**Universidade de São Paulo
Escola Superior de Agricultura “Luis de Queiroz”**

**Identificação de proteínas da região cambial de *Eucalyptus grandis*
por eletroforese bidimensional e espectrometria de massas**

Paola Alejandra Fiorani Celedón

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

Piracicaba

2006

Paola Alejandra Fiorani Celedón
Bióloga

**Identificação de proteínas da região cambial de *Eucalyptus grandis* por
eletroforese bidimensional e espectrometria de massas**

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
2006

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

Celedón, Paola Alejandra Fiorani

Identificação de proteínas da região cambial de *Eucalyptus grandis* por eletroforese bidimensional e espectrometria de massas / Paola Alejandra Fiorani
Celedón. - - Piracicaba, 2006.

112 p. : il.

Tese (Doutorado) - - Escola Superior de Agricultura Luiz de Queiroz, 2006.
Bibliografia.

1. Cambio (botânica) 2. Espectrometria de massas 3. Eucalipto 4. Proteínas de plantas I. Título

CDD 634.9734

“Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor”

AGRADECIMENTOS

Ao departamento de Genética da ESALQ-USP, pela oportunidade de realização desta tese;

Ao CNPq pela concessão de bolsa e auxílios financeiros;

À Companhia Suzano de Papel e Celulose e seus funcionários, em especial ao Dr. Shinitiro Oda, pela concessão do material vegetal utilizado neste projeto, apoio durante as coletas e investimentos e confiança dirigidos ao Laboratório Max Faffer de Genética de Plantas, essenciais ao desenvolvimento desta tese;

Ao Dr. Carlos Alberto Labate pela confiança depositada, apoio e orientação durante o decorrer do curso;

Ao Dr. David Moon e à Dra. Mônica Labate pela amizade, paciência e sugestões para esta tese;

Aos funcionários do Departamento de Genética pelas instruções necessárias ao andamento do curso;

Ao Laboratório do Prof. Dr. Márcio Lambais, em especial ao aluno de doutorado e amigo Simão, pelas sugestões e empréstimo de materiais;

Por último, mas não menos importante, a todos os colegas do laboratório pela atenção e apoio que me dedicaram, e foram fundamentais para a finalização desta tese.

SUMÁRIO

RESUMO.....	6
ABSTRACT	7
LISTA DE FIGURAS	8
LISTA DE TABELAS	10
LISTA DE ABREVIATURAS.....	11
1 INTRODUÇÃO	12
2 REVISAO BIBLIOGRÁFICA.....	16
2.1 Recursos florestais no Brasil	16
2.2 O gênero Eucalyptus no Brasil	17
2.3 Importância econômica do eucalipto	17
2.4 Melhoramento genético do eucalipto.....	18
2.5 Biotecnologia aplicada ao eucalipto	19
2.6 Xilogênese.....	20
2.7 Modelos para o estudo da xilogênese	23
2.8 Proteoma de plantas	24
2.9 Estratégias proteômicas.....	26
2.10 Espectrometria de massas para análise de proteínas.....	28
2.11 Equipamentos	28
2.12 Bioinformática.....	30
2.13 Bancos de ESTs na identificação de proteínas	32
3 MATERIAL E MÉTODOS	33
3.1 Coleta do material	33
3.2 Extração de proteínas	33
3.3 Separação eletroforética de proteínas	35
3.4 Análise de imagem dos géis.....	36
3.5 Digestão in-gel das proteínas.....	37
3.6 Sequenciamento de peptídeos por LC-MS/MS	38
3.7 Bancos de ESTs de Eucalyptus spp. para análise das proteínas	39
3.8 Processamento dos espectros MS/MS	39
3.9 Estratégias de análises dos espectros MS/MS	41

4 RESULTADOS E DISCUSSÃO.....	43
4.1 Técnica de coleta de material vegetal	43
4.2 Eficiência dos protocolos de extração de proteínas	43
4.3 Separação das proteínas em géis bidimensionais	45
4.4 Seleção de spots para sequenciamento	47
4.5 Apresentação dos resultados de espectrometria de massas	47
4.6 Taxas de identificação de proteínas de <i>E. grandis</i>	49
4.7 Perfil do proteoma da madeira	61
4.7.1 Avaliação qualitativa.....	61
4.7.2 Nível de expressão de proteínas.....	62
4.8 Proteínas representadas por vários spots	63
4.9 Sucesso de identificação de proteínas	64
4.10 Relação entre proteínas e nível de transcritos	66
4.11 Proteínas da região cambial de <i>E. grandis</i>	71
4.11.1 Proteínas de resposta a stress.....	71
4.11.2 Síntese e turnover de proteínas	74
4.11.3 Proteínas de estresse oxidativo	75
4.11.4 Metabolismo de nitrogênio e aminoácidos	78
4.11.5 Proteínas do metabolismo de carboidratos	79
4.11.6 Proteínas de parede	81
4.11.7 Metabolismo de energia	83
4.11.8 Sinalização celular	85
4.11.9 Citoesqueleto	85
4.12 Local da síntese de proteínas da região cambial de <i>E. grandis</i>	88
5 CONCLUSÕES	89
REFERÊNCIAS:.....	90

RESUMO

Identificação de proteínas da região cambial de *Eucalyptus grandis* por eletroforese bidimensional e espectrometria de massas

A proteômica expandiu-se dentro da área biológica a partir da década de 90 com o desenvolvimento de técnicas de ionização branda que permitem analisar macromoléculas por espectrometria de massas (MS) e é uma nova estratégia na busca de genes de interesse e de informações sobre o controle da expressão gênica para a manipulação genética de plantas. Na busca por genes diferencialmente expressos durante o desenvolvimento da madeira, foi estabelecida uma plataforma 2D-LC-MS/MS (eletroforese bidimensional seguida de cromatografia líquida e espectrometria de massas em tandem) e identificadas 72 proteínas da região cambial de *Eucalyptus grandis* de árvores de 2 anos e 8 meses. A identificação foi feita por homologia com o banco de proteínas SwissProt, com um banco de ESTs específico de *Eucalyptus spp.* (GENOLYPTUS) e com o NCBI. Os dados de proteínas também foram confrontados com a expressão dos transcritos correspondentes obtidos a partir do mesmo tecido por SAGE (*Serial Analysis of Gene Expression*). Dentre as proteínas identificadas, a maior parcela (24,1 %) tem função relacionada com resposta a estresse, 20,5 % pertence a metabolismo de energia, 17,8 % são componentes estruturais, 14,3 % participam em metabolismo de macromoléculas e 15,2 % em metabolismos básicos como de nitrogênio, aminoácidos, nucleotídeos, lipídeos e metabolismo secundário.

Palavras-chave: *Eucalyptus*, câmbio (botânica), proteínas de plantas, espectrometria de massas

ABSTRACT

Identification of proteins from the cambial region of *Eucalyptus grandis* by bidimensional electrophoresis and mass spectrometry

Proteomics impacted the biological sciences since the 90's after the development of soft ionization techniques that allow to analyse macromolecules by mass spectrometry (MS). It became a new strategy to identify genes and obtain information about the molecular control of gene expression, of importance to genetic manipulation of plants. In order to search for genes involved in wood quality and to understand gene expression during wood development, a 2D-LC-MS/MS system (bidimensional electrophoresis followed by liquid chromatography coupled to mass spectrometry in tandem) was established. 72 proteins from the cambial region of *Eucalyptus grandis* trees at the age of 2 years and 8 months were identified. The MS/MS spectra were processed using the SwissProt databank, an EST (Expressed Sequence Tags) bank of *Eucalyptus* spp. (GENOLYPTUS) and the NCBI. A comparison of gene and protein expression was carried out using a databank constructed in our laboratory using SAGE (Serial Analysis of Gene Expression) obtained from the same cambial tissue. Most identified proteins are related to stress response (24.1 %), 20.5 % participate in energy metabolism, 17.8 % are important to the cell as structural components, 14,3 % are related to macromolecular metabolism and 15,2 % to basic metabolism of nitrogen, amino acids, nucleotides, lipids and secondary metabolism.

Keywords: *Eucalyptus*, cambium (botany), plant proteins, mass spectrometry

LISTA DE FIGURAS

Figura 1 – Modelo de construção do Q-TOF com operação em modo MS/MS. Fonte: Waters.....	30
Figura 2 – Metodologia de coleta de tecidos do xilema em desenvolvimento.....	33
Figura 3 – Parâmetros de processamento dos espectros MS/MS. Programa Global Server v. 2.1.....	40
Figura 4 – Parâmetros para identificação de proteínas com o banco SwissProt v 1.0. pelo programa ProteinLynx v.2.1 Os mesmos parâmetros foram utilizados para buscas no banco GENOLYPTUS, trocando-se apenas o item “Databank – SwissProt 1.0” do primeiro quadro por “Databank-GENOLYPTUS”.....	41
Figura 5 – Fluxograma das estratégias de buscas em bancos de dados para Identificação das proteínas.....	42
Figura 6 – a) Distribuição de peso molecular do extrato de proteínas da região cambial de <i>E. grandis</i> obtida por extração ácida (TCA) e fenólica (PHE); M = Marcador molecular; b) Eficiência de remoção de ácidos nucléicos do extrato de proteínas obtidas por extração ácida (TCA) e fenólica (PHE). M = marcador molecular. As setas apontam o acúmulo de ácidos nucléicos.....	44
Figura 7 – Padrão de distribuição das proteínas de pl 3-10 da região cambial de <i>E. grandis</i> aos 2 anos e 8 meses. Dentro do quadrado verde podem ser vistos alguns dos <i>spots</i> mais corados.....	45
Figura 8 – Padrão de distribuição de proteínas de pl 4-7 da região cambial de <i>E. grandis</i> aos 2 anos e 8 meses de idade. Os números apontam os <i>spots</i> removidos para sequenciamento. Dentro do quadrado verde estão algumas regiões com baixa resolução de <i>spots</i>	46
Figura 9 – Cromatograma de eluição dos peptídeos em gradiente de 45 minutos. A figura superior mostra o tempo de eluição de cada peptídeo.....	48
Figura 10 – Apresentação dos espectros MS/MS dos peptídeos seqüenciados. A figura superior apresenta o tempo de eluição de cada peptídeo e a inferior o espectro de fragmentação íon de m/z = 735,86.....	48
Figura 11 – Apresentação dos resultados após processamento pelo programa ProteinLynx. À esquerda no canto superior: os peptídeos sugeridos após análise automática dos espectros. Quadro superior direito: proteínas com as quais foi encontrada homologia no banco de dados,	

e no centro os íons fragmentados de cada peptídeo. Embaixo o espectro com a distribuição das massas do íon de $m/z = 735,868$49

Figura 12 - Distribuição em categorias funcionais de 112 proteínas identificadas da região cambial de *E. grandis* a partir de árvores de 2 anos e 8 meses.....62

Figura 13 - Proteínas mais intensamente expressas na região cambial de *E. grandis* entre os 112 *spots* identificados. Foi considerada a soma de todas as prováveis isoformas.....63

Figura 14 – Diferentes padrões de distribuição de *spots* representando uma mesma proteína no gel de pH 4-7; a) isoflavona redutase; b) tubulinas α e β ; c) 14-3-3.....64

Figura 15 – Esquema do complexo V da ATP synthase de *E. coli*. As setas vermelhas indicam as enzimas identificadas no xilema em formação de *E. grandis*.....84

Figura 16 – Isoformas de tubulinas encontradas na região cambial de *E. grandis*. As setas vermelhas apontam os *spots* que não puderam ser identificados devido à baixa resolução da região.....87

LISTA DE TABELAS

Tabela 1 – Gradiente de eluição dos peptídeos para sequenciamento por MS/MS.....	38
Tabela 2 – Modificações pós-traducionais consideradas pela ferramenta AutoMod do programa ProteinLynx v 2.1.....	39
Tabela 3 – Proteínas identificadas por MS/MS na região cambial de <i>E. grandis</i>	52
Tabela 4 – Proteínas da região cambial de <i>E. grandis</i> representadas por maior número de <i>spots</i> no gel de pH 4-7.....	64
Tabela 5 – Peptídeos eluídos de proteínas representadas por vários spots. a) 14_33 e b) S-adenosilmetionina sintetase. Cores semelhantes representam peptídeos com alta ou total similaridade. As seqüências sob as tabelas pertencem à isoforma com maior homologia, e foram extraídas do banco do SwissProt. Em vermelho são indicados outros pontos possíveis para clivagem pela tripsina.....	67
Tabela 6 – Influência do tempo de análise por MS no número de peptídeos Identificados por MS/MS.....	68
Tabela 7 – Comparação entre quantificação de RNA e proteínas.....	69

LISTA DE ABREVIATURAS

MS - espectrometria de massas

2D-LC-MS/MS – eletroforese bidimensional seguida de cromatografia líquida e espectrometria de massas em tandem

SAGE - Serial Analysis of Gene Expression – análise seriada da expressão gênica

EST - Expressed Sequence Tags

pI - ponto isoelétrico

PM - peso molecular

MALDI – matrix assisted laser desorption ionization

ESI – electrospray

TOF - time of flight

GC-MS - cromatografia gasosa acoplada à espectrometria de massas

Q-TOF – quadrupole - time of flight

TCA - ácido tricloroacético

DTT – ditioneitol

SDS-PAGE - sodium dodecyl sulfate-poliacrylamide gel electrophoresis

IAA - iodoacetamida

AF - ácido fórmico

ACN – acetonitrila

HSP – heat shock protein

SOD - superóxido dismutase

SAM-S - S-adenosylmethionine synthetase

COMT – caffeic acid 3 O methyltransferase

4CL – 4-coumarate:CoA ligase

(p/v) – peso por volume

(v/v) – volume por volume

1 INTRODUÇÃO

Na área de genética de plantas, esforços significativos têm sido feitos na busca de conhecimento sobre os genes responsáveis por características de interesse agrícola como produtividade, teor nutritivo, resistência a doenças e tolerância a estresses bióticos e abióticos. A identificação dos genes envolvidos nestas características permite que, através de técnicas já bem estabelecidas como clonagem e transformação genética, sejam desenvolvidas novas cultivares com fenótipos mais vantajosos.

A genética de plantas de grande porte vem mais recentemente investindo no melhoramento de espécies florestais. Embora em comparação com as cultivares agrícolas as pesquisas no setor florestal encontrem-se ainda em estágio inicial (Merkle e Dean, 2000), importantes projetos vêm sendo realizados visando o melhoramento das características da madeira, maior crescimento das árvores e resistência a estresses. Um dos maiores interesses em árvores é compreender a xilogênese e a biotecnologia contribui para caracterizar os genes relacionados com o processo. Estudos sobre o tema têm abordado especialmente a síntese de lignina e celulose, metabolismo de açúcares de parede celular, morte celular programada e fatores que determinam o destino celular a partir das células iniciais cambiais. Embora uma ampla gama de publicações esteja disponível, a genética de árvores é uma área ainda pouco explorada se considerada sua complexidade e importância econômica.

O Brasil consumiu em 2004 cerca de 305 milhões de m³ de madeira (SBS, 2005), sendo apenas 48% proveniente de florestas plantadas. O manejo florestal sustentável vem sendo disseminado como alternativa à crescente devastação de áreas nativas como a da Amazônia, que entre 2003 e 2004 atingiu 24.000 Km² (INPE, 2005). A expansão de florestas plantadas pode reduzir estes números uma vez que a maior parte dos desmatamentos visa a comercialização da madeira e a obtenção de novas áreas para a extensão da agricultura e da pecuária. Portanto, uma forma de preservação dos recursos naturais, sem prejuízo econômico de populações dependentes de atividades rurais, é o cultivo de espécies florestais.

O setor florestal brasileiro movimentou em 2004 em torno de USD 25 bilhões ou 4,1% do PIB nacional. A indústria de papel e celulose representou nesse mesmo período

1,9% deste PIB (ABIMCI, 2004). A indústria de papel e celulose está em expansão e a previsão é que até 2010 será o setor de maior investimento de capital nacional e estrangeiro no Brasil, ficando atrás apenas da indústria petroquímica. Entretanto, a previsão é que a demanda continue em ascensão, tornando-se necessário o aumento da produtividade e portanto também dos investimentos no melhoramento genético da espécie. Embora os países da Europa e América do Norte tenham maior interesse em outras espécies de lenhosas (*Populus* e *Pinus*), há diversos outros (Austrália, França, China, Japão, etc) investindo recursos também em *Eucalyptus*. Para que o Brasil não perca a liderança de produção e exportações com esta espécie em um mercado tão competitivo, é fundamental que sejam incentivadas as pesquisas para o melhoramento da qualidade da madeira.

Além do melhoramento clássico, esforços vêm sendo desenvolvidos a nível internacional com ferramentas da biotecnologia com diversas espécies de *Eucalyptus* na busca de informações sobre a genética da formação da madeira. No Brasil foram iniciados nos últimos cinco anos dois grandes projetos em genômica de *Eucalyptus*, o FORESTs (<https://forests.esalq.usp.br>) e o GENOLYPTUS (www.genolyptus.ucb.br). Ambos são resultados de consórcios entre universidades e empresas privadas do setor de papel e celulose sendo que, no projeto GENOLYPTUS, também houve a participação da Embrapa. Os dois projetos formaram grandes e importantes bancos de dados com seqüências do genoma expresso de diferentes espécies e tecidos de *Eucalyptus spp.*

Esta tese faz parte de um terceiro projeto onde duas bases de dados estão sendo montadas: uma de ESTs, obtidas por SAGE, e outra de proteínas, identificadas por uma plataforma 2D-LC-MS/MS, sendo ambas as bases construídas exclusivamente a partir de tecidos da região cambial da madeira de *Eucalyptus* em diversas etapas do desenvolvimento (6 meses, 2 anos e 8 meses, e 6 anos). Neste trabalho estão sendo analisados os tecidos da idade intermediária de 2 anos e 8 meses através de técnicas da proteômica. O uso da proteômica está gerando um conjunto de informações sobre o proteoma diferencialmente expresso durante o desenvolvimento da madeira. Aliado ao banco de ESTs obtidos por SAGE, o proteoma auxiliará a montar um perfil da expressão gênica durante o desenvolvimento do xilema de *Eucalyptus*.

Apesar de atualmente existirem tecnologias que possibilitam o sequenciamento de genomas inteiros, outras estratégias estão se tornando necessárias para compreender a expressão gênica em um contexto espaço-temporal na célula. A perspectiva de desvendar o código genético dos seres vivos parecia a chave para a manipulação das características de interesse em plantas, mas a difusão de novas técnicas tem trazido à tona novos aspectos da expressão gênica. Isto está sendo evidenciado por análises de transcritos e de proteomas inteiros de tecidos que demonstram que, as etapas de *splicing* alternativo (remoção de exons) e as modificações pós-traducionais comuns nas células, dificultam a identificação exata da função gênica baseando-se apenas no sequenciamento de genes. O novo desafio dos geneticistas é compreender as complexas redes reguladoras da expressão do genoma. Possivelmente somente através da integração das várias ferramentas (genômica + transcriciômica + proteômica) será possível dimensionar a complexidade dos sistemas biológicos para encontrar alternativas que permitam maior eficiência na manipulação das características genéticas de plantas.

A proteômica passou a ocupar uma posição de destaque nas ciências biológicas após a difusão da espectrometria de massas (MS) como técnica para análise de biomoléculas, o que permitiu o desenvolvimento de novas estratégias de pesquisa para estudar a expressão gênica através de proteínas. A identificação de proteínas em larga escala, de maneira semelhante ao que é feito com o DNA, tornou possível o estudo de proteomas inteiros de tecidos e compartimentos sub-celulares.

O termo proteoma pode ser definido como o conjunto de proteínas expresso por uma célula ou tecido em uma determinada etapa do desenvolvimento (WILKINS et al., 1995) ou condição fisiológica específica. Enquanto um indivíduo possui um único genoma ele pode ter n tipos de proteomas, pois um único gene pode originar diversos produtos após as etapas de processamento, tornando a identificação de proteínas muito mais complexa e informativa do que a de genes. Através da análise do proteoma de uma célula submetida a determinados estímulos físicos, químicos ou biológicos pode-se inferir como estes estímulos regulam a expressão dos genes, e quais transformações estes produtos sofrem nos diferentes compartimentos celulares. Embora parte destes dados possa ser obtida também por análise de RNA (genoma expresso) a análise

conjunta do proteoma confere informações adicionais relativas ao produto final dos genes como estrutura, local de atuação na célula e ainda confere dados sobre possíveis interações entre o produto de vários genes para a realização de uma determinada função celular. Conforme considerado por Watson et al. (2003) “na busca da compreensão da função gênica e para estudar a biologia holística dos sistemas é inevitável que estudemos os proteomas”.

Espera-se que os dados sobre as proteínas diferencialmente expressas durante o desenvolvimento do xilema, analisadas em conjunto com o genoma expresso, possam contribuir para a compreensão de parte dos eventos genéticos que ocorrem durante o desenvolvimento da madeira de *Eucalyptus*, e estas informações abram novas possibilidades para o melhoramento genético desta espécie.

2 REVISAO BIBLIOGRÁFICA

2.1 Recursos florestais no Brasil

As florestas plantadas têm importância ecológica por constituírem um ecossistema que abriga diversas espécies, por reduzir a devastação de florestas naturais, atuarem como reguladoras de microclimas e por contribuírem para a qualidade da água de microbacias hidrográficas. Além disso, as árvores sequestram CO₂ excedente da atmosfera através de seu metabolismo de incorporação de carbono em compostos orgânicos. O principal produto das plantações florestais, a madeira, é uma energia renovável, e também é considerada um biocombustível que poderá ter grande impacto no futuro.

Apenas 0,5% da cobertura florestal brasileira é representada por florestas plantadas, cerca de 5 milhões de hectares que suprem em torno de 75% das necessidades da indústria nacional. As principais espécies cultivadas são *Eucalyptus* (60% da área plantada) e *Pinus* (36%), e em menor escala também são plantadas acácias, tecas e araucárias (4%).

O setor florestal movimentou em 2004 no Brasil cerca de US\$ 25 bilhões, representando 4,1% do Produto Interno Bruto (PIB) e gerou mais de 6 milhões de empregos diretos e indiretos (ABIMCI, 2004). O setor do agronegócio tem contribuído fortemente para o fortalecimento das exportações brasileiras nos últimos anos, e estão inseridos nesse ramo os produtos derivados do setor florestal. Estes produtos contribuíram com 17% do montante das exportações de 2004, cerca de US\$ 6,8 bilhões de dólares.

A demanda é crescente e o aumento da produtividade não segue o mesmo ritmo. A produção de madeira para uso industrial a partir de florestas plantadas no Brasil está próxima do limite de produção sustentada, o que pode em poucos anos provocar um “apagão florestal” comprometendo o crescimento ou mesmo a manutenção da indústria florestal. Para suprir a crescente demanda torna-se primordial a expansão das áreas florestais no país; de outra forma em poucos anos o Brasil deverá tornar-se importador de madeira de países vizinhos como Chile, Argentina e Uruguai.

2.2 O gênero *Eucalyptus* no Brasil

O gênero *Eucalyptus*, nativo da Austrália e pertencente à família Myrtaceae, possui cerca de 700 espécies (BROOKER, 2000) bem distribuídas em um amplo espectro de variedades de climas e solos.

A introdução do gênero no Brasil foi feita por Edmundo Navarro de Andrade que, no início do século XIX, passou a estudar e plantar diversas espécies de *Eucalyptus*, no Horto Florestal de Rio Claro-SP. Este e vários outros hortos supriram as necessidades da Companhia Paulista de Estradas de Ferro por madeira para lenha e para a construção de dormentes, moirões e postes. De São Paulo as plantações de diversas espécies de eucalipto espalharam-se para outros estados, sendo logo observado que diferentes espécies produziam madeira com qualidade para diversos fins.

As espécies de eucalipto são bem adaptadas a uma grande variedade de solos e temperaturas e as condições edafoclimáticas do Brasil permitiram a disseminação das plantações. A boa adaptação destas espécies, aliada aos avanços nas técnicas de silvicultura, permitiram alcançar rotações de culturas que variam entre 5 e 7 anos e conferem ao país as mais altas taxas de crescimento em plantios florestais no mundo (Carvalho et al., 2005). O Brasil possui e deve explorar portanto a vocação natural que possui para o setor florestal.

2.3 Importância econômica do eucalipto

As principais espécies de eucalipto plantadas no Brasil são *E. grandis*, *E. urophylla*, *E. saligna* e *E. citridora*, bem como alguns híbridos interespecíficos. A produtividade pode alcançar 50m³ de madeira por ha/ano. A madeira de eucalipto é matéria prima para o setor siderúrgico, de móveis, construção civil e fonte para a produção de papel e celulose. Este último setor consome cerca de 50% da produção de madeira em toras originadas de florestas plantadas, que em 2003 quase atingiu a marca de 1 bilhão de m³ (SBS, 2005). Os principais estados produtores de madeira para papel e celulose são São Paulo, Paraná, Bahia, Santa Catarina e Espírito Santo.

No Brasil as espécies mais utilizadas para a produção de papel e celulose são *E. grandis* e *E. urophylla*, além do híbrido entre as duas espécies. O Brasil é o 7º maior produtor mundial de celulose e lidera as exportações de celulose de fibra curta a partir de madeira de eucalipto. Entretanto, para manter essa posição de liderança, o setor de papel e celulose deverá crescer ao redor de 5% ao ano. Para suprir a demanda dos mercados interno e externo, o setor florestal também deverá buscar incrementos de produtividade para abastecer o mercado com madeira para a siderurgia e para a produção de carvão vegetal, serrados, compensados, painéis e móveis.

2.4 Melhoramento genético do eucalipto

A cultura do eucalipto pode ser atacada por uma série de pragas (formigas, cupins e lagartas), doenças como tombamento, podridão das raízes, oídio, ferrugem, entre outras causadas por várias espécies de fungos e vírus (CIESLA et al., 1996). Estas doenças podem se espalhar em novas regiões junto com a planta hospedeira devido à disseminação internacional de germoplasmas, e as perdas podem causar fortes impactos econômicos, especialmente em grandes plantações. O melhoramento clássico procura explorar a diversidade genética natural das várias espécies cruzando e selecionando indivíduos com genótipos resistentes. Com frequência esta seleção é assistida pelo uso de marcadores moleculares que são desenvolvidos para auxiliar a construção de mapas de ligação e a identificação de loci de características quantitativas (QTLs) (BRONDANI, et al., 2002), que podem ser relacionados a resistência a doenças (JUNGHANS et al., 2003), a características da madeira (THAMARUS et al., 2004), tolerância ao frio (VAILLANCOURT et al., 1994), entre várias outras.

Grande parte das plantações de eucalipto são destinadas à produção de celulose e uma grande dificuldade técnica deste processo está na etapa de remoção da lignina, que requer o uso de grande volume de compostos químicos altamente poluidores para o meio ambiente. Uma estratégia buscada pela biotecnologia para reduzir esta dificuldade é a manipulação genética de árvores para obtenção de plantas com menores teores de lignina ou alteração da composição química deste polímero. Transgênicos com reduzida concentração de lignina já foram obtidos por manipulação

de genes da via dos fenilpropanóides, como no caso de *Eucalyptus* (VALÉRIO et al., 2003), de *Populus* (JOUANIN et al., 2000) e de plantas modelo como *Arabidopsis* (MEYER et al., 1998) e tabaco (PIQUEMAL et al., 1998; BLEE et al., 2001). Um efeito decorrente da alteração da composição química da lignina de álamos transgênicos durante testes de polpação foi a redução do volume de produtos químicos necessário para sua remoção (PILATE et al., 2002). Embora algumas alternativas estejam sendo apresentadas, o controle da expressão dos vários genes envolvidos nesta rota biossintética deve ser melhor compreendido para tornar a manipulação genética mais eficiente e evitar a alteração indesejada de genes de outras rotas essenciais (BOERJAN et al., 2003).

2.5 Biotecnologia aplicada ao eucalipto

A biotecnologia vegetal vem ajudando a compreender os mecanismos moleculares envolvidos nos diversos tipos de metabolismos de plantas. Na busca de informações sobre o processo de formação da madeira, algumas das estratégias já empregadas com as diversas espécies de eucalipto foram a identificação e clonagem de genes (BHOORA et al., 2003; POKE et al., 2003; BOSSINGER e LEITCH, 1999), a construção de transgênicos (GONZÁLEZ et al., 2002; SARTORETTO et al., 2002; TOURNIER et al., 2003), otimização de protocolos de regeneração em cultura de tecidos (DIBAX et al., 2005; ALVES et al., 2004; HERVE et al., 2001) e estudos de expressão por *microarrays* e sequenciamento de ESTs (SATO et al., 2001; PAUX et al., 2004).

Alguns projetos de sequenciamento de genomas expressos (ESTs) de *Eucalyptus* já foram realizados por alguns grupos no mundo e, no Brasil, os projetos FORESTs e GENOLYPTUS têm destaque internacional. O FORESTs resultou em um banco de dados com 123.889 ESTs de diferentes espécies e tecidos de *Eucalyptus* (Carrer, 2005). O GENOLYPTUS gerou em sua primeira etapa um banco com 150.000 ESTs, incluindo também diferentes tecidos e espécies de *Eucalyptus*, embora a etapa de sequenciamento deva ser complementada com outros subprojetos que estão em andamento, incluindo a construção de mapas de ligação, análises de expressão por

microarrays, bioinformática, e outros. Variados esforços estão sendo desenvolvidos com *Eucalyptus* utilizando ferramentas genômicas em países como Austrália, França, África do Sul, China, e outros (POKE et al., 2005).

Embora os grandes projetos de sequenciamento permitam a identificação de muitos genes, as pesquisas devem aumentar a compreensão sobre os elementos reguladores que estão envolvidos no controle de sua expressão. Um exemplo de fatores de regulação de transcrição recentemente observado em *Eucalyptus* é o gene *EgMYB2*, pertencente à família dos fatores MYB, comuns em plantas. Através da clonagem deste gene em tabaco foi possível demonstrar sua participação na regulação da síntese de lignina e a relação disto com o padrão de formação da parede celular (GOICOECHEA et al., 2005), ambos eventos importantes na formação da madeira. Este tipo de informação criará novas possibilidades para a manipulação genética na busca de árvores com fenótipos mais favoráveis à cultura.

2.6 Xilogênese

A formação da madeira (xilogênese) é a principal característica que diferencia árvores perenes de plantas herbáceas anuais e resulta do contínuo crescimento secundário do meristema cambial. A madeira tem a importante função de conferir suporte mecânico para o transporte de água e nutrientes das raízes às folhas e é um tecido que, durante sua formação, sofre influência de fatores genéticos e ambientais. Apesar da reconhecida importância econômica, existe pouco conhecimento sobre os fatores genéticos envolvidos na formação deste tecido (PLOMION et al., 2001).

