 1993.

MORAES, F.; HANH, M. G.; BRAGA, M. R. Comparative analysis of leaf cell-wall polysaccharides of *Dialypetalanthus fuscescens* and *Bathysa meridionalis*: evidence of biochemical similarities between Dialypetalanthaceae and Rubiaceae-Cinchonoideae. **Revista Brasileira de Botânica**. v. 24(3), p. 289 – 294. 2001.

PATTATHIL, S.; AVCI, U.; BALDWIN, D.; SWENNES, A. G.; MCGILL, J. A.; POPPER, Z.; BOOTTEN, T.; ALBERT, A.; DAVIS, R. H.; CHENNAREDDY, C.; DONG, R.; O'SHEA, B.; ROSSI, R.; LEOFF, C.; FRESHOUR, G.; NARRA, R.; O'NEIL, M.; YORK, W. S.; HAHN, M. G. A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. **Plant Physiology**. v. 153, p. 514 – 525. 2010.

PETTERSEN, R. C. The Chemical Composition of Wood, *In*: "Advance in Chemistry Series 207: The Chemistry of Solid Wood", Rowell, R., Ed., ASC, Washington,D. C., 57-126. 1984.

RAISKILA, S.; PULKKINEN, M.; LAAKSO, T.; FAGERSTEDT, K.; LOIJA, M.; MAHLBERG, R.; PAAJANEN, L.; RITSCHKOFF, A-N.; SARANPAA, P. FTIR spectroscopic

prediction of Klason and acid soluble lignin variation in Norway spruce cutting clones. **The Finish Society of Forest Science**, v. 41(2), p. 351 – 371. 2007.

RENCORET, J.; GUTIÉRREZ, A.; NIETO, L.; JIMÉNEZ-BARBERO, J.; FAULDS, C. B.; KIM, H.; RALPH, J.; MARTÍNEZ, A. T.; DEL RÍO, J. C. Lignin composition and structure in young versus adult *Eucalyptus globulus* plants. **Plant Physiology**, v. 155, p. 667 – 682. 2010.

RODRIGUES, J.; MEIER, D.; FAIX, O.; PEREIRA, H.; **Journal of Analytical and Applied Pyrolysis**, v. 48 (121). 1999.

SALISBURY, F. B. & ROSS, C. W. **Plant Physiology**. 4. ed. Belmont: Wassworth Publishing Company, 1991. 682p.

SARKANEN, K. V. The chemistry of delignification in pulp bleaching. **Pure and Applied Chemistry**, v. 5 (12), p. 219 – 213. 1962.

SHI, R, SUN, Y-H.; LI, Q.; HEBER, S.; SEDEROFF, R.; CHIANG, V. L. Towards a systems approach for lignin biosynthesis in *Populus trichocarpa*: transcript abundance and specificity of the monolignol biosynthetic genes. **Plant Cell Physiology**, v.51(10), p. 144 – 163. 2010.

THEANDER, O.; WESTERLUND, E. Studies on dietary fiber. 3. Improved procedures for analysis of dietary fiber. **Journal of Agricultural Food Chemistry**, v.34, p.330-336, 1986.

CONCLUSÃO GERAL:

As três espécies aqui estudadas são importantes por possuírem características de interesse econômico para a indústria de papel e celulose. Entretanto essas características são contrastantes em alguns aspectos. *E. globulus* é a espécie conhecida por ter um dos maiores rendimentos, necessitando de 25% menos madeira para uma tonalada de celulose. Entretanto essa espécie não é adaptada às condições climáticas tropicais. Diversos programas de melhoramento genético convencional com essa espécie já foram realizados, mas ainda não alcançando plantio comercial. Já *E. grandis* é conhecida pelo seu rápido crescimento e por ter boa adaptabilidade e resistência a diversos tipos de stress (biótico e abiótico). *E. urophylla* é uma espécie altamente resistente à diferentes patógenos e com alta plasticidade ambiental. O híbrido *E. grandis* X *E. urophylla* é hoje o mais plantado comercialmente no Brasil. As plantações do Brasil têm hoje um rendimento em média de 44 (m³/ha/ano), mas tem o potencial de atingirem 70 de rendimento.

Tanto o melhoramento genético convencional como transgenia são necessários para entender as bases moleculares e químicas que estão por trás dessas diferenças. A super-expressão de genes chave que influenciam na produtividade, quanto modificações na arquitetura da parede celular podem ser alternativas para se alcançar de forma mais rápida o rendimento. Entretanto, **para modificar é preciso entender.**

A figura 24 apresenta uma visão geral de uma célula vegetal destacando as principais vias de biossíntese de celulose, hemicelulose e lignina. Os genes mais expressos em *E. globulus* (círculos rosa) indicam que essa espécie tem o potencial de produzir uma maior quantidade de celulose, em comparação com as outras duas. A diminuição na quantidade de amido observada, também pode ser um indício de que a glicose produzida pode estar sendo desviada para a produção de celulose e hemicelulose. A hemicelulose tem um importante papel durante o processo industrial, pois diminui o tempo requerido no refino da polpa de celulose.

Além disso, como já discutido no capítulo 2, essa espécie, possui uma lignina mais facilmente extraível (maior quantidade de monômeros do tipo S), como também uma parede com arquitetura mais frouxa. Provavelmente pelo fato de a lignina presente não interagir tão forte com as outras moléculas (quantidade e tipos de ligações).

A quantidade de lignina na amostra pode contribuir para a recalcitrância dessas espécies, mas como visto aqui, este não é o principal fator. Em *E. grandis*, espécie com recalcitrância intermediária, observa-se maior expressão dos genes responsáveis pela síntese de lignina (Figura 24, seta vermelha). Não se observou diferença na razão entre os monômeros

ANEXOS

Additional File 1:

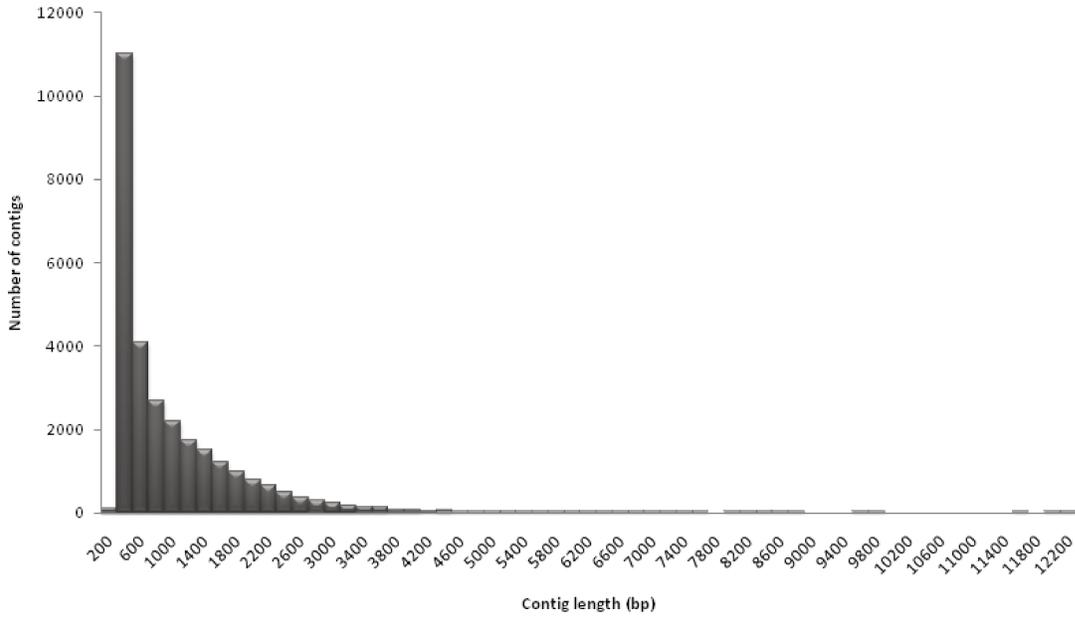


Figure S1: Distribution of contig lengths: 29,292 *EUCANEXT* contigs (min: 200, max: 12,053, mean length: 899.5, n50: 1442bp).

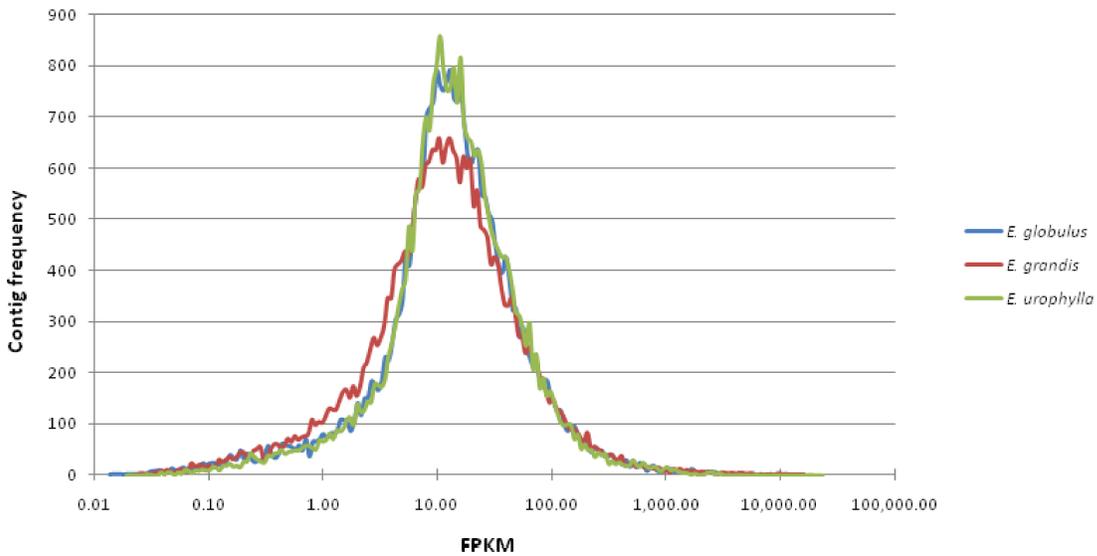


Figure S2: Distribution tail of FPKM values vs. the contig frequency for each xylem library.

Additional File 2:

Table S1: Cell wall genes: FPKM values, fold-changes and p-values for all genes mentioned in the text. Only isoforms with the most significant expression in the xylem were considered. The dark-gray squares indicate the fold-changes of differentially expressed genes with fold changes >2 and p-values <0.01. The light-gray squares are genes that tended to be differential, although the fold-change is 2.0 > x > 1.5 (p-value 0.01). GL: *E. globulus*; GR: *E. grandis*; UR: *E. urophylla*.

| Number | Contig | FPKM | | | Fold-change | | | P_value | | |
|--------|--------------|--------------|--------------|--------------|-------------|-------|-------|---------|-------|-------|
| | | <i>E. gl</i> | <i>E. gr</i> | <i>E. ur</i> | GR/GL | GR/UR | GL/UR | GR/GL | GR/UR | GL/UR |
| 1 | contig_2805 | 577.24 | 610.98 | 507.94 | 1.06 | 1.20 | 1.14 | 0.33 | 0.00 | 0.04 |
| 2 | contig_268 | 565.04 | 689.59 | 525.17 | 1.22 | 1.31 | 1.08 | 0.00 | 0.00 | 0.23 |
| 3 | contig_31 | 536.16 | 387.13 | 403.16 | - | - | 1.33 | 0.00 | 0.57 | 0.00 |
| 4 | Contig102138 | 128.76 | 186.80 | 62.64 | 1.45 | 2.98 | 2.06 | 0.00 | 0.00 | 0.00 |
| 5 | Contig103224 | 126.76 | 121.05 | 46.14 | 1.05 | 2.62 | 2.75 | 0.72 | 0.00 | 0.00 |
| 6 | Contig103901 | 118.60 | 26.19 | 66.91 | 4.53 | 2.55 | 1.77 | 0.00 | 0.00 | 0.00 |
| 7 | contig_5238 | 112.46 | 65.43 | 82.21 | 1.72 | 1.26 | 1.37 | 0.00 | 0.17 | 0.03 |
| 8 | contig_4202 | 111.82 | 57.50 | 70.02 | 1.94 | 1.22 | 1.60 | 0.00 | 0.27 | 0.00 |
| 9 | contig_18438 | 96.60 | 65.72 | 34.66 | 1.47 | 1.90 | 2.79 | 0.02 | 0.00 | 0.00 |
| 10 | Contig102040 | 95.48 | 75.12 | 38.91 | 1.27 | 1.93 | 2.45 | 0.12 | 0.00 | 0.00 |
| 11 | contig_7101 | 88.10 | 48.72 | 88.52 | 1.81 | 1.82 | 1.00 | 0.00 | 0.00 | 0.97 |
| 12 | Contig102834 | 78.67 | 42.71 | 31.58 | 1.84 | 1.35 | 2.49 | 0.00 | 0.20 | 0.00 |
| 13 | Contig103506 | 76.91 | 23.84 | 24.87 | 3.23 | 1.04 | 3.09 | 0.00 | 0.88 | 0.00 |
| 14 | Contig103585 | 60.10 | 12.07 | 17.68 | 4.98 | 1.47 | 3.40 | 0.00 | 0.31 | 0.00 |
| 15 | contig_6128 | 59.04 | 31.69 | 37.51 | 1.86 | 1.18 | 1.57 | 0.00 | 0.49 | 0.03 |
| 16 | Contig102181 | 53.38 | 14.59 | 16.44 | 3.66 | 1.13 | 3.25 | 0.00 | 0.74 | 0.00 |
| 17 | contig_19509 | 52.24 | 47.24 | 53.76 | 1.11 | 1.14 | 1.03 | 0.62 | 0.52 | 0.88 |
| 18 | contig_22590 | 37.36 | 30.61 | 58.50 | 1.22 | 1.91 | 1.57 | 0.42 | 0.00 | 0.03 |
| 19 | contig_18095 | 37.07 | 22.09 | 22.45 | 1.68 | 1.02 | 1.65 | 0.05 | 0.96 | 0.06 |
| 20 | contig_45342 | 35.35 | 27.37 | 39.34 | 1.29 | 1.44 | 1.11 | 0.32 | 0.15 | 0.65 |
| 21 | Contig105625 | 15.53 | 7.51 | 2.05 | 2.07 | 3.67 | 7.59 | 0.10 | 0.09 | 0.00 |
| 22 | Contig104749 | 14.58 | 4.96 | 6.47 | 2.94 | 1.30 | 2.25 | 0.03 | 0.67 | 0.08 |
| 23 | Contig109040 | 14.09 | 6.23 | 1.67 | 2.26 | 3.72 | 8.42 | 0.09 | 0.12 | 0.00 |
| 24 | contig_48170 | 8.02 | 3.68 | 4.21 | 2.18 | 1.15 | 1.90 | 0.22 | 0.86 | 0.29 |
| 25 | contig_45044 | 1.81 | 44.40 | 22.51 | 24.59 | 1.97 | 12.47 | 0.00 | 0.01 | 0.00 |
| 26 | contig_93894 | 0.04 | 11.37 | 9.28 | 306. | 1.23 | - | 0.00 | 0.65 | 0.00 |

