Danielle Gregório Gomes Caldas
Engenheiro Agrônomo

Análise do transcrito da região cambial de *Eucalyptus grandis* e a identificação de genes diferencialmente expressos em árvores juvenis e adultas

Orientador:
Prof. Dr. CARLOS ALBERTO LABATE

**Tese apresentada para obtenção do título de Doutor em
Agronomia. Área de Concentração: Genética e
Melhoramento de Plantas**

Piracicaba
2007

Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP

Caldas, Danielle Gregorio Gomes

Análise do transcriptoma da região cambial de *Eucalyptus grandis* e a identificação de genes diferencialmente expressos em árvores juvenis e adultas / Danielle Gregorio Gomes Caldas. -- Piracicaba, 2007.
152 p. : il.

Tese (Doutorado) -- Escola Superior de Agricultura Luiz de Queiroz, 2007.
Bibliografia.

1. Eucalipto 2. Expressão gênica -- Análise quantitativa 3. Madeira -- Qualidade 4. Melhoramento genético vegetal 5. Xilema -- Formação I. Título

CDD 634.9734

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte -- O autor"

*Dedico esta tese aos meus pais, Elias e Cecilia,
Por terem me ensinado que eu poderia ser tudo que eu desejasse,
Com humildade, ética e dignidade.
Amo Vocês!*

*Ao meu querido Maurício que com amor, carinho
e compreensão, sempre me incentivou e caminhou
de mãos dadas, ao meu lado*

OFEREÇO

AGRADECIMENTOS

Agradeço primeiramente à USP e à ESALQ por me proporcionarem uma formação profissional que dificilmente eu alcançaria em outra instituição;

Ao Prof. Dr. Carlos Alberto Labate, pela oportunidade, confiança, orientação e amizade, essenciais para o meu desenvolvimento profissional e pessoal;

Ao Dr. David Henry Moon, pela amizade, paciência, companheirismo e pelos ensinamentos que levarei por toda minha vida profissional;

Aos queridos amigos Mayra C.C.G. de Carvalho e Raphael Tozelli Carneiro, que compartilharam comigo este projeto, sempre com dedicação, alegria e amizade. Com vocês ao meu lado foi tudo muito mais leve e prazeroso!

À Livia Maria Franceschini, pela amizade, paciência e pela grande contribuição com seus conhecimentos em bioinformática;

Aos amigos de toda hora, Simone e Juliano Bragato, Mayra e Paulo André de Carvalho, Fernanda Salvato, Raphael Carneiro e Ana Letícia Bertolo, Livia e Cristiano, Eduardo de Camargo, e Karem Meireles, por todos os momentos de alegria que passamos juntos;

Aos amigos do Laboratório Max Feffer: Mônica Labate, Guilherme Salvatierra, Lilia Barbosa, Gabriela Conti, Gunta Gutmanis, Daniela Nascimento, Felipe Boaretto, Gisele Lacerda, Inês Faraldo, Paola Celedon, , Alexander de Andrade, Ednilson Rabello e Matheus Bonatto, pelo coleguismo e troca de experiências que contribuíram muito para minha formação;

Aos meus queridos irmãos Denise e Michel e meus familiares, Regina, João e Beatriz, Giuliano e Paula, pelo carinho, incentivo e trocada pelo meu sucesso;

Aos professores, funcionários e colegas do Departamento de Genética da ESALQ/USP;

À Fapesp, pela concessão de minha bolsa de estudos;

À Companhia Suzano de Papel e Celulose pelo apoio financeiro e material biológico;

Enfim, a todos que de alguma forma contribuíram para a realização deste trabalho, o meu
MUITO OBRIGADO!

SUMÁRIO

RESUMO.....	9
ABSTRACT	11
LISTA DE FIGURAS.....	13
LISTA DE TABELAS.....	15
1 INTRODUÇÃO.....	17
2 REVISÃO BIBLIOGRÁFICA.....	21
2.1 O Gênero <i>Eucalyptus</i>	21
2.2 O setor florestal brasileiro e a importância do eucalipto.....	22
2.3 Melhoramento genético e biotecnologia.....	24
2.4 Caracteres de qualidade da madeira.....	27
2.4.1 Madeira Juvenil vs Madeira Adulta.....	28
2.5 A Formação da Madeira.....	31
2.6 O estudo da xilogênese na era das “Ômicas”	35
2.6.1 A técnica SAGE (Análise Serial da Expressão Gênica).....	39
3 MATERIAL E MÉTODOS.....	43
3.1 Material vegetal.....	43
3.2 Extração de RNA total.....	44
3.3 Construção da biblioteca SAGE.....	46
3.3.1 Isolamento do RNA mensageiro e síntese de cDNA.....	46
3.3.2 Digestão do cDNA com a endonuclease <i>Nla</i> III e ligação dos adaptadores.....	48
3.3.3 Digestão com a endonuclease <i>Bsm</i> F I e formação dos dímeros de cDNA.....	49
3.3.4 Otimização das reações de PCR para produção de dímeros em larga escala.....	51
3.3.5 Amplificação em larga escala e purificação dos dímeros de cDNA.....	52
3.3.6 Digestão dos dímeros de cDNA com <i>Nla</i> III.....	53
3.3.7 Polimerização/Concatenação e clonagem das <i>ditags</i>	55
3.4 Sequenciamento dos polímeros/concatenados de <i>ditags</i>	57
3.5 Bioinformática e análise dos dados.....	58
3.5.1 Análise da Qualidade do Sequenciamento.....	58
3.5.2 Extração das SAGE <i>tags</i>	58

3.5.3 Identificação das SAGE <i>tags</i>	60
3.6 RT-PCR em Tempo Real (qRT-PCR).....	64
3.6.1 Seleção de genes e desenho de <i>primers</i>	65
3.6.2 Isolamento de RNAm e Síntese da primeira fita de cDNA.....	67
3.6.3 PCR em Tempo Real.....	68
4 RESULTADOS E DISCUSSAO.....	69
4.1 Análise dos dados gerados pelo SAGE.....	69
4.1.1 Dados Gerais.....	69
4.1.2 Caracterização dos genes por níveis de expressão.....	74
4.1.3 Identificação de transcritos anti-sense.....	77
4.1.4 Identificação de transcritos alternativos.....	80
4.1.5 Perfil transcricional global da região cambial de <i>Eucalyptus grandis</i>	83
4.2 Perfil transcricional durante a Xilogênese.....	87
4.2.1 Divisão Celular.....	87
4.2.2 Expansão Celular.....	92
4.2.3 Síntese de Polissacarídeos e Deposição da Parede Celular Secundária.....	96
4.2.4 Lignificação.....	102
4.2.5 Morte Celular Programada.....	109
4.3 Validação da Biblioteca SAGE por qRT-PCR.....	110
4.4 Análise comparativa da expressão gênica entre árvores juvenis e adultas.....	113
5 CONSIDERAÇÕES FINAIS.....	121
REFERÊNCIAS.....	123
ANEXOS.....	141

RESUMO

Análise do transcrito da região cambial de *Eucalyptus grandis* e a identificação de genes diferencialmente expressos em árvores juvenis e adultas

O Brasil é hoje o maior produtor mundial de papel e celulose à partir de fibra curta de *Eucalyptus*, com um setor florestal bem desenvolvido e estabelecido, que gera muitos empregos e divisas ao país. A madeira utilizada pelas indústrias brasileiras é exclusivamente obtida de florestas plantadas, sendo que 1,27 milhões de hectares correspondem a plantações de *Eucalyptus*. Durante muitos anos, os programas de melhoramento genético exploraram as características de altas taxas de crescimento, produção de biomassa e adaptabilidade deste gênero para alcançar uma alta produtividade, que hoje é de 45 m³/ha/ano. Entretanto, a demanda do mercado nacional e internacional por madeira de *Eucalyptus* é crescente. Para manter a competitividade do setor, torna-se necessária a busca por uma madeira de melhor qualidade, que leve a uma maior eficiência na produção de papel e celulose. O conhecimento dos processos que levam à formação da madeira, um tecido que se diferencia a partir do câmbio vascular, certamente irá contribuir para que essa qualidade superior da madeira seja alcançada. As árvores só apresentam as características desejadas pela indústria quando adultas, entretanto, o *Eucalyptus* é cortado ainda em fase juvenil de desenvolvimento. Portanto, o objetivo deste trabalho foi traçar um perfil transcricional da região cambial de *Eucalyptus grandis* em fase juvenil de desenvolvimento e, com as informações geradas, comparar a expressão de alguns genes envolvidos na formação da madeira entre a região cambial de árvores juvenis e adultas. A técnica SAGE foi utilizada para construção de uma biblioteca que gerou 22660 *tags*, das quais 2942 representavam *tags* únicas com número de cópias igual ou maior que 2. Dessas, 874 puderam ser associadas a uma provável função e classificadas em categorias funcionais. O SAGE permitiu o acesso a genes com baixos níveis de expressão, a identificação de potenciais transcritos anti-sense e alternativos, e revelou a presença de muitas sequências ainda não caracterizadas. Os dados obtidos foram organizados, também, nas principais vias responsáveis por cada fase da diferenciação que ocorre na xilogênese. Os genes que codificam as principais enzimas responsáveis pela expansão celular, deposição de parede celular secundária e lignificação foram encontrados, e seus níveis de expressão indicaram as prováveis vias metabólicas utilizadas pelas células da região cambial de *E. grandis* para produzir madeira juvenil. Os dados de expressão gênica obtidos com o SAGE foram validados através de RT-PCR em tempo real. Vinte e cinco genes envolvidos com a expansão celular, biossíntese de celulose e hemicelulose, deposição de parede secundária, lignificação e metabolismo foram selecionados para uma comparação entre as regiões cambiais de árvores juvenis e adultas de *E. grandis* através de RT-PCR em tempo real. Os resultados mostraram que existe uma expressão diferencial entre idades, sendo que 15 genes foram mais expressos nas árvores juvenis e 6 mais expressos nas árvores adultas. A caracterização quantitativa dos genes expressos na região cambial de *E. grandis* constitui, portanto, uma grande fonte de informação que pode direcionar para diferentes enfoques relacionados ao complexo processo de biossíntese da madeira.

Palavras-chave: *Eucalyptus grandis*; Xilogênese; SAGE; qRT-PCR; Qualidade da madeira; Madeira juvenil vs Madeira adulta

ABSTRACT

Analysis of the transcriptome of the *Eucalyptus grandis* cambial region and the identification of differentially expressed genes in juvenile and mature trees

Currently, Brazil is the largest producer of pulp and paper using short fiber from *Eucalyptus*, with a well developed and established forest sector, that generates many jobs and income for the country. The wood for the Brazilian industries is taken exclusively from planted forests, where 1.27 million hectares correspond to *Eucalyptus* plantations. During many years, genetic improvement programs have explored the high growth rates, biomass production and adaptability of this genus to reach high productivity, approximately 45 m³/ha/year. However, national and international demand for *Eucalyptus* wood is increasing and to ensure the competitiveness of this sector, it is necessary to search for a better quality wood, which allows more efficient pulp and paper production. The knowledge of the processes that lead to the wood formation, a tissue that differentiates from the vascular cambium, will certainly contribute to reach this goal. Trees normally acquire the characteristics desired by industry when adult, but *Eucalyptus* is cut when it is still in juvenile phase of development. Therefore, the objective of this work was to produce a transcriptional profile of the cambial region of *Eucalyptus grandis* in the juvenile phase of development and, with this information, compare the expression of some genes involved in wood formation in the cambial region of juvenile and mature trees. A SAGE library was produced generating 22660 tags, of which 2942 represented unique tags with two or more copies. Of these, 874 could be associated with a probable function and were classified in functional categories. SAGE allowed the access to genes with low expression levels, the identification of potential anti-sense and alternative transcripts, and disclosed the presence of many sequences not yet characterized. The data were also organized to demonstrate the main pathways involved at each differentiation stage during xylogenesis. The genes for the major enzymes involved in cell wall expansion, secondary cell wall deposition and lignification were found and their expression levels indicated the preferential use of alternative branches in some of the metabolic pathways used by the cambial cells of *E. grandis* to produce juvenile wood. The expression data obtained with SAGE was validated through Real Time RT-PCR. Twenty five genes involved with cell wall expansion, cellulose and hemicellulose biosynthesis, secondary wall deposition, lignification and metabolism were selected to compare the cambial regions of juvenile and mature *E. grandis* trees using Real Time RT-PCR. The results showed the existence of a differential expression pattern between the ages, were 15 genes were more expressed in the juvenile trees and 6 genes were more expressed in the mature trees. The quantitative characterization of the genes expressed in the cambial region of *E. grandis* constitutes, therefore, a great information source that might drive the studies to different approaches related to the complex process of wood biosynthesis.

Keywords: *Eucalyptus grandis*; Xilogenesis; SAGE; qRT-PCR; Wood quality; Juvenile wood vs Mature wood

LISTA DE FIGURAS

Figura 1 - Localização da madeira juvenil e adulta na árvore e suas principais características.....	29
Figura 2 - Estrutura tridimensional da parede celular secundária de uma célula do xilema.....	33
Figura 3 - Amostragem do material biológico.....	43
Figura 4 - Géis de agarose com bandas de RNA e do produto de RT-PCR.....	45
Figura 5 - Esquema ilustrativo da metodologia SAGE (Análise Serial da Expressão Gênica).....	48
Figura 6 - Gel de eletroforese dos produtos das reações de otimização da concentração de DNA para PCR.....	52
Figura 7 - Gel de eletroforese dos produtos da amplificação em larga escala dos dímeros de cDNA.....	53
Figura 8 - Gel de eletroforese dos produtos da digestão com <i>Nla</i> III.....	54
Figura 9 - Gel de eletroforese dos produtos da ligação das <i>ditags</i>	55
Figura 10 - Gel de agarose 1,5% para verificação dos insertos clonados.....	57
Figura 11 - Cromatogramas de dois eventos de sequenciamento de um mesmo clone.....	59
Figura 12 - Arquivo de saída do <i>software</i> SAGE2000 TM (versão 4.5) ou <i>Tag list</i> da biblioteca da região cambial de <i>E. grandis</i>	59
Figura 13 - Sequência de um EST utilizado na anotação de uma <i>tag</i>	61

Figura 14 - Ficha “eletrônica” do banco de dados local de genômica funcional de <i>Eucalyptus grandis</i>	64
Figura 15 - Classificação funcional das 2942 tags únicas (genes) da biblioteca da região cambial de árvores de <i>Eucalyptus grandis</i>	72
Figura 16 - Distribuição dos genes de acordo com sua categoria funcional em três níveis de expressão.....	75
Figura 17 - Origem de transcritos anti-sense.....	78
Figura 18 - Esquema do ciclo celular e o processo de indução da divisão através da auxina.....	90
Figura 19 - Ilustração da parede celular primária com seus componentes estruturais e enzimas de modificação.....	93
Figura 20 - Reações de interconverções de açúcares-NDP, síntese de celulose, hemiceluloses e pectinas.....	97
Figura 21 - Ilustração da síntese de microfibrilas de celulose através do complexo protéico presente na membrana celular.....	101
Figura 22 - Síntese e polimerização dos monolignóis.....	104
Figura 23 - Modelo para PCD em TEs proposto por Groover e Jones (1999).....	110
Figura 24 - Curva padrão obtida a partir das leituras de CT de três diluições e três repetições dos produtos de amplificação do gene Isocitrato desidrogenase.....	112
Figura 25 - Quantificação relativa dos níveis de expressão de 25 genes por qRT-PCR entre árvores jovens e adultas.....	118

LISTA DE TABELAS

Tabela 1 - Valores médios do comprimento, largura e espessura da parede das fibras de <i>E. grandis</i> no sentido medula-casca.....	30
Tabela 2 - Otimização das reações de PCR a partir de diferentes concentrações da amostra.....	51
Tabela 3 - Reação de ligação do vetor aos polímeros de <i>ditags</i>	56
Tabela 4 - Classificação Funcional de Rison et al. (2000).....	63
Tabela 5 - <i>Primers</i> específicos para os genes usados na validação da biblioteca SAGE por qRT-PCR.....	65
Tabela 6 - <i>Primers</i> específicos para os genes usados na comparação entre árvores juvenis e adultas por qRT-PCR.....	66
Tabela 7 - Dados gerais obtidos pela biblioteca SAGE da região cambial de <i>E. grandis</i>	69
Tabela 8 - Relação das 50 <i>tags</i> mais expressas e suas prováveis identidades.....	71
Tabela 9 - <i>Tags</i> associadas a transcritos anti-sense verdadeiros identificadas de forma direta, como descrito no texto.....	79
Tabela 10 - <i>Tags</i> associadas a potenciais transcritos anti-sense identificadas de forma indireta, como descrito no texto.....	80
Tabela 11 - Relação das 10 <i>tags</i> mais expressas da biblioteca SAGE associadas a potenciais transcritos alternativos.....	82

Tabela 12 - Perfil transcricional da região cambial de <i>E. grandis</i> baseada em Categorias e Subcategorias Funcionais.....	84
Tabela 13 - Genes envolvidos com respostas a fitormônios encontrados na biblioteca de madeira juvenil, suas SAGE <i>tags</i> e seus números de cópias.....	88
Tabela 14 - Valores de Eficiência de Amplificação (<i>E</i>) para os <i>primers</i> utilizados no experimento de validação dos dados do SAGE.....	111
Tabela 15 - Análise de correlação entre as frequências das SAGE <i>tags</i> (TPT) e os valores de expressão <i>V</i> obtidos através de qRT-PCR utilizando valores absolutos e normalizados com o gene UGPase.....	112
Tabela 16 - Valores de Eficiência de Amplificação (<i>E</i>) e CTs médios obtidos para 25 genes em amostras coletadas da região cambial de <i>E. grandis</i> nas fases juvenil e adulta de desenvolvimento.....	115

1 INTRODUÇÃO

No último século, o gênero *Eucalyptus* tornou-se uma poderosa fonte de matéria-prima para uma diversidade de usos, entre eles madeira para a indústria moveleira, construção civil, geração de energia (através da queima do carvão vegetal), para indústria de papel e celulose, indústrias químicas e farmacêuticas e, mais recentemente, para a produção de biocombustível.

O Brasil é hoje o maior produtor mundial de celulose de fibra curta a partir do *Eucalyptus*, atividade que contribui com cerca 4% do PIB. O estabelecimento dessa espécie foi promovido devido às suas altas taxas de crescimento, que levam a um alto rendimento de biomassa e a sua adaptabilidade a diversos climas e solos, com diferentes espécies sendo plantadas em várias regiões do país. Além disso, as grandes áreas plantadas com florestas de *Eucalyptus* contribuem para conservação do meio ambiente e já estão sendo utilizadas na comercialização de créditos de carbono, gerando uma fonte adicional de lucro ao setor florestal.

Inicialmente, os programas de melhoramento aplicados à produção de papel e celulose buscavam o aumento da produtividade volumétrica das árvores e resistência a fatores bióticos e abióticos. Hoje, com a grande competitividade entre as empresas e com o mercado externo, a palavra-chave para manutenção dessa vantagem competitiva é a “qualidade” da madeira. Portanto, o que se busca hoje é uma madeira que, além de possibilitar um maior rendimento de pasta celulósica, apresente características físico-químico-mecânicas que favoreçam a extração de celulose e diminuam custos de produção, aumentando a qualidade do papel produzido.

A madeira é formada pelo xilema secundário, composto por vasos, fibras e células do parênquima, e suas propriedades físico-químicas são derivadas diretamente da composição e morfologia das paredes celulares presentes em seus componentes estruturais. Todavia, apesar de seu grande valor econômico, muito pouco é conhecido sobre os processos celulares e moleculares envolvidos na formação da madeira, uma vez que a diferenciação celular na xilogênese é muito complexa.

Durante a biogênese da madeira, as células do câmbio vascular se dividem, dando origem a células iniciais de vasos, fibras e parênquima, que sofrem expansão e alongação antes da deposição da parede celular secundária e lignificação; a diferenciação das células de vasos e fibras é finalizada com a morte celular programada. O xilema, portanto é formado por células mortas compostas pela parede celular que as envolvia, cuja composição depende dos processos

de biossíntese e modificações, dirigidas por genes específicos que controlam a síntese macromolecular e determinam a sua estrutura, além de uma série de fatores endógenos e exógenos que podem levar a alterações na expressão desses genes.

Dessa forma, o conhecimento dos genes envolvidos nos diferentes passos que levam à formação da madeira pode indicar pontos específicos do processo que contribuem para o aumento da qualidade da madeira. O desenvolvimento dos projetos de sequenciamento de genomas abriu um grande leque para a identificação de novos genes e caracterização de perfis transcricionais. A genômica funcional tornou possível a análise da expressão de centenas, e em alguns casos, de milhares de genes simultaneamente, permitindo o estudo de populações completas de RNAm em células ou tecidos e indicando o perfil da transcrição global.

Recentemente, a genômica funcional da xilogênese começou a ser estudada em espécies arbóreas (LORENZ e DEAN, 2002; PAUX et al., 2004; FOUCART et al., 2006), fornecendo informações mais precisas quanto aos mecanismos responsáveis principalmente pela deposição da parede celular secundária e lignificação. Muitos fatores de transcrição também têm sido identificados, entre eles MYB e LIM, envolvidos na regulação da lignificação (PAUX et al., 2005).

Entre as técnicas quantitativas aplicadas na análise da expressão gênica está o SAGE (Análise Serial da expressão Gênica), que permite uma caracterização quantitativa altamente acurada do transcrito de uma espécie. Esta técnica baseia-se na construção de uma biblioteca de *tags* (etiquetas) compostas por 10 pb que, teoricamente, representam o conjunto de RNAm presentes num tecido num dado momento e, por gerar uma grande quantidade de informação sem a necessidade de um conhecimento prévio dos genes a serem estudados, foi escolhida para traçar o perfil transcricional da região cambial de árvores de *Eucalyptus grandis* em fase juvenil de desenvolvimento.

Este trabalho faz parte de um projeto maior, desenvolvido pelo Laboratório Max Feffer de Genética de Plantas, cujo objetivo é conhecer os processos moleculares que governam o metabolismo do *Eucalyptus* e a formação da sua madeira, o qual englobou estudos de transcrito e proteoma da região cambial de *Eucalyptus grandis* em três idades diferentes de seu desenvolvimento até a sua idade de corte; o transcrito de folhas resistentes e suscetíveis à ferrugem; a comparação dos perfis transcricionais entre folhas e a região cambial; e mais,

recentemente, vem estudando os perfis protéicos da madeira de tração e de raízes submetidas a estresse hídrico.

Com a obtenção de uma visão global da expressão gênica durante a formação da madeira juvenil, a qual é atualmente utilizada pela indústria de papel e celulose, foi estabelecida uma grande plataforma de conhecimentos. Diferentemente da madeira juvenil, a madeira adulta, que é produzida pelas árvores após muitos anos de desenvolvimento, apresenta uma qualidade muito superior para a produção de papel e polpa celulósica. Por outro lado, o corte das árvores de *Eucalyptus* em idades mais avançadas é financeiramente inviável para a indústria, já que aos sete anos de idade, essas árvores já atingiram um grande volume de biomassa. Com base nestas informações e com o estabelecimento do perfil transcricional da xilogênese em *Eucalyptus*, surgiu a necessidade de investigar se os genes envolvidos neste processo comportam-se diferencialmente entre as regiões cambiais de árvores juvenis e adultas. A descoberta de genes que regulam a formação das características presentes na madeira adulta seria de grande valor, principalmente se suas expressões pudessem ser induzidas e/ou selecionadas em árvores em fase juvenil de desenvolvimento, aliando as altas taxas de crescimento à alta qualidade de madeira.

Dessa forma, este trabalho teve como objetivos: a) traçar o perfil transcricional da região cambial de *Eucalyptus grandis* em fase juvenil de desenvolvimento através da técnica SAGE; b) utilizar o potencial desta técnica para detecção de genes ainda não descritos, identificação de transcritos com baixos níveis de expressão, identificação de transcritos anti-sense e identificação de transcritos alternativos; c) comparar a expressão de 25 genes envolvidos na formação da madeira entre tecidos da região cambial de árvores juvenis e adultas.

2 REVISÃO BIBLIOGRÁFICA

2.1 O Gênero *Eucalyptus*

A primeira espécie deste gênero reconhecida oficialmente foi o *Eucalyptus obliqua*, com a qual se definiu, por volta de 1788, o gênero *Eucalyptus* (L'Herit) e que, devido às suas características, foi incluído na família *Myrtaceae* (COZZO, 1955). Suas plantas jovens apresentam folhas opostas e geralmente sésseis e, quando adultas, alternas e elípticas; suas flores ocorrem em inflorescências axilares do tipo panícula ou umbela e seus frutos são do tipo cápsulas lenhosas com opérculo apical (BONA et al., 2004). O nome deste gênero vem do grego *Eucalyptus*, que significa “bem coberto” devido, justamente, à presença do opérculo de seus frutos, que protege suas sementes (COZZO, 1955).

As espécies de *Eucalyptus* são monóicas, protândricas e preferencialmente alógamas (CARR et al., 1971), são diplóides e, em sua maioria, possuem 22 cromossomos caracterizados por um tamanho extremamente pequeno (RUGGERI, 1961), com um conteúdo de DNA nuclear (2c) estimado entre 500 e 650 Mpb (GRATTAPAGLIA & BRADSHAW, 1994).

O eucalipto é originário da Austrália onde cobre quase 90% de sua área territorial, formando densos maciços florestais nativos, com algumas poucas espécies nativas das ilhas de Java, Filipinas, Timor, Papua Nova Guiné e outras (PRYOR, 1976). Engloba cerca de 700 espécies, sub-espécies, variedades e híbridos naturais, os quais são encontrados desde a forma arbustiva até árvores de grande porte (CRESSWELL et al., 1975). Atualmente, encontra-se amplamente distribuído em países tropicais e sub-tropicais entre latitudes de 9°N e 45°S e altitudes que variam de 30 a 1000 metros, adaptados a diversas condições de clima e solo (BROOKER, 2000).

As primeiras sementes do gênero, coletadas na Austrália, foram levadas para a Europa por volta de 1774 apenas como curiosidade botânica e suas plantas permaneceram em coleções de alguns jardins botânicos, sem nenhuma importância comercial até a metade do século XIX (ANDRADE, 1961). Na Índia, os primeiros plantios ocorreram em 1843 e, por volta de 1856 já haviam plantações de eucaliptos bem desenvolvidas. Na África do Sul, os primeiros ensaios com eucaliptos ocorreram em 1828, na colônia do Cabo. Os primeiros ensaios na Europa, visando à produção comercial datam de 1854, principalmente com o *Eucalyptus globulus*, plantados em

Portugal pelo Barão de Marcellos e por J. M. Eugênio de Almeida. Em 1863 foi introduzido na Espanha e, em 1869, na Itália (O EUCALIPTO, 2001).

Na América do Sul, o Chile foi o primeiro país a introduzir o eucalipto em 1823 seguido pela Argentina, Uruguai e Brasil. No Brasil, os primeiros plantios de eucalipto ocorreram por volta de 1868, no Rio Grande do Sul, por Frederico de Albuquerque e, no Rio de Janeiro, por Pereira da Cunha, com finalidades ornamentais. Porém, o reconhecimento pela introdução econômica da cultura foi creditado a Navarro de Andrade, que iniciou por volta de 1904 pesquisas comparativas entre espécies de eucaliptos e espécies florestais nativas para a Companhia Paulista de Estradas de Ferro (GRUBER, 2006). As primeiras mudas foram produzidas de uma coleção de sementes trazidas de Portugal e outras coletadas no estado de São Paulo, em árvores que serviam de abrigo, quebra-vento e ornamentação. Os resultados iniciais aconselharam a utilização do eucalipto na produção de lenha, dormentes, postes e moirões, devido ao seu rápido crescimento e por apresentar grande resistência às intempéries do ambiente. Por apresentar tantas possibilidades de uso, os plantios foram incorporados gradativamente ao cenário florestal brasileiro (ANDRADE, 1961). De acordo com a Companhia Paulista de Estradas de Ferro (2005), entre 1904 e 1919, Navarro de Andrade introduziu um total de 144 espécies de eucalipto. A partir da década de 50, o eucalipto passou a ser produzido como matéria prima para o abastecimento de fábricas de papel e celulose (TOMASELLI, 2004).

2.2 O setor florestal brasileiro e a importância do eucalipto

A grande variedade de espécies de eucalipto introduzidas no Brasil no início do século XX transformou o Horto Florestal de Rio Claro, em São Paulo, no mais completo banco de germoplasma de eucalipto no Brasil (ANDRADE, 1961) e, rapidamente, estas áreas se tornaram grandes pólos de fornecimento de sementes para os programas de formação de florestas incentivados pelo governo federal brasileiro entre os anos de 1960 e 1970.

Nesta época, os trabalhos florestais se dinamizaram devido ao estímulo concedido pelo governo Castelo Branco, que liberou o investimento em reflorestamentos para pessoas físicas e jurídicas, no valor de até 50% do imposto de renda devido. O setor florestal ganhou uma poderosa fonte para estimular o plantio em larga escala e proporcionar a geração de recursos financeiros suficientes para financiar a expansão da indústria de base florestal (PEREIRA, 2002).

No final da década de 70 e durante os anos 80, foram instalados os primeiros testes de progênes e pomares de sementes para atender a demanda crescente de sementes.

O plantio de eucalipto no Brasil foi consolidado graças ao seu excelente potencial para produção de madeira e por apresentar uma série de características favoráveis, entre elas: alta adaptabilidade a diferentes condições de clima e solo; elevada produção de sementes e facilidade de propagação vegetativa; crescimento rápido, alta produtividade e boa forma do fuste; boa resposta a tratos culturais de manejo e melhoramento genético; e adequação aos mais diferentes usos e aceitação de mercado (MORA e GARCIA, 2000; SILVA, 2005).

Segundo levantamentos da FAO, mais de noventa países possuem programas de reflorestamento com eucalipto, atingindo uma área total equivalente a 18 milhões de hectares. Desses, 3,4 milhões de hectares estão no Brasil (SBS, 2006), onde o eucalipto é cultivado em escala comercial, com produtividade média de 45 m³/ha.ano, podendo atingir 60 m³/ha.ano em plantios clonais estabelecidos em sítios superiores. Entre as espécies mais utilizadas estão *Eucalyptus grandis* (55%), *Eucalyptus saligna* (17%), *Eucalyptus urophylla* (9%), *Eucalyptus viminalis* (2%), híbridos de *E. grandis* x *E. urophylla* (11%) e outras espécies (6%) (O EUCALIPTO, 2001). Os estados com os maiores plantios de eucalipto são Minas Gerais, com produção direcionada ao atendimento da indústria siderúrgica e de carvão vegetal, e São Paulo, Bahia e Espírito Santo, com suas produções utilizadas principalmente nos pátios das indústrias de celulose e papel. Seu ciclo de rotação é de sete anos, quase três vezes menor que o de *Pinus* e mais de oito vezes inferior ao de outras coníferas típicas de países de clima temperado (revisado por MEIRELES, 2006).

A madeira do eucalipto também pode ser utilizada por outros segmentos como matéria prima na produção de uma variedade de produtos, como móveis, madeira serrada, madeira laminada, madeira processada para produção de aglomerados e chapas de fibras e, devido à sua composição química, pode ser usada também pela indústria petroquímica na produção de plástico e outros produtos químicos.

O setor florestal brasileiro gera cerca de 3 milhões de empregos diretos e indiretos e apresentou, em 2005, um faturamento de US\$ 27,8 bilhões (3,5% do PIB nacional), contribuindo com US\$ 4,2 bilhões em impostos. As exportações brasileiras ligadas ao setor atingiram US\$ 7,4 bilhões, correspondendo a aproximadamente 7% do total exportado pelo país (SBS, 2006; LEITE, 2005). Hoje, o Brasil ocupa a sétima posição mundial na produção de papel e celulose e é

o 11º maior fabricante de papel, além de ser o maior produtor de celulose de fibra curta de eucalipto (SBS, 2006).

A expectativa para 2007 é que a produção brasileira atinja 11,75 milhões de toneladas de celulose e 9 milhões de toneladas de papel, obtendo um crescimento de 5,9% e 2,9%, respectivamente, em relação à produção registrada em 2005, o que levará o Brasil a passar da sétima para a sexta posição como produtor mundial de celulose (BRACELPA, 2006).

De acordo com pesquisas recentes, o consumo médio mundial de madeira deve aumentar a uma taxa de 1,2% a 3,4% ao ano nos próximos cinco anos (PERSPECTIVAS, 2003) e, para suprir essa demanda, o setor de papel e celulose apresentou ao governo um programa nacional de investimentos da ordem de US\$ 14,4 bilhões para o período de 2003 a 2012, destinados a expansão da capacidade produtiva e da competitividade das indústrias brasileiras no cenário mundial (BRACELPA, 2006).

Além da importância das áreas reflorestadas de eucalipto para a indústria brasileira, atualmente existe uma destacada atenção voltada para sua contribuição na minimização do efeito estufa, causado principalmente pela emissão de dióxido de carbono e metano na atmosfera. Uma forma encontrada de reverter ou atenuar as previsões de superaquecimento global foi a criação do Protocolo de Kyoto, segundo o qual os países altamente industrializados devem diminuir seus níveis de emissão de gases poluentes na atmosfera. Caso suas metas não sejam atingidas, estes países poderão comprar créditos de carbono de países que não tenham essa obrigatoriedade, ao qual o Brasil se enquadra. Portanto, o país deve se colocar como um grande vendedor de créditos de carbono, no qual espera-se que o faturamento decorrente da venda desses créditos alcance US\$ 13 bilhões em 2007 (BRASIL, 2005).

2.3 Melhoramento genético e biotecnologia

Inicialmente, os programas de melhoramento genético do eucalipto tiveram como objetivo avaliar a capacidade de adaptação de suas espécies em ambientes específicos através de ensaios de progênies em diferentes sítios florestais e assim, estudar a variação da resposta de um genótipo a diferentes condições impostas pelo ambiente, a chamada interação genótipo x ambiente (RESENDE, 1990). Num passo seguinte, os programas de melhoramento passaram a visar o aumento da produtividade volumétrica da madeira (MENK et al., 1989) e, como consequência, os plantios comerciais de eucalipto passaram de 12 m³/ha.ano para uma produtividade média de 45

m³/ha.ano (LEITE, 2005). Entre as estratégias adotadas para o melhoramento do eucalipto estão a seleção massal, obtenção de clones das árvores superiores e seleção de famílias de meio-irmãos para obtenção de populações melhoradas (PEREIRA et al., 2002).

Com o fim do subsídio fiscal concedido pelo governo brasileiro, a partir de 1986 o setor florestal e o melhoramento genético do eucalipto entraram numa nova fase, onde a eficiência nos plantios florestais era a palavra chave. Com isso, as pesquisas básicas de melhoramento foram abandonadas e o enfoque dos produtores passou a ser a propagação de árvores superiores, especialmente de híbridos como, por exemplo, o *E. urograndis* (*E. grandis* x *E. urophylla*) (FERREIRA, 1983).

Embora a obtenção de híbridos já ocorresse desde a década de 70, ele era restrito ao uso da polinização aberta entre espécies compatíveis, que deveriam florescer ao mesmo tempo. Dessa forma, um método de cruzamento controlado que explorava a protandria do eucalipto foi desenvolvido nos anos 80 e levou a uma expansão de híbridos produzidos artificialmente no Brasil (ASSIS, 1985; BRIGATTI et al., 1983; MARTINS e IKEMORI, 1987). Hoje, são usadas técnicas que transformam a protandria numa proginia (AIP - *Artificially Induced Protogyny*) e permitem a realização de cruzamentos em larga escala, possibilitando a multiplicação de famílias superiores de meio-irmãos (ASSIS e MAFIA, 2007). Assim, os programas de melhoramento exploram a variabilidade entre as espécies de *Eucalyptus* adaptadas ao Brasil, que possuem diferentes propriedades de madeira, tolerância a estresses bióticos e abióticos, bem como manifestação de heterose, para alcançar uma complementariedade entre as espécies, aumentando o número de características desejáveis para a produção de indivíduos superiores através da hibridação.

Com o emprego da clonagem na produção de mudas de materiais genéticos superiores foi possível obter ganhos de produtividade pelo uso da herdabilidade no sentido amplo, maximizando os resultados em produtividade volumétrica, em propriedade tecnológica da madeira e em resistência a fatores bióticos e abióticos. A clonagem também possibilitou a produção de matéria prima mais uniforme que, do ponto de vista industrial, representa significativos benefícios, tanto para redução dos custos dos processos industriais, quanto para o aumento da qualidade dos produtos. A microestaquia e a miniestaquia são as técnicas mais modernas para a clonagem ou micropropagação vegetativa do eucalipto em larga escala, e são atualmente, as mais usadas no Brasil (ODA et al., 2007).

Mesmo com toda a contribuição do melhoramento genético do eucalipto para a formação de florestas altamente produtivas, ainda existem características inerentes à biologia das árvores que dificultam esse processo. Entre elas estão: longo ciclo de vida, porte, maturação tardia, polinização cruzada, elevada heterozigosidade, alta complexidade da análise dos descendentes após cruzamentos e retrocruzamentos, dificuldade de conseguir ganhos significativos para caracteres complexos e, necessidade de grandes áreas para plantios experimentais (VALVERDE et al., 2007; CAMPBELL et al., 2003). Assim, a biotecnologia florestal surgiu como uma ferramenta para alcançar progressos de forma rápida, manipular e estudar caracteres até então desconhecidos ou que não podiam ser manipulados através do emprego das técnicas de melhoramento convencional.

Incrementos ainda maiores na produtividade do setor florestal podem ser alcançados com o emprego de métodos biotecnológicos auxiliares aos programas de melhoramento que facilitem a seleção e produção de genótipos superiores. Neste sentido, a genética molecular pode oferecer uma importante alternativa aos problemas comuns do melhoramento clássico de espécies arbóreas através da seleção assistida, que usa marcadores moleculares para alelos relacionados aos genes alvo, e que por serem herdáveis, permitem a seleção precoce de genótipos superiores; e através da transformação genética, que possibilita alterar características genéticas em uma única geração através, por exemplo, da introdução de alelos dominantes (CAMPBELL et al., 2003).

Os marcadores genéticos vêm sendo utilizados em espécies florestais desde os anos 80, quando levou à construção das primeiras versões de mapas genéticos de coníferas baseados em isoenzimas. Os mapas genéticos permitem a localização de genes e regiões genômicas envolvidos no controle de características quantitativas (QTL) e têm sido um recurso muito utilizado pelos programas de melhoramento uma vez que possibilitam a seleção precoce de caracteres identificados somente na fase adulta, redução do tempo para obtenção de indivíduos superiores, monitoramento da herança dos alelos favoráveis, caracterização de populações de melhoramento, discriminação e verificação de clones elite e, ainda, a determinação de paternidade em populações de melhoramento (FERREIRA e GRATTAPAGLIA, 1995). Espécies florestais como *Pinus*, *Populus* e *Eucalyptus* já têm mapas genéticos construídos através da combinação de marcadores dominantes RAPD e AFLP (RITTER et al., 1990; GRATTAPAGLIA e SEDEROFF, 1994), e que foram complementados com informações obtidas a partir de marcadores co-dominantes RFLP (BYRNE et al., 1995), marcadores microssatélites (BRONDANI et al., 1998),

os quais possibilitaram a revelação de uma quantidade muito maior de polimorfismo, permitindo uma maior eficiência na obtenção e validação de QTLs (BRONDANI et al., 1998) e mais recentemente, com marcadores de polimorfismos de base individual (SNPs) (BRONW et al., 2004).

Os métodos de transformação genética oferecem uma importante opção para o melhoramento de árvores pois permite a introdução de um gene, ou genes, desejado em um único passo, na tentativa de se alcançar objetivos específicos. Dessa forma, a transformação genética do eucalipto tem sido o objetivo de vários grupos de pesquisa nos últimos anos, mas poucos trabalhos relatam a obtenção de plantas transgênicas de eucalipto, devido principalmente a sua recalcitrância (SERRANO et al., 1996; MORALEJO et al., 1998). Somente em 2002, Gonzalez e colaboradores conseguiram obter plantas transgênicas para a principal espécie de interesse comercial no Brasil, o *E. grandis*, e para o híbrido *E. grandis* x *E. urophylla* através do sistema de transformação por *Agrobacterium tumefaciens*, assistido por sonicação. Com isto, as características de interesse do eucalipto podem ser geneticamente modificadas, ou até mesmo novas características podem ser introduzidas, aumentando ainda mais a produtividade do setor florestal brasileiro.

2.4 Caracteres de qualidade da madeira

A globalização da economia promoveu um forte aumento na concorrência entre as empresas de base florestal, que passaram a se preocupar e tomar decisões no sentido de tirar o máximo proveito da matéria-prima disponível e de otimizar o processo produtivo. Os objetivos das empresas passam a ser, então, o baixo custo e a alta qualidade, definindo quem permanece no mercado.

A qualidade da madeira é a soma de todas as características e propriedades que afetam o rendimento em produtos finais e sua adequação para as aplicações pretendidas. A qualidade final dos produtos pode ser afetada por características, como: densidade, rigidez, estabilidade, presença de nós, cerne, bolsas de resina, teor de lignina, teor de extrativos e outros (A INDUSTRIALIZAÇÃO, 2007). Assim, conhecendo-se a qualidade da matéria-prima e o processo a ser utilizado, é possível obter a otimização de ambos e do produto final. Na indústria siderúrgica, as características mais importantes da madeira são aquelas que afetam positivamente a qualidade do carvão, tais como elevadas densidades e alto teor de resinas as quais somam maior

poder calorífero ao carvão. Para a indústria moveleira, características como uniformidade no sentido medula-casca, cor, baixa retratibilidade, desrama natural e boa usinabilidade, são desejáveis (FERREIRA, 2003). As características da madeira desejadas para construção civil são a resistência, a rigidez, retidão, estabilidade dimensional, isenção de empenamentos e defeitos, longos comprimentos e durabilidade. Já a madeira destinada à produção de celulose e papel deve apresentar alto conteúdo de celulose, elevada proporção relativa de fibras e baixo teor de lignina e extrativos, visando diminuir o custo de produção e aumentar a qualidade do produto final (POT et al., 2002).