Esta região meristemática é um estreito anel concêntrico de poucas camadas de células indiferenciadas que determina o crescimento em diâmetro do tronco das árvores. No sentido centrífugo as células iniciais cambiais diferenciam em células que compõem o floema e, no sentido centrípeto, o xilema. Os fatores que determinam o destino das células iniciais cambiais durante o desenvolvimento não são totalmente compreendidos, mas tem sido demonstrada a relação da regulação hormonal por ácido indole-3-acético (JACOBS, 1952; SAVIDGE e BARNETT, 1993) e por giberelinas (ERIKSSON et al., 2002) com número e extensão de fibras do xilema.

As células cambiais e suas derivadas têm sido detalhadamente estudadas também por microscopia eletrônica e de imunofluorescência (CHAFFEY, 1999) na busca de informações sobre sua anatomia, composição química e etapas de diferenciação. Espera-se que o conhecimento sobre a composição e organização de microtúbulos e filamentos do citoesqueleto de células da madeira contribua para aumentar o leque de opções para a manipulação genética visando alterações da composição celular e, portanto de seus tecidos (CHAFFEY et al., 2002).

A madeira é composta principalmente de celulose, hemiceluloses, lignina e pectinas. Portanto as rotas metabólicas de síntese destes compostos têm sido amplamente estudadas e vários dos genes envolvidos em sua síntese estão bem caracterizados. A estratégia para alcançar maior conhecimento sobre genes relacionados a esta e outras rotas em árvores tem sido o uso de espécies modelos como *Arabidopsis*, mais bem caracterizadas geneticamente (JOSHI et al., 2004). O primeiro gene da rota da celulose sintase de plantas (*CesA*) foi clonado a cerca de 10 anos de plantas de arroz e algodoeiro (PEAR et al., 1996; ARIOLI et al., 1998). Atualmente sabe-se que em *Arabidopsis* há 10 genes de CSA e ao menos 6 famílias de genes CSL (relacionados ao da celulose sintase). Sabe-se que alguns destes genes estão associados à formação da parede celular primária (*AtCesA1*, *AtCesA2*, *AtCesA3*, *AtCesA5*, *AtCesA6*, *AtCesA9* e *AtCesA10*) enquanto outros (*AtCesA4*, *AtCesA7* e *AtCesA8*) estão associados à formação da parede celular secundária (KALLURI AND JOSHI, 2003; 2004; DOBLIN et al., 2002). Apesar da caracterização de diversos genes da rota faltam conhecimentos sobre os elementos reguladores da expressão (READ e BASIC, 2002). Espera-se que, após o sequenciamento do genoma de *Populus* (STERKY et al., 2004), este importante metabolismo da madeira possa ser estudado dentro da heterogeniedade de sistemas de árvores.

A biossíntese de lignina, o segundo mais abundante polímero na natureza depois da celulose, também tem sido extensivamente estudada (HERRMANN, 1995; DONALDSON, 2001). Os monômeros de lignina são formados pela via dos fenilpropanóides e genes envolvidos nesta via já foram identificados em eucalipto como o da cinnamoyl-CoA redutase (CCR; LACOMBE et al., 1997), da caffeic acid O-methyltransferase (COMT) e cinnamyl alcohol dehydrogenase (CAD) (HAWKINS et al,

2003). O gene da enzima caffeic acid O-methyltransferase também já foi super expresso em *Populus* (JOUANIN et al 2000), resultando em alterados teores de lignina na planta. A lignificação também têm sido estudada em sistemas de células livres de *Populus*, pois a composição da lignina nestas células mostra-se semelhante ao de árvores (STASOLLA et al., 2003).

A identificação de genes, preferencialmente expressos em xilema, pode demonstrar sua atuação na formação do tecido, mas para relacioná-los com as diversas estruturas da madeira é necessário esclarecer de que maneira atuam durante o desenvolvimento. Arabinogalactanas são importantes proteínas relacionadas ao crescimento e desenvolvimento de plantas e são codificadas por diversas famílias de genes, preferencialmente expressas em xilema (LOOPSTRA E SEDEROFF, 1995). O papel destes genes foi estudado em vários tipos de madeira de *Pinus taeda*: xilema maduro e juvenil, madeira de reação e células *in vitro* induzidas para lignificação (YANG et al., 2005). Foi observado que os vários genes expressam-se de maneira diferencial nos vários tecidos demonstrando que, embora pertencentes a uma mesma família, os genes podem ter diferentes funções e diversos controles de expressão dentro de um mesmo tipo de tecido.

Na busca de genes relacionados à qualidade da madeira e características de crescimento, comparou-se o nível de expressão de genes do xilema de três genótipos de *E. globulus* e *E. grandis* com fenótipos contrastantes (KIRST et al., 2002). Dentre os genes avaliados, apenas quatro apresentaram expressão diferencial, o que possivelmente sugere que poucos genes devem ter papel chave na determinação da qualidade da madeira, e que mais relevante pode ser a regulação da expressão.

Embora muitos genes relacionados à xilogênese estejam bem identificados por clonagem ou mapeamento baseado em marcadores moleculares, a complexa regulação de sua expressão demanda mais esforços que esclareçam a rede de controle espaço-temporal dos genes e permitam atingir o desejado nível de manipulação genética para a construção da “planta ideal” para as várias finalidades industriais.

2.7 Modelos para o estudo da xilogênese

Devido à facilidade de manipulação e/ou ao seu valor econômico, algumas espécies tem sido eleitas como modelos para estudos de xilogênese. *Populus* é a árvore até hoje mais estudada dentro do grupo das angiospermas, e é de grande interesse para a silvicultura de países do hemisfério norte. Características como genoma pequeno (550 Mb), facilidade de transformação genética e de regeneração *in vitro* fizeram desta espécie um modelo para estudos de formação da madeira. Recentemente *Populus* tornou-se a primeira árvore a ter seu genoma seqüenciado (STERKY et al, 1998, Brunner et al, 2004).

Também a planta modelo *Arabidopsis thaliana*, aceita como modelo em genética de plantas devido ao seu pequeno porte e curto ciclo de vida, foi totalmente seqüenciada (*The Arabidopsis Genome Initiative*, 2000). A ortologia entre genes de *Arabidopsis* com os de outras espécies é bem observada. Genes de celulose sintase de *Arabidopsis* e *Populus*, associados ao desenvolvimento da parede celular secundária, mostraram 91% de similaridade (SAMUGA E JOSHI, 2002).

Além das espécies citadas, há tipos de células e tecidos que representam bons modelos de estudo de xilogênese, como a madeira de reação (PILATE et al., 2003). Este tipo de tecido forma-se quando um caule ou galho de árvore cresce de maneira acêntrica forçando o desenvolvimento de tecidos mais resistentes que tendem a puxar a estrutura de volta a sua posição original. Regiões da madeira formadas sob este tipo de tensão têm vasos e fibras anatomicamente alterados. A observação da expressão diferencial de genes e proteínas neste tipo de madeira, em relação à normal, permite inferir quais genes estão envolvidos na formação destas estruturas. O nível de expressão de diversos genes de madeira de tensão de *Eucalyptus globulus* relacionados à biossíntese de celulose foram 6,5 vezes mais expressos do que em madeira normal (PAUX et al., 2005). Proteínas de madeira de tensão e de madeira normal de *Eucalyptus gunnii* também foram comparadas e apresentaram expressão diferencial nos tecidos indicando sua possível ligação com o desenvolvimento do xilema (PLOMION et al., 2003).

Os tecidos de árvores, de um modo geral, mostram-se recalcitrantes a muitos processos *in vitro*, como regeneração e transformação genética. Devido a esta dificuldade, sistemas de células em suspensão como de *Zinnia elegans* têm servido para acompanhar a expressão de genes envolvidos em diversos processos celulares como diferenciação de traqueídeos (Fukuda, 1997) e de genes envolvidos em lignificação (YE et al., 1994). Entretanto, este sistema é limitado pela ausência dos efeitos de um sistema heterogêneo de tecidos, onde a comunicação celular pode ser mais um fator regulador do desenvolvimento (CHAFFEY, 1999).

Pinus também é um gênero de alto valor econômico e é utilizado como modelo de estudo de madeira representando as características das coníferas (LEV-YADUN E SEDEROFF, 2000). Allona et al. (1998) sequenciaram 1097 ESTs de *Pinus taeda* e relacionaram várias delas com a formação da parede celular, metabolismo de carbono e síntese de lignina. Entretanto, esta espécie é mais recalcitrante em cultura de tecidos e, apesar de ser uma lenhosa de grande importância econômica para países como os Estados Unidos, tem sido menos utilizada para pesquisas em genética. Apesar da distância filogenética com angiospermas, tem sido sugerido que suas funções sejam estudadas em *Arabidopsis* ou *Populus*, pois há significativa similaridade de sequência em diversos genes ortólogos (YANG et al., 2005).

Apesar dos grandes esforços para sequenciar genomas inteiros (DNA) e expressos (RNA) de árvores e da grande quantidade de informação gerada por estes projetos, o nível de complexidade da regulação gênica da formação da madeira requer o uso de técnicas que permitam avaliar também o controle molecular existente não só durante, mas após a etapa de transcrição. *Splicing* e modificações pós-traducionais são etapas de maturação química que definem o produto de um gene, sua função e local de atuação ao nível subcelular. Estas informações vêm sendo mais recentemente estudadas por análise de proteomas celulares, dentro dos conceitos de uma área de trabalho denominada proteômica.

2.8 Proteoma de plantas

O conhecimento de todo o código genético de uma planta pode ser de pouco valor se não houver informações sobre a regulação espaço-temporal dos genes. Esta regulação

está sendo explorada por um campo denominado genômica funcional onde é dado enfoque ao produto dos genes, sua estrutura, inter-relações e conseqüências no contexto da célula. A proteômica é uma forte ferramenta deste campo e é uma área emergente dentro da genética desde a década passada. A proteômica pode ser definida como o campo da ciência que estuda proteínas, as verdadeiras efetoras das funções celulares, através de tecnologias que geram grande quantidade de dados sobre estrutura e seqüência com enorme precisão e rapidez, permitindo o estudo de todo o proteoma de um tecido.

Embora o proteoma de uma espécie seja bem mais complexo do que o genoma, as ferramentas disponíveis para a análise de proteínas são mais limitadas. A PCR (reação em cadeia da polimerase) permite amplificar uma seqüência de DNA de maneira que é possível dispor de muito material para as análises. Uma técnica equivalente não existe ainda para proteínas, portanto é necessário adaptar os equipamentos a uma sensibilidade compatível com a baixa concentração destas moléculas em amostras biológicas. Devido a estas dificuldades, o estudo de proteínas ganhou impulso apenas na década de 90 com a difusão da espectrometria de massas (MS) nas ciências biológicas, uma vez que tornou-se possível analisar proteínas a níveis tão reduzidos quanto os encontrados na natureza.

Recentemente, foi divulgado o proteoma do arroz (*Oriza sativa* L.) (AGRAWAL & RAKWAL, 2006), uma espécie aceita como modelo para o grupo das monocotiledôneas (GOFF, 1999). Através de diversas técnicas como gel 2D, MS, sequenciamento de Edman e tecnologia multidimensional de identificação de proteínas (MudPit), milhares de proteínas e seus genes foram identificadas a partir de várias cultivares de arroz utilizando órgãos reprodutivos, folhas, raízes, sementes e organelas (IMIN ET AL., 2004; AGRAWAL et al., 2002). Outros importantes cereais como milho e trigo também já tiveram o proteoma de vários tecidos extensivamente analisados, inclusive após serem submetidos a estresses biótico ou abiótico (CHANG et al., 2000; RANIERI ET AL, 1995, MAJOUL et al., 2004).

O primeiro proteoma de árvores foi desenvolvido com *Pinus pinaster*, quando foram caracterizadas 63 proteínas de tecidos fotossintetizantes e do xilema resolvidas em géis 2D e identificadas por MS (COSTA et al., 1999).

Trabalhando com *Populus*, Vander Mijnsbrugge et al. (2000) analisaram proteínas da madeira através de géis 2D e microsequenciamento interno e observaram que embora haja semelhanças no padrão de proteínas expressas em xilema maduro, xilema juvenil e na casca podem ser observadas diferenças no nível de expressão entre os tecidos, demonstrando que existe um controle espacial da expressão destes genes. Também foi observado que, proteínas preferencialmente expressas no xilema, são mais abundantes nos períodos de verão, ficando demonstrado que informações referentes ao efeito de fatores ambientais sobre a expressão de genes pode ser obtida por análise de proteínas. Recentemente foi estabelecida uma comparação entre o proteoma do xilema jovem (3 anos) com o xilema maduro (10 anos) de *Pinus pinaster* (GION et al., 2005), onde algumas proteínas típicas de cada fase do desenvolvimento foram caracterizadas. As proteínas são diferencialmente expressas em uma célula dependendo da idade e das condições fisiológicas de um tecido. Existem diversas estratégias em proteômica que permitem identificar e quantificar proteínas com sensibilidade e rapidez, e estão sendo utilizadas a poucos anos para estudar a expressão gênica de plantas.

2.9 Estratégias proteômicas

Uma estratégia clássica dentro da área de proteômica é o uso de géis bidimensionais para separar centenas ou milhares de proteínas a partir de um extrato celular total (O'Farrel, 1975), seguido da identificação dos *spots* isoladamente por espectrometria de massas. Dois procedimentos são mais comumente adotados: caracterização geral do proteoma com identificação do maior número possível de proteínas ou identificação apenas das proteínas diferencialmente expressas entre duas amostras com características fenotípicas contrastantes, como sensibilidade e resistência a uma doença. Os *spots* com as proteínas de interesse são submetidos à digestão enzimática, sendo tripsina a mais utilizada, e os peptídeos resultantes são analisados por MS.

Os géis bidimensionais são vantajosos na medida em que conferem uma visão global do proteoma de um determinado tecido e permitem a pré-separação das proteínas da amostra por ponto isoelétrico (pI) e peso molecular (PM). Entretanto, há dificuldades associadas a esta técnica como a difícil reprodutibilidade dos resultados de

eletroforese, baixa resolução dos *spots* em algumas faixa de PM e pI, pouca sensibilidade para detecção de proteínas de baixa concentração e pequena representação de proteínas hidrofóbicas. Estas dificuldades técnicas vêm sendo superadas pelo uso de maior número de repetições de géis (RUBINFELD et al., 2003), cuidadosas análises de imagem destes géis (ROGER et al., 2003.), seleção de faixas de pH estreitas que aumentam a resolução da amostra (GÖRG et al., 1999), enriquecimento de frações de proteínas baseado em características químicas (CORTHALS et al., 2000) e aplicação de protocolos específicos para recuperação de proteínas de baixa solubilidade (MÉCHIN et al., 2003).

Os géis 2D têm sido utilizados preferencialmente no caso de comparação de proteomas em diferentes estádios fisiológicos ou do desenvolvimento, onde a observação de proteínas diferencialmente expressas, permite inferir o envolvimento dos genes correspondentes a cada estágio em questão. O proteoma nuclear de *Arabidopsis* foi caracterizado com auxílio de géis 2D e MS, e comparado com o proteoma da mesma organela resultante de estresse por baixa temperatura. Nesse caso 54 proteínas nucleares tiveram expressão diferencial devido ao estresse (BAE et al., 2003). Com árvores de interesse florestal e tecidos da madeira, estas técnicas foram utilizadas para comparar extratos de proteínas do xilema com tecidos fotossintetizantes (COSTA et al., 1999), madeira normal e de tensão (PLOMION et al., 2003) e tecidos somáticos em diversas fases do desenvolvimento (LIPPERT et al., 2005).

Uma estratégia interessante que pode preceder a eletroforese 2D é o fracionamento subcelular onde tecidos homogeneizados são submetidos à centrifugação com gradiente de densidade para separação de organelas e estruturas subcelulares (DREGER, 2003). A partir de cada fração pode ser extraído um grupo diferente de proteínas metabolicamente relacionadas. O fracionamento possibilita a redução da complexidade da amostra nos géis 2D, beneficiando a resolução e aumentando a probabilidade de identificação das proteínas.

Embora a espectrometria de massas tenha sido o ponto culminante no desenvolvimento da proteômica, o desafio da descoberta de novos produtos gênicos somente será alcançado através da união de diversas técnicas que superem as dificuldades intrínsecas a cada material biológico.

2.10 Espectrometria de massas para análise de proteínas

O marco histórico do início da espectrometria de massas data ao século XIX quando da construção do primeiro equipamento por Sir Joseph John Thomson que buscava determinar a relação massa/carga de íons. A técnica tornou-se bem difundida nos meios acadêmicos e por volta da década de 30 já era bem utilizada em pesquisas na área de física. Na década de 50, observou-se que a técnica poderia ser de grande importância científica se utilizada para a análise de moléculas orgânicas, porém a ionização, requisito fundamental para geração do espectro de massas, não era possível devido à maior massa destas moléculas e à sua polaridade (MANN et al., 2001). A técnica sofreu grande limitação dentro das ciências biológicas embora diversas técnicas e equipamentos tenham sido desenvolvidos. Somente no fim da década de 80, foram criadas duas técnicas de ionização para a espectrometria de massas, que permitiam ionizar macromoléculas e produzir o espectro das massas de seus íons. Foram estas a ionização por desorção a laser, ou MALDI (KARAS e HILLENKAMP, 1988) e o *electrospray*, ou ESI (FENN et al, 1989). Este novo processo foi tão relevante para as ciências que seus principais criadores, John Fenn e Koishi Tanaka foram condecorados com o Prêmio Nobel de Química em 2002. A partir da década de 90, a espectrometria tornou-se uma poderosa ferramenta na análise de biomoléculas permitindo alcançar níveis de sensibilidade, rapidez e resolução das amostras a um nível até então inesperado. Desde então, diversos modelos de equipamentos foram desenvolvidos.

2.11 Equipamentos

Um espectrômetro de massas é basicamente composto de três elementos: o ionizador, o analisador de massas e o detector.

As duas formas de ionização mais comumente utilizadas para análise de proteínas são MALDI e ESI que diferem na forma de emissão do feixe de íons sobre a amostra. O termo ionização se refere à passagem da amostra para a fase gasosa mantendo e/ou adquirindo uma carga que permite que ela seja captada pelo equipamento. Na

ionização ESI, a amostra líquida é injetada por um tubo capilar e, ao emergir, atinge um campo de alto potencial. O efeito é a formação de um *spray* de gotículas altamente carregadas (íons) que são encaminhadas a um analisador. Os analisadores mais comuns são do tipo quadrupolo e o tipo “TOF” (“time of flight” ou tempo de voo).

O analisador tipo quadrupolo foi desenvolvido por Paul e Steinwedel (1953) e separa íons filtrando-os pelo uso de um campo elétrico quadrupolar que define suas trajetórias. Somente os íons selecionados desenvolverão uma trajetória que lhes permitirá alcançar o caminho do detector. O outro analisador, do tipo TOF, foi desenvolvido com base em um instrumento inicialmente concebido por Wiley e MacLaren (1955) que separa os íons de acordo com diferenças na velocidade com que percorrem um tubo no vácuo. Íons de menor m/z (relação massa/carga) alcançam primeiro o detector.

Devido à alta velocidade de operação, os dois tipos de analisadores podem ser acoplados à cromatografia gasosa (GC-MS) ou líquida (LC-MS), e identificar o grande volume de moléculas geradas simultaneamente por estas duas técnicas.

Equipamentos híbridos são comuns, como o ESI-Q-TOF, compostos de dois analisadores, um quadrupolo (Q) seguido de um analisador TOF (figura 1). Este sistema possui alta sensibilidade, podendo detectar moléculas em concentrações tão reduzidas quanto picomoles (10^{-12} moles) ou femtomoles (10^{-15} moles). Seu sistema de ionização (ESI) requer um contínuo fluxo líquido de amostra, o que é uma excelente qualidade para ser acoplado a um sistema de cromatografia líquida de alta performance (HPLC) que purifica, concentra e separa as moléculas antes de serem introduzidas no espectrômetro. Os peptídeos são capturados em uma coluna de fase reversa, eluídos por hidrofobicidade e imediatamente injetados na interface do ESI através de um tubo capilar, que permite trabalhar em microescala. Este sistema de baixo volume de injeção favorece que amostras de volumes entre 1 e 5 μ l sejam analisadas, e que mais experimentos sejam realizados a partir da mesma amostra. A eluição por gradiente cria um enriquecimento de classes de peptídeos que favorece sua identificação ao evitar que os mais concentrados mascarem o restante da amostra. A figura 1 mostra um esquema da composição do espectrômetro Q-TOF acoplado a um sistema de HPLC, apontando o caminho da amostra desde sua introdução até a geração dos espectros de massas.

O Q-TOF é um tipo de equipamento que permite realizar a espectrometria de massas em tandem (MS/MS), o que resulta em seqüências de peptídeos, uma estratégia que superou a identificação de proteínas por massas, pois aumenta a probabilidade das identificações. Com a estratégia MS/MS até um único peptídeo pode gerar alta confiabilidade e validar a identidade de uma proteína (Shevchenko et al., 1996).

Figura 1 – Modelo de construção do Q-TOF com operação em modo MS/MS. Fonte: Waters

2.12 Bioinformática

Em 1989, foi sugerido pela primeira vez o uso de massa molecular de peptídeos gerados por MS para identificação de proteínas em bases de dados (HENZEL et al., 1989) e em poucos anos a estratégia tornou-se vastamente difundida em análises de proteínas. Para serem traduzidos em informação biológica os espectros de íons de peptídeos gerados por MS devem ser interpretados de maneira manual ou automática. O modo manual é um processo trabalhoso devido à grande quantidade de espectros gerados por amostra, além de ser necessário um bom grau de conhecimento e experiência. Assim, as análises são preferencialmente realizadas de forma automática. Com esta finalidade diversos tipos de softwares foram especialmente desenvolvidos (RUSSELL et al, 2004) e são freqüentemente atualizados, podendo ser utilizados por

profissionais das áreas biológicas com menor conhecimento dos processos físicos e químicos inerentes à espectrometria de massas.

Para identificação de proteínas, os programas basicamente interpretam os dados de fragmentação dos espectros e convertem esta informação em massas de peptídeos (*fingerprinting*) e/ou seqüências de aminoácidos, dependendo dos recursos do equipamento (SHEVCHENKO et al, 2001). Os softwares realizam digestões teóricas em proteínas depositadas em bancos de dados e comparam estes valores com os dados experimentais. A comparação entre informações teóricas e experimentais gera listas de proteínas com diferentes graus de homologia com o dado experimental. As listas são ordenadas de acordo com uma probabilidade estatística que é determinada pelo método de cálculo do algoritmo e pelos parâmetros de tolerância definidos pelo usuário (CHAMRAD, et al., 2003). A adstringência da busca fica definida por parâmetros como tolerância de erro de massa de cada aminoácido, número de substituições aceito para cada peptídeo, tipos de modificações pós-traducionais a considerar, metodologia usada para gerar os peptídeos (química, enzimática, triptica, etc.).

Existem muitos tipos de programas com algoritmos desenvolvidos para realizar interpretação de espectros MS/MS e buscas em bancos de dados. SEQUEST (YATES et al., 1995) e MASCOT (PERKINS et al., 1999) são dois dos programas mais utilizados, mas a exemplo do que ocorre com a maior parte dos softwares existentes, o seu caráter comercial limita as informações disponíveis sobre os cálculos efetuados pelos seus algoritmos. Uma dificuldade de quase todos os programas reside nos falsos negativos que podem ser gerados quando não é encontrada homologia com peptídeos das bases de dados devido, por exemplo, à presença de modificações pós-traducionais nas seqüências experimentais que alteram sua massa (FRANK et al., 2005). O programa ProteinLynx (Waters, Micromass; DENNY et al., 2002) é um dos programas capaz de processar os dados baseado em uma seqüência de filtros selecionados pelo usuário, podendo ser incluída a busca por modificações. Por outro lado falsos positivos são comumente reduzidos pela adstringência dos parâmetros de busca.

2.13 Bancos de ESTs na identificação de proteínas

Os projetos genoma ganharam impulso durante a década de 80 com o desenvolvimento de tecnologias que possibilitavam o rápido sequenciamento de ácidos nucleicos. Diversas espécies têm seus genomas anotados e um grande volume de seqüências encontra-se publicamente disponível. Atualmente mais de 1700 projetos já foram concluídos ou estão em andamento (BERNAL et al., 2001). Devido à existência de grandes regiões do DNA não codificarem genes, tem sido muito utilizada a alternativa de sequenciar apenas as regiões codificadoras, dando-se origem aos bancos de etiquetas de seqüências expressas (ESTs) ou de genomas expressos.

O uso de bancos de ESTs espécie-específicos está sendo de primordial importância para os projetos proteoma pois resulta em maior probabilidade de identificação de proteínas por espectrometria de massas. Lippert et al. (2005) aumentaram em 29,2% a taxa de identificação de proteínas de embriões somáticos de *Picea glauca*, uma conífera de importância comercial na América do Norte. Da mesma forma, ao analisar o perfil de proteínas de diversos tecidos de madeira de *Pinus pinaster* através de géis 2D e MS, Gion et al. (2005) conseguiram 53% mais proteínas identificadas através de buscas em bases de ESTs de *Pinus pinaster* e de *Pinus taeda*. Embora o uso de bancos de ESTs aumente a taxa de identificação de proteínas, não é observada uma alta correlação qualitativa (ANDERSON E SEILHAMER, 1997) ou quantitativa (GYGI et al., 1999) entre o RNAm e as proteínas relacionadas extraídas de um mesmo tecido. Esse trabalhos mostram que a presença do RNAm não é indicativo da atividade da proteína equivalente na célula, tornando-se necessário o complemento das informações de expressão gênica com o estudo de proteínas. As diferentes taxas de síntese e degradação das duas moléculas pode ser uma das causas da baixa correlação (CARGILE et al, 2004) e esta estabilidade molecular pode ser global ou específica, dependendo do ambiente e talvez do estágio de desenvolvimento. É possível que, somente com a união das diferentes ferramentas da genômica funcional será possível obter um diagnóstico mais próximo dos eventos moleculares que ocorrem nos sistemas biológicos.

3 MATERIAL E MÉTODOS

3.1 Coleta do material

O tecido vegetal utilizado neste trabalho foi definido como região cambial ou xilema em desenvolvimento do caule de árvores de *Eucalyptus grandis*. A coleta foi realizada em campo, na estação de Itapetinga (SP) da Companhia Suzano de Papel e Celulose, em janeiro de 2005. Foram escolhidos aleatoriamente 50 indivíduos de uma progênie de meio-irmãos de 2 anos e 8 meses de idade da matriz M33 de eucalipto.

A técnica da coleta consistiu na retirada de um retângulo de 20 x 15 cm da casca do tronco das árvores seguida de raspagem com lâmina do tecido jovem existente entre a casca e o xilema lignificado (figura 2). O tecido, chamado de xilema em diferenciação (LEV-YADUN e SEDEROFF, 2000), que abrange diferentes camadas da região cambial, foi imediatamente congelado em nitrogênio líquido, parcialmente macerado e acondicionados para o transporte em tubos Falcon de 50 ml em gelo seco. No laboratório, o material foi armazenado em freezer a -80°C até o momento da extração das proteínas.

Figura 2 – Metodologia de coleta de tecidos do xilema em desenvolvimento

3.2 Extração de proteínas

O material vegetal congelado no -80°C foi macerado a um pó fino em nitrogênio líquido com a ajuda de cadinho e pilão. Duas técnicas de extração de proteínas foram

testadas: a extração ácida (DAEMERVAL et al, 1986) e a extração fenólica (HURKMAN E TANAKA, 1986).

A extração ácida consistiu em transferir 2 g do macerado para um tubo Nalgene com 10 ml de tampão composto de 10 % (p/v) de ácido tricloroacético (TCA) e 0,7 % (v/v) de β -mercaptoetanol em acetona. O material foi cuidadosamente homogeneizado e a precipitação das proteínas ocorreu durante 1 h a -20°C . Os tubos foram centrifugados por 15 min a 16000 g a -20°C . O sobrenadante foi descartado e o pellet foi duas vezes lavado com 0,7 % (v/v) de β -mercaptoetanol em acetona, sendo as lavagens seguidas de precipitação das proteínas durante 1 h a -20°C e centrifugação conforme citado. O pellet final foi completamente seco com auxílio de bomba de vácuo a 4°C . Para ressuspender as proteínas do pellet foi utilizado tampão de solubilização consistindo em 7 M uréia, 2 M tiuréia, 0,4 % (v/v) triton-X 100, 4 % (p/v) chaps (Bio-Rad 161-0460), 10 mM de dithiothreitol (DTT; Bio-Rad 161-0611). O pó foi misturado com o tampão em tubos Eppendorf de 1.5 ml, na proporção de 30 $\mu\text{l}/\text{mg}$ de pó. O material foi homogeneizado em Vortex em baixa rotação por 10 segundos e centrifugados por 2 min em microcentrifuga a 5.000 g à temperatura ambiente. O sobrenadante foi transferido para outro tubo e a centrifugação repetida até que a amostra estivesse visivelmente limpa (2 a 3 vezes).

A extração fenólica consistiu em homogeneizar 1,5 g do material vegetal em 10 ml de tampão de extração composto de 1% (p/v) polyvinylpolypirrolidone (PVPP), 50 mM de EDTA, 0,5 M de Tris-HCl pH 7,5, 0,7 M de sacarose, 2 % (p/v) β -mercaptoetanol e 1 mM de phenylmethylsulfonyl fluoride (PMSF). Após 30 min de homogeneização sob agitação constante no gelo, acrescentou-se o mesmo volume de fenol saturado em Tris pH 8.5, seguindo-se mais 30 min de homogeneização nas mesmas condições. O material foi então centrifugado 30 min a 10.000 g a 4°C , e a fase fenólica foi transferida para um novo tubo. O fenol recuperado foi re-extraído com iguais volumes, primeiramente de tampão e em seguida de fenol, sendo homogeneizado por 30 min a cada acréscimo. A re-extração foi seguida de centrifugação nas condições anteriores. A re-extração foi repetida 2 ou 3 vezes, até todos os resíduos serem visivelmente removidos da fase fenólica. As proteínas do fenol recuperado na última re-extração foram precipitadas por 16 h a -20°C pelo acréscimo de 5 volumes de 0.1 M de acetato

de amônio em metanol. No dia seguinte, os tubos foram centrifugados por 30 min a 10.000 *g* e a 4°C, sendo o sobrenadante descartado. O pellet recuperado foi lavado 2-3 vezes com 0.1 M de acetato de amônio em metanol e 1 vez com acetona gelada. A cada etapa as proteínas foram precipitadas por 30 min a -20°C e recuperadas por centrifugação nas mesmas condições anteriores. O pellet final foi seco por 1 ou 2 dias a 4°C e ressuspendido em 1 ml do mesmo tampão de solubilização utilizado na extração ácida. Após cada extração, a qualidade da amostra foi avaliada em eletroforese de gel de poliacrilamida com SDS (SDS-PAGE) (LAEMMLI, 1970) e as proteínas quantificadas pelo método de Bradford (1976).

