

Universidade Federal do Rio Grande do Sul
Centro de Biotecnologia
Programa de Pós-Graduação em Biologia Celular e Molecular

**SELEÇÃO E AVALIAÇÃO DE GENES DE REFERÊNCIA
PARA ESTUDOS DE EXPRESSÃO GÊNICA EM *Eucalyptus***

Fernanda Macedo Bastolla

Orientadora: Célia Regina R.S. Carlini

Co-orientador: Giancarlo Pasquali

Dissertação submetida ao Programa de Pós-Graduação em
Biologia Celular e Molecular da UFRGS como requisito parcial
para a obtenção do grau de Mestre

Porto Alegre
Outubro, 2007

A todos aqueles que me ajudaram a transformar um limão em caipirinha

“I get by with a little help from my friends”

Agradecimentos

Aos meus pais, dos quais herdei a persistência e a personalidade, pelo apoio e pelos sábios e oportunos conselhos.

Ao meu amor, Giuli, pelas tantas idas e vindas para o Campus, força, carinho, colo e paciência (“Will you still need me, will you still feed me when I'm sixty-four?”).

Ao meu orientador, Gian, pelas grandes oportunidades, confiança, ensinamentos e pelo Projeto Genolyptus.

À Marina, siamesa, pela ajuda incondicional, grande força e amizade nos (muitos) momentos complicados e alegres deste trabalho (“Nós conseguimos Cérebro!”).

Ao Prof. Rogério Margis, que além de participar da comissão de acompanhamento foi um grande colaborador neste trabalho, agradeço pelas sugestões, atenção e paciência.

À Profa. Célia pela orientação, sugestões e conselhos.

Ao Daniel Lorenzini pelo grande auxílio, idéias e discussões, e a sempre incansável e alegre Débora, por toda ajuda e ensinamentos.

À Fê Sperb, pela ajuda nas horas difíceis, tanto profissionais quanto pessoais, e pelos conselhos que sempre me trazem à tona, muitas vezes para fora do Mundo de Nikita.

À Dani, pela grande força, pelo amparo também na vida pessoal e pela agradável e indispensável companhia nas horas, horas e horas de estudo.

Ao Fabiano, pelo incentivo e por toda ajuda técnica, psicológica e filosófica (“Thank you dear!”).

Ao Guilherme por toda ajuda, poesia, música e gargalhadas durante seus períodos de vassalagem no início deste trabalho (“The moment I wake up, before I put on my makeup...”).

À Rochele, minha tutora quando ingressei no laboratório e que me iniciou no mundo do Genolyptus, por todo esse ótimo tempo de convivência, pela parceria e momentos divertidos.

À Michèle, sempre vendo o lado bom da coisa ruim, pelo Harry e pelo seu sempre bem-vindo otimismo.

À Andréia, grande exemplo de alegria e disposição, por todo o companheirismo e força.

À Ana e seus racionais conselhos e sábias avaliações psicológicas, pela força e por todo auxílio no laboratório.

À Anne, companheira nos momentos “de última hora” e nos momentos tragicômicos como, por exemplo, na época dos Fab Five (entre outros tantos).

A todas as colegas do Laboratório de Biologia Molecular Vegetal, pela ótima convivência no trabalho que tornou esse tempo todo menos difícil e também pela indispensável parceria nos *happy hours*, almoços e cafezinhos. Obrigada Luisa, Gabi, Sinara...

Ao Bauer, pela sempre disposta ajuda nas tantas coletas na Aracruz.

A todos os pesquisadores do Projeto Genolyptus, às indústrias de Celulose e Papel e ao CNPq.

Índice

| | |
|---|----|
| Resumo | 7 |
| Abstract | 8 |
| 1. Introdução | 9 |
| 1.1 Indústria de celulose e papel | 9 |
| 1.1.1 Importância econômica da madeira | 9 |
| 1.1.2 Empresas de celulose e papel | 10 |
| 1.2 <u>Eucalyptus</u>: características da madeira e melhoramento genético | 12 |
| 1.3 <u>Eucalyptus grandis</u> e <u>Eucalyptus globulus</u> | 14 |
| 1.4 O Processo de gênese da madeira | 15 |
| 1.5 Projeto GENOLYPTUS e os estudos de genômica funcional | 17 |
| 1.6 Microarranjos de DNA | 19 |
| 1.7 PCR quantitativa em tempo real e genes-refêrencia | 19 |
| 2. Objetivos | 23 |
| 2.1 Objetivos específicos | 23 |
| 3. Material e Métodos | 24 |
| 3.1 Experimentos de microarranjos | 24 |
| 3.1.1 Coleta de amostras e extrações de RNA | 24 |
| 3.1.2 Síntese e hibridizações de microarranjos | 25 |
| 3.1.3 Normalização de dados | 25 |
| 3.1.4 Análise estatística pelo teste <i>t</i> | 26 |
| 3.1.5 Análise estatística pelo método “Significance Analysis of Microarrays” (SAM) | 28 |
| 3.1.6 Seleção de genes constitutivos a partir dos microarranjos | 28 |
| 3.1.7 Seleção dos genes diferencialmente expressos entre xilemas de <u>E. globulus</u> e <u>E. grandis</u> | 30 |
| 3.2 Seleção de genes constitutivamente expressos de acordo com a literatura | 31 |
| 3.2.1 Genes constitutivos clássicos | 31 |
| 3.3.2 Novos genes constitutivos de <u>Arabidopsis thaliana</u> | 31 |
| 3.3 RT-PCR quantitativa em tempo real | 32 |

| | |
|--|-----------|
| 3.3.1 Delineamento de primers | 32 |
| 3.3.2 Síntese de cDNA | 32 |
| 3.3.3 qRT-PCR | 32 |
| 4. Resultados e Discussão | 34 |
| 4.1 Híbridizações de microarranjos de DNA visando a seleção de genes-referência constitutivamente expressos em <u>Eucalyptus</u> e diferencialmente expressos entre xilemas de <u>E. grandis</u> e <u>E. globulus</u> | 34 |
| 4.1.1 Preparação de RNAs, cDNAs marcados e montagem dos microarranjos | 34 |
| 4.2 Análises estatísticas visando a validação e interpretação de dados dos microarranjos | 38 |
| 4.3 Genes diferencialmente expressos entre xilemas de <u>E. globulus</u> e <u>E. grandis</u> | 41 |
| 4.4 Genes constitutivos em <u>Eucalyptus</u> | 54 |
| 4.5 qRT-PCR | 58 |
| 4.6 Conclusões e Perspectivas | 64 |
| 5. Referências Bibliográficas | 65 |
| Anexos | 81 |

Resumo

Um dos principais avanços da era pós-genômica consiste na possibilidade da análise da expressão global de genes pela tecnologia de microarranjos de DNA, a qual permite a investigação simultânea do perfil de transcrição de milhares de genes. Com os microarranjos, tornou-se possível comparar os padrões de expressão entre indivíduos de diferentes espécies, de diferentes órgãos/tecidos dentro de uma mesma espécie ou diferentes tecidos submetidos a várias situações e tratamentos e, ainda, analisar os genes de proteínas potencialmente reguladoras da transcrição, os fatores de transcrição. Uma etapa fundamental após a análise dos resultados dos microarranjos consiste na validação experimental dos dados de expressão gênica. Esta validação é realizada pela utilização da técnica de qRT-PCR (“Real Time Quantitative RT-PCR”) um método que, atualmente, apresenta maior sensibilidade e especificidade na análise de transcritos. A qRT-PCR necessita, entretanto, de normalização para uma leitura adequada dos dados. Essa normalização tornou-se bastante específica já que depende da disponibilidade de genes que apresentem transcritos com expressão uniforme em todas as células do organismo ou entre as espécies que estão sendo analisadas, bem como durante várias fases do desenvolvimento e sob diferentes condições ambientais. São os chamados genes-referência ou *housekeeping*. Entretanto, alguns trabalhos vêm constatando que os genes-referência utilizados na era pré-genômica não apresentam expressão estável entre tecidos e genótipos e, por este motivo, não são adequados como controles em qRT-PCR. O objetivo desse trabalho foi investigar a estabilidade de expressão de genes constitutivos frequentemente utilizados como controles internos em estudos de expressão gênica e selecionar novos genes-referência para *Eucalyptus* por meio de experimentos de qRT-PCR para a sua utilização na validação de experimentos de microarranjos do Projeto Genolyptus. Como resultados das análises de qRT-PCR, foi possível observar que para xilemas de *E. globulus* e xilemas e folhas maduras de *E. grandis*, os genes EuC12, At2g28390, EuC10, Histona H2B, Gliceraldeído-3-fosfato-desidrogenase (GAPDH), EuC06 e EuC09 foram os que demonstraram, respectivamente, menor variação em suas expressões, caracterizando-se como constitutivos para estes tecidos. De acordo com os resultados encontrados nas análises com cDNAs derivados de folhas e xilemas de *E. grandis*, somente EuC12 e At2g28390 demonstraram invariabilidade de expressões entre os tecidos. Dois tradicionais genes constitutivos, Histona H2B e GAPDH, apresentaram uma variação considerável nas suas expressões entre os tecidos. Os genes selecionados, servirão como controles internos em todas as qRT-PCRs subsequentes nas avaliações de dados gerados pelos microarranjos de DNA para todos os demais genes sob análise no Projeto Genolyptus, e demais linhas de investigação de *Eucalyptus*. Paralelamente realizou-se, ainda, a seleção e análise de treze genes com expressão diferencial significativa entre tecidos vasculares das espécies de *E. grandis* e *E. globulus*. Estes foram selecionados com base nos resultados de microarranjos do Genolyptus e terão seus padrões de expressão analisados por qRT-PCR.

Abstract

One of the most important advances of post-genomic era is the possibility of global gene expression analysis by DNA microarray technology, which allow verifying transcription profile for thousand genes simultaneously. Microarray allows comparing expression patterns among individuals from distinct species, from distinct organs/tissues within of the same specie or distinct tissues submitted across a wide range of developmental and environmental conditions, beyond genes of proteins mighty involved in transcriptional regulation, the transcription factors. A basic stage of microarray analysis is the experimental validation from gene expression data. This validation is carried out by the use of qRT-PCR (“Real Time Quantitative RT-PCR”) technique witch is the method that currently presents greater sensitivity and specificity in transcript analysis. However, qRT-PCR needs normalization for an accurate data interpretation. This data normalization is specific since it depends on the availability of a gene which displays highly uniform expression in all organism cells, between species under analysis, during various phases of development and under different environmental conditions. They are known as reference or housekeeping genes. However, some studies come evidencing that reference genes normally used in pre-genomic era doesn’t show steady expression between tissues and genotypes and, for this reason, are not suitable as controls in qRT-PCR. The objective of this work was to investigate the expression stability of housekeeping genes often used as internal controls in genetic expression studies and select novel genes for *Eucalyptus* by qRT-PCR for its use in the validation of Genolyptus Project microarray experiments. As results it was possible to identified histone, EuC06, EuC09, EuC10, EuC12, At2g28390 and GAPDH as the most stably expressed genes between *E. globulus* xylems and *E. grandis* xylems and mature leafs, as well pointing them out as control genes for these tissues. In accordance with the results found in the analysis with cDNAs derived from *E. grandis* leaves and xylems, only EuC12 and At2g28390 had demonstrated expression invariability between tissues. We found out that two traditional housekeeping genes, Histone and GAPDH, shown a considerable expression variation between those tissues. According to our analysis the selected genes will serve as internal controls in all subsequent qRT-PCRs, in microarray data evaluation to other genes under investigation in Genolyptus Project as well as another research lines in *Eucalyptus*. At the same time, we selected and analyzed thirteen genes with significant differential expression between *E. grandis* and *E. globulus* vascular tissues. These had been selected on the basis of the results of Genolyptus microarrays. Their expression patterns will be analyzed by qRT-PCR.

1. Introdução

1.1 A indústria de celulose e papel

1.1.1 Importância econômica da madeira

Atualmente, as plantações florestais de rápido crescimento possuem um papel importante na economia mundial e na preservação do ecossistema. A disponibilidade de madeira de plantio, mais barata e homogênea que a das florestas naturais, é uma alternativa eficiente para a demanda de biomassa lenhosa, diminuindo a devastação das florestas nativas. Além disso, as plantações florestais comerciais contribuem na remoção do excesso de dióxido de carbono da atmosfera, reduzindo um dos principais causadores do aquecimento global.

A madeira representa a principal fonte de energia natural renovável e é uma das mais importantes fontes de matéria-prima para a indústria mundial. Uma grande quantidade de madeira é manufaturada para gerar combustível, fibras (para produção de pasta de celulose e diversos tipos de papel) e madeira para serrarias (indústria de construção civil e movelaria). O complexo químico que constitui a madeira (celulose, hemiceluloses, ligninas e pectinas) também faz dela matéria-prima para o que poderia ser, futuramente, uma indústria “ligno-química”, isto é, em substituição à indústria petroquímica na obtenção de material plástico e outros produtos químicos (Plomion *et al.*, 2001). Em média, 70% da madeira maciça utilizada pela indústria moveleira é proveniente de plantios florestais, principalmente de espécies de *Pinus*. O eucalipto vem consolidando-se no segmento de camas e de salas de jantar e estruturas de móveis estofados e, mais recentemente, na fabricação de móveis de jardim para exportação (Sociedade Brasileira de Silvicultura, 2006).

Fora do âmbito industrial, a madeira é amplamente utilizada *in natura* como combustível (“lenha”), sendo a principal fonte de energia térmica para mais de dois bilhões de pessoas ao redor do mundo. A madeira e o carvão vegetal contribuem com mais de 14% da energia primária mundial e têm potencial para, futuramente, substituírem os combustíveis fósseis, ajudando a reduzir o efeito estufa causado por emissões de gases. O cenário sócio-econômico indica um crescimento da demanda por biocombustíveis, e a tendência é que este crescimento aumente exponencialmente nas próximas décadas. Nos

países em desenvolvimento, a dependência desse tipo de combustível é ainda maior, sendo ele responsável por cerca de um-terço da energia total utilizada. A madeira e o carvão são essenciais para o abastecimento de residências urbanas e rurais de baixa renda, além de desempenharem um papel importante também para indústrias de processamento de alimentos. Em algumas sub-regiões da África, por exemplo, cerca de 80% da energia utilizada deriva deste tipo de combustível (FAO, 2006 “Programme on Wood Energy”). Neste contexto, o Brasil é um dos maiores produtores e consumidores de carvão vegetal no mundo. De acordo com a Sociedade Brasileira de Silvicultura, a produção nacional foi de 5,5 milhões de toneladas em 2005, sendo que 2,5 milhões de toneladas (46%) foram oriundos de florestas plantadas. Conseqüentemente, foram consumidos 38,05 milhões de metros cúbicos de lenha na forma de carvão vegetal somente naquele ano (Sociedade Brasileira de Silvicultura, 2006).

O Brasil possui área total absoluta de 851 milhões de hectares (ha). Deste total, 477,7 milhões correspondem a florestas naturais e 5,6 milhões a florestas plantadas. Estas últimas ocupam apenas 0,65% do território nacional e 1% do solo agropecuário. O país conta, ainda, com 61,8 milhões de ha de unidades federais de conservação sob regime de proteção integral (45,5%) e de uso sustentável (54,5%). A área de plantações existentes corresponde a aproximadamente 5,6 milhões de ha, sendo 3,4 milhões de ha com eucalipto, 1,8 milhão de ha com *Pinus* e 326 mil ha com outras espécies, como acácia-negra, gmelina, álamo, seringueira, teca e araucária (BRACELPA, 2006; Sociedade Brasileira de Silvicultura, 2006).

1.1.2 Empresas de celulose e papel

As empresas brasileiras de celulose e papel têm o eucalipto como fonte de matéria-prima e possuem uma das maiores produtividades florestais do mundo, suprindo, de forma eficiente, a demanda de matéria-prima com o fornecimento de produtos de qualidade internacional. Também são elas as responsáveis pela geração de renda, empregos, tributos e bilhões de dólares em exportações para o país. De acordo com levantamento realizado no Ministério de Comércio Exterior (MCEX) pelo Instituto de Pesquisas e Estudos Florestais as vendas de papel e celulose para o exterior representaram cerca de US\$ 4,0 bilhões em 2006. Os principais mercados para exportação da celulose brasileira têm sido a Europa (49%), a Ásia (30%) e a América do Norte (19%). No segmento de papel, os principais

mercados são a América Latina (55%), a Europa (17%), a América do Norte (16%), a Ásia (7%) e a África (5%) (<http://www.ipef.br/estatisticas/#Relatorios>).

Em 2006, a produção brasileira de celulose alcançou 11,1 milhões de toneladas e a de papel, 8,8 milhões de toneladas. Segundo pesquisas de institutos internacionais, o resultado da produção de celulose em 2006 levou o Brasil a passar da sétima para a sexta posição como produtor mundial, ultrapassando o Japão. O valor total da produção do setor de base florestal em 2005 foi de US\$ 27,8 bilhões, ou seja, 3,5% do Produto Interno Bruto (PIB) nacional. Neste valor estão incluídos celulose, papel, madeira industrializada sob todos os processos, móveis, siderurgia a carvão vegetal e produtos florestais não madeireiros (BRACELPA, 2006; Sociedade Brasileira de Silvicultura, 2006).