Entretanto, a maioria dessas características é produzida pela ação conjunta de outros caracteres que influenciam a composição química, física e anatômica da madeira. Essas características podem variar de espécie para espécie, dentro de uma árvore no sentido medula-casca e base-topo, podem variar de acordo com as estações do ano, entre outros. Para a indústria de papel e celulose, as características desejadas parecem estar intimamente ligadas à idade das árvores. A característica de maior importância e que tem despertado grande interesse em pesquisas químicas, físicas e anatômicas é a proporção de madeira madura produzida pelas árvores, uma vez que sendo mais densa, leva a uma madeira de melhor qualidade.

2.4.1 Madeira Juvenil vs Madeira Adulta

A qualidade da madeira juvenil, formada nos primeiros anos de vida das árvores, é comprovadamente diferente daquela apresentada pela madeira adulta. Esse tipo de madeira, quando processado pela indústria, pode acarretar certas características indesejáveis no produto final (BRASIL e FERREIRA, 1979). Entretanto, as árvores de eucalipto que são encaminhadas para industrial de papel e celulose são cortadas quando estão com sete anos de idade, constituída ainda por grande parte de madeira juvenil.

A madeira juvenil ocorre na região central da árvore e apresenta a forma de um cilindro, com diâmetro quase uniforme, desde a base até o topo (EVANS et al., 2000) (Figura 1). De maneira geral, ela apresenta uma elevação dos valores de densidade, comprimento de fibra, etc, da fase juvenil até atingirem a maturidade, onde os valores permanecem mais ou menos constantes. Na fase juvenil a taxa de incorporação de biomassa é crescente, tendendo a se estabilizar, até a árvore atingir a idade adulta (TRUGILHO et al., 1996).

O período de juvenilidade do eucalipto é variável entre as suas espécies, podendo ser superior a 10 anos em alguns *Eucalyptus* plantados na Austrália (OLIVEIRA et al., 2005). Em áreas de plantio de rápido crescimento a madeira juvenil é de importância considerável, pois ela pode constituir uma grande proporção do tronco. Em relação à madeira adulta a madeira juvenil apresenta uma maior quantidade de xilanas e ligninas, menos celulose e glucomananas, menor comprimento de fibra e vasos, menor diâmetro celular, paredes celulares mais finas, menor densidade básica, maiores ângulos microfibrilares e maiores proporções de fibras e menores proporções de vasos (TRUGUILHO et al., 1996).

Figura 1 - Localização da madeira juvenil e adulta na árvore e suas principais características

Oliveira e colaboradores (2005), avaliando teores de umidade e de densidade básica na madeira de sete espécies de eucalipto com idades entre 22 e 26 anos concluíram que, existe uma queda no gradiente de umidade no sentido medula-casca, com exceção do *E. grandis* que mostrou uma distribuição heterogênea. Quanto à densidade básica, os autores concluíram que houve um aumento dessa propriedade na direção medula-casca, com perfis de variação distintos para cada espécie. No caso do *E. grandis*, a madeira de densidade menos variável foi verificada somente a partir de 25 cm de diâmetro do tronco.

Truguilho e colaboradores (1996), avaliando a madeira de *Eucalyptus saligna* aos 12, 24, 36 e 48 meses, concluíram que a densidade básica apresentou elevados e significativos coeficientes de correlação com as características dimensionais da fibra, coeficientes de correlação negativos e significativos com o teor de extrativos totais e lignina, e positivo com o teor de

holocelulose (celulose + hemicelulose). Assim, a composição química da madeira estaria intimamente relacionada à densidade básica, e que, dentre os componentes químicos, os carboidratos são os principais responsáveis pela sua variação.

Quanto às variações nos ângulos microfibrilares, estes têm um profundo efeito nas propriedades da madeira, em particular em sua dureza. Os maiores ângulos microfibrilares da madeira juvenil conferem baixa dureza e dão à árvore a flexibilidade necessária para ela sobreviver a fortes ventos sem se quebrar. Isto significa, entretanto, que a madeira contendo altas proporções de madeira juvenil não é favorável para produção e polpa de celulose (BARNETT e BONHAM, 2004). Ainda, a angulação das microfibrilas de celulose está intimamente ligada com o comprimento das fibras e com a quantidade de celulose depositada, principalmente na camada S2 da parede secundária.

Na Tabela 1 são apresentados os resultados obtidos por TOMAZELLO FILHO (1985) a respeito da variação radial da estrutura anatômica de *Eucalyptus grandis* com 10 anos de idade.

Tabela 1 - Valores médios do comprimento, largura e espessura da parede das fibras de *E. grandis* no sentido medula-casca

Característica	Medula ← Posição → Casca				
	0%	25%	50%	75%	100%
Comprimento da fibra (mm)	0,68	0,85	1,07	1,20	1,32
Largura da fibra (µm)	19,30	18,60	21,23	23,00	24,10
Espessura da parede (µm)	4,60	4,50	5,10	5,40	6,00

Fonte: TOMAZELLO FILHO (1985)

Existem várias hipóteses para a formação da madeira juvenil. PLUMPTRE (1983), estudando *Pinus caribaea*, verificou que a formação da madeira juvenil e a transição para a madeira adulta variam com a constituição genética da árvore, com o clima e com as atividades silviculturais. Em seu trabalho, OLIVEIRA (1997) cita que a auxina encontrada na casca fotossintética (primeira periderme) causa a formação de madeira juvenil ou previne a formação de madeira adulta e que a localização da transição madeira juvenil/adulta está relacionada com a transição entre a casca fotossintética e a não fotossintética (ponto de formação da segunda periderme). Para Bendtsen (1978), a madeira juvenil é o xilema secundário produzido durante a fase jovem da árvore, mais precisamente pelas regiões cambiais que são influenciadas pela atividade dos meristemas apicais.

As variações observadas entre a madeira juvenil e a madeira adulta são o resultado das variações que ocorrem nas células que as originam, denominadas fusiformes iniciais. A estabilização de todas as propriedades da madeira ocorre somente quando as células do câmbio atingirem comprimento máximo, iniciando a formação da madeira caracterizada como adulta (TOMAZELLO FILHO, 1987). Segundo Zobel (1980), citado por Severo et al. (2006), o uso do termo madeira juvenil é de certa forma errôneo, pois não ocorre uma mudança abrupta da madeira juvenil para a adulta, mas sim uma variação das propriedades físicas, químicas, mecânicas e anatômicas da madeira durante um período longo de anos, caracterizando assim a zona juvenil que apresenta propriedades variáveis se comparada com a zona adulta, que as tem de forma constante.

A produção de polpa a partir do lenho juvenil é menor do que a partir do lenho adulto, tanto com base em massa seca quanto com base em massa úmida. O papel fabricado de polpa de lenho juvenil tem maior resistência à tração, ao estouro, à dobra, apresenta folha mais lisa e possui maior densidade aparente (ZOBEL e VAN BUIJTENEN, 1989). A madeira madura, mais homogênea quanto à sua densidade no interior do tronco, comporta-se melhor nas operações de processamento e reflete maior uniformidade nas demais propriedades tecnológicas. As árvores em idade de corte contêm altas proporções de madeira juvenil e, portanto, um menor valor econômico e, embora sejam selecionadas para altas taxas de crescimento inicial, atingindo altos valores em diâmetro, a busca por um aumento na produção de madeira deve concentrar esforços no melhoramento da qualidade da madeira ligada a sua finalidade na indústria, a qual é refletida pela quantidade de madeira adulta presente.

2.5 A Formação da Madeira

A formação da madeira é um processo derivado do crescimento secundário das plantas que, diferentemente do crescimento primário, é produzida a partir de células meristemáticas do câmbio. A madeira, então, nada mais é que o produto do câmbio vascular, e é composta de um conjunto de células mortas do xilema secundário que são acumuladas ao longo do desenvolvimento das espécies lenhosas (LARSON, 1994). O processo responsável pela formação do xilema, portanto, é chamado de xilogênese.

O câmbio vascular é geralmente descrito como uma camada única de células constituída por células iniciais permanentes, as iniciais cambiais. Quando estas células se dividem

periclinalmente para dentro, uma célula permanece no estado meristemático, enquanto a outra, mais interna, se torna uma célula mãe do xilema. A célula mãe do xilema pode, por sua vez, sofrer ainda um limitado número de divisões celulares ou, sem mais divisões, se diferenciar em células de xilema secundário (LARSON, 1994). O xilema secundário maduro é composto de uma série de tipos celulares com forma e funções especializadas, os quais incluem células condutoras, células de suporte e células de estoque. Em angiospermas, as células condutoras são compostas por vasos e as de suporte, pelas fibras, ambas com paredes espessas e alongadas na direção longitudinal do tronco das árvores. Já as células de estoque, as células do parênquima radial e longitudinal, têm paredes finas e geralmente permanecem vivas para estocar e transportar nutrientes por um longo período de tempo (LI et al., 2006).

Como e quando o destino dessas células é estabelecido é um dos mais intrigantes temas da xilogênese. Sugere-se que essas células já são determinadas a assumir formas e funções diferentes em estágios iniciais do desenvolvimento do tecido secundário, contudo, a exata natureza molecular dessas diferenças ainda não está completamente entendida (MELLEROWICZ et al., 2001).

Uma vez formada, a nova célula do xilema passará por um processo de diferenciação que pode ser dividido em quatro etapas principais: divisão celular, expansão celular (elongação e aumento radial), espessamento da parede celular (envolvendo a deposição de celulose, hemicelulose, proteínas da parede e biossíntese de lignina) e morte celular programada (PLOMION et al., 2001).

Após a divisão, as células que ainda são compostas somente de parede celular primária, se expandem, formando uma zona onde os elementos de vasos e as fibras claramente já apresentam características morfológicas diferentes. Durante a expansão, cada tipo celular difere em comprimento, polaridade e tipo de extensão. As fibras e as células do parênquima expandem primariamente na direção radial, com suas paredes expandindo-se uniaxialmente; enquanto os vasos, em adição à expansão radial, podem sofrer um crescimento tangencial. O alongamento das fibras é alcançado pelo crescimento reforçado das pontas das células e requer biogênese local de parede celular. O padrão básico do arranjo entre os diferentes tipos celulares do xilema é bastante conservado, apesar de seus diferentes padrões de expansão (CATESSON, 1989). Ainda, como o grau de expansão varia entre os tipos celulares, devem existir mecanismos que regulam de

maneira diferencial a pressão de turgor e/ou a plasticidade da parede celular (MELLEROWICZ et al., 2001).

Após a expansão celular, todas as células do xilema formam uma parede celular secundária, caracterizada por um espessamento com diferentes padrões, que fornece às células resistência mecânica para suportar o peso da árvore e a pressão negativa gerada pela transpiração. O espessamento padronizado da parede secundária é regulado através da deposição controlada de microfibrilas de celulose (FLADUNG et al., 1991). Enquanto a orientação das microfibrilas na parede primária é aleatória ou longitudinal (FUNADA, 2000), a da parede secundária consiste de três camadas S1, S2 e S3, e é marcada pela formação de um denso arranjo helicoidal, quase transversal (ABE et al., 1997). A primeira camada formada S1 é mais fina (0,1-0,35 μm) e é responsável por apenas 5-10% da espessura da parede celular, apresentando um ângulo reto, quase transversal de microfibrilas (60-80°) que gradualmente muda num sentido horário para um arranjo longitudinal que caracteriza a camada S2. Nesta segunda camada as microfibrilas podem ser depositadas em ângulos que variam entre 5° e 30° e é mais espessa (1-10 μm), representando 75-85% da parede celular. Finalmente, a camada S3 é formada após uma abrupta re-orientação das microfibrilas de volta para as hélices transversais, num ângulo de 60-90° (CHAFFEY et al., 1997; FUNADA, 2000) (Figura 2).

Figura 2 - Estrutura tridimensional da parede celular secundária de uma célula do xilema. Apresentação dos diferentes ângulos de microfibrilas na parede celular primária e em cada camada da parede celular secundária. Retirado de Plomion et al. (2001)

A lignificação das células começa pelos vasos, com as fibras sendo as últimas a entrar nesta fase. A lignina é depositada primeiramente na lamela média, particularmente nos cantos das células, quando estas completam a deposição da camada S1 (TERASHIMA et al., 1993). Após, a lignificação progride internamente durante a formação da camada S2, concomitantemente com a deposição de celulose e hemiceluloses, mas este processo é mais intenso quando a camada S3 está formada (MELLEROWICZ et al., 2001). Quando a lignificação é finalizada, os elementos de vasos e fibras sofrem morte celular programada (PCD), a qual envolve a hidrólise do protoplasto. Este processo leva as células ao final da diferenciação, tornando-as prontas para assumir suas funções. No caso dos vasos, as extremidades das células ainda são perfuradas para formar colunas tubulares, enquanto as fibras apresentam apenas lumens. A madeira, portanto, é composta principalmente por células mortas, com uma espessa parede celular composta basicamente por celulose, hemicelulose, proteínas e lignina.

A composição química das células do xilema difere entre a parede celular primária e a parede celular secundária. As células em fases iniciais de diferenciação, contendo apenas parede celular primária, contêm 47% de pectina, 22% de celulose, 18% de uma matriz de glicanas (hemicelulose), 10% de proteínas e 3% de outros materiais (SIMSON e TIMELL, 1978). As proteínas presentes na parede celular têm função estrutural e as pectinas desempenham papel fundamental na regulação de sua extensibilidade. Em contraste, a madeira com parede celular secundária é composta de cerca de 40-50% de celulose. Esta, está associada a hemiceluloses - polissacarídeos solúveis não celulósicos - que respondem por cerca de 25% do peso seco da madeira. Já a lignina, um polímero fenólico derivado de álcoois hidroxinamil, representa o terceiro maior componente da madeira (25-30%) e se encaixa na matriz de polissacarídeos dando rigidez e coesão ao tecido da madeira, além de formar uma superfície hidrofóbica necessária para o transporte de água (PLOMION et al., 2001).

A celulose é formada pela polimerização de monossacarídeos de β -(1 $\rightarrow$ 4) D-glicose, que através de um complexo de enzimas que sintetizam 36 cadeias lineares desse açúcar, unidas umas às outras, ao longo de seu comprimento, por pontes de hidrogênio, formando uma estrutura em roseta com um diâmetro aproximado de 8 a 10 nm (CARPITA et al., 2000). As hemiceluloses são polissacarídeos sintetizados no complexo de Golgi e caracterizam-se por sua flexibilidade e capacidade de se ligar à cadeia de celulose. Diversos tipos de hemicelulose são encontrados em paredes celulares. Elas possuem uma estrutura básica de resíduos de β -(1 $\rightarrow$ 4) D-glicose, com

cadeias laterais de xilose, arabinose, galactose, manose, entre outros, dependendo da espécie. As pectinas incluem vários tipos de polissacarídeos ricos em ácido galacturônico (GalA) e as proteínas da parede são classificadas de acordo com sua composição predominante de aminoácidos, por exemplo, glicoproteína rica em hidroxiprolina (HRGP), proteína rica em glicina (GRP) e proteína rica em prolina (PRP) (TAIZ e ZEIGER, 2004; REITER, 2002). Finalmente, a lignina é formada pela polimerização de três diferentes álcoois de fenilpropanóides (ou monolignóis): coniferil, coumaril e sinapil (TAIZ e ZEIGER, 2004), que dão origem às unidades G,H e S, respectivamente, diferindo entre si apenas pelo grau de metoxilação (PLOMION et al., 2001).

A xilogênese, portanto, é um exemplo da diferenciação celular em uma forma extremamente complexa. Este processo é controlado por uma enorme variedade de fatores exógenos (fotoperíodo e temperatura) e endógenos (fitormônios) e pela interação entre eles. A presença de famílias multigênicas, a extrema plasticidade do metabolismo envolvido e a expressão coordenada de muitos genes estruturais proporcionam essa enorme complexidade, dificultando o entendimento sobre o processo de formação da madeira (PLOMION et al., 2001).

2.6 O estudo da xilogênese na era das “Ômicas”

As empresas nacionais de papel e celulose contam hoje com as maiores produtividades florestais do mundo, graças às florestas de eucalipto de rápido crescimento e grande produção de biomassa lenhosa. A manutenção dessa vantagem em relação aos outros países vai depender fortemente de investimentos feitos em programas de melhoramento nos próximos anos, que agora, podem contar com o desenvolvimento das redes de genômica e com os avanços da bioinformática. Essas novas ferramentas têm revolucionado profundamente as perspectivas e a velocidade de utilização do conhecimento biológico, uma vez que são eficientes, sofisticadas e precisas na análise e interpretação de grandes quantidades de dados gerados por sequenciamento de genomas e proteomas e de perfis transcricionais.

A genômica pode ser definida como o grupo de estratégias utilizadas na identificação de genes e genomas. Os dados provenientes do sequenciamento de genomas tornam possível o entendimento da estrutura, organização e função dos genes numa espécie, como também, fornecem evidências para estudos evolutivos e das funções gênicas (genômica funcional). A era da genômica florestal começou com projetos de sequenciamentos de ESTs (*Expressed Sequence*

Tags) em *Pinus* e *Populus* e teve como objetivo inicial a identificação de genes específicos de espécies arbóreas (BHALERAO et al., 2003). Diferentemente do sequenciamento de genomas completos, onde o conjunto das sequências de nucleotídeos (expressas ou não) é acessado, este tipo de estudo considera apenas os genes que são expressos.

Os trabalhos pioneiros sequenciaram 5692 ESTs de *Populus* (STERKY et al., 1998) e 1000 ESTs de *Pinus* (ALLONA et al., 1998). Desde então, já estão disponíveis nos bancos de dados públicos (<http://www.ncbi.nlm.nih.gov/>) mais de 359000 ESTs de *Pinus sp.* e mais de 369200 de *Populus sp.* Além disso, já foram depositadas no NCBI 15413 sequências de ESTs de *E. globulus*, *E. camaldulensis* e *E. grandis*. No Brasil, foram criados dois projetos para o sequenciamento de ESTs de *Eucalyptus* financiados pela iniciativa privada e por órgãos públicos de auxílio à pesquisa (*Genolyptus* e *Forests*). Infelizmente, esses dados ainda não foram disponibilizados publicamente, o que restringe a realização de pesquisas e pode interferir no avanço da pesquisa básica e aplicada dessa espécie.

O primeiro sequenciamento de um genoma completo de uma espécie arbórea foi o do *Populus trichocarpa* (TUSKAN et al., 2006). O genoma do cloroplasto de *E. globulus* já está disponível (STEANE, 2005), e o sequenciamento do genoma de espécies de eucalipto de interesse comercial (*E. camaldulensis* e *E. grandis*) já foi iniciado pelo consórcio internacional do genoma de *Eucalyptus* (IEuGC), formado em 2004. Portanto, para os próximos anos haverá uma grande quantidade de dados genômicos dessa espécie.

Com a geração desses dados, foi criada a necessidade da identificação da função dos genes e compreensão de processos biológicos por eles controlados, o que levou ao início da era das “ômicas”. Esta nova era na genética molecular de plantas compreende o estudo de transcritos ou transcritômica, o estudo de proteínas ou proteômica (CHEN e HARMON, 2006), e o estudo dos metabólitos celulares ou metabolômica (BHALLA et al., 2005), com o objetivo maior de traçar perfis funcionais da célula e/ou tecido, fornecendo uma visão holística da interação genes-fenótipo.

De maneira geral, as metodologias utilizadas para estabelecer perfis transcricionais baseiam-se na geração de populações de cDNA a partir da população de RNAm, e podem ser divididas em métodos abertos ou globais de análise, ou em métodos fechados ou limitados de análise (ALBA et al., 2004). Os métodos globais permitem o acesso à potencialmente todos transcritos expressos em um determinado momento, uma vez que informações prévias de

sequências gênicas para a espécie não são requeridas; já os métodos fechados são limitados pela quantidade de informação disponível para a espécie uma vez que utilizam sondas e/ou *primers* específicos para hibridização.

Atualmente, diferentes metodologias estão sendo aplicadas para a análise global do transcrito de uma espécie, entre elas o sequenciamento de grandes quantidades de ESTs (RONNING et al., 2003), o SAGE (VESLJUCULESCU et al., 1995), o *Differential display* (LIANG e PARDEE, 1992), *Supression subtractive hybridization* (SSH) (DIATCHENKO et al., 1996), *Amplified restriction fragment polymorphism* baseado em cDNA (cDNA-AFLP) (BACHEM et al., 1996), e *Massively parallel signature sequencing* (MPSS) (BRENNER et al., 2000). Entre as metodologias restritas a disponibilidade de informações, os microarranjos de oligonucleotídeos ou de cDNA são os mais utilizados. Outra técnica que tem sido bastante aplicada é o qRT-PCR, ou RT-PCR em Tempo Real, usada para quantificar a expressão de genes específicos. Enquanto os microarranjos são ferramentas poderosas para identificar genes diferencialmente regulados sob determinadas condições, o qRT-PCR é um método "mais quantitativo, usado tanto para validar os resultados dessas técnicas de expressão gênica, quanto para fornecer uma maior acurácia e precisão em pesquisas específicas (ROTH, 2002).

O terceiro nível na análise da expressão de genes é a proteômica, que reúne técnicas para o sequenciamento do conjunto de proteínas presentes num organismo. Através desse tipo de estudo é possível obter informações detalhadas sobre a estrutura e função das proteínas em diferentes sistemas biológicos, e verificar se existe correlação entre o nível de expressão dos RNAs e das proteínas, a ocorrência de isoformas e de modificações pós-traducionais (PATTERSON e AEBERSOLD, 2003).

Basicamente, a metodologia para o sequenciamento de proteomas consiste de uma separação prévia dos componentes da amostra através de eletroforese bidimensional ou cromatografia líquida com o objetivo de aumentar a resolução do segundo passo, o sequenciamento propriamente dito. Para isto, a análise e caracterização estrutural das moléculas é realizada através da espectrometria de massas (PIMENTA, 2003).

O controle genético da formação da madeira em árvores vem sendo investigado através da interação das diferentes abordagens envolvendo a genômica, transcritômica, proteômica e a bioinformática, associada ao uso de plantas modelo como a *Zinnia sp*, *Arabidopsis sp* e *Populus sp*. Hertzberg et al. (2001) estabeleceram um padrão hierárquico da expressão gênica de

diferentes zonas de desenvolvimento do xilema de *pópulus* isolando suas células nos diferentes estádios da xilogênese. Usando análise de microarranjos, os autores mostraram que genes que codificam enzimas envolvidas na biossíntese de celulose, bem como um grande número de fatores de transcrição e potenciais reguladores da xilogênese estão sob um rígido controle em cada um desses estágios.

Paux et al. (2004) desenvolveram um experimento dirigido de genômica funcional através da construção de uma biblioteca subtrativa de xilema vs folhas para identificar genes envolvidos no controle da formação da madeira de *Eucalyptus*. As duas principais classes de ESTs preferencialmente expressas em xilema foram relacionadas a sinalização por auxina através da via proteolítica de ubiquitinação e, biossíntese e remodelagem de parede celular.

Mais recentemente, a indução de madeira de tensão foi usada como um sistema modelo para estudar a formação de madeira devido ao seu alto conteúdo de celulose e lignificação reduzida (PAUX et al. 2005, ANDERSSON-GUNNERAS et al. 2006).

Recentemente, Ranik e Myburg (2006) clonaram os genes que codificam as seis celulosas sintases do Eucalipto e, através de análises de PCR em tempo real verificaram que os genes *EgCesA* apresentavam expressões diferenciadas em tecidos em processo de deposição de parede celular primária e secundária. Dessa forma, os genes foram separados em dois grupos, onde *EgCesA1*, *EgCesA2* e *EgCesA3* estão envolvidos na deposição de celulose em xilema e, *EgCesA4* e *EgCesA5* que tiveram níveis maiores de expressão em flores e folhas. O *EgCesA6* foi expresso em níveis extremamente baixos em todos os tecidos.

Pesquisas direcionadas para genes específicos relacionados à formação da madeira também têm sido bastante empregadas, principalmente na obtenção de mutantes ou de transformantes. Jouanin et al. (2000), obtiveram plantas de *Populus* transgênicas que superexpressavam um gene que codifica a enzima ácido cafeico 3-O-metiltransferase (COMT). Devido ao fenômeno do silenciamento gênico, uma das linhagens transgênicas que continha duas cópias apresentou uma atividade da enzima igual a zero, o que resultou na redução do nível de lignina em 17% e no aumento do conteúdo de celulose, levando a um aumento no rendimento de produção de pasta celulósica de 10%.

Nishikubo et al. (2007) estudou os efeitos da expressão de uma xiloglicana endotransglicosilase (XTH34) em linhagens transgênicas de *Populus*. As linhagens com níveis altos para seus transcritos apresentaram uma alta atividade da enzima em tecidos em formação de

madeira e exibiram mudanças na quantidade e no tamanho das cadeias de xiloglicanas e, conseqüentemente, no tamanho das células de fibra.

Mijnsbrugge et al. (2000) analisaram a formação da madeira em *Populus trichocarpa* através da identificação e caracterização sazonal de proteínas em diferentes tecidos. A avaliação das proteínas do xilema jovem, xilema adulto e da casca, colhidos em quatro períodos do ano, demonstrou a existência de certa similaridade entre os diferentes tecidos e entre as diferentes épocas do ano, embora houvesse uma variação em sua abundância, isto é, durante o verão a atividade metabólica foi mais intensa.

Recentemente, Celedon et al. (2007) utilizaram um sistema 2D-LC-MS/MS para identificar e caracterizar proteínas presentes na região cambial de *Eucalyptus grandis* em três idades (6 meses, 3 anos e 6 anos), na tentativa de traçar um perfil protéico dinâmico desse tecido desde o início do desenvolvimento das árvores até a sua idade de corte. No total, foram identificadas 240 proteínas, entre as quais, nove mostraram mudanças significantes em seus níveis de expressão conforme a idade analisada.

2.6.1 A técnica SAGE (Análise Serial da Expressão Gênica)

Com o avanço da genômica funcional tornou-se possível a análise da expressão de centenas, e em alguns casos, milhares de genes simultaneamente. Essas novas técnicas permitem a análise de populações completas de RNAm em células ou tecidos indicando o perfil da transcrição global (STOLLBERG et al., 2000).

A técnica SAGE (*Serial Analysis of Gene Expression*) foi desenvolvida por Velculescu et al. (1995) e é comercializada na forma de *kit* pela Invitrogen®. O SAGE é um método capaz de identificar e quantificar transcritos de genomas eucariotos, num dado momento, levando a uma análise qualitativa altamente precisa da expressão gênica de um grande número de transcritos simultaneamente. Por ser um sistema aberto, esta técnica também permite a identificação de genes ainda desconhecidos. Ainda, ela tem a exclusividade de apresentar o perfil dos transcritos de uma dada amostra de forma digital.

A metodologia resume-se na geração de sequências curtas de cDNA (14 bases) ou *tags* a partir de uma posição única de cada RNAm da amostra, e na concatenação destas sequências para clonagem e sequenciamento (Ver Material e Métodos). A concatenação das *tags* representa um importante passo no processo uma vez que permite que o sequenciamento da biblioteca SAGE

seja realizado mais rapidamente e com maior economia quando comparado ao sequenciamento de bibliotecas de ESTs. Cada sequência gerada pelo SAGE é o produto de múltiplas *tags* ou transcritos sequenciados em uma única reação por clone (SUN et al., 2004).

O *kit* SAGE conta com um *software* que extrai as *tags* dos dados de sequenciamento e contabiliza quantas vezes cada *tag* apareceu na biblioteca. A frequência de cada *tag* reflete diretamente a abundância de cada RNAm presente no tecido amostrado. Além disso, tamanho das *tags* contém informação suficiente para identificar unicamente um gene, já que elas são extraídas de uma posição definida do transcrito. Além disso, a quantidade máxima de *tags* diferentes que podem ser produzidas por um tecido ($4^{10} = 1.048.576$) é muito maior que o número estimado de genes em plantas (COOKE et al., 1996).

Mesmo assim, o ponto crítico do método é o processo denominado *tag mapping*, ou anotação, no qual cada *tag* é associada ao seu transcrito de origem. Para isto, é de extrema importância a disponibilidade de um banco de dados de sequências de ESTs 3' ou de cDNAs completos da espécie que se deseja estudar (LASH et al., 2000). Como as *tags* são produzidas a partir da clivagem de um ponto único na molécula de cDNA, definido pelo último sítio de corte da enzima de restrição que antecede o terminal poli(A) fixado às esferas magnéticas, todo o processo de anotação das *tags* deve ser baseado neste fundamento do método.

Um importante recurso comumente utilizado para a anotação das *tags* é o SAGEmap, disponível no *site* do NCBI (<http://www.ncbi.nlm.nih.gov/sage>). Este programa cruza os dados do SAGE a um conjunto de *tags* extraídas de ESTs e sequências de cDNAs depositadas no NCBI UniGene clusters (<http://www.ncbi.nlm.nih.gov/UniGene/>). Para a extração das *tags*, o SAGEmap utiliza apenas as sequências expressas (ESTs e cDNAs) com a terminação poli(A) ou com o sinal de poliadenilação. As SAGE *tags* são extraídas da porção mais próxima ao terminal 3' de cada sequência (determinada pela presença do sítio de restrição da enzima utilizada na geração das *tags*), e as *tags* “virtuais” podem então ser comparadas ao conjunto de *tags* obtido experimentalmente (PLEASANCE et al., 2003).

Uma vez que o SAGE produz bibliotecas de *tags* com valores absolutos de níveis de expressão, ele também permite a realização de comparações diretas dos padrões de expressão para diferentes tratamentos de um mesmo tecido ou estádios diferentes do desenvolvimento. Adicionalmente, o uso do SAGE permite o acesso a transcritos de baixa abundância, ao quais são facilmente perdidos em outras abordagens (SUN et al., 2004; GIBBINGS et al., 2003) e as *tags*

geradas por ele, podem ser utilizadas como *primers* na identificação de genes ainda desconhecidos, ou como “sondas” para identificar genes desconhecidos em genomas (COEMANS et al., 2005; CHEN et al., 2002).

A eficácia do uso do SAGE para analisar o transcrito de *Arabidopsis* foi maximizada por Robinson et al. (2004). Os autores descreveram um processo de anotação que explora as sequências de genoma completo desse organismo através da obtenção de todas as possíveis *tags* presentes nos dados genômicos. Essas *tags* “virtuais” foram então contrastadas com as *tags* geradas experimentalmente. Em adição, eles também contrastaram suas *tags* com as *tags* virtuais existentes em outros sítios de clivagem da enzima, que não o mais próximo da terminação poli(A). Dessa forma um número muito maior de *tags* pôde ser identificado em relação à forma convencional de identificação, além da geração de evidências da presença de transcritos alternativos. Ainda, as *tags* também foram utilizadas para evidenciar a ocorrência da regulação anti-sense em *Arabidopsis* através da identificação de sequências complementares invertidas entre os dados.

Poucos trabalhos de análise do transcrito de plantas têm sido desenvolvidos através da técnica SAGE (MATSUMURA et al., 1999, 2003; LORENZ e DEAN, 2002; GIBBINGS et al., 2003; COEMANS et al., 2005; BAO et al., 2005), já que para a maioria das espécies ainda existem poucas sequências disponíveis. O trabalho realizado por Matsumura et al. (1999) em arroz foi o primeiro trabalho onde o SAGE foi empregado no estudo da expressão gênica em plantas. Os autores foram os pioneiros também em demonstrar um método mais eficiente e rápido que o reportado no trabalho original de Velculescu et al. (1995) para a identificação de novos genes. Lorenz e Dean (2002) desenvolveram o primeiro estudo de árvores utilizando o SAGE, no qual foram comparadas bibliotecas SAGE criadas a partir do xilema do ápice e da base do tronco de uma árvore de *Pinus taeda* de 10 anos. Para o ápice, foram geradas 85.000 *tags*, representando 27.398 genes, e para a base do tronco, 65.000 *tags* ou 25.983 genes. Os autores identificaram 33 *tags* com níveis de expressão diferentes em pelo menos 5 vezes entre as bibliotecas.

Salvatierra (2006) construiu duas bibliotecas SAGE a partir de folhas de *Eucalyptus* suscetíveis e resistentes à ferrugem (*Puccinia psidii*) e identificou 232 genes diferencialmente expressos entre elas que podem estar envolvidos de alguma forma na resistência à doença. A comparação dos perfis transcricionais entre folhas e a região cambial de *Eucalyptus grandis* foi realizada por Carvalho (2006) e levou a identificação de genes exclusivamente expressos em uma

das bibliotecas e revelando potenciais genes candidatos para busca de promotores tecido-específicos.

Algumas variações da técnica originalmente descrita vêm sendo desenvolvidas para facilitar a identificação precisa de genes através das *tags* e para favorecer a caracterização de novos genes através do uso direto da própria sequência da *tag* como *primers* (MATSUMURA et al., 2003; WEI et al., 2004; GOWDA et al., 2004). Entre eles, o *SuperSAGE* se baseia no uso de uma enzima de restrição tipoIII para gerar *tags* maiores que 25 bases (MATSUMURA et al., 2003), e o *LongSAGE* produz *tags* de 20 bases tanto a partir da região 5' quanto da região 3' dos transcritos, aumentando a especificidade de identificação de genes. Outras modificações foram também propostas para aumentar a eficiência da clonagem e tamanho dos insertos, reduzindo o custo das metodologias que produzem *tags* de maior tamanho (GOWDA et al., 2004).

3 MATERIAL E MÉTODOS

3.1 Material vegetal

O material vegetal utilizado neste trabalho foi gentilmente fornecido pela empresa Suzano Papel e Celulose. As árvores amostradas localizavam-se na cidade de Itapetininga, estado de São Paulo, Brasil (23°35'20''S e 48°03'11''O) na altitude de 656 m e estavam espaçadas em 3 m por 1,5 m. Para construção da biblioteca SAGE foram amostrados 35 indivíduos com 1 ano e 7 meses e 35 com 2 anos e 9 meses de uma população de meio-irmãos de *Eucalyptus grandis* provenientes de um plantio comercial. Para a análise comparativa da expressão gênica entre idades foram amostrados 30 indivíduos dessa mesma população com 3 anos de idade e 30 indivíduos de uma população de matrizes com 22 anos de idade. A região cambial das árvores, que compreende não somente o meristema cambial, mas todas as células em diferenciação e maduras do floema e do xilema, foi coletada utilizando um método não destrutivo através do qual uma janela de 20 x 15 cm da casca das árvores no DAP (diâmetro à altura do peito) é aberta, o tecido exposto do tronco e da casca das árvores é raspado com uma lâmina e imediatamente congelado em Nitrogênio líquido (Figura 3a,b).

Figura 3 - Amostragem do material biológico. (a) Remoção da casca; (b) Raspagem do tecido da região cambial; (c) corte transversal da madeira de *Eucalyptus* antes da remoção da casca; (d) corte transversal da madeira de *Eucalyptus* após a remoção da casca. Escala: barra = 100 μ m. X = xilema, C = região cambial e F = floema

A parte interna da casca também foi raspada para garantir que todo o material meristemático fosse representado na amostragem. Este procedimento foi usado primeiramente por Foucart et al. (2006) que mostrou que grande parte dessas células ficava aderida à casca quando ela era removida. As Figuras 3c e 3d mostram dois cortes transversais de madeira de

Eucalyptus (xilema, região cambial e floema) onde é possível observar as células meristemáticas sendo removidas juntamente com a casca.

No laboratório, todas as amostras foram maceradas em Nitrogênio líquido com auxílio de almofariz e pistilo até a obtenção de um fino pó e reunidas de acordo com a idade. Todo material macerado ficou armazenado em ultrafreezer (-80°C) onde permaneceu até sua utilização.

3.2 Extração de RNA total

Como a molécula de RNA possui uma maior propensão à degradação, todo e qualquer material utilizado para sua obtenção e manipulação, incluindo vidraria, tubos plásticos, reagentes e outros, foi previamente tratado para eliminação de RNAses (ANEXO A).

A extração do RNA total dos tecidos da região cambial foi realizada seguindo um protocolo de extração baseado em fenol, descrito por Salzman et al. (1999) (ANEXO B) com pequenas modificações. A escolha deste protocolo foi em decorrência de ele ter sido descrito como muito eficiente para extração de RNA de tecidos vegetais contendo altos níveis de compostos fenólicos e por sua eficiência já ter sido comprovada em trabalhos anteriores do laboratório, onde foi utilizado para extração de RNA de vários tecidos de eucalipto, principalmente para utilização na construção de bibliotecas de ESTs.

Dois gramas de tecido vegetal macerado são necessários para extração de RNA total através do protocolo escolhido. Assim, para extração do RNA total utilizado na construção da biblioteca SAGE, 1 g de tecido macerado de cada idade (1 ano e 7 meses e 2 anos e 9 meses) foi misturado para obtenção da amostra final; e para o RNA utilizado na comparação de expressão gênica entre idades, 2 g de cada idade (3 e 22 anos) foram utilizados em extrações separadas.

O processo de extração do RNA total teve início com a solubilização do macerado num tampão de extração (ANEXO C) e foi seguida por lavagens com clorofórmio:álcool isoamílico (24:1) até a obtenção de uma interfase límpida, que neste trabalho totalizou três lavagens. As amostras foram então precipitadas em 2X seu volume de etanol e 0,1X seu volume de NaCl (5M) para eliminação dos compostos fenólicos ligados ao PVP do tampão de extração. Outros contaminantes presentes como proteínas e carboidratos foram removidos em passos posteriores através da extração com fenol e precipitação em 0,25X seu volume de LiCl (10 M), respectivamente. Após a última precipitação, o precipitado foi ressuspensionado em água DEPC para

quantificação e avaliação de sua qualidade por espectrofotometria e eletroforese em gel de agarose 1,5%.

Quando este protocolo foi conduzido sem qualquer modificação, o RNA total obtido não apresentava qualidade suficiente para ser utilizado em técnicas moleculares, uma vez que quando submetido a uma reação de RT-PCR teste com o gene que codifica a UDP-glicose pirofosforilase, não gerava nenhum produto de amplificação (dados não mostrados). Originalmente, o protocolo recomenda a solubilização da amostra vegetal em 10 mL do tampão de extração, mas uma série de testes indicou um melhor rendimento das extrações quando as amostras foram solubilizadas em 20 mL. Além disso, os testes indicaram a presença de fortes inibidores de reação no RNA total, que poderiam ser provenientes de compostos fenólicos não eliminados durante sua extração e que podem ligar-se aos ácidos nucleicos, interferindo em reações enzimáticas. Assim, a quantidade de PVP (polivinilpirrolidona solúvel) e β -mercaptoetanol também foi dobrada.

A Figura 4a mostra o gel de agarose no qual pode ser observada a banda equivalente a 5 μ g de RNA total com seu padrão característico bem definido, não havendo sinal de degradação e sem a presença de impurezas ou DNA; a Figura 4b mostra a banda equivalente ao produto obtido da reação de RT-PCR utilizando o RNA total extraído com as modificações descritas acima já incorporadas ao protocolo original.

Figura 4 - Géis de agarose com bandas de RNA e do produto de RT-PCR. (a) RNA total; (b) produto da amplificação por RT-PCR, P = padrão de 1 Kb, I = produto de RT-PCR com o gene UDP-glicose pirofosforilase

A concentração e qualidade de todas as amostras de RNA total foram estimadas por espectrofotometria (Hitachi modelo U3300) nos comprimentos de onda 260 nm e 280 nm. A concentração de RNA das amostras foi estimada em função da equação determinada por Sambrook et al. (1989) como segue: $[RNA] = (A_{260nm} \times Fc \times Fd)/1000 \mu g/\mu L$, onde Fc representa o fator de conversão, que indica a equivalência de 1 unidade de absorvância no

comprimento de onda 260 nm correspondente a 40 µg RNA/mL, e Fd é o fator de diluição da amostra. A ausência de contaminantes nas amostras foi avaliada pela razão entre as absorbâncias ($A_{260\text{nm}}/A_{280\text{nm}}$). Quando a razão variava entre 1,7 a 2,0 o RNA obtido era classificado como de boa qualidade. Todas as amostras de RNA total obtidas foram separadas em alíquotas de 50 µg, precipitadas em uma solução de 1% acetato de sódio 2,5M e 2X o volume de etanol 100%, e armazenadas em ultrafreezer até sua utilização.

3.3 Construção da biblioteca SAGE

A biblioteca SAGE da região cambial de *Eucalyptus grandis* foi construída de acordo com as instruções fornecidas pelo *Kit* I-SAGETM versão E da Invitrogen, que são baseadas na metodologia de análise serial da expressão gênica (SAGE) desenvolvida originalmente por Velculescu et al. (1995). A metodologia descrita no *Kit* inclui algumas modificações que levam a um aumento da eficiência das várias reações enzimáticas necessárias ao processo e permite a construção da biblioteca a partir de quantidades menores de RNA. A Figura 5 apresenta um esquema geral ilustrativo da metodologia SAGE, que será descrita detalhadamente nas seções que seguem.