3.3 Separação eletroforética de proteínas

A mistura de proteínas totais da madeira foi separada por eletroforese bi-dimensional (O'FARREL, 1979). Duas faixas de pH foram utilizadas para a os géis: pH 3-10 não linear (Amersham Biosciences cód.17-123501) para uma caracterização geral da amostra, e pH 4-7 (Amersham Biosciences cód.17-123351), para o sequenciamento de proteínas.

A focalização isoeletrica para as duas faixas de pH foi conduzida em sistema Ettan IPGPhor (Amersham Biosciences) carregando as fitas com 700 µg de proteínas solubilizadas, em 400 µl de tampão de solubilização, com 1% (v/v) de IPG buffer (Amersham Biosciences cód. 17-6000-88 para pH 3-10NL, e cód. 17-6000-86 para pH 4-7). As mesmas condições de focalização foram aplicadas a ambos géis consistindo em 12 h de reidratação ativa a 100 V, seguidas de 1 h a 200 V, 1 h a 400 V, 1 h a 600 V, 1 h a 800 V e 1 h a 1000 V; finalmente cerca de 16 h até ser alcançada a voltagem de 65.000 V. As fitas focalizadas foram armazenadas a -80°C até serem equilibradas antes da 2ª dimensão eletroforética.

A solução de equilíbrio consistiu de 50 mM Tris-HCl pH 6,8, 6 M uréia, 30 % (v/v) glicerol e 2 % (p/v) sodium dodecyl sulfate (SDS). O equilíbrio consistiu em imergir as fitas na solução de equilíbrio com 1 % (p/v) DTT por 15 min e em seguida na mesma solução com 2,5 % (p/v) de iodoacetamida (IAA, Sigma cód. I6125) e traços do corante

bromofenol blue por 15 min. As fitas foram moldadas no alto de um gel vertical com 0.5 % de agarose em tampão de eletroforese (25 mM Tris, 250 mM Glicina, 0,1 % (p/v) SDS) para realizar a segunda dimensão eletroforética. A eletroforese em sistema SDS-PAGE foi realizada em gel de 12% de poliacrilamida com 1,5 mm de espessura usando o sistema Protean II XL Cell (Bio-Rad). A migração das proteínas foi efetuada com amperagem fixa a 30mA por gel, até o corante alcançar a base do gel.

Os géis foram corados com Coomassie Brilliant Blue G-250 (Bio-Rad cód. 161-0406) de acordo com Candiano et al. (2004) e modificações conforme descrito a seguir: 2 lavagens de 1 h em 3 % (v/v) de ácido fosfórico e 50 % (v/v) de etanol; 1 lavagem de 1 h em 2 % (v/v) de ácido fosfórico; e finalmente 16 h em 10 % (v/v) de ácido fosfórico, 10 % (p/v) de sulfato de amônio, 17 % (v/v) de etanol e 0,1 % (p/v) de Coomassie G-250. Todas as etapas foram realizadas sob agitação. O excesso de corante foi removido com água MilliQ e os géis conservados em 15 % (p/v) de sulfato de amônio a 4°C.

3.4 Análise de imagem dos géis

A digitalização das imagens dos géis foi feita em scanner modelo UTA-1100, através do programa LabScan versão 5.0 (Amersham Bioscience) e as imagens analisadas por programa Image Master Platinum versão 5.0 (Amersham Bioscience). Tanto os géis de caracterização (pH 3-10NL) quanto os de sequenciamento (pH 4-7) foram feitos em triplicata. Os *spots* foram determinados em modo automático com os seguintes parâmetros: *smooth* = 2, *Min. Area* = 21 e *saliency* = 200.000. Após este procedimento algumas adições e deleções de *spots* foram feitas manualmente e os géis foram alinhados para verificar a reprodutibilidade entre os *spots*. Para obter uma quantificação relativa de cada proteína dentro da amostra os volumes médios dos *spots* foram obtidos em percentual sobre o total de 523 *spots* detectados. Os pesos moleculares (PM) e pontos isoelétricos (pI) experimentais das proteínas foram determinados automaticamente pelo programa a partir de 5 pontos de referência anotados manualmente, e os teóricos, foram sugeridos pelo banco do SwissProt referentes à proteína com maior homologia com cada dado experimental.

3.5 Digestão *in-gel* das proteínas

Os *spots* selecionados foram recortados do gel (somente a área mais intensamente corada) e descorados por 15 min com 5 % (v/v) de acetonitrila (ACN) em 25 mM de bicarbonato de amônio (uma vez) e por 30 min com 50% (v/v) de ACN em 25 mM de bicarbonato de amônio (duas vezes). Em seguida os *spots* foram desidratados com 100% (v/v) de ACN até ficarem totalmente esbranquiçados (cerca de 10 min) e, após a remoção do líquido excedente, os tubos foram deixados 5 min abertos em capela para evaporação dos resíduos da acetonitrila. Os pedaços de géis foram reidratados com 20 mM de DTT em 50 mM de bicarbonato de amônio (50 a 80 μ l ou o suficiente para ficarem imersos) e incubados por 30 min a 65°C. O excesso de líquido foi removido, substituído por 50 μ l de 55 mM de iodoacetamida em 50 mM de bicarbonato de amônio e os tubos foram incubados por 30 min à temperatura ambiente. O líquido excedente foi removido e os géis lavados por no mínimo 1 h com 50 mM de bicarbonato de amônio. Os pedaços de géis foram completamente desidratados com 100 % de ACN e reidratados com uma solução de 10 ng/ μ l de tripsina (Promega V5111) em 25 mM de bicarbonato de amônio, pH 8,0. Após a reidratação total, o excesso da solução de tripsina foi removido e os géis parcialmente cobertos com 25 mM de bicarbonato de amônio. A digestão foi realizada a 37°C por no máximo 14 h.

Após a digestão enzimática o líquido excedente foi transferido para novo tubo. Para a eluição dos peptídeos, os géis foram cobertos com 30-40 μ l de 60 % (v/v) de metanol com 1 % (v/v) de ácido fórmico (AF) e sonificados por 20 min. (40 KHz e 30 Watts) em banho a 45°C (sonicador Ultra Som, Thronton-IPEC Eletrônica Ltda). Após a sonicação o sobrenadante foi transferido para outro tubo. Este processo foi repetido uma vez com a mesma solução anterior e uma vez com 100 % (v/v) de acetonitrila, nesta última etapa a sonicação foi reduzida a 10 min. Todos os sobrenadantes foram unidos em um mesmo tubo e reduzidos a 30 μ l em concentrador a vácuo (Eppendorf modelo 5301) à temperatura ambiente e transferidos para recipiente próprio para injeção no sistema LC-MS/MS.

3.6 Sequenciamento de peptídeos por LC-MS/MS

O sequenciamento das proteínas foi conduzido em plataforma LC/MSMS com espectrômetro de massas híbrido quadrupolo/"time of flight" Q-TOF Ultima API (Waters, Micromass) acoplado a um sistema *on-line* de HPLC capilar CapIC XE Pump (Waters). A amostra injetada (10 µl) foi desalinizada com auxílio de uma pré-coluna *in-line* C₁₈ Sentry Guard (Waters) e os peptídeos capturados por uma coluna de fase reversa Symmetry C₁₈ de 5 µm (0,32 × 150 mm). Os peptídeos resultantes da digestão *in-gel* foram eluídos da segunda coluna com um gradiente de 45 minutos com os tampões A (5 % (v/v) ACN, 95 % (v/v) H₂O e 0,1 % (v/v) de AF) e B (95 % (v/v) ACN, 5 % (v/v) H₂O e 0,1 (v/v) % de AF) conforme a tabela 1. O fluxo da amostra foi conduzido à fonte do espectrômetro a 5 µl/min nos primeiros 15 min; 2 µl/min entre os minutos 15 e 40 (período de maior eluição dos peptídeos); e 5 µl/min nos 5 min finais para lavagem da coluna. A ionização por electrospray foi realizada com a voltagem da fonte ajustada para 3,0 KV e temperatura do cone igual a 90°C, na presença de 5 psi de Nitrogênio. A energia de colisão por gás Argônio foi mantida em 15 psi. Todos os parâmetros foram definidos através do programa MassLynx v 2.1.

Tabela 1 – Gradiente de eluição dos peptídeos para sequenciamento por MS/MS

Concentração do tampão B	Tempo (minutos)	Concentração do tampão B	Tempo (minutos)
10-15 %	5	45-80 %	5
15-30%	20	80 %	5
35-45%	5	10 %	5

Todas as aquisições foram calibradas com GLU¹-fibrinopeptide B (GFP; Sigma cód. F3261), em concentração igual a 100 fmol/µl e injeção automática de 0,6 µl/min com sistema Nanolock Spray (Waters, Micromass).

3.7 Bancos de ESTs de *Eucalyptus* spp. para análise das proteínas

Dois bancos de seqüências do genoma expresso de várias espécies de *Eucalyptus* foram utilizados: um banco para a identificação de proteínas (GENOLYPTUS) e um banco construído por tecnologia SAGE a partir do mesmo tecido de *E. grandis* para comparação com o nível de transcritos das proteínas correspondentes.

As bibliotecas do GENOLYPTUS utilizadas neste projeto foram construídas separadamente a partir de xilema de *E. grandis*, de *E. globulus*, de *E. pellita* e de *E. urophylla*; misturas de floema de várias espécies; misturas de raízes de várias espécies, folhas jovens de *E. grandis*, folhas maduras de *E. grandis*, flores de *E. grandis* e plântulas de *E. grandis* submetidas e estresse e em condições normais, e tecidos de *E. grandis* infectado por *Puccinia* (fungo patogênico).

As bibliotecas SAGE foram construídas em paralelo a este trabalho por outro grupo em nosso laboratório e geraram 70.000 seqüências de 10 pares de bases denominadas genericamente de tags (atualização outubro/2005).

3.8 Processamento dos espectros MS/MS

Os parâmetros para processamentos dos espectros e identificação das proteínas pelo programa ProteinLynx v 2.1 estão reunidos na tabela 2 e nas figura 3 e 4.

Tabela 2 – Modificações pós-traducionais consideradas pela ferramenta AutoMod do programa ProteinLynx v 2.1

Tipo de modificação	Tipo de modificação	Tipo de modificação
Acetyl K Acetyl N-TERM Amidation C-TERM Biotin K Carbamidomethyl C Carbamyl K Carbamyl N-TERM Carboxymethyl C C-Mannosyl W Deamidation N Deamidation Q Dehydration ST Farnesyl C Flavin-adenine dinucleotide CH	Formyl N-TERM Gamma-carboxyglutamic acid E Geranyl-geranyl C Glycation N-TERM Hydroxyl DKNP Lipoyl K Methyl CDEHKNRQ Methyl N-TERM Methyl C-TERM Myristoyl K Myristoyl N-TERM NIPCAM C O-GlcNac ST Oxidation M	Palmitoyl CST Phosphopantetheine S Phosphoryl STY Propionamide C Pyridoxal phosphate K Pyrrolidone carboxylic acid N-TERM S-pyridylethyl C SMA K SMA N-TERM d0 C d8 C

Figura 3 – Parâmetros de processamento dos espectros MS/MS. Programa Global Server v. 2.1

Figura 4 – Parâmetros para identificação de proteínas com o banco SwissProt v 1.0. pelo programa ProteinLynx v.2.1 Os mesmos parâmetros foram utilizados para buscas no banco GENOLYPTUS, trocando-se apenas o item “Databank – SwissProt 1.0” do primeiro quadro por “Databank - GENOLYPTUS”

3.9 Estratégias de análises dos espectros MS/MS

Todos os espectros foram analisados contra dois bancos de dados para identificação das proteínas: o SwissProt (<http://ca.expasy.org>) e o GENOLYPTUS, utilizando respectivamente os algoritmos e tBlastn do NCBI (ALTSCHUL et al., 1997). Da busca com Blastp resultaram proteínas com homologia com sequencias do SwissProt e foi considerada adequado a probabilidade $\geq 95\%$. Da busca por tBlastn resultaram as seqüências de cDNA dos transcritos que foram então contrastada com a base de

proteínas do NCBI através de Blastx. Nesta etapa foi estabelecido o e-value limiar = 10^{-15} , que mostrava boa homologia. Foram validadas as proteínas que puderam ser confirmadas por uma das buscas conforme mostra a figura 5

Peptídeos sem homologia com a base do SwissProt ou do GENOLYPTUS foram avaliados via “*short, nearly exact matches*”, um recurso disponível no NCBI, e validados se confirmassem com 100% de identidade uma proteína previamente encontrada por algum dos caminhos anteriores.

Figura 5 – Fluxograma das estratégias de buscas em bancos de dados para identificação das proteínas

4 RESULTADOS E DISCUSSÃO

4.1 Técnica de coleta de material vegetal

O processo adotado para coleta de tecidos da região cambial demonstrou ser eficiente para obtenção de proteínas em quantidade adequada às técnicas adotadas para sequenciamento e identificação. As vantagens da técnica de raspagem em uma pequena área do tronco da árvore evita a derrubada e permite:

- obter material de maior número de indivíduos, o que torna a amostra mais representativa da população;
- a continuidade do crescimento em campo, podendo a árvore continuar seu ciclo para corte e comercialização;
- repetir a coleta de material dos mesmos indivíduos em outras épocas, podendo ser monitoradas variações de expressão de proteínas (ou genes) durante o desenvolvimento do indivíduo. Com isto são reduzidas as variações devidas ao ambiente e entre indivíduos.

4.2 Eficiência dos protocolos de extração de proteínas

As duas estratégias de extração de proteínas, ácida e fenólica, apresentaram diferenças no rendimento total de proteína por grama de tecido e quanto à pureza da amostra produzida. O rendimento das extrações foi determinado pelo método de Bradford (BRADFORD, 1976) e a pureza avaliada em géis de acrilamida e agarose. O rendimento de proteínas por extração ácida foi 3 a 4 vezes menor (~500 µg de proteína/g de tecido vegetal) comparado com o rendimento do protocolo com fenol (~1800 µg de proteína/g de tecido vegetal). Conforme já relatado na literatura a extração fenólica produz uma amostra mais pura (HURKMAN e TANAKA, 1986). Ácidos nucleicos, fenóis, lipídios e carboidratos que comumente provocam manchas em géis bidimensionais (Saravanan e Rose, 2004) são mais facilmente removidos da amostra, uma vez que permanecem junto à fase solúvel do tampão, enquanto as proteínas ficam

retidas na fase fenólica. A pureza do extrato obtido pelas duas extrações é comparada na figura 6a, onde nota-se que a quantificação pelo método de Bradford da amostra proveniente da extração ácida foi super estimada possivelmente devido à presença de contaminantes. A comparação da eficiência da remoção de DNA e RNA pelos dois tipos de extração pode ser observada na figura 6b, onde 12 µg de proteínas provenientes das duas extrações foram aplicadas em gel de agarose corado com brometo de etídio. Nota-se a maior intensidade do corante na amostra extraída com TCA devido à permanência de ácidos nucleicos contaminantes.

Figura 6 – a) Distribuição de peso molecular do extrato de proteínas da região cambial de *E. grandis* obtida por extração ácida (TCA) e fenólica (PHE); M = Marcador molecular; b) Eficiência de remoção de ácidos nucleicos do extrato de proteínas obtidas por extração ácida (TCA) e fenólica (PHE). M = marcador molecular. As setas apontam o acúmulo de ácidos nucleicos

Portanto as análises foram conduzidas com amostras obtidas por extração fenólica. A figura 6a também mostra o peso molecular das proteínas da região cambial de *E. grandis* separadas em sistema SDS-PAGE. Esta figura mostra que o maior número de proteínas acumula-se em uma faixa entre 30 e 120 KDa.

4.3 Separação das proteínas em géis bidimensionais

Duas faixas de pH foram utilizadas para separação das proteínas totais em géis bidimensionais: pH 3-10 não linear – NL – (figura 7) e pH 4-7 linear (figura 8). Pode ser observado no gel de pH 3-10NL que grande número de *spots* bem corados localizam-se mais próximos ao pH 3 (delimitados pelas linhas verdes na figura 7), e por isto a região de pH 4-7 foi eleita para separação e sequenciamento das proteínas.

Figura 7 – Padrão de distribuição das proteínas de pl 3-10 da região cambial de *E. grandis* aos 2 anos e 8 meses. Dentro do quadrado verde podem ser vistos alguns dos *spots* mais corados

O protocolo para coloração dos géis com o corante Coomassie Colloidal G-250 foi adequado para identificação de grande número de *spots*. Embora maior número de *spots* poderia ser detectado com o uso de corantes mais sensíveis, como os fluorescentes, possivelmente seriam observadas proteínas em concentrações demasiadamente baixas o que dificultaria a etapa de sequenciamento e identificação por espectrometria de massas.

Os géis foram realizados em triplicata e analisados com programa Image Master versão 5.0 (Amersham Biosciences). Após o procedimento automático de detecção de *spots* e correções manuais, um total de 684 *spots* com reprodutibilidade entre géis foi detectado nas faixas de pH 3-10 e 523 *spots* na faixa de 4-7.

4.4 Seleção de *spots* para sequenciamento

Devido às limitações de sensibilidade da técnica 2D-LC/MSMS, os *spots* foram subjetivamente selecionados, baseado na maior intensidade do corante, para aumentar a probabilidade de identificação, e tratados para sequenciamento da proteína. Foi dada preferência a *spots* com boa reprodutibilidade entre os géis sendo escolhidos os *spots* presentes em pelo menos dois géis conforme procedimento de Hunt et al. (2005). Observou-se que, em muitos casos, o tratamento de um único *spot* não permitiu a recuperação de material suficiente para a identificação por espectrometria de massas. Portanto, com a finalidade de aumentar a concentração dos peptídeos recuperados, os *spots* de mesmo número nos 3 géis foram digeridos em um mesmo tubo.

4.5 Apresentação dos resultados de espectrometria de massas

As figuras 9, 10 e 11 mostram os espectros e cromatogramas resultantes das análises por espectrometria de massas. No exemplo são mostrados os dados referentes ao *spot* 1 (*S-adenosylmethionine-synthetase*).

Figura 9 – Cromatograma de eluição dos peptídeos em gradiente de 45 minutos. A figura superior mostra o tempo de eluição de cada peptídeo

Figura 10 – Apresentação dos espectros MS/MS dos peptídeos sequenciados. A figura superior apresenta o tempo de eluição de cada peptídeo e a inferior o espectro de fragmentação íon de $m/z = 735,86$

Figura 11 – Apresentação dos resultados após processamento pelo programa ProteinLynx. À esquerda no canto superior: os peptídeos sugeridos após análise automática dos espectros. Quadro superior direito: proteínas com as quais foi encontrada homologia no banco de dados, e no centro os íons fragmentados de cada peptídeo. Embaixo o espectro com a distribuição das massas do íon de $m/z = 735,868$

4.6 Taxas de identificação de proteínas de *E. grandis*

Os espectros resultantes do sequenciamento por LC MS/MS de 143 *spots* foram analisados contra uma base de proteínas (Swiss-Prot) e contra uma base de ESTs (GENOLYPTUS). Deste total, 112 *spots* (78,3 %) puderam ser identificados resultando em 72 proteínas diferentes (Tabela 3). Foi considerada identificada a proteína com homologia com as seqüências de pelo menos um dos bancos com os escores descritos na metodologia. Não puderam ser identificados 31 *spots* (21,6%) devido à baixa recuperação de peptídeos ou por falta de homologia entre as seqüências experimentais e as dos bancos usados (2 *spots*).

Conforme bem relatado o uso de bancos de nucleotídeos espécie-específicos aumenta o número de identificações de proteínas (Gion et al., 2005; Lippert et al., 2005). Com o material da região cambial de eucalipto observou-se uma melhoria de 29,5 % nas identificações de proteínas, o que representa 33 *spots* que somente puderam ser identificados devido ao uso do banco GENOLYPTUS. Por outro lado apenas 7 *spots* (6,2%) foram identificados exclusivamente com o banco do SwissProt. O uso do GENOLYPTUS aumentou a taxa de identificação por aumentar o tamanho da sequência a ser contrastada, em alguns casos incluindo sítios de domínios conservados. A validação para o Blastx foi baseada no *e-value* limiar = 10^{-15} determinado pelo algoritmo do NCBI. Os *spots* 43 e 74 ficaram abaixo desse limiar, mas por terem sido encontrados 7 e 3 peptídeos respectivamente nos espectros, é mais provável que o baixo score represente sequências específicas de *Eucalyptus* com pouca homologia com as sequências das bases utilizadas.

Misturas de duas ou três proteínas que co-migraram para os mesmos pontos do gel foram encontradas em 13,3 % dos casos (15 *spots*), sendo considerada apenas a proteína dominante. Uma taxa de 15 % de co-migração foi observada em géis de pH 4-7 do proteoma de xilema de *Pinus* (GION et al, 2005) e 14,9% com o proteoma de endosperma de trigo (VENSEI et al, 2005). A figura 8 apresenta o padrão de distribuição das proteínas no gel de pl 4-7 onde são apontados os 143 *spots* escolhidos para análise.

As proteínas identificadas mostraram similaridade com proteínas bem caracterizadas em outras espécies. Para a maior parte das proteínas foi encontrada homologia com sequências de *Arabidopsis thaliana*, possivelmente em função do grande número de sequências disponíveis para esta espécie e à disponibilidade do genoma totalmente seqüenciado (*The Arabidopsis Genome Initiative*, 2000).

Tabela 3 – Proteínas identificadas por MS/MS na região cambial de *E. grandis*

(continua)

SPOT	PROTEINA	Accession nº (SP)	Espécie com maior homologia	Score SwissProt	Score GENOLYPTUS	PM/pl teorico	PM/pl Experim.	Peptídeos / Modificações
1	<i>S adenosylmethionine synthetase</i> EC 2.5.1.6	P43282	<i>Lycopersicon esculentum</i>	100%	Expect = 0.0	42656 5,8	42656 5,8	(K)TNMVMVFGEITTK(G) (R)VHTVLSTQHDVETVTNDEIAADLK(E) (K)EHVIRPVIPEK(Y) (R)FVIGGPHGDAGLTGR(K) (K)SVVASGLAR(R) (K)ENDFRPGMISINLDLK(R) Methyl C-TERM Deamidation N (16) Methyl N-TERM
2	<i>Isoflavone reductase</i> EC 1.3.1.45	P52577	<i>Solanum tuberosum</i>	100%	Expect = 7e-150	33851 6,2	35310 5,7	(K)VLIGGTGYIGK(F) (K)FIVEASAK(A) (R)QSTVSDPVKQQLVESFK(N) (K)QADVISTVGHMQMADQTK(I) (R)FFPSEFGNDVDR(V) (K)AIFNKEDDIAAFTIK(A) (K)ILYIRPPK(N) (K)NVYSFNELVALWEK(K) (K)YLPPEEQILK(Q) (K)YTTVEEYLENFV(I)
3	<i>Actin</i>	Q96292	<i>Arabidopsis thaliana</i>	99,9%	Expect = 0.0	41876 5,4	41876 5,5	(K)GILTLK(Y) (K)YPIEHGIVSNWDDMEK(I) (K)IWHHTFYNELR(V) (R)VAPEEHPVLLTEAPLNPK(A) CDEHKNRQ (5) (R)TTGIVLDSGDGVSHTVPIYEGYALPHAILR(L) (K)SYELPDGQVITIGAER(F) (R)GYMFTTTAER(E) (K)LAYVALDYEQELETAK(S) (K)NYELPDGQVITIGAER(F) (K)EITALAPSSMK(I) Methyl Oxidation M (3)
4	<i>Glutamine synthetase cytosolic isozyme</i> EC 6.3.1.2	P04770	<i>Phaseolus vulgaris</i>	100%	Expect = 8e-56	39118 5,8	39408 5,5	(K)VIAEIWIWGGSGMDLR(S) (R)TLNGPVSDPAK(L) (K)IFSQPDVVAEVPWYIGIEQEYTLQK(N) (R)YILER(I) (R)ITEIADVLSFDPKIQGDWNGAGAHTNYSTK(S) (R)NEGGYEIHK(K) (R)AAVPSGASTGIYEALRL(D) (K)AVGNINSIIGPALVGK(D) (K)IPLYK(H) (K)LAMQEFMILPVGASSFK(E) (K)EGLELLNTAIK(A) (K)VVIGMDVAASEFYGEDK(S) (K)SYDLNFK(E) (K)VNQIGSVTESIEAVR(M) (K)YNQLLR(I) CDEHKNRQ (6) (R)IEEELGAEVYAGANFR(K) MethylN-term
5	<i>Enolase</i> EC 4.2.1.11	P42896	<i>Ricinus communis</i>	100%	Expect = 1e-143	47883 5,7	54098 5,6	(R)AAVPSGASTGIYEALRL(D) (K)AVGNINSIIGPALVGK(D) (K)IPLYK(H) (K)LAMQEFMILPVGASSFK(E) (K)EGLELLNTAIK(A) (K)VVIGMDVAASEFYGEDK(S) (K)SYDLNFK(E) (K)VNQIGSVTESIEAVR(M) (K)YNQLLR(I) CDEHKNRQ (6) (R)IEEELGAEVYAGANFR(K) Methyl C-term
6	<i>Eukaryotic initiation factor 4A 6</i> EC 3.6.1.-	Q40469	<i>Zea mays</i>	100%	Expect = 9e-96	45849 5,6	43483 5,6	(K)LFVLDEADEMLSR(G) (R)VLITTDLLAR(G) (R)GIYAYGFEPKPSAIQQR(G) (K)GLDVIQQAQSGTGK(T) (R)ALGDYLGVK(V) (R)VFDMLR(R) (R)GFKDQIYDIFQLLPSK(I) (K)DQIYDIFQLLPSK(I) (R)ILSSGVHVVGTPGR(V) (R)VAPEEHPVLLTEAPLNPK(A) (R)TTGIVLDSGDGVSHTVPIYEGYALPHAILR(L) (R)GYSFTTTAER(E) (K)LAYIALDYEQELETAK(S) (K)AEYDESGPSIVHR(K) Oxidation M (10)
7	<i>Actin 104 Fragment</i>	P46258	<i>Solanum tuberosum</i>	100%	Expect = 5e-164	39528 5,8	42077 5,6	(K)LFVLDEADEMLSR(G) (R)VLITTDLLAR(G) (R)GIYAYGFEPKPSAIQQR(G) (K)GLDVIQQAQSGTGK(T) (R)ALGDYLGVK(V) (R)VFDMLR(R) (R)GFKDQIYDIFQLLPSK(I) (K)DQIYDIFQLLPSK(I) (R)ILSSGVHVVGTPGR(V) (R)VAPEEHPVLLTEAPLNPK(A) (R)TTGIVLDSGDGVSHTVPIYEGYALPHAILR(L) (R)GYSFTTTAER(E) (K)LAYIALDYEQELETAK(S) (K)AEYDESGPSIVHR(K)

Tabela 3 – Proteínas identificadas por MS/MS na região cambial de *E. grandis*

(continuação)

SPOT	PROTEINA	Accession nº (SP)	Espécie com maior homologia	Score SwissProt	Score GENOLYPTUS	PM/pl teórico	PM/pl Experim.	Peptídeos / Modificações
8	Caffeic acid 3 O methyltransferase EC 2.1.1.68	P46484	<i>Eucalyptus gunii</i>	100%	Expect = 3e-82	39888 5,7	38723 5,6	(K)AAIELDLLEIMAK(A) (R)LLASYSVLTCTLR(N) (K)ILMESWYYLK(D) (K)GFEGLETVDVGGGTGAVLSMIVAK(Y) (K)GINFDLPHVIEDAPPLPGVK(H) (K)NVIHIDCIMLAHNPGGK(E) (K)NCYDALVNIGK(V) Propionamide C (10) Oxidation M (3) Carbamidomethyl C (7) Carbamidomethyl C (2)
9	S adenosylmethionine synthetase EC 2.5.1.6	P43282	<i>Lycopersicon esculentum</i>	100%	Expect = 5e-59	42581 6,4	42861 5,7	(K)TNMVMVFGEITTK(A) (R)FVIGGPHGDAGLTGR(K) (K)IITDYGGWGAHGGGAFSGK(D) (K)SVVASGLAR(R) (K)VLVNIQQSPDIAQGVHGLSK(K) (R)AIGVPEPLSVFVDYGTGK(I) (R)DDADFTWEVVKPLK(W) (R)VHTVLISTQHDETVTNDKIAADLK(E) (K)TIFHLNPSGR(F) Methyl N-term Methyl C-term
10	Tubulin alpha 1 chain	P33629	<i>Prunus amygdalus</i>	100%	Expect = 0.0	49526 4,9	45621 5,2	(R)QLFHPEQLISGK(E) (K)EIVDLCLDR(I) (K)DVNAAVATIK(T) (R)AFVHWYVGEEMEEGEFSEVK(E) (R)EDLAALEK(D) (K)TVGGDDAFNTFFSETGAGK(H) (R)AVFVDLEPTVIDEVR(T) (R)LVSQVISSLTASLR(F) (R)FDGALNVDVTEFQTNLVPYPR(I) (R)IHFMLSSYAPVISAER(A) (R)AVCMISNSTSVAEVFSR(I) Carbamidomethyl C (6)
11	Tubulin beta 1 chain	P12459	<i>Glycine max</i>	100%	Expect = 7e-64	49991 4,9	49382 5,0	(K)LAVNLIPFPR(L) (R)YLTASAMFR(G) (K)NSSYFVEWIPNNVK(S) (R)VSEQFTVMFK(R) (R)AVLMDLEPGTMDSVR(S) (K)EVDEQMINVQNK(N) (K)MASTFIGNSTSIQEMFR(R)
12	Tubulin beta	P20364	<i>Daucus carota</i>	100%	Expect = 0.0	35788 4,8	49978 5,0	(R)INVYYNEASGGR(Y) (K)GHYTEGAELIDAVLDVVR(K) (R)MMLTFSVFPSPK(V) (R)FPGQLNSDLR(K) (K)LAVNLIPFPR(L) (R)JLHFFMVGFAPLTSR(G) (R)YLTASAMFR(G) (K)NSSYFVEWIPNNVK(S) (R)VSEQFTAMFR(R) Methyl C-term Oxidation M (5)
13	Heat shock 70 kDa	P11143	<i>Zea mays</i>	100%	Expect = 0.0	70603 5,2	70383 5,4	(R)VEIANDQGNR(T) (K)MKEIAEAYLGSTIK(N) (K)EIAEAYLGSTIK(N) (K)DAGVIAGLNVMR(I) (K)ATAGDTHLGGEDFDNR(M) (R)MVNHVFVQEFK(R) (R)FEELNMDLFR(K) (K)EQVFSTYSDNQPGVLIQVYEGER(A) (K)NALENYAYNMR(N) (K)FAAENGVLTVTPK(E) Gamma-carboxyglutamic acid E (3) Oxidation M (11)