As empresas de celulose investem constantemente em atualização tecnológica e possuem um rigoroso padrão de controle ambiental realizado pelo uso renovável dos recursos naturais. As companhias, visando atender ao Programa Brasileiro de Certificação Florestal (Cerflor), buscam desenvolver técnicas e práticas de manejo florestal que sejam capazes de permitir o aumento da produtividade e garantir a minimização dos danos ambientais, de forma a reduzir os impactos sobre o solo, fauna, flora e investir no tratamento de efluentes. Atendendo à legislação florestal, as empresas responsabilizam-se por manter áreas de reservas nativas onde existam áreas plantadas com eucalipto, garantindo grande parte da biodiversidade regional. Para exemplificar, as áreas de terras da empresa Veracel Celulose S.A. somam 164,6 mil ha, distribuídos em dez municípios no sul da Bahia. Deste total, 78,1 mil hectares são destinados ao plantio de eucalipto, o que equivale a 47,5% da área total. A maior parte, que chega a 48%, é destinada à recuperação e preservação ambiental. A empresa Aracruz Celulose S.A. possui a maior área de reservas florestais nativas de Mata Atlântica da Bahia e do Espírito Santo, contribuindo para divulgar conceitos e temas ligados ao meio ambiente. Desta maneira, os plantios são estabelecidos por meio de práticas que visam, além da formação de florestas com boa produtividade, a conservação do solo, da água, da fauna e da flora locais. A empresa Aracruz Celulose S.A. também desenvolve e apóia diversas iniciativas de educação ambiental em parceria com outras entidades (Aracruz, 2007; Veracel, 2007).

1.2 Eucalyptus: características da madeira e melhoramento genético

O gênero *Eucalyptus* (família Myrtaceae) inclui mais de 500 espécies que, somadas a diversos híbridos naturais e artificiais, totalizam mais de 1.500 formas distintas. O eucalipto ocorre principalmente em solos ácidos, estendendo-se de planícies tropicais com temperaturas quentes e úmidas até regiões montanhosas com clima frio e temperado. Algumas espécies possuem altas taxas de crescimento e uma grande adaptabilidade a solos e climas (Turnbull, 1999). Assim, a partir desta diversidade, encontram-se madeiras de características físico-mecânicas e estéticas bastante diferenciadas. O *Eucalyptus* é a arbórea mais amplamente distribuída em todo o planeta, sendo proveniente da Austrália, Nova Guiné, Indonésia e Filipinas. Existem plantações comerciais estabelecidas em mais de 17 milhões de ha na Ásia, África, América do Sul e América do Norte (FAO, 2000). Além de ser conhecido pela sua adaptação a uma grande diversidade de climas e solos, rápido crescimento e alta produtividade, o eucalipto fornece madeira com fibra de excelente qualidade e uma arquitetura altamente favorável em função dos troncos retos e com poucas ramificações (revisado em Poke *et al.*, 2005).

As espécies de *Eucalyptus* são as mais utilizadas no mundo para a exploração comercial de madeira e, atualmente, o Brasil ocupa o primeiro lugar em áreas plantadas desta arbórea para a sua exploração comercial. Em nosso país, o principal uso da madeira do eucalipto está na indústria de celulose e papel e, em segundo lugar, na produção de carvão vegetal para a indústria siderúrgica. Outras utilidades desta madeira incluem a produção de tábuas para a construção civil, postes e serragem para aglomerados. Também há um crescente interesse nos óleos essenciais do eucalipto para a obtenção de compostos aromáticos. O óleo do eucalipto é obtido pela simples destilação de suas folhas mediante o uso de vapor d'água. Por ser uma cultura ligada à extração de folhas, o manejo da área de plantio pode ser conduzido no conceito de uso múltiplo da floresta, ou seja, além da extração das folhas, a atividade pode ser conjugada com a produção de madeira destinada para lenha, mourões, postes e até toras para serraria (Sociedade Brasileira de Silvicultura, 2006).

Para as indústrias papeleiras, a qualidade do produto final, celulose ou papel, irá depender da qualidade da madeira e das condições do processo de fabricação. A madeira é constituída de celulose, hemiceluloses, ligninas, extrativos e de uma pequena fração de

compostos inorgânicos, distribuídos nas diversas camadas que compõem a fibra (Colodette *et al.*, 2004). As propriedades da madeira variam muito entre as espécies e entre indivíduos/genótipos da mesma espécie. A variabilidade ocorre tanto entre diferentes tecidos (proporção de diferentes tipos celulares) bem como em nível celular: tamanho, forma, estrutura da parede celular, textura e composição química (Paux *et al.*, 2005). O comprimento dos traqueídeos ou elementos de vaso, o diâmetro e a espessura da parede celular afetam a resistência e a densidade da madeira. Quando as ligninas são removidas, essas propriedades também irão determinar a resistência, a densidade da polpa e a qualidade dos produtos de papel (Whetten *et al.*, 2001). No processo de polpação química, as ligninas são dissolvidas para liberar o material fibroso constituído basicamente de holocelulose (celulose e hemiceluloses). A dissolução das ligninas na polpação depende do tipo de ligações e dos grupos funcionais presentes em sua estrutura. Em geral, madeiras com maior teor de celulose e menor de hemiceluloses deverão resultar em maior rendimento de polpação. Tendo em vista que o objetivo da polpação é a remoção seletiva das ligninas e a liberação da porção fibrosa da madeira, as madeiras com baixos teores de ligninas e de extrativos propiciam melhor desempenho da polpação em termos de rendimento e consumo de produtos químicos utilizados neste processo (Colodette *et al.*, 2004). A grande importância econômica da madeira propiciou a formação de amplos projetos de seqüenciamento de espécies de *Populus* e *Eucalyptus*. Entretanto, o acúmulo de seqüências representa apenas o primeiro passo em direção às pesquisas a respeito da biologia dessas espécies e de outras arbóreas. Muito pouco se sabe sobre o processo de gênese da madeira, sobre os mecanismos moleculares envolvidos na determinação das características estruturais da madeira, bem como a identificação dos genes, proteínas e fatores que regulam o crescimento secundário.

No último século, o Brasil obteve grandes avanços de ordem tecnológica que levaram o país ao primeiro lugar mundial na produção de fibra curta de celulose e papel. Tais conquistas foram devidas ao desenvolvimento de tecnologia em todos os setores da cultura do eucalipto: produção de sementes e mudas melhoradas, implantação e adubação de florestas, tratamentos silviculturais, colheita de produtos florestais, secagem e processamento mecânico. Os ganhos de produtividade foram significativos e obtidos mediante o desenvolvimento de programas de melhoramento, com a introdução de novas espécies e procedências, de seleção de árvores superiores, obtenção de híbridos e propagação

vegetativa (Sociedade Brasileira de Silvicultura, 2006). Da mesma forma, a grande variabilidade entre as espécies de *Eucalyptus* despertou o interesse biotecnológico no melhoramento genético dessas espécies visando a modificação das propriedades da madeira para uso industrial. O principal foco de estudo tem sido a identificação de genes envolvidos no controle genético da qualidade da madeira, além da resistência a doenças e pragas, resposta a estresses ambientais e tolerância a deficiências nutricionais (Poke *et al.*, 2005).

Dentro desse cenário, duas espécies de eucalipto chamam a atenção: *E. grandis*, por apresentar um rápido crescimento volumétrico e uma boa resistência a pragas e doenças, e *E. globulus*, por possuir um baixo teor de ligninas na madeira. O objetivo principal seria a obtenção de plantas que combinassem as características de ambas as espécies de forma a permitir a obtenção de árvores com um teor mais baixo de lignina para um melhor rendimento na polpação, maior crescimento e maior adaptabilidade ao clima e solos brasileiros.

1.3 *Eucalyptus grandis* e *Eucalyptus globulus*

A mais importante espécie de *Eucalyptus* utilizada na produção de papel e pasta de celulose no Brasil é o *E. grandis*, especialmente na forma do híbrido *E. urophylla* x *E. grandis*, mais conhecido entre os técnicos florestais como “*Urograndis*”, entre outros híbridos. *E. grandis* apresenta um bom desenvolvimento em regiões de clima tropical e subtropical, além de conter um ótimo teor de celulose na madeira (em torno de 48%) e possuir um bom crescimento volumétrico (Colodette *et al.*, 2004; revisado em Poke *et al.*, 2005). De acordo com estudos realizados em 1983 por Brito e *et al.*, no Instituto de Pesquisas e Estudos Florestais (IPEF) e no Departamento de Silvicultura da Escola Superior de Agronomia Luiz de Queiroz (ESALQ/USP), *E. grandis* é uma das espécies de eucalipto que possui o maior volume cilíndrico médio, cerca de 1.105 m³/ha.

A maior restrição encontrada pelas indústrias ao *E. grandis*, entretanto, está nos índices elevados de ligninas em sua madeira, ou seja, cerca de 24% de lignina total (Pereira *et al.*, 2000; Colodette *et al.*, 2004). Existem algumas espécies de eucalipto que apresentam baixos índices de ligninas e que, por esse motivo, deveriam ser escolhidas para o plantio

industrial do eucalipto. Tais espécies, como *E. dunnii* e *E. globulus*, possuem, em média, 22% de lignina total. Seus cultivos, entretanto, apresentam dificuldades de adaptação ao clima e solo brasileiros. Além disso, *E. globulus* tem um crescimento relativamente lento e baixa resistência a pragas e doenças, e *E. dunnii*, apesar de possuir um rápido crescimento volumétrico, conta com um alto teor em produtos extrativos que interferem criticamente na produção da celulose (Eng. Teotônio F. de Assis, comunicação pessoal).

1.4 O Processo de gênese da madeira

A formação da madeira é um processo que se origina a partir do crescimento secundário das plantas. Ao contrário do crescimento primário, o crescimento secundário deriva-se das células meristemáticas presentes no câmbio vascular e no felogênio, além de proporcionar o crescimento em circunferência do caule das plantas. Durante o crescimento secundário, as plantas convertem a maioria dos produtos da fotossíntese em vários polímeros que serão usados para a formação dos tecidos lenhosos (Li *et al.*, 2006). O xilema é composto pelos traqueídeos, vasos e fibras que estão envolvidos no transporte de água e no suporte mecânico da árvore. O processo de diferenciação dessas células ocorre durante a formação do xilema secundário (xilogênese) e leva à formação da madeira. Este processo inicia no câmbio vascular e passa por cinco estágios principais: (i) divisão e (ii) expansão celular, (iii) espessamento da parede celular, (iv) lignificação e (v) morte celular programada. Todos estes estágios são controlados pela expressão coordenada de numerosos genes com funções regulatórias e estruturais, sendo a maioria ainda desconhecida (Hertzberg *et al.*, 2001; Plomion *et al.*, 2001; Yang *et al.*, 2004).

A ação dos genes e das proteínas presentes no xilema secundário em diferenciação irá determinar as propriedades da madeira. A composição e a morfologia da parede celular secundária das células do xilema são os principais fatores responsáveis pela qualidade da madeira (Whetten *et al.*, 2001). A formação da parede celular secundária é comandada pela expressão coordenada de vários genes, envolvidos especificamente na síntese e deposição de polissacarídeos, ligninas, pectinas e proteínas de parede celular (Plomion *et al.*, 2001; Whetten *et al.*, 2001; Yang *et al.*, 2004). Assim, as principais características da madeira irão variar de acordo com o tamanho, a forma e a organização das células no xilema, bem como a estrutura e a composição química da parede celular secundária.

Dentre os estágios que levam à formação da madeira, o principal foco de estudos tem sido o processo de diferenciação do xilema, e os principais alvos são a síntese e a deposição dos principais componentes da parede celular, além da regulação deste processo por fito-hormônios (revisado em Du & Yamamoto, 2007) e fatores de transcrição (revisado em Demura & Fukuda, 2007). Muitos desses genes têm sido descobertos pelo uso de técnicas de genômica funcional em larga escala em espécies como *Eucalyptus* (Paux *et al.*, 2004; Paux *et al.*, 2005; Foucart *et al.*, 2006), *Populus* (Hertzberg *et al.*, 2001; Israelsson *et al.*, 2003; Schrader *et al.*, 2004; Aspeborg *et al.*, 2005; Prassinis *et al.*, 2005; Andersson-Gunnera *et al.*, 2006), *Pinus* (Allona *et al.*, 1998; Whetten *et al.*, 2001; Lorenz & Dean, 2002; Yang *et al.*, 2004) e até *Robinia pseudoacacia* (Falsa Acácia; Yang *et al.*, 2003). Além das espécies que atraem interesse comercial, há vários projetos dedicados tanto ao estudo de tecidos vasculares isolados quanto a traqueídeos em cultura, em plantas-modelo como *Zinnia elegans* (Demura *et al.*, 2002;) e *Arabidopsis thaliana* (Oh *et al.*, 2003; Ko & Han, 2004; Kubo *et al.*, 2005; Persson *et al.*, 2005; Zhao *et al.*, 2005).

Entretanto, mesmo com o advento de vários projetos de seqüenciamento e de genômica funcional, os mecanismos moleculares que levam ao crescimento secundário das plantas ainda precisam ser investigados. Mesmo que as principais vias estruturais já tenham sido descritas, pouco se sabe a respeito da regulação deste processo, tanto em nível celular e molecular quanto de desenvolvimento (Schrader *et al.*, 2004; Li *et al.*, 2006a; Groover & Robischon, 2006). Outro desafio está em compreender de que forma os genes envolvidos no processo da xilogênese irão influenciar o fenótipo final da madeira. Toda essa deficiência de informações deve-se ao fato de que, das supostas centenas de genes que são necessários para a síntese e deposição de polissacarídeos, apenas alguns foram caracterizados até o momento.

Em todos os projetos de seqüenciamento de transcriptomas de plantas, cerca de 40 a 60% das seqüências de genes expressos (ou *expressed sequence tags*, ESTs) correspondem a genes ainda desconhecidos, os denominados “*no hits*”, “*hypothetical protein*”, “*expressed protein*” ou “*putative protein*”, além de outras denominações genéricas, após uma análise manual ou automática utilizando-se a ferramenta BLAST (*Basic Local Alignment Search Tool*) do *National Center for Biotechnology Information* (NCBI, www.ncbi.com). É bastante provável que toda a fração ainda desconhecida a respeito do processo de formação da madeira será esclarecida a partir de novos genes regulatórios e

estruturais que ainda serão caracterizados. O uso da tecnologia de *DNA microarray* ou microarranjos de DNA tem possibilitado a análise do perfil de expressão de milhares de genes de tecidos vasculares e proporcionado avanços importantes na genômica funcional de diferenciação do xilema e de potenciais reguladores dos eventos da xilogênese. Este perfil de expressão está baseado na medida da quantidade de mRNAs disponíveis em um determinado momento ou condição, pela hibridização a fragmentos homólogos de DNA imobilizados em uma superfície sólida.

A finalização do seqüenciamento do genoma de *Populus* está particularmente fornecendo poderosas ferramentas para o entendimento da biologia das espécies arbóreas e auxiliando os programas de melhoramento nesses organismos (Brunner *et al.*, 2004). Visando ampliar o conhecimento sobre a formação e as propriedades da madeira, a descoberta das bases moleculares envolvidas na xilogênese irá servir de base para a seleção de genes candidatos que afetam características importantes como peso específico, ângulo das microfibrilas ou conteúdo dos componentes da parede celular para serem modificados geneticamente (Plomion *et al.*, 2001).

1.5 Projeto GENOLYPTUS e os estudos de genômica funcional

O projeto "GENOLYPTUS: Rede Brasileira de Pesquisa do Genoma de *Eucalyptus*" foi oficialmente lançado em fevereiro de 2002 pelo Ministério da Ciência e Tecnologia, com recursos do Fundo Setorial "Verde-Amarelo" e de empresas de celulose e papel. O projeto tem a participação de 10 centros de pesquisas e 14 empresas e possui definido como objetivo principal o descobrimento, o seqüenciamento, o mapeamento e a determinação da função de genes de importância econômica de espécies de *Eucalyptus*, visando a incorporação de tecnologias de genômica em programas de melhoramento e produção florestal em busca de plantas de *Eucalyptus* mais produtivas, mais resistentes a moléstias e pragas e mais tolerantes a estresses abióticos. O projeto GENOLYPTUS está organizado em oito subprojetos que englobam desde a instalação de uma rede experimental a campo e implementação de tecnologias para a avaliação da qualidade da madeira, passando pela construção de mapas genéticos e físicos para o mapeamento de locos controladores de características quantitativas (ou *quantitative trait loci*, QTLs), além da

investigação da base genética e identificação de genes que conferem resistência a doenças em *Eucalyptus*.

O subprojeto intitulado "Seqüenciamento do Transcriptoma de *Eucalyptus*" teve como metas a construção de bibliotecas de expressão de *Eucalyptus*, a obtenção de pelo menos 30.000 seqüências gênicas únicas ou de 150.000 ESTs válidas, e a organização e manutenção dos estoques de clones de cDNA. Como conclusão, 13 bibliotecas de expressão foram construídas: (1) xilema de *E. grandis*; (2) xilema de *E. globulus*; (3) xilema de *E. pellita*; (4) xilema de *E. urophylla*; (5) mistura de xilemas de nove espécies e híbridos; (6) mistura de floemas de nove espécies e híbridos; (7) folhas maduras de *E. grandis*; (8) folhas jovens de *E. grandis*; (9) folhas jovens de *E. grandis* infectadas a campo pelo fungo *Puccinia*; (10) plântulas *in vitro* de *E. grandis* e (11) mistura de plântulas de *E. grandis* submetidas a 20 diferentes tratamentos; (12) raízes de diferentes espécies e híbridos de *Eucalyptus*; e (13) flores de *E. grandis*. Um total de 96.493 ESTs válidas foram depositadas nas centrais de bioinformática do Projeto GENOLYPTUS, incluindo 21.905 seqüências únicas.

O projeto WoodGenes, ou "Gênese de Madeira em *Eucalyptus*: Genes, Funções, Regulação e Expressão Transgênica", teve início em janeiro de 2005 e está direcionado à exploração da coleção de clones de cDNAs gerados no projeto GENOLYPTUS. O objetivo principal do projeto é a avaliação da função gênica das seqüências isoladas e, especialmente, o potencial dos cDNAs de *Eucalyptus* em proporcionar características benéficas ao processamento da sua madeira por meio da expressão transgênica. Este estudo está direcionado, principalmente, aos genes cujas expressões afetem, direta ou indiretamente, a formação da madeira. Para esta finalidade, foram planejados e estão sendo conduzidos experimentos de microarranjos de DNA.