CesA
super
family

| | | | | | | | | | | | |
|----------|----|--------------|---------|---------|---------|--------|-----------|--------|------|------|------|
| | | | | | | 80 | | 250.33 | | | |
| | 27 | contig_93894 | 0.04 | 11.37 | 9.28 | 306.80 | 1.23 | 250.33 | 0.00 | 0.65 | 0.00 |
| Susy | 1 | contig_8830 | 2292.54 | 2399.56 | 1136.92 | 1.05 | 2.11 | 2.02 | 0.12 | 0.00 | 0.00 |
| | 2 | contig_7488 | 1401.91 | 1992.43 | 723.66 | 1.42 | 2.75 | 1.94 | 0.00 | 0.00 | 0.00 |
| | 3 | contig_69050 | 1040.89 | 657.13 | 461.37 | -1.58 | 1.42 | 2.26 | 0.00 | 0.00 | 0.00 |
| | 4 | contig_40644 | 346.64 | 175.46 | 108.63 | -1.98 | 1.62 | 3.19 | 0.00 | 0.00 | 0.00 |
| | 5 | contig_19008 | 148.66 | 103.45 | 175.63 | -1.44 | -1.70 | -1.18 | 0.00 | 0.00 | 0.13 |
| Expansin | 1 | contig_86402 | 435.75 | 488.55 | 365.18 | 1.12 | 1.34 | 1.19 | 0.08 | 0.00 | 0.01 |
| | 2 | contig_17470 | 124.52 | 114.31 | 85.32 | -1.09 | 1.34 | 1.46 | 0.51 | 0.04 | 0.01 |
| | 3 | Contig102340 | 93.83 | 104.88 | 61.16 | 1.12 | 1.71 | 1.53 | 0.43 | 0.00 | 0.01 |
| | 4 | Contig102226 | 75.07 | 90.32 | 42.94 | 1.20 | 2.10 | 1.75 | 0.24 | 0.00 | 0.00 |
| | 5 | Contig102122 | 73.64 | 65.80 | 32.92 | -1.12 | 2.00 | 2.24 | 0.51 | 0.00 | 0.00 |
| | 6 | Contig103889 | 62.28 | 91.14 | 30.53 | 1.46 | 2.99 | 2.04 | 0.02 | 0.00 | 0.00 |
| | 7 | contig_74595 | 60.46 | 68.78 | 48.37 | 1.14 | 1.42 | 1.25 | 0.47 | 0.06 | 0.25 |
| | 8 | Contig103132 | 58.02 | 66.37 | 35.56 | 1.14 | 1.87 | 1.63 | 0.46 | 0.00 | 0.02 |
| | 9 | contig_4175 | 38.16 | 34.69 | 25.49 | -1.10 | 1.36 | 1.50 | 0.69 | 0.24 | 0.11 |
| | 10 | contig_18464 | 22.16 | 14.82 | 13.20 | -1.50 | 1.12 | 1.68 | 0.23 | 0.77 | 0.14 |
| XTH | 1 | contig_8735 | 1197.75 | 729.13 | 546.19 | -1.64 | 1.33 | 2.19 | 0.00 | 0.00 | 0.00 |
| | 2 | contig_11204 | 342.33 | 250.13 | 195.87 | -1.37 | 1.28 | 1.75 | 0.00 | 0.01 | 0.00 |
| | 3 | contig_69791 | 258.61 | 167.44 | 120.86 | -1.54 | 1.39 | 2.14 | 0.00 | 0.01 | 0.00 |
| | 4 | contig_45421 | 179.83 | 125.60 | 145.04 | -1.43 | -1.15 | 1.24 | 0.00 | 0.24 | 0.05 |
| | 5 | contig_85003 | 117.92 | 70.91 | 74.66 | -1.66 | -1.05 | 1.58 | 0.00 | 0.76 | 0.00 |
| | 6 | contig_17359 | 19.53 | 54.12 | 22.86 | 2.77 | 2.37 | -1.17 | 0.00 | 0.00 | 0.61 |
| | 7 | contig_41740 | 56.39 | 49.73 | 41.03 | -1.13 | 1.21 | 1.37 | 0.52 | 0.36 | 0.12 |
| | 8 | contig_83736 | 38.98 | 35.40 | 5.51 | -1.10 | 6.42 | 7.07 | 0.68 | 0.00 | 0.00 |
| | 9 | contig_33395 | 46.91 | 25.94 | 13.56 | -1.81 | 1.91 | 3.46 | 0.01 | 0.05 | 0.00 |
| | 10 | contig_45566 | 12.11 | 25.42 | 17.10 | 2.10 | 1.49 | -1.41 | 0.03 | 0.21 | 0.36 |
| | 11 | contig_58698 | 3.85 | 17.66 | 10.61 | 4.59 | 1.66 | -2.76 | 0.00 | 0.19 | 0.08 |
| | 12 | Contig100254 | 6.48 | 15.66 | 6.93 | 2.42 | 2.26 | -1.07 | 0.05 | 0.07 | 0.91 |
| | 13 | contig_18863 | 15.98 | 15.02 | 9.93 | -1.06 | 1.51 | 1.61 | 0.86 | 0.32 | 0.24 |
| | 14 | contig_30193 | 21.36 | 9.36 | 22.22 | -2.28 | -2.37 | -1.04 | 0.03 | 0.02 | 0.90 |
| PAL | 1 | contig_10482 | 299.12 | 699.52 | 502.08 | 2.34 | 1.39 | -1.68 | 0.00 | 0.00 | 0.00 |
| | 2 | contig_16352 | 201.02 | 322.84 | 154.33 | 1.61 | 2.09 | 1.30 | 0.00 | 0.00 | 0.01 |
| | 3 | contig_21173 | 207.05 | 256.43 | 185.62 | 1.24 | 1.38 | 1.12 | 0.02 | 0.00 | 0.28 |
| | 4 | contig_31988 | 2.49 | 16.93 | 175.79 | 6.80 | ### ## | -70.5 | 0.00 | 0.00 | 0.00 |

| | | | | | | | | 8 | | | |
|-------------|---|--------------|--------|---------|--------|----------------|-----------|----------------|------|------|------|
| C4H | 1 | contig_3622 | 703.81 | 1533.83 | 993.15 | 2.18 | 1.54 | - 1.41 | 0.00 | 0.00 | 0.00 |
| | 2 | contig_59694 | 242.93 | 68.68 | 55.29 | - 3.54 | 1.24 | 4.39 | 0.00 | 0.23 | 0.00 |
| 4CL | 1 | contig_6919 | 406.31 | 708.91 | 544.80 | 1.74 | 1.30 | - 1.34 | 0.00 | 0.00 | 0.00 |
| | 2 | contig_90495 | 6.23 | 15.22 | 71.73 | 2.44 | - 4.71 | - 11.5 2 | 0.06 | 0.00 | 0.00 |
| | 3 | Contig106944 | 19.13 | 31.46 | 12.79 | 1.64 | 2.46 | 1.49 | 0.09 | 0.00 | 0.27 |
| | 4 | contig_19066 | 38.19 | 14.22 | 31.88 | - 2.69 | - 2.24 | 1.20 | 0.00 | 0.01 | 0.45 |
| | 5 | contig_19280 | 33.17 | 33.26 | 46.76 | 1.00 | - 1.41 | - 1.41 | 0.99 | 0.13 | 0.13 |
| C3H | 1 | contig_84312 | 449.51 | 660.06 | 424.34 | 1.47 | 1.56 | 1.06 | 0.00 | 0.00 | 0.39 |
| | 2 | contig_47438 | 164.89 | 238.74 | 177.38 | 1.45 | 1.35 | - 1.08 | 0.00 | 0.00 | 0.50 |
| | 3 | contig_13477 | 149.89 | 113.71 | 91.42 | - 1.32 | 1.24 | 1.64 | 0.03 | 0.12 | 0.00 |
| | 4 | contig_58358 | 55.26 | 1.10 | 1.23 | - 50.1 0 | - 1.12 | 44.9 0 | 0.00 | 0.95 | 0.00 |
| | 5 | contig_58373 | 93.90 | 2.68 | 3.42 | - 35.0 8 | - 1.28 | 27.4 7 | 0.00 | 0.78 | 0.00 |
| | 6 | Contig103324 | 223.61 | 294.31 | 186.02 | 1.32 | 1.58 | 1.20 | 0.00 | 0.00 | 0.06 |
| | 7 | Contig106858 | 65.11 | 2.57 | 2.42 | - 25.3 1 | 1.06 | 26.8 7 | 0.00 | 0.95 | 0.00 |
| HCT | 1 | contig_9631 | 264.46 | 334.26 | 228.62 | 1.26 | 1.46 | 1.16 | 0.00 | 0.00 | 0.11 |
| | 2 | contig_4162 | 454.83 | 688.15 | 393.80 | 1.51 | 1.75 | 1.15 | 0.00 | 0.00 | 0.04 |
| | 3 | contig_2187 | 785.98 | 1188.56 | 603.75 | 1.51 | 1.97 | 1.30 | 0.00 | 0.00 | 0.00 |
| | 4 | contig_10228 | 373.01 | 812.34 | 428.25 | 2.18 | 1.90 | - 1.15 | 0.00 | 0.00 | 0.05 |
| | 5 | Contig102271 | 142.74 | 337.09 | 163.49 | 2.36 | 2.06 | - 1.15 | 0.00 | 0.00 | 0.24 |
| CCoA OMT | 1 | contig_4162 | 454.83 | 688.15 | 393.80 | 1.51 | 1.75 | 1.15 | 0.00 | 0.00 | 0.04 |
| | 2 | contig_2187 | 785.98 | 1188.56 | 603.75 | 1.51 | 1.97 | 1.30 | 0.00 | 0.00 | 0.00 |
| | 3 | contig_10228 | 373.01 | 812.34 | 428.25 | 2.18 | 1.90 | - 1.15 | 0.00 | 0.00 | 0.05 |
| | 4 | Contig102271 | 142.74 | 337.09 | 163.49 | 2.36 | 2.06 | - 1.15 | 0.00 | 0.00 | 0.24 |
| CCR | 1 | contig_11747 | 259.11 | 515.63 | 371.59 | 1.99 | 1.39 | - 1.43 | 0.00 | 0.00 | 0.00 |
| | 2 | contig_3044 | 45.18 | 35.13 | 68.08 | - 1.29 | - 1.94 | 1.51 | 0.26 | 0.00 | 0.03 |
| | 3 | contig_24128 | 31.22 | 13.01 | 19.37 | - 2.40 | - 1.49 | 1.61 | 0.01 | 0.27 | 0.10 |
| | 4 | contig_11579 | 5.11 | 9.68 | 24.85 | 1.89 | - 2.57 | - 4.86 | 0.25 | 0.01 | 0.00 |
| | 5 | contig_11598 | 13.59 | 41.60 | 50.59 | 3.06 | - 1.22 | - 3.72 | 0.00 | 0.35 | 0.00 |
| | 6 | contig_23371 | 8.65 | 18.38 | 17.92 | 2.12 | 1.03 | - 2.07 | 0.06 | 0.94 | 0.08 |
| CAD | 1 | contig_1247 | 661.61 | 698.17 | 483.47 | 1.06 | 1.44 | 1.37 | 0.32 | 0.00 | 0.00 |
| | 2 | contig_4533 | 173.73 | 110.01 | 72.99 | - 1.58 | 1.51 | 2.38 | 0.00 | 0.01 | 0.00 |
| | 3 | contig_8428 | 102.57 | 69.12 | 55.08 | - 1.48 | 1.25 | 1.86 | 0.01 | 0.21 | 0.00 |