Tabela 3 – Proteínas identificadas por MS/MS na região cambial de *E. grandis*

(continuação)

SPOT	PROTEINA	Accession nº (SP)	Espécie com maior homologia	Score SwissProt	Score GENOLYPTUS	PM/pl teórico	PM/pl Experim.	Peptídeos / Modificações
15	18.5 kDa class I heat shock protein (HSP 18.5)	P05478	<i>Glycine max</i>	99,9%	Expect = 2e-53	18502 5,8	18465 5,7	(R)DFPFPPSSSLTAANFPQTFR(E) (R)ADLPGLK(K) (K)ASMENGVLTVTPK(V) (R)ILQISGER(S) FESANGVLTVVVVK
16	chloroplast small heat shock protein	Q6WHC0	<i>Capsicum frutescens</i>	---	Expect = 1e-56	18178 6,2	18225 5,7	(K)APGGAPANVAIVTR(L) (R)TALAFVTLR(A) (R)EFMFYR(N) (K)LGDDEFGHMLAGILK(T)
19	Fructokinase EC 2.7.1.4	P37829	<i>Solanum tuberosum</i>	100%	Expect = 2e-64	33764 5,8	35104 5,4	(K)VYGPDPYASPK(R) (K)GVEFETVGVLLK(G) (K)LQPFGLSPVVDGDTLTFESR(A) (K)YKQQGTLLGK(T) (R)GLVEQWLEVAHNFHPIYDLVLK(I) (K)IVFGPFGLTPDPK(Q) (K)VLELWPGLY(I)
20	Glutathione S-transferase EC 2.5.1.18	Q9FQD6	<i>Glycine max</i>	---	Expect = 4e-89	24272 5,6	24340 6,0	(K)VYGPDPYASPK(R) (K)GVEFETVGVLLK(G) (K)LQPFGLSPVVDGDTLTFESR(A) (K)YKQQGTLLGK(T) (R)GLVEQWLEVAHNFHPIYDLVLK(I) (K)IVFGPFGLTPDPK(Q) (K)VLELWPGLY(I)
21	Tubulin alpha	P33624	<i>Anemia phyllitidis</i>	100%	Expect = 2e-113	40176 5,1	45403 5,2	(R)SLDIERPTYTNLNR(L) (R)LISQIISSLTSLR(F) (-)LSVDYGK(-) (-)EDLAALTR(-) (R)AVFLDLEPTVIDEVR(S) (K)DVNAAVATIK(T) (R)LVSQVISSLTASLR(F) (R)SLDIERPTYTNLNR(L)
22	Tubulin alfa	P33623	<i>Anemia phyllitidis</i>	99,9%	---	49786 5,0	48559 5,1	(R)AVFLDLEPTVIDEVR(T) (R)LSVDYGK(K) (R)SLDIERPTYTNLNR(L) (R)LVSQVISSLTASLR(F) (K)DVNAAVATIK(T) (R)IDHKFDLMYAK(R) (R)LVSQIISSLTSLR(F)
23	ATP synthase beta chain mitochondrial precursor EC 3.6.3.14	P37399	<i>Daucus carota</i>	100%	Expect = 6e-68	59135 5,6	50339 5,4	(K)TVLIMELINNVAK(A) (R)VGLTGLTVAEHFR(D) (R)IPSAVGYQPTLATDLGGQLER(I) (R)FLSQPFHVAEIFTGAPGKYVELK(E) CDEHKNRQ (18) (K)VVDLLAPYKR(G) (K)GLFGGAGVGK(T)
25	Putative glycine-rich RNA-binding protein	Q6YNS1	<i>Prunus avium</i>	---	Expect = 5e-38	17327 7,8	15969 5,7	(R)AFSQFGDILDSK(I) (R)GFGVTFGNEK(A)
26	143 3 like protein	P46077	<i>Arabidopsis thaliana</i>	99,9%	Expect = 4e-105	30193 4,8	30193 4,6	(R)NLLSVAYK(N) (K)VFYK(M) (R)YLAEFK(S) (K)DAAESTLTAYK(S) (K)DSTLIMQLLR(D) (K)ICDGVLLKLDTR(L)
27	Phosphoglycerate kinase cytosolic EC 2.7.2.3	Q42962	<i>Nicotiana tabacum</i>	99,9%	Expect = 0.0	42364 5,7	41729 6,3	(R)LSELLGVEVK(M) (K)GVIESLLEK(V) (K)GVSLLLPTDVIADK(F) (K)TFSEALDTTK(T) (K)GVTTIIGGDSVAAVEK(A)

Tabela 3 – Proteínas identificadas por MS/MS na região cambial de *E. grandis*

(continuação)

SPOT	PROTEINA	Accession nº (SP)	Espécie com maior homologia	Score SwissProt	Score GENOLYPTUS	PM/pl teórico	PM/pl Experim.	Peptídeos / Modificações
28	2,3-bisphosphoglycerate-independent phosphoglycerate mutase EC 5.4.2.1	P30792	<i>Zea mays</i>	100%	Expect = 0.0	60619 5,3	62778 5,4	(K)ALEYEDFDK(FDR(V) (R)TSGEYLVHNGVR(T) (K)FGHVTFWNGNR(S) (K)ESFESGTLHLIGLLSDGGVHSR(L) (R)LDQLLLLVK(G) Hydroxyl DKNP (6) Deamidation Q (3)
29	Heat shock protein 90 homolog	O43109	<i>Podospira anserina</i>	99,9%	Expect = 5e-54	79322 4,9	81356 5,0	(R)ELVSNASDALDK(I) (K)ADLVNNGTIAR(S)
30	Heat shock 70 kDa	P11143	<i>Zea mays</i>	100%	Expect = 0.0	70603 5,2	70783 5,2	(R)VEIIANDQGNR(T) (R)FEELNMDLFR(K) (K)NAVVTVPAYFNDQSR(Q) (R)TTPSYVAFTDTER(L) (R)GYMFTTAER(E) (K)EITALAPSSMK(I)
31	Actin	Q96483	<i>Lycopersicon esculentum</i>	27,8	Expect = 2e-104	37174 5,3	41920 5,4	(K)LNNAILAEDK(H) (K)NYSVDYIAGGATQNSIR(V) (K)VLPMYDIVFGNETEAR(T) (K)VHGWETDDVEEIALK(I) (R)ITVITQGPDPVVVAEDGK(V) (K)FPVILLPK(E)
32	Adenosine kinase EC 2.7.1.20	Q9SF85	<i>Arabidopsis thaliana</i>	50,4	Expect = 5e-93	37836 5,3	37653 5,4	(K)LNNAILAEDK(H) (K)NYSVDYIAGGATQNSIR(V) (K)VLPMYDIVFGNETEAR(T) (K)VHGWETDDVEEIALK(I) (R)ITVITQGPDPVVVAEDGK(V) (K)FPVILLPK(E)
33	Enolase	P26300	<i>Lycopersicon esculentum</i>	99,9%	Expect = 0.0	47798 5,7	53196 5,6	(K)LNNAILAEDK(H) (K)NYSVDYIAGGATQNSIR(V) (K)VLPMYDIVFGNETEAR(T) (K)VHGWETDDVEEIALK(I) (R)ITVITQGPDPVVVAEDGK(V) (K)FPVILLPK(E)
34	Tubulin alpha	P29511	<i>Arabidopsis thaliana</i>	100%	Expect = 0.0	49537 4,9	46730 5,3	(K)TVGGDDAFNTFFSETGAGK(H) (R)AVFVDLEPTVIDEVR(T) (K)EIVDLCLDR(I) Carbamidomethyl C (6) (R)ILVSQVISSLTASLR(F) Methyl N-term (R)JHFMILSSYAPVISAER(A) Oxidation M (4) (K)DVNAAVGTIK(T)
35	ATP synthase alpha chain mitochondrial EC 3.6.3.14	P05492	<i>Oenothera biennis</i>	99,9%	---	55596 6,2	55545 6,6	(R)AAELTTLESR(I) (R)VISVGDGIAR(V) (R)JVDALGVPIGR(G) (K)AVDSLVPGR(G)
36	14 3 3 like protein	Q26537	<i>Schistosoma mansoni</i>	49,9%	Expect = 4e-113	24607 5,0	27028 4,9	(R)DQYVYLAK(L) (R)LLDTNLVPSAAASESK(V) (K)VFLK(M) (R)YLAEFK(V)
37	14 3 3 like protein	P42653	<i>Vicia faba</i>	100%	Expect = 1e-62	29420 4,7	29729 4,8	(R)NLLSVAYK(N) (K)VFLK(M) (R)YLAEFK(S) (K)DAAESTLTAYK(S) Methyl CDEHKNRQ (1) (K)DSTLIMQLLR(D) Oxidation M (6) (K)AAQDIANTELAPTHPIR(L)
38	Glyoxalase I homolog EC 4.4.1.5	O65398	<i>Arabidopsis thaliana</i>	---	Expect = 1e-65	31928 5,2	36325 5,5	(K)GGTTVIAFVK(D) (K)DPDGYIFELIQR(G) (K)LVIEELGGK(I) (R)QPGPLPGINTK(I) (K)ITSFLDPDGWK(T)

Tabela 3 – Proteínas identificadas por MS/MS na região cambial de *E. grandis*

(continuação)

SPOT	PROTEINA	Accession nº (SP)	Espécie com maior homologia	Score SwissProt	Score GENOLYPTUS	PM/pl teorico	PM/pl Experim.	Peptídeos / Modificações
39	<i>Isoflavona reductase</i> EC 1.3.1.-	P52578	<i>Solanum tuberosum</i>	50,6%	Expect = 2e-110	33851 6,2	35264 5,6	(K)GKLVESFK(N) (K)FIVEASAK(A) (K)VLIIGGTGYIGK(F)
40	<i>Glutamine synthetase cytosolic isozyme</i> EC 6.3.1.2	P24099	<i>Glycine max</i>	12,6%	Expect = 1e-39	38844 5,4	40270 5,7	(R)HETADINTFLWGVANR(G)
41	<i>Isoflavone reductase</i> EC 1.3.1.45	P52575	<i>Medicago sativa</i>	100%	Expect = 3e-118	35455 5,2	34973 5,7	(K)EDDIGTYTIK(A) (K)VVLGDGNVK(A) (K)AGIVEGFK(S) (K)FIVEASAK(S) (R)IVVAIK(A) (R)FLPSEFGNDVDR(V)
42	<i>UTP glucose 1 phosphate uridylyltransferase</i> EC 2.7.7.9	P19595	<i>Solanum tuberosum</i>	100%	Expect = 1e-78	51742 5,7	53375 6,3	(R)YLSGEAQHVDWSK(I) (K)LDALLAKGK(E) Hydroxyl DKNP (7) (K)GGTLISYEGK(V) (K)MEIIPNPK(E) Oxidation M (1) (K)VQLQLETAAGAAIK(F) (R)SNPSNPSIELGPFEK(K) Methyl CDEHKNRQ (2) (K)SIPSIIIDLDSLK(V) Methyl CDEHKNRQ (7) (K)IFNTNNLWVNLK(A) (K)ATSDLLLVQSDLYTLVDGYVIR(N)
43	<i>Cell division cycle protein 48 homolog</i>	P54774	<i>Glycine max</i>	50,9%	Expect = 3e-10	89769 5,2	89769 5,5	(K)DFSTAILER(K) (R)ELVELPLR(H) (K)GILLYGPPGSGK(T) (R)IVSQLLTMDGLK(S) (R)ELQETVQYPVEHPEK(F) (R)LDQLIYIPLPDEDSR(H) (R)GFGSEFR(F)
44	<i>Actin depolymerizing factor 1</i>	Q39250	<i>Arabidopsis thaliana</i>	33,3%	Expect = 8e-37	16112 6,7	17669 5,7	(K)FLELK(A) (R)FIVYK(I) (R)YAVYDFDMTEENVPK(S) (R)IIIAWSPDTSK(V)
46	<i>Caffeoyl CoA O methyltransferase</i> EC 2.1.1.104	Q43095	<i>Populus tremuloides</i>	26,6%	Expect = 1e-117	27866 5,3	26740 5,1	(K)SLLQSDALYQYILETSVYPR(E) (K)ILAMDINR(E) (R)ENYELGLPVIQK(A) (R)LIELVK(V)
48	<i>putative quinone oxidoreductase</i> EC 1.6.5.5	Q94110	<i>Fragaria ananassa</i>	---	Expect = 8e-108	36194 6,6	34479 5,5	(K)LVDPDVGVPK(E) (K)VVAASLNVPDFK(R) (R)ALGFFK(D) (K)VGDVYGDINEK(A) (K)QHGTAEYTAVER(L) (K)LNPFLESQK(I) (K)VVIHPIA(I)
49	<i>14 3 3 like protein</i>	Q41418	<i>Solanum tuberosum</i>	99,9%	Expect = 2e-129	29394 4,7	30372 5,0	(R)YEEMVEFMEK(V) (R)NLLSVAYK(N) Methyl C-term (K)ICDGILSLLESHLVPSASTAESK(V) (R)YLAEFK(T)
50	<i>Triosephosphate isomerase cytosolic</i> EC 5.3.1.1	P21820	<i>Coptis japonica</i>	100%	Expect = 4e-74	27089 5,5	28212 8,1	(K)FFVGGNWK(C) (K)TAYALSQGLK(V) (K)VIACVGETLEQR(E) Carbamidomethyl C (4)
51	<i>Ascorbate_peroxidase</i>	Q41712	<i>Vigna unguiculata</i>	---	Expect = 1e-100	27033 5,6	30265 7,9	(K)SYPTVSEYK(A) (K)HAAELSHGANSGLDVAVR(L) (K)ALLADPVFRPLVEK(Y) (K)YAADEDAFFDYAEHLK(L)

Tabela 3 – Proteínas identificadas por MS/MS na região cambial de *E. grandis*

(continuação)

SPOT	PROTEINA	Accession nº (SP)	Espécie com maior homologia	Score SwissProt	Score GENOLYPTUS	PM/pl teorico	PM/pl Experim.	Peptídeos / Modificações	
52	reversibly glycosylated polypeptide	Q8VWN8	Gossypium hirsutum	---	Expect = 2e-44	40850 6,0	42149 7,3	(K)ASNPFFVNLK(K)	
53	Putative alpha-soluble NSF attachment protein	Q9M5P8	Solanum tuberosum	99,9%	Expect = 1e-116	32477 5,1	35139 5,3	(K)HEDAADLFDK(A) (R)AIDNFER(A) (K)VAQFAAQLEQYQK(A) (K)AIEIYEEIAR(H) (K)ADIVAITNALER(Y) (K)TTLLLR(V)	
55	Caffeic acid 3 O methyltransferase	P46484	Eucalyptus gunnii	100%	Expect = 4e-99	39914 5,5	39748 5,6	(K)AAIELDLLEIMAK(A) (R)LLASYSVLTCTLR(N) (R)LYGLAPVCK(F) (K)NEDGVSI AALNLMNQDK(I) (K)ILMESWYYLK(D) (K)GINFDLPHVIEDAPPLPGVK(H)	Carbamidomethyl C (10) Carbamidomethyl C (8) Oxidation M (13) Oxidation M (3)
57	14 3 3 protein homolog	Q41246	Nicotiana tabacum	100%	Expect = 3e-119	28666 4,9	31138 4,9	(K)LDVELTVEER(N) (R)YLAEFK(A) (K)AYELASTTAEADLSPTHPIR(L) (K)DSTLIMQLLR(D) (R)VISSIEQK(T)	Oxidation M (6) Methyl N-term
58	Soluble inorganic pyrophosphatase	O22537	Oryza sativa	50,2%	Expect = 1e-57	23610 6,0	29694 5,5	(R)VLYSSVVYPHNYGFIPR(T)	
59	Probable protein disulfide isomerase A6 precursor EC 5.3.4.1	O22263	Arabidopsis thaliana	100%	Expect = 2e-160	39496 5,8	39748 5,6	(K)YGVSGYPTIQWFPK(G) (K)SEGDVVIANLDADKYK(D) (K)YGVSGFP TLK(F) (K)AGIVASLDALVK(E)	
60	Phosphoglycerate kinase cytosolic EC 2.7.2.3	Q42962	Nicotiana tabacum	99,9%	Expect = 0.0	42364 5,7	41629 5,9	(R)VDLNVPLDENFNITDDTR(I) (R)LSELLGVEVK(I) (K)LASLADLYVND AFGTAHR(A) (K)GVTTIIGGDSVA AVEK(V) (K)MSHISTGGGASLELLEK(V) (K)KPFAAIVGGSK(V)	Dehydration ST (14)
61	Cysteine synthase EC 4.2.99.8	Q43726	Arabidopsis thaliana	99,9%	Expect = 5e-152	34605 8,8	34479 5,7	(K)LYGIEPVESAILSGGKPGPHK(I) (K)LIVVIFPSFGER(Y) (R)IGYSMTVD AEQK(G) (K)AFGAELVLTDPK(G) (K)DVTELIGK(T)	Deamidation Q (11)
62	Vacuolar ATP synthase catalytic subunit A EC3.6.3.14	O23654	Arabidopsis thaliana	33,3%	Expect = 2e-25	68812 5,1	70553 5,5	(K)LAADTPLLTGQR(V) (R)EASIYTGITIAEYFR(D) (R)LASFYER(A) (R)EDDLNEIVQLVGK(D) (R)NIIHFYNLANQ AVER(A) (R)JLGDLFYR(L) (K)ITYIAPPGQYSLK(D)	
63	Malate dehydrogenase mitochondrial precursor EC 1.1.1.37	Q9ZP06	Arabidopsis thaliana	20,5%	Expect = 6e-78	35804 8,5	38133 8,9	(R)DDLNFNINAGIVK(N)	
65	cpn60 chaperonin	Q05046	Cucurbita maxima	99,5%	Expect = 7e-103	61130 6,3	60278 5,4	(K)GVEDLADAVK(V) (K)EGVITISDGK(T) (K)TLYNELEVVEGMK(L) (K)VLELALK(R) (K)IGVQIIQNALK(T)	

Tabela 3 – Proteínas identificadas por MS/MS na região cambial de *E. grandis*

(continuação)

SPOT	PROTEINA	Accession nº (SP)	Espécie com maior homologia	Score SwissProt	Score GENOLYPTUS	PM/pl teórico	PM/pl Experim.	Peptídeos / Modificações
66	<i>Eukaryotic initiation factor 4A</i> EC 3.6.1.-	Q40469	<i>Nicotiana tabacum</i>	9,5%	Expect = 4e-36	29367 6,5	43374 5,4	(R)GIYAYGFEKPSAIQQR(G) (R)VLITDILLAR(G)
67	<i>Caffeoyl CoA O methyltransferase</i> EC 2.1.1.104	Q43237	<i>Vitis vinifera</i>	100%	Expect = 5e-120	27234 5,3	28414 5,3	(K)ENFELGLPVIQK(A) (R)LIDLK(V) (R)DFVLELNK(A) (K)ILAMDINR(E)
68	<i>Late embryogenesis like protein</i>	O81366	<i>Prunus armeniaca</i>	---	Expect = 1e-113	28955 4,7	36933 4,8	(R)AEIVVDILK(N) (K)LISGLIPDAGTIHAHGEETVK(I) (K)IPVVLIIYDDIK(N) (K)FSFEETVAVLHLK(L) (K)ILIPITFRPK(D)
70	<i>Caffeoyl CoA O methyltransferase</i> EC 2.1.1.104	Q43237	<i>Vitis vinifera</i>	100%	Expect = 3e-139	27234 5,3	28584 5,4	(K)ENFELGLPVIQK(A) (R)LIDLK(V) (R)DFVLELNK(A)
71	<i>thioredoxin peroxidase</i>	Q8SAG2	<i>Capsicum annuum</i>	---	Expect = 2e-77	17406 6,1	19575 5,7	(K)YTHALGLELDLSEK(G)
72	<i>14 3 3 like protein</i>	Q41246	<i>Nicotiana tabacum</i>	100%	Expect = 1e-116	28666 4,9	32766 5,0	(K)LDVELTVEER(N) (R)NVLSVGYK(N) (K)HLIPSSSSGSESTVFYK(M) (R)YLAEFK(A) (K)AYEASSTASTDLAPTHPIR(L)
73	<i>Elongation factor 2</i>	O23755	<i>Beta vulgaris</i>	99,9%	Expect = 0.0	93798 5,9	94379 7,2	(K)FTADELR(A) (R)LWGENFFDPATK(K) (R)TVIWMGK(K) (K)FSVSPVVR(V)
74	<i>Putative TPR-repeat protein</i>	Q84TJ2	<i>Arabidopsis thaliana</i>	---	Expect = 3e-07	64519 5,8	74766 5,8	(R)LGAAYQGLGR(Y) (R)SLAVSIPR(L) (K)ALETYQEGLK(H)
75	<i>Heat shock 70 kDa homolog protein</i>	P08418	<i>Schistosoma mansoni</i>	51,6%	Expect = 1e-104	69875 5,4	73298 5,4	(K)VEIIANDQGNR(T) (K)VQKLLQDFFNGK(E) (K)FELSGIPPAPR(G)
77	<i>Ubiquitin-conjugating enzyme 2</i>	Q45W76	<i>Arachis hypogaea</i>	---	Expect = 1e-06	17823 4,9	17627 7,3	(R)LLEELER(G)
78	<i>Universal stress protein</i>	Q2TN84	<i>Triticum aestivum</i>	---	Expect = 2e-45	17775 6,4	17775 6,2	(K)WAVDNVAR(N) (K)VDPETIDIVSTIAR(Q)
79	<i>putative chalcone isomerase 4</i> EC 5.5.1.6	Q53B72	<i>Glycine max</i>	---	Expect = 5e-66	23495 4,9	26677 5,4	(K)FTAIGVYLDPEIVGHLK(Q) (K)GKPGNVLAEDDDFFDAVASAPVEK(F) (K)GSQYGGLEGAVR(D) (R)LAAEDKYEEEEAALEK(V) (K)VVEFFQSK(Y) (K)WYLGGR(G) (R)GVSETTITSLANTLSEK(L)
81	<i>6-phosphogluconolactonase</i> EC 3.1.1.31	Q84WW2	<i>Arabidopsis thaliana</i>	---	Expect = 8e-47	35644 7,6	29800 5,4	(K)SEWVTFTDSPKPPPER(I)
82	<i>aluminum-induced protein</i>	Q5MCR8	<i>Codonopsis lanceolata</i>	---	Expect = 1e-113	25274 5,4	25527 5,2	(K)SANEVVLVIEAYK(A) (K)STSTLFVASDQYK(V) (K)ITAVPANEEIWGATFK(V) (K)VEGPAVLAATE(I)

Methyl C-term

Tabela 3 – Proteínas identificadas por MS/MS na região cambial de *E. grandis*

(continuação)

SPOT	PROTEINA	Accesion nº (SP)	Espécie com maior homologia	Score SwissProt	Score GENOLYPTUS	PM/pl teórico	PM/pl Experim.	Peptídeos / Modificações
84	14 3 3 protein homolog	Q9SP07	<i>Lilium longiflorum</i>	100%	Expect = 3e-130	29252 4,8	29800 4,9	(K)TVDVEELTVEER(N) (R)NLLSVAYK(N) (R)VISSIEQK(T) (K)VIFYLK(M) (R)YLAEFK(T) (K)SAQDIALAELAPTHPIR(L) (K)DSTLIMQLLR(D) (R)YEEMVEFMEK(V) Methyl N-term Oxidation M (6)
86	60S acidic ribosomal protein	O04204	<i>Arabidopsis thaliana</i>	26%	Expect = 4e-134	33666 5,2	37773 5,3	(K)GTVEIITPVELIK(K) (K)VGSSEAALLAK(L) (K)NVLAVAVETESYSPHADEVK(E)
87	Eukaryotic initiation factor 4A 13	P41382	<i>Nicotiana tabacum</i>	39,8%	Expect = 1e-121	46842 5,4	44847 5,6	(K)GLDVIQQAQSGTGK(T) (R)ALGDYLGVK(V) (R)ILQSGVHVVGTPGR(V) (R)VLITTDLLAR(G) (K)DQIYDIFQLPSK(I)
88	flavanone 3-hydroxylase EC 1.14.11.9	Q6BEH2	<i>Eustoma grandiflorum</i>	---	Expect = 9e-110	40757 5,3	41919 5,7	(K)VVVNYYPK(C)
89	Vacuolar ATP synthase subunit C EC 3.6.3.14	Q9SDS7	<i>Arabidopsis thaliana</i>	99,9%	Expect = 6e-90	42619 5,4	40992 5,6	(R)JFAESILR(H) (R)FNIPNLR(V) (K)FWWDDAK(Y) (K)YPAMAPLR(E) (R)ETVDTIQGVAK(I)
90	Phosphoglycerate kinase cytosolic EC 2.7.2.3	P12783	<i>Triticum aestivum</i>	100%	Expect = 2e-122	42121 5,6	41826 5,8	(K)ELDYLVGAVANPK(K) (K)KPFAIVGGSK(V) (R)VDLNVPLDENFNITDDTR(I) (R)AAVPTIK(Y) (R)JSELLGVEVK(I)
91	UDP-glucose pyrophosphorylase EC 2.7.7.9	Q6Y643	<i>Bambusa oldhamii</i>	---	Expect = 1e-119	51992 5,3	53630 5,7	(K)LEIPDGVVLENK(E)
93	Actin Fragment	P30161	<i>Costaria costata</i>	99,9%	---	37223 5,5	42409 5,4	(R)GVLTLLK(Y) (R)VAPEEHPVLLTEAPLNPK(A) (R)JAGFAGDDVPK(C) (R)GYMFTTAER(E) (K)NYELPDGQVITIGAER(F) Methyl N-term
94	Putative 26S protease regulatory subunit homolog A	O23894	<i>Brassica campestris</i>	35%	Expect = 1e-177	47492 4,9	47715 5,5	(K)GVLLYGPPTGK(T) (R)VVGSEFVQK(Y)
95	C2 domain-containing protein	Q5DVL6	<i>Hordeum vulgare</i>	---	Expect = 1e-46	32667 5,6	42496 5,0	(R)DVLDDAGFGEALK(R) (R)APDPYHASPYPVPATR(E)
96	Sgt1	Q9SE32	<i>Oryza sativa</i>	---	Expect = 5e-12	40952 5,0	42496 5,2	(K)GVAPEDAHVEFGQILSVK(I) (K)INVPSDAYYFQPR(L) (R)YQILSTK(I)
97	Sgt1	Q704T2	<i>Brassica oleracea</i>	---	Expect = 2e-75	39407 5,1	42510 5,3	(R)YQILSTK(I)
98	Heat shock 70 kDa protein mitochondrial precursor	Q01899	<i>Phaseolus vulgaris</i>	99,9%	Expect = 1e-129	72537 5,9	76330 5,4	(K)ETAAYLGK(S) (R)NSADTTIYSIEK(S) (K)ETAAYLGK(S) (K)VQEVSEIFGK(S) (K)ALIDIR(N) (K)AVITVPAYFDNAQR(E)

Tabela 3 – Proteínas identificadas por MS/MS na região cambial de *E. grandis*

(continuação)

SPOT	PROTEINA	Accesion n° (SP)	Espécie com maior homologia	Score SwissProt	Score GENOLYPTUS	PM/pl teórico	PM/pl Experim.	Peptídeos / Modificações
99	<i>GDP dissociation inhibitor</i>	O22402	<i>Nicotiana tabacum</i>	---	Expect = 8e-119	49702 5,4	61147 5,6	(R)VLHTDVT(K)(Y) (K)YLYFK(A) (K)AVDGGYVFNK(G) (K)LYAESLAR(F)
101	<i>Triosephosphate isomerase chloroplast precursor EC 5.3.1.1</i>	Q9SKP6	<i>Arabidopsis thaliana</i>	99,9%	Expect = 2e-59	33345 7,7	27417 5,7	(K)GGAFTGEISVEQLK(D) (K)GPEFATVNSVTSK(K) (K)FFVGGNWK(C) (-)KNLASQYKEPR(-)
102	<i>6-phosphogluconate dehydrogenase EC 1.1.1.44</i>	Q8LG34	<i>Arabidopsis thaliana</i>	---	Expect = 1e-123	53246 5,6	51254 7,3	(K)GFPISVYNR(T) (K)YIEDILVK(V) (K)GWDLTTLGELAR(I) (R)VPANLVQAQR(D)
103	<i>20S proteasome alpha subunit</i>	Q9LSU2	<i>Oryza sativa</i>	100%	Expect = 4e-93	25844 5,4	25075 5,6	(K)KLPSILVDETSVQK(I) (K)EGFEGQISGK(N) (K)NIEIGITDK(K) (R)VLTPEIDYLAEEV(I)
104	<i>putative mitochondrial ATP synthase</i>	Q6L460	<i>Solanum demissum</i>	---	Expect = 4e-80	19807 5,3	18378 5,3	(R)AFSEVNSQLETK(F) (K)EAYDSIEIPK(Y) (K)EIADVQELK(K)
105	<i>Triosephosphate isomerase cytosolic EC 5.3.1.1</i>	P21820	<i>Coptis japonica</i>	100%	Expect = 3e-103	27089 5,5	26868 6,7	(K)TAYALSQGLK(V) (R)EAGATLEVVAQTK(A) (K)WLHANTSPEVAATTR(I)
106	<i>S adenosylmethionine synthetase 3 EC 2.5.1.6</i>	P43282	<i>Lycopersicon esculentum</i>	99,9%	Expect = 6e-166	42652 5,8	44101 6,1	(R)FVIGGPHGDAGLTGR(K) (K)SVVASGLAR(R) Methyl N-term (R)VHTVLSTQHDVETNDEIAADLK(E) (K)TIFHLNPSGR(F)
107	<i>D 3 phosphoglycerate dehydrogenase chloroplast precursor EC 1.1.1.95</i>	O04130	<i>Arabidopsis thaliana</i>	100%	---	66454 5,8	63072 5,6	(K)YVGVS LVGK(T)
108	<i>S adenosylmethionine synthetase 3 EC 2.5.1.6</i>	P31155	<i>Petroselinum crispum</i>	100%	Expect = 4e-98	25652 7,9	43581 5,6	(K)TIFHLNPSGR(F) (R)FVIGGPHGDAGLTGR(K) (K)SVVASGLAR(R) Methyl N-term
109	<i>Alcohol dehydrogenase EC 1.1.1.1</i>	P14673	<i>Solanum tuberosum</i>	16,7%	Expect = 2e-57	41094 5,9	42308 6,7	(K)FGVTEFVNPK(D)
112	<i>Leucoanthocyanidin dioxygenase EC 1.14.11.19</i>	Q8LP73	<i>Vitis labrusca x Vitis vinifera</i>	---	Expect = 3e-67	40193 5,6	42293 7,9	(K)AGEVFFNLPIEEK(E)
113	<i>Isoflavone reductase EC 1.3.1.45</i>	O65002	<i>Betula verrucosa</i>	---	Expect = 1e-53	33153 7,8	36110 6,9	(K)VLVIGGTGYIGK(F) (K)SLGVTLVHGDLYDQESLLNAIK(Q)
114	<i>putative NADPH-dependent mannose 6-phosphate reductase EC 1.1.1.224</i>	Q9SJV2	<i>Arabidopsis thaliana</i>	---	Expect = 2e-131	35010 6,2	34032 6,9	(K)LQLDYLDLYLHFPVATK(H) (R)SIGISNYDIFLTR(D) (R)TVAAQVLR(W)
115	<i>20S proteasome alpha subunit A</i>	O48551	<i>Glycine max</i>	100%	---	27392 5,8	27450 6,4	(R)HITIFSPEGR(L) (K)AAGISIGVR(G) (K)LLDNSTVTHLFPITK(Y) Methyl CDEHKNRQ (4) (K)DQEAINFLEK(K) Methyl CDEHKNRQ (1) (R)VLTDEIDEHLTAISERD(-) (K)FLGLLATGMTADSR(S)