3.3.1 Isolamento do RNA mensageiro e síntese de cDNA

Um total de 50 µg de RNA total foi utilizado na construção da biblioteca SAGE. O primeiro passo da metodologia é o isolamento do RNA mensageiro através de partículas esféricas magnéticas associadas à *primers* oligo(dT) (desoxitimidina). Para isto, 100 µL da solução de partículas magnéticas foram transferidos para um tubo posicionado numa estante magnética onde permaneceram por 2 minutos. Isto faz com que as partículas fiquem concentradas na parede do tubo e o sobrenadante possa ser removido. Em seguida, foram adicionados 500 µL do tampão de lise/ligação (*Lysis/binding buffer* ANEXO D) às partículas, enquanto o RNA total também foi solubilizado em 1000 µL deste tampão. O tubo contendo as partículas magnéticas equilibradas no tampão de lise/ligação foi então novamente posicionado na estante magnética para a remoção do tampão e adição imediata da amostra de RNA total equilibrado no mesmo tampão. O RNA total misturado às partículas no tampão de lise/ligação foi mantido em agitação suave por 30 minutos, à temperatura ambiente, para promover a ligação das moléculas de RNAm da amostra aos oligos(dT) presentes nas partículas magnéticas.

Após a reação de ligação, o tubo foi novamente posicionado na estante magnética para retirada do tampão de lise/ligação e lavagem das partículas associadas ao RNAm. Foram realizadas duas lavagens com o tampão de lavagem A (*Washing buffer* A ANEXO E) e uma com o tampão de lavagem B (*Washing buffer* B ANEXO F). Em seguida, a amostra foi preparada para a síntese da primeira fita das moléculas de cDNA através de 4 passos de lavagens sucessivas com o tampão de síntese da primeira fita (1X *First strand buffer* ANEXO G). No quarto passo da lavagem, o sobrenadante só foi removido no momento em que a mistura de reação da primeira fita de cDNA (ANEXO H) já estava pronta para adição do RNAm.

As moléculas de RNAm ainda ligadas às partículas magnéticas foram então ressuspendidas na solução de reação e mantidas a 37°C por 2 minutos para pré-aquecimento. Em seguida, foram adicionados 3 µL da enzima transcriptase reversa *SuperScriptTM* II (ANEXO I) e o tubo foi mantido em banho maria a 37°C por mais uma hora. Para finalizar a reação de síntese da primeira fita de cDNA o tubo foi colocado no gelo por 2 minutos.

Para a síntese da segunda fita, 465 µL de água DEPC, 150 µL de tampão de síntese da segunda fita (5X *Second strand buffer* ANEXO J), 15 µL de dNTP *mix* (ANEXO K), 5 µL de *E. coli* DNA ligase (ANEXO L), 20 µL de *E. coli* DNA polimerase (ANEXO M) e 5 µL de *E. coli* RNase H (ANEXO N) foram adicionados à solução contendo a primeira fita. Esta reação foi incubada por 2 horas à 16°C e foi interrompida através da adição de 45 µL de EDTA 0,5 M. Com o tubo posicionado na estante magnética, as moléculas de fita dupla de cDNA permaneceram aderidas à sua parede, sendo possível a retirada da mistura utilizada para reação de síntese e a adição de 750 µL do tampão de lavagem C (*Washing buffer* C ANEXO O), pré-aquecido à 75°C para inativar a DNA polimerase. Após 12 minutos, o tampão de lavagem C foi descartado e uma nova lavagem com o mesmo tampão foi realizada. As moléculas de cDNA aderidas às partículas ainda foram lavadas 4 vezes sucessivas com 750 µL do tampão de lavagem D (*Washing buffer* D ANEXO P), e após a última lavagem, foram adicionados 200 µL de 1X tampão 4 (ANEXO Q). Esta amostra foi então transferida para um tubo novo e armazenada a 4°C por 16 horas antes de seguir para a próxima reação.

Figura 5 - Esquema ilustrativo da metodologia SAGE (Análise Serial da Expressão Gênica). O cDNA é sintetizado a partir de RNAm da amostra, ligado a esferas magnéticas; em seguida ele é digerido com a enzima *Nla III*, que reconhece e corta todo sítio CATG; a amostra é dividida em duas partes e dois adaptadores são ligados às extremidades coesivas geradas pela digestão anterior; uma segunda digestão com a enzima *BsmFI* é realizada. Esta enzima reconhece seu sítio na sequência dos adaptadores e cliva a molécula ~ 10 pb a jusante, liberando a cauda poli(A); as amostras são novamente unidas e ligadas para a obtenção de *ditags*; os dímeros são amplificados e novamente digeridos com a enzima *Nla III* para liberação da sequência dos adaptadores; as *ditags* são ligadas, formando um concatenado de 400 a 1000 pb; o produto da ligação é clonado e sequenciado

3.3.2 Digestão do cDNA com a endonuclease *Nla III* e ligação dos adaptadores

A amostra de cDNA solubilizada em 1X tampão 4 foi novamente precipitada com o auxílio da estante magnética e ressuspensa na mistura de reação de digestão contendo 172 µL de LoTE (ANEXO R), 2 µL de 100X BSA (ANEXO S), 20 µL de 10X *buffer* 4 e 6 µL de *Nla III* (ANEXO T). A reação foi mantida por 1 hora a 37°C. Com o auxílio da estante, as moléculas de

cDNA foram lavadas duas vezes com 750 μ L do tampão de lavagem C para inativar a enzima e quatro vezes com 750 μ L do tampão de lavagem D. A enzima *Nla* III reconhece os sítios CATGs presentes nas moléculas de cDNA e clivam sua fita dupla exatamente neste sítio (Figura 5).

Neste momento as moléculas de cDNA são compostas por sua cauda poli(A) e por um fragmento que se estende até o sítio CATG mais próximo de sua extremidade 3'. Para que esses fragmentos pudessem ser amplificados foi realizada a ligação de adaptadores às moléculas de cDNA. Para isto, a amostra de cDNA foi preparada para reação de ligação através de duas lavagens com 150 μ L do tampão de ligação 1X (ANEXO U) e homogeneização em 100 μ L desse mesmo tampão. Os 100 μ L foram divididos em duas partes e transferidos para dois novos tubos A e B, os quais foram mais uma vez para a estante magnética para o descarte do tampão. No gelo, a mistura de ligação dos adaptadores contendo 14 μ L de LoTE, 2 μ L do tampão para ligase 10X concentrado, 1,5 μ L do adaptador A (20 ng/ μ L) no tubo A e 1,5 μ L do adaptador B (20 ng/ μ L) no tubo B, foi adicionada. Os tubos foram aquecidos por 2 minutos a 50°C, resfriados por 15 minutos à temperatura ambiente e imediatamente colocados no gelo para a adição de 2,5 μ L de T4 DNA ligase (ANEXO V).

As amostras foram incubadas por 2 horas a 16°C e, em seguida, lavadas quatro vezes com 500 μ L do tampão de lavagem D. Após estas lavagens, elas ainda foram lavadas duas vezes com 200 μ L do tampão 4 1X concentrado, ressuspensas neste mesmo tampão e armazenadas a 4°C por 16 horas até o próximo passo. Os adaptadores, com aproximadamente 40 pb cada, ligam-se nas extremidades 5' das moléculas de cDNA, imediatamente antes do sítio CATG (Figura 5). Eles contêm extremidades coesivas de 4 pb complementares ao cDNA digerido com a *Nla* III, um sítio de reconhecimento para o Tipo IIS de enzima de restrição na extremidade 3' e um sítio para *primer* (PCR).

3.3.3 Digestão com a endonuclease *BsmF* I e formação dos dímeros de cDNA

Os tubos A e B, contendo os cDNAs ligados a adaptadores A e B respectivamente, foram posicionados no suporte magnético e o sobrenadante descartado. Em seguida, adicionou-se a mistura de digestão contendo 174 μ L de LoTE, 20 μ L do tampão 4 10X concentrado e 4 μ L de BSA 100X concentrado. As amostras foram pré-incubadas a 65°C por 2 minutos antes da adição de 2 μ L da enzima *BsmF* I e de sua incubação 65°C por uma hora. A digestão com *BsmF*I resulta na liberação dos cDNA ligados aos adaptadores das partículas magnéticas. Esta enzima liga-se ao

seu sítio de reconhecimento nos adaptadores e cliva o cDNA 10 a 14 pb a jusante, liberando um fragmento de aproximadamente 50 pb com 4 pb coesivas em sua extremidade 5'. Este fragmento consiste de 40 pb de sequência do adaptador e de 10 a 14 pb de uma sequência derivada de um único transcrito (Figura 5).

Após o término da digestão, os dois tubos foram posicionados na estante magnética e o sobrenadante, desta vez, foi recuperado em dois novos tubos. Um PCR-teste (ANEXO X) foi realizado com potenciais produtos da reação de digestão que possam ter permanecido ligados às partículas magnéticas para testar a eficiência da digestão com *BsmF* I. Como esperado, não houve sinal de banda no gel.

Esses fragmentos de cDNAs seguiram para uma purificação através da adição do mesmo volume de fenol:clorofórmio (1:1) e centrifugação em velocidade máxima por 5 minutos a temperatura ambiente. A fase aquosa de cada tubo foi transferida para dois novos tubos e precipitada com 133 µL de acetato de amônia 7,5 M, 3 µL de *mussel glycogen* e 1000 µL de etanol 100%. O passo de precipitação foi realizado em gelo seco por 10 minutos e o *pellet* recuperado após uma centrifugação a 16000 g, por 40 minutos, a 4°C. Após duas lavagens com etanol 70%, os *pellets* foram ressuspensos em 10 µL de LoTE e incubados a 37°C por 3 minutos para garantir a solubilização.

Em cada tubo, A e B, foram adicionados 5 µL do tampão *Klenow* (ANEXO Y), 1 µL de BSA 100X, 2,5 µL de uma mistura de dNTPs (10 mM cada), 30,5 µL de água DEPC e 1,0 µL da polimerase *Klenow* (ANEXO Z). Esta enzima tem a função de preencher as extremidades 5' dos fragmentos de cDNAs criadas pela digestão com a *BmsF* I. Os tubos foram incubados por 30 minutos a 37°C e seus volumes foram imediatamente reunidos em um único tubo, alcançando um volume final de 100 µL, aos quais foram adicionados 200 µL de LoTE. A mistura de cDNAs foi então purificada utilizando fenol:clorofórmio (1:1) e precipitada em acetato de amônia 7,5 M e etanol 100%, como descrito anteriormente. Durante a extração com fenol:clorofórmio, 100 µL da fase aquosa foram separados e a eles foram adicionados mais 100 µL de água DEPC para serem usados como controle negativo (sem ligase). O controle negativo também passou pelo processo de purificação com etanol como descrito acima.

Para a reação de ligação dos fragmentos, ou seja, formação dos dímeros, os *pellets* dos dois tubos (amostra e controle), provenientes da precipitação em etanol, foram ressuspensos em 1,5 µL de LoTE e incubados a 37°C por 15 minutos para solubilização, e usados diretamente na

reação de ligação para formar as *ditags*, ou dímeros. A mistura de reação de ligação continha 1,25 μL de Tris-HCl 3 mM pH 7,5, 0,75 μL do tampão de Ligase 10X concentrado, 0,75 μL de água DEPC e, no tubo contendo 1,5 μL da amostra, 1 μL de T4 DNA Ligase. No tubo contendo 1,5 μL do controle negativo foi adicionado mais 1 μL de Tris-HCl 3 mM pH 7,5 no lugar da T4 DNA ligase. Ambos tubos foram mantidos a 16°C *overnight*.

3.3.4 Otimização das reações de PCR para produção de dímeros em larga escala

Após o período de incubação, 14 μL de LoTE foram adicionados aos tubos, que seguiram diretamente para as reações de otimização de PCR. Para determinar a melhor diluição do produto proveniente da reação de ligação que resulte em quantidade de *amplicons* suficiente para o passo de polimerização das *ditags*, 3 diferentes diluições da amostra foram utilizadas. A diluição adequada foi obtida através de uma comparação entre a intensidade das bandas referentes aos três produtos de amplificação e de um controle positivo fornecido pelo *kit*, o qual continha uma concentração de 50-250 ng/50 μL , a concentração ótima do produto após 27 ciclos de amplificação. Assim, enquanto o controle negativo foi diluído 1/20, a amostra foi diluída em 1/20, 1/40 e 1/80. As reações de otimização foram realizadas seguindo a Tabela 2.

Tabela 2 - Otimização das reações de PCR a partir de diferentes concentrações da amostra

Regentes	Tubos					
	1	2	3	4	5	6
10X BV Buffer (ANEXO AE)	5 μL	5 μL	5 μL	5 μL	5 μL	5 μL
DMSO	3 μL	3 μL	3 μL	3 μL	3 μL	3 μL
dNTP mix (10 mM cada)	7,5 μL	7,5 μL	7,5 μL	7,5 μL	7,5 μL	7,5 μL
DTP-1 (<i>primer</i> de anelamento ao adaptador, 175 ng/ μL)	2 μL	2 μL	2 μL	2 μL	2 μL	2 μL
DTP-2	2 μL	2 μL	2 μL	2 μL	2 μL	2 μL
Água DEPC	29 μL	29 μL	29 μL	29 μL	29 μL	29 μL
Platinum®Taq DNA polimerase	0,5 μL	0,5 μL	0,5 μL	0,5 μL	0,5 μL	0,5 μL
Amostra 1/20	1 μL	-	-	-	-	-
Amostra 1/40	-	1 μL	-	-	-	-
Amostra 1/80	-	-	1 μL	-	-	-
Controle positivo (<i>Kit</i>)	-	-	-	1 μL	-	-
Controle negativo (sem Ligase) 1/20	-	-	-	-	-	1 μL

Tubos 1-3: diferentes diluições da amostra, 1/20, 1/40 e 1/80; tubo 4: controle positivo fornecido no *kit*; tubo 5: controle negativo (sem DNA) e tubo 6: controle negativo (sem ligase). Foram realizados 27 ciclos seguindo o programa: um ciclo de 2 minutos a 95°C; 27 ciclos de 30 segundos a 95°C, 1 minuto a 55°C e 1 minuto a 70°C; e 1 ciclo de 5 minutos a 70°C

Cinco μL dos produtos de PCR foram separados em gel de poliacrilamida 12% (ANEXO AA), corados com brometo de etídeo (0,5 $\mu\text{g/mL}$) e visualizados sob luz ultravioleta. O produto

de reação que mais se aproximou ao produto do controle positivo foi o obtido com a diluição 1/40 (segunda coluna no gel) (Figura 6).

Figura 6 - Gel de eletroforese dos produtos das reações de otimização da concentração de DNA para PCR. 1: diluição 1/20; 2: diluição 1/40; 3: diluição 1/80; P: padrão de DNA de 1 Kb; C: controle positivo do *kit*; N1: controle negativo (sem cDNA); N2: controle negativo (sem Ligase)

3.3.5 Amplificação em larga escala e purificação dos dímeros de cDNA

Para obtenção de grande quantidade de dímeros, 200 reações de PCR com volume final de 50 μ L cada foram realizadas. Cada reação continha 5 μ L de tampão BV 10X concentrado (ANEXO AB), 3 μ L de DMSO, 7,5 μ L de dNTP *mix* (10 mM de cada), 2 μ L de cada *primer* DTP-1 e DTP-2 (175 ng/ μ L em água DEPC para cada *primer*), 24 μ L de água DEPC, 0,5 μ L de *Platinum@Taq* DNA polimerase e 6 μ L da amostra diluída 1/40 vezes. O programa de PCR utilizado foi o mesmo descrito na Tabela 2. Após o término da amplificação, os produtos das 200 reações foram reunidos em um único tubo para subsequente purificação e precipitação dos dímeros com fenol:clorofórmio (1:1) e acetato de amônia 7,5 M/ etanol 100% conforme já descrito em outras etapas. O *pellet* resultante da precipitação foi ressuscitado em 300 μ L de LoTE e incubado por 10 minutos a 37°C.

Em seguida, foi realizada a purificação do volume total de dímeros utilizando 9 poços de um gel de poliacrilamida 12 %. Como os *primers* utilizados na amplificação eram específicos para os adaptadores A e B, somente os dímeros, ou *ditags*, formados por ambos foram amplificados. Isto garantiu que dímeros provenientes de ligações errôneas, como aqueles formados por apenas um dos dois adaptadores ou aqueles formados por apenas adaptadores, não fossem incluídos na amostra amplificada. Assim, o tamanho esperado dos dímeros obtidos das duas populações iniciais de cDNA (separadas inicialmente em A e B) deveria ser de

aproximadamente 102 pb (41 pb de cada adaptador + 10 pb de uma molécula de cDNA + 10 pb da outra molécula de cDNA). Na figura 7 é possível observar a presença de uma banda mais forte de aproximadamente 100 pb e outras duas mais fracas logo abaixo de aproximadamente 80 e 60 pb, correspondendo a produtos com falha na ligação dos adaptadores.

Figura 7 - Gel de eletroforese dos produtos da amplificação em larga escala dos dímeros de cDNA. Todas as bandas de 100 pb (círculo) foram excisadas do gel para geração das *ditags*. 1 – 9, Amostras de dímeros de cDNA; P, Padrão de DNA

Dessa forma, as bandas correspondentes a 100 pb foram recortadas do gel usando um bisturi e colocadas em 9 tubos de 0,7 mL (furados no fundo) encaixados em tubos de 1,7 mL e centrifugados a 14000 g por 3 minutos, o que levou ao rompimento da porção de gel em frações menores facilitando assim a eluição dos fragmentos de DNA. Os fragmentos de gel foram dissolvidos em 150 μ L de LoTE e acetato de amônia 7,5 M na proporção 125:25.

Após uma homogeneização, as amostras foram incubadas por 2 horas a 65°C. Após este período, as amostras foram combinadas três a três e passadas em uma coluna S.N.A.PTM através de uma centrifugação a velocidade máxima por 2 minutos. A solução eluída (aproximadamente 300 μ L) foi precipitada com acetato de amônia 7,5 M/ etanol 100% e os *pellets* resultantes foram ressuspensos em 14 μ L de LoTE, reunidos em um único tubo com volume final de aproximadamente 126 μ L, e armazenados a -20°C *overnight*.

3.3.6 Digestão dos dímeros de cDNA com *Nla* III

Como os adaptadores utilizados já exerceram suas funções, esta etapa da metodologia permite a liberação dos mesmos para obtenção das *ditags* propriamente ditas. Como eles foram

ligados através do sítio de reconhecimento da enzima *Nla* III, essa mesma enzima é utilizada para excluí-los.

Assim, os 126 μL de amostra foram divididos em três tubos, aos quais foi adicionada a mistura de reação de digestão contendo: 15 μL de 10X tampão 4, 2 μL de 100X BSA, 12 μL de *Nla* III e 79 μL de água DEPC. Os três tubos foram incubados a 37°C por 2 horas e seus volumes finais, após este período, foram ajustados para 200 μL com LoTE. Novamente foi realizada uma purificação e precipitação, agora das *ditags*, usando fenol:clorofórmio (1:1) e acetato de amônia 7,5 M/ etanol 100%, respectivamente, conforme já descrito em passos anteriores. Os *pellets* dos três tubos foram ressuspensos em LoTE e combinados em um único tubo com volume final de 32 μL . Ao final da digestão, espera-se que as *ditags* sejam compostas de 28 pb, equivalentes a duas *tags* de 10 pb mais duas seqüências referentes ao sítio CATG.

Foi realizada, então, uma separação do produto da digestão em gel de poliacrilamida 12%, onde os fragmentos de 28 pb puderam ser separados de outros produtos não digeridos. Para isto, 7 μL do tampão de amostra TBE (ANEXO AC) foram adicionados em cada tubo contendo 32 μL da solução de *ditags* e, todo volume foi aplicado em três colunas no gel (Figura 8).

Figura 8 - Gel de eletroforese dos produtos da digestão com *Nla* III. 1 a 3, produtos da digestão com *Nla* III; P, Padrão de DNA

Para excisão das bandas de 28 pb, o gel foi corado com brometo de etídio (0,5 $\mu\text{g/mL}$) e visualizado sob luz UV. Os fragmentos de gel foram quebrados em pequenos pedaços usando o mesmo procedimento já descrito para a purificação dos dímeros de cDNA, homogeneizados em uma solução contendo 150 μL de LoTE/ acetato de amônia 7,5 M (125:25) e incubados a 37°C *overnight*.

3.3.7 Polimerização/Concatenação e clonagem das *ditags*

Com a obtenção de uma população de *ditags*, que teoricamente representa a população inicial de RNAm, o último passo na construção da biblioteca SAGE é a concatenação desses dímeros para que possam ser seqüenciados.

Para isto, a solução eluída do gel foi purificada em colunas S.N.A.P.TM. O volume resultante foi dividido em alíquotas de 200 µL distribuídas em diferentes tubos para precipitação com etanol. A cada tubo foram adicionados 90 µL de acetato de amônia 7,5 M, 3 µL de *muscel glycogen* e 850 µL de etanol 100%. Após precipitadas, as amostras foram centrifugadas a velocidade máxima por 30 minutos a 4°C e os *pellets* resultantes foram lavados com etanol 70% e ressuspensos em um único tubo com 7,75 µL de LoTE. A este volume, foram adicionados, então, 1 µL do tampão de ligase 10X concentrado e 1,25 µL de T4 DNA ligase, que permaneceu incubado por 3 horas a 16°C. Após a incubação, foram adicionados 12 µl do tampão de amostra TBE 5X concentrado à mistura de reação que foi então mantida por mais 10 minutos a 65°C. Todo o volume da amostra foi submetido a eletroforese em gel de poliacrilamida 8% em um único poço. No gel já corado com brometo de etídio foi possível visualizar um rastro que se inicia com um produto de aproximadamente 100 pb indo até 1 Kb (Figura 9).

Figura 9 - Gel de eletroforese dos produtos da ligação das *ditags*. I, Concatenado das *ditags*; P, Padrão de DNA. Região selecionada: concatenados de 500 a 1000 pb

Como demonstrado na Figura 9, fragmentos contendo polímeros entre 500 pb - 1 kb foram excisados do gel e eluídos como descrito anteriormente para purificação dos dímeros de cDNA. Cada fração eluída do gel foi solubilizada em dois tubos contendo 200 µL de LoTE/acetato de amônia 7,5 M (5:1) e incubada a 65°C por 2 horas. O conteúdo dos dois tubos (de cada amostra) foi purificado em uma coluna S.N.A.P.TM, centrifugado por 2 minutos a

velocidade máxima, e a solução purificada foi dividida em volumes de 200 μL para extração e precipitação dos concatenados de *ditags*. Os passos de extração e precipitação foram realizados conforme já descrito em passos anteriores e o *pellet* resultante foi solubilizado em 6 μL de LoTE.

Antes da clonagem dos concatenados no vetor pZErO[®]-1, foi necessário fazer sua linearização através da enzima *Sph*I. Para isto 2 μL do vetor (1 $\mu\text{g}/\mu\text{L}$ ANEXO AD) foram adicionados à seguinte mistura de reação: 2,5 μL do tampão 2 10X concentrado (ANEXO AE), 1,4 μL da enzima *Sph* I (ANEXO AF) e 19,1 μL de água estéril. A solução foi incubada a 37°C por 30 minutos, e depois o vetor linearizado foi purificado e precipitado. O *pellet* resultante foi ressuspensionado em 60 μL de LoTE. Os 6 μL resultantes da concatenação das *ditags* foram colocados em contato com o vetor conforme mostrado na Tabela 3. Ainda, a reação de ligação foi realizada com dois controles negativos. Através da concatenação das *ditags*, é possível produzir grandes polímeros compostos de uma sequência de *ditags* intercaladas pelo sítio de clivagem CATG (Figura 5).

Tabela 3 - Reação de ligação do vetor aos polímeros de *ditags*

Reagentes	Amostra	Controle negativo (sem DNA)	Controle negativo (sem ligase)
Concatâmeros	6 μl	-	-
pZErO [®] -1/ <i>Sph</i> I	1 μl	1 μl	1 μl
Tampão de ligase 10X	1 μl	1 μl	1 μl
T4 DNA ligase	2 μl	2 μl	-
Água estéril	-	6 μl	8 μl

As três reações esquematizadas na Tabela 3 foram incubadas por 3 horas a 16°C. Após o término da reação o volume dos tubos foi ajustado para 200 μL , os quais foram purificados e precipitados usando fenol:clorofórmio (1:1) e etanol 100%. Seus *pellets* foram ressuspensionados em 12 μL de LoTE e utilizados na transformação das células de *E. coli* para construção da biblioteca SAGE.

Numa cubeta de 0,2 cm foram misturados 50 μL de células de *E. coli* eletrocompetentes *One Shot* TOP10 (Invitrogen) e 2 μL da mistura de ligação contendo o vetor com o inserto. Com um pulso elétrico de 2,5 kV por 4 mili-segundos (eletroporador: Micro Pulser[™] da Bio-Rad) aplicado a esta mistura, as células puderam ser transformadas. Após a eletroporação, 250 μL de meio SOC (ANEXO AG) foram adicionados à mistura, que foi mantida sob agitação constante (200 rpm) a 37°C por 1 hora. Após, mais 750 μL do mesmo meio foram adicionados e, a solução resultante foi dividida em volumes de 100 μL e distribuída em placas contendo meio LB baixo

sal sólido (ANEXO AH) com o antibiótico ZeocinaTM (50 µg/mL ANEXO AI) como agente de seleção.

As placas foram mantidas a 37°C *overnight* e no dia seguinte pôde-se observar um grande número de colônias somente nas placas com as amostras positivas (aproximadamente 150 colônias/placa). Como cada colônia representa um clone diferente, algumas colônias foram selecionadas para verificação da porcentagem de clones que continham um inserto e do tamanho dos fragmentos inseridos. Assim, um pontinho das colônias selecionadas foi retirado e aplicado diretamente na seguinte mistura de reação: 2,5 µL de tampão de PCR 10X, 0,75 µL de MgCl₂ 50 mM, 0,5 µL de cada dNTP, 1 µL de cada *primer* para o M13, 0,2 µL de Taq DNA polimerase e 17,55 µL de água, totalizando uma reação de 25 µL. A Figura 10 mostra o gel de eletroforese dos produtos amplificados, indicando que a maior parte deles contém insertos e que estes estão dentro do tamanho esperado. Na sequência, os clones foram encaminhados para o sequenciamento.

Figura 10 - Gel de agarose 1,5% para verificação dos insertos clonados. P, padrão de DNA

3.4 Sequenciamento dos polímeros/concatenados de *ditags*

A extração de DNA plasmidial e o sequenciamento dos plasmídeos recombinantes foram realizados em colaboração com o laboratório de Biologia Celular e Molecular do CENA-USP sob supervisão da Dra. Siu Mui Tsai. As extrações foram realizadas usando a metodologia *PureLink Plasmid Purification System* (Invitrogen) e o sequenciamento foi realizado através do sistema *BigDye® Terminator* usando o *ABI PRISM® 3100 Genetic Analyzer* (Applied Biosystems). No total foram sequenciados 736 clones da biblioteca SAGE construída a partir da região cambial de *E. grandis* em fase juvenil de desenvolvimento.

3.5 Bioinformática e análise dos dados

Após o sequenciamento, uma grande quantidade de dados é gerada na forma de cromatogramas, os quais são transformados em seqüências de nucleotídeos e estas passam por uma série de análises que permitem desde a extração e contagem das *tags* presentes na biblioteca até a sua associação a um EST. Assim, todos os critérios adotados na geração dos dados bem como na sua análise serão apresentados nas seções que seguem.

3.5.1 Análise da Qualidade do Sequenciamento

Os dados de sequenciamento gerados foram avaliados quanto a sua confiabilidade através do programa *Phred* (EWING et al., 1998). Este programa tem a capacidade de ler os sinais de fluorescência dos cromatogramas gerados pelo sequenciador e determinar as posições ideais dos picos de cada base. Depois, ele compara as localizações dos picos detectados com as dos picos idealizados, quantifica a qualidade do sinal (baseando-se na intensidade do sinal, espaçamento entre os picos e na presença de picos mal resolvidos em volta da base em questão) e atribui valores de qualidade às bases preditas nos arquivos de saída. O valor de qualidade está relacionado com a probabilidade de erro nas atribuições das bases através da fórmula: $q = -10 \times \log(p)$, onde p é a probabilidade de erro estimado (EWING et al., 1998). Para o presente trabalho foram utilizados os clones que apresentaram mais de 400 bases com qualidade igual ou superior a 20 que, segundo este programa, significa que 1 a cada 100 bases sequenciadas possa estar errada, ou que a probabilidade de atribuição de bases corretas é de 99%. Clones com seqüências de qualidade inferior foram re-sequenciados até a obtenção de seqüências com qualidade adequada. A Figura 11 mostra dois cromatogramas provenientes de um mesmo clone que apresentaram qualidades de sequenciamento bem diferentes. O primeiro continha picos muito mal definidos, que passaram a apresentar um padrão muito melhor após ser re-sequenciado.

3.5.2 Extração das SAGE *tags*

Para a extração das SAGE *tags* dos dados de sequenciamento, o kit I-SAGE fornece o *software* SAGE2000TM versão 4.5 (<http://www.invitrogen.com/sage>). Este programa importa os arquivos de seqüências com a extensão .seq, identifica os sítios de restrição da enzima *Nla* III (CATG), extrai as *tags* usando os sítios reconhecidos como referência e contabiliza a ocorrência de cada *tag* na biblioteca. Ao final, o programa lança um arquivo de saída ou *tag list* contendo as

sequências de 10 bases das *tags* seguidas do seu nível de expressão, ou seja, o número de vezes em que aquela *tag* foi encontrada na biblioteca (Figura 12).

Figura 11 - Cromatogramas de dois eventos de sequenciamento de um mesmo clone. (a) Cromatograma produzido por um sequenciamento de baixa qualidade; (b) Cromatograma referente ao re-sequenciamento do mesmo clone visto em (a), agora com alta qualidade

Figura 12 - Arquivo de saída do *software* SAGE2000™ (versão 4.5) ou *Tag list* da biblioteca da região cambial de *E. grandis*. A *Tag list* contém uma lista das *tags* sequenciadas na biblioteca com seu respectivo número de cópias

As *tags* lançadas nos arquivos de saída do *software* SAGE2000™ foram então utilizadas para encontrar as sequências de cDNA ou ESTs que provavelmente lhes deram origem e assim, estabelecer o perfil de expressão do tecido vegetal aqui estudado.

3.5.3 Identificação das SAGE *tags*

Somente as *tags* com número de cópias igual ou maior a dois foram utilizadas para a anotação de suas prováveis identidades, já que de acordo com vários autores, aquelas sequenciadas apenas uma vez (1 cópia) poderiam ser provenientes de erros de sequenciamento.

Um banco de dados local contendo ESTs e cDNAs completos de *Eucalyptus* foi construído. Grande parte dos ESTs desse banco de dados é proveniente do projeto Genolyptus, que foi lançado em 2002 como uma parceria entre 12 empresas do setor de papel e celulose, 7 Universidades, a Embrapa e o Ministério da Ciência e Tecnologia e levou ao sequenciamento de 70893 *reads*. No laboratório Max Feffer de Genética de Plantas foi realizada a montagem dessas sequências através do pacote de softwares phred/phrap/consed (www.phrap.org), resultando na obtenção de 23292 *contigs* representativos de amostras de folhas, plântulas, raízes, xilema, floema e flores das espécies de eucalipto *E. globulus*, *E. grandis*, *E. pellita* e *E. urophylla*. O restante é composto de ESTs e cDNAs completos de alguns genes de *Eucalyptus* disponíveis publicamente no GenBank (www.ncbi.nlm.nih.gov/). A maioria dos ESTs é proveniente de quatro bibliotecas previamente montadas e depositadas no GenBank, em formato fasta: biblioteca de cDNA de xilema em diferenciação de *E. gunnii*; biblioteca subtrativa de gravidade vs microgravidade de *E. globulus*; ESTs de bibliotecas subtrativas de madeira juvenil vs madura e madura vs juvenil de *E. globulus*; e biblioteca subtrativa de xilema vs floema de *E. gunnii* (nº de seus GIs no NCBI: 103484112-103472139, 84490923-84490662 e 51555734-5155511).

A busca por similaridade entre as SAGE *tags* e os ESTs foi feita localmente através da ferramenta de bioinformática BlastN, a qual compara sequências de nucleotídeos com uma base de dados também formada por sequências de nucleotídeos (ALTSCHUL et al., 1997). Como resultado é fornecida uma lista de ESTs que contêm a sequência usada para busca, nesse caso, as 4 bases do sítio CATG seguidas pelas 10 bases que compõem uma determinada *tag*. Quando essas buscas foram feitas, esta era a única informação apresentada como saída e era necessário analisar individualmente cada sequência de EST para verificar se ela continha cauda poli(A) e se haviam outros sítios CATG após a *tag*, já que o principal critério para identificação de uma