| | | | | | | | | | | | |
|------|---|--------------|---------|---------|---------|------|-------|--------|-------|-------|-------|
| | 4 | contig_10547 | 33.84 | 72.93 | 38.27 | 2.15 | 1.91 | 1.13 | 0.00 | 0.00 | 0.60 |
| | 5 | contig_90513 | 56.31 | 48.88 | 49.52 | 1.15 | 1.01 | 1.14 | 0.47 | 0.95 | 0.51 |
| | 6 | contig_15056 | 68.87 | 63.88 | 48.02 | 1.08 | 1.33 | 1.43 | 0.67 | 0.14 | 0.05 |
| | 7 | contig_76006 | 4.50 | 18.59 | 6.19 | 4.13 | 3.00 | 1.38 | 0.00 | 0.01 | 0.62 |
| F5H | 1 | contig_390 | 960.17 | 1299.64 | 928.29 | 1.35 | 1.40 | 1.03 | 0.00 | 0.00 | 0.46 |
| COMT | 1 | contig_2396 | 2388.00 | 2970.26 | 2218.05 | 1.24 | 1.34 | 1.08 | 0.00 | 0.00 | 0.01 |
| | 2 | contig_18001 | 0.00 | 266.12 | 5.65 | * | 47.10 | * | 0.00 | 0.00 | 0.02 |
| | 3 | contig_2083 | 1628.25 | 2776.69 | 1385.56 | 1.71 | 2.00 | 1.18 | 0.00 | 0.00 | 0.00 |
| | 4 | contig_2422 | 3102.48 | 4708.53 | 2863.33 | 1.52 | 1.64 | 1.08 | 0.00 | 0.00 | 0.00 |
| | 5 | contig_7294 | 2902.53 | 4478.07 | 2517.46 | 1.54 | 1.78 | 1.15 | 0.00 | 0.00 | 0.00 |
| | 6 | Contig102789 | 1273.98 | 1391.42 | 926.29 | 1.09 | 1.50 | 1.38 | 0.02 | 0.00 | 0.00 |
| | 7 | contig_94134 | 26.03 | 18.17 | 27.64 | 1.43 | 1.52 | 1.06 | 0.24 | 0.17 | 0.83 |
| CHS | 8 | Contig106079 | 41.57 | 0.00 | 0.26 | * | * | 160.58 | 0.00 | 0.84 | 0.00 |
| | 1 | contig_18717 | 45.75 | 52.13 | 289.87 | 1.14 | 5.56 | 6.34 | 0.52 | 0.00 | 0.00 |
| | 2 | contig_34592 | 29.17 | 61.50 | 215.96 | 2.11 | 3.51 | 7.40 | 0.00 | 0.00 | 0.00 |
| | 3 | contig_25107 | 16.72 | 35.24 | 327.71 | 2.11 | 9.30 | 19.60 | 0.01 | 0.00 | 0.00 |
| CHI | 4 | Contig103947 | 22.50 | 43.14 | 334.83 | 1.92 | 7.76 | 14.88 | 0.01 | 0.00 | 0.00 |
| | 1 | contig_13058 | 52.18 | 74.40 | 223.92 | 1.43 | 3.01 | 4.29 | 0.05 | 0.00 | 0.00 |
| F3H | 2 | contig_6799 | 10.96 | 31.59 | 142.18 | 2.88 | 4.50 | 12.98 | 0.00 | 0.00 | 0.00 |
| | 1 | contig_65221 | 22.52 | 33.29 | 105.23 | 1.48 | 3.16 | 4.67 | 0.15 | 0.00 | 0.00 |
| DRF | 1 | contig_13871 | 54.69 | 83.85 | 275.24 | 1.53 | 3.28 | 5.03 | 0.01 | 0.00 | 0.00 |
| | 2 | contig_28580 | 46.39 | 63.67 | 259.64 | 1.37 | 4.08 | 5.60 | 0.10 | 0.00 | 0.00 |
| | 3 | Contig103338 | 31.56 | 32.80 | 167.92 | 1.04 | 5.12 | 5.32 | 0.88 | 0.00 | 0.00 |
| ANS | 1 | contig_10593 | 19.71 | 45.61 | 288.90 | 2.31 | 6.33 | 14.66 | 0.001 | 0.000 | 0.000 |

Table S2: Other genes: FPKM values, fold-changes and p-values for all genes mentioned in the text, however only isoforms with the most significant expression in the xylems are considered. The dark-gray squares indicate the fold-changes of differentially expressed genes with fold changes >2 and p-values<0.01. The light-gray squares are genes that tended to be differential, although the fold-change is 2.0>x>1.5 (p-value 0.01)). GL: *E. globulus*; GR: *E. grandis*; UR: *E. urophylla*.

| | Contig | FPKM | | | Fold-change | | | P_value | | |
|-----|-------------|--------------|--------------|--------------|-------------|-------|-------|---------|-------|-------|
| | | <i>E. gl</i> | <i>E. gr</i> | <i>E. ur</i> | GR/GL | GR/UR | GL/UR | GR/GL | GR/UR | GL/UR |
| Myb | contig_5413 | 170.79 | 262.08 | 208.15 | 1.53 | 1.26 | -1.22 | 1E-05 | 1E-02 | 6E-02 |

| | | | | | | | | | | |
|---------|--------------|--------|--------|--------|--------|--------|--------|-------|-------|-------|
| | Contig102139 | 131.99 | 254.29 | 219.23 | 1.93 | 1.16 | -1.66 | 4E-10 | 1E-01 | 3E-06 |
| | Contig102021 | 137.61 | 145.54 | 121.76 | 1.06 | 1.20 | 1.13 | 6E-01 | 1E-01 | 3E-01 |
| | contig_9940 | 40.59 | 69.30 | 75.39 | 1.71 | -1.09 | -1.86 | 6E-03 | 6E-01 | 1E-03 |
| | contig_10425 | 45.18 | 67.01 | 34.16 | 1.48 | 1.96 | 1.32 | 4E-02 | 1E-03 | 2E-01 |
| | contig_22918 | 37.24 | 66.48 | 39.05 | 1.79 | 1.70 | -1.05 | 4E-03 | 8E-03 | 8E-01 |
| | contig_16088 | 50.42 | 60.78 | 87.97 | 1.21 | -1.45 | -1.74 | 3E-01 | 3E-02 | 1E-03 |
| | contig_11920 | 43.23 | 57.74 | 75.68 | 1.34 | -1.31 | -1.75 | 2E-01 | 1E-01 | 3E-03 |
| | contig_43728 | 77.33 | 56.67 | 48.56 | -1.36 | 1.17 | 1.59 | 7E-02 | 4E-01 | 1E-02 |
| | contig_16528 | 38.48 | 49.29 | 41.76 | 1.28 | 1.18 | -1.09 | 3E-01 | 4E-01 | 7E-01 |
| | contig_8730 | 63.15 | 46.93 | 33.99 | -1.35 | 1.38 | 1.86 | 1E-01 | 2E-01 | 3E-03 |
| | contig_15480 | 76.88 | 44.26 | 58.30 | -1.74 | -1.32 | 1.32 | 3E-03 | 2E-01 | 1E-01 |
| | contig_45561 | 24.98 | 40.00 | 34.40 | 1.60 | 1.16 | -1.38 | 6E-02 | 5E-01 | 2E-01 |
| | Contig102288 | 12.60 | 38.46 | 15.56 | 3.05 | 2.47 | -1.23 | 3E-04 | 2E-03 | 6E-01 |
| | contig_7769 | 30.99 | 37.96 | 34.03 | 1.23 | 1.12 | -1.10 | 4E-01 | 6E-01 | 7E-01 |
| | contig_18879 | 25.19 | 34.11 | 44.46 | 1.35 | -1.30 | -1.76 | 3E-01 | 2E-01 | 2E-02 |
| | Contig103368 | 22.39 | 33.76 | 45.00 | 1.51 | -1.33 | -2.01 | 1E-01 | 2E-01 | 6E-03 |
| | Contig108162 | 1.73 | 5.75 | 41.08 | 3.32 | -7.15 | -23.72 | 2E-01 | 7E-08 | 6E-11 |
| | Contig108033 | 2.10 | 4.82 | 34.23 | 2.30 | -7.10 | -16.31 | 3E-01 | 1E-06 | 1E-08 |
| | contig_19923 | 26.19 | 32.35 | 22.30 | 1.23 | 1.45 | 1.17 | 4E-01 | 2E-01 | 6E-01 |
| | contig_15147 | 21.18 | 29.44 | 30.41 | 1.39 | -1.03 | -1.44 | 2E-01 | 9E-01 | 2E-01 |
| | contig_50581 | 16.42 | 26.04 | 23.43 | 1.59 | 1.11 | -1.43 | 1E-01 | 7E-01 | 3E-01 |
| | Contig108358 | 19.04 | 23.35 | 16.47 | 1.23 | 1.42 | 1.16 | 5E-01 | 3E-01 | 7E-01 |
| | Contig107059 | 15.68 | 22.20 | 13.79 | 1.42 | 1.61 | 1.14 | 3E-01 | 2E-01 | 7E-01 |
| | Contig108312 | 12.57 | 21.76 | 16.74 | 1.73 | 1.30 | -1.33 | 1E-01 | 4E-01 | 4E-01 |
| | contig_10798 | 15.31 | 21.73 | 11.90 | 1.42 | 1.83 | 1.29 | 3E-01 | 9E-02 | 5E-01 |
| | contig_19952 | 12.47 | 18.39 | 24.45 | 1.47 | -1.33 | -1.96 | 3E-01 | 4E-01 | 5E-02 |
| | contig_37214 | 0.19 | 31.62 | 0.31 | 170.24 | 103.03 | -1.65 | 6E-10 | 8E-10 | 9E-01 |
| | contig_63637 | 16.57 | 28.50 | 33.10 | 1.72 | -1.16 | -2.00 | 8E-02 | 6E-01 | 2E-02 |
| | Contig106626 | 9.95 | 27.72 | 30.55 | 2.79 | -1.10 | -3.07 | 4E-03 | 7E-01 | 1E-03 |
| | Contig108099 | 12.21 | 16.39 | 19.13 | 1.34 | -1.17 | -1.57 | 4E-01 | 7E-01 | 2E-01 |
| | contig_52761 | 23.22 | 14.66 | 21.12 | -1.58 | -1.44 | 1.10 | 2E-01 | 3E-01 | 8E-01 |
| WRKY | Contig109446 | 7.08 | 8.26 | 8.12 | 1.17 | 1.02 | -1.15 | 8E-01 | 1E+00 | 8E-01 |
| | Contig104297 | 6.89 | 8.00 | 13.05 | 1.16 | -1.63 | -1.89 | 8E-01 | 3E-01 | 2E-01 |
| | contig_71856 | 3.07 | 7.74 | 8.50 | 2.52 | -1.10 | -2.77 | 2E-01 | 9E-01 | 1E-01 |
| | contig_41240 | 1.61 | 5.65 | 3.71 | 3.50 | 1.52 | -2.30 | 2E-01 | 5E-01 | 4E-01 |
| | contig_82298 | 0.20 | 2.74 | 0.73 | 13.87 | 3.73 | -3.72 | 2E-01 | 3E-01 | 7E-01 |
| | contig_92892 | 0.91 | 1.65 | 20.24 | 1.82 | -12.24 | -22.31 | 7E-01 | 3E-05 | 8E-06 |
| AP2/ERF | contig_30684 | 19.20 | 37.12 | 31.12 | 1.93 | 1.19 | -1.62 | 2E-02 | 5E-01 | 9E-02 |
| | contig_7595 | 22.27 | 32.25 | 40.64 | 1.45 | -1.26 | -1.83 | 2E-01 | 3E-01 | 2E-02 |
| | contig_21551 | 18.08 | 25.39 | 28.06 | 1.40 | -1.11 | -1.55 | 3E-01 | 7E-01 | 1E-01 |
| | contig_57724 | 14.43 | 21.85 | 19.74 | 1.51 | 1.11 | -1.37 | 2E-01 | 7E-01 | 4E-01 |