Recentemente, foi aprovado o projeto para o seqüenciamento na íntegra do genoma do eucalipto junto ao *Joint Genome Institute* (JGI) ligado ao Departamento de Energia do governo dos EUA (USDOE). O projeto foi organizado por pesquisadores dos EUA, Brasil e África do Sul, e conta com participantes da Austrália, Portugal, França e Suécia. Todo o trabalho computacional de anotação e de montagem do genoma será realizado nos EUA e os dados serão públicos. A espécie selecionada foi *E. grandis*, chamado de eucalipto tropical, por ser a espécie que melhor adapta-se ao clima e solo brasileiros e a mais cultivada no mundo, como inicialmente referido. Essa conquista é de extrema importância

para o Brasil e para o aperfeiçoamento das tecnologias de genômica para os programas de melhoramento silviculturais.

1.6 Microarranjos de DNA

Como citado anteriormente, o uso da tecnologia de microarranjos tem possibilitado a análise do perfil de expressão de genes de tecidos vasculares e tem levado a importantes avanços nos estudos de genômica funcional em relação ao processo de diferenciação do xilema e dos potenciais reguladores dos eventos da xilogênese. A tecnologia de microarranjos vem revolucionando a biologia molecular e tem se tornado uma importante ferramenta para a análise de expressão gênica em larga escala. Com esta técnica, milhares de sondas (cDNAs ou oligonucleotídeos sintéticos) imobilizadas mecanicamente em uma lâmina são hibridizadas simultaneamente com várias moléculas de DNA. Isso permite medir a quantidade momentânea individual de mRNA derivada de milhares de genes simultaneamente, possibilitando analisar padrões de expressão gênica globais. Além disso, os microarranjos têm permitido também analisar os níveis de mRNAs de genes em resposta a estresses bióticos e abióticos (Kawaura *et al.*, 2006; Weber *et al.*, 2006) e comparar transcriptomas vegetais (Ma *et al.*, 2005). Diversas ferramentas de bioinformática e informações sobre seqüências em bancos de dados públicos como o “*The Arabidopsis Information Resource*” (TAIR, <http://www.arabidopsis.org>) e do genoma de *Populus* (*PopulusDB*, <http://genome.jgi-psf.org/Poptr1/>), entre outros, têm facilitado a análise dos dados obtidos com essa técnica (Ree *et al.*, 2006).

1.7 PCR Quantitativa em tempo real e genes-refêrencia

Uma etapa fundamental da análise de resultados dos microarranjos consiste na validação dos dados de expressão gênica. Esta validação é realizada pela utilização da técnica conhecida por qRT-PCR (do inglês, *real-time quantitative RT-PCR*). A técnica de qRT-PCR tem como base o processo de transcrição reversa (RT) seguida da reação em cadeia da DNA-polimerase (PCR) com a incorporação de moléculas fluorescentes

covalentemente ligadas ou não a nucleotídeos, as quais podem ser quantificadas durante a cinética da reação (em “tempo real”). Os produtos formados são monitorados a cada ciclo de reação, o que permite uma detecção rápida e específica dos produtos de amplificação. Atualmente, é o método que apresenta maior sensibilidade e especificidade na análise de transcritos (Gachon *et al.*, 2004). A técnica de qRT-PCR necessita, entretanto, de normalização para uma leitura adequada dos dados. A normalização se faz necessária já que podem existir fatores como a própria variabilidade inerente do RNA, protocolos de extração que podem co-purificar inibidores e discrepâncias de eficiência entre as amostras durante o processo de transcrição reversa e da PCR (Bustin *et al.*, 2005).

Para a normalização dos dados obtidos com a análise por qRT-PCR, são necessários genes que apresentem transcritos que possuam uma expressão uniforme em uma maioria das células do organismo ou entre as espécies que estão sendo analisadas, assim como durante várias fases do desenvolvimento e sob diferentes condições ambientais. São estes os chamados genes constitutivos ou *housekeeping genes*. Estes genes são escolhidos com base no conhecimento ou provável papel nos processos celulares básicos como, por exemplo, manutenção da estrutura celular ou metabolismo primário. Alguns dos mais conhecidos e utilizados genes-referência em estudos com plantas e animais incluem aqueles codificadores do rRNA 18S, gliceraldeído-3-fosfato desidrogenase (GAPDH), fator de alongamento 1 α (EF-1 α), actina (ACT), α -tubulina (TUA) e β -tubulina (TUB) (revisado em Czechowski *et al.*, 2005). Durante anos, estes genes foram utilizados como referência em experimentos de Northern *blots* e RT-PCR convencionais. Entretanto, trabalhos têm constatado que esses genes utilizados na era pré-genômica não apresentam expressão estável entre tecidos e genótipos e, por esse motivo, não são adequados como controles a métodos mais sensíveis como a qRT-PCR.

Iskandar e *et al.*, (2004) avaliaram a expressão de TUA, TUB, GAPDH e rRNA 25S em diversos tecidos e genótipos de cana-de-açúcar. Entre os genes estudados, os que melhor apresentaram um perfil constitutivo foram os codificadores do rRNA 25S e da GAPDH.

Brunner *et al* (2004) utilizaram a técnica de RT-PCR para analisar a expressão de 10 genes constitutivos comumente usados nos estudos de expressão gênica de *Populus* (híbrido *P. trichocarpa* x *P. deltoides*) em vários tecidos coletados em diferentes estádios e em diferentes períodos do ano. Os genes estudados apresentaram expressão variável

durante o desenvolvimento sazonal e a idade das árvores. Os genes constitutivos escolhidos para o estudo representavam diferentes classes funcionais e famílias gênicas: ubiquitinas (UBQ), actinas, tubulinas, ciclofilina citosólica (CYC), peptidilprolil-isomerase, fator de iniciação da tradução, fator de alongamento da tradução e rRNA. Entre todos os genes testados, apenas os codificadores de poliubiquitina (UBQ11) e TUA apresentaram níveis estáveis de expressão. Neste trabalho, os autores recomendaram ainda a utilização de oligonucleotídeos iniciadores (*primers*) que permitam a amplificação de dois ou mais membros de uma mesma família gênica. Segundo eles, o nível de expressão combinado entre os membros leva a um melhor nível de expressão e estabilidade, além de evitar um processo mais caro e trabalhoso ao utilizar múltiplos genes como controle. Os *primers* para famílias gênicas irão dar origem a múltiplos *amplicons* provenientes dos diferentes membros que serão do mesmo tamanho e apresentarão a mesma eficiência na qRT-PCR. Por esses motivos, membros de uma mesma família podem resultar em um controle interno mais estável do que um gene de apenas uma cópia.

Nicot *et al.*, (2005) testaram a expressão de sete genes constitutivos durante estresses bióticos e abióticos em batatas pelo método de qRT-PCR. Da mesma forma que os demais autores, eles incluíram no estudo *primers* para TUB, CYC, ACT, EF-1 α , rRNA 18S e, ainda, para adenina-fosforribosiltransferase (APRT) e ribonucleoproteína citoplasmática L2. Neste trabalho, apenas EF-1 α manteve sua expressão estável em todas as condições de estresses.

Jain *et al.*, (2006) analisaram dez genes constitutivos em plantas de arroz para diferentes tecidos e órgãos em diferentes estádios do desenvolvimento, além de plântulas submetidas a várias condições de estresses e tratamentos hormonais. O método utilizado foi também qRT-PCR. Os genes testados foram os codificadores de ACT11, da enzima de conjugação à ubiquitina E2 (UBC), de eEF-1 α (EF-1 α eucariótico), GAPDH, TUB, eIF-4a, UBQ10, UBQ5, rRNA 18S e 25S. Os resultados indicaram expressão diferencial nas amostras analisadas. Os genes mais estáveis foram UBQ5 e eEF-1 α .

Até o momento, o trabalho mais completo para genes-referência em estudos de expressão gênica em plantas foi realizado por Czechowski *et al.*, (2005) em *Arabidopsis*. A constitutividade de genes foi estudada com base em uma grande quantidade de dados obtidos em trabalhos anteriores usando arranjos que representavam 23.500 genes e incluíram análises de expressão para a maioria dos genes sob um amplo espectro de

condições ambientais e estádios de desenvolvimento. De fato, 79 diferentes tecidos, órgãos, estádios do desenvolvimento e genótipos; raízes e brotos submetidos a estresses bióticos e abióticos; plântulas submetidas a tratamentos com diferentes qualidades e durações de luz; e plântulas tratadas com sete diferentes fito-hormônios foram avaliados por estes autores. A partir dos dados obtidos e por meio de análises estatísticas, os autores testaram a constitutividade de alguns genes-referência tradicionais (alguns já estudados em trabalhos anteriores): ACT2, TUB6, EF-1 α , UBQ10 e GAPDH. Todos estes demonstraram diferenças de expressão. UBQ10 apresentou estabilidade em todos os estádios de desenvolvimento analisados, exceto entre as sementes, assim como GAPDH, que não se apresentou estável nas amostras de pólen e sementes. O resultado das análises de todo o conjunto de genes levou à indicação de 100 genes constitutivos subdivididos em diversas categorias (Czechowski *et al.*, 2005: Tabela Suplementar 1). Dezoito novos genes apresentaram-se estáveis em todas as condições e foram validados para a normalização de qRT-PCR em *Arabidopsis*.

Conforme os trabalhos citados acima, genes-referência com expressão estável em um organismo podem não ser adequados para a normalização da expressão gênica em outro organismo, e ainda sob um determinado número de tratamentos e estádios de desenvolvimento. Por este motivo, Jain *et al.*, (2006) recomendaram que os genes constitutivos devem ser validados antes da sua utilização como controles internos em qRT-PCRs. A única espécie arbórea a apresentar esse tipo de estudo de validação de genes constitutivos é *Populus*, como citado anteriormente. Para *Eucalyptus*, até o momento, não há qualquer estudo de normalização com genes-referência.

2. *Objetivos*

Dois objetivos gerais foram propostos à execução do presente trabalho:

(i) Investigar a estabilidade de expressão de genes potencialmente constitutivos freqüentemente utilizados como controles internos em estudos de expressão gênica e confirmá-los como referências adequadas a estes estudos ou selecionar novos genes-referência de expressão constitutiva para *Eucalyptus*.

(ii) Identificar e selecionar genes de mais alta expressão diferencial entre xilema de *Eucalyptus grandis* e *Eucalyptus globulus* como possíveis justificativas das diferenças observadas entre as qualidades das madeiras destas espécies.

2.1 *Objetivos específicos*

- Selecionar pelo menos três genes constitutivos clássicos de vegetais e testar suas expressões em *Eucalyptus* por qRT-PCR;
- Avaliar em *Eucalyptus*, por qRT-PCR, pelo menos quatro novos genes constitutivos dentre os descritos por Czechowski *et al.* (2005) como bons controles internos em *Arabidopsis thaliana*;
- Identificar, por microarranjos de DNA, outros genes candidatos a referências de expressão constitutiva em *Eucalyptus* e validar tal padrão de expressão por qRT-PCR;
- A partir dos dados confirmatórios da expressão constitutiva, validar os resultados de microarranjos realizados entre folhas e xilema de *E. grandis* e entre xilemas de *E. grandis* e *E. globulus*;
- Paralelamente, identificar e analisar, por microarranjos de DNA, pelo menos dez genes que possuam expressão diferencial significativa entre xilemas de *E. grandis* e *E. globulus*.

3. Material e Métodos

3.1 Experimentos de microarranjos

3.1.1 Coleta de amostras e extrações de RNA

Amostras de RNA total foram preparadas a partir de xilemas de árvores de quatro anos de idade das espécies *E. grandis* e *E. globulus* plantadas em 2002 a campo no Hortoflorestal Barbanegra da empresa Aracruz Celulose S.A., em Barra do Ribeiro, RS. As amostras foram coletadas de dois indivíduos- clones derivados de duas matrizes geneticamente não relacionadas (filhas de pais distintos) de *E. grandis* e de *E. globulus*. De dois dos clones de uma das matrizes de *E. grandis*, também foram coletadas folhas maduras, isto é, totalmente expandidas (Tabela 1). As extrações de RNA total foram realizadas utilizando-se o *kit* PureLink Plant RNA Reagent (Invitrogen) de acordo com o protocolo “*Small Scale RNA Isolation*”, fornecido pelo fabricante. Amostras de 40 µg de RNAs totais foram tratadas com 40 U da enzima DNase I (Fermentas), conforme sugerido pelo fabricante, e a qualidade e a quantidade das amostras foram posteriormente verificadas por eletroforese em gel de agarose a 1,5%, tampão TAE 1X (Tris 4 mM; ácido bórico 2 mM; EDTA 0,1 mM) e por espectrofotometria (leitura de absorbâncias a 260 nm e a 280 nm). Cerca de 40 µg de cada amostra foram desidratados em “*speedvac*” e enviados à empresa NimbleGen Systems Inc. (Reykjavik, Islândia) acondicionados em microtubos de 0,5 mL e sob temperatura ambiente para a realização da síntese de cDNAs marcados e hibridizações de microarranjos.

Tabela 1. Identidade das amostras biológicas.

| NOME DA AMOSTRA | ESPÉCIE | TECIDO | CLONE |
|-----------------|--------------------|----------------|-------|
| GR-XY-A1 | <i>E. grandis</i> | Xilema | A1 |
| GR-XY-A2 | | | A2 |
| GR-XY-B1 | | | B1 |
| GR-XY-B2 | | | B2 |
| GR-ML-A1 | <i>E. grandis</i> | Folhas Maduras | B1 |
| GR-ML-A2 | | | B2 |
| GL-XY-A1 | <i>E. globulus</i> | Xilema | A1 |
| GL-XY-A2 | | | A2 |
| GL-XY-B1 | | | B1 |
| GL-XY-B2 | | | B2 |

3.1.2 Síntese e hibridizações de microarranjos

A empresa NimbleGen Systems Inc. foi contratada para a execução dos experimentos de microarranjos de DNA no âmbito do Projeto Genolyptus. Um total de 21.442 seqüências derivadas dos 21.905 agrupamentos (*clusters*) e singletos (*singletons*) de ESTs obtidos no Projeto Genolyptus foram encaminhadas à empresa pelo Prof. Dr. Georgios J. Pappas Jr. (Universidade Católica de Brasília, DF) para a projeção e síntese de nove oligonucleotídeos sintéticos de 50 unidades de bases (*50mer*) para cada seqüência-consenso de agrupamentos e singletos, distribuídas ao longo de cada seqüência e com temperaturas médias de desnaturação próximas. As nove sondas foram duplicadas nos *chips*, perfazendo um total de 385.856 sondas (*features*) por *chip*. Um total de 10 *chips* idênticos foi produzido segundo protocolos da própria empresa. A partir das amostras de RNA total encaminhadas à NimbleGen, análises de qualidade e quantidade foram conduzidas para confirmar seus bons estados de conservação e cDNAs foram sintetizados com a marcação pelo grupo fluoróforo Cy3. Hibridizações, lavagens, *scanning*, coleta de dados e normalizações preliminares dos dados foram realizados conforme metodologias da NimbleGen. Versões resumidas destes protocolos e suas características podem ser acessadas em <http://www.nimblegen.com/products/exp/index.html>.

3.1.3 Normalização de dados

A normalização de dados das hibridizações dos dez microarranjos foi realizada pela NimbleGen pelo método denominado “*Robust Multi-Array Analysis*” (RMA, Irizarry *et al.*, 2003). Os valores referentes à intensidade de sinal capturado pelo scanner para cada uma das 385.856 sondas imobilizadas nos *chips* foram analisados e os valores de intensidade referentes às nove sondas transformados em um único valor de intensidade final. Ao final da análise pelo método RMA, é resultante um valor numérico relativo à quantidade momentânea de mRNA da amostra-alvo (*steady-state mRNA*) para cada um dos genes, em escala Log2. O método de análise RMA consiste das seguintes etapas: eliminação da fluorescência basal, normalização e transformação dos dados (Irizarry *et al.*, 2003).

3.1.4 Análise estatística pelo teste *t*

A análise inicial dos dados normalizados das hibridizações de microarranjos enviados pela NimbleGen foi realizada em conjunto com o Prof. Dr. Alexandre Coelho (Universidade Federal de Goiás, Goiânia, GO) utilizando-se o Teste *t* de Student, com vistas à avaliação da robustez e validação dos dados. A codificação utilizada em todas as análises foi a seguinte:

Egr - *Eucalyptus grandis*
Egl - *Eucalyptus globulus*

cA - matriz A
cB - matriz B

r1 - clone 1*
r2 - clone 2

L - folha
X - xilema

b1 - bloco 1 (do slide)**
b2 - bloco 2 (do slide)

*Os clones foram considerados replicatas biológicas.