SAGE *tag* é sua presença imediatamente após o último sítio CATG da extremidade 3'. A partir do mês de março de 2007, com o auxílio de Livia Maria Franceschini (bioinformata do laboratório Max Feffer de Genética de Plantas), essa busca pôde ser automatizada. Foram criados mais campos para refinar a busca, os quais permitem escolher como referência somente ESTs que contenham ou não cauda poli(A), escolher o sítio desejado para a busca da *tag*, o da extremidade 3' ou em outros, e o banco de sequências a ser utilizado, Genolyptus ou de ESTs e cDNAs públicos. A Figura 13 mostra a sequência do EST encontrado para a *tag* mais expressa desta biblioteca.

```
>xylemEguniict980110
GCAATGATGTGGACAGGGTGCATGCTGTGGAGCCAGCGAAGTCTGCTTTTGAATTGAAGG
CCCAGATCCGCCGTGCCGTGGAGGCAGCAGGCATCCCTTACACCTACGTCCCATGCGGCT
GCTTCGCCGGCTACTTCCTCCCAACACTGGCGCAGCCAGGAGTCACTGCTCCTCCGAAGG
ACAAAGTCACCGTCATGGGTGACGGAAATGCAAAGGCAATTTTCAACAAGGAAGATGACA
TTGCGGCCTTCACCATCAAGGCTGTGGATGATCCGAGATCGCTGAACAAGATCCTTTACA
TCAGGCCTCCTAAGAACGTTTACTCATTCATGAGCTTGTTGCCTTGTTGGGAGAAGAAAA
TTGGCAAGACCCTCGAGAAGATTTACCTTCCTGAAGAGCAAATCCTGAAGCAAATCCAGG
AGTCCCCAATTCCCATCAATGTCATATTAGCAGTGAACCATTCAATCTTTGTAAAGGGCG
ACGGTGCCAATTTTGAGATCGAGGAGTCATTTGGTGTGCGAGGCTTCTGAGCTGTACCCAG
ATGTGAAGTACACTACAGTGGAAGAAATACCTCGAAAATTTTGTCTAAATTAAGGCCCATGC
GTCTCCTGTTCTTCAAGGAGTGAGTTACCGTGACTCGAGTGGACAGTTGATATGTATTAA
AAGGCTGTACACCTAAAGAATATCAAAGGTCACGGTCTTATTTAGAATTGTCTCTGATGT
CATATTCTTCTTGGTCTTCTTGGACATGTATTTGCTTTTCTTTGCTGTGGTATCCCATGAA
TTTCCCAGGGTTGTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

Figura 13 - Sequência de um EST utilizado na anotação de uma *tag*. Em negrito estão todos os possíveis sítios de restrição da enzima *Nla* III, a sequência da *tag* específica, utilizada no BlastN está sublinhada. Na extremidade 3' da sequência é possível também observar a presença de um terminal poli(A)

Depois de encontrado um EST representativo para uma determinada *tag*, o próximo passo foi utilizar sua sequência para busca por similaridades no banco de dados internacional NCBI através da ferramenta BlastX e do banco de dados nr, que compara sequências de nucleotídeos traduzidas em seus seis quadros de leitura possíveis com sequências de proteínas (ALTSCHUL et al., 1997). Como resultado, o BlastX libera uma lista de alinhamentos ordenados de acordo com o grau de similaridade encontrado entre a sequência inserida na busca e as proteínas. A medida de confiabilidade do alinhamento é dada pelos valores de *E*, que indicam a probabilidade de um dado alinhamento de pontuação igual ou maior ocorrer ao acaso numa dada base de dados. Ou seja, quanto maior o valor de *E* maior a probabilidade do alinhamento ser um produto do acaso (ALTSCHUL et al., 1997). Dessa forma, para este trabalho, apenas alinhamentos com valores de

$E \leq e^{-05}$ foram utilizados na anotação das *tags*. Hoje, o BlastX também está sendo feito automaticamente, indicando inclusive a direção do alinhamento da sequência de EST com a proteína encontrada.

Outras ferramentas da bioinformática também foram utilizadas para auxiliar a anotação das SAGE *tags*, entre elas, o ORF Finder (<http://www.ncbi.nlm.nih.gov>), que localiza na sequência de nucleotídeos (neste caso do EST) seus possíveis quadros abertos de leitura, traduz para proteínas e busca domínios funcionais (COG e Pfam) através do BlastP; e o ClustalW (<http://www.ebi.ac.uk/clustalw/>), utilizado principalmente na análise de diferentes *tags* anotadas que apresentaram a mesma função putativa para verificar a ocorrência de isoformas gênicas. Também foi construído um banco de *tags* “virtuais” extraídas após a obtenção de sequências complementares invertidas dos ESTs contidos no banco de dados local, para identificação de potenciais *tags* representando transcritos anti-sense. Foi realizada, ainda, uma busca por possíveis transcritos alternativos. Para isto, as *tags* que não puderam ser associadas a nenhum EST ou gene de *Eucalyptus*, normal ou anti-sense, foram buscadas em outros sítios CATG das sequências contidas no banco de dados local, que não o seu convencional, o mais próximo da extremidade 3'. Como as sequências que continham as *tags* “alternativas” também apresentavam uma *tag* em seu sítio CATG 3', foi verificado se essa segunda *tag* existia na biblioteca e, quando existisse, o resultado do BlastX deveria ser o mesmo para ambas.

Durante a anotação das *tags*, além da sua associação a uma provável função, foram realizadas também pesquisas no banco de dados KEGG e em referências bibliográficas para uma melhor interpretação de suas funções biológicas. Para a obtenção de uma análise funcional dos genes identificados na região cambial de *E. grandis*, eles foram classificados de acordo com suas funções em categorias, seguindo a classificação originalmente proposta por Rison et al. (2000) (Tabela 4), com pequenas modificações. As subcategorias Transporte de Vesículas e Fatores Transportadores foram incluídas na categoria Transporte, e a subcategoria Citoesqueleto, contendo outras duas subcategorias, foi incluída na categoria Estrutura e Organização da Estrutura.

Tabela 4 - Classificação Funcional de Rison et al. (2000). Os números representam as funções-chave ou principais

1 Metabolismo	4 Estrutura e organização da estrutura celular
1.1 Energia	4.1 Envoltórios celulares/Membranas
1.1.1 Metabolismo autotrófico	4.1.1 Membrana celular
1.1.2 Metabolismo de carbono	4.1.2 Parede celular
1.1.3 Transferência de energia/ATP	4.2 Matrix extracelular
1.2 Macromoléculas	4.2.1 Estruturas de superfície
1.2.1 Polissacarídeos/Lipossacarídeos	4.3 Ribossomos
1.2.2 Proteoglicanas/Glicoproteínas	4.3.1 RNAs ribossomais
1.2.3 Lipídeos/Lipoproteínas	4.3.2 Proteínas ribossomais
1.3 Pequenas moléculas	4.4 Citoesqueleto
1.3.1 Metabolismo de aminoácidos	4.5 Relacionados ao cromossomo
1.3.2 Metabolismo de Nitrogênio	4.6 Relacionados às organelas
1.3.3 Metabolismo de nucleotídeos	
1.3.4 Metabolismo de fósforo	5 Vias de Informação
1.3.5 Metabolismo de carboidratos	5.1 Relacionada ao DNA
1.3.6 Metabolismo de Lipídeos	5.1.1 Síntese e Replicação
1.3.7 Vitaminas/co-fatores/grupos protéticos	5.1.2 Restrição/modificação/reparo
1.3.8 Metabolismo Secundário	5.1.3 Recombinação
	5.1.4 Degradação
2 Processos	5.2 Relacionada ao RNA
2.1 Processos celulares	5.2.1 Síntese
2.1.1 Divisão celular	5.2.2 Transcrição
2.1.2 Tradução de sinais	5.2.3 Modificação
2.1.3 Direcionamento de proteínas	5.2.4 Degradação
2.1.4 Regulação celular	5.3 Relacionada a proteínas
2.2 Processos do organismo	5.3.1 Síntese de proteínas
2.2.1 Adaptação	5.3.2 RNAt
2.2.2 Respostas de proteção/detoxificação	5.3.3 Relacionada à tradução
2.2.3 Respostas a estímulos	5.3.4 Modificação/fosforilação
	5.3.5 Modificação/estrutura quaternária
3 Transporte	5.3.6 Degradação
3.1 Grandes moléculas	
3.1.1 Proteínas	6 Outros
3.1.2 Ácidos nucleicos	6.1 Não caracterizados/hipotéticos
3.2 Pequenas moléculas	6.2 Sequência Expressa/Sem anotação
3.2.1 Canais de íons/porinas	6.3 Sem registro
3.2.2 Transportadores de açúcares e carboidratos	
3.2.3 Transportadores de aminoácidos	
3.2.4 Transportadores de nucleotídeos/nucleosídeos	
3.2.5 Fatores Transportadores	
3.2.6 Transporte de Vesículas	

Todas as análises realizadas para a anotação de cada *tag* geraram grandes quantidades de informações, que foram armazenadas num banco de dados local na forma de “fichas eletrônicas”. Assim, cada *tag* foi registrada numa ficha que continha sua sequência, seu número de cópias, o tecido de origem, sua identificação pelo BlastX e ORF Finder, sua classificação funcional e a sequência do EST que a identificou (Figura 14). Para ter acesso a uma determinada ficha basta fazer uma busca com uma das seguintes opções: sequência da *tag*, biblioteca (existem outras

bibliotecas SAGE além da descrita neste trabalho), resultado do BlastX, domínio funcional ou categoria.

TAG: Delete: Tag Delete?

Registered Groups: Groups that TAG is registered:

Description:

Libraries where the Tag is found:

Blast Result:

Category:

Sequence:

Figura 14 - Ficha “eletrônica” do banco de dados local de genômica funcional de *Eucalyptus grandis*

3.6 RT-PCR em Tempo Real (qRT-PCR)

A quantificação do nível de expressão através de qRT-PCR foi utilizada em dois experimentos. O primeiro foi a validação dos dados obtidos pela biblioteca SAGE e, o segundo, foi a comparação do nível de expressão de alguns genes selecionados entre as regiões cambiais de árvores de *E. grandis* juvenis (3 anos) e adultas (22 anos). Para ambos os experimentos o RNA total foi extraído e o RNAm foi purificado, sendo a reação de RT-PCR conduzida em dois passos (síntese da primeira fita de cDNA separada da amplificação) utilizando *primers* específicos. O material e a metodologia utilizados serão descritos nas próximas sub-seções.

3.6.1 Seleção de genes e desenho de *primers*

Para o experimento de validação da biblioteca SAGE foram selecionados 14 genes com diferentes níveis de expressão (número de cópias) entre as *tags* (Tabela 5).

Tabela 5 - *Primers* específicos para os genes usados na validação da biblioteca SAGE por qRT-PCR. Para cada gene a tag que o representa, seu número de cópias e o *gi* da sequência utilizada como molde para construção de seu par de *primers* também são informados

Genes/tags	N° cópias	gi(NCBI)	Pares de <i>primers</i>
Glutamina sintetase (GS1) TATATCGTCG	66	103473664	L: CGTCAATTAGGGTTGGGAGA R: ATTCCCCACAAGGACAACAA
Cafeil-CoA O-metiltransferase (CCoAOMT) AACCTAGAAA	48	3319277	L: TCTTCCTGATGACGGAAAGATT R: ATAGGGTGTTCGTCGTAGCCAAT
S-adenosilmetionina sintetase (SAMS) GTTCGCCGCT	34	103482634	L: TGTCCGTCTTCGTTGATTCTTA R: GGCTAACCAATTACCCAATGAG
Sacarose sintase (SuSy) TGACTAACCT	29	80973755	L: GAAATCCTTCCACCCTGAGAT R: TTCATCTCAGACTGCTCTTCCA
Álcool cinamil desidrogenase (CAD) CCCTCTGTTG	24	2984652	L: GTTTGTGGTGAAAATCCCAGAT R: ATTGACTTCCTCCCAAGCATAA
Isocitrato desidrogenase (IDH) TACTCAGATG	20	1750379	L: CTGTTGAGTCTGGCAAGATGAC R: CATTTAATTCTCCCAACAAA
UDP-glicose pirofosforilase (UGPase) TCGGTCTTAT	15	103475065	L: TTCCAGTGAAAGCAACTTCAGA R: AAGATGCTTCAAACGGAATTGT
Piruvato quinase (PK) TCCTGGATAC	12	103478670	L: GTCAGTGGTGGTCCCTGTTT R: CCTTCACTTCACGATGCAGA
Pectinesterase GGAAAGGCAG	9	103474979	L: AGCTTCGCTTCACGAATTTATC R: ATCCGTCAAATTCATTCCATTC
4-coumarato: coenzima A ligase (4CL) CTATTGTAAT	8	73665528	L: ACCGCGAATACCATAGACAAAG R: ATTGCTTGATTTCGTCCTCAGT
UDP-xilose sintase 4 (UXS4) TCATTATCAA	5	103473775	L: GAGGAAGCCCGATATCACAA R: TGTGCCGGTACTGAAGATTG
UDP-arabinose 4-epimerase TATTGTAAAC	5	28395528	L: CACACCGGATGGTACTTGTG R: GTGTATTTCCGCCGTCCAGTT
Alcohol desidrogenase (ADH) TATTTCTG	3	103474114	L: GTGCCCCATAAAGATGCTGT R: TCCGCTTCTTTACTCCTCCA
Xilana 1,4-beta-xilosidase TAGCCAAACT	2	103474759	L: AGAAGGTGTTCTTGGCCTCA R: AGAGGGAGTTCTTGCGAACA

Para o experimento de comparação entre as regiões cambiais de árvores juvenis e adultas foram escolhidos 25 genes (mais um gene de referência) envolvidos na modificação da parede celular, biossíntese e deposição de parede celular secundária e lignificação (Tabela 6). A escolha dos genes foi baseada nas diferenças químicas e estruturais descritas entre esses dois tipos de madeira e, para ambos os experimentos, na disponibilidade de sequências públicas de ESTs, uma vez que para fins de publicação, as sequências de ESTs do banco de dados do Genolytus não poderiam ser utilizadas.

Tabela 6 - *Primers* específicos para os genes usados na comparação entre árvores juvenis e adultas por qRT-PCR. Para cada gene é apresentado também o gi da sequência utilizada como molde para construção de seu par de *primers*

Genes	gi(NCBI)	Pares de <i>primers</i>
Poligalacturonase	32163864	L: AACCTGACGAGCCACATGAC R: TGTGTGGTTAAGAAGCTTTTGG
Pectato Liase	88659681	L: ACACCAGCGGCATATCTAAT R: CGAATGTTTCTACAATGTCAAT
Pectinesterase	103474979	L: AGCTTCGCTTCACGAATTTATC R: ATCCGTCAAATTCATTCCATTC
Xiloglucana Endotransglicosilase 2 (XTH2)	70779682	L: CTACAGGAAGTTCCAGGCCACT R: TCTTTTTCTTCTCTAGAATCGC
Xiloglucana Endotransglicosilase 3 (XTH3)	70779684	L: TGCAAGGACACGGCGCGGTTCC R: GGAAGATATCAAAGAACACGGAGA
Xilanase	88659657	L: ATAATCCAGTGGGACCAATCAC R: TCTTTCCTGAGAGCCAGGTATC
UDP-glicose pirofosforilase (UGPase)	103475065	L: TTCCAGTGAAAGCAACTTCAGA R: AAGATGCTTCAAACGGAATTGT
UDP-xilose sintase 3 (UXS3)	103473755	L: CGGATATCTCCAAAGCCAAG R: GACGAGCACAGAATGGGATT
Arabinose 4-epimerase	28395528	L: CACACCGGATGGTACTTGTG R: GTGTATTTCGCCGTCCAGTT
Sacarose sintase (SuSy)	80973755	L: GAAATCCTTCCACCCTGAGAT R: TTCATCTCAGACTGCTCTTCCA
Endo 1,4- β glucanase (KOR)	70779690	L: CATTTCACAATCAGACCAGCAT R: TGTAGCTCATTTTCCGAGGATT
Celulose sintase 1	67003906	L: CGATCGCAGTGATCGATATG R: TCCAAGGTTGAAGATGGCGGCATT
Celulose sintase 2	67003908	L: ACATGTGATTGCTGGCCTTC R: TTCTGAGACATGAGCGATGA
Celulose sintase 3	67003910	L: CTGTGATTGCTGCCCCATGCT R: CGCCACCCTGTTCCATCAAA
α -Tubulina 2	103482259	L: TCTCCCTCATCACCATCCTC R: CTGGGTTCAAGTGCGGTATT
α -Tubulina 5	84143217	L: TTCAAATGCGGCATCAACTA R: TTCCTCCTCCTCATCACCAC
α -Tubulina 6	103475503	L: AAGCGTGACTTTGTGCACTG R: CACGCAGAGGTGTCTTTGAA
β -Tubulina 8	103481142	L: GTTGAATTGCCGATGAAGGT R: TGGATTTGCACCACTCACAT
β -Tubulina 3	103477591	L: ACCCAACCATGAAGAAGTGC R: TCCCTTCCCCTAAGGTGTCT
Fenilalanina amônia liase (PAL)	5690432	L: CCTGATCTCATCAAGGAAGACC R: GAGCAGGTCCCTCTCACAGTAG
4-coumarato: coenzima A ligase (4CL)	73665528	L: ACCGCGAATACCATAGACAAAG R: ATTGCTTGATTTCGTCTCAGT
Hidroxicinamil transferase (HCT)	84490724	L: CTCCAGCCTGATCTCTCAGC R: GGATTATCATTATATTGCCTAGC
Ferulato 5-hidroxilase (F5H)	6012982	L: AGTATGGCGGGATCTTCCAC R: CGGTCATACGTGAGGTAGCTG
Cafeil O-metil transferase (COMT)	437776	L: AAATCCTCATGGAAAGCTGGTA R: GTCGAAGTTGATCCCTTTTCATC
Cafeil-CoA O-metiltransferase (CCoAOMT)	3319277	L: TCTTCCTGATGACGGAAAGATT R: ATAGGGTGTTGTCGTAGCCAAT
Álcool desidrogenase	103474114	L: GTGCCCCATAAAGATGCTGT R: TCCGCTTCTTTACTCCTCCA

As sequências selecionadas foram inseridas no programa primer3 (Rozen e Skaletsky, 2000), disponível na Internet (<http://frodo.wi.mit.edu/cgi-bin/primer3/primer3>), para construção dos *primers* específicos. Todos os *primers* foram desenhados para que tivessem um tamanho entre 18 e 25 pb (tamanho ótimo = 20), uma temperatura de dissociação entre 57 e 63°C (t° ótima = 60°C) e que gerassem um produto entre 250 a 350 pb. As Tabelas 5 e 6 mostram os *primers* desenvolvidos para os genes usados na validação do SAGE e na comparação entre idades, respectivamente.

3.6.2 Isolamento de RNAm e Síntese da primeira fita de cDNA

As alíquotas de 50 µg de RNA total armazenadas no ultrafreezer foram utilizadas para a purificação de seus RNAm. Para a validação da biblioteca SAGE foi utilizado RNA total da mesma amostra usada para sua construção; já para o experimento de comparação entre idades foram usados RNA totais de outras amostras (árvores de 3 e 22 anos). Para o isolamento do RNAm, foi utilizado o *kit* Dynabeads[®] mRNA purification (Dynal Biotech), o qual é baseado numa separação magnética. Ele contém microesferas com Oligo (dT)₂₅ ligados a sua superfície que capturam e purificam poli(A) + RNA.

Cada 200 µL de microesferas são capazes de purificar 75 µg de RNA total, produzindo até 2 µg de RNAm. Assim, foram conduzidas várias purificações de 75 µg de acordo com as instruções do fabricante, que de maneira geral, segue o mesmo princípio do isolamento de RNAm descrito no primeiro passo da construção da biblioteca SAGE. Cada 50 µg de RNA purificado foram recuperados em 10,5 µL de Tris-HCl 10 mM, e 3,5 µL do RNAm resultante foram imediatamente usados nas reações de síntese da primeira fita de cDNA.

A primeira fita de cDNA de cada gene (Tabela 5 e 6) foi sintetizada através do *kit* SuperScript[™] One-Step RT-PCR Platinum[®] *Taq* (Invitogen), em uma mistura de reação contendo 6,25 µL de Tampão 2X concentrado (contendo 0,4 mM de cada dNTP e 2,4 mM de MgSO₂), 1,25 µL de cada *primer* (400 nM), 0,25 µL de RT/Platinum[®] *Taq* e 3,5 µL de RNAm, totalizando 12,5 µL de volume final. Os tubos com as misturas de reação foram colocados num termociclador com a seguinte programação: 1 ciclo de 50°C por 30 minutos e 94°C por 2 minutos para síntese de cDNA e pré-desnaturação da enzima transcriptase reversa; e 10 ciclos de 94°C durante 45 segundos, 55°C durante 1 minuto e 72°C durante 2 minutos e 30 segundos para amplificação dos fragmentos.

3.6.3 PCR em Tempo Real

A mistura de reação usada para qPCR continha 400 nM de cada *primer*, 12,5 µL de SYBR green (SYBR Green real-time RT-PCR SuperMix-UDG) e 3 µL de cDNA, num volume final de 25 µL. Para os 14 genes usados no experimento de validação do SAGE foram usadas três diluições (0,1, 0,01 e 0,001 vez) e três repetições e, para os genes usados no experimento de comparação entre idades foram necessárias quatro diluições (0,1, 0,01, 0,001 e 0,0001 vez) e três repetições. As reações foram realizadas no equipamento iQ5 (BioRad) com a seguinte programação: 1 ciclo de 95°C por 15 minutos para desnaturação da RT/Taq; 40 ciclos de 95°C por 20 segundos, 54° a 62°C (dependendo dos *primers*) por 30 segundos e 72°C por 30 segundos para amplificação dos fragmentos, com leitura da fluorescência em 72°C; 1 ciclo de 95°C por 30 segundos, 1 ciclo de 55°C por 30 segundos; e 1 ciclo para construção de curvas de dissociação dos *primers* com a temperatura se elevando de 55°C até 95°C e leituras da fluorescência a cada 1 minuto.

Os CTs médios (*Threshold cycle*) e as Eficiências de amplificação (*E*) de cada gene nas três diferentes amostras de cDNA (SAGE, 3 e 22 anos) foram obtidos com o auxílio do programa *Relative Expression Software Tool - 384 = REST-384* © - *version 1* desenvolvido por Pfaffl. As Eficiências (*E*) foram calculadas a partir da curva gerada pelas diluições dos cDNAs, enquanto os CTs foram obtidos com a média aritmética das três repetições de cada reação.

Um valor arbitrário de expressão *V* foi calculado para cada gene, nas diferentes amostras, através da equação proposta por Pfaffl (2001) que consiste de: $V = E^{-CT}$. Este valor de expressão *V* estimado, foi então normalizado com o gene UGPase (UDP-glicose pirofosforilase) no experimento de validação do SAGE e com um gene que codifica uma α -tubulina no experimento de comparação entre as idades. Estes genes foram escolhidos como genes normalizadores por serem aqueles que apresentaram a menor variação na análise feita pelo programa *NormFinder* (<http://www.mdl.dk/publicationsnormfinder.htm>) (dados não mostrados). No primeiro caso, os valores absolutos e normalizados de expressão para os 14 genes foram submetidos a uma análise de correlação (Pearson) com as frequências relativas de suas respectivas *tags* (TPT – *Tags per thousand*) observadas na biblioteca SAGE. Já para estimar a diferença no nível de expressão entre as árvores juvenis e adultas, os valores de expressão *V* normalizados, foram divididos entre si ($V_{\text{juvenil}}/V_{\text{adulta}}$ e $V_{\text{adulta}}/V_{\text{juvenil}}$) para obtenção de uma relação que indicasse quantas vezes um determinado gene é mais expresso numa idade que na outra.

4 RESULTADOS E DISCUSSÃO

4.1 Análise dos dados gerados pelo SAGE

Por ser uma técnica com a capacidade de resumir o transcrito de um organismo em sua forma mais simples (10 pb) e ser altamente sensível sem perder a especificidade, o SAGE além de gerar perfis de expressão gênica de forma qualitativa e quantitativa, numa investigação mais profunda ele também leva à identificação de novos genes, de genes com baixo nível de expressão, de transcritos alternativos e transcritos anti-sense. Assim, neste tópico, será realizada uma extensa investigação de todos os resultados que esta técnica pôde gerar para a biblioteca da região cambial de árvores juvenis de *Eucalyptus grandis*.

4.1.1 Dados Gerais

No total foram sequenciados 736 clones provenientes da região cambial que, após terem seus arquivos de cromatogramas processados pelo software SAGE2000, geraram uma biblioteca SAGE com 22660 *tags*. Apesar de todos os cromatogramas terem passado por uma análise no programa Phred, o que assegurou a obtenção de sequências com alto nível de qualidade, todas as SAGE *tags* que apresentaram somente uma cópia foram excluídas deste estudo, um procedimento comum na análise de bibliotecas SAGE uma vez que estas *tags* podem ser resultantes de erros de sequenciamento. Desta forma, 15385 *tags* com duas ou mais cópias que totalizaram 2942 *tags* únicas ou genes foram usadas para traçar o perfil transcricional da região cambial de árvores juvenis de *E. grandis* (Tabela 7).

Tabela 7 - Dados gerais obtidos pela biblioteca SAGE da região cambial de *Eucalyptus grandis*

Resumo dos dados	
Nº de clones sequenciados	736
Nº de total de <i>tags</i> extraídas	22660
Eficiência (Nº de <i>tags</i> / Nº de clones)	30,79
Nº de <i>tags</i> ≥ 2 cópias	15385
Nº de genes ≥ 2 cópias	2942
Nº de genes identificados	874
Nº de genes com função desconhecida (não caracterizados e sequências expressas)	295
Nº de <i>tags</i> sem homologia com ESTs de <i>Eucalyptus</i>	1773

Quando as sequências de 10 bases das SAGE *tags* antecedidas pelo sítio *Nla* III (CATG) foram contrastadas contra o banco de dados local (ESTs do Genolyptus e sequências públicas de

de *Eucalyptus*), 1169 puderam ser associadas a uma sequência de *Eucalyptus* e destas, 874 puderam ter uma provável função identificada através da ferramenta de busca BlastX do NCBI (Tabela 7). Das 295 *tags* restantes, 152 foram associadas a sequências de proteínas sem função conhecida e 143 não encontraram similaridade no NCBI. A Tabela 8 traz a relação das 50 *tags* mais expressas encontradas na biblioteca SAGE da região cambial de *E. grandis*.

A *tag* mais abundante nesta biblioteca corresponde ao RNAm do gene que codifica uma *allergenic isoflavone reductase*, enzima pertencente a uma família de redutases envolvida na síntese de fitoalexinas. As fitoalexinas, por sua vez, agem na proteção do tecido vegetal contra a invasão por fungos. Por outro lado, a sequência de EST correspondente a essa *tag* também apresenta alta similaridade ao gene *phenylcoumaran benzylic ether reductase* de *Populus*, que juntamente com outras redutases dependentes de NADPH formam a classe de enzimas responsáveis pela biossíntese de isoflavonóides e lignanas (SHOJI et al., 2002). A expressão da isoflavona redutase tem sido correlacionada positivamente com a expressão de enzimas-chave na biossíntese de lignina como PAL (Fenilalanina amônia liase), COMT (Ácido cafeico O-metiltransferase) e CCoAOMT (Cafeil Coenzima-A O-metiltransferase) (VINCENT et al., 2005) e com o aumento do crescimento secundário em árvores de *populus* transgênicas (ISRAELSSON et al., 2003).

A segunda *tag* mais expressa representa um EST com alta similaridade a uma proteína arabinogalactana de *populus Pop14A9* (Tabela 8). As AGPs (proteína arabinogalactana) são proteínas estruturais encontradas na parede celular primária e com função importante na mediação das interações célula-célula durante a diferenciação vascular (MAJEWSKA-SAWKA e NOTHNAGEL, 2000). Yang et al. (2005) analisaram a expressão de 11 AGPs e *AGP-like genes* de *Pinus taeda* em diferentes tecidos sob diferentes condições e mostraram que o padrão de expressão desses genes era variável, sugerindo diferentes funções para cada AGP. A AGP de *Pinus* similar à *Pop14A9* foi altamente e preferencialmente expressa no xilema em diferenciação quando comparada à sua expressão no floema ou nas folhas. Da mesma forma, o gene ortólogo de *Arabidopsis AtFLA12* foi expresso em caule e não em folhas, levando os autores a sugerirem que este gene desempenha uma função similar em diferentes espécies, especialmente durante a xilogênese.

Tabela 8 - Relação das 50 *tags* mais expressas e suas prováveis identidades. À sequência segue seu número de cópias, o gene que ela corresponde (quando existir), o número de identificação do gene no ncbi e o valor *E* com que o gene foi encontrado

TAG	nº cópias	Identificação Putativa (Blastx)	gi	E
AATTTCCAG	321	Allergenic isoflavone reductase-like protein	10764491	2e-145
TACTTATTTG	160	Arabinogalactan protein Pop14A9	5919185	2e-50
CAGTTAAAAG	145	Tag sem homologia com ESTs de Eucalyptus	-	-
ATTAGTGGTG	137	PIP3 (PLASMA MEMBRANE INTRINSIC PROTEIN 3)	15236267	2e-148
GGATTGCATA	117	ADP-ribosylation factor 1	56311394	9e-100
TGTGTGTTTG	100	EST sem homologia com o ncbi	-	-
CTGGGGGGTT	91	EST sem homologia com o ncbi	-	-
AAATCACAGG	87	Unknown protein	21536535	9e-17
TTGTGGATTG	82	Beta-tubulin 1	23452315	0.0
AATTGCTATC	78	Tubulin alpha-6 chain (TUA6)	2244853	2e-58
GAAGAATCGA	71	Harpin-induced 1	92898246	1e-47
GTATTTCGAGT	70	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic	120676	3e-21
TATATCGTCG	66	Glutamine synthetase, cytosolic	13924490	9e-20
GTGTCTCCAG	61	Hypothetical protein	84579145	3e-09
AGACGGAATA	60	Adenosine kinase	92880884	2e-113
GGAACCTCCG	60	Profilin (Allergen Fra a 4)	85701214	2e-65
TTCGGTTTCC	59	Tag sem homologia com ESTs de Eucalyptus	-	-
AGGTTTCTTG	56	Putative membrane protein	50920439	1e-50
CTAATCTATC	53	Wheat adenosylhomocysteinase-like protein	29367605	1e-154
AGCAGTTTAA	52	Tag sem homologia com ESTs de Eucalyptus	-	-
ATACACGCGA	51	Tag sem homologia com ESTs de Eucalyptus	-	-
TGTGCTGTAG	50	CDC2 protein kinases	3482933	3e-74
ATAAGAGTTT	49	Translation initiation factor	15236989	4e-55
AACCTAGAAA	48	Caffeoyl-CoA O-methyltransferase	1934859	3e-138
TGTTTGTTAA	48	Tag sem homologia com ESTs de Eucalyptus	-	-
TTGGCTTGTT	47	Unknown protein	34913256	1e-29
CTCCTCGTTG	46	Cytochrome b5 isoform Cb5-C	50844677	8e-62
TAAGTACCGG	44	Putative ripening-related protein	7406663	5e-53
AGAAATGTGA	42	Tag sem homologia com ESTs de Eucalyptus	-	-
ATTGCTGTT	42	Blue copper protein	21553614	8e-35
GCCGTTCTTA	42	rRNA intron-encoded homing endonuclea	13171103	1e-25
TTTGGAATTT	42	Fiber protein, Fb34	38016521	9e-71
CTCTTCCGGG	41	5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase	8134569	0.0
GTGTCGGGTT	38	Eukaryotic initiation factor 5A (2)	19702	8e-83
GGATAACATC	37	Hypothetical protein	84579145	5e-13
AAAGCGATGT	36	Auxin-repressed protein	71040687	1e-31
AGGCTCTTTA	36	Mob1/phocein family protein	32470635	1e-118
GAATCAAAAT	36	Cellulose synthase 1	67003907	2e-144
GAACGCTATC	35	Cysteine protease	2677828	2e-170
TAAGCTATCA	35	17.9 kDa heat-shock protein	11990130	6e-23
ATGTTTGAGT	34	Tag sem homologia com ESTs de Eucalyptus	-	-
CTTGCTTGCG	34	Tubulin beta chain	34909054	2e-139
GTTCGCCGCT	34	S-adenosyl methionine synthase-like	78191442	6e-84
TGAGATTGAT	33	EST sem homologia com o ncbi	70878528	1.3
ATAACTTGAC	32	EST sem homologia com o ncbi	41946856	1.5
ATCACCAATG	32	Ubiquitin; Apoptosis regulator Bcl-2 protein, BAG	92881534	9e-75
ATGCAATACA	32	Sem homologia com ESTs de Eucalyptus	-	-
GGTGTGTTGGA	31	RAN2; GTP binding	15241245	4e-73
ACCAGGGCCG	30	Fasciclin-like AGP 13	47717929	4e-26
TTGTTTTTCT	30	Hydroxymethyltransferase	2244749	0.0

Além das duas primeiras *tags* terem sido identificadas como genes envolvidos no metabolismo de parede celular, entre as *tags* mais expressas também foram identificados genes relacionados diretamente com biossíntese e organização de parede celular secundária, entre eles uma celulose sintase, uma cafeil-CoA O-metiltransferase e três tubulinas, como também genes envolvidos no crescimento (proteína reprimida por auxina), divisão celular (CDC2 proteína quinase) e morte celular programada (cisteína protease), evidenciando a identidade do tecido coletado para construção desta biblioteca.

O passo seguinte à busca por homologia e consequente atribuição de funções às *tags* únicas (genes) foi a classificação de cada gene em categorias de acordo com sua funcionalidade. A Figura 15 mostra tanto a porcentagem de *tags* únicas (genes) que puderam ser identificadas, quanto a porcentagem que estas representam em cada Categoria Funcional. Mesmo utilizando bancos de dados específicos para a espécie, com grande quantidade de ESTs provenientes de tecido cambial ou xilemático, 60% das 2942 *tags* únicas não puderam ser associadas a um EST ou gene de *Eucalyptus* conhecido. Um dos possíveis motivos é que, apesar de um grande número de ESTs presentes no banco de dados do Genolyptus, ele foi construído a partir de várias bibliotecas de diferentes tecidos, mas que como em outros bancos de dados de ESTs, não consegue representar ESTs de baixa expressão. Outro possível explicação é que a cauda poli(A) dos ESTs pode não ser sequenciada adequadamente, inviabilizando a identificação das *tags*.

Figura 15 - Classificação funcional das 2942 *tags* únicas (genes) da biblioteca da região cambial de árvores de *Eucalyptus grandis*

A origem das *tags* não anotadas vem sendo muito debatida na literatura, uma vez que na maioria dos trabalhos envolvendo a técnica SAGE mais da metade das SAGE *tags* incluem-se nesta classificação. Wang (2007), em sua revisão a respeito dos dados produzidos pelo SAGE, propõe duas explicações para a origem das *tags* não anotadas: primeiro, as *tags* não anotadas poderiam ser resultantes de um acúmulo de erros de sequenciamento, mas este problema afetaria principalmente *tags* com baixo número de cópias e a probabilidade do mesmo erro ocorrer na mesma *tag* em clones diferentes é praticamente nula; segundo, as *tags* não anotadas poderiam compreender *tags* verdadeiras representando transcritos ainda desconhecidos mas detectados pelo SAGE devido à sua alta sensibilidade. Esta segunda explicação é baseada na ocorrência de muitas *tags* com baixo nível de expressão que podem ser associadas a um transcrito ou gene conhecido e no fato de que os dados do SAGE têm sido gerados a partir de sequenciamentos com alto controle de qualidade.

A análise desta biblioteca mostra um resultado semelhante ao discutido acima, onde 546 SAGE *tags* representadas por 3 e 2 cópias puderam ser associadas a um ESTs ou gene de *Eucalyptus*. Este valor compreende 46,7% de todas as *tags* anotadas na biblioteca e 29,15% de todas as *tags* contidas neste intervalo de expressão.

Numa análise preliminar da distribuição dos genes nas cinco Categorias Funcionais, foi observado que, de maneira geral, as células amostradas para a construção desta biblioteca estão voltadas para a produção de RNAs mensageiros de forma quase igualitária em quatro das cinco categorias (Figura 15). A categoria Metabolismo e Energia foi representada pela maior quantidade de genes, um resultado esperado, já que estes genes são responsáveis pelo metabolismo básico da célula e por sua manutenção. A segunda categoria mais abundante foi a de Estrutura e Organização da Estrutura, que entre outras subcategorias contém aquelas responsáveis pela biossíntese de Parede Celular e componentes do Citoesqueleto. Na sequência, aparecem as categorias responsáveis pelas Vias de Informação e Processos Celulares, as quais contém genes de grande importância na regulação celular, tradução de sinais e regulação da expressão gênica. Já os genes envolvidos no transporte celular, como transporte de grandes e pequenas moléculas, transporte vesicular e fatores de transporte, aparecem com menor representatividade, o que não significa uma menor importância dessa classe para as células estudadas neste trabalho. Para uma melhor compreensão da atuação destes grupos de genes e suas interações, estas categorias serão analisadas de forma mais abrangente nas próximas seções.

4.1.2 Caracterização dos genes por níveis de expressão

Como já observado por inúmeros trabalhos desenvolvidos com a técnica SAGE, este não foi diferente e apresentou um padrão de expressão onde poucos genes foram encontrados com alto número de cópias e a grande maioria, com baixo número de cópias. A questão que surge é: quais são os genes ou, os grupos de genes (Categorias Funcionais) expressos nestes diferentes níveis e qual a consequência disto?

Para responder a estas perguntas, os genes foram divididos em três intervalos arbitrários de expressão de acordo com seu número de cópias. Aqueles que apresentaram número de cópias acima de 20, foram classificados como tendo alto nível de expressão; aqueles com número de cópias entre 20 e 5, como nível intermediário de expressão; e aqueles com número de cópias entre 5 e 2, como baixo nível de expressão. Ainda, dentro de cada intervalo, os genes foram separados de acordo com as Categorias Funcionais com o objetivo de analisar a proporção dessas categorias em cada intervalo de expressão (Figura 16). O intervalo superior foi escolhido de forma que pelo menos um gene fosse representado em cada categoria, e o intervalo inferior foi determinado para o número de cópias a partir do qual ocorria uma inversão nos níveis de expressão das categorias Processos Celulares e Estrutura e Organização da Estrutura. É importante ressaltar que esta classificação dos genes em intervalos de expressão, como também os limites dos intervalos usados foram obtidos para a biblioteca em questão, não podendo ser extrapolada para outras análises, uma vez que foi baseada especificamente em seu tamanho e no número de cópias apresentadas pelas *tags* (ou genes).

A Figura 16 mostra os perfis obtidos para as categorias funcionais em cada intervalo de expressão onde nota-se claramente que ocorrem mudanças no perfil obtido em função da classe de expressão considerada. Entre os genes altamente expressos a categoria funcional mais abundante é a de Estrutura e Organização da Estrutura com 32,89% dos genes, seguida por Metabolismo e Energia com 30,04% dos genes presentes neste intervalo, enquanto que as outras categorias estão representadas com um número bem menor de genes. Estes dados são consistentes com o material vegetal amostrado para construção da biblioteca SAGE, uma vez que a diferenciação nessas células deve exigir a ação de muitos genes ligados a processos de metabolismo geral, que normalmente são mais expressos, e à biossíntese de parede celular, os quais devem estar fortemente ativados para o aumento da biomassa das árvores.

Entretanto, conforme o nível de expressão vai diminuindo, ou seja, quando genes com menores números de cópias são analisados - grande vantagem oferecida pela técnica SAGE - verifica-se que esta situação se inverte e as categorias Processos Celulares e Vias de Informação tornam-se mais abundantes que a de Estrutura e Organização da Estrutura (Figura 16).

Figura 16 - Distribuição dos genes de acordo com sua categoria funcional em três níveis de expressão: alto (n° de cópias ≥ 20), intermediário ($20 < n^{\circ}$ de cópias > 5 cópias) e baixo ($5 \leq n^{\circ}$ de cópias ≥ 2)

Entre os genes com baixo nível de expressão classificados em Processos Celulares destacam-se as calmodulinas, proteínas 14-3-3, nodulinas, proteínas induzidas por Harpin e genes que respondem a fitormônios. Calmodulinas são proteínas capazes de transferir Ca^{2+} para outra proteína alvo. O Ca^{2+} é um mensageiro essencial para o início de mudanças fisiológicas em resposta a diversos estímulos em células eucarióticas e sua ligação em uma proteína resulta em mudanças conformacionais que modulam sua atividade ou sua habilidade de interagir com outras proteínas (BOONBURAPONG e BUABOOCHA, 2007). Kaplan et al. (2006) analisaram os tipos de mudança transcricionais que ocorrem rapidamente em resposta a um aumento no Ca^{2+}

citossólico. Os principais genes ativados foram fatores de transcrição, como aqueles que respondem a etileno, fatores do tipo zíper de leucina e dedo de zinco, MYB e B3, e outras proteínas sinalizadoras, como RAB-GTPases e proteínas induzidas por fitormônios. Os fatores de transcrição MYB foram relacionados por alguns autores (PAUX et al., 2005; DELUC et al., 2006) como reguladores da lignificação, e têm sido encontrados numa série de trabalhos (KARPINSKA et al. 2004; PAUX et al. 2005).

As proteínas 14-3-3 também são de grande importância na transmissão de sinais mas, diferente das calmodulinas que causam mudanças diretas nas proteínas-alvo, elas parecem atuar sobre proteínas já fosforiladas, tendo a função de completar a modificação da proteína-alvo (FERL, 2004). Sua função na região cambial ainda não foi descrita, mas estas proteínas também foram encontradas por Celedon et al. (2007) no mesmo tecido utilizado para construção desta biblioteca, sugerindo que elas estejam relacionadas a processos de desenvolvimento e/ou diferenciação celular. Já as nodulinas, proteínas conhecidamente envolvidas com nodulação em raízes, foram encontradas por Moreau et al. (2005) em fibras mortas de *Populus* desempenhando um possível papel na sinalização inicial da morte celular das fibras de madeira. Da mesma forma, proteínas induzidas por Harpin têm sido relacionadas como sinalizadores em processos de resposta hipersensível (RH), quando da invasão do tecido por um patógeno ou durante a senescência do tecido (TAKAHASHI et al., 2004), ambos envolvendo modificações na parede celular e processo de morte celular programada.

Ainda, na categoria Vias de Informação foi encontrada uma grande quantidade de fatores de transcrição e genes envolvidos na modificação pós-traducional, como também na degradação de proteínas, com número de cópias entre 5 e 2. Entre eles estão fatores de transcrição LIM, que exercem papel regulatório na biossíntese de lignina (HUANG et al., 2007), *heat shock proteins*, chaperonas moleculares responsáveis pelo dobramento de proteínas que podem levar a diferentes conformações e conseqüentemente a uma variedade de atividades e especificidade, e proteínas da via de ubiquitinação e cisteína proteases, envolvidas na degradação de proteínas e na morte celular programada. A grande diversidade encontrada neste grupo de genes pode explicar a baixa correlação geralmente verificada entre os transcritos e proteínas identificados num mesmo material biológico. O coeficiente de correlação entre os níveis de expressão obtidos através do SAGE e os volumes das proteínas extraídas do caule de *E. grandis* com seis meses de idade, calculado por Carneiro (2007), também foi muito baixo ($R^2 = 0,1037$).