Tabela 3 – Proteínas identificadas por MS/MS na região cambial de *E. grandis*

(conclusão)

SPOT	PROTEINA	Accession nº (SP)	Espécie com maior homologia	Score SwissProt	Score GENOLYPTUS	PM/pl teórico	PM/pl Experim.	Peptídeos / Modificações
116	20S proteasome beta subunit	O82531	<i>Petunia hybrida</i>	---	Expect = 2e-112	24617 6,3	27028 8,3	(K)TVFVSATER(D)
117	Superoxide dismutase [Mn], mitochondrial precursor EC 1.15.1.1	Q9SM64	<i>Prunus persica</i>	---	Expect = 1e-108	25455 8,6	25075 7,3	(K)LVVETTSNQDPLVTK(G)
118	20S proteasome alpha subunit G-1	O23715	<i>Arabidopsis thaliana</i>	50%	Expect = 2e-52	27377 5,9	27978 7,0	(R)VFQIEYAAK(A)
120	F1-ATP synthase delta subunit	Q9SXX1	<i>Ipomoea batatas</i>	---	Expect=1e-62	27145 9,6	25285 6,7	(K)ANTLDKVESEFVDLVEASER(S)
125	T complex protein 1 epsilon subunit	P40412	<i>Avena sativa</i>	33,3%	---	58899 5,5	20554 5,3	(K)QQQILLATQVVK(M)
126	Elongation factor 2	P28996	<i>Chlorella kessleri</i>	95,4%	---	94115 5,8	59991 5,5	(K)FSVSPVVR(V) (K)ILSEEFQWQDLAK(K)
131	Isoflavone reductase like protein EC 1.3.1.45	Q3KN68	<i>Vitis vinifera</i>	---	Expect = 1e-145	33885 5,8	31659 5,6	(K)VLVIGGTGYIGK(F) (R)VYVPEEQVLK(N)
133	Late embryogenesis abundant protein	O80576	<i>Arabidopsis thaliana</i>	---	Expect = 7e-123	36036 4,7	36994 4,7	(K)VKDFIQDIGEK(I) (R)AEIVVDILIK(N) (K)IPVVLIIYDDVK(N) (K)FSFEETVAVLHLK(L) (K)LENKDNDFELGLK(D) (K)ILIPITFRPK(D)
135	20S proteasome alpha subunit B	O23708	<i>Arabidopsis thaliana</i>	100%	Expect = 3e-95	25701 5,5	28720 6,0	(K)LVQIEHALTAVGSGQTSLGK(A) (R)LYKEPIPTQLVR(E) (K)EPIPTQLVR(E) (R)VLTPAEIDDYLAEEV(-)
136	Vacuolar ATP synthase subunit B EC 3.6.3.14	P11574	<i>Arabidopsis thaliana</i>	100%	Expect = 2e-28	54107 5,0	52864 5,3	(R)TVSGVAGPLVLDK(V) (K)YQEIYNIR(L) (K)AVVQVFEGTSGIDNK(F) (K)FTTVQFTGEVLK(T) (K)TPVSLDMLGR(I) (R)QIYPPINVLPSSLR(L) (K)SAIGEGMTR(K)
138	Pyruvate dehydrogenase E1 component beta subunit EC1.2.4.1	P52904	<i>Pisum sativum</i>	100%	Expect = 3e-179	38793 5,9	37152 5,5	(R)TIFQSLDLAWTLR(I) (K)TLDQYYSR(D) (K)SNYMSAGQISVPIVFR(G)
140	Alpha soluble NSF attachment protein	Q9M5P8	<i>Solanum tuberosum</i>	100%	Expect = 1e-116	32477 5,1	33078 5,4	(K)LNSWGLFGSK(H) (K)HEDAADLFDK(A) (K)DIGELYESDQNLER(A) (K)AIEIYEEIAR(H) (K)ADIVAITNALER(Y) (R)YQELDPTFSGTR(E) (K)JTLLLR(V)
141	Profilin	Q64LH2	<i>Ambrosia artemisiifolia</i>	---	Expect = 4e-65	14083 4,7	93930	(R)LGDYLVQGF(I)
142	Tubulin beta	P33632	<i>Anemia phyllitidis</i>	4,3%	Expect = 2e-69	27438 5,0	48984 5,1	(K)LAVNLIPFPR(L)
143	Ascorbate peroxidase	Q42459	<i>Spinacia oleracea</i>	---	Expect = 1e-17	27625 5,4	26550 5,5	(K)ALLTDPVFRPLVEK(Y)

4.7 Perfil do proteoma da madeira

4.7.1 Avaliação qualitativa

As proteínas identificadas foram classificadas em categorias funcionais baseado no trabalho de Rison et al. (2000) e a distribuição está apresentada na figura 12. Pode ser observado que grande parte das proteínas seqüenciadas da região cambial de *E. grandis* participam em processos celulares como resposta a estresse, em metabolismo de energia como metabolismo de glicose e transporte de elétrons, na formação de componentes estruturais de membrana plasmática, parede celular e citoesqueleto. O metabolismo macromolecular inclui proteínas relacionadas com síntese e degradação de DNA, RNA e proteínas. O metabolismo de aminoácidos, carboidratos, nucleotídeos, lipídeos e o metabolismo secundário está representado no item “Metabolismo”. Conforme a literatura, proteínas hidrofóbicas, de baixo peso molecular e baixa concentração são fracamente representadas em géis 2D (SKYLAS et al. 2005). Estes tipos de proteínas da região cambial de *E. grandis* tiveram pequena representação neste primeiro trabalho de sequenciamento do proteoma de xilema jovem de eucalipto. Para visualizá-las será necessário desenvolver protocolos específico e estratégias de enriquecimento (SANTONI et al., 2000). Foram identificadas proteínas predominantemente citoplasmáticas, mas também de mitocôndrias (ATP sintase mitocondrial nos spot 23, 35, 104, 120) e cloroplasto (*chloroplast small heat shock protein* e *D 3 phosphoglycerate dehydrogenase chloroplast precursor*, nos spots 16 e 107 respectivamente).

Algumas modificações pós-traducionais foram detectadas automaticamente pelo programa ProteinLynx e as mais comuns foram: oxidação de resíduos de metionina, metilação, deamidação, acrilamidação de resíduos de cisteína (*propionamide C*) e alquilação por iodoacetamida (*Carbamidomethyl C*). Oxidações e acrilamidações estão entre as modificações mais comumente encontradas em proteínas sequenciadas por MS (SCHEVCHENKO et al., 2001). Deve-se salientar que acrilamidações e alquilações são decorrentes da manipulação da amostra e não de modificações que ocorrem na célula. Já oxidações, metilações e deamidações podem ser originadas no tecido da planta e deverão ser confirmadas experimentalmente.

Figura 12 – Distribuição em categorias funcionais de 112 proteínas identificadas da região cambial de *E. grandis* a partir de árvores de 2 anos e 8 meses

4.7.2 Nível de expressão de proteínas

Os géis 2D são comumente realizados para comparações qualitativas entre tratamentos, onde diferenças na expressão de proteínas (presença e ausência) permitem relacioná-las a algum processo metabólico (ANDERSON E ANDERSON, 1998). A quantificação de proteínas por géis também pode ser efetuada, porém as estimativas deixam a desejar quanto à precisão dos resultados devido em parte à diferente afinidade das proteínas pelos corantes (NEUHOFF et al., 1990). Para estimativas de concentração mais precisas outras técnicas são preferidas como ICAT (GYGI et al., 1999b) que utiliza isótopos marcados.

Para estabelecer uma quantificação relativa das proteínas identificadas em relação ao total de proteínas de pI entre 4 e 7 da região cambial de *E. grandis*, foi comparado o percentual do volume médio dos *spots* com repetitibilidade nos três géis dessa faixa de pH. A determinação dos volumes foi obtida por análise automática com o programa Image Master v. 5.0 (Amersham Biosciences) e permitiu obter uma visão global da participação de cada proteína nesse tecido. Os resultados mostram que, entre os *spots*

identificados, as 10 proteínas mais intensamente expressas foram *S-adenosylmethionine synthetase*, actina, tubulina alfa, isoflavona redutase, *heat shock protein 70*, tubulina beta, *14-3-3*, *enolase*, *glutamine synthetase* e *Caffeic acid 3 O methyltransferase* (figura 13). As proteínas mais intensamente expressas na região cambial de *Eucalyptus* estão em acordo com observações feitas em xilema de outras espécies (VANDER MIJINSBRUGGE et al., 2000; Gion et al, 2005). Várias destas proteínas foram representadas por mais de um *spot* e, nestes casos, os volumes foram somados devido à impossibilidade de verificar se representam proteínas com diferentes funções.

Figura 13 – Proteínas mais intensamente expressas na região cambial de *E. grandis* entre os 112 *spots* identificados. Foi considerada a soma de todas as prováveis isoformas

4.8 Proteínas representadas por vários *spots*

As proteínas representadas por mais de um *spot* no gel localizaram-se comumente em regiões adjacentes, e vários padrões de distribuição foram observados (figura 14). Embora não esteja claro o que estas repetições representam, na literatura especula-se que possam ser devidas a:

- variações alélicas de um mesmo gene (isoenzimas),

- modificações pós-traducionais de um mesmo gene (KHOUDOLI et al., 2004),
- a produtos de genes parálogos,
- a eventos provenientes de splicing alternativo
- a uma fração de uma proteína degradada ou modificada devido à manipulação e preparo da amostra (HERBERT et al., 2001).

A figura 12 mostra que mais de uma das alternativas acima podem estar acontecendo, uma vez que observam-se variações de PM e/ou pl entre os *spots*. As proteínas encontradas em maior número de regiões no gel estão representadas na tabela 4.

Figura 14 – Diferentes padrões de distribuição de *spots* representando uma mesma proteína no gel de pH 4-7; a) isoflavona redutase; b) tubulinas α e β ; c) 14-3-3

Tabela 4 – Proteínas da região cambial de *E. grandis* representadas por maior número de *spots* no gel de pH 4-7

Proteína	Spot	Proteína	Spot
1433	26, 36, 37, 49, 57, 72, 84	HSP70	13, 30, 75, 98
Actin	3, 7, 31, 93	Isoflavona redutase	2, 39, 41, 113, 131
Tubulina α	10, 21, 22, 34	CCoAMT	46, 67, 70
Tubulina β	11, 12, 142	s-adenosilmetionina sintetase	1, 9, 106, 108

4.9 Sucesso de identificação de proteínas

A taxa de sucesso de identificação das proteínas foi determinada por 3 fatores:

- 1) concentração proteína/*spot*;
- 2) estrutura primária da proteína;

3) tempo de análise por MS

A influência da concentração das proteínas na taxa de identificação é bem relatada na literatura (BARNES et al., 2004) e torna a técnica 2D-MSMS *per se* limitante para a realização do proteoma completo de um tecido. Partindo de 523 proteínas teóricas (*spots*) detectáveis por análise de imagem nos géis de pH 4-7, apenas 112 (21,4%) puderam ser identificadas. Para melhorar esta eficiência uma alternativa é o pré-fracionamento da amostra antes da separação eletroforética para evitar a saturação do gel com as proteínas mais abundantes (DREGGER, 2003). Entretanto alguns *spots* bastante corados, aparentemente com alta concentração de proteínas (ex: 14, 17, 24 e 132) também não resultaram em qualquer identificação. Nestes casos a não identificação pode ser devida à:

- presença de contaminantes que dificultem a acessibilidade da tripsina durante a digestão.
- influência da sequência da proteína dificultando a digestão enzimática e/ou a eluição dos peptídeos (p.ex. efeito prolina).

Moléculas contaminantes podem ter origem no próprio tecido vegetal, como polissacarídeos e lipídeos, que podem formar complexos com proteínas e além de diminuir a solubilidade da proteína alteram seu PM e pl. Também comuns em tecidos de plantas são os compostos fenólicos e podem modificar as proteínas por reações oxidativas (SARAVAN E ROSE, 2004). Embora o Coomassie G-250 seja muito eficiente para ligação com proteínas (CANDIANO et al., 2004), a hipótese de contaminantes deve ser considerada. Também é bem relatada a ligação irreversível das proteínas por ligações cruzadas com a malha da acrilamida (GALVANI et al., 2000). Para reduzir esta possibilidade todos os géis foram polimerizados por 24 h antes de utilizados para a segunda dimensão eletroforética. Entretanto pequenas regiões do gel que demoram mais a polimerizar podem reter os peptídeos durante a eluição.

Foram avaliados os peptídeos seqüenciados de algumas das proteínas representadas por vários *spots* no gel para verificar a possibilidade de uma relação entre a estrutura primária da proteína com a probabilidade de identificação. Proteínas homólogas à 14-3-3 foram encontradas em 7 *spots* diferentes (26, 36, 37, 49, 57, 72 e 84) e a S-adenosilmetionina sintetase em 4. A tabela 5 aponta os peptídeos eluídos destas duas

proteínas, bem como a posição destes peptídeos dentro da sequência total da proteína com maior homologia extraída do banco SwissProt. Observa-se que certas seqüências foram identificadas com maior freqüência nos vários *spots* eluídos separadamente. É bem observado que embora os sítios de acesso da tripsina sejam as regiões C-terminal de lisina e arginina, a presença de prolina na região carboxílica destes dois resíduos praticamente inibem a digestão (SNYDER, 2000), bem como a presença de resíduos ácidos em um dos lados da ponte potencialmente clivável (WILKINSON, 1986). Ao menos em alguns casos a sequência dos resíduos pode estar sendo um obstáculo à identificação das proteínas (SHEVCHENKO et al., 2000), como no caso de proteínas hidrofóbicas ou muito básicas.

Já foi observado que quanto menor a taxa de fluxo da amostra conduzida ao electrospray melhor a intensidade do sinal dos íons (Schlosser, 2004). A velocidade de eluição dos peptídeos foi reduzido de 5 para 2 $\mu\text{l}/\text{min}$ durante o período de maior eluição dos peptídeos (minutos 15 a 40). Esta redução aumentou o número de peptídeos identificados pelo espectrômetro. Os *spots* 15 a 22 não foram identificados pelo fluxo fixo a 5 $\mu\text{l}/\text{min}$, e os *spots* 1, 4, 5, 6, 7, 8, 10 e 12 todos muito intensos, foram identificados com maior número de peptídeos após redução do fluxo (tabela 6).

4.10 Relação entre proteínas e nível de transcritos

Algumas observações puderam ser feitas ao comparar o perfil de expressão das proteínas da região cambial e o dos transcritos seqüenciados a partir do mesmo tecido pela tecnologia SAGE (tabela 7). Não foi possível estabelecer uma correlação entre as expressões de RNA e da proteína correspondente, o que já foi bem observado em outros organismos (GYGI et al., 1999, CHEN et al., 2002). É interessante notar na tabela 7 que, enquanto algumas proteínas foram intensamente expressas, nenhuma tag correspondente foi encontrada por SAGE no mesmo tecido, portanto parece estar havendo uma regulação diferencial da expressão de RNA e proteínas.

Tabela 5 – Peptídeos eluídos de proteínas representadas por vários spots. a) 14_33 e b) *S-adenosilmetionina sintetase*. Cores semelhantes representam peptídeos com alta ou total similaridade. As seqüências sob as tabelas pertencem à isoforma com maior homologia, e foram extraídas do banco do SwissProt. Em vermelho são indicados outros pontos possíveis para clivagem pela tripsina

a) 14332_ARATH 14-3-3-like; protein GF14 omega (Q01525)

26	36	37	49	57	72	84
(K)DAAESTLTAYK(S)	(R)LLDTNLVPSAAAS ESK(V)	(K)DAAESTLTAYK(S)	(K)ICDGILSLLESHLV PSASTAESK(V)		K)HLIPSSSSGESTVFYY K(M)	(K)EAAESTLLAYK(S)
(K)ICDGVLLDTR(L)		(K)AAQDIANTELAPT HPIR(L)			(K)AYEAASSTASTD THPIR(L)	(K)SAQDIALAELAPTHPIR
(K)VFYLK(M)	(K)VFYLK(M)	(K)VFYLK(M)	(R)YEEMVEFMEK(V)			(K)VFYLK(M)
(R)NLLSVAYK(N)	(R)NLLSVAYK(N)	(R)NLLSVAYK(N)	(R)NLLSVAYK(N)		(R)NVLGVYK(N)	(R)NLLSVAYK(N)
(R)YLAEFK(T)		(R)YLAEFK(S)	(R)YLAEFK(T)	(K)LDVELTVEER(N)	(K)LDVELTVEER(N)	(K)TVDVELTVEER(N)
(K)DSTLMQLLR(D)		(K)DSTLMQLLR(D)		(R)YLAEFK(S)	(R)YLAEFK(S)	(R)YLAEFK(S)
				(K)DSTLMQLLR(D)	(K)DSTLMQLLR(D)	(K)DSTLMQLLR(D)
				(R)VISSIEQK(T)	(R)IISSIEQK(E)	(R)IISSVEQK(E)

MASGEEFVYMAKLAEQAEYEEMVEFMEKVSAAVDGDELTVERNLLSVAYKNVIGARRASWRIISSIEQKEESRGNDHVTAREYRSKIETELSGICDGILKLL
DSRLIPAAASGDSKVYFLKMKGDYHRYLAEFKTGGERKDAAEHTLAAYKSAQDIANAELAPTHPIRLGLALNFSVFYYEILNSPDRACNLAKQAFDEAIAELDTLGEE
SYKDSTLMQLLRDNLTLWTSDMQDDAADEIKAAAAPPTTEEQQ

b) *S-adenosilmetionina sintetase*

1	9	108	106
(K)TNMVMVFGEITTK(G)	(K)TNMVMVFGEITTK(A)		
(K)SVVASGLAR(R)	(K)SVVASGLAR(R)	(K)SVVASGLAR(R)	(K)SVVASGLAR(R)
(K)ENDFRPGMLMSINLDLK(R)	(K)TIFHLNPSGR(F)	(K)TIFHLNPSGR(F)	(K)TIFHLNPSGR(F)
(K)EHVIRPVIPEK(Y)			
(R)FVIGGPHGDAGLTGR(K)	(R)FVIGGPHGDAGLTGR(K)	(R)FVIGGPHGDAGLTGR(K)	(R)FVIGGPHGDAGLTGR(K)
(R)VHTVLISTQHDETVTNDKIAADLK(E)	(R)VHTVLISTQHDETVTNDKIAADLK(E)		(R)VHTVLISTQHDETVTNDKIAADLK(E)

METFLFTSESVDNEGHPDLCDQVSDAILDACLEQDPESKAVACETCTKTNMVMVFGEITTKATVDYEKIVEDTCRGIGFVSADVGLDADNCKVLVNI
EQQSPDIAQGVHGLTKKPEEIGAGDQGHMFGYATDETPELMPLTHVLATKLGAKLTVRKNTCPWLPPDGKTQVTVEYKNDNGAMVPIRVHT
VLSTQHDETVTNDKIAADLK(E)HVIKPVIPAKYLDENTIFHLNPSGRFVIGGPHGDAGLTGRKIIIDTYGGWGAHGGGAFSGKDPKVDKSGAYIVR
QAAKSVVASGLARRCIVQVSYAIGVAEPLSVFVDYKKTGTIPDKDILVLKENDFRPGMMSINLDLLKGGNYRYQKTAAYGHFGKDDPDFTWETV

Tabela 6 – Influência do tempo de análise por MS no número de peptídeos identificados por MS/MS

Spot	Fluxo contínuo a 5 µl/min	Fluxo reduzido 2 µl/min	Spot	Fluxo contínuo a 5 µl/min	Fluxo reduzido 2 µl/min
1	(K)TNMVMVFGEITTK(G) (R)VHTVLISTQHDETVTND AADLK(E)	(K)TNMVMVFGEITTK(G) (R)VHTVLISTQHDETVTND EIAADLK(E) (K)EHVIRPVIPEK(Y) (R)FVIGGPHGDAGLTGR(K) (K)SIVASGLAR(R) (K)ENFDFRPGMISINLDLK(R)	7	(K)YPIEHGIVSNWDDMEK(I) (R)VAPEEHPVLLTEAPLNPK(A) (R)TTGIVLDSDGVSHTVPIYEGYALPH ILR(L) (R)GYSFTTTAER(E) (K)LAYIALDYEQELETAK(S) (K)AEYDESGPSIVHR(K)	(R)VAPEEHPVLLTEAPLNPK(A) (R)TTGIVLDSDGVSHTVPIYEGYALPH ILR(L) (R)GYSFTTTAER(E) (K)LAYIALDYEQELETAK(S) (K)AEYDESGPSIVHR(K)
4	(K)IIAEYIWIGGSGMDLR(S) (K)WNYDGSSTGQAPGEDS EVILYPQAIFK(D) LTELALTVLSFDPK	(K)IIAEYIWIGGSGMDLR(S) (R)TLNGPVS DPAK(L) (K)IFSQPDVVAEVPWYGIEQEYTLQK(N)) (R)YILER(I) (R)ITEIADVLSFDPKPIQGDWNGAGAH TNYSTK(S) (R)NEGGYEIIK(K)	8	(K)AAIELDLLEIMAK(A) (K)ILMESWYYLK(D) (K)DAVLEGGIPFNK(A) (K)GINFDLPHVIEDAPPLPGVK(H)	(K)AAIELDLLEIMAK(A) (R)LLASYSVLTCTLR(N) (K)ILMESWYYLK(D) (K)GFEGLETVVDVGGGTGAVLSMIVAK(Y) (K)GINFDLPHVIEDAPPLPGVK(H) (K)NVIHIDCIMLAHNPGGK(E) (K)NCYDALVNIGK(V)
5	(R)AAVPSGASTGIYEALR(D) (K)EGLELLNTAIK(A)	(R)AAVPSGASTGIYEALR(D) (K)AVGNINSIIIPALVGK(D) (K)IPLYK(H) (K)LAMQEFMILPVGASSFK(E) (K)EGLELLNTAIK(A) (K)VVIGMDVAASEFYGEDK(S) (K)SYDLNFK(E) (K)VNQIGSVTESIEAVR(M) (K)YNQLLR(I) (R)IEEELGAEAVYAGANFR(K)	10	(K)TVGGGDDAFNTFFSETGAGK(H) (R)AVFVDLEPTVIDEVR(T) (R)QLFHPEQLISGK(E) (R)SLDIERPTYTNLNR(L) (R)LVSQVISSLTASLR(F) (R)IHFMLSSYAPVISA(EK)(A) (K)AYHEQLSVAEITNSAFEPSSMMAK(C) (R)AFVHWYVGEEMEEGEFSEAR(E)	(K)FDLAMYAK(R) (R)EDLALEK(D) (K)DVNAAVATIK(T) (R)LVSQVISSLTASLR(F) (R)AVFVDLEPTVIDEVR(T) (K)TVGGGDDAFNTFFSETGAGK(H) (R)QLFHPEQLISGK(E) (R)IHFMLSSYAPVISA(EK)(A) (R)LVSQVISSLTASLR(F) (R)AFVHWYVGEEMEEGEFSEAR(E) (R)FDGALNVDTFQTNLVPYPR(I) (R)INVYYNEASGGR(Y) (K)GHYTEGAELIDAVLDVVR(K) (R)MMLTFSVFPSPK(V) (R)FPGQLNSDLR(K) (K)LAVNLIPFPR(L) (R)LHFFMVGFAPLTSR(G) (R)YLTASAMFR(G) (K)NSSFVFEWIPNNVK(S) (R)VSEQFTAMFR(R)
6	(R)VLITDLLAR(G)	(R)GIYAYGFEKPSAIQQR(G) (K)GLDVIQQAQSGTGK(T) (R)ALGDYLGVK(V) (R)ILSSGVHVVVGTPGR(V) (R)VFDMLR(R) (K)IFVLDEADEMLSR(G) (R)GFKDQIYDIFQLLPSK(I) (K)DQIYDIFQLLPSK(I)	12	(K)LAVNLIPFPR(L) (K)NSSFVFEWIPNNVK(S)	

Tabela 7 – Comparação entre quantificação de RNA e proteínas

(continua)

Proteína	Nº spots	Volume relativo dos spots	Tipos de tags	Nº de tags	Proteína	Nº spots	Volume relativo dos spots	Tipos de tags	Nº de tags
<i>143 3 like protein</i>	7	2,31	3	34	<i>Glyoxalase</i>	1	0,50	1	6
<i>18.5 kDa class I heat shock protein</i>	1	1,20	--	--	<i>Heat shock 90 homolog</i>	1	0,32	1	2
<i>2,3-bisphosphoglycerate-independent phosphoglycerate mutase</i>	1	0,62	1	3	<i>Heat shock 70 kDa</i>	3	2,91	4	6
<i>20S proteasome alpha subunit</i>	3	0,44	2	6	<i>Heat shock 70 kDa protein mitochondrial precursor</i>	1	0,21	--	--
<i>20S Proteasome beta subunit</i>	2	0,08	1	3	<i>Isoflavone reductase</i>	5	3,36	3	268
<i>60S acidic ribosomal protein</i>	1	0,12	5	21	<i>Late embryogenesis abundant protein</i>	2	0,57	2	15
<i>6-phosphogluconate dehydrogenase</i>	2	0,92	--	--	<i>Leucoanthocyanidin dioxygenase</i>	1	0,06	1	6
<i>6-phosphogluconolactonase</i>	1	0,07	--	--	<i>Malate dehydrogenase mitochondrial precursor</i>	1	0,04	1	5
<i>Actin</i>	4	4,69	4	29	<i>Phosphoglycerate kinase cytosolic</i>	3	1,35	--	--
<i>Actin depolymerizing factor 1</i>	1	0,41	2	8	<i>Probable protein disulfide isomerase A6 precursor</i>	1	0,22	--	--
<i>Adenosine kinase</i>	1	0,50	1	50	<i>Profilin</i>	1	0,41	1	49
<i>Alcohol dehydrogenase</i>	1	0,05	3	11	<i>Putative 26S protease regulatory subunit homolog A</i>	1	0,21	2	3
<i>aluminum-induced protein</i>	1	0,15	1	15	<i>Putative alpha-soluble NSF attachment protein</i>	2	0,62	1	1
<i>Ascorbate peroxidases</i>	2	0,20	--	--	<i>Putative glycine-rich RNA-binding protein</i>	1	0,67	3	19
<i>ATP synthase alpha chain mitochondrial</i>	1	0,22	--	--	<i>putative NADPH-dependent mannose 6-phosphate reductase</i>	1	0,07	--	--
<i>ATP synthase beta chain mitochondrial</i>	1	1,21	1	2	<i>putative quinone oxidoreductase</i>	1	0,18	2	5
<i>Putative mitochondrial ATP synthase</i>	1	0,06	1	25	<i>Putative TPR-repeat protein</i>	1	0,35	1	5
<i>C2 domain-containing protein</i>	1	0,11	2	2	<i>Pyruvate dehydrogenase E1 component beta subunit</i>	1	0,25	--	--
<i>Caffeic acid 3 O methyltransferase</i>	2	1,38	1	68	<i>reversibly glycosylated polypeptide</i>	1	0,27	1	4
<i>Caffeoyl CoA O methyltransferase</i>	3	0,74	2	61	<i>S adenosylmethionine synthetase</i>	4	5,27	2	31
<i>Cell division cycle protein</i>	1	0,31	--	--	<i>Sgt1</i>	2	0,42	--	--
<i>Chalcone-flavanone isomerase</i>	1	0,09	1	2	<i>Soluble inorganic pyrophosphatase</i>	1	0,27	1	2
<i>chloroplast small heat shock protein</i>	1	0,70	--	--	<i>Superoxide dismutase [Mn], mitochondrial precursor</i>	1	0,08	--	--

Tabela 7 – Comparação entre quantificação de RNA e proteínas

(conclusão)

Proteína	Nº spots	Volume relativo dos spots	Tipos de tags	Nº de tags	Proteína	Nº spots	Volume relativo dos spots	Tipos de tags	Nº de tags
<i>cpn60 chaperonin</i>	1	0,28	1	1	<i>T complex protein 1 epsilon subunit</i>	1	0,21	--	--
<i>Cysteine synthase chloroplast precursor</i>	1	0,18	--	--	<i>thioredoxin peroxidase</i>	1	0,11	1	21
<i>D 3 phosphoglycerate dehydrogenase chloroplast precursor</i>	1	0,17	--	--	<i>Triosephosphate isomerase chloroplast precursor</i>	1	0,12	1	2
<i>Elongation factor 2</i>	2	0,34	--	--	<i>Triosephosphate isomerase cytosolic</i>	2	0,35	2	4
<i>Enolase</i>	2	2,01	1	2	<i>Tubulin alpha</i>	4	5,17	3	101
<i>Eukaryotic initiation factor 4A 13</i>	3	0,43	--	--	<i>Tubulin beta</i>	3	2,36	11	165
<i>Eukaryotic initiation factor 4A 6</i>	1	0,59	1	1	<i>UTP glucose 1 phosphate uridylyltransferase</i>	1	0,59	1	12
<i>F1-ATP synthase delta subunit</i>	1	0,12	--	--	<i>Ubiquitin-conjugating enzyme 2</i>	1	0,15	2	23
<i>flavanone 3-hydroxylase</i>	1	0,14	--	--	<i>UDP-glucose pyrophosphorylase</i>	1	0,73	3	29
<i>Fructokinase</i>	1	1,09	1	6	<i>Universal stress protein</i>	1	0,27	4	8
<i>GDP dissociation inhibitor</i>	1	0,25	--	--	<i>Vacuolar ATP synthase catalytic subunit A</i>	1	0,28	--	--
<i>Glutamine synthetase cytosolic isozyme</i>	2	1,74	3	66	<i>Vacuolar ATP synthase subunit B</i>	1	0,31	--	--
<i>Glutathione S-transferase</i>	1	0,52	4	32	<i>Vacuolar ATP synthase subunit C</i>	1	0,13	3	12

Deve-se levar em conta que a amostragem de SAGE é muito maior do que a de proteínas, pois apenas uma pequena porcentagem do gel de pH 4-7 foi seqüenciada, e o gel corresponde a apenas uma faixa de pI do proteoma total do tecido.