**Replicatas experimentais.

1) Comparação dos níveis de expressão entre folhas e xilemas da matriz A de *E.*

grandis:

Grupo A:

Egr_cA_r1_L_b1
Egr_cA_r1_L_b2
Egr_cA_r2_L_b1
Egr_cA_r2_L_b2

Grupo B:

Egr_cA_r1_X_b1
Egr_cA_r1_X_b2
Egr_cA_r2_X_b1
Egr_cA_r2_X_b2

2) Comparação entre xilemas *E. grandis* e *E. globulus*:

Grupo A:

Egr_cA_r1_X_b1
Egr_cA_r1_X_b2
Egr_cA_r2_X_b1
Egr_cA_r2_X_b2
Egr_cB_r1_X_b1
Egr_cB_r1_X_b2
Egr_cB_r2_X_b1
Egr_cB_r2_X_b2

Grupo B:

Egl_cA_r1_X_b1
Egl_cA_r1_X_b2
Egl_cA_r2_X_b1
Egl_cA_r2_X_b2
Egl_cB_r1_X_b1
Egl_cB_r1_X_b2
Egl_cB_r2_X_b1
Egl_cB_r2_X_b2

3) Comparação entre xilemas dos clones de *E. grandis*:

Grupo A:

Egr_cA_r1_X_b1
Egr_cA_r1_X_b2
Egr_cA_r2_X_b1
Egr_cA_r2_X_b2

Grupo B:

Egr_cB_r1_X_b1
Egr_cB_r1_X_b2
Egr_cB_r2_X_b1
Egr_cB_r2_X_b2

4) Comparação entre xilemas das matrizes de *E. globulus*:

Grupo A:

Egl_cA_r1_X_b1
Egl_cA_r1_X_b2
Egl_cA_r2_X_b1
Egl_cA_r2_X_b2

Grupo B:

Egl_cB_r1_X_b1

Egl_cB_r1_X_b2
Egl_cB_r2_X_b1
Egl_cB_r2_X_b2

3.1.5 Análise estatística pelo método “Significance Analysis of Microarrays” (SAM)

Após o tratamento estatístico e validação dos dados normalizados dos microarranjos, a seleção dos genes diferencialmente expressos foi realizada pelo método SAM ou, em língua portuguesa, “Análise da Significância de Microarranjos” (Tusher *et al.*, 2001). Este método é geralmente utilizado nas análises de dados de microarranjos realizados pela NimbleGen (Li & Waldbieser, 2006; Pischke *et al.*, 2006). O algoritmo SAM faz parte do pacote de ferramentas de bioinformática para análise de microarranjos denominado *MultiExperiment Viewer* (MeV; Saeed *et al.*, 2003) disponibilizado livremente (<http://www.tm4.org/mev.html>) pelo *The Institute for Genomic Research* (TIGR)

3.1.6 Seleção de genes constitutivos a partir dos microarranjos

Para encontrar os genes constitutivos, a análise final dos resultados dos microarranjos foi realizada utilizando-se todos os dados de hibridizações de cDNAs derivados dos tecidos vegetais e formando cinco grupos. Os parâmetros utilizados conforme descrito em Tusher *et al.* (2001) foram (i) número de permutações, 100; (ii) valor de delta, 0,9; (iii) valor de FDR de 0%; e (iv) *fold change* de 2. As seguintes designações dos grupos foram utilizadas:

Para a análise com SAM, foram criados cinco grupos, totalizando 20 valores de expressão gênica diferentes, onde:

Grupo1. Folhas da matriz A de *E. grandis*. Estão representados os dois clones (ou replicatas biológicas) e cada um dos valores dos blocos do *slide*.

Egr_cA_r1_L_b1
Egr_cA_r1_L_b2
Egr_cA_r2_L_b1
Egr_cA_r2_L_b2

Grupo 2. Xilema da matriz A de *E. grandis*: os dois clones e cada um dos valores dos blocos do *slide*.

Egr_cA_r1_X_b1
Egr_cA_r1_X_b2
Egr_cA_r2_X_b1
Egr_cA_r2_X_b2

Grupo 3. Xilema da matriz B de *E. grandis*: os dois clones e cada um dos valores dos blocos do *slide*.

Egr_cB_r1_X_b1
Egr_cB_r1_X_b2
Egr_cB_r2_X_b1
Egr_cB_r2_X_b2

Grupo 4. Xilema da matriz A de *E. globulus*: os dois clones e cada um dos valores dos blocos do *slide*.

Egl_cA_r1_X_b1
Egl_cA_r1_X_b2
Egl_cA_r2_X_b1
Egl_cA_r2_X_b2

Grupo 5. Xilema da matriz B de *E. globulus*: os dois clones e cada um dos valores dos blocos do *slide*.

Egl_cB_r1_X_b1
Egl_cB_r1_X_b2
Egl_cB_r2_X_b1
Egl_cB_r2_X_b2

O critério utilizado na seleção foi o coeficiente de variação entre os valores de expressão de cada amostra. Esse valor deveria ser menor que 0,2, ou seja, menos de 20% de variação nas expressões de um mesmo gene entre diferentes tecidos. Foi realizada, ainda, uma análise por BLASTx no banco de dados de seqüências peptídicas de plantas, tendo sido descartados aqueles genes que apresentaram identidade de “*no hits*”.

3.1.7 Seleção dos genes diferencialmente expressos entre xilemas de *E. globulus* e *E. grandis*

Para selecionar os genes diferencialmente expressos entre xilemas de *E. globulus* e *E. grandis*, os parâmetros utilizados foram (i) número de permutações, 12870; (ii) valor de delta, 3.29; (iii) valor de FDR de 0%; e (iv) *fold change* de 2. As seguintes designações dos grupos foram utilizadas:

Grupo 1. Xilemas das matrizes A e B de *E. grandis*, incluindo os clones e os blocos do *slide*.

Egr_cA_r1_X_b1
Egr_cA_r1_X_b2
Egr_cA_r2_X_b1
Egr_cA_r2_X_b2
Egr_cB_r1_X_b1
Egr_cB_r1_X_b2
Egr_cB_r2_X_b1
Egr_cB_r2_X_b2

Grupo 2. Xilemas das matrizes A e B de *E. globulus*, incluindo os clones e os blocos do *slide*.

Egl_cA_r1_X_b1
Egl_cA_r1_X_b2
Egl_cA_r2_X_b1
Egl_cA_r2_X_b2
Egl_cB_r1_X_b1
Egl_cB_r1_X_b2
Egl_cB_r2_X_b1
Egl_cB_r2_X_b2

A anotação foi refeita com a *Basic Local Alignment and Search Tool* (BLAST), utilizando-se as seqüências deduzidas de peptídeos dos 60 genes candidatos contra o banco de seqüências protéicas derivadas de plantas (BLASTx) disponíveis pelo sítio do *National Center for Biotechnology Information* (NCBI, <http://www.ncbi.nlm.nih.gov/blast/>) na *internet*.

3.2 Seleção de genes constitutivamente expressos de acordo com a literatura

3.2.1 Genes constitutivos clássicos

Foram escolhidos quatro genes constitutivos normalmente usados em experimentos de Northern blots e RT-PCRs convencionais. Os genes escolhidos foram aqueles codificadores da histona H2B (acesso no EMBL-GenBank código AJ010592), da ribonucleoproteína L23A (Rib, código AL078621) e gliceraldeído-3-fosfato-desidrogenase (GAPDH, AF140546). A S-adenosil-metionina-sintetase (SAM, AY635050) também foi incluída nas análises por ter apresentado perfil constitutivo em diferentes bibliotecas de cDNA em análises anteriores.

3.3.2 Novos genes constitutivos de Arabidopsis thaliana

De acordo com o estudo realizado por Czechowski *et al.* (2005), foram selecionados quatro novos genes constitutivos descritos para *Arabidopsis*. Seqüências dos possíveis ortólogos de eucalipto foram identificadas por meio de um BLASTn contra a base de dados do Genolyptus. Apenas seqüências que apresentaram uma alta homologia foram selecionadas e *primers* foram projetados nas regiões mais conservadas. Os genes candidatos incluídos na seleção referem-se à (i) serina/treonina proteína fosfatase PP2A2 (Serine/Threonine protein phosphatase, PP2A2, código At1g59830); proteína da família SAND (*SAND family protein*, código At2g28390); proteína semelhante às da família TIP41 (*TIP41-like family protein*, código At4g34270) e proteína expressa em *A. thaliana* (*expressed protein in A. thaliana*, código At4g33380).

3.3 RT-PCR Quantitativa em tempo real (Real-time Quantitative RT-PCR ou qRT-PCR)

3.3.1 Delineamento de primers

Os *primers* para as análises de RT-PCR quantitativa em tempo real (qRT-PCR) foram desenhados utilizando-se a ferramenta criada pela “*GenScript Corporation’s Real-Time PCR Primer Design*” (www.genscript.com/ssl-bin/app/primer). Por meio da comparação das seqüências agrupadas, os *primers* foram projetados de modo a serem o mais específico possível para os genes selecionados. Foram selecionados apenas aqueles resultantes em amplicons com mais de 100 pb e com temperaturas médias de hibridização em torno de 60 °C. Todos os *primers* foram inicialmente testados via qRT-PCR e a especificidade da amplificação foi verificada pela análise das curvas de dissociação pelo programa *7500 System SDS Software* (Applied Biosystems).

3.3.2 Síntese de cDNA

As reações de síntese dos cDNAs foram realizadas a partir de 300 ng de mRNAs e oligo-dT_{30(V)} (SMART CDS II A, RW Genes/Clontech). Para a síntese, utilizou-se, por reação, 200 unidades da enzima M-MLV RT (Promega) e o respectivo tampão de reação, 5 mM de desoxirribonucleosídeos trifosfatados (dNTPs), 10 pmol do oligo-dT. A reação foi realizada com uma temperatura inicial de 70 °C durante 5 minutos e temperatura de extensão de 40 °C por 1 hora. Para cada um dos tecidos, foi executado uma única reação de síntese de cDNA.

3.3.3 qRT-PCR

Os cDNAs utilizados foram diluídos 50 vezes em água ultrapura. As mesmas alíquotas foram utilizadas em todos os experimentos de qRT-PCR. Foram realizadas quadruplicatas experimentais e foram incluídos controles sem cDNA para cada par de *primers*. As qRT-PCRs foram efetuadas no equipamento *ABI Prism 7500 Detection System* (Applied Biosystems) utilizando-se parâmetros de amplificação recomendados pelo

fabricante (5 minutos a 94 °C; 40 ciclos de 94 °C por 15 segundos, 60 °C por 10 segundos, 72 °C por 15 segundos e 60 °C por 35 segundos). As reações de amplificação dos cDNAs foram realizadas em um volume final de 20 µL contendo 4 pmol dos *primers* gene-específicos, 0,5 mM de dNTPs, 0,05 µL da enzima *Platinum Taq* (Invitrogen) com seu respectivo tampão de reação, SYBR Green 1X e 10 µL dos cDNAs 1:50. As medidas dos produtos de amplificação foram realizadas pela incorporação do marcador fluorescente *SYBR Green* (Applied Biosystems) na dupla fita de DNA. A especificidade das amplificações e o ponto de corte C_T (*threshold cycles*) foram determinados utilizando-se o programa *7500 System SDS Software* (Applied Biosystems). A quantificação foi realizada pelo método $2^{-\Delta\Delta C_T}$, utilizado para calcular mudanças relativas na expressão de genes avaliados por qRT-PCR (Livak & Schmittgen, 2001). O grau da expressão relativa foi calculado a partir das eficiências das reações e do desvio do C_T da amostra-alvo *versus* o controle. Para tanto, e por se tratar de genes constitutivos, foi utilizado como normalizador, inicialmente, o gene de ribonucleoproteína (Rib).

4. Resultados e Discussão

Dois objetivos principais orientaram a realização deste trabalho. O primeiro visou analisar a estabilidade de expressão de genes potencialmente constitutivos, incluindo aqueles freqüentemente utilizados como controles internos em estudos de expressão gênica e potenciais novos genes, e a confirmação destes como referências adequadas aos estudos em *Eucalyptus*. Em segundo lugar, foi realizada também a identificação e a seleção de genes que apresentaram a mais alta expressão diferencial entre xilemas de *Eucalyptus grandis* e *Eucalyptus globulus*, como possíveis justificativas das diferenças observadas entre as qualidades das madeiras destas espécies. Visando atingir estes objetivos, fundamentalmente três conjuntos de experimentos e análises foram conduzidos: hibridizações de microarranjos de DNA; bioinformática; e qRT-PCR.

4.1. Hibridizações de microarranjos de DNA visando a seleção de genes-referência constitutivamente expressos em Eucalyptus e diferencialmente expressos entre xilemas de E. grandis e E. globulus

4.1.1 Preparação de RNAs, cDNAs marcados e montagem dos microarranjos

Amostras de RNA total utilizadas nas análises de microarranjos e qRT-PCRs foram preparadas a partir de xilemas de árvores de quatro anos de idade das espécies *E. grandis* e *E. globulus* plantadas a campo no Hortoflorestal Barbanegra da empresa Aracruz Celulose S.A. As amostras foram coletadas de dois clones derivados de duas matrizes geneticamente não relacionadas, ou seja, derivadas de pais distintos. Assim, os clones de cada matriz representam a mais perfeita réplica biológica possível em condições de campo, já que possuem genomas idênticos, mesma idade e encontram-se submetidos às mesmas condições de solo, clima e agentes biológicos. De dois dos clones de uma das matrizes de *E. grandis*, também foram coletadas folhas maduras, isto é, totalmente expandidas visando à investigação, por microarranjos, das diferenças entre folhas e xilemas dessa espécie. As identidades das amostras estão descritas na Tabela 1 da Seção 3.1.1 (Material e Métodos).

As extrações de RNA foram realizadas com o *kit PureLink Plant RNA Reagent* (Invitrogen) de acordo com o protocolo fornecido pelo fabricante e renderam uma boa quantidade e qualidade de RNA total. A integridade das amostras de RNA pode ser observada na Figura 1, onde a localização e a intensidade das bandas referentes ao RNA ribossômico 28S e ao RNA ribossômico 18S são indicadoras da baixíssima ou ausente degradação das amostras (Ausubel *et al.*, 1997). A análise da qualidade e da quantidade dos RNAs também foi realizada por leitura em espectrofotômetro nos comprimentos de onda de 260 nm e de 280 nm. Os resultados demonstrados na Tabela 2 confirmaram os baixos índices de polissacarídeos, compostos fenólicos e proteínas contaminantes, determinados pelas razões das absorbâncias Abs_{260}/Abs_{280} superiores a 1,8 (Ausubel *et al.*, 1997). Também foi possível confirmar a alta concentração das amostras de RNA total, principalmente daquelas derivadas de xilema dos clones de *E. globulus* (G7 a G10), em torno de 1,5 a 2,5 μg de RNA.

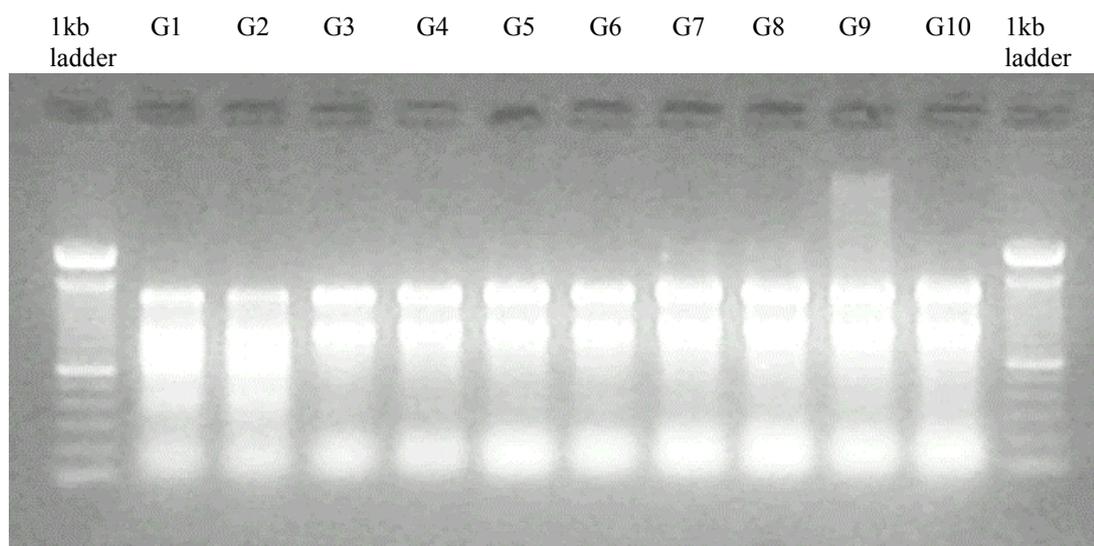


Figura 1. Imagem do resultado da eletroforese em gel de agarose a 1,5% em tampão TAE 1X corado com 0,01 mg/L de brometo de etídio. Foram aplicadas no gel amostras de 5 μL (derivados de volumes totais de 330 μL) de RNA total extraído de folhas maduras de dois clones de *E. grandis* (G1 e G2); xilemas de quatro clones de *E. grandis* (G3-G6); e xilemas de quatro clones de *E. globulus* (G7-G10). Como marcador de tamanhos de fragmentos de DNA foi utilizado o *GeneRuler 1 kb DNA Ladder* (Fermentas).

A qualidade das amostras foi confirmada por re-análises realizadas por técnicos da NimbleGen utilizando o equipamento baseado em eletroforese capilar *Bioanalyzer* (*Agilent Technologies Inc.*). O resultado está representado na Figura 2. As múltiplas bandas de

RNA foliar diferentes dos rRNAs 28S e 18S visualizadas nas canaletas 1 e 2, representam rRNAs de cloroplastos; 16S e 23S (Patterson & Smillie, 1971; Smart *et al.*, 1999; Tattersall *et al.*, 2005; Wang *et al.*, 2005). É possível observar também mRNAs de genes altamente expressos em tecidos fotossintetizantes como a ribulose-bifosfato descarboxilase (RUBISCO) (Krishnan *et al.*, 1986; Smart *et al.*, 1999).

Tabela 2. Resultados das avaliações espectrofotométricas das amostras de RNA total preparadas para ensaios de hibridização de microarranjos de DNA.

| Amostras de RNA | Abs ₂₆₀ | Abs ₂₈₀ | 260/280 | Concentração (ng/μL) | 20 μg (μL) |
|-----------------|--------------------|--------------------|---------|----------------------|------------|
| G1 | 0.053 | 0.000 | ----- | 212 | 94.339 |
| G2 | 0.077 | 0.013 | 5.923 | 308 | 64.935 |
| G3 | 0.117 | 0.036 | 3.250 | 468 | 42.735 |
| G4 | 0.130 | 0.043 | 3.023 | 520 | 38.461 |
| G5 | 0.211 | 0.081 | 2.604 | 844 | 23.693 |
| G6 | 0.181 | 0.062 | 2.919 | 724 | 27.624 |
| G7 | 0.634 | 0.284 | 2.232 | 2.536 | 7.886 |
| G8 | 0.448 | 0.212 | 2.301 | 1.952 | 10.246 |
| G9 | 0.354 | 0.144 | 2.458 | 1.416 | 14.124 |
| G10 | 0.409 | 0.196 | 2.420 | 1.636 | 12.225 |

A partir das amostras de RNA total encaminhadas, os passos seguintes foram realizados pela NimbleGen e incluem a síntese de cDNAs com a marcação com o grupo fluoróforo Cy3, hibridizações, lavagens, *scanning*, coleta de dados e normalizações preliminares dos dados, conforme metodologias da empresa.