De maneira geral, o que estes resultados mostram é que existem grupos de genes cuja atividade requer um alto nível de expressão e grupos para os quais poucas cópias são suficientes. A maioria dos autores que utiliza a técnica SAGE para traçar perfis transcricionais deixa de explorar o intervalo de dados dos genes com baixos níveis de expressão para dispensar maiores esforços na caracterização dos genes mais expressos da biblioteca. No entanto, a análise desses genes revelou informações importantes a respeito da regulação das células da região cambial do Eucalipto e mostrou que os genes ou grupos de genes com baixos níveis de expressão não são menos importantes que aqueles com altos níveis. Ao contrário, é a categoria funcional a que eles pertencem que não necessita de um grande número de moléculas de um mesmo transcrito e sim de uma variedade de genes que desencadeiam diferentes respostas de regulação, restritas em função do espaço/tempo, e que por isso, geralmente têm um tempo de vida útil muito curto, fato que durante a construção da biblioteca SAGE torna seus RNAs mensageiros menos disponíveis para amostragem.

4.1.3 Identificação de transcritos anti-sense

Pesquisas recentes têm sugerido que RNAs anti-sense podem realizar um papel regulatório de forma semelhante aos miRNAs (REINHART et al., 2002) e a técnica SAGE tem se mostrado eficiente em sua detecção. Robinson et al. (2004) observaram 259 genes com transcritos anti-sense em *Arabidopsis* através do SAGE, dos quais 164 foram comprovados por outras técnicas, microarranjos de DNA (ou cDNA) e MPSS (*Massively Parallel Signature Sequencing*). A busca por transcritos anti-sense não somente leva a um aumento no número de *tags* identificadas, mas indica quais genes podem estar sendo regulados negativamente no genoma de um organismo.

A origem dos transcritos anti-sense pode ocorrer de três maneiras, segundo Luther (2005). A primeira considera a possibilidade da transcrição de um pseudo-gene, forma inativa de um gene verdadeiro, que pode ter duas variantes: um arranjo invertido de *introns* e *exons* na fita 3' → 5', ou um arranjo normal na fita 5' → 3'. A segunda é a transcrição do gene a partir de sua fita oposta e, a terceira, é a transcrição de um transcrito normal por uma RNA polimerase dependente de RNA, já no citoplasma (Figura 17).

Figura 17 - Origem de transcritos anti-sense. a) transcrição de dois tipos de pseudogenes A e B; b) transcrição do gene a partir da fita oposta do gene; c) transcrição de um transcrito sense por uma RNAPolimerase dependente de RNA, no citoplasma. Retirado de Luther, 2005

As *tags* referentes a transcritos anti-sense presentes neste estudo foram anotadas quando as 14 bases (CATG + *tag*) se alinhavam positivamente a um EST de *Eucalyptus* e seu alinhamento com a proteína no BlastX era invertido (*frame* = -); ou quando as 14 bases se alinhavam de maneira invertida a um EST mas seu alinhamento no BlastX era positivo (*frame* = +). No primeiro caso, a anotação foi feita de forma direta, já que o banco de ESTs utilizado contém esse tipo de sequência em seus dados. Já no segundo caso, todas as *tags* com número de cópias maior ou igual a 2 e que ainda não tinham sido identificadas (não foram associadas a um EST) foram comparadas com todas as *tags* que poderiam ser provenientes de um EST anti-sense (*tags* virtuais extraídas de sequências de ESTs invertidas). Entretanto, as *tags* detectadas dessa maneira foram consideradas como potenciais anti-sense, já que seus verdadeiros ESTs não estão presentes no banco de dados.

Assim, 51 *tags* puderam ser detectadas usando o primeiro critério de análise, e 62 *tags* usando a segunda forma de análise. A Tabela 9 contém as *tags* que puderam ser identificadas da forma direta e que alinharam com uma sequência conhecida de proteína com $E\ value \leq e^{-5}$, enquanto que na Tabela 10 estão listadas as *tags* que foram identificadas através de um EST anti-sense “virtual”, também com $E\ value \leq e^{-5}$. Como pode ser observado, a maioria das *tags* anti-sense (reais e potenciais) foi detectada entre aquelas com baixo número de cópias, um resultado esperado já que moléculas anti-sense realizam funções regulatórias nas células. Nas duas tabelas (9 e 10) é possível notar a presença de diferentes transcritos anti-sense para genes que desempenham alguma função nos processos que levam à formação da madeira como modificação da parede primária, citoesqueleto e lignificação, evidenciando que esses processos sofrem um tipo muito refinado de regulação, ainda pouco conhecida. Outra observação interessante foi que vários anti-sense apresentaram similaridade a proteínas de função ainda desconhecida ou nem puderam ser associados a qualquer sequência protéica depositada nos bancos públicos (dados não mostrados), o que representa uma ótima oportunidade para a identificação de novos genes que necessitam deste tipo de regulação na região cambial de *Eucalyptus grandis*.

Tabela 9 - Tags associadas a transcritos anti-sense verdadeiros identificadas de forma direta, como descrito no texto

Tag	Nº cópias	Identificação (Blastx)	gi	Frame
GATCTCCGAT	25	Dirigent-like protein pdir10	88771143	-2
CCCTATGGAT	24	Emp24/gp25l/p24 family protein	15223027	-2
GGGAATATGC	12	Ribosomal protein S3	108802679	-2
CGGGCCGATC	10	Cytochrome P450 like	1545805	-1
TITTCACTCA	5	Plasma membrane H ⁺ -atpase	31580857	-1
TAGACCCATA	4	Unnamed protein product	116060704	-3
AGAATCGATC	3	Hypothetical protein	41052959	-1
GCGTTGCCCC	3	Zinc-binding protein	18410510	-3
GGTTAGTCGA	3	Cytochrome P450 monooxygenase	38640722	-2
CTAGCTTTGA	2	Hypothetical protein	140065720	-3
ATCAAGAAGT	2	Heat shock protein binding	145340151	-3
GACTGTATAA	2	Porin	5031279	-2
TGGGAGACAG	2	Hypothetical protein	116060704	-2
GTGTCCAACA	2	Unknown protein	42572449	-2
TAAAAAGTAC	2	Ring domain containing protein	41059804	-2

Tabela 10 - Tags associadas a potenciais transcritos anti-sense identificadas de forma indireta, como descrito no texto

Tag	Nº cópias	Identificação (Blastx)	gi	Frame
GCGTTGCAAA	45	NADH-ubiquinone oxidoreductase-related	18396430	+1
AAAAGTGAAA	44	Fasciclin-like AGP 15	47717933	+1
TCACATTACA	9	ADP-ribosylation factor	39653273	+1
GCAACTCCTG	7	Trans-cinnamate 4-hydroxylase	67983271	+3
AACAAATGGC	6	Germin-like protein	2961201	+2
GTCAGTGCCC	6	Hypothetical protein	147777142	+1
TAAAGAAGCT	5	Hypothetical protein	144923180	+3
AGATTTACTT	4	Phenylalanine ammonia lyase	5690433	+2
TCGTCCGGTC	4	Glutamine synthetase	2811030	+2
GAAACAAGAA	4	Extensin-like protein	21360370	+1
GTGTTCTGT	4	Putative ADP,ATP carrier-like protein	84468306	+2
ATCGTGTCAA	3	Spl15 (spirall-like5)	18416160	+3
CAAATGCCAC	3	ADP-ribosylation factor-like protein	89475524	+1
CACTGCCATA	3	Actin depolymerizing factor 2	89212812	+1
GCCAAAAAAA	3	Remorin (pp34)	34925093	+1
TCTGGCGGTG	3	Extensin like protein	124484377	+1
AATTCTTCCA	3	Heavy metal transport/detoxification protein	92898246	+2
AAACAGAGAT	2	Mitochondrial processing peptidase beta	12802327	+3
ACCAACTTGT	2	Cobalamine-independent methionine synthase	974782	+3
AGCGGCAGGC	2	Dtdp-glucose 4-6-dehydratase	6469141	+1
AGGAACCAGG	2	Hypothetical protein	125536996	+1
AGTAGGAATC	2	Thioredoxin h-like protein	24637231	+1
CAGATCTCGA	2	Caffeoyl-coa O-methyltransferase 2	30580381	+2
TCTCTCTCTC	2	Aluminum-induced protein	56606534	+1
TGAACCCGCC	2	Cytokinin binding protein CBP57	2160322	+1
TGGAAGTGCA	2	Fasciclin-like AGP 13	47717929	+2
TTGGAGGCTG	2	Putative glutamine synthase	20269121	+2
AAGTGATTCT	2	Alpha tubulin	56481443	+1
ACAAAATCTC	2	Histone H2A	124360158	+3
ACATCAAAGC	2	Hypothetical protein	125559328	+2
ACATCTTAAA	2	Ribosomal protein S5	92891212	+2
CATCAGTTCT	2	KED (wound-responsive gene)	8096269	+3
CATCCGAAAA	2	Hin1 like protein	42557351	+2
CCACTCCCGA	2	NC domain-containing protein	18415178	+3
CTAATGTATA	2	Hypothetical protein	147826550	+2
GGGCTTGGAT	2	Putative senescence-associated protein	13359451	+1
GTATCTCAGA	2	Ribosomal protein S5	92891212	+1
GTTCTGAAGG	2	Coatomer subunit beta-1	75283240	+1
TGAGCAACAA	2	Protein kinase	92871857	+3
TGGACAATCA	2	Glyceraldehyde-3-phosphate dehydrogenase	106879589	+2
TTCTGCAGCC	2	Sar1 (synaptobrevin-related protein 1)	79324082	+1

4.1.4 Identificação de transcritos alternativos

Ao contrário das pesquisas envolvendo a regulação da expressão gênica através da produção de transcritos anti-sense, o *splicing*, ou processamento alternativo do RNAm nos organismos superiores vem sendo discutido há algum tempo pelos pesquisadores, os quais têm demonstrado tratar-se de um processo de grande importância, principalmente no que diz respeito

a geração de variabilidade. Segundo Robinson et al. (2004), a precisão com que um gene é identificado através do SAGE permite que *tags* sejam também associadas a genes únicos quando estão presentes em outras posições, ou sítios CATG que não o último da extremidade 3' de um transcrito e, dessa forma, a técnica tem o potencial de identificar eventos de processamento diferencial não caracterizados, aumentando o número de *tags* que podem ser anotadas.

Assim, todas as *tags* com número de cópias maior ou igual a 2, que não foram associadas a nenhum EST através da forma direta de anotação e que não foram identificadas como transcritos anti-sense, foram usadas na busca por potenciais transcritos alternativos. Mesmo durante o processo normal de identificação das *tags* já foi possível observar, através de alinhamentos, a presença de vários transcritos alternativos derivados de processamento diferencial da extremidade 3' do RNAm, uma vez que as sequências contidas no banco de dados local construído para esta análise possuem este tipo de informação.

Através da busca por *tags* presentes em outros sítios que não o último da extremidade 3', foi possível a localização das mesmas em 239 sequências do banco de dados local, sendo que em vários casos, uma mesma *tag* foi encontrada em sítios diferentes, de sequências diferentes. Assim, os 239 potenciais transcritos alternativos foram representados, na verdade, por 182 diferentes *tags*. Para garantir que uma mesma *tag*, presente em posições diferentes de duas ou mais sequências, representasse o mesmo gene, BlastX foi realizado e, somente as sequências que levavam a um mesmo resultado foram consideradas. Dessa forma, das 182 diferentes *tags* com transcritos alternativos, 117 puderam ser associadas a uma proteína conhecida através do BlastX ($E\text{ value} \leq e^{-5}$).

Um exemplo da identificação de dois transcritos alternativos para um mesmo gene através de duas sequências de ESTs, contendo a mesma *tag* em sítios diferentes, está demonstrado a seguir. Analisando as sequências é possível notar que, na primeira, a *tag* alvo (em vermelho) está localizada em seu sítio convencional, enquanto na segunda, existe um fragmento a mais, o qual contém mais dois sítios CATG, fazendo com que a *tag* alvo seja encontrada numa posição mais distante da extremidade 3'. Além disso, é interessante observar que, neste caso, a sequência do gene parece conter um poli(A) interno (em azul), o que deve permitir um processamento alternativo de seu RNAm, levando a produção de dois transcritos diferentes.

TAG: TAATGTCAGC

>xylemEguniICT983243

CGGATCTCAACAGCATCACCCGAGTCTTCAGCTTCGACTACAAGCCTCTCATCGCTTACGAATCCATAAACGATACGG
 GCATGTTCTACGGAATGAAGTTTTACAATGATCTACTGATGGAAGCAGGGCCAGAAGGAAATGTACAGTCCGAAGTGC
 TACTTCAGAAAGACAAGAACACATTTACATTCAAACAAGGATGGGCATTTCCAAGAAAAGTATACTTCAATGGAGATG
 AATGTATGCTTCCCCACCTGATCAATACCCATTTCTGCCGAAGTCCGCCCAAGTGAATCTAGTAGCCTTGTCAACAC
 TCGTGACCTCCATGGCTATGTCGATACTCCTTATTCACAAATCTCCATAGAAGCAGATGCATTTTCGGATTTTTTTTAA
 AAGCAAGGCCATATGCAAGCAATTTAATACGTACGCTGCCCTCATCAGGTAAATATCATACAACTTTCCGGACCT
 CTCCAGACACGGCTTT**CATGTAATGTCAGCAAAAAAAAAAAAAAAAAA**

BlastX: ref|NP_197067.2| COBL4/IRX6 (COBRA-LIKE4); hydrolase 184 3e-45

>xylemEguniICT985517

CGGCAGTCCGAAGTGCTACTTCAGAAAGACAAGAACACATTTACATTCAAACAAGGATGGGCATTTCCAAGAAAAGTA
 TACTTCAATGGAGATGAATGTATGCTTCCCCACCTGATCAATACCCATTTCTGCCGAAGTCCGCCCAAGTGAATCTA
 GTAGCCTTGTCAACACTCGTGACCTCCATGGCTATGTCGATACTCCTTATTCACAAATCTCCATAGAAGCAGATGCAT
 TTCGGATTTTTTTTTTAAAGCAAGGCCATATGCAAGCAATTTAATACGTACGCTGCCCTCATCAGGTAAATATCATAC
 AAATTTCCGGACCTCTCCAGACACGGCTTT**CATGTAATGTCAGCAAAAGAAAAAAGAA**CGTGCTTGTAAAAAGAA
 TCATGCAAAAGCTTGCTCTCAAGTAGTAACAGTTTTGGGTGTCTGTGT**CATG**CAATACTGAGTACCACTGCTTCAACT
 GTCCATTTCTATATTCTGTATGGTTTGACAAAAAAAAAAAAAAAAAAAA

BlastX: ref|NP_197067.2| COBL4/IRX6 (COBRA-LIKE4); hydrolase 106 1e-21

A Tabela 11 contém os dados referentes às 10 *tags* mais expressas que representam potenciais transcritos alternativos e que puderam ser associadas a uma provável função.

Tabela 11 – Relação das 10 *tags* mais expressas da biblioteca SAGE associadas a potenciais transcritos alternativos, seus números de cópias e as posições em que se encontram nas sequências de ESTs que as identificaram

Tag	Nº cópias	Posição*	Identificação (BlastX)	gi
TACTTGAGCT	25	3	Wound induced protein-like	37724583
TTGTTAGTGC	11	2	Cytokinin binding protein CBP57	2160322
TCGAGCTACT	10	2	ATP synthase epsilon chain, mitochondrial	15217996
TTGAAGTCGA	10	2	Fiber protein Fb15	30841448
GGCTGTATTA	9	2	NADPH:quinone oxidoreductase	15232856
TATTIGCTTT	9	2	Allergenic isoflavone reductase-like	10764491
TGTTGTGGGC	7	2	Adenylate kinase	92873599
ACCCATTTCC	6	3	Zinc finger, RING-type	92885424
GTCAGTCCCC	6	2	Ubiquitin; Apoptosis regulator Bcl-2 protein	92881534
ACTGACTTGG	6	2	Secretory carrier membrane protein	15222550

* o número das posições é referente à posição do sítio CATG que a *tag* se encontra em relação à cauda poli(A) na sequência do EST

Os transcritos alternativos encontrados diretamente durante a anotação normal das *tags* e que codificam genes importantes para o processo de xilogênese serão discutidos em seções posteriores.

4.1.5 Perfil transcricional global da região cambial de *Eucalyptus grandis*

As análises discutidas nos tópicos anteriores basicamente abordaram os dados de forma qualitativa, entretanto a principal característica do SAGE é a quantificação do nível de expressão de um grande número de genes simultaneamente. Assim, além de serem classificados em Categorias Funcionais, todos os genes encontrados nesta biblioteca também foram divididos em subcategorias e, além da quantidade de genes presentes em cada uma, o número total de transcritos (número de cópias) que eles somavam também foi considerado, o que possibilitou uma comparação da expressão entre grupos de genes funcionalmente relacionados, principalmente quando a relação número de transcritos/gene foi avaliada.

Os genes contidos nas categorias *Metabolismo e Energia*, *Transporte* e *Estrutura e Organização da Estrutura* apresentaram maiores relações número de transcritos/gene (Tabela 12). Numa análise qualitativa preliminar onde foram considerados somente os genes identificados nesta biblioteca (Figura 15), a categoria *Transporte* teve a menor representatividade entre as categorias, mas, quando o nível de expressão desse grupo de genes foi levado em conta, ficou claro que eles possuem um grau de importância similar às categorias *Metabolismo e Energia* e *Estrutura e Organização da Estrutura* (Tabela 12), uma vez que, de acordo com a relação número de transcritos/gene, as células estão produzindo quantidades muito próximas de RNAm destes grupos de genes funcionalmente distintos.

Esta análise também mostrou que as diferenças nos níveis de expressão nas subcategorias foram maiores que aquelas observadas entre as Categorias Funcionais principais.

Dentro da categoria *Metabolismo e Energia* as subcategorias *Metabolismo de Nitrogênio* e *Metabolismo Secundário* destacaram-se por apresentar os maiores número de transcritos/gene (Tabela 12). Apesar dos poucos genes encontrados nesta classe, a Glutamina sintetase (GS1) parece ser de grande importância para as células no processo de formação da madeira estudado neste trabalho, já que duas GS1 e um precursor de GS foram responsáveis por 74,75% do número total de cópias incluídos nesta subcategoria. Dentro de *Metabolismo Secundário* quatro Isoflavonas redutases foram observadas, sendo que três delas tinham baixos níveis de expressão. Entretanto a quarta – o gene mais expresso da biblioteca – representou sozinho 86,75 % do total de cópias nesta subcategoria.

Tabela 12 - Perfil transcricional da região cambial de *E. grandis* baseada em Categorias e Subcategorias Funcionais

Categorias Funcionais	Genes	Σ do nº de transcritos (cópias)	nº de transcritos/ Gene
1. Metabolismo e Energia			
1.1 - Metabolismo Autotrófico e Energia (ATP)	95	758	7,98
1.2 - Metabolismo de Carbono	2	15	7,50
1.3 - Metabolismo de Carboidratos	38	255	6,71
1.4 - Metabolismo de Nitrogênio	6	99	16,5
1.5 - Metabolismo de Aminoácidos	36	341	9,47
1.6 - Metabolismo de Lipídio	23	121	5,26
1.7 - Metabolismo de Pequenas Moléculas	8	81	10,12
1.8 - Metabolismo Secundário	10	370	37,00
Total	218	2040	9,36
2. Processos Celulares			
2.1 - Divisão Celular	15	170	11,33
2.2 - Tradução de Sinais	44	242	5,50
2.3 - Regulação Celular	53	305	5,75
2.4 - Respostas de Proteção / Destoxificação	46	348	7,56
2.5 - Respostas a Estímulos	16	111	6,94
Total	174	1176	6,76
3. Transporte			
3.1 - Grandes Moléculas	19	80	4,21
3.2 - Pequenas Moléculas	20	250	12,50
3.3 - Vesicular	40	427	10,67
3.4 - Kinesinas / Fatores de Transporte	6	42	7,00
Total	85	799	9,40
4. Estrutura e Organização da Estrutura			
4.1 - Envelope Celular/ Membrana	12	68	5,66
4.2 - Parede Celular	74	969	13,09
4.3 - Citoesqueleto	33	482	14,60
4.4 - RNAs/Proteínas Ribossomais	82	451	5,50
4.5 - Cromossomos	4	14	3,5
Total	205	1984	9,68
5. Vias de Informação			
5.1 - Metabolismo de DNA (síntese/ degradação/ modificação)	7	32	4,57
5.2 - Metabolismo de RNA (síntese/ degradação/ modificação)	50	271	5,42
5.3 - Metabolismo de PROTEÍNA (síntese/ degradação/ modificação)	135	934	6,92
Total	192	1237	6,44
TOTAL IDENTIFICADO	874		

A categoria *Processos Celulares* foi principalmente representada por genes relacionados com a Divisão Celular, que revelou a maior relação número de transcritos/gene da categoria (Tabela 12). Este resultado é um forte indicativo de que, além das células de xilema em diferenciação, a coleta do material vegetal permitiu também a amostragem de células meristemáticas em processo de divisão. Em adição, os dados sugerem ainda que a divisão celular nesta região deve estar ocorrendo numa alta taxa. Por tratar-se da região cambial de árvores em fase juvenil de desenvolvimento, este resultado era esperado uma vez que a alta taxa de crescimento é uma das principais características de árvores jovens. Cato et al. (2006) avaliaram,

entre outras características, a taxa de crescimento radial durante a formação da madeira de *Pinus radiata* na base (madeira adulta) e no ápice (madeira juvenil) das árvores. A taxa de divisão celular encontrada pelos autores foi 3,3 vezes maior no ápice que na base das árvores. Entre os genes com maiores níveis de expressão identificados nesta biblioteca estão duas proteínas kinases CDC2, três TCTP (*Translationally controlled tumour protein*) e quatro Skp1 (*suppressor of kinetochore protein 1*). CDC2 são kinases dependentes de ciclinas que desempenham função regulatória-chave para a entrada da célula em Mitose e foram descritas preferencialmente expressas no meristema cambial de *Populus* (HERTZBERG et al., 2001). Já as Spk1 são proteínas do centrômero requeridas para o progresso do ciclo celular (CONNELLY e HIETER, 1996) e, apesar de ter sido sugerido que proteínas TCTP sejam importantes no crescimento e divisão em sistemas mamíferos (BOMMER e THIELE, 2004), sua exata função no sistema vegetal ainda é desconhecida. Em ervilhas, a expressão deste gene foi observada em células da ponta de raízes em divisão, sugerindo que estes genes, assim como seus homólogos em mamíferos, estejam envolvidos em proliferação celular (WOO e HAWES, 1997). Assim, a destacada presença de TCTPs nesta biblioteca sugere que estas proteínas também tenham importância na região cambial de *Eucalyptus grandis*.

Os genes relacionados ao Transporte de Pequenas Moléculas e Transporte Vesicular apresentaram os maiores níveis de expressão da categoria *Transporte* (Tabela 12). Entre os genes envolvidos no Transporte de Pequenas Moléculas as MIP (*Membrane Intrinsic Proteins* ou Proteínas Intrínsecas de Membrana) foram as mais abundantes em tipo e quantidade de transcritos, representando 78% do número total de transcritos presentes nesta subcategoria. Essas proteínas intrínsecas de membrana facilitam o transporte de água e solutos neutros através da camada dupla de lipídios e acredita-se que são importantes no processo de divisão e expansão celular e para propriedades do transporte de água em resposta a condições ambientais (BANSAL e SANKARARAMAKRISHNAN, 2007). Ainda, as aquaporinas, um tipo de MIP, podem funcionar no transporte de longa distância no xilema e floema (CHAUMONT et al., 1998). Outro tipo de Transporte fortemente representado nesta biblioteca é o de Vesículas. As *GTP-binding proteins* foram as mais expressas entre os genes identificados desta subcategoria, sendo representadas por diferentes membros dessa família como *Ras/Ran* envolvidos em proliferação celular, *Rab* e *Arf* em tráfico de membranas e *Rac* relacionado com polimerização de actinas (VERNOUD et al., 2003). Esta subcategoria tem ganhado atenção dos pesquisadores que

estudam o processo de formação de madeira uma vez que os polissacarídeos da parede celular sintetizados no complexo de Golgi e os complexos protéicos envolvidos na síntese de celulose são exportados através de vesículas (ODA et al., 2005).

A categoria funcional mais expressa e também a de maior importância para a biossíntese de madeira, a categoria *Estrutura e Organização da Estrutura*, foi principalmente representada por genes responsáveis pela organização do citoesqueleto e pelo metabolismo da parede celular (Tabela 12). Entre os genes classificados na subcategoria Citoesqueleto foram observadas 4 α -tubulinas e 8 β -tubulinas responsáveis por 64% da quantidade de transcritos presentes nesta subcategoria. O citoesqueleto regula a polaridade, a morfologia e a movimentação intracelular por estar envolvido em diversos eventos, como tráfico de vesículas, movimento de organelas, respostas a estresses bióticos e abióticos, transmissão de sinais e biossíntese de parede celular. A deposição padronizada da parede celular secundária é um dos principais eventos do desenvolvimento das células de xilema, o qual é precisamente e dinamicamente regulado pelo citoesqueleto (ODA e HASEZAWA, 2006). Assim, ao mesmo tempo em que seus componentes atuam no transporte do complexo de biossíntese de celulose, eles também agem diretamente na deposição de celulose, uma vez que são encontrados logo abaixo da membrana plasmática em associação à esse complexo (BARNETT e BONHAM, 2004).

Os genes envolvidos na biossíntese da parede celular também apresentaram um maior nível de expressão. Nesta subcategoria estão incluídos genes que codificam as principais enzimas envolvidas na síntese de celulose, hemicelulose e lignina, enzimas de modificação da parede celular, e para proteínas estruturais da parede. As outras subcategorias contidas em *Estrutura e Organização da Estrutura* apresentaram níveis bem menores de expressão (número de cópias/gene) (Tabela 12) quando comparadas ao Citoesqueleto e à Parede Celular, resultado esperado já que os genes pertencentes a estas categorias são responsáveis pelo principal processo realizado por células de xilema em diferenciação. Dessa forma, foi criado um tópico específico para discutir a expressão dos genes envolvidos na formação da madeira através das principais vias bioquímicas que levam à xilogênese.

De acordo com a Tabela 12, a categoria com menor nível de expressão foi a de *Vias de Informação*, onde suas subcategorias apresentaram níveis de expressão igualmente baixos quando comparados aos observados nas outras subcategorias já discutidas, com uma média de 5,63 transcritos/gene. Este resultado complementa a discussão gerada pela seção anterior a respeito

dos genes com baixo número de cópias sugerindo que estes genes são principalmente voltados para regulação celular e regulação da expressão gênica.

4.2 Perfil transcricional durante a Xilogênese

Com a construção da biblioteca da região cambial de *Eucalyptus grandis* em fase juvenil de desenvolvimento foi possível traçar um perfil global da expressão dos genes identificados, o qual forneceu importantes informações a respeito do metabolismo geral dessas células, do metabolismo específico que as conduz à formação da madeira e como todo esse sistema pode ser regulado.

A condição perene que caracteriza as árvores é garantida pela renovação regular do xilema e floema funcionais através de processos específicos de diferenciação das células a partir do câmbio vascular. Nesta seção os dados da biblioteca SAGE serão usados para traçar os perfis transcricionais de cada fase da xilogênese, desde a divisão e extensão das células, deposição de parede celular secundária e lignificação, até a morte celular programada dos elementos de vaso e fibras.

4.2.1 Divisão Celular

O início da diferenciação que dá origem a células funcionais de xilema ocorre com a divisão das células do câmbio, que do lado de fora produzem células-mãe do floema e, do lado de dentro, células-mãe do xilema. As células-mãe, por sua vez, dividem-se gerando diferentes células derivativas que, no xilema, dão origem às células do parênquima radial e longitudinal, às fibras e aos vasos. Este padrão de desenvolvimento, ou seja, os tipos e a quantidade de células produzidas, é controlado principalmente pela auxina, que promove a divisão e a diferenciação celular (TAIZ e ZEIGER, 2004).

O gradiente de auxina (longitudinal e radial) formado na região cambial responsável por essa padronização ocorre devido ao seu transporte polar através das células do câmbio e do parênquima radial (TAIZ e ZEIGER, 2004; UGGLA, et al., 1996). A transmissão de sinais de auxina inicia-se com a ativação de fatores de transcrição ARF (*Auxin Response Factor* ou Fator de Resposta à Auxina). Na ausência de auxina, esses fatores formam heterodímeros com proteínas Aux/AIA; na presença de auxina essas proteínas são alvo de um complexo SCF e são degradadas através da via de ubiquitinação. Assim, os fatores de transcrição ARF são liberados

e ativam genes de resposta à auxina, relacionados ao crescimento e desenvolvimento (GRAY, 2004) (Figura 18).

Apesar de sua função na inibição da via de sinalização da auxina, o gene Aux/AIA também é um dos genes ativados pelo ARF ativo, juntamente com o gene SAUR e GH3, os quais levam à expressão de outros genes responsáveis diretamente pela resposta à presença de auxina (TAIZ e ZEIGER, 2004). A análise dos dados gerados pela biblioteca SAGE identificou a presença de dois genes Aux/AIA e dois de SAUR, todos em baixo nível de expressão. Também foram identificados um gene induzido pela auxina e um promotor de crescimento, com níveis similarmente baixos de expressão. Por outro lado, foram encontrados quatro genes reprimidos pela auxina (ARP) apresentando maior quantidade de transcritos (Tabela 13). Uma vez que ARPs são reprimidos pela presença de auxina e estes foram encontrados em altos níveis, as células da região cambial coletadas para construção desta biblioteca provavelmente estavam recebendo baixas concentrações do fitormônio. Uggle et al. (1996) analisaram a distribuição radial de auxina em árvores de *Pinus sylvestris* e verificaram que esta concentração diminui do câmbio para as zonas tardias de desenvolvimento do xilema, alcançando níveis mínimos a partir da zona onde os traqueídeos sofrem deposição da parede celular secundária. Como a camada de células em divisão celular é bem mais fina que as camadas seguintes de diferenciação, os RNAs mensageiros dessas camadas devem estar presentes em maior quantidade na amostra utilizada para este estudo, o que explicaria a baixa abundância observada para os genes da via de sinalização de auxina.

Tabela 13 - Genes envolvidos com respostas a fitormônios encontrados na biblioteca de madeira juvenil, suas SAGE tags e seus números de cópias. Números entre parênteses são os respectivos GIs que os identificam no NCBI

Tag	Nº cópias	Identificação Putativa	gi
AAAGCGATGT	36	Auxin-repressed protein - ARP	71040687
AGAGGCTGCT	18	Auxin-repressed protein like-protein - ARP2	13430203
AGAGGCTGGT	14	Auxin-repressed protein-like protein - ARP2	62526583
CTTGCAGCA	8	Auxin-repressed protein-like protein - ARP1	62526583
CAATATGTGG	5	Auxin-induced putative CPI2 domain-containing	55736085
GCCCTATTAG	5	Auxin growth promotor protein	15144512
AGTCTTTGTT	4	Aux/IAA protein	20269051
AAGAGTTGCG	2	Aux/IAA protein	29465672
GTCGAGTTCA	3	Auxin responsive SAUR protein	87240820
TGAATATAGT	3	Auxin responsive SAUR protein	92897061
TTAGTGAGTC	3	Ethylene response factor) - ERF	62526569
CCAGAGCTTC	2	Abscisic acid-induced-like protein	21537153
GCAAGAAGCA	2	Abscisic stress ripening-like protein	16588758
AACATTTTCC	2	Gibberellin regulated protein	92872830
ATCTCCAGCA	2	Gibberellin regulated protein	92890175
TAGACTATTG	2	Cop9 (Constitutive Photomorphogenic 9)	1256771

A família gênica ARP (*Auxin-repressed protein*) é específica de plantas superiores e é regulada pós-transcricionalmente pela auxina através de suas regiões UTRs. Em mandioca (*Manihot esculenta*), apesar das regiões codificadoras disponíveis para ARP1 e ARP2 serem idênticas, Reilly et al. (2007) identificaram uma diferença em sua região 3'UTR, com uma inserção de ~ 40 pb no gene ARP1 em relação ao ARP2. Através do alinhamento das quatro sequências de ESTs de *Eucalyptus* identificadas como ARPs foi possível verificar que três deles (em *itálico* na Tabela 13) apresentavam alta similaridade (> 92%). Curiosamente, um dos três ESTs (*tag* CTTTGCAGCA) apresentou uma inserção de 36 pb em relação aos outros dois (*tags* AGAGGCTCCT e AGAGGCTGGT) em sua extremidade 3', que continha mais um sítio CATG, permitindo que estes transcritos diferentes para o gene ARP fossem identificados nesta biblioteca. Uma outra curiosidade observada com o alinhamento dessas sequências foi que, embora os dois ESTs identificados como ARP2 apresentassem entre si uma alta similaridade (97%), foi detectada uma diferença justamente nos 10 pb correspondentes às suas *tags*.

Outros genes regulados por Giberelina, Etileno e Ácido Abscísico também foram identificados e, como aqueles ativados por auxinas, também apresentaram baixos níveis de expressão. Estes fitormônios também regulam a atividade cambial. Ridoutt et al. (1996) mostraram que giberelinas, principalmente a GA1, agem diretamente na divisão das células do câmbio em *Eucalyptus globulus*. As quantidades de etileno são altas quando o câmbio está ativo (SAVIDGE, 1988) e o efeito do ácido abscísico nestas células ainda não foi totalmente esclarecido, mas deve funcionar de maneira inibitória na taxa de divisão periclinal (LACHAUD et al., 1999).

Apesar da camada de células em divisão celular ser muito fina, a biblioteca SAGE foi capaz de detectar uma série de genes envolvidos diretamente no ciclo celular. A Figura 18 ilustra o ciclo celular com os genes localizados em seus pontos específicos de atuação desde a percepção do sinal de auxina até a divisão celular propriamente dita.

identificada nenhuma DNA polimerase, quatro *tags* representando Histonas, proteínas que precisam ser produzidas durante a replicação de DNA, foram encontradas.

Durante a fase S e G2 o complexo Mitótico “Ciclina M-CDK” é sintetizado e mantém-se inativo devido a uma fosforilação num sítio inibitório até que a síntese de DNA seja completa. Quando ele é ativado, promove a entrada da célula em Mitose fosforilando proteínas responsáveis por cada fase da divisão celular. A CDK que se liga às ciclinas da mitose é a cdc2, para a qual foi possível identificar duas *tags*, uma delas com o maior nível de expressão entre os genes incluídos em todo o processo, indicando que nas células da região cambial a Mitose ocorre ativamente.

Outra indicação disto é a presença de um gene *Xklp*, com 16 cópias (Figura 18), uma kinase requerida para localização de proteínas motoras nos microtúbulos e redistribuição destas do cinetócoro para zonas intermediárias do fuso durante a anáfase (CONNELLY e HIETER, 1996).

Ainda durante a Mitose, um complexo chamado APC (*Anaphase Promoting Complex*) com função de ubiquitina ligase (E3), poliubiquitina proteínas regulatórias, marcando-as para degradação no proteossomo, permitindo o avanço das fases da divisão celular e, ao final da anáfase, poliubiquitina também as ciclinas da Mitose. Duas *tags* representando ACP foram encontradas, com baixo nível de expressão (Figura 18).

Com uma maior concentração de fosfatases, o término da Mitose ocorre com a remoção de fosfatos das proteínas de divisão, permitindo a descondensação dos cromossomos separados, a formação de novos envelopes nucleares, a reassimilação do Complexo de Golgi e a divisão do citoplasma (LODISH, 2004). Um gene envolvido com a finalização da Mitose, denominado *Mob1* apresentou um alto nível de expressão em relação aos outros genes do ciclo celular nesta biblioteca (Figura 18). Este gene é requerido para a citocinese e separação celular e é importante na manutenção da ploidia da célula (LUCA e WINEY, 1998).

Foi interessante notar que os transcritos mais abundantes foram referentes a genes com função na entrada da célula em Mitose, na separação dos cromossomos e na citocinese, confirmando a identidade das células meristemáticas coletadas e sugerindo que estas são fases de grande importância na geração das células que darão origem ao xilema.

4.2.2 Expansão Celular

Ultimamente o processo de expansão celular que ocorre durante a xilogênese tem despertado o interesse de pesquisadores e da indústria de papel e celulose já que através dele é definido o tamanho final das células presentes no xilema, incluindo as fibras. Uma fibra de madeira de *E. grandis* em fase juvenil de desenvolvimento, por exemplo, tem a metade do comprimento de uma fibra produzida por uma árvore adulta (TOMAZELO FILHO, 1985). A extensão celular durante a xilogênese ocorre na parede celular primária das células através de enzimas de modificação que “remodelam” a rede de polissacarídeos e proteínas que a compõem e permitem a inclusão de novos componentes.

Segundo a revisão feita por Mellerowicz et al. (2001), a composição química da parede celular primária de *Populus*, espécie modelo para estudos de xilogênese, contém 47% de pectina, composta por galacturonato, galactose, arabinose e raminose (5:2:2:1); 22% de celulose; 18% de hemicelulose, principalmente xilanas, xiloglicanas e glicomananas; e 10% de proteínas estruturais. Entre as proteínas estruturais de parede primária estão as proteínas arabinogalactanas e as extensinas. A Figura 19 mostra a composição da parede primária e as SAGE tags referentes a proteínas e enzimas envolvidas com os processos de modificação e extensão encontradas na biblioteca da região cambial juvenil de *Eucalyptus grandis*.

As tags referentes aos genes que codificam as proteínas estruturais de parede do tipo arabinogalactana foram mais abundantes que as extensinas, tanto em número de genes quanto em nível de expressão (número de cópias). A tag mais expressa entre elas, referente a uma AGP15 de *Arabidopsis* (*AtFLAI2*), apresentou 160 cópias; as outras tags representavam AGP11, AGP12 e AGP13, totalizando 51, 11 e 55 cópias cada uma respectivamente (Figura 19). Três tags referentes às AGP11 (CTTTGAAAAG, CTTTGAAAAA e CTTTGAGAAG) apresentaram sequências muito parecidas e, quando seus respectivos ESTs foram alinhados foi verificado que estas tags estavam localizadas na região poli(A) de transcritos praticamente idênticos. Enquanto a tag CTTTGAAAAA alinha-se perfeitamente com três ESTs públicos de xilema em desenvolvimento, as outras duas alinham-se com dois diferentes *Contigs* do banco de dados do Genolyptus, o que pode indicar a ocorrência de mutações nesta região do RNAm da AGP11.

Figura 19 - Parede celular primária com seus componentes estruturais e enzimas de modificação. Estão representadas as proteínas e enzimas que puderam ser associadas a *tags*. *Tags* precedidas por um asterisco (*) representam isoformas de um mesmo gene. Para cada *tag* é fornecida uma ilustração colorida quanto ao seu nível de expressão. A cor azul indica baixo nível de expressão, a cor laranja, nível intermediário de expressão, e a vermelha, alto nível de expressão. AGP, proteína arabinogalactana

Entre as enzimas de modificação identificadas neste estudo, aquelas envolvidas na degradação de pectinas foram as mais abundantes (Figura 19), possivelmente devido à maior proporção destes polissacarídeos em relação às hemiceluloses. O mecanismo de “afrouxamento” da parede primária através das pectinas depende da metilação/desmetilação destas moléculas, regulada pela enzima Pectinametilesterase, ou Pectinesterase (PME/PE). Quando as substâncias pécticas são desmetiladas pela PME, tornam-se acessíveis às pectinases e podem ser degradadas. Cinco *tags* com diferentes níveis de expressão representando PMEs foram encontradas nesta biblioteca (Figura 19) e o alinhamento das sequências dos ESTs que elas representam mostrou que pelo menos três delas representam genes diferentes.

Curiosamente, as duas *tags* mais abundantes para PMEs possuem sequências quase idênticas, a não ser por uma única base diferente entre elas (Figura 19); e a similaridade entre

seus respectivos ESTs foi relativamente alta (66%). Com o alinhamento, foi possível verificar que as sequências desses ESTs possuem a região 3' muito diferente uma da outra, com uma delas contendo 181 pb a mais. Entretanto, o sítio CATG mais próximo da extremidade 3' é o mesmo para as duas sequências. Dessa forma, foi realizada uma pesquisa no banco de dados local de ESTs de *Eucalyptus* para verificar a procedência dessas sequências e foi constatado que uma delas foi montada a partir de 5 *reads* do Genolyptus e a outra, proveniente de um EST público, aparece em outros 6 ESTs que representam este mesmo gene. Portanto, as duas SAGE *tags* para PME mais expressas representam dois transcritos provavelmente obtidos de um mesmo gene, os quais foram identificados através de uma única base diferente na SAGE *tag*. Este resultado evidencia mais uma propriedade da técnica SAGE, a capacidade de identificar SNPs na sequência de 10 pb de uma SAGE *tag*.

Ainda, como as pectinas podem ser compostas por diferentes açúcares (formando os tipos Homogalacturonano, Ramnogalacturonano I e Ramnogalacturonano II (WILLATS et al., 2001)), estes diferentes genes identificados na biblioteca da região cambial juvenil de *E. grandis* para a enzima PME poderiam produzir diferentes isoformas com afinidades diferentes para os substratos contidos nas pectinas, porém um estudo muito mais detalhado precisaria ser realizado para confirmar esta possibilidade.