| | | | | | | | | | | | |
|--------------|--------------|---------|---------|---------|--------|--------|--------|--------|--------|--------|-------|
| | contig_52085 | 5.01 | 16.79 | 14.23 | 3.35 | 1.18 | -2.84 | 1E-02 | 7E-01 | 4E-02 | |
| | contig_47385 | 7.76 | 12.57 | 9.91 | 1.62 | 1.27 | -1.28 | 3E-01 | 6E-01 | 6E-01 | |
| | contig_19853 | 3.84 | 10.42 | 10.41 | 2.72 | 1.00 | -2.71 | 9E-02 | 1E+00 | 9E-02 | |
| | contig_22710 | 12.38 | 8.69 | 9.92 | -1.42 | -1.14 | 1.25 | 4E-01 | 8E-01 | 6E-01 | |
| | Contig101903 | 11.14 | 8.28 | 5.65 | -1.35 | 1.47 | 1.97 | 5E-01 | 5E-01 | 2E-01 | |
| | contig_38719 | 2.13 | 2.38 | 2.17 | 1.12 | 1.10 | -1.02 | 9E-01 | 9E-01 | 1E+00 | |
| | contig_21245 | 5.43 | 1.57 | 5.91 | -3.45 | -3.76 | -1.09 | 2E-01 | 1E-01 | 9E-01 | |
| | contig_50337 | 1.45 | 1.35 | 3.19 | -1.07 | -2.37 | -2.20 | 1E+00 | 4E-01 | 5E-01 | |
| | contig_14329 | 0.00 | 0.00 | 0.00 | - | - | - | 1E+00 | 1E+00 | 1E+00 | |
| | contig_68262 | 0.00 | 0.00 | 0.00 | - | - | - | 1E+00 | 1E+00 | 1E+00 | |
| | contig_69519 | 0.00 | 0.00 | 0.00 | - | - | - | 1E+00 | 1E+00 | 1E+00 | |
| | contig_75755 | 0.00 | 0.00 | 0.00 | - | - | - | 1E+00 | 1E+00 | 1E+00 | |
| | contig_93102 | 0.00 | 0.00 | 0.00 | - | - | - | 1E+00 | 1E+00 | 1E+00 | |
| Ubiquitins | contig_4429 | 3709.59 | 3632.34 | 6056.74 | -1.02 | -1.67 | -1.63 | 4E-01 | 2E-135 | 7E-126 | |
| | Contig102729 | 3595.97 | 2207.14 | 2506.83 | -1.63 | -1.14 | 1.43 | 6E-75 | 1E-05 | 2E-44 | |
| | contig_26643 | 1588.94 | 2109.21 | 2493.01 | 1.33 | -1.18 | -1.57 | 1E-17 | 2E-08 | 8E-46 | |
| | contig_2028 | 930.02 | 2905.75 | 2635.16 | 3.12 | 1.10 | -2.83 | 3E-234 | 3E-04 | 9E-187 | |
| | contig_21904 | 590.31 | 194.02 | 424.46 | -3.04 | -2.19 | 1.39 | 2E-47 | 7E-21 | 2E-07 | |
| | contig_44184 | 535.41 | 1274.27 | 1317.29 | 2.38 | -1.03 | -2.46 | 2E-69 | 4E-01 | 5E-76 | |
| | contig_8452 | 279.42 | 332.89 | 235.46 | 1.19 | 1.41 | 1.19 | 3E-02 | 4E-05 | 5E-02 | |
| | contig_2117 | 184.46 | 317.64 | 463.11 | 1.72 | -1.46 | -2.51 | 2E-09 | 2E-07 | 1E-28 | |
| | contig_10068 | 158.68 | 394.17 | 372.63 | 2.48 | 1.06 | -2.35 | 3E-24 | 4E-01 | 5E-21 | |
| | contig_6805 | 158.01 | 180.26 | 117.05 | 1.14 | 1.54 | 1.35 | 2E-01 | 2E-04 | 1E-02 | |
| | contig_18951 | 148.64 | 153.04 | 310.07 | 1.03 | -2.03 | -2.09 | 8E-01 | 2E-13 | 3E-14 | |
| | contig_25422 | 140.26 | 279.20 | 194.26 | 1.99 | 1.44 | -1.39 | 8E-12 | 9E-05 | 3E-03 | |
| | Contig102611 | 19.38 | 678.41 | 956.18 | 35.01 | -1.41 | -49.34 | 2E-173 | 6E-12 | 4E-254 | |
| | contig_28426 | 105.79 | 419.78 | 254.95 | 3.97 | 1.65 | -2.41 | 2E-45 | 2E-10 | 2E-15 | |
| | contig_42814 | 103.66 | 185.71 | 271.43 | 1.79 | -1.46 | -2.62 | 1E-06 | 6E-05 | 1E-18 | |
| | contig_20016 | 82.20 | 159.07 | 294.78 | 1.94 | -1.85 | -3.59 | 6E-07 | 1E-10 | 2E-29 | |
| | contig_6900 | 54.18 | 223.49 | 192.06 | 4.12 | 1.16 | -3.54 | 8E-26 | 1E-01 | 2E-19 | |
| | Contig103530 | 41.63 | 180.17 | 225.86 | 4.33 | -1.25 | -5.43 | 6E-22 | 2E-02 | 5E-32 | |
| | contig_12497 | 41.46 | 232.84 | 275.62 | 5.62 | -1.18 | -6.65 | 1E-33 | 6E-02 | 8E-44 | |
| | contig_9381 | 15.11 | 205.79 | 261.61 | 13.62 | -1.27 | -17.31 | 3E-44 | 1E-02 | 2E-59 | |
| contig_47153 | 10.37 | 161.53 | 101.56 | 15.58 | 1.59 | -9.79 | 3E-36 | 2E-04 | 3E-20 | | |
| contig_8172 | 5.08 | 204.20 | 268.52 | 40.17 | -1.32 | -52.82 | 5E-54 | 3E-03 | 8E-73 | | |
| contig_16984 | 4.70 | 233.53 | 150.79 | 49.71 | 1.55 | -32.10 | 4E-63 | 2E-05 | 4E-39 | | |
| Resistance | Contig108116 | 100.62 | 0.37 | 10.34 | - | 272.13 | -27.97 | 9.73 | 2E-30 | 2E-03 | 5E-20 |
| | Contig103867 | 78.24 | 2.79 | 7.76 | -28.06 | -2.78 | 10.08 | 2E-20 | 1E-01 | 5E-16 | |
| | Contig103531 | 64.89 | 5.34 | 2.45 | -12.15 | 2.18 | 26.52 | 2E-14 | 3E-01 | 5E-17 | |

| | | | | | | | | | |
|--------------|-------|-------|-------|---------|--------|--------|-------|-------|-------|
| Contig107748 | 55.62 | 0.86 | 0.35 | -65.04 | 2.41 | 156.95 | 3E-16 | 7E-01 | 7E-17 |
| contig_82116 | 50.51 | 41.26 | 22.00 | -1.22 | 1.88 | 2.30 | 3E-01 | 2E-02 | 8E-04 |
| Contig101269 | 38.55 | 4.56 | 13.91 | -8.45 | -3.05 | 2.77 | 5E-08 | 3E-02 | 6E-04 |
| Contig107092 | 38.47 | 3.38 | 13.47 | -11.38 | -3.98 | 2.86 | 8E-09 | 1E-02 | 5E-04 |
| Contig107792 | 34.67 | 4.56 | 11.65 | -7.60 | -2.55 | 2.98 | 5E-07 | 8E-02 | 6E-04 |
| contig_13453 | 34.37 | 4.97 | 25.47 | -6.91 | -5.12 | 1.35 | 1E-06 | 1E-04 | 3E-01 |
| Contig107196 | 32.81 | 5.01 | 11.52 | -6.54 | -2.30 | 2.85 | 3E-06 | 1E-01 | 1E-03 |
| Contig108369 | 31.86 | 1.38 | 0.59 | -23.10 | 2.33 | 53.78 | 1E-08 | 7E-01 | 2E-09 |
| Contig104613 | 30.35 | 5.00 | 5.18 | -6.07 | -1.04 | 5.86 | 1E-05 | 1E+00 | 1E-05 |
| contig_40237 | 30.29 | 10.31 | 18.53 | -2.94 | -1.80 | 1.63 | 2E-03 | 1E-01 | 9E-02 |
| Contig106748 | 28.77 | 3.41 | 17.67 | -8.44 | -5.18 | 1.63 | 3E-06 | 2E-03 | 1E-01 |
| Contig100032 | 26.81 | 3.18 | 12.79 | -8.43 | -4.02 | 2.10 | 7E-06 | 2E-02 | 3E-02 |
| Contig104947 | 25.94 | 1.94 | 1.12 | -13.38 | 1.74 | 23.24 | 2E-06 | 7E-01 | 3E-07 |
| Contig105649 | 24.78 | 1.01 | 13.12 | -24.45 | -12.95 | 1.89 | 5E-07 | 9E-04 | 6E-02 |
| Contig105657 | 24.64 | 0.12 | 0.12 | -204.83 | -1.04 | 197.70 | 6E-08 | 1E+00 | 6E-08 |
| Contig104195 | 24.48 | 8.80 | 6.94 | -2.78 | 1.27 | 3.53 | 7E-03 | 7E-01 | 2E-03 |
| contig_54735 | 23.75 | 2.42 | 3.99 | -9.80 | -1.65 | 5.95 | 1E-05 | 6E-01 | 1E-04 |
| contig_46067 | 23.50 | 4.60 | 3.75 | -5.10 | 1.23 | 6.26 | 3E-04 | 8E-01 | 1E-04 |
| contig_85010 | 23.20 | 8.35 | 4.18 | -2.78 | 2.00 | 5.55 | 8E-03 | 3E-01 | 2E-04 |
| Contig104434 | 22.97 | 0.00 | 0.00 | - | - | 0.00 | 1E-07 | 1E+00 | 1E-07 |
| Contig108770 | 22.24 | 0.00 | 0.45 | - | - | 49.22 | 2E-07 | 7E-01 | 7E-07 |
| Contig104653 | 21.77 | 0.08 | 0.00 | -272.55 | 0.00 | 0.00 | 4E-07 | 9E-01 | 3E-07 |
| contig_26547 | 19.12 | 4.24 | 7.66 | -4.51 | -1.81 | 2.50 | 2E-03 | 3E-01 | 3E-02 |
| contig_42861 | 18.67 | 2.81 | 9.19 | -6.64 | -3.27 | 2.03 | 5E-04 | 7E-02 | 8E-02 |
| contig_33919 | 18.66 | 3.36 | 7.47 | -5.56 | -2.22 | 2.50 | 9E-04 | 2E-01 | 3E-02 |
| Contig101071 | 18.30 | 4.00 | 10.23 | -4.57 | -2.56 | 1.79 | 2E-03 | 1E-01 | 1E-01 |
| Contig100754 | 17.72 | 1.88 | 5.28 | -9.44 | -2.81 | 3.36 | 2E-04 | 2E-01 | 1E-02 |
| Contig104595 | 17.18 | 2.63 | 0.74 | -6.54 | 3.54 | 23.13 | 9E-04 | 4E-01 | 4E-05 |
| Contig109432 | 16.79 | 1.76 | 1.56 | -9.55 | 1.13 | 10.76 | 3E-04 | 9E-01 | 2E-04 |
| Contig109148 | 16.34 | 0.78 | 2.43 | -20.90 | -3.11 | 6.72 | 8E-05 | 4E-01 | 1E-03 |
| Contig101059 | 16.04 | 3.18 | 6.40 | -5.04 | -2.01 | 2.50 | 3E-03 | 3E-01 | 4E-02 |
| Contig109620 | 15.51 | 1.27 | 8.83 | -12.24 | -6.97 | 1.76 | 3E-04 | 2E-02 | 2E-01 |
| Contig109256 | 15.14 | 0.00 | 0.00 | - | - | - | 3E-05 | 1E+00 | 3E-05 |
| contig_51588 | 14.25 | 2.16 | 8.58 | -6.60 | -3.97 | 1.66 | 3E-03 | 6E-02 | 2E-01 |
| contig_33686 | 13.82 | 0.54 | 2.81 | -25.50 | -5.18 | 4.92 | 2E-04 | 3E-01 | 7E-03 |
| contig_17551 | 13.56 | 2.80 | 2.36 | -4.84 | 1.19 | 5.74 | 8E-03 | 9E-01 | 5E-03 |
| contig_37889 | 12.41 | 2.16 | 7.64 | -5.74 | -3.54 | 1.62 | 7E-03 | 9E-02 | 3E-01 |
| Contig109775 | 11.62 | 1.39 | 4.86 | -8.36 | -3.50 | 2.39 | 4E-03 | 2E-01 | 1E-01 |
| contig_14675 | 11.37 | 1.01 | 6.96 | -11.29 | -6.91 | 1.63 | 3E-03 | 4E-02 | 3E-01 |
| Contig109244 | 11.30 | 0.22 | 1.81 | -51.84 | -8.29 | 6.25 | 7E-04 | 4E-01 | 9E-03 |
| contig_26515 | 11.03 | 1.02 | 4.41 | -10.80 | -4.32 | 2.50 | 3E-03 | 2E-01 | 1E-01 |