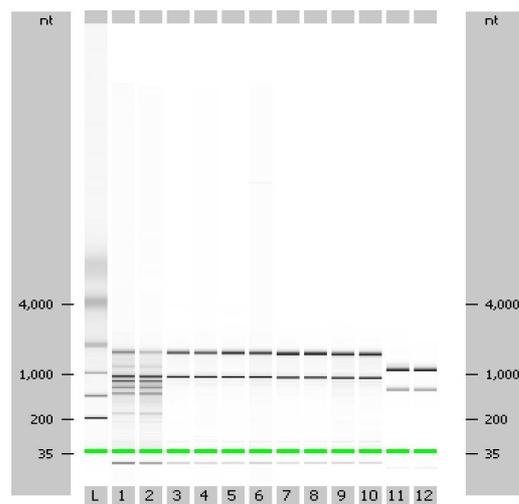


Figura 2. Imagem derivada da eletroforese capilar de fração das amostras de RNA total de *Eucalyptus* enviadas à NimbleGen e executada por técnicos daquela empresa com o equipamento *Bioanalyzer* (*Agilent Technologies Inc.*). A ordem de amostras aplicadas foi: (L) Marcador de tamanhos de fragmentos de DNA (origem não informada); (1 a 10) amostras G1 a G10, conforme legenda da Figura 1; (11 e 12) Controles de RNAs ribossômicos de eucariotos (origem não informada).

A Figura 3 (Clarke & Zhu, 2006) ilustra, de forma geral, o processo experimental e os subsequentes passos de análise dos dados de um experimento de microarranjo. O processo experimental, que inicia na definição dos objetivos, projeção do experimento e seleção do material biológico apropriado para testar a hipótese de interesse em nível transcricional (nesse caso, diferenças entre os xilemas de *E. grandis* e *E. globulus* e entre folhas e xilemas de *E. grandis*) foi realizado pelo nosso grupo. A análise dos dados foi realizada por ambas as partes e consistiu em duas etapas: (i) preparação e processamento dos dados e (ii) identificação dos genes candidatos e interpretação biológica. Os passos de processamento dos dados visaram eliminar o *background*, quantificar os sinais de fluorescência, examinar a qualidade dos dados e normalizá-los para a análise. As etapas de identificação dos genes candidatos e interpretação de dados visaram o reconhecimento dos padrões de expressão gênica e a compreensão do significado biológico de tais padrões (Clarke & Zhu, 2006).

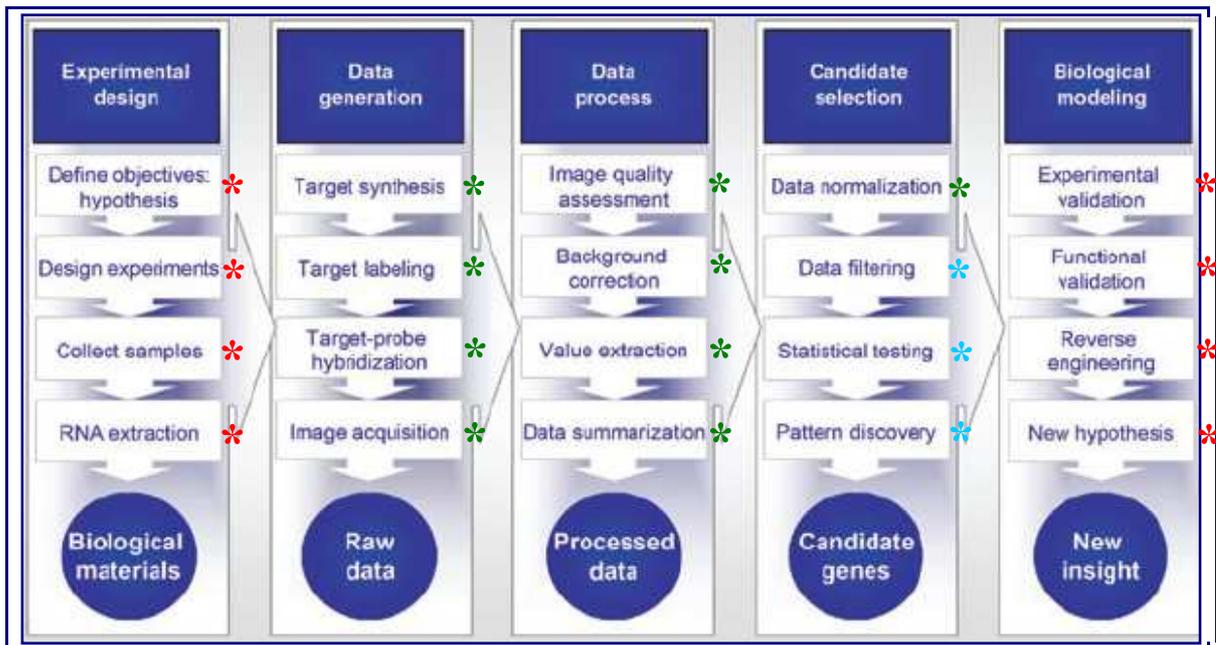


Figura 3. Etapas da confecção e análise dos resultados dos microarranjos, segundo Clarke e Zhu (2006). Assinalado em verde estão as etapas realizadas pela NimbleGen, em vermelho as etapas realizadas pelo grupo integrante do Projeto Genolyptus e, em azul, aquelas realizadas por ambos.

Na arquitetura dos microarranjos, cada gene foi avaliado por um conjunto de nove sondas, em duplicata, as quais apresentam distintas constantes de afinidade pelo alvo e, conseqüentemente, diferentes níveis de intensidade do sinal. Para que a medida final de intensidade represente uma quantificação do nível de mRNA presente na amostra hibridizada, é necessária a transformação dos valores de intensidade, referentes às sondas, em um único valor de intensidade final (Irizarry *et al.*, 2003).

4.2. Análises estatísticas visando a validação e interpretação de dados dos microarranjos

Análises foram realizadas a fim de avaliar a consistência dos dados dos microarranjos fornecidos pela plataforma da NimbleGen. Para tanto, foram realizadas comparações entre diversos grupos de amostras pelo Teste *t* (resultados não apresentados).

A consistência dos dados foi comprovada em termos da baixa variação observada entre os blocos do mesmo *slide* e entre as réplicas biológicas, ou seja, com uma elevada reprodutibilidade. Assim, como resultado, foi possível validar a robustez dos microarranjos realizados pela NimbleGen.

Conforme descrito na Seção “Material e Métodos”, as análises dos dados normalizados enviados pela NimbleGen Systems Inc. referentes aos dez microarranjos de DNA foram realizadas pelo método SAM (Tusher *et al.*, 2001). O algoritmo SAM faz parte do pacote MeV, disponibilizado pelo TIGR, e é uma alternativa ao Teste *t* de Student para busca de expressão diferencial, já que mais de duas amostras podem ser avaliadas simultaneamente. O método consiste no cálculo de uma estatística de expressão diferencial que é uma modificação do Teste *t* e se baseia na razão entre a mudança na expressão gênica entre duas condições e o erro padrão das medições. O algoritmo SAM permite identificar genes com mudanças estatísticas significantes na expressão por meio da realização de conjuntos de Testes *t* gene-específicos. Genes com valores maiores do que o ponto de corte são considerados potencialmente significantes. A porcentagem desses genes identificados ao acaso é a denominada “falsa razão de descobertas” (*False Discovery Rate*, ou FDR). O algoritmo gerou um gráfico interativo onde estão os valores observados *versus* os valores esperados (baseados nos dados de permutação ou *d-values*, Figura 4). O ponto de corte pôde ser ajustado para identificar menores ou maiores conjuntos de genes de acordo com o valor do parâmetro delta e as FDRs foram calculadas para cada conjunto. O valor de delta foi ajustado manualmente no gráfico e significa a distância vertical (em unidades de gráfico) da linha sólida que representa o grau de inclinação (quando o observado é igual ao esperado) e entre duas linhas paralelas pontilhadas que representam a região dentro da qual os genes são considerados com expressão diferencial significativa. Foi utilizado, ainda, o critério de *fold change* que determina o número de vezes que um gene é mais ou menos expresso em relação a outro (MeV Manual, 2006).

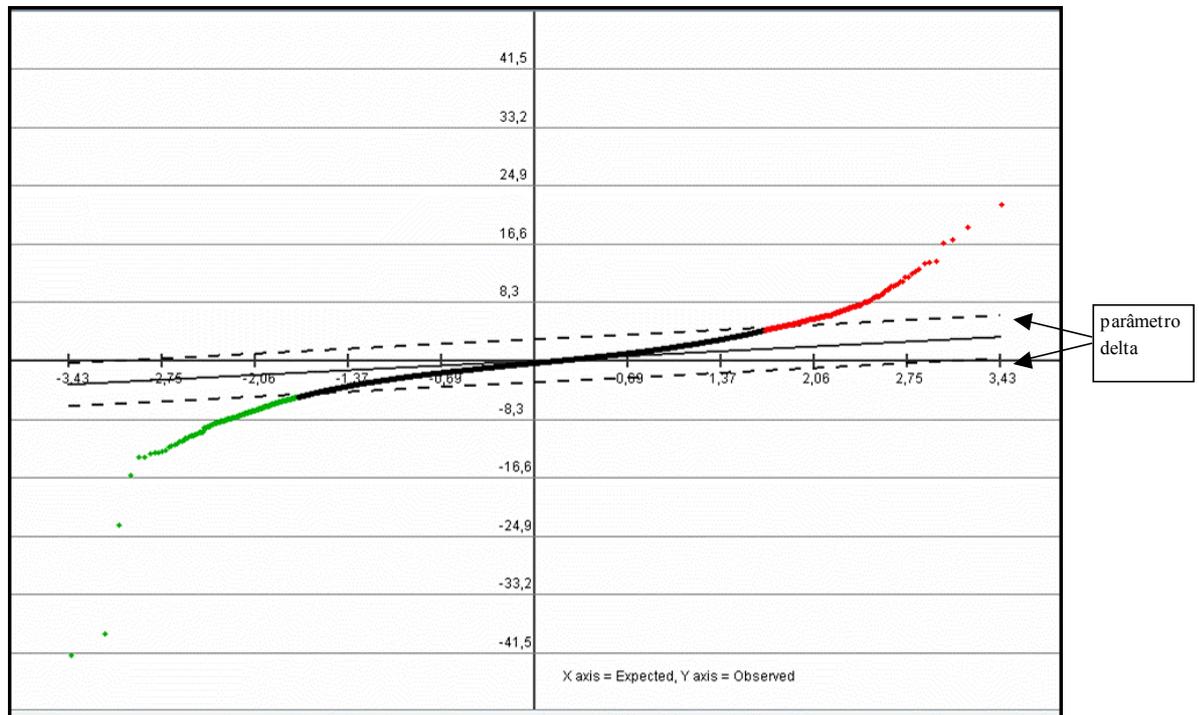


Figura 4. Gráfico gerado pelo algoritmo SAM na análise dos genes mais diferencialmente expressos entre xilemas de *E. grandis* e *E. globulus*. No eixo "x" estão representados os valores esperados e no eixo "y" os valores observados. A linha sólida representa a situação observado = esperado. Em cor verde estão representados os genes mais expressos em xilema de *E. grandis*. Em cor vermelha estão identificados os genes mais expressos em xilema de *E. globulus*. Os pontos de cor negra localizados no centro do gráfico representam os supostos genes com expressão menos diferencial, isto é, os genes com expressão menos variável (constitutiva) entre os xilemas analisados.

No presente trabalho, dois diferentes tipos de análise foram realizadas: (i) a busca por genes diferencialmente expressos entre xilemas de *E. globulus* e *E. grandis* (Seção 1.3); e (ii) a seleção de genes com expressão constitutiva em *Eucalyptus* (Seção 1.4).

4.3 Genes diferencialmente expressos entre xilemas de *E. globulus* e *E. grandis*

Com vistas a identificar os genes mais diferencialmente expressos entre xilemas de *E. globulus* e *E. grandis*, em busca de justificativa, pelo menos parcial, das razões das diferenças tão significativas entre a qualidade da madeira destas arbóreas (vide a Seção 1.3, Introdução), foram comparados os dados normalizados dos microarranjos hibridizados com cDNAs derivados dos xilemas das matrizes A e B de *E. grandis* e os xilemas das matrizes A e B de *E. globulus*, incluindo, para ambos, ambos os clones (réplicas biológicas) e os dois blocos de cada *slide* (réplicas técnicas ou experimentais):

Grupo 1: *E. grandis*

Egr_cA_r1_X_b1
Egr_cA_r1_X_b2
Egr_cA_r2_X_b1
Egr_cA_r2_X_b2
Egr_cB_r1_X_b1
Egr_cB_r1_X_b2
Egr_cB_r2_X_b1
Egr_cB_r2_X_b2

Grupo 2: *E. globulus*

Egl_cA_r1_X_b1
Egl_cA_r1_X_b2
Egl_cA_r2_X_b1
Egl_cA_r2_X_b2
Egl_cB_r1_X_b1
Egl_cB_r1_X_b2
Egl_cB_r2_X_b1
Egl_cB_r2_X_b2

Foram realizadas 12.870 permutações entre os dois grupos e foi utilizado um *fold change* de 2, ou seja, apenas genes com valor de expressão diferencial mínimo de duas vezes foram selecionados. Como resultado desta análise, foram obtidos 471 genes com expressão maior em *E. globulus* e 427 genes com maior expressão em *E. grandis*. Desse total de aproximadamente 900 genes, com base nos maiores valores de *fold change*, foram selecionados os 30 genes candidatos mais expressos em xilema de *E. grandis* e os 30 mais expressos em xilema de *E. globulus* e seus perfis de expressão analisados *in silico* (Figura 5). A seguir, uma busca pelos clones de cDNA potencialmente mais completos e representativos de cada um dos agrupamentos representados pelos 60 supostos genes foi realizada no banco de dados de ESTs do Projeto Genolyptus. Em seguida, conforme já referido na Seção “Material e Métodos”, a anotação foi refeita utilizando-se a ferramenta BLASTx utilizando-se as seqüências deduzidas de peptídeos dos 60 genes candidatos contra o banco de seqüências protéicas derivadas de plantas disponíveis pelo NCBI (Tabelas 3 e 4). Uma segunda seleção foi realizada baseada nas identidades moleculares e

homologias dos genes candidatos e seus prováveis papéis no metabolismo em células vegetais. No total, treze genes foram selecionados, seu perfil de expressão analisado *in silico* e *primers* para qRT-PCR foram projetados. Finalmente, treze genes foram selecionados e *primers* para qRT-PCR foram projetados (Tabelas 5 e 6). Uma análise *in silico* do perfil de expressão de cada um dos genes foi realizada com base em gráficos de expressão gerados pelo algoritmo SAM (Anexo 1). Por meio dessa análise, foi possível verificar o alto nível de expressão diferencial entre os genes selecionados. A confirmação ou não do caráter “diferencialmente expresso” dos genes-candidatos selecionados deverá ainda ser realizada por qRT-PCR.

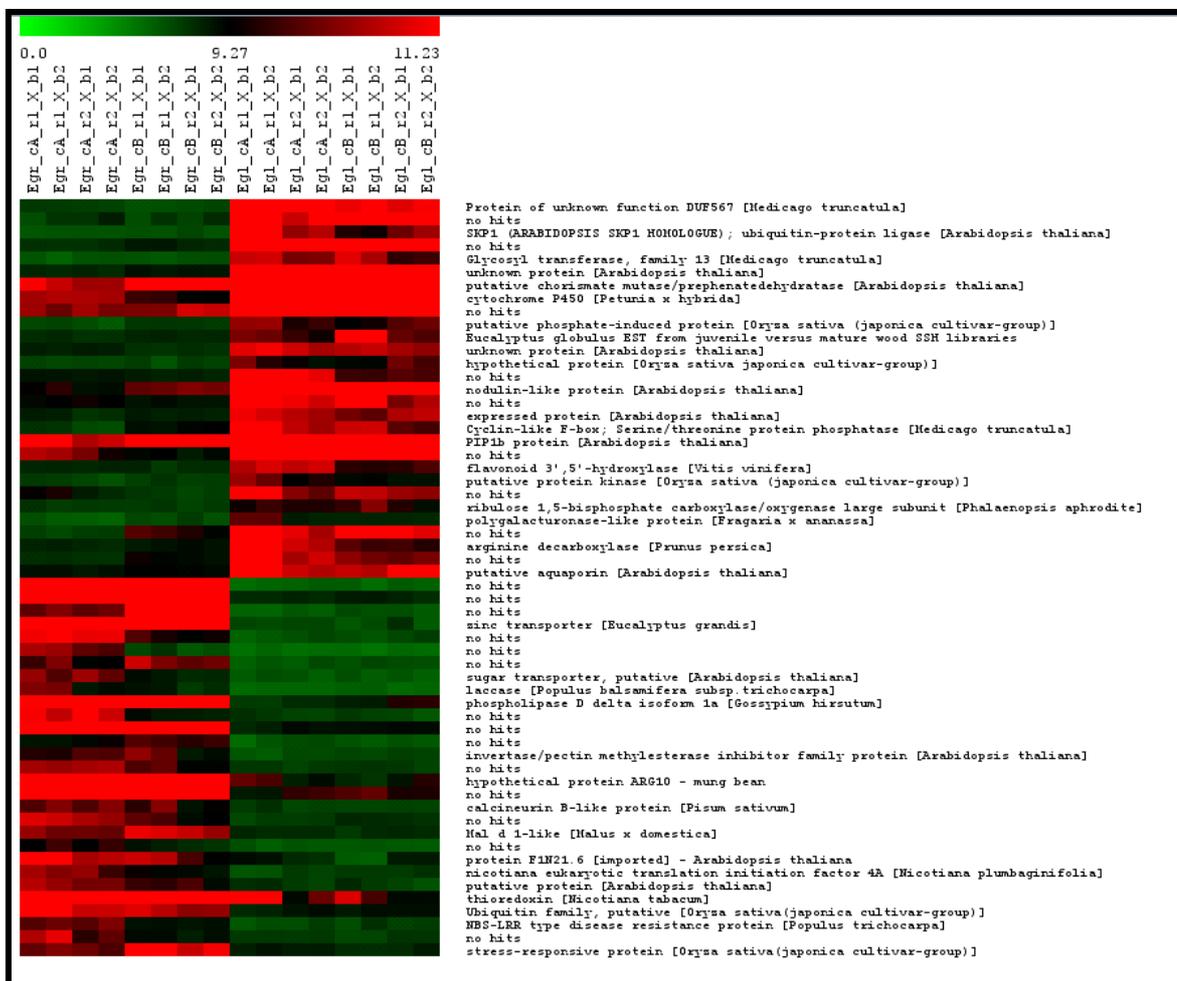


Figura 5. Perfil de expressão dos 60 genes mais diferencialmente expressos entre xilemas de *E. globulus* e *E. grandis* de acordo com análises realizadas pelo método SAM. Em verde estão representados os genes com menor expressão e em vermelho aqueles com maior expressão. A nomenclatura referida nas diferentes canaletas na vertical estão definidas na Seção 3.2.1.3 (Material e Métodos).