Além das PMEs, poligalacturonases e pectato liases, as principais enzimas de degradação de pectinas, também puderam ser identificadas nesta biblioteca. Duas *tags* representando diferentes genes para poligalacturonases foram observadas com nível médio de expressão, enquanto três *tags* representando pectato liases apresentaram níveis médio e baixo de expressão (Figura 19). As poligalacturonases são capazes de hidrolisar o polímero poligalacturonato, gerando produtos saturados, enquanto as liases clivam pontes internas do polímero, produzindo resíduos insaturados (BEKRI et al., 1999). Entre as *tags* referentes a pectato liases, duas delas (CGATTGTGTG e ATATATCACC) correspondem a transcritos alternativos de terminação, sendo que o transcrito representado pela primeira *tag* possui 65 pb a mais que o outro (com menor expressão).

De maneira geral, os níveis de expressão para PMEs nesta biblioteca foram maiores do que para as duas enzimas de degradação de pectinas. A PME possui duas formas de ação: dependendo das propriedades da parede celular ela pode agir aleatoriamente sobre as pectinas promovendo a ação de hidrolases dependentes de pH como as Poligalacturonases e contribuir

para o “afrouxamento” da parede primária; ou pode agir linearmente, dando origem a grupos carboxil livres que interagem com Ca^{2+} , levando à rigidificação da parede (MICHELI, 2001). Assim, não somente as pectinases dependem da ação de PMEs durante a expansão celular, mas também as células no final da expansão, incluídas nesta biblioteca, dependem delas para se tornarem rígidas, justificando sua elevada presença na região cambial de *Eucalyptus*.

A rede de celulose e xiloglicanas que compõe a parede primária de dicotiledôneas é mantida principalmente por pontes de Hidrogênio entre as cadeias 1,4- β -D-glicanas das microfibrilas e o esqueleto de glicanas da hemicelulose (DARLEY et al., 2001). A enzima Xiloglicana Endotransglicosilase (XET) catalisa a clivagem de cadeias de xiloglicana e une suas extremidades a novas cadeias de hemicelulose que estão sendo depositadas, promovendo também a expansão da parede primária. Três *tags* representando esta enzima foram encontradas nesta biblioteca e sua importância durante a xilogênese vem sendo comprovada por diversos pesquisadores. Paux et al. (2004) e Foucart et al. (2005) observaram a presença desta enzima em bibliotecas subtrativas de *Eucalyptus* de xilema vs folhas e xilema vs floema, respectivamente.

Enquanto a XET corta e liga cadeias de xiloglicanas, existe ainda uma outra enzima denominada Expansina que se intercala entre as microfibrilas de celulose e os polímeros de xiloglicanas e rompe as pontes de Hidrogênio entre elas (MELLEROWICZ et al., 2001), agindo como um tipo de “lubrificante molecular” que permite aos polímeros deslizar uns sobre os outros (DARLEY et al., 2001). Analisando os dados da biblioteca SAGE, foi interessante notar que além de duas α -expansinas, esta biblioteca também gerou duas *tags* representando β -expansinas, que têm sido relacionadas a gramíneas por serem os principais componentes de seus pólenes. Ambos tipos de expansinas compreendem grandes famílias multigênicas, podendo cada isoforma agir sobre diferentes polímeros e, sua expressão ser regulada de acordo com a fase de desenvolvimento da célula ou por sinais hormonais (DARLEY et al., 2001). Entretanto, somente dois genes para cada tipo desta enzima, com baixos níveis de expressão, foram identificados neste estudo, mesmo com o banco de dados de ESTs de *Eucalyptus* construído para análise desta biblioteca contendo 6 α -expansinas. Estes dados, em conjunto com os discutidos no início desta seção, indicam que em *Eucalyptus grandis* em fase juvenil de desenvolvimento a expansão da parede celular primária durante a xilogênese pode estar ocorrendo preferencialmente através da hidrólise de pectinas ou que o tecido amostrado continha mais células na fase final da expansão, sofrendo rigidificação, caracterizado pela redução de expansinas.

4.2.3 Síntese de Polissacarídeos e deposição da Parede Celular Secundária

A síntese dos polissacarídeos que compõem a parede celular primária e secundária é realizada a partir do precursor comum UDP-D-glicose, o qual pode ser dirigido para a síntese de hemiceluloses e pectinas e para a síntese de celulose. A Figura 20 apresenta todas as *tags* encontradas nesta biblioteca envolvidas no metabolismo desses polissacarídeos.

Hemiceluloses e pectinas representam cerca de 65% da parede primária e 26-36% da parede secundária (MELLEROWICZ et al., 2001) e a síntese de seus polissacarídeos é resultante de uma série de reações de desidrogenação, descarboxilação e epimerização que utilizam nucleosídeos 5'-açúcar-difosfato (NDPs) a partir de UDP-D-glicose (REITER e VANZIN, 2001). O ácido glucurônico, ou UDP-glucuronato, é o precursor comum de ambos componentes da parede e pode ser produzido a partir da UDP-D-glicose através da enzima UDP-glicose desidrogenase ou, alternativamente, do mio-inositol pela mio-inositol oxigenase (MIOX). Entretanto, nenhuma *tag* representando as enzimas-chave da via do mio-inositol foi identificada, como também não existe nenhuma sequência similar a MIOX no banco de ESTs de *Eucalyptus* utilizado para a análise desta biblioteca. Além disso, a predominância da UDP-glicose desidrogenase em meristemas, raízes e tecidos vasculares de *Arabidopsis* foi demonstrada por Seitz e colaboradores (2000), enquanto em hipocótilos jovens e cotilédones predomina a via do mio-inositol.

De maneira geral, as enzimas responsáveis pelas reações de interconversões entre os açúcares-NDP apresentaram níveis baixos a médios de expressão (Figura 20). As duas enzimas de maior representatividade foram a UDP-glicose desidrogenase (UGDH) e a UDP-glucuronato descarboxilase (UXS), responsáveis pela produção de UDP-glucuronato e UDP-xilose, respectivamente. Entretanto, somente uma *tag* representando UGDH com 14 cópias foi encontrada, enquanto a UXS foi representada por quatro *tags* que totalizaram 38 cópias. Este padrão de expressão é consistente com resultados anteriores de Dalessandro e Northcote (1977) que apontaram uma baixa atividade da UGDH comparada a outras enzimas responsáveis por reações subsequentes da via de interconversão de açúcares. Ainda, entre as três isoformas de UGDH presentes em *Populus*, somente uma é ativada durante a formação da parede celular secundária (HERTZBERG et al., 2001; JOHANSSON et al., 2002) e a *tag* encontrada nesta biblioteca é similar a esta isoforma (gi 61645910).

Figura 20 - Reações de interconversões de açúcares-NDP, síntese de celulose, hemiceluloses e pectinas. A ilustração traz todas as *tags* que puderam ser identificadas nesta vias, acompanhadas por setas cheias. Setas tracejadas indicam passos enzimáticos para os quais nenhuma *tag* foi encontrada. *Tags* precedidas por um asterisco (*) representam isoformas de um mesmo gene. Para cada *tag* é fornecida uma ilustração colorida quanto ao seu nível de expressão. A cor azul indica baixo nível de expressão, a cor laranja, nível intermediário de expressão, e a vermelha, alto nível de expressão. GT8D, glicosil transferase da família 8D

A enzima mais importante na produção de xilose, açúcar predominante na fração hemicelulose de *Eucalyptus* e principal componente das xilanas e xiloglicanas, é a UXS. Além disso, arabinose também pode ser sintetizado a partir de UDP-D-xilose. Entre as quatro *tags* identificadas como UXS, duas são similares a isoforma associada à membrana do Golgi *UXS4* de *Arabidopsis* e, outras duas, a isoforma solúvel *UXS5*. Uma análise mais detalhada dos ESTs associados às *tags* mostrou que as duas *UXS4* foram produzidas através de *splicing* 3' alternativo, uma vez que a *tag* "TCATTATCAA" está presente em ambos ESTs, 27 e 146 pb da cauda poli(A). Além disso, a análise dos outros dois ESTs revelou baixa similaridade entre eles, sugerindo que estes são genes distintos em *Eucalyptus grandis*. Um dado interessante é que a

UDP-D-xilose pode regular o fluxo de açúcares-NDPs para deposição da parede celular através de uma inibição por *feedback* negativo da UGDH e da UDP-glicose pirofosforilase (UGPase) (HARPER e BAR-PELED, 2002). Então, a alta expressão das UXSS observada pode estar regulando o fluxo de UDP-D-glicose na região cambial de *Eucalyptus grandis*, funcionando no direcionamento da via como um passo limitante para a síntese dos polissacarídeos da parede celular.

Enzimas envolvidas na produção de rafinose e ramnose também foram identificadas nesta biblioteca. A glicose 4,6-desidratase foi representada por apenas 3 cópias da *tag* que a representa enquanto a galactinol sintase, identificada por quatro *tags*, totalizou 20 cópias. A ramnose pode fazer parte da fração das pectinas da parede celular secundária (MELLEROWICZ et al., 2000) enquanto que a rafinose pertence a uma família de oligossacarídeos (RFOs) que são sintetizados por tecidos vegetais em formação como uma forma transportável e/ou estoque de açúcares (SPRENGER e KELLER, 2000). Esses dados sugerem que, ao invés de estocar açúcar como amido e transportá-lo como sacarose, a região cambial de *Eucalyptus grandis* pode fazê-lo através da rafinose.

Além da via de produção de UDP-D-xilose, outra importante ramificação a partir de UDP-D-glucuronato resulta na síntese de pectinas. Apesar de sua importância, já discutida na seção “Expansão Celular”, poucos genes responsáveis pela biossíntese de pectinas têm sido descritos. Somente uma *tag* com apenas 5 cópias pôde ser associada a um desses genes, o qual codifica uma glicosiltransferase de *Populus PttGTD8* (Figura 20). Este gene pertence ao mesmo subgrupo do gene QUASIMODO 1 (QUA1) de *Arabidopsis*, que tem sido relacionado a biossíntese de pectina devido a uma deficiência em homogalacturonana apresentada por mutantes QUA1 (BOUTON et al., 2002). Muito provavelmente outras *tags* representando enzimas do tipo galacturonosiltransferase, que convertem UDP-D-galacturonato em pectina, não puderam ser identificadas devido à falta de ESTs de *Eucalyptus* para seus genes. Esta hipótese pôde ser reforçada com uma busca no banco de dados local de ESTs de *Eucalyptus* por oito diferentes galacturonosiltransferases de *Arabidopsis* (GAUT), na qual foi verificado que existem apenas dois ESTs similares a GAUT3 e GAUT12 mas que não possuem cauda poli(A), pré-requisito imprescindível para identificação das *tags*.

Resumidamente, o nível de expressão das *tags* para as enzimas envolvidas na síntese de glicanas e pectinas foi relativamente baixo quando comparado ao nível de expressão das *tags*

para as enzimas relacionadas à síntese de celulose, sugerindo que a maior parte das células da região cambial das árvores jovens amostradas estava em crescimento secundário ativo.

A parede celular secundária é caracterizada principalmente pela deposição padronizada de microfibrilas de celulose, para as quais o UDP-D glicose também é o precursor. A principal via de produção deste substrato nas células da região cambial de *Eucalyptus grandis* parece ocorrer pela sacarose, através da sacarose sintase (SUSY), cujo nível de expressão de sua *tag* foi maior que o encontrado para a *tag* referente a UDP-glicose pirofosforilase (Figura 20). Segundo Joshi *et al.* (2004), a sacarose sintase possui a importante função de direcionar UDP-glicose diretamente para as enzimas celulose sintase do complexo, evitando uma competição pelo *pool* celular de UDP-glicose e possibilitando, consequentemente, as elevadas taxas de síntese de celulose observadas durante o crescimento secundário. A única *tag* similar ao gene *SUSY1* de *Arabidopsis* identificada nesta biblioteca apresentou 29 cópias (Figura 20). Entretanto, não é possível afirmar com certeza se a *tag* identificada é referente à *SUSY1* ou *SUSY2* de *Arabidopsis*, pois as duas sequências diferem na região 5' e são muito similares na região 3', a partir da qual as SAGE *tags* são produzidas. Ambas *SUSY1* e *SUSY2* são representadas pelos transcritos relacionados ao metabolismo de carboidratos mais abundantes em *Populus* (GEISLER-LEE *et al.*, 2006), e são altamente expressos durante a formação de madeira de tensão nesta espécie (ANDERSON-GUNNERÅS *et al.*, 2006).

A alta expressão de transcritos para o gene da Sacarose sintase associada ao fato de que nenhuma *tag* referente às enzimas envolvidas na síntese e degradação de amido foi identificada nesta biblioteca é um forte indicativo de que na região cambial de árvores de *Eucalyptus grandis* na fase juvenil de desenvolvimento a atividade da sacarose sintase é fundamental para produzir UDP-glicose e que o *pool* de UDP-glicose produzido está sendo direcionado para a síntese de celulose e não para a síntese de compostos de reserva. Este resultado corrobora resultados anteriores que reportam a expressão de enzimas relacionadas à síntese de amido somente nas células cambiais em dormência, em gemas e órgãos de reserva, e não em células cambiais ativas (GEISLER-LEE *et al.*, 2006).

Recentemente foram identificados seis genes para celulose sintase em *Eucalyptus grandis* sendo três relacionados à síntese de parede primária (*EgCesA4*, *EgCesA5* e *EgCesA6*), e três mais expressos em xilema, relacionados à síntese de parede secundária (*EgCesA1*, *EgCesA2* e *EgCesA3*) (RANIK e MYBURG, 2006). Com exceção do *EgCesA6*, esta biblioteca produziu *tags*

para todos os outros genes (Figura 20). No entanto, enquanto uma única *tag* foi identificada para *EgCesA2*, *EgCesA3* e *EgCesA5*, os genes *EgCesA1* e *EgCesA4* foram representados por duas *tags* diferentes cada um. Através do alinhamento de seus respectivos ESTs foi possível observar que nos dois casos um único gene produz duas isoformas distintas devido a splicing 3'. Ainda, foi identificado um gene *PtCslA1* (*cellulose synthase-like* de *Pinus*), com apenas duas cópias, que não mostrou similaridade a nenhuma celulose sintase de *Eucalyptus* conhecida. Na verdade, esta proteína similar à celulose sintase contém um motivo D,D,D,QXXRW que provavelmente reflete uma divergência em sua função. Segundo Doblin et al. (2002), é possível que proteínas Csl estejam envolvidas na síntese de polissacarídeos com ligações do tipo β como calose, xiloglicana, glucuronoarabinosilana, glucomanana, galactomanana ou galactoglucomanana, e possivelmente celulose.

A *tag* “GAATCAAAAT” referente a uma das isoformas para *EgCesA1* foi a mais abundante entre todas as outras *EgCesA*, apresentando 36 cópias. As demais *tags* apresentaram número relativamente baixo de cópias, onde os genes relacionados à parede secundária apresentaram 2 e 8 cópias e os genes de parede primária 3 a 10 cópias (Figura 20). Quando a soma dos níveis de expressão para essas duas celulose sintase foi calculada, observou-se os mesmos perfis transcricionais descritos por RANIK e MYBURG (2006). Entretanto, enquanto no trabalho citado os genes de parede secundária possuem níveis de expressão igualmente altos, nesta biblioteca *EgCesA2* e *EgCesA3* apresentaram níveis similares ao gene *EgCesA4*, relacionado à parede primária. Além disso, esses autores verificaram que o gene de maior expressão em xilema de *Eucalyptus grandis* é o *EgCesA3* e não *EgCesA1* como observado nesta biblioteca. Uma hipótese que explicaria esta diferença entre os resultados é a divergência na estrutura genética entre os indivíduos da população natural, utilizados no trabalho de Ranik e Myburg, e os indivíduos da população utilizada para construção desta biblioteca, provenientes de um plantio comercial melhorado geneticamente para a produção de papel e celulose. Resultados anteriores de RANIK et al. (2006) também suportam esta hipótese. Ranik e seus colaboradores realizaram um estudo do perfil transcricional de plantas altamente produtivas de *Eucalyptus* destinadas à produção de papel e celulose, e identificaram entre os genes de maior expressão, transcritos para subunidades de CESA similares a *PtrCesA1* e *AtCesA8*, os quais segundo a análise de alinhamento de sequências de aminoácidos, realizadas por RANIK e MYBURG (2006), são similares à *EgCesA1*.

5

Figura 21 - Ilustração da síntese de microfibrilas de celulose através do complexo protéico presente na membrana celular. PM, Membrana Plasmática; UDPG, UDP-Glicose; SUSY, Sacarose sintase; CESA, Celulose sintase A

A deposição padronizada da parede celular secundária é dirigida pelos microtúbulos, os quais localizam-se diretamente abaixo da parede secundária, com uma orientação paralela às microfibrilas de celulose. Assim, eles são os responsáveis não somente pela deposição das microfibrilas de celulose na forma de rosetas, mas também pelo ângulo que estas são depositadas, principalmente na camada S2 da parede celular secundária. Diferentes α e β -tubulinas e actinas foram encontradas apresentando altos, médios e baixos níveis de expressão nesta biblioteca. Uma análise mais aprofundada desses genes poderia indicar se existe especificidade entre eles para os diferentes tipos celulares (vasos, fibras e parênquima), diferentes estádios de desenvolvimento ou ainda entre as diferentes madeiras que podem ser produzidas (por exemplo, de acordo com a estação climática ou com o estresse sofrido). Spokevicius et al. (2007) mostraram que uma β -tubulina de *Eucalyptus* (EgTUB1) determina a orientação das microfibrilas de celulose em paredes secundárias de fibras, uma vez que o ângulo de deposição das microfibrilas foi significativamente alterado como consequência de sua superexpressão.

Finalmente, durante a maturação do xilema, ligninas e hemiceluloses fixam as microfibrilas resultando num material composto, rígido e resistente, envolvido por uma matriz amorfa e resistente à compressão (PILATE et al., 2002).

4.2.4 Lignificação

Aproximadamente 19 a 20% da parede celular de células maduras do xilema do *Populus* correspondem a lignina (MacDOUGALL et al., 1993). Os polímeros fenólicos de lignina são formados a partir da polimerização de três diferentes álcoois hidroxicinamóil ou monolignóis, *p*-coumaril, coniferil e sinapil, os quais dão origem às subunidades do polímero de lignina H, G e S, respectivamente.

Os precursores da lignina são produzidos a partir da via dos fenilpropanóides, que inicia-se com a desaminação da Fenilalanina pela enzima fenilalanina amônia liase (PAL), gerando ácido cinâmico. Este ácido é então convertido a ácido coumárico através de uma segunda enzima denominada cinamato 4-hidroxilase (C4H) e, na sequência, encaminhado para a síntese de tioésteres, precursores dos monolignóis. A Figura 22 ilustra todas as etapas que levam à produção de lignina, desde a entrada da fenilalanina na via dos fenilpropanóides até a polimerização dos monolignóis.

Todas as enzimas necessárias para a produção de lignina foram representadas nesta biblioteca (Figura 22), indicando que o material coletado estava em pleno crescimento secundário. Duas *tags* para o gene da PAL foram identificadas, sendo que o EST referente à *tag* mais expressa mostrou alta similaridade com o gene *AtPAL1*, enquanto o da *tag* menos expressa foi similar aos genes *AtPAL1* e *AtPAL2*. Entretanto, o alinhamento desses ESTs indicou que eles representam genes distintos. É válido ressaltar que, no genoma de *Arabidopsis*, estes dois genes possuem um papel mais importante durante a lignificação vascular do que os quatro outros identificados nesta espécie (RAES et al., 2003) e, portanto, podem ser específicos e/ou predominantes em tecidos que estão em processo de xilogênese. Da mesma forma, a *tag* identificada como C4H apresentou um nível de expressão semelhante ao da PAL (Figura 22) e seu EST correspondente foi similar ao gene *PtreC4H2* de *Populus*, o qual apresentou uma maior especificidade no xilema do que o gene *PtreC4H1* e, além disso, é fracamente expresso no floema (LU et al., 2006).

Na sequência da via de produção da lignina, o gene da enzima 4-coumarato: coenzima A ligase (4CL) produz tioésteres-CoA a partir de ácidos hidroxicinâmicos que, nesta biblioteca, foi identificado por uma *tag* apresentando 8 cópias (Figura 22), com alta similaridade ao gene *4CL1* de *Eucalyptus camaldulensis*. Como o gene *C4H*, este gene também mostrou-se preferencialmente expresso no xilema quando comparado a um outro gene, o *4CL2* (ALLINA et al., 1998).

As enzimas PAL, C4H e 4CL estão presentes em plantas como pequenas famílias gênicas com isoformas que exibem expressão tecido-específica e que vem sendo correlacionadas a funções distintas no metabolismo dos fenilpropanóides em cada tecido (KAO et al., 2002; RAES et al., 2003; LU et al., 2006). Embora estas enzimas realizem as reações centrais do metabolismo dos fenilpropanóides, já que catalisam os passos iniciais da via, nenhuma alteração nos seus níveis de expressão foi observada durante a formação de madeira de tração (caracterizada pela redução no conteúdo de lignina) em *Eucalyptus* e *Populus*. Ao contrário, foram observadas reduções no número de transcritos para as enzimas que catalisam passos posteriores, específicos em cada braço da via que leva à formação das subunidades H, S e G de lignina (PAUX et al., 2005; ANDERSSON-GUNNERÅS et al., 2006).

Figura 22 - Síntese e polimerização dos monolignóis. A ilustração traz todas as *tags* identificadas nesta via acompanhadas por setas pretas. Setas na cor cinza indicam passos enzimáticos para os quais nenhuma *tag* foi encontrada e as setas tracejadas representam passos de conversão catalisados por F5H e COMT. *Tags* precedidas por um asterisco (*) representam isoformas de um mesmo gene. Para cada *tag* é fornecida uma ilustração colorida quanto ao seu nível de expressão. A cor azul indica baixo nível de expressão, a cor laranja, nível intermediário de expressão, e a vermelha, alto nível de expressão. PAL, fenilalanina amônia-liase; C4H, cinamato 4-hidroxilase; 4CL, 4-coumarato: coenzima A ligase; CCR, cinamil-coenzima A redutase; CAD, cinamil álcool desidrogenase; C3H, coumarato 3-hidroxilase; HCT, hidroxicinamil-CoA shiquimato/quinato hidroxicinamil-transferase; COMT, cafeato/5-hidroxi coniferaldeído O-metiltransferase; CCoAOMT, cafeil-coenzima A O-metiltransferase; F5H, ferulato/coniferaldeído 5-hidroxilase

A madeira das angiospermas é composta por grandes quantidades de unidades G e S de lignina, enquanto unidades H são adicionadas em pequenas quantidades predominantemente na lamela média e nos cantos das células. As unidades G são principalmente depositadas na parede celular secundária e as S nos estágios finais da lignificação (LEWIS, 1999). Corroborando com estes resultados, nesta biblioteca foi observada uma maior expressão entre as *tags* que representam as enzimas responsáveis pelos passos mais específicos da via do que entre aquelas referentes a enzimas das reações iniciais (PAL e C4H) e das reações direcionadas à síntese de

unidades H (4CL, cinamil-CoA redutase (CCR) e cinamil álcool desidrogenase (CAD)) (Figura 22). Assim, as *tags* com maiores níveis de expressão correspondem a ácido cafeico O-metiltransferase (COMT), cafeil-CoA O-metiltransferase (CCoAOMT), S-adenosil metionina sintase (SAMS) e lacase. Reforçando os resultados obtidos neste estudo, PAUX et al. (2005) sugeriram que os genes das enzimas 4CL, CCR e CAD formam um *cluster* comum de controle transcricional enquanto COMT e CCoAOMT formam um outro *cluster* de co-regulação.

As duas *tags* identificadas para CCR apresentaram uma baixa expressão (5 cópias cada) e tiveram alta similaridade ao gene CCR de *Eucalyptus gunnii*, único gene para a enzima identificado até então na espécie (LACOMBE et al., 1997) e a uma CCR de *Arabidopsis*. Para a CAD, duas *tags* com níveis de expressão bem diferentes foram encontradas. O EST que identificou a *tag* mais expressa apresentou similaridade a uma CAD de *Eucalyptus globulus* enquanto o outro, foi similar a uma CAD de *E. gunnii*. Entretanto, o alinhamento dos ESTs não produz nenhuma região de identidade indicando que provavelmente, as duas *tags* representam sequências transcritas a partir de genes diferentes em *E. grandis*.

O passo determinante para o direcionamento de p-coumaril-CoA para a síntese de unidades G e S é sua conversão ao ácido *p*-coumaril shiquimico/ácido quinico pela Hidroxicinamoil-transferase (HCT), já que a Coumarato 3-hidroxilase (C3H) não é capaz de utilizar *p*-coumaril-CoA como substrato (SCHÖCH et al., 2001). Tem sido demonstrado que *p*-coumaril shiquimato e *p*-coumaril quinato são importantes intermediários na via dos fenilpropanóides, com a enzima HCT agindo antes e depois da C3H na produção de cafeil-CoA (HOFFMANN et al., 2004). Uma única *tag* com 8 cópias foi identificada como HCT enquanto uma *tag* quase três vezes mais expressa referente a C3H apresentou similaridade ao gene CYP98A3 de *Arabidopsis* cuja expressão é mais evidente em células vasculares em lignificação (NAIR et al., 2002). Alternativamente, C3H também pode agir sobre o ácido *p*-coumárico produzindo ácido cafeico que, por sua vez, pode ser direcionado para produção de ferulato pela enzima COMT, ou para produção de cafeil-CoA pela 4CL (Figura 22).

Durante a biossíntese dos monolignóis, reações que envolvem metilação são requeridas em duas posições do anel aromático da unidade de monolignol. Uma ocorre na posição 3 e a outra na posição 5. A metilação na posição 3 leva à formação de unidades G enquanto a dupla metilação nas posições 3 e 5 levam à formação de unidades S. A metilação da posição 3 ocorre sobre o intermediário CoA éster e é catalisada pela CCoAOMT, enquanto a metilação na posição

5 ocorre sobre o intermediário aldeído sendo catalisada pela COMT (Revisado em LI et al., 2006). Apesar de inicialmente se acreditar que a COMT somente convertia Ácido Cafeico em Ferulato (DIXON, 2001), evidências posteriores demonstraram que a COMT catalisa preferencialmente a conversão de 5-hidroxiferulato, 5-hidroxiconiferaldeído e 5-hidroxiconiferil álcool em ácido sinápico, sinapaldeído e álcool sinapil, respectivamente, e, portanto, sua atividade ocorre preferencialmente sobre os produtos da reação catalisada pela enzima ferulato/coniferaldeído 5-hidroxilase (F5H) (PARVATHI et al., 2001). As *tags* para F5H e COMT, entretanto, apresentaram padrões de expressão bastante distintos. A *tag* referente a F5H apresentou 7 cópias, enquanto 25 cópias foram observadas para a *tag* da COMT nesta biblioteca. Estes resultados sugerem que provavelmente, a transcrição de F5H e COMT não seja co-regulada. Da mesma forma, proteínas F5H não estiveram presentes durante a regeneração cambial em *Populus*, enquanto isoformas protéicas de COMT foram detectadas em todos os estádios avaliados (JUAN et al., 2006).

A CCoAOMT é uma enzima chave na produção dos monolignóis, metilando Cafeil-CoA e gerando Feruloil-CoA, o que justifica os altos níveis de expressão de transcritos para CCoAOMT encontrados em células de *Populus* durante o processo de lignificação (ZHONG et al., 2000). Além disso, proteínas CCoAOMT também foram identificadas como umas das mais abundantes durante a diferenciação do xilema em *Populus* (VANDER MIJNSBRUGGE et al., 2000). Os dados gerados a partir da biblioteca da região cambial de *E. grandis*, desenvolvida no presente estudo, mostraram que a CCoAOMT teve o maior nível de expressão entre todos os genes identificados na via de biossíntese de lignina (Figura 22). As duas *tags* encontradas para CCoAOMT representaram ESTs com alta similaridade aos genes CCoAOMT de *Eucalyptus gunnii* e *E. globulus* sendo a de menor expressão (25 cópias) similar a CCoAOMT2 de *E. globulus* e a de maior expressão (48 cópias) similar a *EgCCoAOMT1*. A importância dessa rota catalizada pela CCoAOMT foi demonstrada através da repressão da enzima em plantas de Tabaco e *Populus* transgênicas, que levou a uma alteração na relação S/G de lignina e a uma significativa redução em seu conteúdo total (ZHONG et al., 1998; ZHONG et al., 2000). Como o nível de expressão da CCoAOMT observado nesta biblioteca é bem maior que o da COMT, a síntese das unidades S e G provavelmente está sendo realizada preferencialmente via Cafeil-CoA nas árvores de *E. grandis* em fase juvenil de desenvolvimento.

A adição de grupos metil em posições específicas dos monolignóis leva a um grande consumo desses grupamentos nas células do xilema durante a lignificação. Para isto, tanto na reação catalisada por COMT quanto na reação catalisada por CCoAOMT, a S-adenosilmetionina (SAM) é utilizada como doadora de grupos metil (CANTÓN et al., 2005) e a enzima que catalisa sua síntese é a S-adenosilmetionina sintase (SAMS). Apesar de ser considerada uma enzima do tipo *housekeeping* envolvida no metabolismo de metionina, transcritos para SAMS têm sido encontrados em elevada abundância no xilema durante o crescimento secundário em *Populus*, *Arabidopsis* e outras espécies (STERKY et al., 1998; VANDER MIJNSBRUGGE et al., 2000; OH et al., 2003; JUAN et al., 2006). Foi interessante notar que as enzimas CCoAOMT e SAMS apresentaram níveis de expressão similarmente altos nesta biblioteca, o que poderia indicar a presença de um sistema regulatório comum para seus genes, embora ainda não tenha sido descrita nenhuma evidência para isto. De acordo com Cantón et al. (2005), SAMS pode funcionar localmente fornecendo grupamentos metil que são consumidos por CCoAOMT e COMT, garantindo a manutenção de altas taxas de lignificação que são observadas durante o crescimento secundário. Quatro *tags* para SAMS que totalizaram 68 cópias foram observadas nesta biblioteca e, o alinhamento de seus ESTs mostrou que as duas mais expressas, com alta similaridade ao gene *SAM-2* de *Arabidopsis*, possuem um SNP na sequência de 10 pb da tag. O EST referente a *tag* com 12 cópias teve alta similaridade ao gene *SAM-1* de *Arabidopsis* e o menos expresso ao gene *SAM-3*. Ainda, confirmando a importância desta enzima na formação do xilema/lignificação, três *spots* representando SAMS foram identificados através do sequenciamento de proteínas extraídas do mesmo material vegetal utilizado para construção desta biblioteca (CELEDON et al., 2007).

O último passo na síntese de lignina é a polimerização dos monolignóis que pode ser conduzida por peroxidases e/ou lacases (BAUCHER et al., 2003). Enquanto cinco *tags* representando Lacases foram encontradas nesta biblioteca, nenhuma *tag* para Peroxidase foi observada, mesmo existindo duas peroxidases de *E. globulus* no banco de dados de ESTs utilizado para identificação das *tags*. Em adição a esses resultados, a importância das lacases durante o crescimento secundário tem sido demonstrada em espécies de *Populus*, *Pinus* e *Eucalyptus* (ALLONA et al., 1998; PAUX et al., 2004; ANDERSSON-GUNNERÅS et al., 2006). Por outro lado, nenhuma redução significativa na expressão da peroxidase foi observada em madeira de tensão em *Populus* (caracterizada por um decréscimo na quantidade de lignina),

enquanto o gene *lac3* para lacase foi co-regulado com genes envolvidos na biossíntese de lignina (ANDERSSON-GUNNERAS et al., 2006).

Recentemente foi verificada uma indução do gene *Dirigent protein* coincidente com a lignificação (HERTZBERG et al., 2001) e foi proposto que estas proteínas podem atuar na polimerização dos monômeros de lignina uma vez que elas se ligam a sítios localizados principalmente na camada S1 da parede secundária e mediam as ligações entre os polímeros (BURLAT et al., 2001). Corroborando com esta hipótese, foram encontradas três *tags* para *dirigent proteins* nesta biblioteca, totalizando 35 cópias (Figura 22). As duas *tags* com menor expressão representaram isoformas similares ao gene *pDIR18* de *Picea glauca*, enquanto a *tag* de maior expressão representou um EST similar ao gene *pDIR10*. Este gene está relacionado à síntese de compostos fenólicos (lignanais e ligninas) de resistência ao ataque de insetos em coníferas (RALPH et al., 2006). O alinhamento dos ESTs mostrou que as *tags* para o gene *pDIR18* representam transcritos produzidos por um processamento alternativo 3'. Portanto, possivelmente diferentes genes e diferentes isoformas da proteína DIR estejam atuando na polimerização de monolignóis durante a lignificação em *E. grandis*.

Ainda, foram encontradas três *tags* representando proteínas do tipo *Germin-like protein* (GLP) que totalizaram 37 cópias (dados não mostrados). Juntamente com as germins, as GLPs formam uma família caracterizada por duas funções enzimáticas diferentes, desempenhando uma atividade oxalato redutase e superóxido dismutase. Ambas atividades resultam na produção de H_2O_2 que pode ser direcionado para a polimerização oxidativa dos monolignóis durante a lignificação (MATHIEU et al., 2006). Dessa forma, a presença destas *tags* nesta biblioteca pode indicar uma ligação entre a expressão desse gene na região cambial de *E. grandis* e um possível papel na lignificação.

A lignificação intensa da parede celular dos elementos de vaso e fibras que ocorre durante o crescimento secundário libera grandes quantidades de amônia através da desaminação da fenilalanina pela PAL. Este grande consumo de N poderia causar uma deficiência nutricional na planta caso todo N consumido fosse suprido apenas pela assimilação direta através da redução de nitrato. Ao contrário disso, nas células em crescimento secundário ocorre um importante mecanismo de reciclagem de N onde o N liberado é localmente reciclado e reincorporado à glutamina através da glutamina sintetase (GS) (KEYS, 2006; CÁNTON et al., 2005). Os resultados obtidos desta biblioteca indicam uma provável importância da glutamina sintetase

durante a xilogênese já que detectaram a presença de uma *tag* referente a isoforma citosólica GS1 com nível muito alto de expressão (Figura 22), a qual ocupou o décimo terceiro lugar entre as 50 *tags* mais expressas da biblioteca da região cambial de *E. grandis*. Além disso, esse resultado é reforçado pelos resultados obtidos por Celedon et al. (2007) que mostraram a presença desta proteína neste mesmo material vegetal. Outra *tag*, com baixo nível de expressão, também foi identificada como GS1, mas o alinhamento dos ESTs das duas *tags* não apresentou nenhuma região de identidade entre as sequências, sugerindo que em *E. grandis*, ao menos dois genes GS1 estão presentes.

4.2.5 Morte Celular Programada

Poucos trabalhos envolvendo a Morte Celular Programada (PCD) que ocorre durante a xilogênese têm sido publicados, sendo a maioria sobre PCD de elementos traqueários (TE), estrutura correspondente aos vasos das angiospermas. Foi somente em 2005 que Moreau e colaboradores investigaram a PCD em células de fibra de árvores de *Populus*, diferente daquela ocorrida nos vasos, que é muito mais rápida, ocorrendo a cerca de 100-150 μm da região cambial. Nesta região as células de fibras ainda estão terminando sua expansão e iniciando a deposição de parede secundária. Esses autores puderam correlacionar uma série de genes envolvidos com PCD e autólise celular em plantas tais como cisteína protease, *SLP2* (subtilisin serina protease), *Endo93*, nodulina *MtN21*, reguladores de apoptose *Bcl-2* e *Bax inhibitor-1*, os quais também foram encontrados na biblioteca da região cambial de *E. grandis*.

Apesar da identificação de vários genes relacionados à PCD de fibras do xilema, o mecanismo pelo qual ela ocorre ainda não foi descrito, com a exceção de que ocorre um desaparecimento gradual do citoplasma até a sua autólise. Groover e Jones (1999) propuseram um modelo para diferenciação de TEs que pressupõe uma atividade coordenada entre a síntese da parede celular e a PCD. Segundo esse modelo, durante a síntese da parede celular secundária ocorre também a produção de enzimas hidrolíticas; a síntese dos precursores da parede celular é acompanhada pela secreção de uma protease de 40 kD (serina protease), levando a um aumento da atividade proteolítica na matriz extracelular. Seis *tags* representando serina proteases do tipo subtilase foram encontradas na biblioteca SAGE, totalizando 43 cópias (Figura 23). Ao mesmo tempo nucleases, proteases e fosfatases ácidas se acumulam no vacúolo das células, que sofre um aumento em seu tamanho. Uma nuclease e uma fosfatase ácida foram identificadas neste estudo,

enquanto para a cisteína protease foram encontradas 9 *tags* com 73 cópias no total (Figura 23). Quando a quantidade de proteases na matriz extracelular atinge uma taxa crítica, a síntese de parede secundária é finalizada e a morte celular é ativada por um influxo de Ca^{2+} , que causa um colapso do vacúolo e uma invasão do citoplasma pelas enzimas hidrolíticas, que atacam organelas e ácidos nucleicos, resultando na degradação do conteúdo celular.

Figura 23 - Modelo para PCD em TE proposto por Groover e Jones (1999). As *tags* identificadas como genes envolvidos neste processo estão representadas. A cor azul indica baixo nível de expressão, a cor laranja, nível intermediário de expressão, e a vermelha, alto nível de expressão

4.3 Validação da Biblioteca SAGE por qRT-PCR

A validação de bibliotecas SAGE, assim como de outras técnicas de estudo de perfis transcricionais, tem sido realizada por muitos autores através de qRT-PCR, técnica capaz de monitorar o acúmulo de produtos amplificados ao final de cada ciclo através de leituras de fluorescência. Essa fluorescência é liberada por um corante adicionado à mistura de reação (neste caso o *SYBR green*) que se intercala com a fita dupla de DNA e, durante a reação de PCR, ela aumenta exponencialmente. Assim, o ciclo de PCR em que a fluorescência cruza um limiar ou *threshold* pode ser relacionado à quantidade inicial de moléculas na reação; consequentemente,

quanto mais moléculas iniciais uma amostra tiver, menor será o número de ciclos (CT) necessários para ultrapassar este limite para a leitura da fluorescência, e vice-versa.