| | | | | | | | | | |
|--------------|-------|-------|--------|--------|--------|---------|-------|-------|-------|
| Contig103237 | 7.03 | 85.47 | 20.57 | 12.16 | 4.16 | -2.93 | 2E-18 | 8E-11 | 1E-02 |
| Contig103016 | 10.54 | 60.40 | 27.13 | 5.73 | 2.23 | -2.57 | 7E-10 | 3E-04 | 7E-03 |
| Contig106835 | 0.00 | 47.09 | 0.00 | 0.00 | 0.00 | - | 7E-15 | 7E-15 | 1E+00 |
| Contig106455 | 0.24 | 34.93 | 0.40 | 143.53 | 86.87 | -1.65 | 7E-11 | 1E-10 | 9E-01 |
| contig_6481 | 4.43 | 30.56 | 2.14 | 6.89 | 14.30 | 2.07 | 5E-06 | 1E-07 | 4E-01 |
| contig_67819 | 12.48 | 27.02 | 6.17 | 2.16 | 4.38 | 2.02 | 2E-02 | 2E-04 | 2E-01 |
| Contig100430 | 0.64 | 23.55 | 3.23 | 36.97 | 7.29 | -5.07 | 5E-07 | 5E-05 | 2E-01 |
| Contig104914 | 0.00 | 21.78 | 3.16 | 0.00 | 6.89 | | 3E-07 | 1E-04 | 1E-01 |
| Contig108963 | 2.90 | 16.82 | 11.62 | 5.80 | 1.45 | -4.00 | 1E-03 | 3E-01 | 2E-02 |
| Contig104568 | 3.31 | 16.09 | 11.80 | 4.86 | 1.36 | -3.56 | 3E-03 | 4E-01 | 3E-02 |
| contig_2929 | 65.42 | 65.48 | 141.11 | 1.00 | -2.15 | -2.16 | 1E+00 | 1E-07 | 1E-07 |
| contig_15186 | 0.79 | 31.10 | 84.60 | 39.61 | -2.72 | -107.77 | 4E-09 | 5E-07 | 7E-25 |
| contig_6270 | 8.18 | 39.52 | 81.49 | 4.83 | -2.06 | -9.97 | 3E-06 | 1E-04 | 1E-16 |
| contig_70903 | 28.51 | 60.38 | 68.81 | 2.12 | -1.14 | -2.41 | 7E-04 | 5E-01 | 4E-05 |
| contig_56829 | 25.67 | 12.30 | 62.85 | -2.09 | -5.11 | -2.45 | 3E-02 | 2E-09 | 7E-05 |
| Contig103290 | 31.01 | 26.80 | 58.48 | -1.16 | -2.18 | -1.89 | 6E-01 | 6E-04 | 4E-03 |
| contig_50194 | 12.63 | 28.49 | 56.75 | 2.26 | -1.99 | -4.49 | 1E-02 | 2E-03 | 5E-08 |
| contig_6412 | 0.28 | 8.65 | 53.42 | 31.10 | -6.18 | -192.09 | 4E-03 | 3E-09 | 2E-16 |
| Contig106633 | 0.11 | 0.14 | 44.32 | 1.20 | - | -390.37 | 1E+00 | 7E-14 | 7E-14 |
| Contig103849 | 12.37 | 16.59 | 42.42 | 1.34 | -2.56 | -3.43 | 4E-01 | 7E-04 | 4E-05 |
| contig_16305 | 35.17 | 19.55 | 40.83 | -1.80 | -2.09 | -1.16 | 4E-02 | 6E-03 | 5E-01 |
| Contig106532 | 0.00 | 5.76 | 37.56 | 0.00 | -6.52 | ##### | 2E-02 | 5E-07 | 5E-12 |
| Contig101829 | 4.46 | 6.02 | 27.80 | 1.35 | -4.62 | -6.23 | 6E-01 | 1E-04 | 2E-05 |
| contig_18015 | 4.52 | 3.71 | 26.91 | -1.22 | -7.25 | -5.95 | 8E-01 | 1E-05 | 4E-05 |
| contig_3732 | 1.91 | 1.97 | 26.68 | 1.03 | -13.52 | -13.97 | 1E+00 | 1E-06 | 9E-07 |
| Contig106770 | 11.08 | 3.46 | 25.36 | -3.20 | -7.34 | -2.29 | 5E-02 | 2E-05 | 2E-02 |
| Contig101870 | 7.59 | 0.17 | 25.05 | -43.47 | - | -3.30 | 7E-03 | 5E-08 | 2E-03 |
| contig_21657 | 8.95 | 10.08 | 24.68 | 1.13 | -2.45 | -2.76 | 8E-01 | 1E-02 | 7E-03 |
| contig_20672 | 4.95 | 8.88 | 23.61 | 1.79 | -2.66 | -4.77 | 3E-01 | 1E-02 | 4E-04 |
| Contig100083 | 11.42 | 2.84 | 23.46 | -4.02 | -8.26 | -2.05 | 2E-02 | 3E-05 | 4E-02 |
| Contig100530 | 6.99 | 5.97 | 23.30 | -1.17 | -3.90 | -3.33 | 8E-01 | 1E-03 | 3E-03 |
| Contig101496 | 9.70 | 3.78 | 22.21 | -2.56 | -5.87 | -2.29 | 1E-01 | 2E-04 | 3E-02 |
| Contig104419 | 2.32 | 0.87 | 21.89 | -2.68 | -25.28 | -9.45 | 5E-01 | 2E-06 | 3E-05 |
| contig_6058 | 11.31 | 4.56 | 19.14 | -2.48 | -4.20 | -1.69 | 1E-01 | 3E-03 | 2E-01 |
| Contig109794 | 0.21 | 1.78 | 18.99 | 8.37 | -10.66 | -89.23 | 4E-01 | 9E-05 | 3E-06 |
| contig_87186 | 4.50 | 3.14 | 18.82 | -1.43 | -5.99 | -4.18 | 6E-01 | 6E-04 | 3E-03 |
| Contig107447 | 14.58 | 4.29 | 17.51 | -3.40 | -4.08 | -1.20 | 2E-02 | 4E-03 | 6E-01 |
| contig_54946 | 6.92 | 5.71 | 16.82 | -1.21 | -2.94 | -2.43 | 7E-01 | 2E-02 | 4E-02 |
| Contig110189 | 0.00 | 0.00 | 16.72 | - | | | 1E+00 | 9E-06 | 9E-06 |
| Contig105994 | 0.00 | 0.00 | 15.08 | - | | | 1E+00 | 3E-05 | 3E-05 |
| contig_53632 | 9.04 | 3.68 | 15.01 | -2.46 | -4.08 | -1.66 | 1E-01 | 9E-03 | 2E-01 |

| | | | | | | | | | | |
|-------------------|--------------|--------|---------|---------|---------|--------|--------|-------|--------|--------|
| | Contig104379 | 1.30 | 8.56 | 14.78 | 6.58 | -1.73 | -11.36 | 2E-02 | 2E-01 | 5E-04 |
| | Contig108957 | 3.46 | 1.09 | 14.23 | -3.18 | -13.05 | -4.11 | 3E-01 | 5E-04 | 1E-02 |
| | contig_35936 | 1.83 | 0.88 | 11.66 | -2.08 | -13.23 | -6.36 | 6E-01 | 2E-03 | 7E-03 |
| | contig_50856 | 5.14 | 0.74 | 10.99 | -6.90 | -14.76 | -2.14 | 9E-02 | 2E-03 | 2E-01 |
| | contig_78942 | 1.48 | 3.93 | 10.53 | 2.65 | -2.68 | -7.11 | 3E-01 | 9E-02 | 9E-03 |
| | contig_47126 | 8.55 | 0.69 | 10.24 | -12.37 | -14.82 | -1.20 | 1E-02 | 3E-03 | 7E-01 |
| | contig_78643 | 1.28 | 5.44 | 10.23 | 4.26 | -1.88 | -8.01 | 1E-01 | 2E-01 | 8E-03 |
| Nitrate reductase | contig_94502 | 20.33 | 0.85 | 0.95 | -23.86 | -1.12 | 21.38 | 6E-06 | 1E+00 | 8E-06 |
| | Contig109467 | 16.81 | 0.00 | 1.28 | * | * | 13.18 | 9E-06 | 4E-01 | 1E-04 |
| Nitrite reductase | Contig102001 | 178.76 | 0.66 | 1.16 | -272.90 | | -1.77 | 8E-01 | 2E+02 | 3E-52 |
| Sarcosine oxidase | Contig101499 | 29.75 | 0.00 | 0.00 | * | * | | 1E-09 | 1E+00 | 1E-09 |
| | Contig104147 | 23.54 | 0.00 | 0.21 | * | * | 114.58 | 8E-08 | 9E-01 | 2E-07 |
| | Contig105585 | 22.73 | 0.00 | 0.00 | * | * | * | 1E-07 | 1E+00 | 1E-07 |
| | Contig105636 | 5.83 | 7.15 | 10.19 | 1.23 | -1.42 | -1.75 | 7E-01 | 5E-01 | 3E-01 |
| | Contig110164 | 6.79 | 5.80 | 7.21 | -1.17 | -1.24 | -1.06 | 8E-01 | 7E-01 | 9E-01 |
| | contig_2908 | 790.93 | 1722.98 | 2728.42 | 2.18 | -1.58 | -3.45 | 6E-79 | 1E-51 | 3E-247 |
| | Contig103967 | 692.77 | 1545.49 | 2506.46 | 2.23 | -1.62 | -3.62 | 2E-74 | 6E-52 | 2E-239 |
| smHSP | contig_128 | 613.93 | 1301.72 | 2077.53 | 2.12 | -1.60 | -3.38 | 8E-57 | 6E-41 | 1E-184 |
| | contig_11271 | 337.90 | 741.81 | 1184.31 | 2.20 | -1.60 | -3.50 | 2E-35 | 4E-24 | 2E-110 |
| | contig_2772 | 108.34 | 37.24 | 267.46 | 3.97 | 3.08 | -1.29 | 3E-23 | 5E-18 | 2E-01 |
| | contig_15520 | 104.44 | 182.00 | 403.68 | 1.74 | -2.22 | -3.87 | 4E-06 | 2E-20 | 1E-42 |
| | contig_2918 | 94.90 | 173.32 | 841.82 | 1.83 | -4.86 | -8.87 | 1E-06 | 4E-106 | 1E-150 |
| | contig_1248 | 82.74 | 43.37 | 185.58 | 2.12 | -3.67 | -7.80 | 3E-08 | 1E-57 | 2E-98 |
| | contig_4094 | 77.78 | 120.90 | 223.42 | 2.63 | -1.59 | -4.19 | 1E-10 | 9E-06 | 6E-27 |
| | contig_22318 | 74.55 | 158.41 | 581.20 | 1.55 | -1.85 | -2.87 | 2E-03 | 3E-08 | 1E-17 |
| | contig_48196 | 56.95 | 65.18 | 125.39 | 3.52 | -2.02 | -7.11 | 5E-12 | 2E-10 | 2E-38 |
| | contig_10896 | 55.08 | 144.97 | 230.74 | 3.34 | -1.94 | -6.49 | 4E-08 | 1E-06 | 4E-24 |
| | contig_33928 | 52.27 | 207.74 | 67.38 | 4.72 | -1.94 | -9.13 | 4E-10 | 3E-06 | 3E-26 |
| | contig_22032 | 40.16 | 51.90 | 33.20 | 2.52 | -1.92 | -4.85 | 3E-05 | 1E-05 | 3E-17 |
| | contig_9893 | 38.96 | 57.43 | 60.07 | 1.14 | -1.92 | -2.20 | 5E-01 | 1E-05 | 3E-07 |
| | contig_1920 | 36.29 | 48.72 | 77.99 | 2.70 | -1.74 | -4.71 | 2E-05 | 4E-04 | 1E-14 |
| | contig_87909 | 32.55 | 114.42 | 231.50 | 2.01 | -2.05 | -4.12 | 1E-03 | 2E-06 | 3E-15 |
| | contig_7267 | 23.00 | 76.89 | 149.19 | 1.86 | -1.98 | -3.69 | 1E-02 | 2E-04 | 8E-10 |
| | contig_2148 | 30.82 | 61.86 | 126.89 | 1.47 | -1.05 | -1.54 | 6E-02 | 8E-01 | 3E-02 |
| | contig_68530 | 26.14 | 65.93 | 126.89 | 1.79 | -1.84 | -3.30 | 1E-02 | 5E-04 | 4E-09 |
| | contig_8316 | 15.13 | 71.37 | 138.19 | 1.13 | -1.59 | -1.80 | 6E-01 | 5E-02 | 1E-02 |
| | contig_14617 | 29.02 | 45.43 | 45.72 | 1.29 | 1.56 | 1.21 | 2E-01 | 4E-02 | 4E-01 |
| | contig_23565 | 26.93 | 48.11 | 88.76 | 13.37 | 2.88 | -4.65 | 2E-11 | 7E-05 | 3E-03 |
| | contig_1040 | 26.29 | 29.76 | 47.35 | 1.34 | -1.60 | -2.15 | 2E-01 | 9E-03 | 9E-05 |
| | contig_40340 | 23.34 | 43.46 | 86.11 | 1.57 | -1.01 | -1.58 | 6E-02 | 1E+00 | 5E-02 |
| | contig_19664 | 23.12 | 62.40 | 108.82 | 14.71 | 3.02 | -4.87 | 8E-11 | 8E-05 | 5E-03 |

| | | | | | | | | | |
|--------------|-------|-------|-------|-------|--------|---------|-------|-------|-------|
| contig_5572 | 22.10 | 40.83 | 70.21 | -1.91 | -4.28 | -2.24 | 4E-04 | 2E-22 | 2E-10 |
| contig_1749 | 19.54 | 28.32 | 55.56 | 1.85 | -1.72 | -3.18 | 2E-02 | 5E-03 | 3E-07 |
| contig_12998 | 17.92 | 17.11 | 48.90 | -2.91 | -7.18 | -2.47 | 2E-09 | 2E-44 | 8E-17 |
| contig_2313 | 11.83 | 9.89 | 39.96 | 1.45 | -1.96 | -2.84 | 2E-01 | 3E-03 | 3E-05 |
| contig_23555 | 11.68 | 24.22 | 45.68 | 2.07 | -1.89 | -3.91 | 4E-02 | 1E-02 | 4E-06 |
| contig_8237 | 10.52 | 12.74 | 21.43 | -1.05 | -2.86 | -2.73 | 9E-01 | 8E-05 | 1E-04 |
| contig_41197 | 10.07 | 7.15 | 12.54 | 1.21 | -1.68 | -2.04 | 7E-01 | 1E-01 | 6E-02 |
| contig_26247 | 9.54 | 10.88 | 9.42 | 1.14 | 1.16 | 1.01 | 8E-01 | 8E-01 | 1E+00 |
| contig_8921 | 3.65 | 48.73 | 16.94 | -1.20 | -4.04 | -3.38 | 7E-01 | 1E-05 | 7E-05 |
| contig_11998 | 3.42 | 1.08 | 12.50 | 9.67 | -4.33 | -41.83 | 1E-02 | 1E-05 | 7E-11 |
| contig_13145 | 3.05 | 44.94 | 14.88 | 3.74 | 1.10 | -3.41 | 6E-02 | 9E-01 | 9E-02 |
| contig_27300 | 0.90 | 8.73 | 37.77 | -3.18 | -11.60 | -3.65 | 3E-01 | 2E-03 | 2E-02 |
| contig_4304 | 0.04 | 1.04 | 9.60 | 23.92 | -9.22 | -220.59 | 5E-01 | 9E-03 | 1E-03 |
| contig_2047 | 0.00 | 0.86 | 9.47 | 0.00 | -11.00 | ##### | 6E-01 | 7E-03 | 1E-03 |