Tabela 3. Valor de *fold change* para os genes mais diferencialmente expressos em xilema de *E. globulus* frente ao xilema de *E. grandis*. A identidade molecular dos genes foi atribuída por BLASTx. Estão representados, também, os valores de identidade (*e-value*) e o percentual de homologia.

| Fold Change | Identidade (BLASTX) |
|-------------|--|
| 25,522 | Protein of unknown function DUF567 [Medicago truncatula] Length = 215 Score = 287 bits (734), Expect = 8e-77 Identities = 134/196 (68%), Positives = 166/196 (84%) |
| 21,904 | no hit |
| 19,106 | SKP1 (ARABIDOPSIS SKP1 HOMOLOGUE); ubiquitin-protein ligase [Arabidopsis thaliana] Length = 155 Score = 186 bits (471), Expect = 2e-46 Identities = 105/206 (50%), Positives = 135/206 (65%) |
| 18,722 | no hit |
| 17,720 | Glycosyl transferase, family 13 [Medicago truncatula] Score = 49.7 bits (117), Expect = 2e-04 Identities = 21/36 (58%), Positives = 28/36 (77%), Gaps = 0/36 (0%) |
| 14,427 | unknown protein [Arabidopsis thaliana] Score = 87.8 bits (216), Expect = 1e-16 Identities = 45/66 (68%), Positives = 56/66 (84%), Gaps = 2/66 (3%) |
| 8,902 | putative chorismate mutase/prephenatedehydratase [Arabidopsis thaliana] Length = 424 Score = 53.9 bits (128), Expect = 3e-07 Identities = 25/27 (92%), Positives = 26/27 (96%) |
| 8,175 | cytochrome P450 [Petunia x hybrida] Length = 506 Score = 199 bits (507), Expect = 1e-50 Identities = 94/110 (85%), Positives = 102/110 (92%) |
| 8,147 | no hits |
| 7,560 | putative phosphate-induced protein [Oryza sativa (japonica cultivar-group)] Score = 133 bits (335), Expect = 6e-30 Identities = 83/180 (46%), Positives = 109/180 (60%), Gaps = 18/180 (10%) |
| 7,472 | Eucalyptus globulus EST from juvenile versus mature wood SSH libraries Length=208 Score = 123 bits (264), Expect = 5e-26 Identities = 52/61 (85%), Positives = 53/61 (86%), Gaps = 0/61 (0%) |
| 7,266 | unknown protein [Arabidopsis thaliana] Length = 476 Score = 293 bits (749), Expect = 1e-78 Identities = 147/274 (53%), Positives = 187/274 (68%), Gaps = 4/274 (1%) |
| 6,501 | hypothetical protein [Oryza sativa japonica cultivar-group] Length = 243 Score = 149 bits (377), Expect = 2e-35 Identities = 93/227 (40%), Positives = 119/227 (52%), Gaps = 3/227 (1%) |
| 6,258 | no hit |
| 6,205 | nodulin-like protein [Arabidopsis thaliana] Length = 389 Score = 179 bits (454), Expect = 1e-44 Identities = 91/121 (75%), Positives = 100/121 (82%) |
| 5,961 | no hit |
| 5,883 | expressed protein [Arabidopsis thaliana] Length = 177 Score = 265 bits (677), Expect = 2e-70 Identities = 130/174 (74%), Positives = 138/174 (79%), Gaps = 3/174 (1%) |
| 5,750 | Cyclin-like F-box; Serine/threonine protein phosphatase, BSU1 [Medicago truncatula] Length = 341 Score = 199 bits (507), Expect = 1e-50 Identities = 101/175 (57%), Positives = 121/175 (69%), Gaps = 5/175 (2%) |
| 5,738 | putative aquaporin PIP1-2 [Vitis berlandieri x Vitis rupestris] Length = 286 Score = 348 bits (892), Expect = 3e-95 Identities = 174/211 (82%), Positives = 184/211 (87%) |
| 5,733 | no hits |

Tabela 4. Valor de *fold change* para os genes mais diferencialmente expressos em xilema de *E. grandis* frente ao xilema de *E. globulus*. A identidade molecular dos genes foi atribuída por BLASTx. Estão representados, também, os valores de identidade (*e-value*) e o percentual de homologia.

| Fold Change | Identidade (BLASTX) |
|-------------|---|
| 263,308 | no hits |
| 54,598 | no hits |
| 50,000 | no hits |
| 49,241 | zinc transporter [Eucalyptus grandis] Length = 421 Score = 68.6 bits (166), Expect = 1e-11 Identities = 32/35 (91%), Positives = 35/35 (100%) |
| 16,183 | no hits |
| 15,245 | no hits |
| 12,488 | no hits |
| 10,401 | sugar transporter, putative [Arabidopsis thaliana] Length = 470 Score = 194 bits (494), Expect = 4e-49 Identities = 90/125 (72%), Positives = 107/125 (85%) |
| 10,162 | laccase [Populus balsamifera subsp.trichocarpa] Length = 407 Score = 67.4 bits (163), Expect = 3e-11 Identities = 29/30 (96%), Positives = 29/30 (96%) |
| 10,147 | phospholipase D delta isoform 1a [Gossypium hirsutum] Length = 854 Score = 53.9 bits (128), Expect(2) = 1e-07 Identities = 26/54 (48%), Positives = 32/54 (59%) |
| 9,331 | no hits |
| 9,095 | no hits |
| 8,698 | no hits |
| 8,038 | invertase/pectin methylesterase inhibitor family protein [Arabidopsis thaliana] Length = 203 Score = 212 bits (540), Expect = 3e-54 Identities = 105/169 (62%), Positives = 136/169 (80%) |
| 7,513 | no hits |
| 7,471 | hypothetical protein ARG10 - mung bean Length = 237 Score = 110 bits (275), Expect = 3e-24 Identities = 52/57 (91%), Positives = 54/57 (94%) |
| 7,258 | no hits |
| 7,127 | calcineurin B-like protein [Pisum sativum] Length = 225 Score = 188 bits (477), Expect = 3e-47 Identities = 95/111 (85%), Positives = 101/111 (90%), Gaps = 7/111 (6%) |
| 7,086 | no hits |
| 6,869 | Mal d 1-like [Malus x domestica] Length = 163 Score = 234 bits (597), Expect = 4e-61 Identities = 111/162 (68%), Positives = 135/162 (83%), Gaps = 3/162 (1%) |
| 6,820 | no hits |
| 6,777 | protein F1N21.6 [imported] - Arabidopsis thaliana Length = 901 Score = 70.5 bits (171), Expect = 5e-12 Identities = 32/69 (46%), Positives = 46/69 (66%), Gaps = 1/69 (1%) |
| 6,696 | nicotiana eukaryotic translation initiation factor 4A [Nicotiana plumbaginifolia] Length = 391 Score = 76.6 bits (187), Expect(2) = 2e-28 Identities = 33/48 (68%), Positives = 42/48 (87%) |

Tabela 5. Relação de nomes dos supostos genes selecionados com base na maior expressão diferencial em xilema de *E. globulus* frente ao xilema de *E. grandis*. Os valores de identidade atribuídos por BLASTx estão representados pelo valor *E*, ou *E value*. Na Tabela estão representados, também, as seqüências de *primers* indicadas pelo programa *GenScript Corporation's Real-Time PCR Primer Design* e o tamanho esperado dos amplicons para cada seqüência gênica considerada.

| Código NimbleGen | Identidade (BLASTx) | e-value | Amplicon | Forward/Reverse |
|-------------------|---|---------|----------|---|
| LYPTUS00S00009226 | suposta corismato mutase [<i>Arabidopsis thaliana</i>] | 3e-17 | 116 | 5' GATCACGAGAGCAAGACCAC/ 5' CCCTCCTCCTCTTCTTCT |
| LYPTUS00S00003979 | citocromo P450 [<i>Petunia x hybrida</i>] | 1e-50 | 132 | 5' CTCGTTGACTGGAAGCTG/ 5' GGTACATATGCGCTAACGGAA |
| LYPTUS00S00007693 | proteína desconhecida [<i>Arabidopsis thaliana</i>] | 1e-78 | 136 | 5' CTATTTCTGGAGCTCGGCAT/ 5' TGAAGTACAAGCGGTAGGGA |
| LYPTUS00S00000212 | proteína tipo nodulina [<i>Arabidopsis thaliana</i>] | 1e-44 | 122 | 5' GATAGAGAAAGTGCGGTTGGA/ 5' GCTTCGGTTGTATATGACG |
| LYPTUS00S00009542 | proteína expressa [<i>Arabidopsis thaliana</i>] | 2e-70 | 114 | 5' CTACGAGGAGAATGGTGGGT/ 5' TTGTATGGAATTCCTCTTGCC |
| LYPTUS00S00003656 | suposta aquaporina PIP1-2 [<i>Vitis berlandieri</i> x <i>Vitis rupestris</i>] | 3e-95 | 101 | 5' GGTATCTCTGGTGGACACATCA/ 5' CACTGCATGACCATGTAGAACA |
| LYPTUS00S00020347 | flavonóide 3',5'-hidroxilase [<i>Vitis vinifera</i>] | 7e-65 | 106 | 5' CCCTAGACATCAACTTCTCCAA/ 5' AACITCCACCTCAGGCCATA |
| LYPTUS00S00016145 | proteína tipo poligalacturonase [<i>Fragaria x ananassa</i>] | 5e-53 | 151 | 5' GCAGACGTGGAACACCAA/ 5' GCCACTCCGTCATGTCTCT |
| LYPTUS00S00004269 | arginina descarboxilase [<i>Prunus persica</i>] | 3e-94 | 132 | 5' TATTCTGCAGCCATGAGACG/ 5' ACACAGGCCATCAACATCAG |
| LYPTUS00S00007340 | suposta aquaporina [<i>Arabidopsis thaliana</i>] | 1e-93 | 104 | 5' GTGGTGGCCGTGATGATA/ 5' GATCGACAACGGTATGATG |

Tabela 6. Relação de nomes dos supostos genes selecionados com base na maior expressão diferencial em xilema de *E. grandis* frente ao xilema de *E. globulus*. Os valores de identidade atribuídos por BLASTx estão representados pelo valor *E*, ou *E value*. Na Tabela estão representados, também, as seqüências de *primers* indicadas pelo programa *GenScript Corporation's Real-Time PCR Primer Design* e o tamanho esperado dos amplicons para cada seqüência gênica considerada.

| Código NimbleGen | Identidade (BLASTx) | e-value | Amplicon | Forward/Reverse |
|-------------------|--|---------|----------|--|
| LYPTUS00S00007771 | proteína tipo calcineurina B [<i>Pisum sativum</i>] | 3e-47 | 131 | 5' TTGAAGATTGTGCCGAGAAA/ 5' GAATCCATCTCACATTCACAACA |
| LYPTUS00S00006040 | tioredoxina [<i>Nicotiana tabacum</i>] | 2e-42 | 107 | 5' CGAGGCAATGCCAACTTT/ 5' CAGTCGCGTGTGTTTATC |
| LYPTUS00S00003795 | família de proteínas inibitórias de invertase/pectina metil-esterase [<i>Arabidopsis thaliana</i>] | 3e-54 | 142 | 5' GGAACGCGATACTGACCATT/ 5' CGTAGCTGGCTAGGTATGATTG |

Como citado acima, os genes candidatos foram selecionados de acordo com as identidades moleculares, sua função no metabolismo das células vegetais e, ainda, o possível papel na formação da madeira. Abaixo, uma breve descrição das atividades desses genes.

A corismato mutase é a primeira enzima do ramo da via do chiquimato que leva à biossíntese dos aminoácidos aromáticos fenilalanina e tirosina (Eberhard *et al.*, 1993). Os aminoácidos aromáticos fenilalanina, tirosina e triptofano possuem um duplo papel nas

plantas: eles são requeridos na biossíntese de proteínas e precursores de metabólitos aromáticos secundários, incluindo o hormônio auxina, que é um regulador de crescimento da planta (ácido 3-indolacético, AIA), alcalóides antimicrobianos, flavonóides e compostos polifenólicos tais como as ligninas que são requeridas para a estrutura das paredes celulares vegetais. Diferentemente dos metabólitos primários, os metabólitos secundários não são considerados necessários para a sobrevivência celular *per se*, mas são essenciais ao desenvolvimento normal da planta. Alguns desses compostos são produzidos em resposta a sinais ambientais como parte de uma resposta de defesa geral (Mobley et al., 1999). A lignificação da parede celular está relacionada ao mecanismo de defesa da planta durante a interação planta-patógeno. O engrossamento da parede secundária do xilema diminui a extensão da membrana que é exposta às células bacterianas, reduzindo o acesso ao xilema parenquimático. Assim, paredes lignificadas no vaso do xilema podem funcionar como uma barreira à transferência de compostos produzidos pela bactéria para a célula vegetal, como fatores de virulência e avirulência (Hilaire *et al.*, 2001).

Existem 272 genes de citocromo P450 (incluindo 26 pseudogenes) no genoma de *Arabidopsis*, o que classifica estas proteínas como uma das maiores famílias de proteínas de plantas. Acredita-se que o grande número de indivíduos nessa família se deva à ocorrência de duplicação gênica e conversão, e seja resultante da necessidade das plantas de se adaptar ao ambiente e a responder a patógenos. As P450 muitas vezes possuem menos de 20% de identidade entre si e catalisam diversas reações. As suas funções biológicas vão desde a biossíntese de macromoléculas estruturais como lignina, cutina e suberina, à síntese e catabolismo de todos os tipos de hormônios ou moléculas sinalizadoras, de pigmentos e compostos de defesa e o metabolismo de xenobióticos (Werck-Reichhart *et al.*, 2002). As P450 são encontradas mediando processos fundamentais como a síntese e modificação de esteróis, modificação de terpenos cíclicos nas vias do ácido abscísico, brassinoesteróides e giberelinas, hidroxilação de carotenóides, modificação de produtos da via do chiquimato e intermediários. A biossíntese da lignina está acoplada ao metabolismo dos fenilpropanóides, apresentando enzimas compartilhadas com outros processos metabólicos tais como fenilalanina amônio liase (PAL), cinamato 4-hidroxilase (C4H) e ácido caféico O-metiltransferase (COMT) assim como enzimas específicas como cinamoil-CoA redutase (CCR) e cinamil álcool desidrogenase (CAD) (Boerjan *et al.*, 2003). Entre as enzimas dependentes de P450 na via dos fenilpropanóides

encontram-se cinamato 4-hidroxilase (C4H), ferulato 5-hidroxilase (F5H) e p-cumarato 3-hidroxilase (Franke *et al.*, 2002).

A interação entre as leguminosas e bactérias Gram-negativas do gênero *Rhizobium* leva ao desenvolvimento de um órgão inteiramente novo na planta, o nódulo da raiz, conhecido por ser responsável pela fixação simbiótica do nitrogênio. Vários genes são exclusivamente ou predominantemente expressos no nódulo. Os produtos gênicos correspondentes são chamados de nodulinas. Algumas nodulinas estão envolvidas na morfogênese do nódulo e na infecção bacteriana e cerca de 50% das nodulinas de expressão primária representam proteínas ricas em prolina (PRPs), a maioria provavelmente envolvidas na biossíntese da parede celular das plantas (revisado em Schröder *et al.*, 1997). Moreau *et al.* (2005) afirmaram que a alta expressão de proteínas do tipo nodulina nas fibras mortas do xilema indica uma importante função das nodulinas no processo de maturação das fibras do xilema, sinalizando a morte das fibras celulares, bem como nos mecanismos responsáveis pelo controle hormonal, mobilização de nutrientes, regulação da integridade vacuolar e autólise das fibras mortas. Sterky *et al.* (1998) analisaram ESTs provenientes de tecidos formadores de madeira (xilema e câmbio) em *Populus* e a nodulina apresentou-se como um dos transcritos mais abundantes nas bibliotecas analisadas, indicando fortemente seu papel na diferenciação de tecidos vasculares. Paux *et al.* (2004) isolaram genes preferencialmente expressos durante a formação da madeira em *Eucalyptus* pela análise do perfil de transcrição utilizando arranjos de DNA e análise de variância (ANOVA) para identificar genes diferencialmente expressos entre xilema secundário e folhas e encontraram a nodulina como diferencialmente expressa no xilema. Mijnsbrugge *et al.* (2000) encontraram a nodulina entre os cinco genes mais altamente expressos na análise de proteínas preferencialmente produzidas em xilemas em desenvolvimento.