Para o cálculo dos valores de expressão V de cada gene na mesma amostra utilizada para construção da biblioteca SAGE, os valores de CT adotados foram aqueles obtidos com a diluição 0,001 vez e as Eficiências de Amplificação (E) foram obtidas do programa *Relative Expression Software Tool - 384 = REST-384 © - version 1* (Tabela 14). Segundo Roth (2002) a eficiência da reação de PCR deve estar entre 0,8 e 1,0 ou 80 e 100%; para o cálculo de E em qPCR os valores de correlação obtidos com a curva de diluições são somados à eficiência da reação ($E = r + E$). O valor ótimo para E é igual a 2, que significa que a cada ciclo a quantidade de moléculas presentes na reação deve dobrar.

Tabela 14 - Valores de Eficiência de Amplificação (E) para os *primers* utilizados no experimento de validação dos dados do SAGE. E , eficiência de amplificação; r , correlação; E , eficiência de amplificação calculada

Gene	E	r	E
UDP-glicose pirofosforilase (UGPase)	0,96	0,99	1,95
Pectinesterase	0,94	1,00	1,94
Cafeil-CoA O-metiltransferase (CCoAOMT)	0,97	1,00	1,97
S-adenosilmetionina sintetase (SAMS)	0,93	1,00	1,93
Álcool cinamil desidrogenase (CAD)	0,91	1,00	1,91
4-coumarato: coenzima A ligase (4CL)	0,88	1,00	1,88
Isocitrato desidrogenase (IDH)	0,98	1,00	1,98
Xilana 1,4-beta-xilosidase	0,94	0,98	1,92
UDP-xilose sintase 4 (UXS4)	0,66	0,97	1,63
Sacarose sintase (SuSy)	1,38	0,98	2,36
UDP-arabinose 4-epimerase	0,84	0,98	1,82
Alcohol desidrogenase (ADH)	0,88	0,99	1,87
Piruvato quinase (PK)	1,00	1,00	2,00
Glutamina sintetase (GS1)	0,86	0,95	1,81

A Figura 24 mostra a curva obtida com as três diluições e as três repetições do gene Isocitrato desidrogenase, na qual pode-se observar uma eficiência (E) igual a 97,8% e um coeficiente de correlação entre os valores de CT e as diluições igual a 0,996. De maneira geral, os valores de E se enquadraram neste intervalo. Ainda, através de uma cuidadosa análise das curvas de dissociação foi possível observar que todos os produtos apresentaram somente um pico, garantindo a especificidade dos *primers* para os genes alvo do experimento.

Neste estudo os resultados obtidos com a técnica SAGE puderam ser validados através de qRT-PCR. Tanto os valores de expressão V absolutos, quanto os normalizados foram positivamente e significativamente ($p < 0,01$) correlacionados com as frequências das SAGE *tags* (Tabela 15), onde genes representados por *tags* com maiores frequências (TPT) apresentaram menores CTs/maiores valores de expressão V , e vice-versa.

Figura 24 - Curva padrão obtida a partir das leituras de CT de três diluições e três repetições dos produtos de amplificação do gene Isocitrato desidrogenase

Tabela 15 - Análise de correlação entre as frequências das SAGE *tags* (TPT) e os valores de expressão V obtidos através de qRT-PCR utilizando valores absolutos e normalizados com o gene UGPase. CT, *threshold cycle*; DP, desvio padrão; TPT, *tags* por mil

Gene	CT $\pm$ DP	Absoluto		Normalizado	
		TPT	V ($\times 10^{-5}$)	TPT	V ($\times 10^{-5}$)
UGPase	13,65 $\pm$ 0,09	0,6619	10,9908	1	1
GS1	12,81 $\pm$ 0,10	2,9126	50,0193	4,4000	4,5510
CCoAOMT	12,87 $\pm$ 0,18	2,1183	16,2593	3,2000	1,4793
SAMS	12,51 $\pm$ 0,25	1,5004	26,8247	2,2667	2,4406
SuSy	9,77 $\pm$ 0,01	1,2798	22,6352	1,9333	2,0595
CAD	14,03 $\pm$ 0,12	1,0591	11,3831	1,6000	1,0357
IDH	13,58 $\pm$ 0,34	0,8826	9,3411	1,3333	0,8499
PK	14,26 $\pm$ 0,07	0,5296	5,0864	0,8000	0,4628
Pectinesterase	15,52 $\pm$ 0,04	0,3972	3,4076	0,6000	0,3100
4CL	13,61 $\pm$ 0,02	0,3530	18,5306	0,5333	1,6860
UXS4	25,89 $\pm$ 0,34	0,2206	0,3205	0,3333	0,0292
UDP-arabinose 4-epimerase	16,06 $\pm$ 0,05	0,2206	6,6686	0,3333	0,6067
ADH	14,39 $\pm$ 0,15	0,1324	12,2748	0,2000	1,1168
β -Xilosidase	20,87 $\pm$ 0,63	0,0883	0,1226	0,1333	0,0111
		$r(12) = 0,8456, p < 0,01$		$r(12) = 0,8456, p < 0,01$	

4.4 Análise comparativa da expressão gênica entre árvores juvenis e adultas

A análise da biblioteca SAGE da região cambial de *Eucalyptus grandis* em fase juvenil de desenvolvimento levou à caracterização qualitativa e quantitativa dos principais genes envolvidos na xilogênese dessa espécie. Entretanto, a região cambial dessas árvores está produzindo madeira do tipo juvenil, que em relação à madeira adulta apresenta grandes diferenças químicas, físicas e anatômicas. Entre elas, maiores teores de ligninas, menos celulose, menor densidade básica, menor comprimento de fibras e vasos, menor diâmetro celular, paredes celulares mais finas, maiores ângulos fibrilares, entre outras. Enquanto a madeira juvenil é produzida nos primeiros anos de vida da árvore, a madeira adulta é aquela que se forma nas árvores em idades mais avançadas.

A indústria de papel e celulose faz o corte do eucalipto por volta dos sete anos de idade, quando as árvores ainda estão em fase de transição entre a produção de madeira juvenil e adulta. Portanto, a madeira que segue para a indústria possui grande proporção de madeira juvenil e características muito variáveis, o que interfere na eficiência de produção da polpa celulósica. Por outro lado, a madeira do tipo adulta apresenta características bem definidas e estáveis, principalmente com relação a sua densidade, o que levaria a um melhor rendimento na produção de polpa celulósica.

Com a análise da expressão global dos genes presentes na região cambial de *E. grandis* surgiu então, a necessidade de verificar quais genes, ou grupos de genes, poderiam ser responsáveis por características tão diferentes observadas em árvores nestas distintas fases de desenvolvimento.

Oliveira et al. (2005), estudando a densidade básica na madeira de *E. grandis* com 25 anos de idade, demonstraram que seus valores tornam-se menos variáveis somente quando as árvores atingem um diâmetro (DAP) de 25 cm e que o teor de umidade desta madeira, intimamente ligado a densidade básica, além de se manter alto, tende a se estabilizar somente com 30 cm de diâmetro (DAP). Assim, 25 genes envolvidos com metabolismo geral, modificação de parede primária, biossíntese de polissacarídeos, deposição de parede secundária e lignificação foram selecionados para uma comparação de seus níveis de expressão na região cambial de árvores de *E. grandis* com 3 anos (juvenis) e com 22 anos (adultas) de idade, as quais apresentaram um DAP acima de 30 cm.

Foram conduzidas reações de RT-PCR em Tempo Real da mesma forma que para o experimento de validação dos dados do SAGE, com a exceção de que foram utilizadas 4 diluições para a construção das curvas-padrões e não três, e os cálculos dos valores de expressão V foram feitos com os CTs observados para a quarta diluição de cDNA. Este procedimento foi necessário uma vez que foi verificada uma alta concentração de cDNA na primeira diluição, o que interferiu nos cálculos dos valores de Eficiência (E) e poderia prejudicar as comparações entre as amostras já que o fator limitante durante a reação de amplificação deveria ser a quantidade de cDNA inicial.

Os valores de E e CTs obtidos para os 26 genes (25 genes alvo e 1 gene de referência) nas duas amostras estão apresentados na Tabela 16. Roth (2002) verificou que o método mais rigoroso para comparações relativas de níveis de expressão é o cálculo da E para cada amostra, já que ela não varia somente em função dos *primers*, mas também em função do material amostrado. Apesar da quantificação do RNA total e utilização das mesmas quantidades iniciais para a síntese de cDNA, é necessária ainda uma normalização dos resultados através de um gene de referência, ou normalizador (sombreado na Tabela 16), uma vez que tem sido reportado que a Eficiência da síntese de cDNA pode variar entre 5% e 90% devido à alta sensibilidade da enzima transcriptase reversa a sais e compostos orgânicos que podem estar presentes como resíduos da extração do RNA (FREEMAN et al., 1999).

A análise das curvas de dissociação revelou que somente um produto específico de DNA estava sendo amplificado para cada jogo de *primers* e, a partir dos valores de E e CTs foi possível calcular os valores de expressão V , normalizados com o gene que codifica uma α -tubulina, similar a α -tubulina 5 de *Arabidopsis*. Foi então calculada a relação existente entre os valores de expressão V de cada gene entre as duas idades, o que resultou no número de vezes que determinado gene era mais expresso numa idade ou em outra, e com estas razões foram construídos gráficos para melhor visualização (Figura 25).

Tabela 16 - Valores de Eficiência de Amplificação (*E*) e CTs médios obtidos para 25 genes em amostras coletadas da região cambial de *E. grandis* nas fases juvenil e adulta de desenvolvimento. *E*, eficiência de amplificação; *r*, correlação; *E*, eficiência de amplificação calculada; CT, *threshold cycle*; s, desvio padrão

Gene	<i>E</i>		<i>r</i>		<i>E</i>		CT $\pm$ s*	
	juvenil	adulta	juvenil	adulta	juvenil	adulta	juvenil	adulta
Poligalacturonase	0,68	0,73	0,95	0,96	1,63	1,69	19,940 $\pm$ 0,098489	25,197 $\pm$ 0,143643
Pectato Liase	0,83	0,80	1,00	1,00	1,83	1,80	24,015 $\pm$ 0,077782	28,785 $\pm$ 0,318198
Pectinesterase	0,92	0,81	0,93	0,99	1,85	1,80	23,520 $\pm$ 0,056569	26,060 $\pm$ 0,579828
Xiloglicana Endotransglicosilase 2 (XTH2)	0,99	0,99	0,90	0,92	1,89	1,91	21,780 $\pm$ 0	22,065 $\pm$ 0,06364
Xiloglicana Endotransglicosilase 3 (XTH3)	0,79	0,75	1,00	1,00	1,79	1,75	24,560 $\pm$ 0,056569	20,805 $\pm$ 0,13435
Xilanase	0,86	0,85	0,94	0,96	1,80	1,81	27,020 $\pm$ 0,028284	29,995 $\pm$ 0,247487
UDP-glicose pirofosforilase (UGPase)	0,94	0,93	0,97	0,98	1,91	1,91	17,350 $\pm$ 0,20664	20,490 $\pm$ 0,334515
UDP-xilose sintase 3 (UXS3)	0,70	0,73	1,00	1,00	1,70	1,73	24,790 $\pm$ 0	23,880 $\pm$ 0,028284
Arabinose 4-epimerase	0,91	1,06	0,93	0,88	1,84	1,94	25,885 $\pm$ 0,007071	23,300 $\pm$ 0,084853
Sacarose sintase (SuSy)	0,89	0,78	0,93	0,99	1,82	1,77	17,985 $\pm$ 0,219203	22,570 $\pm$ 0,098995
Endo 1,4- β glicanase (KOR)	0,78	0,81	0,97	0,98	1,75	1,79	20,720 $\pm$ 0,028284	25,480 $\pm$ 0
Celulose sintase 1	0,97	1,01	0,94	0,92	1,91	1,93	18,930 $\pm$ 0,070711	22,455 $\pm$ 0,120208
Celulose sintase 2	1,07	0,98	0,91	0,98	1,98	1,96	17,865 $\pm$ 0,035355	22,290 $\pm$ 0,028284
Celulose sintase 3	1,09	0,95	0,91	0,97	2,00	1,92	18,490 $\pm$ 0,042426	23,395 $\pm$ 0,445477
α -Tubulina 2	0,90	1,10	0,93	0,93	1,83	2,03	25,200 $\pm$ 0,183848	24,180 $\pm$ 0,028284
α-Tubulina 5	0,97	0,95	0,91	0,94	1,88	1,89	20,720 $\pm$ 0	21,650 $\pm$ 0,056569
α -Tubulina 6	0,98	0,98	0,98	0,99	1,96	1,97	23,110 $\pm$ 0,098995	21,190 $\pm$ 0,113137
β -Tubulina 8	0,93	0,97	0,96	0,93	1,89	1,90	26,135 $\pm$ 0,205061	21,930 $\pm$ 0,070711
β -Tubulina 3	0,89	1,21	0,96	0,93	1,85	2,14	24,860 $\pm$ 0	20,825 $\pm$ 0,021213
Fenilalanina amônia liase (PAL)	0,79	0,78	0,94	0,97	1,73	1,75	26,820 $\pm$ 0,155563	27,630 $\pm$ 0,127279
4-coumarato: coenzima A ligase (4CL)	0,78	0,93	0,97	0,96	1,75	1,89	20,540 $\pm$ 0	22,135 $\pm$ 0,007071
Hidroxicinamil transferase (HCT)	0,86	0,87	0,94	1,00	1,80	1,87	21,705 $\pm$ 0,106066	26,500 $\pm$ 0,098995
Ferulato 5-hidroxilase (F5H)	0,93	0,90	0,94	1,00	1,87	1,9	24,130 $\pm$ 0,650538	27,860 $\pm$ 0
Cafeil O-metil transferase (COMT)	0,86	0,84	0,96	0,98	1,82	1,82	19,420 $\pm$ 0,084853	22,700 $\pm$ 0,296985
Cafeil-CoA O-metiltransferase (CCoAOMT)	0,90	0,84	0,93	0,99	1,83	1,83	19,985 $\pm$ 0,007071	22,490 $\pm$ 0,19799
Álcool desidrogenase	0,97	0,80	0,93	0,92	1,90	1,72	27,243 $\pm$ 0,102632	21,803 $\pm$ 0,176163

* O desvio padrão de cada CT foi calculado com base nas três repetições da diluição 0,0001 vez

O maior teor de lignina descrito para as árvores em fase juvenil de desenvolvimento foi confirmado através dos resultados obtidos a partir de qRT-PCR, onde todos os genes envolvidos na síntese de lignina selecionados tiveram maiores níveis de expressão na região cambial das árvores juvenis (Figura 25a). A menor diferença observada para o gene que codifica a PAL pode ser explicada pelo fato de que esta enzima é essencial para a via dos fenilpropanóides, os quais, além de serem os precursores dos monolignóis que compõem a lignina, também podem participar

em outras vias do metabolismo secundário das plantas e, portanto, devem ter seus níveis mantidos independentemente do nível de lignificação apresentado por árvores em diferentes estádios de desenvolvimento.

Entre os genes envolvidos no processo de modificação, ou remodelagem, da parede celular selecionados, dois estão envolvidos na degradação e um na modificação de pectinas, polissacarídeos presentes em grande quantidade na parede celular primária, e três na hidrólise de xilanas e xiloglicanas, principais componentes da hemicelulose em *Eucalyptus*. Como uma das principais características observadas em madeira adulta é o maior comprimento e largura das fibras, esperava-se um aumento nos níveis de expressão dessas enzimas, já que através de suas atividades, novos polissacarídeos poderiam ser incorporados à parede primária e, conseqüentemente, levariam a um aumento no tamanho das células cambiais que dão origem às fibras. Entretanto, a quantificação relativa dos níveis de expressão desses genes mostrou que a poligalacturonase, a pectato liase e uma xilanase foram 13, 5,5 e 3,36 vezes, respectivamente, mais expressas nas árvores jovens que nas adultas (Figura 25b). Se a degradação de pectinas é maior nas árvores jovens e estas possuem células de menor tamanho, talvez este não seja o mecanismo determinante para o aumento tanto no comprimento, quanto na largura das fibras da madeira adulta.

Por outro lado, enquanto a XTH2 apresentou níveis próximos de expressão entre as idades, a XTH3 foi 28,7 vezes mais expressa nas árvores adultas que nas jovens, sugerindo a existência de uma predominância enzimática no remodelamento da parede celular que ocorre durante a formação da madeira adulta, a qual possui fibras mais longas que a madeira juvenil. Na parede primária, as xiloglicanas são quebradas em suas ligações com as microfibrilas pela XET, ou XTH, e essa desconecção permite uma movimentação das microfibrilas, levando a parede a se expandir. Além disso, é a XTH que faz a ligação de novas xiloglicanas na matriz da parede. Em adição, a atividade da XTH foi demonstrada estar ocorrendo em células sofrendo processo de formação de parede secundária, com níveis maiores na região entre a parede primária e a camada S1 da parede secundária (BOURQUIN et al., 2002). Através dessas observações, foi sugerido um novo papel para as XTH, no qual elas estariam ligando xiloglicanas à matriz da parede celular mesmo após terminada sua expansão, reestruturando e reforçando a parede primária ao mesmo tempo que a secundária é depositada e, uma vez que os ângulos de deposição das microfibrilas é diferente entre essas duas camadas, as XTH também estariam auxiliando a ligação entre elas.

A expressão dos genes que codificam a UXS3 e a Arabinose 4-epimerase, enzimas envolvidas na biossíntese das pentoses xilose e arabinose foi analisada. Ambos mostraram-se mais expressos nas árvores adultas numa razão de 2,1 e 2,8 vezes mais, respectivamente (Figura 25c), indicando uma maior participação desses genes na síntese de hemiceluloses nessas árvores. A xilose é o principal açúcar da hemicelulose de *Eucalyptus*, já que sua madeira é rica em xiloglicanas, e a arabinose pode ser usada tanto para produção de arabinoxilanas quanto formar cadeias laterais de pectina.

Já, para todos os genes envolvidos com o complexo de biossíntese de celulose da parede secundária, entre eles Susy, KOR, CesA1, CesA2 e CesA3, os maiores níveis de expressão foram observados nas árvores jovens (Figura 25c). Este foi um resultado inesperado já que as fibras da madeira adulta caracterizam-se por uma maior quantidade de celulose em sua parede secundária, apresentando, inclusive uma maior espessura da parede e menor lúmen celular. O mesmo comportamento foi verificado para a UGPase, outra enzima responsável pelo fornecimento de UDP-glicose à célula. Diferente da Susy, que parece ter um papel específico na geração de UDP-glicose para a síntese de celulose, a UGPase é considerada uma enzima de metabolismo primário de carboidratos e portanto, as árvores jovens podem estar sob uma maior taxa metabólica. De fato, árvores em fase juvenil de desenvolvimento possuem uma alta taxa de crescimento e em seus primeiros anos de vida focalizam sua energia para uma rápida produção de biomassa. Como as fibras são compostas basicamente de uma parede celular contendo microfibrilas de celulose, os maiores níveis de expressão observados para estes genes podem ser o reflexo das altas taxas de produção de celulose para um grande número de células em diferenciação, o que não implica necessariamente numa maior quantidade de celulose por fibra.

Figura 25 - Quantificação relativa dos níveis de expressão de 25 genes por qRT-PCR entre árvores jovens e adultas. (a) resultados referentes a genes envolvidos na modificação da parede celular; (b) resultados referentes a genes envolvidos na lignificação; (c) resultados referentes a genes envolvidos na biossíntese de polissacarídeos; (d) resultado referente ao gene que codifica a enzima álcool desidrogenase. Poligal., Poligalacturonase; PE, Pectinesterase; XET2, Xiloglicana endotransglicosilase 2; XET3, Xiloglicana endotransglicosilase 3; PAL, fenilalanina amônia-liase; 4CL, 4-coumarato: coenzima A ligase; HCT, hidroxicinamil-CoA shiquimato/quinato hidroxicinamil-transferase; F5H, ferulato/coniferaldeído 5-hidroxilase; COMT, cafeato/5-hidroxi coniferaldeído O-metiltransferase; CCoAOMT, cafeil-coenzima A O-metiltransferase; UGPase, UDP-glicose pirofosforilase; UXS3, UDP-xilose sintase 3; Susy, Sacarose sintase; KOR, Endo 1,4- β glicanase; CesA1, Celulose sintase 1; CesA2, Celulose sintase 2; CesA3, Celulose sintase 3

Uma característica muito importante na comparação entre fibras de madeira juvenil e adulta é o ângulo de deposição das microfibrilas de celulose na camada S2 da parede secundária, o qual é definido como o ângulo que as microfibrilas formam com o eixo da fibra quando elas são incorporadas. Este ângulo tem grande influência sobre as características da madeira, em particular sua dureza, ou densidade, e é inversamente proporcional ao comprimento das fibras, isto é, quanto mais comprida uma fibra, menor será o ângulo de deposição de suas microfibrilas e consequentemente ela terá uma maior densidade básica (Barnett e Bonham, 2004). Esta deposição padronizada de celulose na camada S2 da parede secundária está associada com o comportamento dos microtúbulos que direcionam a deposição das cadeias de celulose recém formadas (ODA et al., 2005). Para verificar se estes componentes do citoesqueleto estão sob diferentes níveis de expressão em árvores juvenis e adultas, 4 sequências encontradas na biblioteca subtrativa de ESTs de madeira juvenil vs adulta e adulta vs juvenil, disponível publicamente no NCBI, foram selecionadas para este estudo. Entre elas estão 2 α -tubulinas e 2 β -tubulinas, as quais apresentaram alta similaridade aos genes α -Tubulina 2 e 6 e β -Tubulina 3 e 8 de *Arabidopsis*, respectivamente. Através do qRT-PCR foi possível observar um padrão de expressão muito interessante entre estes genes, onde a α -Tubulina 2 teve maior expressão nas árvores jovens e a α -Tubulina 6 foi 6,6 vezes maior nas árvores adultas. Da mesma forma, uma das β -Tubulinas, a β -Tubulina 8 apresentou um nível de expressão 26,1 vezes maior nas árvores adultas, enquanto que a outra praticamente não mostrou diferença entre as idades (Figura 25c). Estes resultados sugerem que tipos específicos de tubulinas podem atuar na deposição das microfibrilas durante a xilogênese nas fases de desenvolvimento do *Eucalyptus grandis* aqui estudadas, produzindo madeira com distintas características físicas e anatômicas.

A maior e mais intrigante diferença nos níveis de expressão entre as regiões cambiais de árvores juvenis e adultas foi verificada para o gene que codifica uma álcool desidrogenase, que foi 714,8 vezes maior nas árvores adultas (Figura 25d). As células da região cambial, ao contrário dos tecidos fotossintetizantes, apresentam um intenso metabolismo heterotrófico a partir do qual é produzida a energia necessária para a manutenção de suas elevadas taxas de crescimento, observadas principalmente durante o crescimento secundário. Entretanto, neste tecido existe uma baixa pressão de oxigênio imposta pela barreira física criada pela casca e pelo seu consumo durante a respiração. Dessa forma, a energia necessária nesta região poderia ser gerada através da fermentação alcoólica (KIMMERER e STRINGER, 1988). Uma vez que a casca das árvores

adultas é muito mais espessa que a de árvores juvenis e que sua madeira é muito mais densa, a necessidade de produzir energia através da fermentação alcoólica também deve ser muito maior nestas células. Além disso, Gartner (1996) levantou a hipótese de que a produção de madeira juvenil é influenciada pela presença da casca fotossintetizante, existente em muitas espécies arbóreas, e que a transição de madeira juvenil para madeira adulta ocorre concomitantemente com a transição da casca fotossintetizante para a casca não fotossintetizante. Segundo Pfanz et al. (2002), o tronco das árvores possui casca fotossintetizante e cloroplastos funcionais em células do parênquima radial e da medula capazes de usar o CO_2 interno e a luz penetrante para produção de açúcar. Esta re-fixação interna de CO_2 pode compensar em 60-90% a perda de carbono que ocorre durante a respiração, produzindo O_2 . Se nas árvores adultas a casca já não pode realizar esse processo, provavelmente a respiração anaeróbica tenha que ser aumentada, o que ficou evidenciado pelos resultados obtidos.

Em conjunto, esses resultados indicam que existem genes específicos agindo durante a fase adulta de desenvolvimento das células cambiais, onde a XTH3 poderia ser a enzima responsável pelo remodelamento diferencial da parede celular, a qual ao se expandir mais que as fibras juvenis, estaria usando os polissacarídeos provenientes do UDP-glucuronato, através da ação da UXS3, para alongar as cadeias de hemicelulose e ligá-las às microfibrilas de celulose. Ao mesmo tempo em que essas células atingem um maior comprimento e maior diâmetro, uma α -tubulina 6 e uma β -tubulina 8 devem estar induzindo uma mudança no ângulo de deposição das microfibrilas de celulose na parede celular secundária das árvores adultas, permitindo a deposição de uma maior quantidade de celulose e, conseqüentemente, uma redução dos espaços interfibrilares. Assim, seria necessária uma menor quantidade de lignina para preencher esses espaços, como sugerido pelos resultados obtidos. Ainda, como o metabolismo das árvores adultas é bem menor do que o das juvenis, a deposição de celulose deve estar ocorrendo de forma mais lenta durante a diferenciação das células e, por isto, não haveria a necessidade de um aumento na expressão dos genes envolvidos neste processo. Finalmente, para garantir a produção de energia necessária à manutenção e desenvolvimento dessas células, as árvores adultas parecem precisar muito mais da fonte de ATP via fermentação alcoólica do que as juvenis, que ainda têm a capacidade de fixar CO_2 através da fotossíntese e produzir O_2 para ser usado durante a respiração.

5 CONSIDERAÇÕES FINAIS

A construção da biblioteca SAGE a partir da região cambial de *Eucalyptus grandis* em fase juvenil de desenvolvimento foi bastante eficiente na obtenção do perfil transcricional deste tecido. Ao todo, foram geradas 22660 SAGE *tags*, que representaram 2942 diferentes transcritos com número de cópias maior ou igual a 2. Desses, 874 puderam ser associados a uma provável função e classificados funcionalmente, o que revelou uma alta expressão para muitos genes envolvidos com a biossíntese da madeira, além de genes que participam de processos regulatórios e do metabolismo geral das plantas. Com essas informações, foi possível analisar o perfil transcricional dos genes envolvidos diretamente com os estádios da diferenciação que ocorrem durante a xilogênese e identificar pontos determinantes de suas vias metabólicas, onde existem grupos de genes apresentando diferentes padrões de expressão.

A técnica SAGE, utilizada neste trabalho, ainda proporcionou a exploração de dados que dificilmente são encontrados por outras técnicas da genômica funcional, como a detecção de genes desconhecidos, genes com baixo nível de expressão, transcritos anti-sense e transcritos alternativos. Assim, as 143 *tags* únicas, ou genes, classificadas na categoria 6.2 Sequência expressa/Sem anotação, podem representar genes potencialmente específicos de *Eucalyptus* e, 152 *tags* classificadas na categoria 6.1 Não caracterizados/hipotéticos e 1773 *tags* que não foram ao menos associadas a uma sequência de *Eucalyptus*, sugerem que ainda existem muitos genes que podem ter um papel importante no metabolismo do tecido aqui estudado, e que precisam ser caracterizados. Entre as *tags* que mostraram um baixo número de cópias, 565 foram associadas a uma provável função, permitindo a identificação de grupos de genes que ocorrem num baixo nível de expressão. A busca por sequências complementares invertidas e por *tags* presentes em sítios CATG alternativos levou à identificação de 103 potenciais transcritos anti-sense e 182 potenciais transcritos alternativos, respectivamente. Esses dois procedimentos tornaram possível a associação de mais 285 *tags* a sequências de *Eucalyptus*, das quais 173 apresentaram similaridade a proteínas de função conhecida. Portanto, as 1169 *tags* que foram inicialmente associadas a um EST, passaram a totalizar 1454 *tags* (aumento de 40% para 49,4% das *tags*) e, as 874 para as quais prováveis funções foram atribuídas, passaram a ser 1047, o que corresponde a 36% do total das *tags* analisadas.

O conjunto dos resultados produzidos a partir do SAGE possibilitou o estabelecimento de uma grande plataforma de conhecimentos, a qual tem como finalidade apontar para investigações aplicadas, servindo como uma base de dados referente à formação da madeira de *Eucalyptus grandis*. Assim, os dados do transcritoma apontaram para um estudo envolvendo a qualidade da madeira de *E. grandis*, o qual comparou o padrão de expressão de 25 genes entre a região cambial de árvores juvenis e adultas através de qRT-PCR e permitiu a identificação daqueles diferencialmente expressos. Entre eles, 15 genes envolvidos com lignificação, síntese de celulose e degradação de pectinas, foram mais expressos na amostra de árvores juvenis e 6 genes envolvidos com a síntese de hemicelulose (UXS3 e arabinose 4-epimerase), modificação de parede celular (XTH3), deposição de microfibrilas de celulose (α -tubulina 6 e β -tubulina 8) e fermentação alcoólica, ou respiração anaeróbica (álcool desidrogenase), foram mais expressos na amostra de árvores adultas. De maneira geral, esses resultados indicaram que as características apresentadas pelas árvores adultas, que conferem maior qualidade à madeira, podem ser resultantes da ação conjunta de genes que regulam a expansão celular através do remodelamento da parede primária e o ângulo de deposição das microfibrilas de celulose, além genes que garantem o suprimento de energia para esse tecido.

REFERÊNCIAS

- A INDUSTRIALIZAÇÃO de produtos madeireiros. **AMBIENTE BRASIL - FLORESTAL**. Disponível em: < <http://www.ambientebrasil.com.br/composer.php3?base=/florestalo=/index.html&conteud./florestal/industrializacao.html> >. Acesso em: 1 jul. 2007.
- ABE, H.; FUNADA, R.; THTANI, J.; FUKAZAWA, K. Changes in the arrangement of cellulose microfibrils associated with the cessation of cell expansion in tracheids. **Trees**, Berlin, v. 11, p. 328-332, 1997.
- ALBA, R.; FEI, Z.; PAYTON, P.; LIU, Y.; MOORE, S.L.; DEBBIE, P.; COHN, J.; D'ASCENZO, M.; GORDON, J.S.; ROSE, J.K.C.; MARTIN, G.; TANKSLEY, S.D.; BOUZAYEN, M.; JAHN, M.M.; GIOVANNONI, G. ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. **The Plant Journal**, Oxford, v. 39, n. 5, p. 697-714, 2004.
- ALLINA, S.M.; PRI-HADASH, A.; THEILMANN, D. A.; ELLIS, B. E.; DOUGLAS, C. J. 4-Coumarate:coenzyme A ligase in hybrid poplar. Properties of native enzymes, cDNA cloning, and analysis of recombinant enzymes. **Plant Physiology**, Lancaster, v. 116, n. 2, p. 743-754, 1998.
- ALLONA, I.; QUINN, M.; SHOOP, E.; SWOPE, K.; CYR, S.S.; CARLIS, J.; RIEDL, J.; RETZEL, E.; CAMPBELL, M.M.; SEDEROFF, R.; WHETTEN, R.W. Analysis of xylem formation in pine by cDNA sequencing. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 95, p. 9693-9698, 1998.
- ALTSCHUL, S.F.; MADDEN, T.L.; SCHAFER, A.A.; ZHANG, J.H.; ZHANG, Z.; MILLER, W.; LIPMAN, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, London, v. 25, n. 17, p. 3389-3402, 1997.
- ANDERSSON-GUNNERÅS, 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**, Oxford, v. 45, p. 144-165, 2006.
- ANDRADE, E.N. **O Eucalipto**. 2.ed. Jundiaí: Companhia Paulista de Estradas de Ferro, 1961. 667 p.
- ASSIS, T.F. Hibridação na Acesita Energética S.A. **IPEF**, Piracicaba, p. 21-38, 1985. Suplemento. Apresentado no SIMPÓSIO SOBRE TENDÊNCIAS DO DESENVOLVIMENTO FLORESTAL BRASILEIRO, 1985, Piracicaba.
- ASSIS, T.F.; MAFIA, R.G. Hibridação e clonagem de *Eucalyptus*. In: BORÉM, A. (Ed.) **Biotecnologia Florestal**. Viçosa: UFV, 2007, p. 93-121.

BACHEM, C.; VAN DER HOEVEN, R.; DE BRUIJN, S.; VREUGDENHIL, D.; ZABEAU, M.; VISSER, R. Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. **The Plant Journal**, Oxford, v. 9, p. 745–753, 1996.

BANSAL, A.; SANKARARAMAKRISHNAN, R. Homology modeling of major intrinsic proteins in rice, maize and Arabidopsis: Comparative analysis of transmembrane helix association and aromatic/arginine selectivity filters. **BMC Structural Biology**, London, v. 7, p.27, 2007.

BAO, J.; LEE, S.; CHEN, C.; ZHANG, X.; ZHANG, Y.; LIU, S.; CLARK, T.; WANG, J.; CAO, M.; YANG, H.; WANG, S.M.; YU, J. Serial analysis of gene expression study of a hybrid rice strain (LYP9) and its parental cultivars. **Plant Physiology**, Lancaster, v.138, p. 1216-1231, 2005.

BARNETT, J.R. e BONHAM, V.A. Cellulose microfibril angle in the cell wall of wood fibres. **Biological Reviews**, United Kingdom, v. 79, p. 461-472, 2004.

BAUCHER, M.; HALPIN, C.; PETIT-CONIL, M.; BOERJAN, W. Lignin: Genetic engineering and impact on pulping. **Critical reviews in biochemistry and molecular biology**, Boca Raton, v. 38, p. 305-350, 2003.

BEKRI, M.A.; DESAIR, J.; KEIJERS, V.; PROOST, P.; SEARLE-VAN LEEUWEN, M.; VANDERLEYDEN, J., VANDE BROEK, A. Azospirillum irakense produces a novel type of pectate lyase. **Journal of Bacteriology**, Baltimore, v. 181, n. 8, p. 2440-2447, 1999.

BENDTSEN, B.A. Properties of wood from improved and intensively managed trees. **Forest Products Journal**, Madison, v. 28, n. 10, p. 61-72, 1978.

BHALERAO, R.; KESKITALO, J.; STERKY, F.; ERLANDSSON, R.; BJORKBACKA, H.; BIRVE, S.J.; KARLSSON, J.; GARDESTROM, P.; GUSTAFSSON, P.; LUNDEBERG, J.; JANSSEN, S. Gene expression in autumn leaves. **Plant Physiology**, Lancaster, v.131, n. 2, p. 430-442, 2003.

BHALLA, R.; NARASIMHAN, K.; SWARUP, S. Metabolomics and its role in understanding cellular responses in plants. **Plant Cell Reports**, Berlin, v. 24, n. 10, p. 562-71, 2005.

BOMMER, U.A.; THIELE, B.J. The translationally controlled tumour protein (TCTP). **The international journal of biochemistry & cell biology**, Exeter, v. 36, p. 379-85, 2004.

BONA, C.; BOEGER, M.A.; SANTOS, G. de O. **Guia ilustrado de anatomia vegetal**. Ribeirão Preto: Holos, 2004. 80 p.

BOONBURAPONG, B.; BUABOOCHA, T. Genome-wide identification and analyses of the rice calmodulin and related potential calcium sensor proteins. **BMC Plant Biology**, London, v. 7, p. 1471-2229, 2007.

BOURQUIN, V.; NISHIKUBO, N.; ABE, H.; BRUMER, H.; DENMAN, S.; EKLUND, M.; CHRISTIERNIN, M.; TEERI, T.T.; SUNDBERG, B.; MELLEROWICZ, E.J. Xyloglucan endotransglycosylases have a function during the formation of secondary cell walls of vascular tissues. **The Plant Cell**, Rockville, v. 14, p. 3073-3088, 2002.

BOUTON, S.; LEBOEUF, E.; MOUILLE, G.; LEYDECKER, M.; TALBOTEC, J.; GRANIER, F.; LAHAYE, M.; HÖFTE, H.; TRUONG, H. *QUASIMODO1* Encodes a putative membrane-bound glycosyltransferase required for normal pectin synthesis and cell adhesion in *Arabidopsis*. **The Plant Cell**, Rockville, v. 14, p. 2577-2590, 2002.

BRACELPA. **Desempenho do setor em 2006 e projeção para 2007**. Disponível em: < <http://www.bracelpa.com.br> > . Acesso em: 23 maio 2007.

BRASIL se prepara para o mercado de redução de emissões. **Revista da Madeira**, Curitiba, v. 15, n. 93, nov. 2005. Disponível em: < <http://www.remade.com.br/revista> > Acesso em: 10 jun. 2007.

BRASIL, M.A.M.; FERREIRA, M. Características das fibras de madeira de *Eucalyptus grandis* Hill ex maiden, aos 3 anos de idade. **IPEF**, Piracicaba, n. 19, p. 80-97, dez. 1979.

BRENNER, S.; JOHNSON, M.; BRIDGHAM, J.; GOLDA, G.; LLOYD, D.H.; JOHNSON, D. LUO, S.; McCURDY, S.; FOY, M.; EWAN, M.; ROTH, R.; GEORGE, D.; ELETR, S.; ALBRECHT, G.; VERMAAS, E.; WILLIAMS, S.R.; MOON, K.; BURCHAM, T.; PALLAS, M.; DUBRIDGE, R.B.; KIRCHNER, J.; FEARON, K.; MAO, J.; CORCORAN, K. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. **Nature Biotechnology**, New York, v.18, n. 6, p. 630-4, 2000.

BRIGATTI, R.A.; FERREIRA, M.; BEIG, O.; FREITAS, M. Polinização controlada em *Eucalyptus urophylla* - um programa desenvolvido pela Champion Papel e Celulose S.A. **Silvicultura**, São Paulo, v. 8, n. 28, p. 213-215, 1983.

BRONDANI, R.P.V.; BRONDANI, C.; TARCHINI, R.; GRATTAPAGLIA, D. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. **Theoretical and Applied Genetics**, Berlin, v. 97, p. 816-827, 1998.

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

BROWN, G.R.; GILL, G.P.; KUNTZ, R.J.; LANGLEY, C.H.; NEALE, D.B. Nucleotide diversity and linkage disequilibrium in loblolly pine. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 101, p. 15255-15260, 2004.

BURLAT, V.; KWON, M.; DAVIN, L. B.; LEWIS, N. G. Dirigent proteins and dirigent sites in lignifying tissues. **Phytochemistry**, Oxford, v. 57, p. 883-897, 2001.

BYRNE, M.; MURRELL, J.C.; ALLEN, B.; MORAN, G.F. An integrated genetic linkage map for eucalypts using RFLP, RAPD and isozyme markers. **Theoretical and Applied Genetics**, Berlin, v. 91, p. 869-875, 1995.

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**, Oxford, v. 1, p. 141-154, 2003.

CANTÓN, F. R; SUÁREZ, M. F.; CÁNOVAS, F. M. Molecular aspects of nitrogen mobilization and recycling in trees. **Photosynthesis Research**, Hague, v. 83, p. 265-278, 2005.

CARNEIRO, R.T. **Análise serial da expressão gênica do caule de plantas de *Eucalyptus grandis* com 6 meses de idade**. 2007. 182p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2007.

CARPITA, N.C.; GIBEAUT, D.M. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. **The Plant Journal**, Oxford, v.3, p. 1-30, 1993.

CARPITA, N.C.; McCANN, M. The cell wall. In: BUCHANAN, B.B.; GRUISSEM, W.; JONES, R.L. (Eds.). **Biochemistry & Molecular Biology of Plants**, Rockville: American Society of Plant Physiologists, 2000, cap. 2, p. 52-108.

CARR, S.G.M.; CARR, D.J.; ROSS, F.L. Male flowers in *Eucalyptus*. **Australian Journal of Botany**, Melbourne, v.19, n.1, p. 73-83, 1971.

CARVALHO, M.C.C.G. **Caracterização de perfis transcricionais de folhas e da região cambial de *Eucalyptus grandis* usando o SAGE**. 2006. 153p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2007.

CATESSON, A.M. Specific characters of vessel primary walls during the early stages of wood differentiation. **Biology of the Cell**, Bethesda, v. 67, p. 221-226, 1989.

CATO, S.; McMILLAN, L.; DONALDSON, L.; RICHARDSON, T.; ECHT, C.; GARDNER, R. Wood formation from the base to the crown in *Pinus radiata*: gradients of tracheid wall thickness, wood density, radial growth rate and gene expression. **Plant Molecular Biology**, New York, v. 60, p. 565-581, 2006.

CELEDON, P.A.F.; ANDRADE, A.; MEIRELES, K.G.X.; CARVALHO, M.C.C.G.; CALDAS, D.G.G.; MOON, D.H.; CARNEIRO, R.T.; FRANCASCHINI, L.M.; ODA, S.; LABATE, C.A. Proteomic analysis of the cambial region in juvenile *Eucalyptus grandis* at three ages. **Proteomics**, Weinheim, v. 7, p. 2258-2274, 2007.

CHAFFEY, N. BARLOW, P.; BARNETT, J. Cortical microtubules rearrange during differentiation of vascular cambial derivatives, microfilaments do not. **Trees – Structure and Functions**, Berlin, v.11, p. 333-341, 1997.

CHAUMONT, F.; BARRIEU, F.; HERMAN, E.M.; CHRISPEELS, M.J. Characterization of a Maize Tonoplast Aquaporin Expressed in Zones of Cell Division and Elongation1. **Plant Physiology**, Lancaster, v. 117, p. 1143–1152, 1998.

CHEN, J.; SUN, M.; LEE, S.; ZHOU, G.; ROWLEY, J.; WANG, S. Identifying novel transcripts and novel genes in the human genome by using novel SAGE tags. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 99, p. 12257-12262, 2002.

CHEN, S.; HARMON, A.C. Advances in plant proteomics. **Proteomics**, Weinheim, v. 6, n. 20, p. 5504-16, 2006.

COEMANS, B.; MATSUMURA, H.; TERAUCHI, R.; REMY, S.; SWENNEN, R.; SAGI, L. SuperSAGE combined with PCR walking allows global gene expression profiling of banana (*Musa acuminata*), a non-model organism. **Theoretical and Applied Genetics**, Berlin, v. 111, n. 6, p. 1118-26, 2005.

COMPANHIA PAULISTA DE ESTRADAS DE FERRO. **O eucalipto**. Disponível em: <<http://www.geocities.com/motorcity/street/2376/eucalipto.html>>. Acesso em: 01 jul. 2006.

CONNELLY, C.; HIETER, P. Budding Yeast SKP1 Encodes an Evolutionarily Conserved Kinetochore Protein Required for Cell Cycle Progression. **Cell**, Cambridge, v. 86, p. 275–285, 1996.

COOKE, R.; RAYNAL, M.; LAUDIE, M. Further progress towards a catalogue of all *Arabidopsis* genes: analysis of a set of 5000 non-redundant ESTs. **The Plant Journal**, Oxford, v. 9, p. 101-124, 1996.

COZZO, D. **Eucalyptus y eucaliptotecnia**. Buenos Aires: El Atenero Editorial, 1955. 393 p.

CRESSWELL, R.; NITSCH, C. Organ culture of *Eucalyptus grandis* L. **Planta**, Berlin, v. 125, p. 87-90, 1975.

DALESSANDRO, G.; NORTHCOTE, D.H. Possible control sites of polysaccharide synthesis during cell growth and wall expansion of pea seedlings (*Pisum sativum* L.). **Planta**, Berlin, v. 134, p. 39-44, 1977.

DARLEY, C.P.; FORRESTER, A.M.; McQUEEN-MASON, S.J. The molecular basis of plant cell wall. **Plant Molecular biology**, New York, v. 47, p. 179-195, 2001.

DELUC, L.; BARRIEU, F.; MARCHIVE, C.; LAUVERGEAT, V.; DECENDIT, A.; RICHARD, T.; CARDE, J.P.; MERILLON, J.M.; HAMDI, S. Characterization of a grapevine R2R3-MYB transcription factor that regulates the phenylpropanoid pathway. **Plant Physiology**, Lancaster, v. 140, p. 499-511, 2006.