Additional File 3:

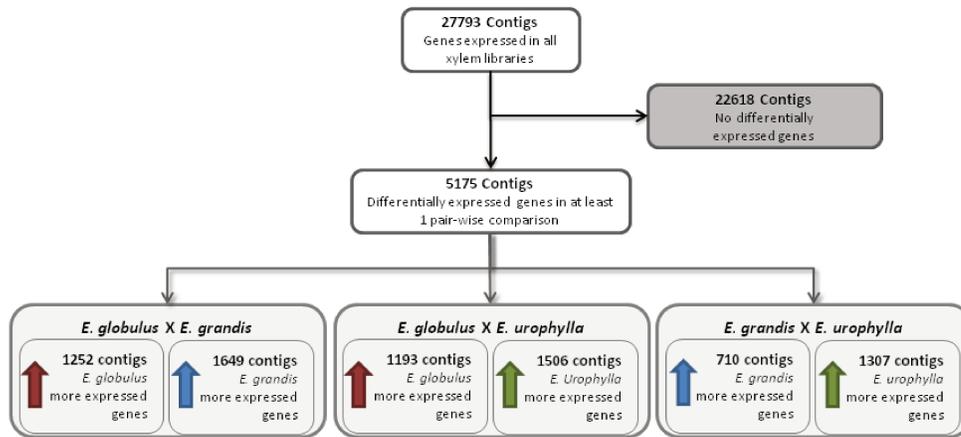


Figure S3: Flow diagram of genes expressed in all xylem libraries (group a, Figure 1). The genes were separated as being “non-differentially expressed” and “differentially expressed”. The differentially expressed genes were analyzed by pair-wise comparisons between species.

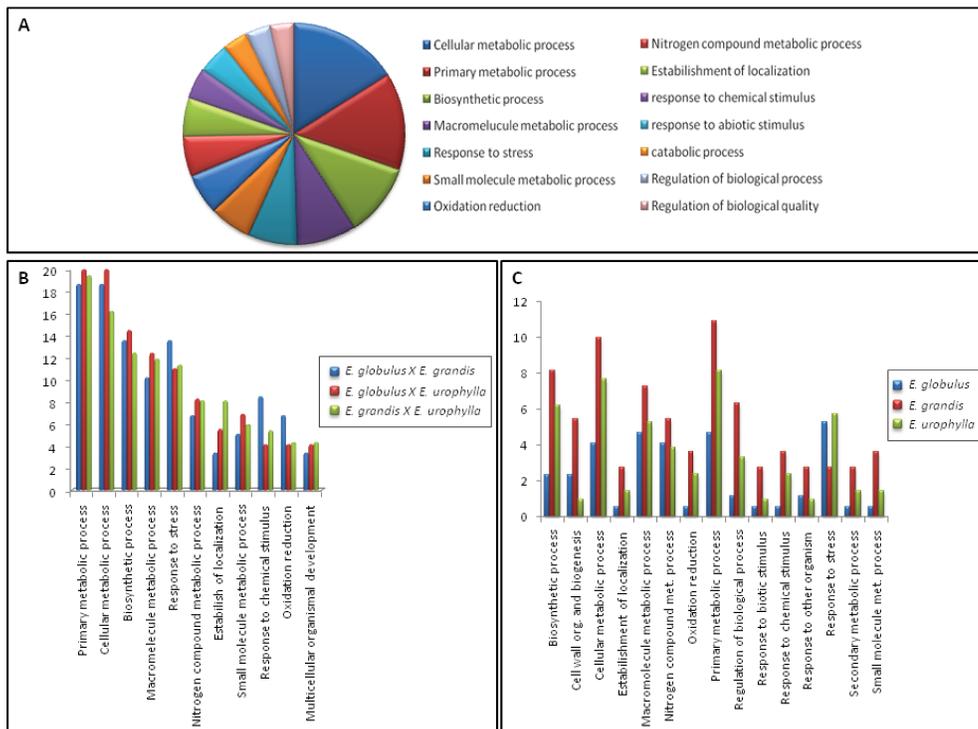


Figure S4: GO categories at Biological Process level 3. A: GO for the top 200 non-differentially expressed genes; B: Representative GO categories of genes shared by only two species; C: Representative GO categories of genes expressed in only one species. B and C: The percentage of contigs in each GO category related to the total number of known function contigs is present on the y-axis.

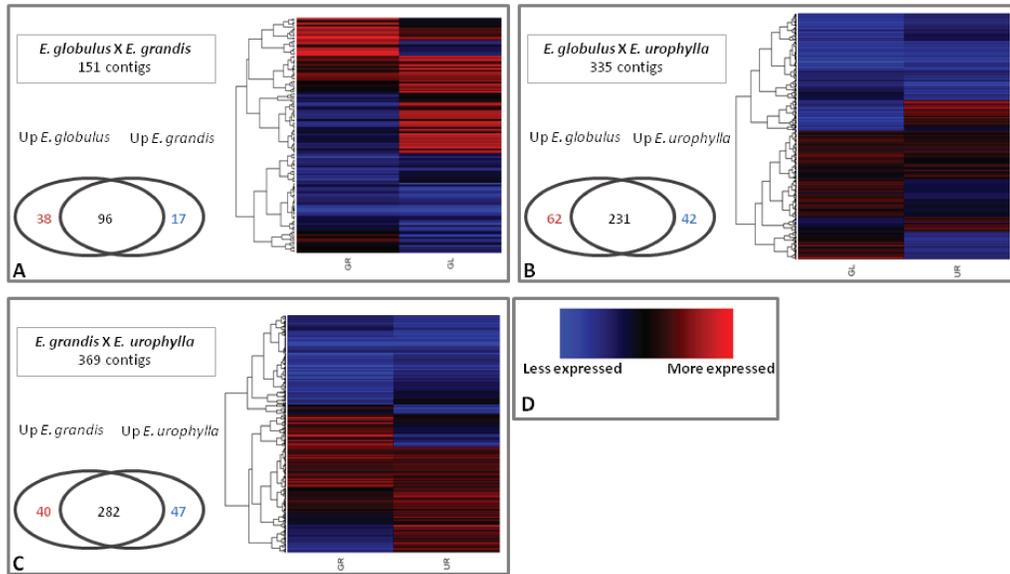


Figure S5: Heatmaps of gene expression profiles used for pair-wise comparisons (groups ‘b’, ‘c’ and ‘d’ in Figure 2) and Venn diagrams showing genes that are differentially and non-differentially expressed among the species. **A:** Gene expression profiling comparison of *E. globulus* and *E. grandis*. **B:** Gene expression profiling comparison of *E. globulus* and *E. urophylla*. **C:** Gene expression profiling comparison of *E. grandis* and *E. urophylla*. **D:** Heatmaps pallet key. GL: *E. globulus*, GR: *E. grandis*, UR: *E. urophylla*.

CAPÍTULO 1.2: Avaliação do efeito da super-expressão de genes chave de *Eucalyptus* em planta modelo (*Arabidopsis thaliana*)

10. Introdução

Diferentes espécies de *Eucalyptus* são reconhecidas por suas características superiores em termos de crescimento, qualidade da madeira e resistência a diferentes tipos de stress. *E. grandis* possui uma certa adaptabilidade a diferentes ambientes mas é principalmente reconhecido por suas propriedades de crescimento, enquanto *E. globulus* possui melhores propriedades da madeira para a indústria de papel e celulose (Grattapaglia & Kirst, 2008). Já *E. urophylla* é a espécie com maior plasticidade ambiental (Potts & Pederick, 2000).

Tais características são, provavelmente, dirigidas pela expressão coordenada de numerosos genes estruturais envolvidos nos processos de xilogênese e de genes regulatórios. A análise de bancos de sequências expressas (ESTs), focando nas comparações entre genes expressos em diferentes tecidos da mesma espécie ou do mesmo tecido em diferentes espécies, permite inferências sobre processos de desenvolvimento e de porque algumas espécies possuem propriedades mais favoráveis, como maior densidade da madeira ou produção de polpa (Poke *et al.*, 2005).

A importância de genes da família celulose sintase na formação da parede celular já foi demonstrada para diversas espécies de plantas (Taylor *et al.*, 1999; Hamann *et al.*, 2004; Brown *et al.*, 2004) Uma mutação de uma isoforma do gene dessa família causa uma redução significativa no conteúdo de celulose total da planta, formando elementos de vasos colapsados. Do mesmo modo, plantas transformadas com o gene sacarose sintase (*Susy*) sob controle do promotor 35S apresentaram um aumento no espessamento de suas paredes celulares (Coleman *et al.*, 2009).

Esses genes também demonstraram seu potencial papel na qualidade da madeira de *Eucalyptus* (capítulo 1), já que a maioria das isoformas dessa família possui expressão aumentada em *E. globulus*, espécie com melhor qualidade da madeira para indústrias de papel e celulose. Além disso, outros genes investigados no capítulo 1 são potenciais para a

contribuição das características diferenciais de madeira encontrada entre as três espécies de *Eucalyptus* estudadas. Dentre estes, nitrato desidrogenase e um gene com função desconhecida foram selecionados.

No capítulo 1, foi desmostrado que a maior parte das diferentes características observadas entre as espécies é provavelmente dirigida por diversos genes ainda com função desconhecida. Dentre os muitos genes com função desconhecida encontrados, aqueles com valor de FPKM alto e aqueles com expressão diferencial entre as espécies merecem destaque, pois podem estar desempenhando um papel importante para determinar as características das espécies, como é o caso do gene denominado UNK875.

Apesar de não possuírem um perfil de expressão regular entre as isoformas, outros genes foram selecionados pelo seu potencial papel no aumento da biomassa: glutamina sintetase, gene importante no metabolismo de nitrogênio e *fasciclin-like arabinogalactan protein* (chamado aqui de *FascArab*) afeta a deposição de celulose e a integridade da matriz (MacMillan *et al.*, 2010). Diversos experimentos mostraram que plantas suplementadas com nitrogênio produzem uma parede celular alterada em relação às quantidades de celulose e lignina produzidas (Pitre *et al.*, 2007, Camargo *et al.*, não publicado). No capítulo 1, demonstramos que *E. globulus* possui genes desta via super-expressos em relação às outras espécies, sugerindo que não só a suplementação de nitrogênio tem a capacidade de influenciar a formação da parede, mas também a capacidade de metabolizar o nitrogênio. Espera-se assim, que a super-expressão desse gene em *A. thaliana* possa melhorar a eficiência de utilização do nitrogênio, causando modificações na quantidade de celulose produzida.

Embora a formação da madeira seja um processo evidentemente característico de árvores, muitas plantas herbáceas, incluindo *Arabidopsis thaliana*, desenvolvem o câmbio vascular e formam xilema secundário. Assim, *Arabidopsis* pode ser considerada como um modelo para o processo de desenvolvimento tanto durante a fase primária como na fase secundária (Nieminen *et al.*, 2004).

Em alguns estudos demonstrou-se que o hipocótilo de *Arabidopsis* passa por um crescimento secundário substancial como resultado da atividade do câmbio vascular, o qual pode ser facilmente avaliado por secção manual combinada com técnicas histológicas. Em particular, um grande número de fibras do xilema secundário e elementos de vaso são produzidos e possuem características estruturais e ultra-estruturais similares com àquelas encontradas em angiospermas. Estas descobertas aliadas às numerosas vantagens de se trabalhar com um sistema geneticamente tratável como *Arabidopsis*, oferecem possibilidades

de análise da regulação gênica do desenvolvimento do xilema secundário (Chaffey *et al.*, 2002).

Desse modo, escolheu-se, nesse trabalho, *A. thaliana* como planta modelo para estudos do efeito da super-expressão de genes relacionados com a qualidade da madeira.

O uso de plantas transgênicas associadas à análise da expressão global tem trazido grande impacto para o entendimento da biologia do xilema. Espera-se desse modo que os resultados desse trabalho possam contribuir para o entendimento dos processos de formação da madeira a fim de orientar o seu melhoramento, bem como correlacionar a composição da madeira com seu perfil de expressão. O aumento da qualidade da madeira, e conseqüentemente sua produtividade tem implicações econômicas significativas, principalmente no que diz respeito ao setor de papel e celulose.