As aquaporinas são poros protéicos que facilitam a difusão passiva da água através das membranas celulares. As aquaporinas de plantas são divididas em grupos baseados na similaridade das seqüências de aminoácidos: proteínas intrínsecas da membrana plasmática (PIPs), proteínas intrínsecas do tonoplasto e proteínas intrínsecas do tipo nodulinas (NIPs) (Aroca *et al.*, 2006). As células em expansão aumentam seu volume principalmente pela aquisição de água e diversos trabalhos demonstram o papel das aquaporinas na divisão celular. As aquaporinas são abundantes, também, em células meristemáticas e pequenos

vacúolos. Chaumont *et al.* (1998), por meio da técnica de hibridização *in situ*, demonstraram uma alta expressão de genes de aquaporinas em meristemas e zonas de crescimento celular como raízes laterais e primórdios foliares. A alta expressão em meristemas e em células em expansão sugere que as aquaporinas são necessárias para a biogênese dos vacúolos e para suportar o rápido influxo de água nestes durante a expansão celular. Em busca da identificação de potenciais reguladores da atividade meristemática no câmbio de *Populus*, a análise do perfil de expressão de meristemas formadores de madeira permitiu identificar vários genes representando aquaporinas em zonas específicas de expansão com padrões de expressão similar àqueles de enzimas como a pectina metil-esterase (Schrader *et al.*, 2004). Quatro genes de aquaporinas tonoplásticas foram descritos como altamente expressos em estádios iniciais e intermediários de desenvolvimento e em índices muito baixos nos estádios maduros, o que confirma a hipótese de uma regulação coordenada da expansão celular durante o crescimento secundário. O aumento do volume e tamanho celulares dependem do relaxamento da parede celular e expansão da célula dirigido pelo influxo de água nos vacúolos. Este padrão de crescimento celular é controlado por uma série de enzimas hidrolíticas da parede celular e pelas aquaporinas (Schäffner *et al.*, 1998). Vários trabalhos sugerem o papel de aquaporinas no controle da expansão celular em plantas, relacionado a expressão de genes para aquaporinas com as zonas de alongamento, divisão e expansão celular em vários tecidos (Ludevid *et al.*, 1992; Chaumont *et al.*, 1998; Balk & de Boer, 1999; O'Brien *et al.*, 2002; Siefritz *et al.*, 2004; Hachez *et al.*, 2006).

A via de síntese dos fenilpropanóides, exclusiva dos vegetais, catalisa a conversão da fenilalanina em uma grande variedade de metabólitos fenólicos secundários, que incluem ligninas, ésteres sinápticos, estilbenos e flavonóides. Estes compostos possuem diversos papéis em processos essenciais nas plantas incluindo crescimento, defesa contra patógenos em diferentes espécies e reprodução (revisado em Winkel-Shirley, 1999). Há também evidências de que os flavonóides atuam como protetores contra a radiação prejudicial às plantas, como agentes capazes de ligar fitotoxinas e com papel na regulação do transporte de auxinas (Saslowky & Winkel-Shirley, 2001). Nos vegetais, as reações de hidroxilação promovidas pela enzima flavonóide-3',5'-hidroxilase (F3'5'H) estão ligadas à biossíntese dos flavonóides (Kaltenbach *et al.*, 1999). Saslowky & Winkel-Shirley (2001) apresentaram a hipótese de que enzimas como a cinamato-4-hidroxilase (C4H) e a F3'5'H

funcionam como âncoras no recrutamento de outras enzimas do metabolismo de fenilpropanóides e flavonóides em complexos de membrana. Diferentes trabalhos relacionaram o fluxo apical-basal de auxinas e a síntese de fenilpropanóides com a diferenciação dos tecidos vasculares. Várias evidências sugerem que o transporte de auxinas e a síntese de fenilpropanóides estão diretamente ligados e co-localizados nos tecidos vasculares (Besseau *et al.*, 2007). De acordo com Yang *et al.* (2003), os flavonóides também são requeridos para a estrutura do xilema. Os autores descrevem o papel dos flavonóides e especialmente da F3'5'H no processo de formação da madeira em *Robinia pseudoacacia*. O alburno (*sapwood*) é a porção mais externa do xilema e contém as células vivas enquanto o cerne (*hardwood*) é a parte mais interna do lenho das árvores, definida como o núcleo morto do eixo lenhoso, tendo como função apenas suporte à árvore, enquanto que o alburno, ou xilema jovem, possui função de condução e de reserva de água e nutrientes. O xilema jovem é gradualmente convertido a cerne inativo. As células do parênquima da madeira passam por diversas mudanças metabólicas e produzem uma grande quantidade de extrativos do cerne como compostos fenólicos, ligninas e substâncias aromáticas que se acumulam nos vasos. Os materiais de reserva nas células do parênquima do alburno são usados para a formação da madeira. A formação do cerne é acompanhada por uma série de alterações no metabolismo, como a senescência. Os autores analisaram os perfis de expressão em troncos de árvores de *R. pseudoacacia* analisando a casca, o alburno e os tecidos da zona de transição de alburno para cerne. O padrão de expressão gênica na zona de transição demonstrou-se relacionada ao metabolismo secundário envolvido na via de biossíntese dos flavonóides. A expressão do gene da F3'5'H teve um grande aumento na zona de transição. O alto nível de expressão de genes relacionados à biossíntese de flavonóides na zona de transição está de acordo com o escurecimento da madeira na região do cerne, que é mediada por estes genes. Outros autores também sugerem o papel da F3'5'H na biossíntese de flavonóides durante a formação da madeira (de Vetten, *et al.*, 1999; Beritognolo *et al.*, 2002).

Ainda com respeito a F3'5'H, é interessante notar que outro dos genes com alta expressão no xilema de *E. globulus* foi o de uma citocromo P450, conforme descrito anteriormente. A F3'5'H é uma enzima redutora ligada à membrana e dependente de P450. Assim, a demonstração da conexão destes dois genes tão diferencialmente expressos entre

os xilemas das espécies de *Eucalyptus* analisadas seria ainda mais promissora na busca dos elementos responsáveis pelas diferenças nas qualidades das madeiras das duas espécies.

Em plantas, as poligalacturonases (PGs) encontram-se expressas em uma ampla variedade de tecidos e estádios de desenvolvimento. As PGs estão associadas com a maturação de frutos, a abscisão de folhas e frutos, a maturação de grãos de pólen, a defesa contra patógenos e em processos de expansão, crescimento e xilogênese (revisado em Atkinson *et al.*, 2002). Torki *et al.* (1999) analisaram a expressão de membros da família de PGs em *Arabidopsis* e observaram um novo tipo de PG expressando-se em tecidos jovens em processo de desenvolvimento como plântulas e raízes, o que sugere o papel desta enzima nas modificações da parede celular observadas durante o processo de alongamento e expansão que ocorre nesses tecidos. As PGs estão entre as pectinases ou enzimas pectinolíticas mais conhecidas. Estas enzimas degradam as substâncias pecticas constituintes da lamela média e são as primeiras a serem produzidas durante o processo da penetração da parede celular e da morte da célula, um dos estágios da formação do xilema secundário (ver Secção 1.4 da Introdução). De acordo com Hadfield & Bennett (1998), a localização de proteínas e de mRNA de PGs em sistemas vasculares em desenvolvimento indica o papel destas enzimas no processo de diferenciação do xilema secundário e separação da parede celular primária dos vasos do xilema. De fato, diversos trabalhos têm constatado o papel das PGs na diferenciação dos traqueídeos e na xilogênese, em diferentes espécies (Hertzberg *et al.*, 2001; Milioni *et al.*, 2001; Nakashima *et al.*, 2004; Yokohama & Nishitani, 2006).

A arginina descarboxilase (ADC) é a enzima-chave na biossíntese de poliaminas em plantas (Bortolotti *et al.*, 2004). A ADC é ativa em células embrionárias, em células em processo de alongamento e submetidas a diferentes situações de estresse (Flores, 1991). Utilizando anticorpos policlonais, Bortolotti *et al.* (2004) demonstraram a presença da ADC em flores, sementes, caules, folhas e raízes. Em tecidos fotossintetizadores, esta enzima localiza-se principalmente nos cloroplastos e nos tecidos não-fotossintetizadores, no núcleo. Pérez-Amador & Carbonell (2005) propuseram o papel da ADC na regulação da expansão celular e nos processos metabólicos secundários como a biossíntese de alcalóides. Os autores analisaram a atividade da ADC em resposta a giberelinas exógenas em ervilhas. As giberelinas induziram um aumento na atividade da ADC, precedendo o aumento das células em tecidos com crescimento ativo. Hummel *et al.* (2004) analisaram

as atividades de dois promotores de ADC em *Arabidopsis* em situações de germinação de sementes, desenvolvimento de raízes e em resposta ao frio e constataram seu envolvimento na regulação da transcrição da ADC nestas situações. Vuosku *et al.* (2006) comprovaram a presença de mRNA e proteínas ADC em diferentes fases da mitose na embriogênese zigótica em *Pinus sylvestris*. Ambos encontraram-se localizados em células em processo de divisão de meristemas embrionários. Esses resultados sugerem um papel essencial da ADC na mitose de células vegetais.

Nas plantas, o íon cálcio está presente em todo o espaço intracelular e é um mensageiro secundário envolvido em diversas vias de sinalização. Variações nas concentrações citosólicas de Ca^{2+} reúnem um grande conjunto de respostas e sinais, como o estabelecimento de respostas de defesa vegetal (Lecourieux *et al.*, 2006). O cálcio possui um importante papel na regulação da expressão gênica e muitos outros processos como sinalização em resposta ao estresse abiótico. Apesar de os mecanismos moleculares envolvendo o papel do cálcio em funções celulares não estar totalmente elucidado, muitos estímulos externos incluindo luz e vários fatores de estresses bióticos e abióticos podem levar a mudanças no nível do Ca^{2+} celular, afetando o crescimento e o desenvolvimento da planta (Mahajan *et al.*, 2006). Muitos processos inerentes de crescimento e desenvolvimento como a prolongação das raízes e a formação do tubo polínico são controlados pelo Ca^{2+} (Franklin-Tong *et al.*, 1997). As proteínas semelhantes à calcineurina B (do inglês, *calcineurin B-like proteins* ou CBLs) são pequenas proteínas que contêm múltiplos domínios de ligação ao Ca^{2+} , transmitindo o sinal pela interação com proteínas e regulando a sua atividade (Luan *et al.*, 2002). O cálcio liga-se as CBLs resultando na ativação da quinases envolvidas em respostas a estresses por excesso salino e em sinalização por ABA (revisado em Hrabak *et al.*, 2003). As quinases reguladas por este íon provêm pontos de interconexão para múltiplas vias sinalizadoras que integram sinais do íon em todos os aspectos de desenvolvimento e crescimento vegetal. Os sinais de cálcio têm também um importante papel na nas funções “nutricionais”, incluindo o correto funcionamento de vias secretoras e a integridade funcional das paredes celulares (Harper *et al.*, 2004). Por regular o fluxo interno de Ca^{2+} , as CBLs são também responsáveis por importantes eventos como os movimentos das plantas (fototropismo e fechamento e abertura de estômatos) e as mudanças massivas no influxo de cálcio e potássio, alterando o turgor das célula. Finalmente, as CBLs também foram descritas como vitais ao controle do

fluxo interno de íons nas raízes, processo crucial para o crescimento, desenvolvimento e sobrevivência das plantas (Trewavas, 1999).

As tioredoxinas são bissulfeto redutases que regulam o estado redox de proteínas-alvo. Embora as tioredoxinas vegetais apresentem uma grande diversidade não encontrada em outros organismos, muitos dos seus papéis fisiológicos ainda não foram determinados. Uma das funções mais conhecidas das tioredoxinas é o seu envolvimento na tolerância das plantas ao estresse oxidativo. Enquanto a maioria dos organismos possui um baixo número de tioredoxinas, as plantas apresentam diversos tipos desta enzima. No genoma de *Arabidopsis* foram identificados 42 genes potencialmente codificadores destas enzimas (Meyer *et al.*, 2005). Algumas tioredoxinas apresentam funções altamente específicas como, por exemplo, a HCF164, envolvida na biogênese do complexo da citocromo *b₆f* em *Arabidopsis*, presente nas membranas tilacóides de plantas superiores, e possuindo uma posição central no transporte de elétrons (Lennartz *et al.*, 2001). As tioredoxinas presentes no cloroplasto regulam a atividade de enzimas envolvidas no metabolismo do carbono (Ruelland & Miginiac-Maslow, 1999). Um trabalho realizado por Wong *et al.* (2002) com plantas de cevada superexpressando um gene codificador de tioredoxina demonstrou a participação de uma tioredoxina citosólica na mobilização das reservas de sementes durante a germinação. Análises por duplo-híbrido mostraram um clone de tioredoxina interagindo com o domínio de ligação ao zinco de um clone de celulose sintase, sugerindo que essas proteínas podem estar envolvidas na modulação redox desta enzima (Revisado em Doblin *et al.*, 2002). Estudo realizado por Potikha *et al.* (1999), analisando o envolvimento do peróxido de hidrogênio na diferenciação das paredes celulares secundárias em fibras de algodão, concluíram que um ambiente oxidativo favorece a síntese da celulose. Isso é afirmado com base na observação da ocorrência de um forte aumento na produção do peróxido de hidrogênio durante a formação da parede celular secundária em fibras de algodão. Estudo realizado em 2002 por Kurek *et al.* fornece evidências de que a dimerização das subunidades catalíticas da celulose sintase é sujeita à modulação redox. Condições oxidativas favorecem a formação de complexos da celulose sintase e a síntese da celulose, enquanto monômeros da celulose sintase estão sujeitos à degradação em situações de redução. Uma vez mais, a demonstração da participação da citocromo P450, com expressão também aumentada em xilemas de *E. globulus*, nos processos redox mediados por tioredoxinas seria muito interessante.

A pectina metilesterase (PME) é a primeira enzima a atuar na pectina, um dos principais componentes da parede celular. A atividade da PME é responsável pela produção de pectinas com diferentes propriedades estruturais e funcionais, o que indica o importante papel desta proteína na fisiologia das plantas (Giovane *et al.*, 2004). A PME encontra-se ligada à parede celular e está envolvida no processo de maturação de frutos e na expansão da parede durante o crescimento da célula. O papel da enzima na extensão celular consiste, essencialmente, na diminuição do pH local que, por sua vez, ativa as enzimas envolvidas na autólise da parede celular (revisado em Balestrieri *et al.*, 1990). As PMEs catalisam a desmetilesterificação das pectinas. O grau da desmetilesterificação determinará a solidez da parede celular. Assim, processos fisiológicos requeridos para o rearranjo da arquitetura da parede celular são afetados diretamente pela atividade da PME (Micheli, 2001). A regulação da atividade da PME é realizada pela expressão diferencial de várias isoformas em diferentes tecidos e estádios do desenvolvimento, e por modificações sutis no pH da parede celular. A família de proteínas inibitórias de invertase/pectina metilesterase (PMEI) é constituída por glicoproteínas denominadas de invertases que inibem a atividade da PME pela formação de um complexo 1:1 reversível (PME-PMEI). A estabilidade desse complexo é altamente influenciada pelo pH, o que indica, mais uma vez, que a modulação da atividade da PME está diretamente ligada ao pH. A regulação também está na modulação da afinidade entre a enzima e seu inibidor (Giovane *et al.*, 2004). Nos vegetais, muitas enzimas modificadoras de carboidratos têm sua atividade modulada por proteínas inibitórias. A PMEI está envolvida em muitos processos fisiológicos requeridos na remodelação da parede celular. As invertases vacuolares e de parede celular possuem importantes papéis no metabolismo da sacarose em plantas superiores. A sacarose e seus compostos derivados hexose, glicose e frutose, liberados pela atividade das invertases, são importantes como metabólitos e também como sinais que influenciam a expressão de diversos genes. A regulação da atividade das invertases vacuolares e/ou de parede celular é crucial para diversos processos do desenvolvimento, incluindo a defesa contra patógenos (revisado em Scognamiglio *et al.*, 2003). A PMEI também possui um importante papel na maturação de frutos, senescência e crescimento do tubo polínico (revisado em Juge, 2007). Entretanto, o efeito bioquímico da PMEI no metabolismo da parede celular e seus papéis fisiológicos durante o desenvolvimento da plantas ainda não estão totalmente esclarecidos (Irifune *et al.*, 2004).

Pouco se sabe, também, a respeito da ocorrência da PMEI em plantas e a sua função *in vivo* (Juge, 2007).

4.4. Genes Constitutivos em Eucalyptus

Como citado anteriormente, a validação experimental dos dados de microarranjos é normalmente realizada pela técnica de qRT-PCR. A normalização dos dados obtidos com a análise por qRT-PCR requer genes que apresentem transcritos com expressão uniforme na maioria das células do organismo ou entre as espécies que estão sendo analisadas, assim como durante várias fases do desenvolvimento e sob diferentes condições ambientais. São os chamados genes constitutivos. Diversos trabalhos têm constatado que genes utilizados na era pré-genômica não apresentam expressão estável entre tecidos e genótipos (vide Seção 1.7, Introdução). Em um trabalho pioneiro realizado em *Arabidopsis*, Czechowski *et al.* (2005) descreveram 18 novos genes constitutivos para esta espécie. Para *Eucalyptus*, até o momento, não há qualquer estudo de investigação da estabilidade de genes-referência, nem a descrição de novos genes para a normalização em estudos de qRT-PCR.