Outras técnicas deverão ser empregadas para identificar maior número de proteínas e quantificá-las com maior eficiência para que esta comparação possa representar os controles de síntese de macromoléculas na célula.

Acredita-se que a célula possua mecanismos que equilibram as taxas de síntese e degradação para as moléculas e que estas taxas possam ser diferentemente reguladas para RNAs e proteínas, conferindo-lhes meia-vida diferentes (CARGILE et al, 2004). Ainda a comparação do número de moléculas de RNA e proteínas pode não fazer muito sentido uma vez que grande número de polipeptídeos podem ser gerados a partir de uma molécula de RNA. Alguns autores ressaltam que a comparação também pode ser obscurecida pela etapa de processamentos pós-traducional pois o acréscimo de modificações cria maior diversidade no proteoma (SKYLAS et al., 2005). A confirmação de como ocorre esta rede de controle é um interessante desafio a ser esclarecido sobre os sistemas biológicos.

Se cada tipo de tag do SAGE corresponde a um gene diferente, deverá ser confirmado por PCR e então poderão ser correlacionadas ao número de isoformas de proteínas identificadas no proteoma.

4.11 Proteínas da região cambial de *E. grandis*

4.11.1 Proteínas de resposta a estress

Todos os organismos possuem proteínas de resposta a estresse e muitas são altamente conservadas. Este conjunto de proteínas comuns a todos os organismos celulares compõem o chamado proteoma mínimo de estresse (KÜLTZ, 2005) e inclui algumas das proteínas que foram identificadas na região cambial de eucalipto como tioredoxina, superóxido dismutase, quinona oxidoreductase, *6-phosphogluconate dehydrogenase* e *chaperonina HSP60*.

A capacidade de resposta da célula aos diversos tipos de estresse depende não só do genoma como também do proteoma, portanto é inerente não apenas à espécie mas também ao tecido e à etapa do desenvolvimento. A observação de que a época do ano está relacionada com períodos de maior infestação por doenças reforça essa teoria. A resposta ao dano da célula pode ser observada pelo aumento do nível de transcritos de alguns genes mas em outros casos ocorre uma alteração na reciclagem de proteínas e modificações pós-traducionais. Portanto importantes informações sobre os sistemas de defesa da célula devem ser estudadas também do ponto de vista dos proteomas.

Estresse celular pode ser causado por fatores bióticos e abióticos e normalmente danificam macromoléculas como DNA, proteínas ou lipídeos de membranas. Quando um sinal de estresse é deflagrado o nível de proteínas envolvidas na defesa eleva-se ou estas tornam-se ativas, e portanto atuam comumente em sistemas como reparo de DNA, reorganização de estrutura de proteínas desnaturadas ou remoção de radicais livres do meio celular (KÜLTZ, 2005).

O reparo de proteínas pode ser efetuado por enzimas que organizam a estrutura molecular, removem oxigênio reativo ou induzem à proteólise. A *Cpn60* é uma proteínas de localização mitocondrial que está relacionada com o correto dobramento e manutenção da estrutura de moléculas destinadas à importação pela organela. Em caso de estresse, e havendo desnaturação parcial de polipeptídeos, esta proteína age na reorganização de sua estrutura. *Cpn60* é uma proteína com homólogas em várias espécies e é induzida por estresse térmico (TSUGEKI et al., 1992). Uma *Cpn60* foi identificada no xilema em formação de *Eucalyptus*, representando 0,28% das proteínas totais na faixa de pI 4-7.

Foi encontrado um expressivo número de proteínas de resposta a estresse térmico, da família chamada *HSP* (*heat shock proteins*), como *HSP70* (*spots* 13, 30, 75 e 98), *HSP90* (*spot* 29) e *HSPs* de baixo peso molecular (*spots* 15 e 16). Estas proteínas também são chamadas de “chaperonas” pelo seu papel principal de proteção da estrutura de proteínas nos diversos sub-compartimentos das células eucarióticas, mas também estão envolvidas com outras funções. As *HSPs* são induzidas ou mais intensamente expressas em resposta a altas e baixas temperaturas (ANDERSON et al., 1994; GUY et al., 1998). A *HSP90* também foi localizada junto à região do

fragmoplasto, um centro organizador de microtúbulos durante a telófase (PETRASEK et al., 1998), e foi postulado que sua função na região possa ser de transmissora de sinais hormonais e do ambiente dirigindo a organização dos microtúbulos (GARDINER et al., 2001). *HSPs* de baixo peso molecular, semelhantes às encontradas nos *spots* 15 e 16, também têm sido associadas com atividade de proteção (VAN MONTFORT et al., 2001) e manutenção de integridade de membrana durante variações térmicas (TSVETKOVA, et al., 2002). Buscas por transcritos destas proteínas foram conduzidas no banco do FORESTS e diversos *clusters* corresponderam a RNAs destas classes de chaperonas (CAGLIARI et al., 2005).

Dois *spots* (96 e 97) do gel de proteínas de eucalipto foram identificados como contendo a proteína *SGT1*, cuja função é associada com complexos multiprotéicos. Sua sequência e estrutura de domínio (*tetratricopeptide*) assemelha-se à das *HSP90* e *HSP70*, portanto foi deduzido que sua função seja semelhante à destas proteínas de defesa (PEART e SHIRASU, 2002). Os 3 peptídeos de *SGT1* da região cambial de eucalipto (*spot* 96) obtiveram apenas 55% de alinhamento com uma proteína homóloga de *Oriza sativa* e também somente foi identificada através do banco específico GENOLYPTUS, o que pode indicar a especificidade de suas sequências.

Outra proteína de estresse encontrada na região cambial de *E. grandis* é a *universal stress protein* (USP), identificada através de 2 peptídeos e somente pela base gênero específica do GENOLYPTUS. É uma proteína bastante difundida entre procariontes, mas também encontrada em fungos, vírus e plantas, entretanto ausente em animais. Existem diversas proteínas com similaridade de sequências com USP e, devido a sua presença em diversos grupos taxonômicos, acredita-se que componham uma família com um gene ancestral muito antigo (ARAVIND et al., 2002). Em *E. coli* é bem relacionada com a sobrevivência da célula em situações de limitação de nutrientes e durante a fase estacionária (NYSTROM E NEIDHARDT, 1994). Em plantas de arroz foi demonstrado que o gene *OsUsp1* possui um papel importante para a adaptação das plantas a condições de baixa disponibilidade de oxigênio nas culturas submersas (SAUTER et al., 2002), sendo neste caso sua expressão regulada por influência de etileno. Análises do genoma de *Arabidopsis* mostraram a presença de 44 USPs, todas com similaridade de sequências com um domínio de USPA de bactérias (KERK et al.,

2003). Sua função no xilema em formação de *E. grandis* deve ser definida mas não é uma surpresa, uma vez que é tão conservada entre as espécies.

Foi identificada nos tecidos da região cambial de *E. grandis* uma proteína induzida por alumínio (*spot* 82) e a *GDP dissociation inhibitor* (*spot* 99), uma proteína que media um mecanismo de resposta à presença deste metal (EZAKI et al., 2000). É relatado que a toxidez causada por alumínio (Al) aumenta a peroxidação de fosfolípidos e proteínas das membranas celulares, levando a um reduzido crescimento principalmente das raízes (YAMAMOTO et al., 2001). Devido a sua atividade antioxidante, a *glutathione S-transferase*, também identificada neste trabalho, é uma das alternativas que a planta possui para conferir resistência ao Al (EZAKI et al., 1995). Estas proteínas merecem maior atenção pois podem ser de importância para o melhoramento do *Eucalyptus* por conferir tolerância ao alumínio em solos ácidos.

4.11.2 Síntese e turnover de proteínas

Várias das proteínas identificadas na região cambial de *E. grandis* podem ser relacionadas com proteólise. A ubiquitinação é um sistema de interação molecular que promove alterações estruturais em proteínas responsáveis por vários processos celulares. A alteração estrutural desencadeia uma transdução de sinais que pode resultar em proteólise, como no caso da atuação de sistemas de defesa da planta. Proteínas poliubiquitinadas são com frequência escoltadas para o *proteassoma* 26S para serem degradadas. O *proteassoma* 26S é um complexo proteolítico composto de cerca de 50 subunidades, encontrado no citoplasma e no núcleo. Sua função é a clivagem de ligações peptídicas de maneira altamente sítio-específicas.

As subunidades alfa e beta do *proteassoma* 20S, semelhantes ao 26S, foram encontradas em 5 pontos do gel (*spots* 103, 115, 116, 118 e 135), e também têm uma atividade proteolítica dependente de ATP. O *proteassoma* 20S é formado pela interação de proteínas SNARES com fatores NSF (*N-ethylmaleimide-sensitive*) (ROTHMAN, 1994), durante os processos de fusão de vesículas (XUE e ZHANG, 2002). O fator NSF também foi encontrado no *spot* 53 deste trabalho. As v-SNAREs são proteínas receptoras encontradas nas membranas de vesículas e t-SNAREs localizam-se na membrana alvo. Alguns fatores NSF podem ter função de ATPase e outros agem como

estimuladores desta função (HAYNES et al., 1998). É compreensível que estas proteínas tenham sido encontradas na região cambial analisada uma vez que as plantas possuem um sistema de endomembranas complexo (revisado em SANDERFOOT e RAIKHEL, 1999). Da mesma maneira pode-se assumir que o sistema de tráfego de vesículas envolvidas com a distribuição de metabólitos entre os diversos compartimentos subcelulares e o exterior da célula deva ser muito ativo em tecidos responsáveis pelo trânsito de assimilados na planta como o floema.

Ainda o *spot 77* (*ubiquitin-conjugating enzyme 2*) pode ser incluído no grupo das proteínas que agem por interação molecular mediando a degradação seletiva de proteínas de meia-vida curta ou com defeitos de seqüência ou estrutura.

Também relacionadas à biosíntese de proteínas foram encontradas as proteínas *60S acidic ribosomal* e fatores de iniciação e elongadores da tradução (*spots* 6, 66, 73, 87 e 126). Juntamente com os proteossomas 20 e 26S constituem um grupo responsável pelo *turnover* de proteínas da região cambial.

4.11.3 Proteínas de estresse oxidativo

Radicais reativos de oxigênio tóxicos são produzidos na célula por efeitos ambientais como baixas temperaturas, alta intensidade luminosa, estresse hídrico e produtos químicos. Proteínas encarregadas da remoção de radicais tóxicos em decorrência de estresse oxidativo foram identificadas no proteoma de *Eucalyptus*, como ascorbato peroxidase, superóxido dismutase, *thioredoxine peroxidase* e *glutathione S-transferase*. Muitas oxidoreduases que agem em resposta a estresse são também componentes de metabolismos básicos como glicólise, metabolismo do piruvato (*glyoxalase I*) ou no metabolismo de aminoácidos (*glutathione S-transferase*).

Uma superóxido dismutase mitocondrial (SOD) foi identificada em um único *spot* e mostrou maior homologia com a mesma proteína de pessegueiro (*Prunus persica*). SODs são classificadas em três classes de acordo com seu cofator metálico: SOD Fe, comuns no citosol e cloroplastos, SOD Mn, encontrada em procariontes e mitocôndria e SOD Cu/Zn, mais comuns em células eucariontes. No xilema jovem de *Eucalyptus* um único *spots* foi identificado como SOD-Mn e representou 0,08% das proteínas no gel de

pH 4-7. Níveis mais altos de RNAm e de proteínas de SOD-Mn já foram relacionados com tecidos juvenis e meristemáticos de pessegueiro, e a expressão também parece ligada ao estágio do desenvolvimento (BAGNOLI et al, 2002).

Ascorbato peroxidase é uma proteína que atua na destoxificação de peróxido de hidrogênio. Na região cambial do eucalipto esta enzima foi encontrada em dois *spots* (51 e 143) de PM entre 27 e 30 KDa, porém com pls diferentes. A observação desta proteína em faixas diferentes de pH também foi feita por Méchin et al. (2004) ao investigar o proteoma de endosperma de milho. Estes autores encontraram 4 possíveis isoformas da ascorbato peroxidase, entretanto a partir de uma amostragem de 496 *spots* identificados. É possível que, aumentando a sensibilidade da técnica, mais isoformas desta proteína sejam encontradas também em eucalipto.

A *glyoxalase I* (metabolismo do piruvato) e II são enzimas que catalizam a destoxificação celular de metilglioxal, um composto citotóxico que reage com DNA e proteínas causando alterações estruturais. A via da *glyoxalase* é dependente de glutatona (GSH) e ocorre conforme a reação:

A *glyoxalase I*, que representou 0,5% das proteínas de pl 4-7 da região cambial de *E. grandis*, têm sido associada com o aumento da resistência de plantas a estresse salino (BAJAJ et al., 1999), ao frio e à seca (SEKI et al., 2001) uma vez que o metilglioxal é encontrado em altas concentrações como resultado do estresse. A *glyoxalase* também foi relacionada com o crescimento e divisão celular (PAULUS et al., 1993) pois acredita-se que o S-D lactoyl GSH, um produto da via da *glyoxalase*, desempenha um papel importante na montagem dos microtúbulos (HOOPER et al., 1988). Estes dados estão em conformidade com a observação da presença da *glyoxalase* nos tecidos meristemáticos do xilema de *Eucalyptus*, que têm alta taxa de divisão celular, processo

dependente da organização de microtúbulos para organização de fuso mitótico e citoesqueleto.

A *thioredoxin peroxidase* (TRX) é outra classe multigênica de proteínas relacionadas com proteção contra a oxidação. Elas reduzem compostos celulares através de reações que requerem a presença de um grupamento tiol e são encontradas em plantas no citosol, cloroplastos e mitocôndrias. Através da análise de clusters de ESTs de *Eucalyptus* do banco FORESTs, foram identificados vários transcritos desta proteína (BARBOSA E MARINHO, 2005). Os níveis mais altos de expressão da proteína foram detectados em plântulas expostas à luz, folhas com deficiência dos elementos fósforo e boro, e folhas colonizadas por *Thyrinteina* (uma lagarta desfolhadora). Por outro lado os menores níveis foram encontrados neste trabalho em plântulas crescidas no escuro e raízes. Tecidos de madeira apresentaram níveis intermediários entre estes grupos. Neste trabalho o sequenciamento de proteínas por MS/MS da região cambial de *E. grandis* mostrou também uma tioredoxina identificada através das bibliotecas de xilema, floema e de plântulas do GENOLYPTUS.

Glutathione S-transferase (GST) é uma enzima candidata à produção de transgênicos pelo seu papel de proteção contra agentes xenobióticos e produtos do metabolismo oxidativo. Também participa da produção de produtos do metabolismo secundário (LI et al., 1997). Em plantas, existem cinco grupos de GSTs baseados em sua seqüência de aminoácidos: Theta, Zeta, Omega, Phi e Tau (EDWARDS et al., 2000), com identidade de seqüências inferior a 10%. A alta variabilidade das seqüências leva a crer que as diferentes isoformas participam em diversas funções. Já foi relatado que em várias espécies de plantas há uma expressão gene-específica e variável de GST durante o desenvolvimento, em resposta a fatores ambientais, químicos ou a patógenos (MARRS, 1996). Em *Arabidopsis* existem 53 genes de GSTs e já foi observado que o ácido salicílico promove o aumento na expressão de pelo menos 4 proteínas desta classe (SAPPL et al., 2004.) e de um número maior destes genes quando são analisados os transcritos (WAGNER et al., 2002) demonstrando que podem haver diferentes níveis de regulação durante as etapas de transcrição e tradução dos genes de GST. Possivelmente, por ser uma enzima importante para a proteção da planta, existem isoforma que são menos dependentes de fatores ambientais, sendo constantemente

sintetizadas para participar em outras vias do metabolismo básico celular. Dentre as proteínas identificadas na região cambial de *Eucalyptus* foi encontrado um único local no gel correspondendo à GST (spot 20).

Late embryogenesis like protein (LEA) são proteínas encontradas em alta concentração em tecidos com tolerância a estresse hídrico, como sementes, e estresse salino (XU et al., 1996). Têm sido localizadas no citoplasma, núcleo, mitocôndrias e cloroplastos de plantas. LEAs também foram localizadas no floema de *Ricinus communis* e relacionadas com o transporte de micronutrientes (KRUGER et al., 2002).

4.11.4 Metabolismo de nitrogênio e aminoácidos

A *S-adenosylmethionine synthetase* (SAM-S) foi a proteína mais intensamente expressa entre as proteínas analisadas do xilema de eucalipto. É a enzima que cataliza a formação de *S-adenosylmethionine* (SAM) a partir de metionina e ATP. SAM atua como principal doadora de grupamentos metil para uma série de compostos na célula como proteínas, polissacarídeos e lipídeos. Na região cambial de eucalipto SAM-S foi representada por 4 spots (1, 9, 106 e 108). A partir da quantificação relativa destas proteínas observou-se que SAM-S foi a mais intensamente expressa. Altos níveis de expressão de SAM-S também foram relatados em xilema em desenvolvimento de *Populus* (STERKY et al., 1998) e *Pinus* (ALLONA et al., 1998).

Adenosine kinase (ADK) catalisa a síntese de adenina monofosfato a partir de adenosina e ATP. Plantas mutantes de *Arabidopsis* deficientes para ADK têm aspecto compacto apresentando encurtamento dos brotos primários, estames, hipocótilos e raízes (MOFFATT et al., 2002). Devido à relação de ADK com a metilação para a célula (por reciclagem de adenosina), foi postulado que a falta de ADK nos tecidos reduz a disponibilidade de SAM. Foi também postulado que ADK tem relação com sistemas de defesa uma vez que interage especificamente com proteínas virais em *Nicotiana benthamiana* e poderia agir como um ativadora indireta de outras proteínas kinases (WANG et al., 2003).

Glutamina sintetase forma parte do metabolismo de nitrogênio e catalisa a síntese de glutamina a partir de glutamato e amônia, com requerimento de ATP. O metabolismo do

nitrogênio é crucial para as plantas para a síntese de proteínas e por metabolizar amônia, que em altas concentrações, torna-se tóxica para as células e pode levar à morte da planta (EVSTIGNEEVA et al., 2003). Por outro lado a escassez de nitrogênio nos tecidos pode aumentar a sensibilidade a patógenos (SOLOMON et al., 2003), além de reduzir o crescimento. No proteoma analisado da região cambial de *E. grandis* a concentração de glutamina sintetase foi das mais expressivas (1,74%).

Relacionadas com síntese de aminoácidos foram identificadas por MS/MS o precursor cloroplastidial da proteína *D-3-phosphoglycerate dehydrogenase*, que participa da via de síntese de glicina, serina e treonina, e a *cisteína sintetase* que catalisa a formação deste aminoácido.

4.11.5 Proteínas do metabolismo de carboidratos

A glicólise é uma via do metabolismo primário da célula e o principal processo de uso energético da glicose, portanto não é de surpreender que estas proteínas estejam fortemente representadas em um tecido com alto requerimento energético como o meristema cambial. Várias enzimas da via glicolítica foram identificadas, entre elas *triosephosphato isomerase* (citossólica e plastidial, nos spots 105, 50 e 101), enolase (spots 5 e 33), *phosphoglycerate kinase* (spots 27, 60 e 90) e *2,3-bisphosphoglycerate-independent phosphoglycerate mutase* (spot 28). Além de duas moléculas de ATP a glicólise produz duas moléculas de piruvato que é oxidado nas mitocôndrias pela piruvato desidrogenase, uma enzima que também foi encontrada neste trabalho (spot 138). Também pertencendo ao metabolismo do piruvato e do ciclo do citrato foi sequenciado o precursor mitocondrial da malato desidrogenase, que cataliza a síntese de oxaloacetato. É interessante observar que a fosfoglicerato mutase encontrada na região cambial é independente do cofator *2,3-bisphosphoglycerate*, como é característico em plantas (HUANG et al., 1993). Esta proteína também foi identificada na seiva do floema de *Ricinus communis* por 2D/MS (BARNES et al., 2004).

Ainda pertencendo ao metabolismo de carboidratos, porém à via alternativa das pentose fosfato, foram encontradas a *6-phosphogluconolactonase* e a *6-phosphogluconate dehydrogenase*. Esta via é uma alternativa para a oxidação da

glicose a pentoses importantes para a célula, como a D-ribose indispensável para a formação de ácidos nucleicos. Em plantas a via também é importante por direcionar o fluxo de carbono para a síntese de aminoácidos aromáticos, flavonóides e ligninas (HERRMAN e WEAVER, 1999). Apesar de ter sido observada alta expressão relativa (0,92%) da *6-phosphogluconate dehydrogenase* do proteoma do xilema jovem de eucalipto nenhuma tag correspondente foi observada e nem tampouco para a 6-phosphogluconolactonase.

UDP-glucose pyrophosphorylase (UGPase) é uma enzima que cataliza a fosforilação de glicose-1-fosfato para a produção de UDP-glicose, um açúcar que fornece grupamentos glicosídicos para serem incorporados com outras macromoléculas importantes para a célula como glicolipídeos, hemiceluloses e ácidos urônicos. Está envolvido com a síntese de sacarose e celulose e tem a alternativa de produzir ou utilizar UDP-glicose dependendo da atividade do tecido. Por exemplo, em células não fotossintetizantes liga-se à degradação da sacarose para fornecer fluxo de carbono para a síntese de amido (KLECZKOWSKI, 1994). O fato de estar ligada à síntese de celulose torna a UGPase economicamente importante e interessante para a manipulação genética. Entretanto pouco sabe-se sobre a regulação molecular de sua expressão. Para compreender a influência da expressão da UGPase na síntese de carboidratos, o gene foi clonado com promotor da enzima 4-coumarate:CoA ligase (4CL) que é preferencialmente expresso em caules (COLEMAN et al., 2006). Embora o conteúdo de açúcar solúvel e o crescimento da planta tenham sido maior do que em plantas controle, não foi possível incrementar os teores de celulose por massa de tecido. Já foi observado que, devido à complexidade das rotas metabólicas, as plantas têm enorme capacidade de adaptação para a manutenção do seu metabolismo (BAUCHER et al., 1996), reequilibrando níveis de enzimas, substratos e produtos finais. Portanto maior conhecimento e pesquisas se fazem necessário para obter sucesso na alteração genética de vias como a de carboidratos em plantas.

Na região cambial de *E. grandis* duas proteínas UGPase foram seqüenciadas (42 e 91) apresentando um índice considerável de expressão (0,73% de volume relativo). Sua presença nos tecidos em formação do xilema é coerente com sua função para a síntese

de celulose que deve ser incorporada às paredes celulares do xilema e do floema, além de seu papel nas outras importantes vias sintéticas mencionadas.

Ainda participando das rotas metabólicas dos carboidratos (metabolismo de frutose e manose) para manutenção de um fluxo de carbono na direção da síntese de amido, foram identificadas a *fructokinase*, que cataliza a conversão de frutose em frutose-6-fosfato, e a *putative NADPH-dependent mannose 6-phosphate reductase*, que cataliza a conversão de D-mannitol 1-phosphate em D-mannose 6-phosphate.

Alcohol dehydrogenase e *pyruvate dehydrogenase* podem ser encontradas na matriz mitocodrial participando da gliconeogênese, passo subsequente à glicólise. Estas duas enzimas também foram encontradas no xilema em desenvolvimento de *E. grandis* e reforçam o número de proteínas de metabolismo de carboidratos.

4.11.6 Proteínas de parede

A lignina é o segundo polímero mais abundante encontrado na madeira e é uma molécula essencial na formação de parede celular secundária do xilema. Na região cambial de eucalipto foram encontradas 2 enzimas ligadas à sua síntese: *caffeic acid-3-O methyltransferase* (COMT) e *caffeoyl-CoA O-methyltransferase* (CCoAOMT), que também foram encontradas no proteoma de xilema em formação de *Pinus* (GION et al., 2005).

A biossíntese de flavonóides é ativa no xilema em desenvolvimento de árvores uma vez que é a via de síntese da lignina, importante para a formação da parede celular. Em concordância com isso foram encontradas as enzimas *flavanone 3-hydroxylase*, *putative chalcone isomerase 4*, *leucoanthocyanidin dioxygenase* e, em alta concentração a *isoflavona reductase* (3,36 % de volume relativo), todas relacionada com síntese de isoflavonóides (Tiemann et al., 1991). A isoflavona reductase foi a proteína que apresentou maior expressão em géis 2D de proteínas de xilema e casca de *Populus* (VANDER MIJINSBRUGGE et al., 2000).

Relacionada com a formação de glicoproteínas, importantes para a parede celular, foi identificada uma *glycine-rich RNA-binding protein* (GRP) pertencente à classe IV de proteínas com um domínio rico em resíduos de glicina (FUSARO et al., 2001). Foram identificadas em plantas e possuem duas regiões conservadas: uma C-terminal rica em

glicina, além de dois sítios conservados na região N-terminal necessários para o reconhecimento de seqüências específicas do RNA. Além de sítios para ligação com outras proteínas e RNA, também foi demonstrada sua afinidade por fitas simples de DNA (DUNN et al., 1996). Diversos sinais externos e internos à célula regulam a expressão destas classes de proteínas como ataque de patógenos, ácido abscísico, estresse hídrico, temperatura, o relógio circadiano e o desenvolvimento (revisado por SACHETTO-MARTINS et al., 2000).

No banco do FORESTs foram identificados diversos transcritos de GRPs pertencentes à classe IV, inclusive nos tecidos da madeira de *Eucalyptus* (BOCCA, et al. 2005), o que confirma sua presença nestas espécies. Em acordo com essas observações, este trabalho mostrou a presença de GRP com domínio para RNA no proteoma do xilema em formação de *E. grandis*.

A *disulfide-isomerase* (DI) cataliza as trocas de grupamento tiol e disulfeto para a formação de pontes de enxofre e estabilização da estrutura terciária e quaternária de proteínas que são processadas pelo sistema de endomembranas como glicoproteínas (FREEDMAN, 1989). Esta proteína está presente no lúmen do retículo endoplasmático. A atividade de DI foi testada por imunolocalização em alfafa nos hipocótilos e caules entre 5-10 semanas e observou-se que, enquanto nos hipocótilos a proteína estava presente na epiderme e sub-epiderme, nos caules mais maduros a atividade ficou restrita à zona cambial (SHORROSH et al, 1993). Em caules e pecíolos de soja a presença de DI também restringiu-se à zona cambial e região adjacente do floema. Foi sugerido neste trabalho que a presença da enzima pode refletir a atividade metabólica dos tipos celulares estando presente em tecidos com maior requerimento por proteínas que necessitam de processamento no retículo endoplasmático. Isto está em acordo com a detecção de uma DI na região cambial de *Eucalyptus*. Estudos mais detalhados seriam necessário para confirmar se sua presença restringe-se à região do floema como encontrado no trabalho de Shorrosh et al (1993), mas certamente é uma proteína que merece atenção devido ao seu papel sobre a estrutura espacial das proteínas e por participar na adição de grupamentos glicosídicos durante o processamento pós-traducionais de proteínas. Proteínas glicosiladas são de fundamental importância para a constituição de proteínas de parede celular.

A *reversibly glycosylated polypeptide* (RGP) pertence a uma família de proteínas altamente conservadas e específicas de plantas. É uma proteína periférica de membrana, portanto liga-se à superfície destas normalmente por associação com outras proteínas integrais. É sintetizada nos polissomos do citosol e associa-se à membrana do Complexo de Golgi como uma proteína periférica (DHUGAA et al., 1997) sendo também incorporada aos plasmodesmos pelo deslocamento de vesículas do Golgi (SAGI et al., 2005). As RGP são autoglicosiladas na presença de UDP-glicose, UDP-xilose e UDP-galactose (DHUGGA et al., 1997). Podem estar indiretamente relacionados ao metabolismo de polissacarídeos de parede transferindo açúcares nucleotídeos a supostos transportadores que atuam nos plasmodesmos (SAGI et al., 2005).

A suposta função como proteínas carreadoras de grupamentos glicosídicos destinados à parede celular torna-a uma proteína de interesse para o melhoramento da madeira, e sua função na formação destes tecidos deverá ser mais bem caracterizada. Foi identificada uma RGP neste trabalho a partir de um único peptídeo (KASNPVFVNLKK) e com a ajuda de bibliotecas específicas de xilema e floema do banco GENOLYPTUS, sendo encontrado 86% de identidade com sequências da mesma proteína de *Gossypium hirsutum*.

4.11.7 Metabolismo de energia

Em um tecido com alta atividade metabólica é alto o requerimento de energia. Em concordância com isso, algumas sub-unidades do complexo de poro da ATP sintase mitocondrial foram localizadas. A figura 15 mostra as proteínas do processo de fosforilação oxidativa identificadas neste trabalho e sua provável localização na membrana e matriz mitocondrial. Este complexo produz ATP a partir de ADP devido ao gradiente de prótons formado através da membrana. As subunidades alfa e beta da porção imersa na matriz mitocondrial foram sequenciadas sendo que a cadeia alfa tem função catalítica e a cadeia beta funciona como subunidade reguladora. A proteína *F1-ATP synthase delta* (spot 120) é outro membro deste complexo que foi identificada por sequenciamento MS/MS.