DIATCHENKO, L.; LAU, Y.F.; CAMPBELL, A.P.; CHENCHIK, A.; MOQADAM, F.; HUANG, B.; LUKYANOV, S.; LUKYANOV, K.; GURSKAYA, N.; SVERDLOV, E.D.; SIEBERT, P.D. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 93, n. 12, p. 6025-30, 1996.

DIXON, R.A. Natural products and plant disease resistance. **Nature**, London, v. 411, p. 843-847, 2001.

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

EWING, B.; HILLIER, L.; WENDL, M.C.; GREEN, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. **Genome Research**, Cold Spring Harbor, v. 8, n. 3, p.175-185, 1998.

EVANS II, J.W.; SENFT, J.F.; GREEN, D.W. Juvenile wood effect in red alder: analysis of physical and mechanical data to delineate juvenile and mature wood zones. **Forests Products Journal**, Madison, v.50, n.7/8, p.75-87, 2000.

FAO. FOOD AND AGRICULTURE ORGANIZATION. **Global Forest resources assessment 2000 - main report**. FAO Forestry paper. Disponível em: < <http://www.fao.org/forestry/fo/fra/main/index.jsp> > Acesso em: 31 maio 2007.

FERL, R.J. 14-3-3 proteins: regulation of signal-induced events. **Physiologia Plantarum**, Lund, v. 120, p. 173-178, 2004.

FERREIRA, S. **Produção e qualidade da madeira serrada de clones de híbridos de *Eucalyptus ssp.*** 2003. 71 p. Dissertação (Mestrado em Ciências Florestais) – Universidade Federal de Lavras, Lavras, 2003.

FERREIRA, M. Melhoramento florestal e silvicultura intensiva com eucalipto. **Silvicultura**, São Paulo, v. 31, p. 5-11, 1983.

FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 2 ed. Brasília: EMBRAPA-CENARGEN, 1995. p. 220.

FLADUNG, M.; BOSSINGER, G.; ROEB, G.W.; SALAMINI, F. Correlated alterations in leaf and flower morphology and rate of leaf photosynthesis in a *midribless* (mbl) mutant of *Panicum maximum* Jaq. **Planta**, Berlin, v. 184, p. 356-361, 1991.

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**, London, v. 170, n. 4, p. 739-752, 2006.

FREEMAN, W.M.; WALKER, S.J.; VRANA, K.E. Quantitative RT-PCR: pitfalls and potential. **Biotechniques**, Natick, v. 26, p. 112-122, 1999.

FUNADA, R. Control of wood structure. In: NICK, P. (Ed.). **Plant Microtubules**, Springer-Verlag, Berlin, p. 51-82, 2000.

GARTNER, B.L. Does photosynthetic bark have a role in the production of core vs. outer wood? **Wood and Fiber Science**, Hanover, v. 28, n. 1, p. 53-61, 1996.

GEISLER-LEE, J.; GEISLER, M.; COUTINHO, P. M.; SEGERMAN, B.; NISHIKUBO, N.; TAKAHASHI, J.; ASPEBORG, H.; DJERBI, S.; MASTER, E.; ANDERSSON-GUNNERÅS, S.; SUNDBERG, B.; KARPINSKI, S.; TEERI, T. T.; KLECZKOWSKI, L. A.; HENRISSAT, B.; MELLEROWICZ, E. J.. Poplar carbohydrate-active enzymes. Gene identification and expression analysis. **Plant Physiology**, Lancaster, v. 140, p. 946-962, 2006.

GIBBINGS, J.G.; COOK, B.P.; DUFAULT, M.R.; MADDEN, S.L.; JHURI, S.; TURNBULL, C. J.; DUNWELL, J.M. Global transcript of rice leaf and seed using SAGE technology. **Plant Biotechnology Journal**, Oxford, v. 1, p. 271-285, 2003.

GONZALEZ, E.R.; de ANDRADE, A.; BERTOLO, A.L.; LACERDA, G.C.; CARNEIRO, R.T.; DEFAVARI, V.A.P.; LABATE, M.T.V.; LABATE, C.A. Production of transgenic *Eucalyptus grandis* x *E. urophylla* using the sonication-assisted *Agrobacterium* transformation (SAAT) system. **Functional Plant Biology**, Collingwood, v. 29, n. 1, p. 97-102, 2002.

GOWDA, M.; JANTASURIYARAT, C.; DEAN, R.A.; WANG, G.L. Robust-LongSAGE (RL-SAGE): a substantially improved LongSAGE method for gene discovery and transcriptome analysis. **Plant Physiology**, Lancaster, v. 134, n. 3, p. 890-897, 2004.

GRATTAPAGLIA, D.; BRADSHAW, J.R. Nuclear DNA content of commercially important *Eucalyptus* species and hybrids. **Canadian Journal of Forest Research**, Ottawa, v. 24, n. 5, p. 1074-1078, may 1994.

GRATTAPAGLIA, D.; SEDEROFF, R.R. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. **Genetics**, Austin, v. 137, p. 1121-1137, 1994.

GRAY, W.M. Hormonal Regulation of Plant Growth and Development. **PLoS Biology**, San Francisco, v. 2, n. 9, p. e311, 2004.

GROOVER, A.; JONES, A.M. Tracheary element differentiation uses a novel mechanism coordinating programmed cell death and secondary cell wall synthesis. **Plant Physiology**, Lancaster, v.119, p. 375-384, 1999.

GRUBER, Y.B.G.; TAVARES, A.C.S.; DETOMINI, E.R.; COELHO, R.D. Eucalipto irrigado tem de produzir mais: quanto é preciso aumentar a produtividade para custear cada sistema de irrigação. In: INSTITUTO FNP. **Agrianual 2006**: anuário da agricultura brasileira. São Paulo, 2006. p. 318-321.

HARPER, A.D.; BAR-PELED, M. Biosynthesis of UDP-Xylose. Cloning and characterization of a novel *Arabidopsis* gene family, *UXS*, encoding soluble and putative membrane-bound UDP-Glucuronic acid decarboxylase isoforms. **Plant Physiology**, Lancaster, v. 130, p. 2188–2198, 2002.

HERTZBERG, M.; ASPEBORG, H.; SCHRADER, J.; ANDERSSON, A.; ERLANDSSON, R.; BLOMQVIST, K.; BHALERAO, R.; UHLÉN, M.; TEERI, 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**, Washington, v. 98, n. 25, p. 14732-14737, 2001.

HOFFMANN, L.; BESSEAU, S.; GEOFFROY, P.; RITZENTHALER, C.; MEYER, D.; LAPIERRE, C.; POLLET, B.; LEGRAND, M. Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinic acid hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. **Plant Cell**, Rockville, v. 16, p. 1446-1465, 2004.

HUANG, R.Z.; WU, X.L.; JIANG, C.Z.; HU, Z.H.; HE, H.Y.; LIOU, Z.H.; XU, Y.H.; JIN, W.; ZHENG, T.; CHEN, J.Q. Isolation, characterization and expression of a gene from *Brassica napus* encoding a LIM-domain protein. **Plant Physiology and Biochemistry**, Paris, 2007. Disponível em: < <http://www.level1diet.com/research/id/322580> > Acesso em: 25 jun. 2007.

ISRAELSSON, M.; ERIKSSON, M. E.; HERTZBERG, M.; ASPEBORG, H.; NILSSON, P.; MORITZ, T. Changes in gene expression in the wood-forming tissue of transgenic hybrid aspen with increased secondary growth. **Plant Molecular Biology**, New York, v. 52, n. 4, p. 893-903, 2003.

JOHANSSON, H.; STERKY, F.; AMINI, B.; LUNDEBERG, J.; KLECZKOWSKI, L.A. Molecular cloning and characterization of a cDNA encoding poplar UDP-glucose dehydrogenase, a key gene of hemicellulose/pectin formation. **Biochimica et biophysica acta**, Amsterdam, v. 1576, p. 53-58, 2002.

JOSHI, C.P.; BHANDARI, S.; RANJAN, P.; KALLURI, U.C.; LIANG, X.; FUJINO, T.; SAMUGA, A. Genomics of cellulose biosynthesis in poplars. **The New phytologist**, London, v. 164, p. 53-61, 2004.

JOUANIN, L.; GOUJON, T.; de NADAI, V.; MARTIN, I.; VALLET, C.; POLLET, B.; YOSHINGA, A.; CHABBERT, B.; PETTIT-CONIL, M.; LAPIERRE, C. Lignification in transgenic Poplars with extremely reduced caffeic acid O-methyltransferase activity. **Plant Physiology**, Lancaster, v. 23, p. 1363-1373, 2000.

JUAN, D.; HONG-LI, X.; DE-QIANG, Z.; MIN-JIE, W.; YING-ZHANG, L.; KE-MING, C.; MENG-ZHU, L. Regeneration of the secondary vascular system in poplar as a novel system to investigate gene expression by a proteomic approach. **Proteomics**, Weinheim, v. 6, p. 881-895, 2006.

KAO, Y.Y.; HARDING, S.A.; TSAI, C.J. Differential expression of two distinct phenylalanine ammonia-lyase genes in condensed tannin-accumulating and lignifying cells of quaking aspen. **Plant Physiology**, Lancaster, v. 130, p. 796-807, 2002.

KAPLAN, B.; DAVYDOV, O.; KNIGHT, H.; GALON, Y.; KNIGHT, M.R.; FLUHR, R.; FROMM, H. Rapid transcriptome changes induced by cytosolic Ca²⁺ transients reveal ABRE-related sequences as Ca²⁺-responsive cis elements in Arabidopsis. **The Plant Cell**, Rockville, v. 18, p. 2733-2748, 2006.

KARPINSKA, B.; KARLSSON, M.; SRIVASTAVA, M.; STENBERG, A.; SCHRADER, J.; STERKY, F.; BHALERAO, R.; WINGSLE, G. MYB transcription factors are differentially expressed and regulated during secondary vascular tissue development in hybrid aspen. **Plant Molecular Biology**, New York, v. 56, p. 255-270, 2004.

KEYS, A. J. The re-assimilation of ammonia produced by photorespiration and the nitrogen economy of C₃ higher plants. **Photosynthesis Research**, Hague, v.87, p. 165-175, 2006.

KIMMERER, T.W.; STRINGER, M.A. Alcohol Dehydrogenase and Ethanol in the Stems of Trees: Evidence for Anaerobic Metabolism in the Vascular Cambium. **Plant Physiology**, Lancaster, v. 87, n. 3, p. 693-697, July, 1988.

LACHAUD, S.; CATESSON, A.M.; BONNEMAIN, J.L. Structure and functions of the vascular cambium. **Comptes rendus de l'Académie des sciences. Série III, Sciences de la vie**, Paris, v. 322, n. 8, p. 633-650, 1999.

LACOMBE, E.; HAWKINS, S.; VAN DOORSSELAERE, J.; PIQUEMAL, J.; GOFFNER, D.; POEYDOMENGE, O.; BOUDET, A. M.; GRIMA-PETTENATI, J. Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. **The Plant Journal**, Oxford, v. 11, n. 3, p. 429-441, 1997.

LARSON, P.R. **The vascular cambium: development and structure**. Berlin: Springer-Verlag, 1994. 351p.

LASH, A.E.; TOLSTOSHEV, C.M.; WAGNER, L.; SCHULER, G.D.; STRAUSBERG, R.L.; RIGGINS, G.J.; ALTSCHUL, S.F. SAGEmap: A public gene expression resource. **Genome Research**, Cold Spring Harbor, v. 10, p. 1051-1060, 2000.

LEITE, N.B. Avanços da silvicultura brasileira são significativos. **Visão Agrícola**, Piracicaba, n. 4, p. 58-61, 2005.

LEWIS, N.G.A. 20th century roller coaster ride: a short account of lignification. **Current Opinion in Plant Biology**, London, v. 2, n. 2, p. 153-162, 1999.

LI, L.; LU, S.; CHIANG, V. A genomic and molecular view of wood formation. **Critical Reviews in Plant Sciences**, Boca Raton, v. 25, p. 215-233, 2006.

LIANG, P.; PARDEE, A. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. **Science**, Washington, v. 257, p. 967-971, 1992.

LODISH, H.; BERK, A.; MATSUDAIRA, P.; KAISER, C.A.; KRIEGER, M.; SCOTT, M.P.; ZIPURSKY, L.; DARNELL, J. **Molecular Cell Biology**. 5. ed. New York: W.H. Freeman and Company, 2004. 973 p.

LORENZ, W.W., DEAN, J.F.D. SAGE Profiling and demonstration of differential gene expression along the axial developmental gradient xylem in loblolly pine (*Pinus taeda*). **Tree Physiology**, Victoria, v. 22, n. 5, p. 301-310, 2002.

LU, S.; ZHOU, Y.; LI, L.; CHIANG, V. L. Distinct roles of cinnamate 4-hydroxylase genes in *Populus*. **Plant & Cell Physiology**, Kyoto, v. 47, n. 7, p. 905-914, 2006.

LUCA, F.C.; WINEY, M. MOB1, an essential yeast gene required for completion of mitosis and maintenance of ploidy. **Molecular Biology of the Cell**, Bethesda, v. 9, p. 29-46, 1998.

MACDOUGALL, G.J.; MORRISON, I.M.; STEWART, D.; WEYERS, J.D.B.; HILLMAN, J.R. Plant fibers: botany, chemistry and processing for industrial use. **Journal of the science of food and agriculture**, London, v. 62, p. 1683-1691, 1993.

MAJEWSKA-SAWKA, A.; NOTHNAGEL, E.A. The multiple roles of arabinogalactan proteins in plant development. **Plant Physiology**, Melbourne, v. 122, p. 3-9, 2000.

MARTINS, F.C.G.; IKEMORI, Y.K. Produção de híbridos de eucalipto na Aracruz. In: REUNIÃO SOBRE TÉCNICAS PARA PRODUÇÃO DE HÍBRIDOS, 1987, Piracicaba. **Anais...** Piracicaba: IPEF, 1987, p. 48-60.

MATHIEU, M.; LELU-WALTER, M.A.; BLERVACQ, A.S.; DAVID, H.; HAWKINS, S.; NEUTELINGS, G. Germin-like genes are expressed during somatic embryogenesis and early development of conifers. **Plant Molecular Biology**, New York, v. 61, p. 615-627, 2006.

MATSUMURA, H.; NIRASAWA, S.; TERAUCHI, R. Transcript profiling in rice (*Oryza sativa* L) seedlings using the serial analysis of gene expression (SAGE). **The Plant Journal**, Oxford, v. 20, n. 6, p. 719-726, 1999.

MATSUMURA, H.; REICH, S.; ITO, A.; SAITOH, H.; KAMOUN, S.; WINTER, P.; KALS, G.; REUTER, M.; KRÜGER, D.H.; TERAUCHI, R. Gene expression analysis of plant host-pathogen interactions by SuperSAGE. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 100, n. 26, p. 15718-15723, 2003.

- MEIRELES, K.G.X. **Identificação de proteínas expressas na região cambial de *Eucalyptus grandis*, por espectrometria de massa**. 2006. 117p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura “Luiz de Querioz”, Universidade de São Paulo, Piracicaba, 2006.
- MELLEROWICZ, E.J.; BAUCHER, M.; SUNDBERG, B.; BOERJAN, W. Unravelling cell wall formation in the woody dicot stem. **Plant Molecular Biology**, New York, v. 47, p. 239-274, 2001.
- MENK, A.L. de M.; ODA, S.; VENKOVSKY, R. Problemas no melhoramento genético clássico do eucalipto em função da alta intensidade de seleção. **IPEF**, Piracicaba, n. 41/42, p. 8-17, jan./dez. 1989.
- MICHELI, F. Pectin methylesterases: cell wall enzymes with important roles in plant physiology. **Trends in Plant Science**, Kidlington, v. 6, n. 9, p. 414-419, 2001.
- MORA, A.; GARCIA, C.H. **A cultura do eucalipto no Brasil**. São Paulo: Sociedade Brasileira de Silvicultura, 2000. 112 p.
- MORALEJO, M.; ROCHANGE, F.; BOUDET, A.M.; TEULIÈRES, C. Generation of transgenic *Eucalyptus globulus* plantlets through *Agrobacterium tumefaciens* mediated transformation. **Australian Journal of Plant Physiology**, Melbourne, v. 25, p. 207-212, 1998.
- MOREAU, C.; AKSENOV, N.; LORENZO, M. G.; SEGERMAN, B.; FUNK, C.; NILSSON, P.; JANSSEN, S.; TUOMINEN, H. A genomic approach to investigate developmental cell death in woody tissues of *Populus* trees. **Genome Biology**, London, v. 6, n. 4, p. 1-14, 2005.
- NAIR, R.B.; XIA, Q.; KARTHA, C.J.; KURYLO, E.; HIRJI, R.N.; DATLA, R.; SELVARAJ, G. Arabidopsis CYP98A3 mediating aromatic 3-hydroxylation. Developmental regulation of the gene, and expression in yeast. **Plant Physiology**, Lancaster, v. 130, p. 210–220, 2002.
- NISHIKUBO, N.; ENDO, S.; TAKAHASHI, J.; BANASIAK, A.; AWANO, T.; ANDERSON, A.; STALBRAND, H.; FUNADA, R.; PIENS, K.; BRUMER, H.; TEERI, T.T.; HAYASHI, T.; SUNDBERG, B.; MELLEROWICZ, E.J. Novel Functions of XETs in poplar wood development. In: IUFRO-TREE BIOTECHNOLOGY MEETING, 2007, Açores. **Anais...** Açores: IUFRO, 2007. Disponível em: <<http://www.itqb.unl.pt/iufro2007/>>. Acesso em: 30 jun. 2007.
- ODA, S.; MELLO, E.J.; SILVA, J.F.; SOUZA, I.C.G. Melhoramento Florestal. In: BORÉM, A. (Ed.). **Biotecnologia Florestal**. Viçosa: Universidade Federal de Viçosa, 2007, cap.6, p. 123-142.
- ODA, Y.; HASEZAWA, S. Cytoskeletal organization during xylem cell differentiation. **Journal of Plant Research**, Tokyo, v. 119, p. 167-177, 2006.
- ODA, Y.; MIMURA, T.; HASEZAWA, S. Regulation of secondary cell wall development by cortical microtubules during tracheary element differentiation in Arabidopsis cell suspensions. **Plant Physiology**, Lancaster, n. 137, p. 1027-1036, 2005.

O EUCALIPTO e suas origens. **Revista da Madeira**, Curitiba, v. 11, n. 59, set. 2001. Disponível em: < <http://www.remade.com.br/revista>>. Acesso em: 12 jun. 2007.

OH, S.; PARK, S.; HAN, K. Transcriptional regulation of secondary growth in *Arabidopsis thaliana*. **Journal of Experimental Botany**, Oxford, v. 54, n. 393, p. 2709-2722, 2003.

OLIVEIRA, J.T.S. **Caracterização da Madeira de eucalipto para construção civil**. 1997. 429p. Tese (Doutorado em Engenharia) - Escola Politécnica, Universidade de São Paulo, São Paulo, 1997.

OLIVEIRA, J.T.S.; HELLMEISTER, J.C.; TOMAZELLO FILHO, M. Variação do teor de umidade e da densidade básica na madeira de sete espécies de eucalipto. **Revista Árvore**, Viçosa, v. 29, n. 1, p. 115-127, 2005.

PARVATHI, K.; CHEN, F.; GUO, D.; BLOUNT, J. W.; DIXON, R. A. Substrate preferences of O-methyltransferases in alfalfa suggest new pathways for 3-O-methylation of monolignols. **The Plant Journal**, Oxford, v. 25, n. 2, p. 193-202, 2001.

PATERSON, S.D; AEBERSOLD, R.H. Proteomics: the first decade and beyond. **Nature genetics**, New York, v. 33, p. 311-323, 2003.

PAUX, E.; CAROCHA, V.; MARQUES, C.; DE SOUSA, A.M.; BORRALHO, N.; SIVADON, P.; GRIMA-PETTENATI, J. Transcript profiling of *Eucalyptus* xylem genes during tension wood formation. **New Phytologist**, London, v. 167, p. 89-100, 2005.

PAUX, E.; TAMASLOUKHT, M.; LADOUCE, N.; SIVADON, P.; GRIMA-PETTENATI, J. Identification of genes preferentially expressed during wood formation in *Eucalyptus*. **Plant Molecular Biology**, The Hague, v. 55, p. 263-280, 2004.

PENG, L.; KAWAGOE, Y.; HOGAN, P.; DELMER, D. Sitosterol-beta-glucoside as primer for cellulose synthesis in plants. **Science**, Washington, v. 295, n. 5552, p. 147-150, 2002.

PEREIRA, H. dos S. Estado actual de la información sobre instituciones forestales. In: FAO. **Estado de la información forestal en Brasil**. Santiago: FAO, 2002. cap. 4, p. 40-75. (FAO. Monografía de Países, 3).

PERSPECTIVAS do setor florestal brasileiro. **Revista da Madeira**, Curitiba, n. 75, ano 13, ago. 2003. Disponível em: <http://www.remade.com.br/revista> > Acesso em: 17 jun. 2007.

PFÄFFL, M.W. A new mathematical model for relative quantification in real-time RT-PCR. **Nucleic Acid Research**, London, v. 29, n. 9, p. 2002-2007, 2001.

PFANZ, H.; ASCHAN, G.; LANGENFELD-HEYSER, R.; WITTMANN, C.; LOOSE, M. Ecology and ecophysiology of tree stems: cortical and wood photosynthesis. **Die Naturwissenschaften**, Berlin, v. 89, n. 4, p. 147-62, 2002.

PILATE, G.; GUINEY, E.; HOLT, K.; PETIT-CONIL, M.; LAPIERRE, C.; LEPLE, J. C.; POLLET, B.; MILA, I.; WEBSTER, E. A.; MARSTORP, H. G.; HOPKINS, D. W.; JOUANIN, L.; BOERJAN, W.; SCHUCH, W.; CORNU, D.; HALPIN, C. Field and pulping performances of transgenic trees with altered lignification. **Nature Biotechnology**, New York, v. 20, p. 607-612, 2002.

PIMENTA, A.M.C. Os desafios do proteoma. **Ciência Hoje**, Rio de Janeiro, v. 32, n. 192, p. 16-22, 2003.

PLEASANCE, E.D.; MARRA, M.A.; JONES, S.J.M. Assessment of SAGE in transcript identification. **Genome Research**, Cold Spring Harbor, v. 13, p. 1203-1215, 2003.

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

POT, D.; CHANTRE, G.; ROZENBERG, P.; RODRIGUES, J.C.; JONES, G.L.; PEREIRA, H.; HANINRUP, B.; CAHALAN, C.; PLOMION, C. Genetic control of pulp and timber properties in maritime pine (*Pinus pinaster* Ait.). **Annals of Forest Science**, Paris, v. 59, n. 5/6, p. 563-575, 2002.

PRYOR, L.D. **The Biology of *Eucalyptus***. London: Edward Arnold, 1976. 82 p.

RAES, J.; ROHDE, A.; CHRISTENSEN, J. H.; PEER, Y. van der; BOERJAN, W. Genome-wide characterization of the lignification toolbox in *Arabidopsis*. **Plant Physiology**, Lancaster, v. 133, p. 1051-1071, 2003.

RALPH, S.; PARK, J.Y.; BOHLMANN, J.; MANSFIELD, S.D. Dirigent proteins in conifer defense: gene discovery, phylogeny, and differential wound- and insect-induced expression of a family of DIR and DIR-like genes in spruce (*Picea* spp.). **Plant Molecular Biology**, The Hague, v. 60, n. 1, p. 21-40, 2006.

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

RANIK, M.; CREUX, N. M.; MYBURG, A. A. Within-tree transcriptome profiling in wood-forming tissues of a fast-growing *Eucalyptus* tree. **Tree Physiology**, Victoria, v. 26, p. 265-375, 2006.

REILLY, K.; BERNAL, D.; CORTES, D.F.; GOMEZ-VASQUEZ, R.; TOHME, J.; BEECHING, J.R. Towards identifying the full set of genes expressed during cassava post-harvest physiological deterioration. **Plant Molecular Biology**, New York, v. 64, p. 187-203, 2007.

REINHART, B.J.; WEINSTEIN, E.G.; RHOADES, M.W.; BARTEL, B.; BARTEL, D.P. MicroRNAs in Plants. **Genes & Development**, New York, v. 16, p. 1616-1626, 2002.

REITER, W.D. Biosynthesis and properties of the plant cell wall. **Current Opinion in Plant Biology**, London, v. 5, p. 536-542, 2002.

REITER, W.; VANZIN, G. F. Molecular genetics of nucleotide sugar interconversions pathways in plants. **Plant Molecular Biology**, New York, v. 47, p. 95-113, 2001.

RESENDE, M.D.V. de; VENCOSKY, R. Condução e utilização de bancos de conservação genética de espécies de eucalipto. In: CONGRESSO FLORESTAL BRASILEIRO, 6, 1990, São Paulo. **Anais...** São Paulo: Sociedade Brasileira de Silvicultura, 1990, p. 434-439.

RIDOUTT, B.G.; PHARIS, R.P.; SANDS, R. Fibre length and giberelins A1 and A20 are decreased in *Eucalyptus globules* by acylcyclohexanedione injected into the stem. **Physiologia Plantarum**, Lund, v. 96, p. 559-566, 1996.

RISON, S.C.; HODGMAN, T.C.; THORNTON, J.M. Comparison of functional annotation schemes for genomes. **Functional & Integrative Genomics**, Berlin, v.1, n. 1, p. 56-69, 2000.

RITTER, E.; GEBHARDT, C.; SALAMINI, F. Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. **Genetics**, Austin, v. 125, p. 645-654, 1990.

ROBINSON, S.J.; CRAM, D.J.; LEWIS, C.T.; PARKIN, I.A.P. Maximizing the efficacy of SAGE analysis identifies novel transcripts in *Arabidopsis*. **Bioinformatics**, Oxford, v. 136, p. 3223-3233, 2004.

RONNING, C.M.; STEGALKINA, S.S.; ASCENZI, R.A.; BOUGRI, O.; HART, A.L.; UTTERBACH, T.R.; VANAKEN, S.E.; RIEDMULLER, S.B.; WHITE, J.A.; CHO, J.; PERTEA, G.M.; LEE, Y.; KARAMYCHEVA, S.; SULTANA, R.; TSAI, J.; QUACKENBUSH, J.; GRIFFITHS, H.M.; RESTREPO, S.; SMART, C.D.; FRY, W.E.; VAN DER HOEVEN, R.; TANKSLEY, S.; ZHANG, P.; JIN, H.; YAMAMOTO, M.L.; BAKER, B.J.; BUELL, C.R. Comparative analyses of potato expressed sequence tag libraries. **Plant Physiology**, Lancaster, v. 131, p. 419-429, 2003.

ROTH, C.M. Quantifying Gene Expression. **Current Issues in Molecular Biology**, Wymondham, v. 4, p. 93-100, 2002.

ROZEN, S.; SKALETSKY, H. Primer3 on the WWW for general users and for biologist programmers. In: KRAWETZ, S.; MISENER, S. (Eds) **Bioinformatics Methods and Protocols: Methods in Molecular Biology**. Totowa: Humana Press, 2000, p. 365-386.

RUGGIERI, P. Contributo alla carilogia del genere *Eucalyptus* (Myrtaceae). **Caryologia**, Firenze, v. 14, p. 111-120, 1961.

SALVATIERRA, G.R. **Identificação de genes diferencialmente expressos em árvores de *Eucalyptus grandis* susceptíveis e resistentes à *Puccinia psiidi***. 2006 154p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2006.

SALZMAN, R.A.; FUJITA, T.; ZHU-SALZMAN, K.; HASEGAWA, P.M.; BRESSAN, R.A. An improved RNA isolation method for plant tissues containing high levels of phenolic compounds or carbohydrates. **Plant molecular biology reporter / ISPMB**, New York, v. 17, p. 11-17, 1999.

SAMBROOK, J. Preparation of reagents and buffers used in molecular cloning. In: FRITSCH, E.F.; MANIATIS, T. (Eds.) **Molecular cloning: a laboratory manual** (2. ed.). New York: Cold Spring Harbor Laboratory Press, 1989. p. 12, Appendix B.

SAVIDGE, R.A. Auxin and ethylene regulation of diameter growth in trees. **Tree Physiology**, Victoria, v. 4, p. 401-414, 1988.

SCHOCH, G.; GOEPFERT, S.; MORANT, M.; HEHN, A.; MEYER, D.; ULLMANN, P.; WERCK-REICHHART, D. CYP98A3 from *Arabidopsis thaliana* is a 3-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. **The Journal of Biological Chemistry**, Baltimore, v. 276, n. 39, p. 36566–36574, 2001.

SEITZ, B.; KLOS, C.; WURM, M.; TENHAKEN, R. Matrix polysaccharide precursors in *Arabidopsis* cell walls are synthesized by alternate pathways with organ-specific expression patterns. **The Plant Journal**, Oxford, v. 21, p. 537-546, 2000.

SERRANO, L.; ROCHANGE, F.; SEMBLAT, J.P.; MARQUE, C.; TEULIERES, C.; BOUDET, A.M. Genetic transformation of *Eucalyptus globules* through biolistics: complementary development of procedures for organogenesis from zygotic embryos and stable transformation of corresponding proliferating tissue. **Journal of Experimental Botany**, Lancaster, v. 47, n. 295, p. 285-290, 1996.

SEVERO, E.T.D.; CALONEGO, F.W.; SANSÍGOLO, C.A. Composição química da madeira de *Eucalyptus citriodora* em função das direções estruturais. **Silva Lusitana**, Lisboa, v. 14, n. 1, p. 113-126, 2006.

SHOJI, T.; WINZ, R.; IWASE, T.; NAKAJIMA, K.; YAMADA Y, HASHIMOTO T. Expression patterns of two tobacco isoflavone reductase-like genes and their possible roles in secondary metabolism in tobacco. **Plant Molecular Biology**, New York, v. 50, n. 3, p. 427-440, 2002.

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

SIMSON, B.W.; TIMELL, T.E. Polissaccharides in cambial tissues of *Populus tremuloides* and *Tilia Americana*. Isolation, fractionation and chemical composition of cambial tissues. **Cellulose Chemistry and Technology**, Romenia, v. 12, p. 39-50.

SOCIEDADE BRASILEIRA DE SILVICULTURA. As plantações de eucalipto no Brasil. Disponível em: < http://www.sbs.org.br/destaques_plantacoesnobrasil.htm >. Acesso em: 18 mar. 2006.

SPOKEVICIUS, A.V.; SOUTHERTON, S.G.; COLLEEN ; MACMILLAN,C.P.; QIU,D.; GAN, S.; TIBBITS, J.F.G.; MORAN, G.F.; BOSSINGER, G. β -tubulin affects cellulose microfibril orientation in plant secondary fibre cell walls. **The Plant Journal**, Oxford, in press, doi: 10.1111/j.1365-313X.2007.03176, 2007.

SPRENGER, N.; KELLER, F. Allocation of a raffinose family oligosaccharide to transport and storage pools in *Ajuga reptans*: the roles of two distinct galactinol synthases. **The Plant Journal**, Oxford, v. 21, n. 3, p. 249-258, 2000.

STEANE, D.A. Complete Nucleotide Sequence of the Chloroplast Genome from the Tasmanian Blue Gum, *Eucalyptus globulus* (Myrtaceae). **DNA research**, Tokyo, v. 12, n. 3, p. 215-220, 2005.

STERKY, F.; REGAN, S.; KARLSSON, J.; HERTZBERG, M.; ROHDE, A.; HOLMBERG, A.; AMINI, B.; BHALERAO, R.; LARSSON, M.; VILLARROEL, R.; VAN MONTAGU, M.; SANDBERG, G.; OLSSON, O.; TEERI, T.T.; BOERJAN, W.; GUSTAFSSON, P.; UHLÉN, M.; SUNDBERG, B.; LUNDERBERG, J. Gene discovery in the wood-forming tissues of poplar: analysis of 5,692 expressed sequence tags. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 95, p. 1330-1335, 1998.

STOLLBERG, J.; URSCHITZ, J.; URBAN, Z.; BOYD, C.D. A Quantitative Evaluation of SAGE. **Genome Research**, Cold Spring Harbor, v. 10, p. 1241-1248, 2000.

SUN, M.; ZHOU, G.; LEE, S.; CHEN, J.; SHI, R.Z.; WANG, S.M. SAGE is far more sensitive than EST for detecting low-abundance transcripts. **BMC Genomics**, London, v. 5, n. 1, p.1-4, 2004.

TAIZ, L.; ZEIGER, E. Paredes celulares: estrutura, biogênese e expansão. I:____. **Fisiologia Vegetal**, Porto Alegre: Artmed, 2004. cap. 15, p. 339-364.

TAKAHASHI, Y.; BERBERICH, T.; YAMASHITA, K.; UEHARA, Y.; MIYAZAKI, A.; KUSANO, T. Identification of tobacco HIN1 and two closely related genes as spermine-responsive genes and their differential expression during the Tobacco mosaic virus -induced hypersensitive response and during leaf- and flower-senescence. **Plant Molecular Biology**, New York, v. 54, p. 613-622, 2004.

TERASHIMA, N.; FUKUSHIMA, K.; HE, L.F.; TAKABE, K. Comprehensive model of the lignified plant cell wall. In: JUNG, H.G. (Ed.). **Forage Cell Wall Structure and Digestibility**, Madison: ASA-CSSA-SSSA, 1993, p. 247-270.

TOMASELLI, I. El Sector Forestal Brasileño en el 2020. In: _____. **Informe Nacional Brasil: Estudio de tendencias y perspectivas del sector forestal en América Latina**. Rome: FAO – MMA Brasil, 2004. cap. 5. Disponível em: <http://www.fao.org/documents/show_cdr.asp?url_file=/docrep/007/j3032s/j3032s00.htm>. Acesso em: 06 jun. 2007.

TOMAZELLO FILHO, M. Variação radial da densidade básica e da estrutura anatômica da madeira do *Eucalyptus globulus*, *E. pellita* e *E. acmenoides*. **IPEF**, Piracicaba, n. 36, p. 35-42, 1987.

TOMAZELLO FILHO, M. Variação radial da densidade básica e da estrutura anatômica da madeira do *Eucalyptus saligna* e *E. grandis*. **IPEF**, Piracicaba, n. 29, p. 37-45, 1985.

TRUGILHO, P.F.; LIMA, J.T.; MENDES, L.M.. Influência da idade nas características físico-químicas e anatômicas da madeira de *Eucalyptus saligna*. **CERNE**, Lavras, v. 2, n. 1, p. 1-152. 1996.

TUSKAN, A.; DIFAZIO, S.; JANSSON, S.; BOHLMANN, J.; GRIGORIEV, I., HELLSTEN, U.; PUTNAM, N.; RALPH, S.; ROMBAUTS, S.; SALAMOV, A.; SCHEIN, J.; STERCK, L.; AERTS, A.; BHALERAO, R.R.; BHALERAO, R.P.; BLAUDEZ, D.; BOERJAN, W.; BRUN, A.; BRUNNER, A.; BUSOV, V.; CAMPBELL, M.; CARLSON, J.; CHALOT, M.; CHAPMAN, J.; CHEN, G.L.; COOPER, D.; COUTINHO, P.M.; COUTURIER, J.; COVERT, S.; CRONK, Q.; CUNNINGHAM, R.; DAVIS, J.; DEGROEVE, S.; DE JARDIN, A.; DEPAMPHILIS, C.; DETTER, J.; DIRKS, B.; DUBCHAK, I.; DUPLESSIS, S.; EHLTING, J.; ELLIS, B.; GENDLER, K.; GOODSTEIN, D.; GRIBSKOV, M.; GRIMWOOD, J.; GROOVER, A.; GUNTER, L.; HAMBERGER, B.; HEINZE, B.; HELARIUTTA, Y.; HENRISSAT, B.; HOLLIGAN, D.; HOLT, R.; HUANG, W.; ISLAM-FARIDI, N.; JONES, S.; JONES-RHOADES, M.; JORGENSEN, R.; JOSHI, C.; KANGASJA, J.; KARLSSON, J.; KELLEHER, C.; KIRKPATRICK, R.; KIRST, M.; KOHLER, A.; KALLURI, U.; LARIMER, F.; LEEBENS-MACK, J.; LEPLE, J. C.; LOCASCIO, P.; LOU, Y.; LUCAS, S.; MARTIN, F.; MONTANINI, B.; NAPOLI, C.; NELSON, D.R.; NELSON, C.; NIEMINEN, K.; NILSSON, O.; PEREDA, V.; PETER, G.; PHILIPPE, R.; PILATE, G.; POLIAKOV, A.; RAZUMOVSKAYA, J.; RICHARDSON, P.; RINALDI, C.; RITLAND, K.; ROUZE, P.; RYABOY, D.; SCHMUTZ, J.; SCHRADER, J.; SEGERMAN, B.; SHIN, H.; SIDDIQUI, A.; STERKY, F.; TERRY, A.; TSAI, C. J.; UBERBACHER, E.; UNNEBERG, P.; VAHALA, J.; WALL, K.; WESSLER, S.; YANG, G.; YIN, T.; DOUGLAS, C.; MARRA, M.; SANDBERG, G.; VAN DE PEER, Y.; ROKHSA, D. The Genome of Black Cottonwood, *Populus trichocarpa* (Torr. & Gray). **Science**, Washington, v. 313, p. 1596-1604, 2006.

UGGLA, C.; MORITZ, T.; SANDBERG, G.; SUNDBERG, B. Auxin as a positional signal in pattern formation in plant. **Proceedings of the National Academy of Sciences of the United States of América**, Washington, v.93, p. 9282-9286, 1996.

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.

VANDER-MIJNSBRUGGE, K.; MEYERMANS, H.; VAN MONTAGU, M.; BAUW, G.; BOERJAN, W. Wood formation in poplar: identification, characterization, and seasonal variation of xylem proteins. **Planta**, Berlin, v. 210, n. 4, p. 589-598, 2000.

VELCULESCU, V.E.; ZHANG, L.; VOGELSTEIN, B.; KINZLER, K.W. Serial analysis of gene expression. **Science**, Washington, v. 270, p. 484-487, 1995.

VERNOUD, V.; HORTON, A.C.; YANG, Z.; NIELSEN, E. Analysis of the small GTPase gene superfamily of Arabidopsis. **Plant Physiology**, Lancaster, v. 131, p. 1191-1208, 2003.

VINCENT, D.; LAPIERRE, C.; POLLET, B.; CORNIC, G.; NEGRONI, L.; ZIVY, M. Water deficits affect caffeate O-methyltransferase, lignification, and related enzymes in maize leaves. A proteomic investigation. **Plant Physiology**, Lancaster, v. 137, n. 3, p. 949-960, 2005.

WANG, S.M. Understanding SAGE data. **Trends in Genetics**, Amsterdam, v. 23, p. 42-50, 2007.

WEI, C.L.; PATRICK, N.G.; CHIU, K.P.; WONG, C.H.; ANG, C.C.; LIPOVICH, L.; LIU, E.T.; RUAN, Y. 5' Long serial analysis of gene expression (LongSAGE) and 3' LongSAGE for transcriptome characterization and genome annotation. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 101, n. 32, p. 11701-6, 2004.

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

WOO, H.H.; HAWES, M.C. Cloning of genes whose expression is correlated with mitosis and localized in dividing cells in root caps of *Pisum sativum* L. **Plant Molecular Biology**, New York, v. 35, p. 1045-1051, 1997.

YANG, S-H., WANG, H.; SATHYAN, P.; STASOLLA, C.; LOOPSTRA, C.A. Real-time RT-PCR of loblolly pine (*Pinus taeda*) arabinogalactan-protein and arabinogalactan-protein-like genes. **Physiologia Plantarum**, Lund, v. 124, p. 91-106, 2005.

ZOBEL, B.J. Inherent differences affecting wood quality in fast-grown plantations. In: IUFRO CONFERENCE, 1980, São Paulo, **Anais...** São Paulo: Oxford/IUFRO, p 169-188.

ZOBEL, B.J.; VAN BUIJTENEN, J.P. **Wood variation: its causes and control**. New York, Springer-Verlag, 1989, 363p.

ZHONG, R.; MORRISON III, W.H.; HIMMELSBACH, D.S.; POOLE II, F.L.; YE, Z. Essential role of caffeoyl coenzyme A O-methyltransferase in lignin biosynthesis in woody poplar plants. **Plant Physiology**, Lancaster, v. 124, p. 563-577, 2000.

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

ANEXOS

ANEXO A - Tratamento do material para extração de RNA

1) Tratamento dos tubos:

- a) Lavagem dos tubos com extran;
- b) Lavagem com SDS 2% em forno a 70°C por 2 horas;
- c) Lavagem com álcool comercial em agitador por 20 minutos;
- d) Lavagem com água oxigenada 3% em agitador por 20 minutos;
- e) Lavagem com água DEPC 0,01% em agitador por 20 minutos;
- f) Autoclavar os tubos por 15 minutos.

2) Preparo de reagentes e vidrarias:

Todos reagentes devem ser preparados com água DEPC 0,1% e autoclavados

As vidrarias devem ser tratadas em forno à 180°C por 8 horas

ANEXO B - Protocolo de extração de RNA total (SALZMAN et al., 1999)

- a) Macerar 2g de tecido e adicionar 10mL do tampão de extração;
- b) Agitar vigorosamente por 1 minuto;
- c) Adicionar 1 volume de clorofórmio:álcool isoamílico (24:1) e agitar por 20';
- d) Centrifugar à 16000 x g por 10 minutos à 4°C;
- e) Transferir a fase superior para outro tubo;
- f) Repetir os passos (c) e (d) até a interfase estar limpa;
- g) Remover o sobrenadante para um novo tubo; e;
- h) Adicionar 2X vol. de etanol 100% e 0,1X vol. de NaCl 5 M;
- i) Precipitar *overnight* à -2°C;
- j) Centrifugar à 16000 x g por 10 minutos à 4°C;
- k) Ressuspender o *pellet* em 10mL de água DEPC 0,1%;
- l) Se necessário, centrifugar por 5' à 20000 x g e recuperar o sobrenadante;
- m) Adicionar 1X fenol saturado em TE (pH 8,0): clorofórmio : álcool isoamílico (25:24:1) e agitar por 10 minutos à temperatura ambiente;
- n) Centrifugar à 13000 x g por 10 minutos à temperatura ambiente;
- o) Transferir a fase superior a um novo tubo;
- p) Repetir os passos (i) e (j) até a interfase estar limpa;

- q) Adicionar 2X o vol. de álcool 100% e 0,1X o vol. de NaCl 5 M;
- r) Precipitar *overnight*;
- s) Centrifugar à 16000 x g por 15 minutos à 4°C;
- t) Ressuspender o *pellet* em 500 µL de água DEPC 0,1%.
- u) Completar o volume para 1 mL com água DEPC e adicionar 333 µM de LiCl 8 M;
- v) Precipitar por 3 horas;
- x) Centrifugar 20 minutos à 12000 x g à 4°C;
- w) Lavar o *pellet* com álcool 80%;
- y) Ressuspender o *pellet* em um volume apropriado de água DEPC 0,1% (50 µL).

ANEXO C - Tampão de extração (SALZMAN et al., 1999)

Tiocinato de guanidina	4 M
Tris-HCl pH 8,0	100 mM
Citrato de sódio pH 8,0	25 mM
Sarcosina N-lauril	0,5%
PVP solúvel	1%
2-mercaptoetanol	2%
Água DEPC 0,1 %	Completar o volume

ANEXO D - Lysis/Binding buffer (I-SAGETM)

Tris-HCl pH 7,5	100 mM
LiCl	500 mM
EDTA	10 mM
Duodecil sulfato de lítio	1%
DTT (DI- Tiotreitol)	5 mM

ANEXO E - Washing buffer A (I-SAGETM)

Tris-HCl pH 7,5	10 mM
LiCl	0,15 M
EDTA	1 mM
Duodecil sulfato de lítio	0,1%
Glicogênio	10 µg/ mL

ANEXO F - Washing buffer B (I-SAGETM)

Tris-HCl pH 7,5	10 mM
LiCl	150 mM
EDTA	1 mM
Glicogênio	10 µg/ mL

ANEXO G - 1X First Strand buffer (I-SAGETM)

Tris-HCl pH 8,3	50 mM
KCl	75 mM
MgCl ₂	3 mM
Duodecil sulfato de lítio	0,1%
<i>Mussel glycogen</i>	10 µg/ mL

ANEXO H - Reação de síntese da primeira fita de cDNA (I-SAGETM)

5X First Strand buffer	18 µL
RnaseOUT TM	1 µL
Água DEPC	54,5 µL
0,1 M DTT (DI - Tiotreitol)	9 µL
dNTP Mix (10 mM de cada)	4,5 µL

ANEXO I - RT SuperScriptTM II (Transcriptase reversa) (I-SAGETM) 200 U/ μ L em:

Tris-HCl pH 7,5	20 mM
NaCl	100 mM
EDTA	0,1 mM
DTT	1 mM
Nonidet – 40 (v/v)	0,01%
glicerol (p/v)	50%

ANEXO J -5X Second Strand buffer (I-SAGETM)

Tris-HCl pH 6,9	100 mM
KCl	450 mM
MgCl ₂	23 mM
β – NAD ⁺	0,075 mM
Sulfato de amônia	50 mM

ANEXO K - dNTP Mix (I-SAGETM)

dATP	10 mM
dGTP	10 mM
dCTP	10 mM
dTTP	10 mM

Em água DEPC

ANEXO L - E. coli DNA ligase (I-SAGETM) 10 U/ μ L em:

Tris-HCl pH 7,4	10 mM
KCl	50 mM
EDTA	0,1 mM
DTT	1 mM
BSA	200 μ g/mL
Glicerol (p/v)	50%
Triton [®] X-100	0,1%

ANEXO M - *E. coli* DNA polimerase (I-SAGETM) 10 U/μL em:

Fostato de potássio pH 7,4	50 mM
KCl	100 mM
DTT	1 mM
Glicerol (p/v)	50%

ANEXO N - *E. coli* RNase H (I-SAGETM) 2 U/μL em:

Tris-HCl pH 7,5	20 mM
KCl	100 mM
EDTA	0,1 mM
DTT	0,1 mM
BSA	50 μg/mL
Glicerol (p/v)	50%
MgCl ₂	10 mM

ANEXO O - Washing buffer C (I-SAGETM)

Tris-HCl pH 7,5	50 mM
NaCl	1 M
EDTA	0,5 mM
SDS (Duodecil Sulfato de Sódio)	1 %
<i>Mussel glycogen</i>	10 μg/mL

ANEXO P - Washing buffer D (I-SAGETM)

Tris-HCl pH 7,5	5 mM
NaCl	1 M
EDTA	0,5 mM
BSA	200 μg/mL

ANEXO Q - 1X Buffer 4 (I-SAGE™)

Tris-acetato pH 7,9	20 mM
Acetato de magnésio	10 mM
Acetato de potássio	50 mM
DTT	1 mM

ANEXO R - LoTE (I-SAGE™)

Tris-HCl pH 7,5	3 mM
EDTA pH 7,5	0,2 mM

ANEXO S -100X BSA

BSA	10 mg/mL
Fosfato de potássio pH 7	20 mM
NaCl	50 mM
EDTA	0,1 mM
Glicerol (p/v)	5%

ANEXO T - Nla III (I-SAGE™) 10 U/μL em:

KCl	200 mM
Tris-HCl pH 7,5	10 mM
EDTA	0,1 mM
DTT	1 mM
100X BSA	500 μg/mL
Glicerol (p/v)	50%

ANEXO U -1X Ligase buffer (I-SAGE™)

Tris-HCl pH 7,5	6 mM	ATP	0,1 mM
NaCl	5 mM	DTT	2 mM
MgCl ₂	6 mM	Espermidina	1 mM
BSA	0,1 mg/mL		
β - mercaptoetanol	7 mM		

ANEXO V - *T4 DNA ligase (I-SAGE™)* 4 U/μL em:

Tris-HCl pH 7,5	10 mM
Cloreto de potássio	50 mM
DTT	1 mM
Glicerol (p/v)	50%

ANEXO W - *BsmF I (I-SAGE™)* 2 U/μL em:

NaCl	50 mM
Tris-HCl pH 7,4	10 mM
EDTA	0,1 mM
DTT	1 mM
BSA	200 μg/mL
Glicerol (p/v)	50%

ANEXO X - PCR teste – Verificação da digestão com *BsmF I*

	1	2	3	4
10X BV <i>buffer</i>	5 μL	5 μL	5 μL	5 μL
DMSO	3 μL	3 μL	3 μL	3 μL
dNTP <i>mix</i>	5 μL	5 μL	5 μL	5 μL
DTP-1 ou DTP-2	2μL	2μL	2μL	2μL
UDGH anti-sense <i>primer</i>	2μL	2μL	2μL	2μL
<i>Platinum®Taq</i> DNA polimerase	0,5 μL	0,5 μL	0,5 μL	0,5 μL
Água DEPC	32 μL	32 μL	32 μL	32 μL
Amostra do tubo A	-	-	0,5 μL	-
Amostra do tubo B	-	-	-	0,5 μL
Controle positivo A (antes da digestão)	0,5 μL	-	-	-
Controle positivo B (antes da digestão)	-	0,5 μL	-	-

ANEXO Y -10X Klenow buffer (I-SAGE™)

Tris-HCl pH 8,0	500 mM
NaCl	500 mM
MgCl ₂	100 mM

ANEXO Z - Klenow polimerase (I-SAGE™) 3-9 U/μL em:

Fosfato de potássio pH 7,0	50 mM
KCl	100 mM
DTT	1 mM
Glicerol (p/v)	50%

ANEXO AA - Gel de poliacrilamida 12%

40% Acrilamida/bisacrilamida	12,8 mL
Tampão de corrida TBE 5X	8 mL
TEMED	16 μL
Água deionizada	18,8 mL
Persulfato de amônio 10%	0,4 mL

ANEXO AB - Tampão BV 10X concentrado

Sulfato de amônio	166 mM
Tris-HCl pH 8,8	670 mM
MgCl ₂	67 mM
β - mercaptoetanol	100 mM

ANEXO AC - Tampão de amostra 5X TBE

Tris base	18 mM
Ácido Bórico	18 mM
EDTA (livre de ácidos)	0,4 mM
Ficoll™ tipo 400	3%
Azul de bromofenol	0,02%
Xileno cianol	0,02%

ANEXO AD - pZErO[®] - I (I-SAGETM)

pZErO[®] - 1 1 µg/µL em tampão TE pH 7,5

ANEXO AE -10X Tampão 2

Tris-HCl pH 7,9 100 mM

MgCl₂ 10 mM

NaCl 50 mM

DTT 1 mM

ANEXO AF - *Sph* I (I-SAGETM) 5 U/μL em:

NaCl 100 mM

Tris-HCl pH 7,4 10 mM

EDTA 0,1 mM

DTT 1 mM

BSA 100 $\mu\text{g/mL}$

Glicerol (p/v)	50%
----------------	-----

ANEXO AG - Meio SOC

Triptona 2%

Extrato de levedura	0,5%
---------------------	------

MgCl₂ 10 mM

NaCl 10 mM

KCl 2,5 mM

MgSO₄ 10 mM

Glicose 20 mM

ANEXO AH - Meio LB sólido pH 7,5

Triptona 1%

Extrato de levedura	0,5%
---------------------	------

NaCl	0,5%
------	------

Agar	1,5%
------	------

ANEXO AI - ZeocinTM (I-SAGETM) (antibiótico)

ZeocinTM 100 mg/mL em água