11. Metodologia

11.1 Material vegetal

As amostras de xilema de três espécies (*E. globulus*, *E. grandis* e *E. urophylla*) para análises de expressão gênica (capítulo 1) e clonagem de genes foram coletadas nas plantações da empresa International Paper do Brasil em Mogi-Guaçu (Figura 25).



Figura 25: Método de coleta de amostras de xilema de eucalipto. A: Tronco de *Eucalyptus grandis*. B: Retirada da casca. C: Raspagem do material com lâmina. D: Congelamento imediato do material em nitrogênio líquido.



Figura 26: Cultivo de *Arabidopsis thaliana* em casa de vegetação.

As plantas de *Arabidopsis thaliana* Col. (transformadas e selvagens) para análise de genes de interesse estão sendo cultivadas em sala apropriada no Laboratório de Genômica e Expressão (IB, UNICAMP) (Figura 26). As plantas estão sendo mantidas com temperatura controlada (21-24 °C), com luz constante (intensidade luminosa de 40 $\mu\text{mol}/\text{m}^2/\text{s}$).

As sementes tratadas (em água deionizada por 15 minutos, depois em hipoclorito de sódio 0,3% por 8 minutos seguido de 8 sucessivas lavagens em água deionizada) foram plantadas (cerca de 6 por vaso) em substrato e argila expandida (vermiculita) na proporção de 1:1. As plantas foram regadas com solução nutritiva 3 vezes por semana. Após a transformação, as plantas foram regadas com água destilada.

Para solução nutritiva 3 soluções estoque são preparadas: 1) **Solução de Micronutrientes** (2860 mg/L de H_3BO_3 ; 1810 mg/L de $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$; 100 mg/L de ZnCl_2 ; 20 mg/L $\text{H}_2\text{MoO}_4 \cdot \text{H}_2\text{O}$; 40 mg/L CuCl_2). 2) **Solução de Fe-EDTA** (24,9 g/L de $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$; 26,1 g/L de EDTA). 3) **Solução nutriente (10X)** (50 mmol/L KNO_3 ; 25 mmol/L KH_2PO_4 ; 40 mmol/L MgSO_4 ; 20 mmol/L $\text{Ca}(\text{NO}_3)_2$). A solução nutriente foi diluída para 1X e adicionou-se 1mL/L de Solução de Micronutrientes e 1mL/L Solução de Fe-EDTA.

11.2 Extração de RNA

O RNA do xilema coletado foi extraído de acordo com protocolo descrito por Zeng & Yang (2002), seguindo algumas alterações proposta por Provost *et al.* (2007). Como descrito a seguir:

As amostras foram pulverizadas maceradas em nitrogênio líquido utilizando almofariz e pistilo e transferidas imediatamente para microtubo contendo 1 mL tampão de extração (Tris-HCl (pH 8,0) 100 mM; EDTA 25 mM; NaCl 2 M; PVP 2%; Spermidina 0,05%; β -mercapto etanol

2% e CTAB 2%) pré-aquecido à 65°C. As amostras foram homogenizadas no vortex por 15 segundos e incubadas por 10 minutos a 65°C e homogenizadas por inversão a cada 3 minutos.

Três extrações com 700 µL de CIA (clorofórmio: álcool isoamílico – 24:1) foram realizadas. Após a adição do CIA, as amostras foram homogenizadas por inversão por 1 minuto, e centrifugadas à 4°C (10000g, 10 minutos), o sobrenadante foi transferido para novo tubo para nova extração com CIA. Após a terceira extração, o sobrenadante foi transferido para novo tubo e ¼ do volume de cloreto de lítio 10 M foi adicionado. As amostras foram homogenizadas e incubadas à 4°C *overnight*.

Após 12 horas, aproximadamente, as amostras foram centrifugadas (10000g, 20 minutos) à 4°C e o sobrenadante foi descartado. O *pellet* seco foi ressuscitado em 500µL de tampão STE (Tris-HCl (pH 8,0) 10 mM; EDTA 1 mM; NaCl 0,1 M) pré-aquecido à 65°C. Nesse passo, três tubos de cada amostra foi ressuscitado junto. Novamente realizou-se uma extração com 400 µL de CIA. Após a centrifugação como descrito acima, transferiu-se os sobrenadantes para novos tubos e esses foram mantidos no gelo. Nos tubos dos quais os sobrenadantes foram retirados, adicionou-se mais 150 µL de STE (aquecido). Repetiu-se a centrifugação e o sobrenadante foi adicionado ao tubo mantido no gelo.

As amostras foram precipitadas com 100 µL de NaCl 5 M e 2 volumes de etanol 100% gelado. Após homogenização por inversão, as amostras foram incubadas por 2 horas à -20°C ou 30 minutos à -80°C.

As amostras foram peletadas por centrifugação à 4°C (10000 g, 20 minutos), descartando-se o sobrenadante. O *pellet* foi lavado com 400 µL de etanol 70% gelado. O etanol foi descartado após centrifugação. O *pellet* seco foi ressuscitado em 12µL de água deionizada autoclavada (tratada com DEPC).

Após avaliação da qualidade por gel desnaturante de agarose e quantificação por espectrometria (Nanodrop) (utilizando 2 repetições para cada amostra – total de 2 µL para cada quantificação), este material foi armazenado a -80°C.

11.3 Síntese de cDNA

O RNA mensageiro foi isolado a partir de 75µg de RNA total extraído de xilema, usando o kit D Dynabeads® Magnetic Beads (Life Technologies). A síntese do cDNA foi realizada utilizando-se o kit SuperScript™ II Reverse Transcriptase (Invitrogen), de acordo com as normas do fabricante.

11.4 Sequenciamento de cDNA – RACE

O gene denominado UNK875 não possuía sequência completa nos bancos de dados disponíveis. Dessa forma, um sequenciamento utilizando a técnica RACE (rapid amplification of cDNA ends) foi realizado para esse gene. A sequência disponível para esse gene não possuía sequências de iniciação e terminação. Para tanto, os Kits “5’ RACE System for Rapid Amplification of cDNA Ends” e “3’ RACE System for Rapid Amplification of cDNA Ends” da Invitrogen foram utilizados. Vários primers foram desenhados ao longo da sequência a fim de garantir a amplificação do gene completo.

A montagem com as sequências geradas pelo RACE gerou um contig de 999pb. Novos *primers* foram desenhados para validar a montagem realizada. Após essa confirmação, novos primers de clonagem para este gene foram desenhados.

11.5 Real-Time PCR

O gene com função desconhecida (UNK875) teve sua expressão avaliada entre o xilema das três espécies via Real-Time PCR. As reações foram realizadas em placas de 96 poços usando o reagente *SYBR Green PCR master mix* (Applied Biosystems) (conforme protocolo do fabricante). Nenhuma amplificação não-específica foi observada nas curvas de dissociação. Os valores de expressão relativa foram calculados de acordo com a equação apresentada por Pfaffl (2001). Para tal, dois genes normalizadores foram utilizados: um gene de poligalacturonase e um de transportador de açúcar. A condição de referência usada foi obtida a partir de amostras de cDNA de folhas de *E. grandis*. Os *primers* utilizados estão descritos na tabela 3.

Tabela 3: Primers utilizados para avaliação da expressão por Real-Time PCR

| Gene | Primer Foward | Primer reverse |
|-------------------------|----------------------|-----------------------|
| UNK875 | GATTCTTTCAGGCTGCTTGG | TAGGCGTTGAGGAGGAAAGA |
| Poligalacturonase | GAGGTCGAGAAGGGTAATT | AGAGAATCGATGCTGAAGAA |
| Transportador de açúcar | AATCGTCGTTGAGGTTACAC | CTTTCGTTGTTGATTTCTCC |

11.6 Clonagem de genes de interesse

Os *primers* para clonagem foram desenhados a partir da sequência gênica depositada nos bancos de dados disponíveis acrescidos de adaptadores para vetores de clonagem

apropriados para os seguintes genes: celulose sintase 1 e 2 (*CesA1* e *CesA2*); sacarose sintase 1 e 3 (*Susy 1* e *3*); glutamina sintetase (*GS*); nitrato redutase (*NR*); *Fasciclin-like arabinogalactan protein* (*FascArab*); gene com função desconhecida (UNK875). Após uma análise de restrição das sequências a serem clonadas, apenas para os genes *CesA* escolheu-se o vetor pCambia1302 (pCambia.org), para os demais pFP101 (Bensmihen et al., 2004).

Os vetores carregando a sequência de interesse foram transformados em cepas de *E. coli* (DH10B) para propagação e seleção de vetores plasmidiais. A extração do DNA por lise alcalina foi realizada para isolamento dos vetores, os quais foram sequenciados e digeridos para confirmação da inserção do gene (Figura 27).

Após confirmação, os vetores foram inseridos em *Agrobacterium tumefaciens* cepa GV 3101, como descrito a seguir: o vetor foi inserido em um tubo contendo 100 µL de suspensão de bactérias competentes, incubado no gelo por 30 minutos. As células foram incubadas em nitrogênio líquido até total solidificação e imediatamente transferidas para incubação a 37 °C por 5 minutos, 1 mL de meio LB foi adicionado e este material foi incubado por 2 horas a 28 °C. Após o crescimento, cerca de 500 µL da suspensão de células foi plaqueado em LB ágar contendo o antibiótico apropriado para a seleção do vetor (espectinomicina para pFP101 e canamicina para pCambia1302).

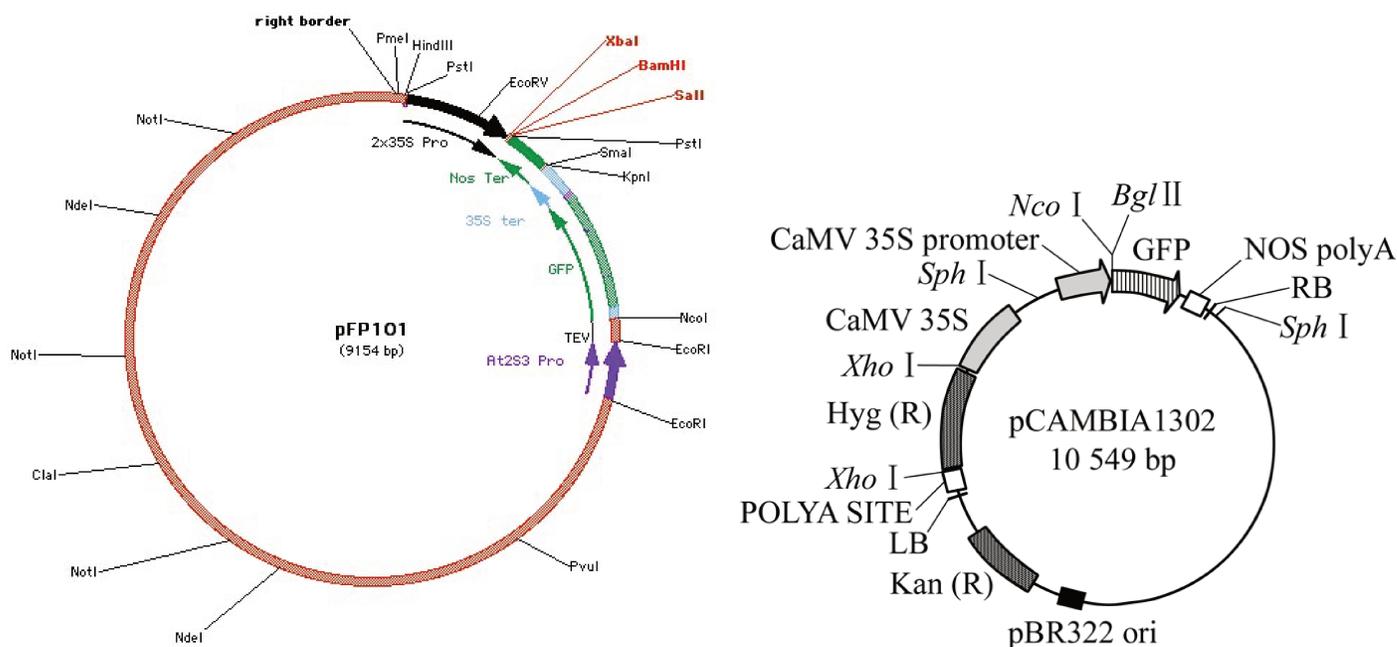


Figura 27: Vetores utilizados na clonagem: pFP101 e pCambia1302.

11.7 Transformação genética de *Arabidopsis thaliana*

As inflorescências de *A. thaliana* foram borrifadas com solução de *A. tumefaciens* carregando o vetor descrito contendo o gene selecionado (Bechtold *et al.*, 1993 com modificações de Clough & Bent (1998). Para a suspensão de *A. tumefaciens*, preparou-se uma pré-cultura em 10mL de LB líquido com os antibióticos adequados por 16 horas (28 °C / 150 rpm), a partir de cultura em placa ou glicerol. No dia seguinte, a pré-cultura foi inoculada novamente em meio LB (fator de diluição 100x) com os respectivos antibióticos, até a A₆₀₀ entre 0,8 e 1,0 ser atingida (28 °C / 150 rpm). O inóculo foi centrifugado a 3000 rpm, 4°C por 40 minutos, o sobrenadante foi descartado. O precipitado foi ressuscitado em solução de transformação (solução de sacarose 5% autoclavada - 1/3 do volume inicial da cultura, e Silwet L-77 (0,03%) foi adicionado imediatamente antes de borrifar a suspensão de agrobactéria. Este procedimento foi repetido por duas semanas.