4.11.8 Sinalização celular

Muitas proteínas têm seu estado funcional regulado por transdução de sinal mediado por fosforilação. Em alguns casos a efetivação desta nova atividade requer além da fosforilação a interação com outras proteínas como a 14-3-3. Esta proteína tem sido relacionada com diversas funções envolvendo associação molecular para o transporte de proteínas dentro da célula (RITTINGER et al., 1999), ligação com proteínas kinases (MARU e WHITE, 1992) e respostas a estress (JARILLO et al., 1994). A 1433 foi a proteína encontrada em maior número de *spots* (7) no gel de eucalipto.

Também foram seqüenciados três peptídeos (*spot* 74) onde foi encontrada homologia com o domínio *tetratricopeptide* (TPR) conhecido por mediar interações proteína-proteínas, e relacionar-se com o funcionamento de *chaperones*, ciclo celular, transcrição e transporte de proteínas (BLATCH et al., 1999). Em *Arabidopsis* foi identificado o gene *mgo3* com um domínio TPR que possui papel fundamental nas divisões celulares meristemáticas. O mutante para este gene apresentou desorganização estrutural e funcional nos meristemas apicais de brotos e de raízes (GUYOMARCH et al., 2004). Sendo a região cambial de *Eucalyptus* rica em células meristemáticas, pode-se inferir que o papel desta proteína neste tecido esteja relacionado com a organização das divisões celulares destas células.

Outro domínio identificado e relacionado a interações entre proteínas e comum a fosfolipases de plantas é o C2 (encontrado no *spot* 95) na região N-terminal, que possui aproximadamente 130 resíduos de aminoácidos (WANG, 2000). Este domínio é comum a proteínas envolvidas com sinalização interna na célula. Comumente esse domínio possui afinidade por ligação a íons Ca^{++} que estimulam sua atividade enzimática. Fosfolipases são enzimas ligadas à biogênese de membranas e importantes para viabilidade celular e homeostase.

4.11.9 Citoesqueleto

Actinas compõem os microtúbulos, elementos essenciais para o citoesqueleto responsável pela forma da célula, sustentação dos elementos presentes no citoplasma

e composição do fragmoplasto durante a divisão celular da célula vegetal. Os monômeros de actina estão em constante polimerização e despolimerização dependendo da necessidade da célula, portanto não é surpresa que também tenham sido encontradas proteínas ligadas a esta reciclagem como o *actin depolymerizing factor* (*spot* 44) e a *profilina*. Esta última proteína também foi encontrada no floema de plântulas de *Ricinus communis* (SCHOBERT et al., 2000), e é um forte regulador da polimerização de actina (STAIGER et al., 1997) quando em baixas concentrações. Ambos fatores têm afinidade pelos monômeros de actina e são totalmente esperados em um tecido com intensa taxa de divisão celular, como a região do meristema cambial de *E. grandis*. As três proteínas também foram identificadas neste tecido por SAGE. A alta concentração de profilina foi observada em exudatos de floema em *R. communis* (SCHOBERT et al., 2000) sendo sugerido que ela possa estar sendo produzida nas células companheiras para ser distribuída para outros tecidos da planta ou mantida em altas concentrações na seiva do floema para impedir a polimerização da actina nos líquidos circulantes.

Tubulinas α e β são codificadas por famílias de genes presentes em animais e plantas. Estas proteínas constituem a unidade estrutural dos microtúbulos, um heterodímero constituído de uma molécula α -tubulina (*TUA*) e uma β -tubulina (*TUB*), necessário em processos celulares como divisão celular e morfogênese. Através da análise de transcritos foram encontrados nove genes de *TUB* em *Arabidopsis* com seqüências altamente conservadas (SNUSTAD et al., 1992). Já nesta mesma planta as seqüências dos genes de *TUA* são bastante divergentes (KOPCZAK et al., 1992). A divergência de seqüências de genes dentro de um genoma e entre espécies é interessante porque pode refletir diferentes propriedades funcionais. Na região cambial de *E. grandis* puderam ser identificadas 4 α -tubulinas (*spots* 10, 21, 22 e 34) e 3 β -tubulinas (*spots* 11, 12 e 142). A figura 16 mostra que existem possivelmente mais duas isoformas que não puderam ser removidas do gel em função da pouca resolução entre os *spots*. Possivelmente representam α -tubulinas, o que estaria em acordo com as observações feitas em outras espécies onde maior número de genes foram encontrados.

A separação desta amostra em uma faixa mais estreita de pH durante a focalização seria uma alternativa para esclarecer o número de tubulinas expressas na região

cambial de eucalipto. Estes genes estão entre os mais expressos no tecido analisado, uma zona de alta taxa de divisão celular, confirmando seu importante papel na formação de tecidos do xilema.

Figura 16 – Isoformas de tubulinas encontradas na região cambial de *E. grandis*. As setas vermelhas apontam os spots que não puderam ser identificados devido à baixa resolução da região

T complex protein 1 epsilon subunit, também chamada complexo CCT, corresponde a uma chaperona de alto peso molecular de localização citoplasmática encarregada do dobramento de proteínas. Esta proteína tem sido melhor caracterizada em animais e está relacionada com afinidade por actina e tubulina para organização do citoesqueleto (STERNLICHT et al., 1993) e formação do fuso mitótico em *Saccharomyces cerevisiae* (URSIC & CULBERSON, 1991). Em coleóptilo de milho foi observada a deposição pronunciada de CCT em regiões de arranjos de microfibrilas, especialmente em células de protoxilema (HIMMELSPACH et al., 1997) caracterizadas por espessamento secundário de parede celular. Em células de tabaco também foi sugerido que a CCT desempenhe um papel no controle espacial da deposição da parede celular, em especial em células em divisão, na região do fragmoplasto (NICK et al, 2000). A presença desta proteína na região cambial de *Eucalyptus* confirma os relatos da literatura e despertam também para a função das CCT que podem constituir uma nova opção para a manipulação genética de células visando alterações de parede celular que reflitam na qualidade da madeira.

4.12 Local da síntese de proteínas da região cambial de *E. grandis*

O câmbio vascular meristemático produz células fusiformes destinadas à formação das células do xilema e do floema. Os elementos crivados são uma destas células presentes no floema. Durante a maturação de elementos crivados do floema e dos elementos de vaso do xilema para formação dos vasos condutores, ocorre a degradação do citoplasma e organelas. A presença de proteínas e moléculas na região cambial de *E. grandis* que são sintetizadas unicamente em organelas está de acordo com a hipótese de que células companheiras ricas em citoplasma sejam as responsáveis pela síntese destes assimilados (Raven, 1991). Esta hipótese é sustentada pela observação do transporte de proteínas desde as células companheiras para os elementos crivados através de plasmodesmas (SAKUTH et al, 1993). Foram observados plastídeos em células fusiformes cambiais de *Pinus* (RENSING et al., 2000), e junto à membrana plasmática de elementos crivados (KNOBLAUCH E VAN BEL, 1998), o que demonstra que a presença de moléculas exclusivamente sintetizadas em cloroplastos podem ter sido sintetizadas na própria região cambial e não apenas ser provenientes do metabolismo foliar e transportadas pelo floema.

Análises de exudatos de floema de *R. communis*, trigo e arroz mostraram a presença de centenas de proteínas, incluindo a presença de proteínas fosforiladas (HAYASHI et al., 2000). Estes autores postularam que proteínas fosforiladas podem estar presentes na seiva do floema para atuarem como transdutoras de sinal a longas distâncias ao longo da planta.

A localização exata da síntese e acúmulo das proteínas da região cambial de *E. grandis*, células cambiais iniciais, derivadas ou células em diferenciação do xilema ou do floema, poderá ser esclarecida com uso de outras técnicas como hibridização *in situ*.

5 CONCLUSÕES

- ☞ A técnica 2D-LC-MS/MS foi adequada para a identificação de apenas um pequeno grupo de proteínas, tornando-se necessário enriquecimento de frações para obter um banco de proteínas mais representativo do tecido analisado.
- ☞ O proteoma da região cambial de *Eucalyptus grandis* aos 2 anos e 8 meses mostra que as proteínas relacionadas a estress estão em maior concentração em relação a outros grupos metabólicos, embora algumas destas proteínas também façam parte de metabolismos básicos da célula.
- ☞ Não foi encontrado correlação entre os níveis de expressão de proteínas e de seus RNAs correspondentes (SAGE), possivelmente porque as duas moléculas possuem diferentes velocidades de turnover e devido a eventos de processamento pós-traducional.

REFERÊNCIAS:

AGRAVAL, G.K.; RAKWAL, R.; TAMOGAMI, S.; YONEMURA, M.; KUBO, A.; SAJI, H. Chitosan activates defense/stress response(s) in the leaves of *Oryza sativa* seedlings. **Plant Physiology and Biochemistry**, Amsterdam, v.40, p.1061-1069, 2002.

AGRAWAL, G.K.; RAKWAL, R. Rice proteomics: A cornerstone for cereal food crop proteomes. **Mass Spectrometry Reviews**, New York, v.,5, n.1, p.1-53, Jan-Feb, 2006.

ALLONA, I.; QUINN, M.; SHOOP, E.; SWOPE, K.; st CYR, S.; CARLIS, J.; RIEDL, J.; RETZEL, E.; CAMPBELL, M.M.; SEDEROFF, R.R.; WHETTEN, R.W. Analysis of xylem formation in pine by cDNA sequencing. **Proceedings of National Academy of Science**, Washington, v.95, p.9693-9698, 1998.

ALTSCHUL, S.F.; MADDEN, T.L.; SHCAFFER, A.A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D.J. "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs". **Nucleic Acids Research**, Oxford, v.25, p.3389-3402, 1997.

ALVES, E.; XAVIER, A.; OTONI, W.C. Organogenesis of the leaf explant of *Eucalyptus grandis* x *E. urophylla* clones. **Pesquisa Agropecuária Brasileira**, Brasília, v.39, n.5, p.421-430, 2004.

ANDERSON, J.V.; LI, Q.-B.; HASKELL, D.W.; GUY, C.L. Structural organization of the spinach endoplasmic reticulum-luminal 70-kilodalton heat-shock cognate gene and expression of 70-kilodalton heat-shock genes during cold acclimation. **Plant Physiology**, Baltimore, v.104, p.1359–1370, 1994.

ANDERSON, L.; SEILHAMER, J.A. Comparison of selected mRNA and protein abundances in human liver. **Electrophoresis**, NewYork, v.19, n.3-4, p.533-537, 1997.

ANDERSON, N.L.; ANDERSON, NG. Proteome an proteomics: new technologies, new concepts and new words. **Electrophoresis**, NewYork, v.19, p.1853-1861, 1998.

ARAVIND, L.; ANANTHARAMAN, V.; KOONIN, E.V. Monophyly of class I aminoacyl tRNA synthetase, USPA, ETFP, photolyase, and PP-ATPase nucleotide-binding domains: implications for protein evolution in the RNA world. **Proteins Structure and Functional Genetics**, New York, v.48, p.1–14, 2002.

ARIOLI, T.; PENG, L.; BETZNER, A.S.; BURN, J.; WITTKE, W.; HERTH, W.; CAMILLERI, C.; HOFTE, H.; PLAZINSKI, J.; BIRCH, R.; CORK, A.; GLOVER, J.; REDMOND, J.; WILLIANSO, R.E. Molecular analysis of cellulose biosynthesis in *Arabidopsis*. **Science**, Washington, v.279, p.717-720, 1998.

ASSOCIAÇÃO BRASILEIRA DA INDÚSTRIA DE MADEIRA PROCESSADA MECANICAMENTE (ABIMCI). Estudo setorial, 2004. Disponível em : <<http://www.abimci.com.br>>. Acesso em 10 fev. 2006.

BAE, M.S.; CHO, E.J.; CHOI, E.Y.; PARK, O.K. Analysis of the *Arabidopsis* nuclear proteome and its response to cold stress. **The Plant Journal**, Oxford, v.36: p.652-663, 2003.

BAGNOLI, F.; GIANNINO, D.; CAPARRINI, S.; CAMUSSI, A.; MARIOTTI, D.; RACCHI, M.L. Molecular cloning, characterization and expression of a manganese superoxide dismutase gene from peach (*Prunus persica* [L.] Batsch). **Molecular Genetics and Genomics**, Berlin, v.267, n.3, p.321-8, 2002.

BAJAJ, S.; TARGOLLI, J.; LIU, L.F.; HO, T.H.D.; WU, R. Transgenic approaches to increase dehydration-stress tolerance in plants. **Molecular Breeding**, Amsterdam, v.5, p. 493–503, 1999.

BARBOSA, A.E.; MARINHO, P. *In silico* analysis of *Eucalyptus* thioredoxins. **Genetics and Molecular Biology**, Ribeirão Preto, v.28, n.3, p.539-547, 2005.

BARNES, A.; BALE, J.; CONSTANTINIDOU, C.; ASHTON, P.; JONES, A.; PRITCHARD, J. Determining protein identity from sieve element sap in *Ricinus communis* L. by quadrupole time of flight (Q-TOF) mass spectrometry. **Journal of Experimental Botany**, Oxford, v.55, n.402, p.1473–1481, 2004.

BAUCHER, M.; CHABBERT, B.; PILATE, C.; VAN DOORSSELAERE, J; TOLLIER, MT; PETIT-CONIL, M; CORNU, D; MONTIES, B; VAN MONTAGU M; INZE, D; JOUANIN, L; BOERJAN, W. Red Xylem and Higher Lignin Extractability by Down-Regulating a Cinnamyl Alcohol Dehydrogenase in *Poplar*. **Plant Physiology**, Baltimore, v.12, p.1479-1490, 1996.

BERNAL, A.; EAR, U.; KYRPIDES, N. Genomes OnLine Database (GOLD): a monitor of genome projects world-wide. **Nucleic Acids Research**, Oxford, v.29, p.126-127, 2001.

BHOORA R, BERGER DK, WINGFIELD MJ, WIGFIELD, B.D. PCR cloning by genome walking of a complete polygalacturonase-inhibiting protein gene from *Eucalyptus grandis*. **South African Journal of Science**, Grahamstown, v.99, n.9, p.410-412, 2003.

BLATCH, G.L.; LÄSSLE, M. The Tetratricopeptide repeat: a structural motif mediating protein-protein interactions. **BioEssays**, New York, v.21, p.:932–939, r 1999.

BLEE, K.; CHOI, J.W.; O'CONNELL, A.P.; JUPE, S.C.; SCHUCH, W.; LEWIS, N.G.; BOLWELL, G.P. Antisense and sense expression of cDNA coding for CYP73A15, a class II cinnamate 4-hydroxylase, leads to a delays and reduced production of lignin in tobacco. **Phytochemistry**, Amsterdam, v.57, p.1159-1166, 2001.

BOCCA, S.N.; MAGIOLI, C.; MANGEON, A.; JUNQUEIRA, R.M.; CARDEAL, V.; MARGIS, R.; SACHETTO-MARTINS, G. Survey of glycine-rich proteins (GRPs) in the *Eucalyptus* expressed sequence tab database (ForEST). **Genetics and Molecular Biology**, Ribeirão Preto, v.28, n.3, p.608-624, 2005.

BOERJAN, W.; RALPH, J.; BAUCHER, M. Lignin biosynthesis. **Annual Review in Plant Biology**, Palo Alto, v.54, p.519-546, 2003.

BOSSINGER, G.E.; LEITCH, M.A. *HEC1*, a Knotted 1-type homeobox gene from *Eucalyptus* and the molecular control of cambial activity. **Journal of Experimental Botany**, Oxford, v.50, p.23-24, 1999.

BRADFORD, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, Burnt Mill, v.72, p.248-254, 1976.

BRONDANI, R.P.V.; BRONDANI, C.; GRATTAPAGLIA, D. Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. **Molecular Genetics and Genomics**, Amsterdam, v.267, n.3, p.338-347, 2002.

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

BRUNNER, A.M.; BUSOV, V.B.; STRAUSS, S.H. Poplar genome sequence: functional genomics in an ecologically dominant plant species. **Trends Plant Science**, Amsterdam, v.9: p.49-56, 2004.

CAGLIARI, T.C.; TIROLI, A.O.; BORGES, J.C.; RAMOS, C.H. Identification and in silico expression pattern analysis of *Eucalyptus* expressed sequencing tags (ESTs) encoding molecular chaperones. **Genetics and Molecular Biology**, Ribeirão Preto, v.28, n.3 (suppl), p.520-528, 2005.

CANDIANO, G.; BRUSCHI, M.; MUSANTE, L.; SANTUCCI, L.; GHIGGERI, G.M.; CARNEMOLLA, B.; ORECCHIA, P.; ZARDI, L.; RIGHETTI, P.G. Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. **Electrophoresis**, New York, v.25, p.1327-1333, 2004.

CARGILE, B.J.; BUNDY, J.L.; GRUNDEN, A.M.; STEPHENSON JR, J.L. Synthesis/degradation ratio mass specrometry for measuring relative dynamic protein turnover. **Analytical Chemistry**, Washington, v.76, p.86-97, 2004.

CARRER, H. Seeing the FORESTs for the trees. **Genetics an Molecular Biology**, Riberião Preto, v.28, n.3 (suppl), p.i-ii, 2005.

CARVALHO, R.M.M.A.; SOARES, T.S.; VALVERDE, S.R. Caracterização do setor florestal: uma abordagem comparativa com outros setores da economia. **Ciência Florestal**, Santa Maria, v.15, n.1, p.105-118, 2005.

CHAFFEY N. Cambium: old challenges – new opportunities. **Trees – structure an function**, Berlin, v.13, p.138-151, 1999.

CHAFFEY, N.; CHOLEWA, E.; REGAN, S.; SUNDBERG, J. Secondary xylem development in *Arabidopsis*: a model for wood formation. **Physiologia Plantarum**, Copenhagen, v.114, p.594-600, 2002.

CHAMRAD DC, KOERTING G, GOBOM J, THIELE H, KLIOSE J, MEYER HE, BLUEGGEL M. Interpretation of mass spectrometry data for high-throughput proteomics. **Analitics Bioanalitics Chemistry**. , v.376, n.7, p.1014-1022, 2003.

CHANG, W.W.P.; HUANG, L.; SHEN, M.; WEBSTER, C.; BURLINGAME, A.L.; ROBERTS, J.K. Patterns of protein synthesis and tolerance of anoxia in root tips of maize seedlings acclimated to a low-oxygen environment, and identification of proteins by mass spectrometry. **Plant Physiology**, Baltimore, v.122, p.295-317, 2000.

CHEN, G.; GHARIB, T.G.; HUANG, C.C.; TAYLOR, J.M.; MISEK, D.E.; KARDIA, S.L.; GIORDANO, T.J.; LANNETTONI, M.D.; ORRINGER, M.B.; HANASH, S.M.; BEER, D.G. Discordant protein and mRNA expression in lung adenocarcinomas. **Molecular & Cellular Proteomics**, Bethesda, v.14, p.304-313, 2002.

CIESLA, W.M.; DIEKMANN, M.; PUTTER, C.A.J. **FAO/IPGRI Technical Guidelines for the Safe Movement of Germplasm**. n. 17. *Eucalyptus* spp. Food and Agriculture Organization of the United Nations, Rome/International Plant Genetic Resources Institute, Rome, 1996. ISBN 92-9043-161-X. Disponível em: <http://www.ipgri.cgiar.org/publications/pubfile+.asp?ID_PUB=406>. Acessado em: 19 fev. 2006.

COLEMAN, H.D.; ELLIS, D.D.; GILBERT, M.; MANSFIELD, S.D. Up-regulation of sucrose synthase and UDP-glucose pyrophosphorylase impacts plant growth and metabolism. **Plant Biotechnology Journal**, Bristol, v.4, p.87-101, 2006.

CORTHALS, G.L.; WASINGER, V.C.; HOCHSTRASSER, D.F.; SANCHEZ, J.C. The dynamic range of protein expression: a challenge for proteomic research. **Electrophoresis**, New York, v.21, p.1104-1115, 2000.

COSTA, P.; PIONNEAU, C.; BAUW, G.; DUBOS, C.; BAHRMANN, N.; KREMER, A.; FRIGERIO, J.M.; PLOMION, C. Separation and characterization of needle and xylem maritime pine proteins. **Electrophoresis**, New York, v.20, p.1098-1108, 1999.

DAMERVAL, C.; DE VIENNE, D.; ZIVY, M.; THIELLEMENT, H. Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins. **Electrophoresis**, New York, v.7, p.52-54, 1986.

DENNY, R.; NEESON, K.; RENNIE, C.; RICHARDSON, K.; LEICESTER, S.; SWAINSTON, N.; WORROLL, J.; YOUNG, P. The use of search workflows in peptide assignment from MS/MS data. **Association of biomolecular resource facilities**, ABRF'02: Biomolecular technologies: tools for discovery in proteomics and genomics, Austin, Texas, March 9-12, 2002.

DHUGGA, K.S.; TIWARI, S.C.; RAY, P.M. A reversibly glycosylated polypeptide (RGP1) possibly involved in plant cell wall synthesis: Purification, gene cloning, and trans-Golgi localization. **Proceedings of National Academy of Science**, Washington, v.94, p.7679-7684, July, 1997.

DIBAX, R.; EISFELD, C.deL.; CUQUEL, F.L.; KOEHLER, H.; QUOIRIN, M. Regeneração de plantas de *Eucalyptus camaldulensis* a partir das explantes cotiledonares. **Scientia Agricola**, Piracicaba, v.62; n.4, p.406-412, July/Aug, 2005.

DOBLIN, M.S.; KUREK, I.; JACOB_WILK, D.; DELMER, D.P. Cellulose biosynthesis in plants: from genes to rosettes. **Plant & Cell Physiology**, Oxford, v.43, p.1407-1420, 2002.

DONALDSON, L.A. Lignification and lignin topochemistry – an ultrastructural view. **Phytochemistry**, Amsterdam, v.57, p.859-873, 2001.

DREGER, M. Subcellular proteomics. **Mass Spectrometry Reviews**, New York, v.22, p.27-56, 2003.

DUNN, M.A.; BROWN, K.; LIGHTOWLERS, R.; HUGHES, M.A. A low temperature responsive gene from barley encodes a protein with single stranded nucleic acid binding activity which is phosphorylated *in vitro*. **Plant Molecular Biology**, New York, v.30, p.947-959, 1996.

EDWARDS, R.; DIXON, D.P.; WALNOT, V. Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. **Trends Pharmacology Science**, Amsterdam, v.5, p.193-198, 2000.

ERIKSSON, M.E.; MORITZ, T. Daylength and spatial expression of a gibberellin 20-oxidase isolated from hybrid aspen (*Populus tremula* L. x *P. tremuloides* Michx.). **Planta**, Amsterdam, v.214, p.920-930, 2002.

EVSTIGNEEVA, Z.G.; SOLOV'EVA, N.A.; SIDEL'NIKOVA, L.I. Methionine sulfoximine and phosphinothricin: a review of their herbicidal activity and effects on glutamine synthetase. **Applied Biochemistry and Microbiology**, Moscow, v.39, p.539-543, 2003.

EZAKI, B.; GARDNER, R.C.; EZAKI, Y.; MATSUMOTO, H. Expression of aluminum-induced genes in transgenic Arabidopsis plants can ameliorate aluminum stress and/or oxidative stress. **Plant Physiology**, Baltimore, v.122, p.657–665, 2000.

EZAKI, B.; YAMAMOTO, Y.; MATSUMOTO, H. Cloning and sequencing of the cDNAs induced by aluminum treatment and Pi starvation in tobacco cultured cells. **Physiologia Plantarum**, Copenhagen, v.93, p.11–18, 1995.

Fenn, J.B.; Mann, M.C.K.; Wong, S.F.; Whitehouse, C.M. Electrospray ionization for mass spectrometry of large biomolecules. **Sciences**, Washington, v.246, p.64-71, 1989.
 FORGAC, M. Structure and Properties of the Vacuolar (H⁺)-ATPases. **Journal of Biological Chemistry**, Bethesda, v.274, n.9, p.12951-12954, May 7, 1989.

FRANK, A.; TANNER, S.; BAFNA, V.; PEVZNER, P. Peptide sequence tags for fast database search in mass-spectrometry. *Journal of proteome research*, *on-line*, 2005.

FREEDMAN, R.B. Protein disulfide isomerase: multiple roles in the modification of nascent secretory proteins. **Cell**, Cambridge, v.57, p.1069-1072, 1989.

FUKUDA, H. Thriceary element differentiation. **The Plant Cell**, Baltimore, v.9, p.1147-1156, 1997.

FUSARO, A.; MANGEON, A.; JUNQUEIRA, R.M.; ROCHA, C.A.B.; COUTINHO, T.C.; MARGIS, R.; SACHETTO-MARTINS, G. Classification, expression pattern and comparative analysis of sugarcane expressed sequences tags (ESTs) encoding glycine-rich proteins (GRPs). **Genetics and Molecular Biology**, Ribeirão Preto, v.24, n.1-4, p.263-273, 2001.

GALVANI, M, HAMDAN, M.; RIGHETTI, P.G. **Electrophoresis**, New York, v.21, p.3684-3692, 2000.

GARDINER, J.C.; HARPER, J.D.I.; WEERAKOON, N.D.; COLLINGS, D.A.; RITCHIE, S.; GILROY, S.; CYR, R.J.; MARC, J. A 90-kD Phospholipase D from Tobacco Binds to Microtubules and the Plasma Membrane. **The Plant Cell**, Baltimore, v.13, p.2143–2158, Sep, 2001.

GION, J.M.; LALANNE, C.; LE PROVOST, G.; FERRY-DUMAZET, H.; PAIVA, J.; CHAUMEIL, P.; FRIGERIO, J.M.; BRACH, J.; BARRE, A.; DE DARUVAR, A.; CLAVEROL, S.; BONNEU, M.; SOMMERER, N.; NEGRONI, L.; PLOMION, C. The proteome of maritime pine wood forming tissue. **Proteomics**, New York, v.5, p.3731-3751, 2005.

GOFF, A.S. Rice as a model for cereal genomics. **Current Opinion in Plant Biology**, Amsterdam, v.2, p.86-89, 1999.

GOICOECHEA, M.; LACOMBE, E.; LEGAY, S.; MIHALJEVIC, S.; RECH, P.; JAUNEAU, A.; LAPIERRE, C.; et al.. *EgMYB2*, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. **The Plant Journal**, Oxford, v. 43, p.553-567, 2005.

GONZÁLEZ, E.R.; DE ANDRADE, A.; BERTOLO, A.L.; LACERDA, G.C.; CARNEIRO, R.T.; DEFÁVARI, 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, p.97–102, 2002.

GÖRG, A.; OBERMAIER, C.; BOGUTH, G.; WEISS, W. Recent developments in two-dimensional gel electrophoresis with immobilized pH gradients: Wide pH gradients up to pH 12, longer separation distances and simplified procedures. **Electrophoresis**, New York, v.20, p.712-717, 1999.

GUY, C.L.; HASKELL, D.; LI, Q.-B. Association of proteins with molecular chaperones at low temperature: evidence for the cold labile proteins in spinach. **Cryobiology**, New York, v.36, p.301–314, 1998.

GUYOMARCH, S.; VERNOUX, T.; TRAAS, J.; ZHOU, D.X.; DELARUE, M. MGOUN3, an *Arabidopsis* gene with Tetratricopeptide-Repeat-related motifs, regulates meristem cellular organization. **Journal of Experimental Botany**, Oxford, v.55, n.397, p.673-684, 2004.

GYGI, S.P.; RIST, B.; GERBER, B.; TURECEK, et al.. **Natural Biotechnology**, v.17, p.994-999, 1999.

GYGI, S.P.; ROCHON, Y.; FRANZA, B.R.; AEBERSOLD, R. Correlation between protein and mRNA abundance in yeast. **Molecular and Cellular Biology**, Washington, v.19, n.3, p.1720-30, 1999b.

HAWKINS, S.; BOUDET, A.; GRIMA-PETTENATI, J. Characterisation of caffeic acid O-methyltransferase and cinnamyl alcohol dehydrogenase gene expression patterns by *in situ* hybridisation in *Eucalyptus gunnii* Hook plantlets. **Plant Science**, Amsterdam, v.164, n.2, p.165-173, Feb, 2003.

HAYASHI, H.; FUKUDA, A.; SUZUI, N.; FUJIMAKI, S. Proteins in the sieve element-companion cell complexes: their detection, localization and possible functions. **Australian Journal of Plant Physiology**, Melbourne, v.27, p.489-496, 2000.

HAYNES, L.P., BARNARD, R.J., MORGAN, A. AND BURGOYNE, R.D. Stimulation of NSF ATPase activity during t-SNARE priming. **FEBS Letter**, v.436, p.1-5, 1998.

HENZEL, W.J.; STULTS, J.T.; WATANABE, C. **Proceedings of the Third Symposium of the Protein Society**, Seattle, WA, 1989.

HERBERT, B.; GALVANI, M.; HAMDAN, M.; OLIVIERI, E.; MACCARTHY, J.; PEDERSEN, S.; RIGHETTI, P.G. Reduction and alkylation of proteins in preparation of two-dimensional map analysis: why, when and how?. **Electrophoresis**, New York, v.22, p.2046-2057, 2001.

HERRMANN, K.M. The shikimate pathway: Early steps in the biosynthesis of aromatic compounds. **The Plant Cell**, Baltimore, v.7, p.907-919, 1995.

HERRMANN, K.M.; WEAVER, L.M. The shikimate pathway. **Annual Review of Plant Physiology and Plant Molecular Biology**, Palo Alto, v.50, p.473-503, 1999.

HERVE, P.; JAUNEAU, A.; PAQUES, M.; MARIEN, J.N.; BOUDET, A.M.; TEULIERES C. A procedure for shoot organogenesis in vitro from leaves and nodes of an elite *Eucalyptus gunnii* clone: comparative histology. **Plant Science**, Amsterdam, v.161, n.4, p.645-653, SEP, 2001.

HIMMELSPACH, R.; NICK, P.; SCHIFER, E.; EHMANN, B. Developmental and light-dependent changes of the cytosolic chaperonin containing TCP-1 (CCT) subunits in maize seedlings, and the localization in coleoptiles. **The Plant Journal**, Oxford, v.12, p.1299-1310, 1997.

HOOPER, N.I.; TISDALE, M.J.; THORNALLEY, P.J. Modification of the glyoxalase system in human HL60 promyelocytic leukemia cells during differentiation to neutrophils *in vitro*. **Biochimica et Biophysica Acta**, Amsterdam, v.966, p.362–369, 1988.

HUANG, Y.; BLAKELEY, S.D.; MCALEESE, S.M.; FOTHERGILL-GILMORE, L.A.; DENNIS, D.T. Higher plant cofactor-independent phosphoglyceromutase: purification, molecular characterization and expression. **Plant Molecular Biology**, New York, v.23, p.1039–1053, 1993.