11.8 Seleção de transformantes

O vetor pFP101 permite a seleção de sementes transformadas de *A. thaliana* pela da expressão de GFP (green fluorescent protein), pois contém um promotor específico de semente (*At2S3*) regulando a expressão desse gene. As sementes foram analisadas em microscópio de epifluorescência Olympus BX51 equipado com filtro de fluorescência azul.

Após a seleção das primeiras sementes transformadas, estas foram plantadas. As plantas transformadas foram autocruzadas até obtenção de plantas homozigotas.

12. Resultados Preliminares

12.1 Amplificação do gene com função desconhecida

O gene aqui denominado UNK875 não possui função conhecida nos bancos de dados de sequências públicas. O gene apresentou expressão significativamente diferencial em *E. grandis* em relação às outras espécies nas bibliotecas de RNAseq (Capítulo 1 e figura 28A). Entretanto, quando a sequência do mesmo foi investigada, não foi encontrado o sítio de iniciação (ATG) e nem uma ORF (*open reading frame*) para o tamanho do gene encontrado.

Deste modo, foi necessário utilizar-se a técnica RACE (*Rapid Amplification of cDNA Ends*) para obter-se a sequência completa. Após a RACE, as sequências obtidas foram montadas e confirmadas via amplificação por PCR (Figura 28B) e Real-Time PCR (Figura 28C).

A expressão esperada do gene foi confirmada via Real-Time, mas observou-se que a sequência, apesar de possuir cauda poli-A, não apresentava ORF. Assim, sugere-se aqui que esse gene possa ser um possível RNA com função regulatória (*long non-coding RNA*). Entretanto, essa classe de RNA ainda está em fase inicial de estudo (Huang *et al.*, 2012) e novos experimentos devem ser realizados para correlacionar o gene aqui encontrado com tal classe.

Este gene está em processo de clonagem e transformação em organismo modelo para avaliar o efeito de sua super-expressão.

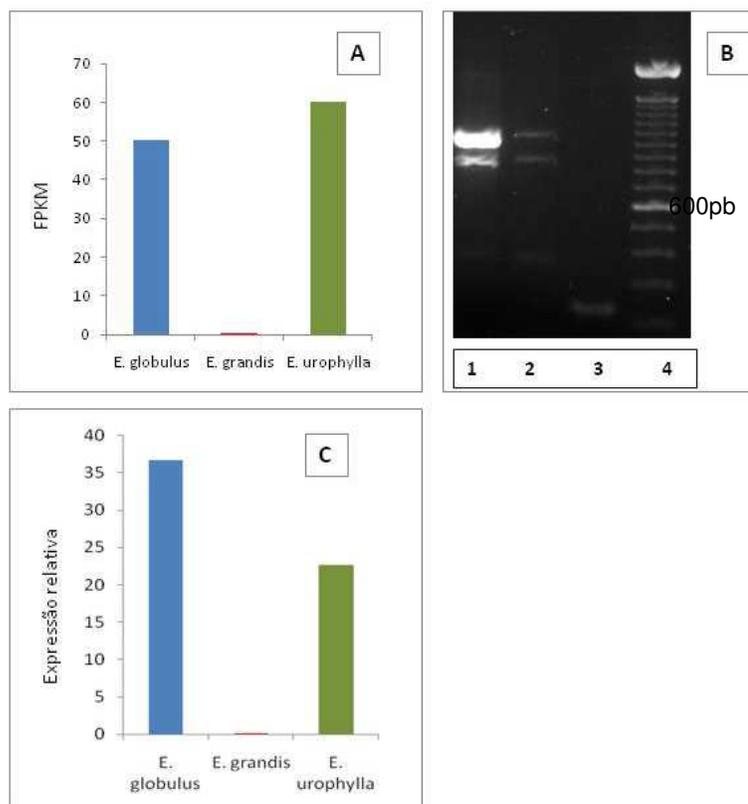


Figura 28: Expressão do gene UNK875 em bibliotecas de RNAseq (A) (Capítulo 1), Real-Time PCR (C) e PCR do cDNA (B): 1 e 2- PCR do cDNA de *E. globulus*, 3 – controle negativo, 4- Marcador de peso molecular 100bp (Invitrogen).

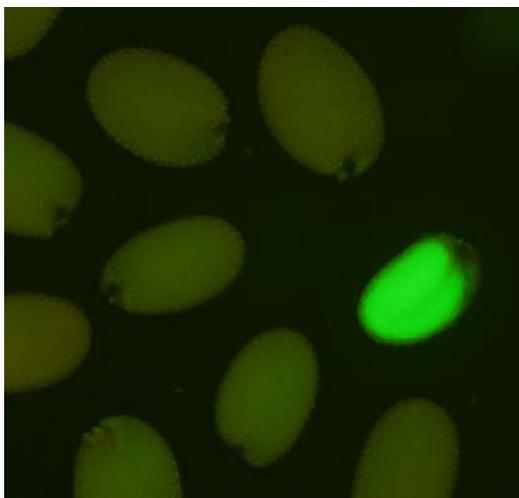
12.2 Clonagem de genes de interesse

Em *Eucalyptus*, três genes *CesA* sintetizam celulose em paredes celulares secundárias (Ranik & Myburg, 2006). Neste trabalho, apenas *CesA1* e *CesA2* foram utilizados para clonagem. A clonagem e super-expressão do gene *CesA3* foi realizada em um trabalho de iniciação científica associado a este (Processo N^o: 2009/51049-4) (ainda em fase de obtenção do homocigoto). O gene *CesA1* já foi amplificado, sequenciado porém ainda não foi clonado no vetor. Já para o gene *CesA2*, a amplificação não foi obtida. Novos primers de clonagem estão sendo desenhados.

Para os genes glutamina sintetase e *Susy3* todas as etapas de amplificação dos genes, digestão com enzimas de restrição apropriadas, ligação nos devidos vetores e transformação de *A. thaliana* já foram realizadas. Outros genes selecionados estão em diferentes etapas deste processo (Tabela 4).

Tabela 4: Resumo do andamento dos experimentos para obtenção de transformantes de *A. thaliana*.

| Gene | Clonagem pGEM | Clonagem | Transformação AGRO | Transformação <i>A. thaliana</i> | Resultados |
|-----------|---------------|--------------|--------------------|----------------------------------|--------------|
| Susy 1 | OK | OK | X | | |
| Susy 3 | OK | OK | OK | OK | Em andamento |
| CesA1 | OK | X | | | |
| FascArab | Em andamento | | | | |
| GlutSynth | OK | OK | OK | OK | Em andamento |
| Nitrato R | Em andamento | | | | |
| UNK 875 | OK | Em andamento | | | |



As sementes das plantas de *A. thaliana* transformadas com os genes citados foram coletadas e analisadas em microscópio para confirmação da transformação, conforme figura 29. As sementes transformadas já foram novamente plantadas para obtenção das plantas homocigotas.

Figura 29: Imagem de sementes de *A. thaliana* borrifadas com *Agrobacterium* contendo gene glutamina sintetase obtida com microscópio com fluorescência. Filtro azul, aumento de 8X.

13. Perspectivas

As sementes transformadas contêm o gene de interesse inserido aleatoriamente em um cromossomo. Tais plantas das sementes selecionadas já foram plantadas e submetidas à autofecundação para obtenção de um transformante homozigoto. Nas próximas etapas as sementes destas plantas contendo o gene transformado em homozigose serão novamente germinadas para avaliação do efeito fenotípico e molecular da super-expressão.

14. Referências bibliográficas

BECHTOLD, N., ELLIS, J., PELLETIER, G. In plant *Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. **Comptes Rendus de Lacademie des Sciences Serie III-Sciences de la Vie-Life Sciences**, v. 316, p. 1194-1199. 1993.

BENSMIHEN, S.; TO, A.; LAMBERT, G.; KROJ, T.; GIRAUDAT, J.; PARCY, F. Analysis of an activated *ABI5* allele using a new selection method for transgenic *Arabidopsis* seeds. **FEBS Letters**, v. 561, p. 127 – 131. 2004.

BROWN, G. R.; GILL, G. P.; KUNTZ, R. J.; LANGLEY, C. H. & NEALE, D. Nucleotide diversity and linkage disequilibrium in loblolly pine. **Proceedings of the National Academy of Science of the Unites States of America – PNAS**, v. 19 (42), p. 15255-15260. 2004.

CHAFFEY, N.; CHOLEWA, E.; REGAN, S; SUNDBERG, B. Secondary xylem development in *Arabidopsis*: a model for wood formation. **Physiologia Plantarum**, v. 114 (4), p. 594 – 600. 2002.

CLOUGH, S.J. & BENT, A.F. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. **The Plant Journal**, v. 16, p. 735-43. 1998.

COLEMAN, H. D.; YAN, J. & MANSFIELD, S. D. Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. **Proceedings of the National Academy of Science of the Unites States of America – PNAS**, v. 106 (31), p. 13118 – 13123. 2009.

GRATTAPAGLIA, D. & KIRST, M. *Eucalyptus* applied genomics: from gene sequences to breeding tools. **New Phytologist**, v. 179, p. 911 – 929. 2008.

HAMANN, T.; OSBORNE, E.; YOUNGS, H. L.; MISSON, J.; NUSSAUME, L. & SOMERVILLE, C. Global expression analysis of *CESA* AND *CSL* genes in *Arabidopsis*. **Cellulose**, v.11, p. 279-286. 2004.

HUANG, Y.; LIU, N.; WANG, J. P.; WANG, Y. Q.; YU, X. L.; WANG, Z. B.; CHENG, X. C.; ZOU, Q. Regulatory long non-coding RNA and its functions. **Journal of Physiology and Biochemistry**, 2012.

MACMILLAN, C. P.; MANSFIELD, S. D.; STACHURSKI, Z. H.; EVANS, R. & SOUTHERTON, S. G. fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in *Arabidopsis* and *Eucalyptus*. **The Plant journal**, v. 62, p. 689 – 703. 2010.

NIEMINEM, K. M.; KAUPPINEN, L.; HELARIUTTA, Y. A weed for wood? *Arabidopsis* as a genetic model for xylem development. **Plant physiology**, v. 135, p. 653 – 659. 2004.

PFAFFL, M. W. A new mathematical model for relative quantification in real-time RT-PCR. **Nucleic Acids Research**, v.29 (900), p. 2002 – 2007. 2001.

PITRE, F.E.; COOKE, J. E. K.; MACKAY, J.J. Short-term effects of nitrogen availability on wood formation and fiber properties in hybrid poplar. **Trees**, v.21, p. 249 – 259. 2007.

POKE, F. S., VAILLANCOURT, R. E., POTTS, B. M.; REID, J. B. Genomic research in Eucalpto. **Genetica**, v. 125, p. 79 - 101. 2005.

POTTS, B. M. & PEDERICK, L. A: Morphology, phylogeny, origin, distribution and genetic diversity of *Eucalyptus*. In **Diseases and pathogens of *Eucalyptus***. 1st edition. Edited by Keane PJ, Kile GA, Podger FD, Brown BN. Collingwood: CSIRO Publishing: 2000. 11- 34.

PROVOST, G. L.; HERRERA, R., PAIVA, J.A., CHAUMEIL, P., SALIN, F.; PLOMION, C. A micromethod for high throughput RNA extraction in forest trees. **Biological Research**, v.40, p. 291 – 297. 2007.

TAYLOR, N. G.; SCHEIBLE, W. CUTLER, S.; SOMERVILLE, C. R.; TURNER, S. R. The *irregular xylem3* locus of *Arabidopsis* encodes a cellulose synthase required for secondary cell wall synthesis. **The Plant Cell**, v.11, p. 769-779.1999.

ZENG, Y. & YANG, T. RNA isolation from highly viscous samples rich in polyphenols and polysaccharides. **Plant Molecular Biology Reporter**, v.20, p. 417 – 417. 2002.

DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha tese de Doutorado intitulada:

Caracterização do transcriptoma e parede celular de três espécies de Eucalyptus com importância industrial.

() não se enquadra no § 3º do Artigo 1º da Informação CCPG 01/08, referente a bioética e biossegurança.

Tem autorização da(s) seguinte(s) Comissão(ões):

(x) CIBio – Comissão Interna de Biossegurança , projeto No. 2011-03, Instituição: Instituto de Biologia (IB/UNICAMP).

() CEUA – Comissão de Ética no Uso de Animais , projeto No. _____, Instituição: _____

() CEP - Comissão de Ética em Pesquisa, protocolo No. _____, Instituição: _____

** Caso a Comissão seja externa ao IB/UNICAMP, anexar o comprovante de autorização dada ao trabalho. Se a autorização não tiver sido dada diretamente ao trabalho de tese ou dissertação, deverá ser anexado também um comprovante do vínculo do trabalho do aluno com o que constar no documento de autorização apresentado.*



Aluno: Marcela Mendes Salazar



Orientador: Gonçalo Amarante Guimarães Pereira

Para uso da Comissão ou Comitê pertinente:

Deferido () Indeferido

Carimbo e assinatura



Prof. Dr. MARCELO LANCELOTTI
Presidente da Comissão Interna de Biossegurança
Instituto de Biologia - UNICAMP

Para uso da Comissão ou Comitê pertinente:

() Deferido () Indeferido

Carimbo e assinatura