HUNT, S.M.N.; THOMAS, M.R.; SEBASTIAN L.T.; SUSANNE, K.; PEDERSEN, S.K.; HARCOURT, R.L.; SLOANE, A.J.; WILKINS, M.R. Optimal replication and the importance of experimental design for gel-based quantitative proteomics. **Journal of Proteome Research**, Boston, v.4, p.809-819, 2005.

HURKMAN, T.; TANAKA, C.K. Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. **Plant Physiology**, Baltimore, v.81, p.802-806, 1986.

IMIN, N.; KERIM, T.; ROLFE, B.G.; WEINMAN, J.J. Effect of early cold stress on the maturation of rice anthers. **Proteomics**, New York, v.4, n.7, p.1873-1882, 2004.

INSTITUO NACIONAL DE PESQUISAS ESPACIAIS (INPE). Relatório de acompanhamento anual 2004. Disponível em: http://www.inpe.br/gestão/arquivo/RelatorioAcompanhamentoAnual2004_INPE.pdf Acesso em 11 fev 2006.

JACOBS, W.P. The role of auxin in differentiation of xylem around a wound. **American Journal of Botany**, St. Louis, v.39, p.301-309, 1952.

JARILLO, J.A.; CAPEL, J.; LEYVA, A.; MARTINEZ-ZAPATER, J.M.; SALINAS, J. Two related low-temperature-inducible genes of *Arabidopsis* encode proteins showing high homology to 14-3-3 proteins, a family of putative kinase regulators. **Plant Molecular Biology**, New York, v.25, n.693-704, 1994.

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

JOUANIN, L.; GOUJON, T.; DE NADAI, V.; MARTIN, M.T.; MILA, I.; VALLET, C.; POLLET, B.; YOSHINAGA, A.; CHABBERT, B.; PETIT-CONIL, M.; LAPIERRE, C. Lignification in transgenic *Poplars* with extremely reduced caffeic acid O-methyltransferase activity. **Plant Physiology**, Baltimore, v.123, p.1363-1373, 2000.

JUNGHANS, D.T.; ALFENAS, A.C.; BROMMONSCHENKEL, S.H.; ODA S.; MELLO, E.J.; GRATTAPAGLIA, D. Resistance to rust (*Puccinia psidii* Winter) in *Eucalyptus*: mode of inheritance and mapping of a major gene with RAPD markers. **Theoretical and Applied Genetics**, Berlin, v.108, n.1, p.175-180, 2003.

KALLURI, U.C.; JOSHI C.P. Differential expression patterns of two cellulose syntase genes are associated wit primary and secondary cell wall development in aspen trees. **Planta**, Amsterdam, v.220, p.47-55, 2004.

KALLURI, U.C.; JOSHI CP. Isolation and characterization of a new, full-length cellulose synthase cDNA, PrrCesA5 from developing xylem of aspen trees. **Journal of experimental botany**, Oxford, v.54, n.390, p.2187-2188, 2003.

KARAS, M.; HILLENKAMP, F. **Analytical Chemistry**, Washington, v.60, p.2299-2301, 1988.

KEGG. Disponível em: <<http://www.genome.jp/kegg/>>. Acessado em: fev. 2006.

KERK, D.; BULGRIEN, J.; SMITH, D.W.; GRIBSKOV, M. Arabidopsis Proteins Containing Similarity to the Universal Stress Protein Domain of Bacteria. **Plant Physiology**, Baltimore, v.131, p.1209–1219, 2003.

KHOUDOLI, G.A.; PORTER, I.M.; BLOW, J.J.; SWEDLOW, J.R. Optimization of the two-dimensional gel electrophoresis protocol using the Taguchi approach. **Proteome Science**, v.2, 12., 2004.

KIRST, M.; CAROCHA, V.; MARQUES, C.M.; VAN ZYL, L.; MYBURG, A.; SEDEREOFF, R. Individual genotypes of eucalypts show striking variation in xylem gene expresssion. In: **Plant, Animal & Microbe Genomes X Conference**, Abstract P548, San Diego, CA, Jan, 12-16, 2002. Disponível em: <http://www.intl-pag.org/10/abstracts/PAGX_P548.html>. Acesso em: 25 jan. 2006.

KLECZKOWSKI, L.A. Glucose activation and metabolism through UDP-glucose pyrophosphorylase in plants. **Phytochemistry**, Amsterdam, v.37, p.1507-1515, 1994.

KNOBLAUCH, M.; van BEL, A.J.E. Sieve Tubes in Action. **The Plant Cell**, Baltimore, v.10, p.35–50, 1998.

KOPCZAK, S.D.; HAAS, N.A.; HUSSEY, P.J.; SILFLOW, C.D.; SNUSTAD, D.P. The small genome of *Arabidopsis* contains at least six expressed α -tubulin genes. **The Plant Cell**, Baltimore, v.4, p.539-547, 1992.

KRUGER, C.; BERKOWITZ, O.; STEPHAN, U.W.; HELL, R. A Metal-binding Member of the Late Embryogenesis Abundant Protein Family Transports Iron in the Phloem of *Ricinus communis* L. **The Journal Of Biological Chemistry**, Bethesda, v.277, n.28, p.25062–25069, 2002.

LACOMBE, E.; HAWKINS, S.; DOORSSELAERE, J.V.; 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, p.429-441, 1997.

LAEMMLI, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4 **Nature**, London, v.227, p.680-685, 1970.

LEE, J-Y; LEE, D.H. Use of Serial Analysis of Gene Expression Technology to Reveal Changes in Gene Expression in *Arabidopsis* Pollen Undergoing Cold Stress. **Plant Physiology**, Baltimore, v.132, p.517–529, 2003.

LEV-YADUN, S.; RONALD SEDEROFF, R. Pines as Model Gymnosperms to Study Evolution, Wood Formation, and Perennial Growth. **Journal of Plant Growth Regulation** 19:290–305, (2000).

LI, Z.S.; ALFENITO, M.; REA, P.A.; WALBOT, V.; DIXON, R.A. Vacuolar uptake of the phytoalexin medicarpin by the glutathione conjugate pump. **Phytochemistry**, Amsterdam, v.45, p.689-693, 1997.

LIPPERT, D.; ZHUANG, J.; RALPH, S.; ELLIS, D.E.; GILBERT, M.; OLAFSON, R.; RITLAND, K.; ELLIS, B.; DOUGLAS, C.J.; BOHLMANN, J. Proteome analysis of early somatic embryogenesis in *Picea glauca*. **Proteomics**, New York, v.5, p.461-473, 2005.

LOOPSTRA, C.A.; SEDEROFF, R.R. Xylem-specific gene expression in loblolly pine. **Plant Molecular Biology**, New York, v.27, p.277-291, 1995.

MAJOUL T, BANCEL E, TRIBOI E, BEM HAMIDA, J.; BRANLARD G. Proteomic analysis of the effect of heat stress on hexaploid wheat grain: Characterization of heat-responsive proteins from non-prolamins fraction. **Proteomics**, New York, v.4, p.505-513, 2004.

MANN, M.; HENDRICKSON, R.C.; PANDEY, A. Analysis of proteins and proteomes by mass spectrometry. *Annual Review Biochemistry*, v.70, p.437-473, 2001.

MARRS, K.A. The function and regulation of glutathione S-transferases in plants. **Annual Review of Plant Physiology and Plant Molecular Biology**, Palo Alto, v.47, p.127-158, 1996.

MARU, Y.; WHITE, O.N. The BCR gene encodes a novel serine/threonine kinase activity within a single exon. **Cell**, Cambridge, v.67, p.459-468, 1992.

MATTHEW, G.D.; GOULD, R.M.; VARDIMON, L. A single glutamine synthetase gene produces tissue-specific subcellular localization by alternative splicing. **FEBS Letters**, Oxford, v.579, p.5527–5534, 2005.

MÉCHIN V., BALLIAU T., CHÂTEAU-JOUBERT S., DAVANTURE, M.; LANGELLA, O.; NEGRONI, L.; PRIOUL, J.L.; THEEVNOT, C.; ZIVY, M.; DAMERVAL, C. A two-dimensional proteome map of maize endosperm. **Phytochemistry**, Amsterdam, v.65, p.1609-1618, 2004.

MECHIN, V.; CONSOLI, L.; LE GUILLOUX, M.; DAMERVAL, C. An efficient solubilization buffer for plant proteins focused in immobilized pH gradients. **Proteomics**, New York, v.3, p.1299-1302, 2003.

MERKLE, S.A.; DEAN, J.F.D. Forest tree biotechnology. **Current Opinions in Biotechnology**, Amsterdam, v.11, p.298–302; 2000.

MEYER, K.; SHIRLEY, A.M.; CUSUMANO, J.C.; BELL-LELONG, D.A.; CHAPPLE, C. Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. **Proceedings of the National Academy of Sciences**, Washington,, v.95, p.6619-6623, 1998.

MOFFATT BA, STEVENS YY, ALLEN MS, SNIDER, J.D.; PEREIRA, L.A.; TODOROVA, M.I.; SUMMERS, P.S.; WERETILNYK, E.A.; MARTIN-MCCAFFREY, L.; WAGNER, C. Adenosine Kinase Deficiency Is Associated with Developmental Abnormalities and Reduced Transmethylation. **Plant Physiology**, Baltimore, v.128, p.812–821, 2002.

NEUHOFF, V.; STAMM, R.; PARDOWITZ, I.; AROLD, N. ET AL. **Electrophoresis**, v. 11, p.101-117, 1990.

NICK, P.; HEUING, A.; EHMANN, B. Plant chaperonins: a role in microtubule-dependent wall formation? **Protoplasma**, Leipzig, v.211, p.234-244, 2000.

NYSTROM, T.; NEIDHARDT, F.C. Expression and role of the universal stress protein, UspA, of *Escherichia coli* during growth arrest. **Molecular Microbiology**, Oxford, v.11, p.537–544, 1994.

O'FARREL. High Resolution Two-Dimensional Electrophoresis of Proteins. **The Journal of Biological Chemistry**, Bethesda, v.250, n.10, p. 4007-4021, 1975.

PARKER, K.C.; GARRELS, J.I.; HINES, W.; BUTLER, E.M.; MCKEE, A.H.; PATTERSON, D.; MARTIN, S. Identification of yeast proteins from two-dimensional gels: working out *spot* cross-contamination. **Electrophoresis**, New York, v.19, 1920–1932, 1998.

PAUL, W.; STEINWEDEL, H. Ein neues Massenspektrometer ohne Magnetfeld. **Zeitschrift für Naturforschung**, Weisbaden, v.8A, p.448-450, 1953.

PAULUS, C.; KOLLNER, B.; JACOBSEN, H.-J. Physiological and biochemical characterization of glyoxalase I: a general marker for cell proliferation from a soybean cell suspension. **Planta**, Amsterdam, v.189, p.561–566, 1993.

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**, Oxford, 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**, New York, v.55, n.2, p.263-280, MAY, 2004.

PEAR, J.R.; KAWAGOE, Y.; SCHRECKENGOST, W.E.; DELMER, D.P.; STALKER, D.M. Higher plants contain homologs of the bacterial *celA* genes encoding the catalytic subunit of cellulose synthase. **Proceedings of the National Academy of Sciences**, Washington, v.93, p.12637-12642, 1996.

PEART, J.R.; SHIRASU, K. To degrade or not to degrade: the emerging question in plant disease resistance. **Physiology and Molecular Plant Pathology**, Amsterdam, v.61, p.73-76, 2002.

PERKINS, D.N.; PAPPIN, D.J.; CREASY, D.M.; COTTRELL, J.S. Probability-based protein identification by searching sequences databases using mass spectrometry data. **Electrophoresis**, New York, v.20, n.18, p.6551-6567, 1999.

PETRASEK, J.; FREUDENREICH, A.; HEUING, A.; OPATRNY, Z.; NICK, P. Heat-shock protein 90 is associated with microtubules in tobacco cells. **Protoplasma**, Leipzig, v.202, p.161-174, 1998.

PILATE G, DÉJARDIN A, LAURANS F.; LEPLÉ, J.C. Tension wood as a model for functional genomics of wood formation. **New Phytologist**, Oxford, v.164, p.63-72, 2003.

PILATE, G.; GUINEY, E.; HOLT, K.; PETIT-CONIL, M.; LAPIERRE, C.; LEPLÉ, J.C.; POLLET, B.; MILA, I.; WEBSTER, 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**, London, v.20, p.607-612, 2002.

PIQUEMAL, J.; LAPIERRE, C.; MYTON, K.; O'CONNEL, A.; SCHUCH, W.; GRIMA-PETTENATI, J.; BOUDET, A.M. Down-regulation of cinnamoyl-CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. **The Plant Journal**, Oxford, v.13, p.71-83, 1998.

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

PLOMION, C.; PIONNEAU, C.; BAILLÈRES, H. Analysis of protein expression along the normal to tensin wood gradient in *Eucalyptus gunnii*. **Holzforschung**, Berlin, v.57, p.353-358, 2003.

POKE, F.S.; VAILLANCOURT, R.E.; ELLIOTT, R.C.; REID, J.B. Sequence variation in two lignin biosynthesis genes, cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase 2 (CAD2). **Molecular Breeding**, Amsterdam, v.12, n.2, p.107-118, SEP, 2003.

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

RANIERI, A.; CASTAGNA, A.; BINI, L.; SOLDATINI, G.F. Two-dimensional electrophoresis of thylakoid protein patterns in two wheat cultivars with different sensitivity to sulfur dioxide. **Electrophoresis**, New York, v.16, p.1301-1304, 1995.

RAVEN, J.A. Long-term functioning of enucleate sieve elements: possible mechanisms of damage avoidance and damage repair. **Plant, Cell & Environment**, New York, v.14, p.139-146, 1991.

READ, S.M.; BASIC, T. Prime time for cellulose. **Science**, Washington, v.295, p.59-60, 2002.

REDE BRASILEIRA DE PESQUISA DO GENOMA DO *EUCALYPTUS* (GENOLYPTUS). Disponível em: <www.genolyptus.ucb.br>.

RENSING, K.H.; SAMUELS, A.L.; SAVIDGE R.A. Ultrastructure of vascular cambial cell cytokinesis in pine seedlings preserved by cryofixation and substitution. **Protoplasma**, Leipzig, v.220, p.39-49, 2000.

RISON, S.C.G.; HODGMAN, T.C.; THORNTON, J.M. Comparison of functional annotation schemes for genomes. **Funct integr genomics**, v.1, p.56-69, 2000.

RITTINGER K, BUDMAN J, XU J; VOLINIA, S.; CANTLEY, L.C.; SMERDON, S.J.; GAMBLIN, S.J.; YAFFE, M.B. Structural analysis of 14-3-3 phosphopeptide complexes identifies a dual role for the nuclear export signal of 14-3-3 in ligand binding. **Molecular Cell**, Cambridge, v.4, p.153-166, 1999.

ROGERS, M.; GRAHAM, J.; TONGE, R.P. Using statistical image models for objective evaluation of *spot* detection in two-dimensional gels. **Proteomics**, New York, v.3, p.879-886, 2003.

ROTHMAN, J.E. Mechanisms of intracellular protein transport. **Nature**, London, v.372, p.55–63, 1994.

RUBINFELD, A.; KEREN-LEHRER, T.; HADAS, G.; SMILANSKY, Z.. Hierarchical analysis of large-scale two-dimensional gel electrophoresis experiments. **Proteomics**, New York, v.3, p.1930-1935, 2003.

RUSSEL, S.A.; OLD, W.; RESING, K.A.; HUNTER, L. Proteomic informatics. **International review of neurobiology**, vol.61, 2004. Disponível em: <<http://compbio.uchsc.edu/Hunter/proteomic-informatics.pdf>>. Acessado em: jun. 2004.

SACHETTO-MARTINS, G.; FRANCO, L.O.; DE OLIVEIRA, D.E. Plant glycine-rich proteins: a family or just proteins with a common motif? **Biochimica et Biophysica Acta**, Amsterdam, v.1492, p.1-14, 2000.

SAGI, G.; KATZ, A.; GENOUNE-GELBART, D; EPEL, B.L. Class 1 Reversibly Glycosylated Polypeptides Are Plasmodesmal-Associated Proteins Delivered to Plasmodesmata via the Golgi apparatus. **The Plant Cell**, Baltimore, v.17, p.1788–1800, June, 2005.

SAKUTH, T.; SCHOBERT, C.; PECSVARADI, A.; EICHHOLZ, A.; KOMOR, E.; ORLICH, G. Specific proteins in the sieve tube exudate of *Ricinus communis* L. seedlings: separation, characterization and *in-vivo* labeling. **Planta**, Amsterdam, v.191, p.207–213, 1993.

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

SANDERFOOT, A.A.; RAIKHEL, N.V. The Specificity of Vesicle Trafficking: Coat Proteins and SNAREs. **The Plant Cell**, Baltimore, v.11, p.629–641, April, 1999.

SANTONI, V.; MOLLOY, M.; RABILLOUD, T. Membrane proteins and proteomics: un amour impossible? *Electrophoresis*, New York, v.21, p.1054-1070, 2000.

SAPPL, P.G.; OÑATE-SANCHEZ, L.; SINGH, K.B.; MILLAR, A.H. Proteomic analysis of glutathione S-transferases of *Arabidopsis thaliana* reveals differential salicylic acid-induced expression of the plant-specific phi and tau classes. **Plant Molecular Biology**, New York, 54, 205–219, 2004.

SARAVAN, R.S.; ROSE, J.K.C. A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *Proteomics*, New York, v.4, p.2522-2532, 2004.

SARTORETTO, L.M.; CID, L.P.B.; BRASILEIRO, A.C.M. Biolistic transformation of *Eucalyptus grandis* x *E. urophylla* callus. **Functional Plant Biology**, Collingwood, v.29, n.8, p.917-924, 2002.

SATO, S., K. HORIKIRI, K. MIYASHITA, N. ISHIGE, T. HIBINO & T. KAWAZU, 2001. Analysis of wood development with a genomic approach: Eucalyptus ESTs and TAC genomic library, Abstract P_01_59, in **Plant & Animal Genome IX Conference**, San Diego, CA, Jan. 13–17.

SAUTER, M.; RZEWUSKI, G.; MARWEDEL, T.; LORBIECKE, R. The novel ethylene-regulated gene OsUsp1 from rice encodes a member of a plant protein family related to prokaryotic universal stress proteins. **Journal of Experimental Botany**, Oxford, v.53, n.379, p.2325-2331, 2002.

SAVIDGE, R.A.; BARNETT, J.R. Protoplasmic changes in cambial cells induced by a tracheid-differentiation factor from pine needles. **Journal of Experimental Botany**, Oxford, v.44, p.395-404, 1993.

SCHLOSSER, A. The smaller, the better – Peptide Analysis with nanoLC-MS/MS Using 25-µm-i.d. Columns. **Abstract WPJ163**, 2004. Disponível em: <<http://www.charite.de/immunologie/research/MSLab/downloads.htm>>. Acesso em: 20 ago. 2005.

SCHOBERT, C.; GOTTSCHALK, M.; KOVAR, D.R.; STAIGER, C.J.; YOO, B.C.; LUCAS, W.J. Characterization of *Ricinus communis* phloem profilin, RcPRO1. **Plant Molecular Biology**, New York, v.42, p.719–730, 2000.

SEKI, M., NARUSAKA, M., ABE, H., KASUGA, M.; YAMAGUCHI-SHINAZAKI, K.; CARNINCI, P.; HAYASHIZAKI, Y.; SHINOZAKI, K. Monitoring the Expression Pattern of 1300 *Arabidopsis* Genes under Drought and Cold Stresses by Using a Full-Length cDNA Microarray. **The Plant Cell**, Baltimore, v.13, p.61–72, 2001.

SHEVCHENKO, A.; LOBODA, A.; SHEVCHENKO, A.; ENS, W.; STANDING, K.G. MALDI Quadrupole time-of-flight mass spectrometry: a powerful tool for proteomic research. **Analytical Chemistry**, Washington, v.72, p.2132-2141, 2000.

SHEVCHENKO, A.; SUNYAEV, S.; LOBODA, A.; SHEVCHENKO, A.; BORK, P.; ENS, W.; STANDING, K.G. Charting the Proteomes of Organisms with Unsequenced Genomes by MALDI-QTOF Mass Spectrometry and BLAST Homology Searching. **Analytical Chemistry**, Washington, v. 73, p.1917-1926, 2001.

SHEVCHENKO, A.; WILM, M.; VORM, O.; MANN, M. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. **Analytical Chemistry**, Washington, v.68, p.850-858, 1996.

SHORROSH, B.S.; SUBRAMANIAM, J.; SCHUBERT, K.R.; DIXON, R.A. Expression and localization of plant protein disulfide Isomerase. **Plant Physiology**, Baltimore, v.103, p.719-726, 1993.

SKYLAS, D.J.; VAN DYK, D.; WRIGLEY, C.W. Proteomics of wheat grain. *Journal of cereal sciences*, v.41, p.165-179, 2005.

SNUSTAD, D.P.; HAAS, N.A.; KOPCZAK, S.D.; SILFLOW, C.D. The small genome of *Arabidopsis* contains at least nine expressed β -tubulin genes. **The Plant Cell**, Baltimore, v.4, p.549-556, 1992.

SNYDER, A.P.. Interpretation peptide and protein mass spectra. In:_____. **Interpreting protein mass spectra: a comprehensive resource**. New York: Oxford University Press, 2000b. p.205.

SOCIEDADE BRASILEIRA DE SILVICULTURA (SBS). **O setor florestal brasileiro: fatos e números do setor florestal brasileiro**. São Paulo, 2005.

SOLOMON, P.S.; TAN, K.C.; OLIVER, R.P. The nutrient supply of pathogenic fungi; a fertile field of study. **Molecular Plant Pathology**, Oxford, v. 4, p.203–210, 2003.

STAIGER, C.J.; GIBBON, B.C.; KOVAR, D.R.; ZONIA, L.E. Profilin and actin-depolymerizing factor: modulators of actin organization in plants. **Trends in Plant Science**, Cambridge, v.2, p.275–281, 1997.

STASOLLA, C.; EGERSTDOTTER, U.; SCOTT, J.; KADLA, J.; SEDEROFF, R.; VAN ZYL, L. Lignin composition on wild type and cinnamyl alcohol dehydrogenase deficient mutant cultured cells of *Pinus taeda*. **Plant Physiology and Biochemistry**, Amsterdam, v.41, p.439-445, 2003.

STERKY, F.; BHALERAO, R.R.; UNNEBERG, P.; SEGERMAN, B.; NILSSON, P.; BRUNNER, A.M.; CHARBONNEL-CAMPAA, L.; LINDVALL, J.J.; TANDRE, K.; STRAUSS, S.H.; SUNDBERG, B.; GUSTAFSSON, P.; UHLÉN, M.; BHALERAO, R. >P.; NILSSON, O.; SANDBERG, G.; KARLSSON, J.; LUNDEBERG, J.; JANSSEN, S. A Populus EST resource for plant functional genomics. **Proceedings of the National Academy of Sciences**, Washington, v.101, n.38, 13951-13956, 2004.

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.; LUNDEBERG, J. Gene discovery in the wood forming tissues of poplar: analysis of 5,692 expressed sequence tags. **Proceedings of the National Academy of Sciences**, Washington, v.95, p.13330-13335, 1998.

STERNLICHT, H.; FARR, G.W.; STERNLICHT, M.L.; DRISCOLL, J.K.; WILLISON, K.; YAFFE, M.B. The t-complex polypeptide 1 complex is a chaperonin for tubulin and actin in vivo. **Proceedings of the National Academy of Sciences**, Washington, v.90, n.20, p.9422-9426, Oct, 1993.

STEVENS, T. H.; FORGAC, M. Structure, function and regulation of the vacuolar (H⁺)-ATPase) **Annual Review of Cell and Developmental Biology**, Palo Alto, v.13, p.779–808, 1997.

THAMARUS, K.A.; GROOM, K.; BRADLEY, A.D.; RAYMOND, C.A.; MORAN, G.F. Identification of quantitative trait loci for wood and fibre properties in two full-sib pedigrees of *Eucalyptus globulus*. **Theoretical and Applied Genetics**, Berlin, v.109, n.4, p.856-864, 2004.

THE *ARABIDOPSIS* GENOME INITIATIVE. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. **Nature**, London, v.408, p.796-815, 2000.

TIEMANN, K.; INZÉ, D.; VAN MONTAGU, M.; BARZ, W. Pterocarpan phytoalexin biosynthesis in elicitor-challenged chickpea (*Cicer arietinum* L.) cell cultures: purification, characterization and cDNA cloning of NADPH: iso⁻avone oxidoreductase. **European Journal of Biochemistry**, Berlin, v.200, p.751±758, 1991.

TOURNIER, V.; GRAT, S.; MARQUE, C.; EL KAYAL, W.; PENCHEL, R.; DE ANDRADADE, G.; BOUDET, A.M.; TEULIERES, C. An efficient procedure to stably introduce genes into an economically important pulp tree (*Eucalyptus grandis* x *Eucalyptus urophylla*). **Transgenic Research**, London, v.12, n.4, p.403-411, 2003.

TSUGEKI, R.; MORI, H.; NISHIMURA, M. Purification, cDNA cloning and Northern-blot analysis of mitochondrial chaperonin 60 from pumpkin cotyledons. **European Journal of Biochemistry**, Berlin, v. 209, p.453-458, Oct 1, 1992.

TSVETKOVA, N.M.; HORVÁTH, I.; TOROK, Z.; WOLKERS, W.F.; BALOGI, Z.; SHIGAPOVA, N.; CROWE, L.M.; TABLIN, F.; VIERLING, E.; CROWE, J.H.; VIGH, L. Small heatshock proteins regulate membrane lipid polymorphism. **Proceedings of National Academy of Science**, Washington, v.99, p.13504–13509, 2002.

URSIC, D.; CULBERSON, M. R. The yeast homolog to mouse Tcp-1 affects microtubule-mediated processes. **Molecular and Cellular Biology**, Washington, v.11, p.2629-2640, 1991.

VAILLANCOURT, R.E.; POTTS, B.M.; MANSON, A.; REID, J.B. Detection of QTLs in a *Eucalyptus gunnii* x *E. globulus* F2 using a RAPD linkage map. **Proceedings of the International Wood Biotechnology Symposium**, Tokyo, 1994, 63-64, 1994.

VALERIO, L.; CARTER, D.; RODRIGUES, J.C.; TOURNIER, V.; GOMINHO, J.; MARQUE, C.; BOUDET, A.M.; MAUNDERS, M.; PEREIRA, H.; TEULIÈRES, C. Down-regulation of cinnamyl Alcohol Dehydrogenase, a lignification enzyme, in *Eucalyptus camaldulensis*. **Molecular Breeding**, v.12, n.2, p.157-167, 2003.

VAN MONTFORT, R.; SLINGSBY, C.; VIERLING, E. Structure and function of the small heat shock protein/alpha-crystallin family of molecular chaperones. **Advances in Protein Chemistry**, San Diego, v.59, p.105–156, 2001.

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

VANDER MIJNSBRUGGE, K.; VAN MONTAGU, M.; INZE, D.; BOERJAN, W. Tissue-specific expression conferred by the S-adenosyl-L-methionine synthetase promoter of *Arabidopsis thaliana* in transgenic poplar. **Plant & Cell Physiology**, Oxford, v.37, n.8, p.1108-15, 1996.

VENSEL, W.H.; TANAKA, C.K.; CAI, N.; WONG, J.H.; BUCHANAN, B.B.; HURKMAN, W.J. Developmental changes in the metabolic protein profiles of wheat endosperm. **Proteomics**, New York, v.5, p.1594-1611, 2005.

WAGNER, U.; EDWARDS, R.; DIXON, D.P.; MAUCH, F. Probing the diversity of the *Arabidopsis* glutathione S-transferase gene family. **Plant Molecular Biology**, New York, v.49, p.515–532, 2002.

WANG X. Multiple forms of phospholipase D in plants: the gene family, catalytic and regulatory properties, and cellular functions. **Progress in Lipid Research**, Oxford, v.39, p.109–149, 2000.

WANG, H.; HAO, L.; SHUNG, C.Y.; SUNTER, G.; BISARO, D.M. Adenosine kinase is inactivated by geminivirus AL2 and L2 proteins. **The Plant Cell**, Baltimore, v.15, p.3020–3032, 2003.

WATSON, M.P.; ASIRVATHAM, V.S.; WANG, L.; SUMNER, L.W. Mapping the proteome of barrel medic (*Medicago truncatula*). **Plant Physiology**, Baltimore, v.131, p.1104-1123, 2003.

WILEY, W.C.; MACLAREN, I.H. Time-of-flight mass spectrometer with improved resolution. **The Review of Scientific Instruments**, New York, v.26, p.1150, 1995.

WILKINS, M.R.; SANCHEZ, J.C.; GOOLEY, A.A.; APPEL, R.D.; HUMPHERY-SMITH, I.; HOCHSTRASSER, D.F.; WILLIAMS, K.L. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. **Biotechnology & Genetic Engineering Reviews**, New castle, v.13, p.19-50, 1995.

Wilkinson, J.M. Fragmentation of polypeptides by enzymic methods. In: _____. Practical protein chemistry: a handbook. New York: John Wiley and Sons, 1989.

XU D, DUAN X, WANG B, HONG, B.; HO, T.H.D.; WU, R. Expression of a late embryogenesis abundant protein gene, *HVA7*, from barley confers tolerance to water deficit and salt stress in transgenic rice. **Plant Physiology**, Baltimore, v.110, p.249-257, 1996.

XUE, M.; ZHANG, B. Do SNARE proteins confer specificity for vesicle fusion? **Proceedings of the National Academy of Sciences**, Washington, v.99, n.21, p.13359–13361, 2002.

YAMAMOTO, Y.; KOBAYASHI, Y.; MATSUMOTO, H. Lipid peroxidation is an early symptom triggered by aluminum, but not the primary cause of elongation inhibition in pea roots. **Plant Physiology**, Baltimore, v.125, p.199–208, 2001.

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

Yates, J.R.3rd; ENG, J.R.; McCORMACK, A.L.; SCHIELTZ, D. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. **Analytical Chemistry**, Washington, v.67, n.2, p.1426-1436, 1995.

YE, Z.H.; KNEUSEL, R.E.; MATERN, U.; VARNER, J.E. An alternative methylation pathway in lignin biosynthesis in *Zinnia*. **The Plant Cell**, Baltimore, v.6, p.1427-1439, 1994.

ZHOU, R.; KROCZYNSKA, B.; HAYMAN, G.T.; MIERNYK, J.A. AtJ2, an *Arabidopsis* homolog of *Escherichia coli* dnaJ. **Plant Physiology**, Baltimore, v.108, p.821-822, 1995.