Dessa forma, três diferentes etapas foram utilizadas na seleção dos potenciais genes constitutivos: (i) escolha de genes constitutivos clássicos; (ii) análise dos novos genes descritos para *Arabidopsis* em estudo realizado por Czechowski *et al.* (2005); e (iii) seleção de genes candidatos com base nas análises dos resultados das hibridizações dos microarranjos. Utilizando os parâmetros definidos na Seção 3.2.1.2 do Material e Métodos para as análises baseadas em SAM, 21 genes representando diferentes classes funcionais e famílias gênicas foram selecionados e seus perfis de expressão gênica em alguns tecidos e espécies de eucalipto foram examinados *in silico* e por qRT-PCR.

Três genes notoriamente constitutivos e normalmente usados como referências em experimentos de Northern *blots* e RT-PCRs convencionais foram escolhidos baseados em diversos trabalhos da literatura: histona H2B (Hays & Skinner, 2001; Devaux *et al.*, 2006; Sterky *et al.*, 2004), Ribonucleoproteína L23A (Mrusek *et al.*, 2005), GAPDH (van Hoof *et al.*, 2001; García-Maroto *et al.*, 2002; Dambrauskas *et al.*, 2003).

Além desses, foram selecionados, na base de ESTs do Genolyptus, quatro novos genes constitutivos, potenciais ortólogos àqueles descritos para *Arabidopsis* (Czechowski

et al., 2005) com vistas a confirmar o caráter invariável de suas expressões. As seqüências dos ortólogos para eucalipto foram identificadas por meio de um BLASTn e apenas aquelas seqüências que apresentaram altas homologias foram selecionadas. Os *primers* para qRT-PCR foram projetados nas regiões mais conservadas, conforme descrito na Seção 3.2.1.2 de Material e Métodos. Os genes selecionados incluem uma serina/treonina proteína fosfatase PP2A2; uma proteína da família SAND; uma proteína semelhante às da família TIP41; e uma proteína expressa de *A. thaliana* com função ainda desconhecida (Tabela 7).

Tabela 7. Relação de nomes dos novos genes consitutivos descritos para *Arabidopsis* (Czechowski *et al.*, 2005) e o seu código nessa espécie. Na Tabela estão representados, também, o código da EST homóloga no banco de dados do Genolyptus e as seqüências de *primers* indicadas pelo programa *GenScript Corporation's Real-Time PCR Primer Design* e o tamanho esperado dos amplicons para cada seqüência gênica considerada.

| Código <i>Arabidopsis</i> | EST homóloga (Código Genolyptus) | Identidade em <i>Arabidopsis</i> | Amplicon | Forward/Reverse |
|---------------------------|-------------------------------------|--|----------|--|
| At1g59830 | EUPE-XY-001-012-B06-RS.R | serina/treonina proteína fosfatase PP2A2 | 139 | 5' CCTCATGAAGGACCAATGTG/ 5' GAGTGAGACCAATTGGTGTGG |
| At2g28390 | EUGR-TS-002-009-F11-RS.R | proteína da família SAND | 143 | 5' CCATTCAACACTCTCCGACA/ 5' TGTGTGACCCAGCAGAGTAAT |
| At4g34270 | EUSP-FX-001-004-B11-BA.R | proteína semelhante às da família TIP41 | 148 | 5' TGTTGGTTTCTTCTCTTGCG/ 5' AAGCTTTGTACGTCGCCTCT |
| At4g33380 | EUSP-FX-001-045-A12-CN.R | proteína expressa de <i>A. thaliana</i> | 113 | 5' CAGGCTAAGTATGCCAGACG/ 5' CAGCACTCCTGGTTCTTGTG |

As PP2A2 incluem uma família de fosfatases protéicas em resíduos de serina/treonina, atuando como enzimas regulatórias que possuem um importante papel no controle da fosforilação de proteínas e regulação no desenvolvimento, morfologia e ciclo celular (revisado em Janssens & Goris, 2001). A família SAND inclui proteínas altamente conservadas em todas as espécies de eucariotos e estão envolvidas no processo de sinalização vacuolar em plantas (revisado em Cottage *et al.*, 2004). As proteínas da família TIP41 regulam negativamente a via de sinalização da proteína TOR (do inglês *target of rapamycin*), uma quinase que controla o crescimento celular em resposta a diferentes nutrientes (revisado em Jacinto *et al.*, 2001).

Utilizando-se os resultados das análises realizadas com o algoritmo SAM sobre dados dos microarranjos, foram selecionados 13 genes com potencial função constitutiva. Após a análise das ESTs dos agrupamentos correspondentes no banco de dados do Genolyptus, foram projetados *primers* nas regiões mais conservadas (Tabela 8). No Anexo

2, estão representados os gráficos que ilustram os alinhamentos realizados com a sequência do *Contig* utilizado na projeção dos oligonucleotídeos, sintetizados pela NimbleGen, para hibridização dos microarranjos. Para cada gene está representada, também, a sequência consenso e os locais de projeção dos *primers* utilizados no qRT-PCR. O perfil de expressão dos genes candidatos foi avaliado *in silico* pela análise dos gráficos gerados pelo programa SAM (Anexo 3). Os valores estão representados em Log₂ e indicam a diferença de expressão relativa entre os vinte valores citados anteriormente. A análise dos gráficos permitiu verificar o perfil de variação (ou não) das expressões dos genes entre as diferentes espécies, entre os diferentes tecidos e entre as diferentes replicatas biológicas e experimentais. A atribuição de uma potencial identidade por BLASTx aos genes selecionados tornou possível verificar que estes incluem diferentes classes funcionais e famílias gênicas, e que estes genes candidatos podem estar envolvidos em diferentes processos celulares. Entretanto, como os valores de identidades (*E values*, Tabela 8) não foram suficientemente baixos, isto é, foram superiores a 1×10^{-40} para a maioria dos genes, essa classificação deve ser considerada apenas especulativa.

Tabela 8. Relação de nomes dos supostos genes constitutivos para *Eucalyptus* selecionados com base nos resultados de microarranjos e o seu coeficiente de variação (CV). Os valores de identidade atribuídos por BLASTx estão representados pelo valor *E*, ou *E value*. Na Tabela estão representados, também, as seqüências de *primers* indicadas pelo programa *GenScript Corporation's Real-Time PCR Primer Design* e o tamanho esperado dos amplicons para cada seqüência gênica considerada.

| Nome | Código Genolyptus | Identidade (Blastx) | CV | Amplicon | Forward/Reverse |
|-------|--------------------|---|------|----------|---|
| EuC01 | 6348 CL2321Contig1 | suposto fator de <i>splicing</i> 7, rico em arginina/serina [<i>Oryza sativa</i> (japônica)] [e-value=9e-40] | 0,19 | 117 | 5' GGTGAGCCTGGACACTTTG/ 5' CTCTGCGGTATCTTGGACT |
| EuC02 | 5068 CL1688Contig1 | proteína da família RING <i>finger</i> [<i>Arabidopsis thaliana</i>] [e-value=3e-35] | 0,16 | 97 | 5' TGTTTCGACTGCAACATCTG/ 5' GGAGCCATTGTAGATGCAG |
| EuC03 | 1999 CL5129Contig1 | poliubiquitina [<i>Pinus sylvestris</i>] [e-value=8e-40] | 0,2 | 95 | 5' GACCAGCAGAGGCTCATCTT/ 5' ACAAGGTGAAGGGTGGACTC |
| EuC04 | 6216 CL1366Contig1 | proteína rica em leucinas [<i>Triticum aestivum</i>] [e-value=2e-81] | 0,2 | 97 | 5' AAGGCCTTTGCTAAGCTGAA/ 5' GGTCAGAGAGGGTGGTGAAT |
| EuC05 | 7694 CL3598Contig1 | BTB/POZ [<i>Medicago truncatula</i>] [e-value=3e-99] | 0,2 | 115 | 5' AACCTTGTAGCTGTCATGCG/ 5' CTCCTTCTCCCGACCTTG |
| EuC06 | 6220 CL1393Contig1 | proteína de ligação a aminoacil-tRNAs [<i>Arabidopsis thaliana</i>] [e-value=0.0] | 0,18 | 121 | 5' CATGGGATGAATTTGTGGAA/ 5' TTACATGCTCCCATCTCACC |
| EuC07 | 3874 CL1679Contig1 | terceira subunidade da RNA Polimerase II [<i>Arabidopsis thaliana</i>] [e-value=e-111] | 0,19 | 108 | 5' CATGTCCATGCGCTTCTC/ 5' GTCTGGTCGGTCATGCACT |
| EuC08 | 1814 CL3832Contig1 | helicase ligante aos módulos DEAD/DEAH [<i>Medicago truncatula</i>] [e-value=5e-64] | 0,18 | 108 | 5' CTTATGCTGCACAGGTGGTT/ 5' CTTCTTCTCCCGACCTTG |
| EuC09 | 5137 CL2173Contig1 | proteína expressa [<i>Arabidopsis thaliana</i>] [e-value=7e-47] | 0,17 | 98 | 5' GGTCCAACCCTTACATTTGG/ 5' TTAGCCTTGGCTTCTCTCC |
| EuC10 | 7845 CL4653Contig1 | proteína hipotética [<i>Arabidopsis thaliana</i>] [e-value=1e-30] | 0,2 | 110 | 5' AGGAGTCTTCGAGCTTCC/ 5' CAGCACGGACACCTGATAAA |
| EuC11 | 5015 CL1310Contig1 | fator de transcrição [<i>Arabidopsis thaliana</i>] [e-value=7e-56] | 0,2 | 129 | 5' TTCTGTGTCTGTGCTGCTG/ 5' GCTGTGCTTCTTCTTCTT |
| EuC12 | 2898 CL3104Contig1 | suposta proteína de ligação a RNAs [<i>Arabidopsis thaliana</i>] [e-value=2e-88] | 0,19 | 114 | 5' GCGTGGTTCTTGGATCACTA/ 5' TGGTGACAAAGTCAGGTGCT |
| EuC13 | 5407 CL4106Contig1 | serina carboxilase [<i>Arabidopsis thaliana</i>] [e-value=3e-120] | 0,18 | 100 | 5' TTTATTGATTACCCGTTGG/ 5' ATGCCAGTGAATCTCAACA |

A função molecular de cada um dos genes, discutida abaixo, foi atribuída de acordo com o código GO (*Gene Ontology*) por pesquisa realizada no *site* do TAIR (*The Arabidopsis Information Resource*) em “*GO Annotations*” (<http://www.arabidopsis.org/tools/bulk/go/index.jsp>). Foi observado que muitos dos genes possuem provável papel no controle de diferentes níveis da expressão como foi o caso das identidades de (i) fator de transcrição para EuC11; (ii) de RING *finger* para EuC02, um membro da família de proteínas que interagem seletivamente com outras proteínas por meio de íons de zinco; (iii) de helicases ligantes aos módulos DEAD/DEAH (*DEAD/DEAH box helicase*) para EuC08, isto é, uma potencial proteína com atividade de helicase dependente de ATP; (iv) de terceira subunidade da RNA Polimerase II para EuC07; (v) de fator de *splicing* para EuC01; (vi) de proteína de ligação a aminoacil-tRNAs para EuC06; e (vii) de proteína de ligação a RNAs para EuC12. Para os demais genes candidatos, foram encontradas duas proteínas de função desconhecida em *Arabidopsis* para

EuC09 (viii) e EuC10 (ix), além da (x) identidade de enzima para poliubiquitinas (EuC03), uma família envolvida na ubiquitinação de proteínas e outras moléculas durante processos diferentes processos catabólicos; (xi) uma proteína rica em leucinas para EuC04, envolvida em transmissão de sinal e ligação a complexos protéicos; (xii) uma identidade de serina carboxilase para EuC13, envolvida em proteólise; e (xiii) de BTB/POZ para EuC05, classe de proteínas que atuam tanto na regulação da transcrição como na interação com outras proteínas e, por isso, têm papel em diversos processos biológicos (revisado em Weber *et al.*, 2005) como, por exemplo, na morfogênese de folhas (Ha *et al.*, 2004), na resposta a patógenos em *Arabidopsis* (Rochon *et al.*, 2006) e na via de regulação da expressão de genes responsivos a hormônios em cevada (Woodger *et al.*, 2004).

4.5 qRT-PCR

Os *primers* projetados para todas as seqüências selecionadas anteriormente tiveram as suas especificidades de amplificação avaliadas diretamente por qRT-PCRs utilizando-se cDNAs de *Eucalyptus*. As curvas de dissociação geradas nas análises por qRT-PCRs foram utilizadas na confirmação das especificidades dos *primers* (Anexo 4) e, também, na determinação experimental das temperaturas médias de anelamento (T_m). As curvas de dissociação permitiram detectar produtos inespecíficos de PCR e dímeros de *primers*, servindo para a otimização da concentração dos *primers* nas reações. A especificidade da reação é observada pela presença de um único pico de amplificação. A ocorrência de mais de um pico, por outro lado, revela a presença de produtos inespecíficos, o que requer a otimização da reação.

Como resultados dessas análises de curvas de dissociação, os *primers* projetados para EuC01, EuC08 e EuC13 não permitiram a amplificação de quaisquer produtos com os cDNAs testados e foram excluídos das avaliações. As qRT-PCRs com os *primers* para At1g59830, EuC07 e EuC11 apresentaram mais de um produto de amplificação, indicando a inespecificidade de anelamento dos mesmos sob as condições testadas. Para estes, foram realizados novos experimentos com gradientes de temperatura, já que a temperatura de anelamento padrão utilizada nas qRT-PCRs foi de 60 °C. As temperaturas testadas foram as de 54 °C; 56,4 °C; 62,2 °C e 64,8 °C. Como resultado das otimizações, apenas os

primers para At1g59830 permitiram a amplificação de produtos específicos sob as temperaturas de 54 °C; 56,4 °C e 62,2°C (dados não constam). Este *primer* será incluído nas análises posteriores. Os demais *primers* EuC07 e EuC11 não apresentaram resultados positivos e foram excluídos das análises posteriores por qRT-PCR. Para todos os demais conjuntos de *primers*, um único pico de amplificação foi observado nas análises das curvas de dissociação, comprovando a alta especificidade dos mesmos para os genes-alvo sob as condições utilizadas nas qRT-PCRs (Anexo 4).

O nível de expressão dos genes selecionados foi avaliado em dez amostras biológicas derivadas de árvores de quatro anos de idade plantadas a campo em áreas do Hortoflorestal Barbanegra da empresa Aracruz Celulose S.A., em Barra do Ribeiro, RS. As árvores amostradas tratavam-se dos mesmos indivíduos utilizados nas análises de microarranjos de DNA conforme descrito na Seção 3.1.1 de Material e Métodos. Os cDNAs produzidos a partir de amostras de RNA total da amostra GR-XY-A2 não apresentaram boa reprodutibilidade e este e sua replicata biológica foram excluídos das análises. Como resultados das análises de qRT-PCR, foi possível observar que para xilemas de *E. globulus* e xilemas e folhas maduras de *E. grandis*, os genes His, EuC06, EuC09, EuC10, EuC12, At2g28390 e GAPDH foram os que demonstraram menor variação em suas expressões, caracterizando-se como constitutivos para estes tecidos. Levando-se em conta os baixos valores de desvio padrão, é possível afirmar que EuC10, EuC12 e At2g28390 foram os melhores genes candidatos para servirem como referências em análises da expressão gênica em *Eucalyptus* (Figura 6), pelo menos para os tecidos e espécies considerados. O gene EuC06 tem identidade potencial de uma proteína de ligação a aminoacil-tRNAs, EuC12, como a de uma proteína de ligação a RNAs e EuC09 e EuC10 codificam proteínas desconhecidas, isto é, ainda não descritas na literatura científica e tampouco encontradas nos bancos de dados públicos.

De acordo com os resultados encontrados nas qRT-PCRs realizadas com cDNAs derivados de folhas e xilemas de *E. grandis*, somente EuC12 e At2g28390 demonstraram invariabilidade de expressões entre os tecidos (Figura 7). Dois tradicionais genes constitutivos, His e GAPDH, apresentaram uma variação considerável nas suas expressões entre os tecidos. Este resultado contraria o obtido por Iskandar *et al.* (2004) em cana-de-açúcar, que encontraram GAPDH como um dos melhores genes constitutivos entre diversos tecidos e genótipos. Jain *et al.* (2006), entretanto, também contestam a

constitutividade de GAPDH na análise de genes constitutivos em arroz com diferentes tecidos e órgãos e em diferentes estádios do desenvolvimento. O mesmo resultado foi obtido por Czechowski *et al.* (2005) em *Arabidopsis*. Uma grande variação foi observada para At4g33380, um dos genes indicados por Czechowski *et al.* (2005) como constitutivo em *Arabidopsis*. Isso pode ser devido ao fato de que o gene ortólogo em eucalipto não seja correspondente ao exato parálogo de *Arabidopsis*. Um alinhamento realizado entre as seqüências derivadas de ambas as espécies revelou uma identidade de 80%. Embora alto, este grau de identidade não garante tratar-se do mesmo ortólogo e, de qualquer forma, não foi encontrado no banco de dados do Genolyptus a seqüência de cDNA completa para este gene para análises mais aprofundadas de homologia.

Dos 13 genes que demonstraram padrão de expressão estável nos microarranjos e que foram analisados por qRT-PCR, seis mostraram padrão constitutivo. A qRT-PCR é uma técnica mais refinada no que diz respeito à análise de transcritos do que os microarranjos. De acordo com Dallas *et al.* (2005) a correlação entre os dados de qRT-PCR e de microarranjos é, na maioria das vezes, robusta. É fundamental que as sondas do qRT-PCR reconheçam os mesmos transcritos das sondas dos microarranjos. Dessa forma, é necessário realizar, ainda, uma análise para verificar se os locais de anelamento dos *primers* de qRT-PCR coincidem com os locais onde as nove sondas dos microarranjos foram projetadas. Esta análise poderá justificar as diferenças que foram encontradas entre ambos os experimentos, pois se a localização dos *primers* diferir daquelas das sondas, há a

GENES CONSTITUTIVOS - *E. grandis* x *E. globulus*

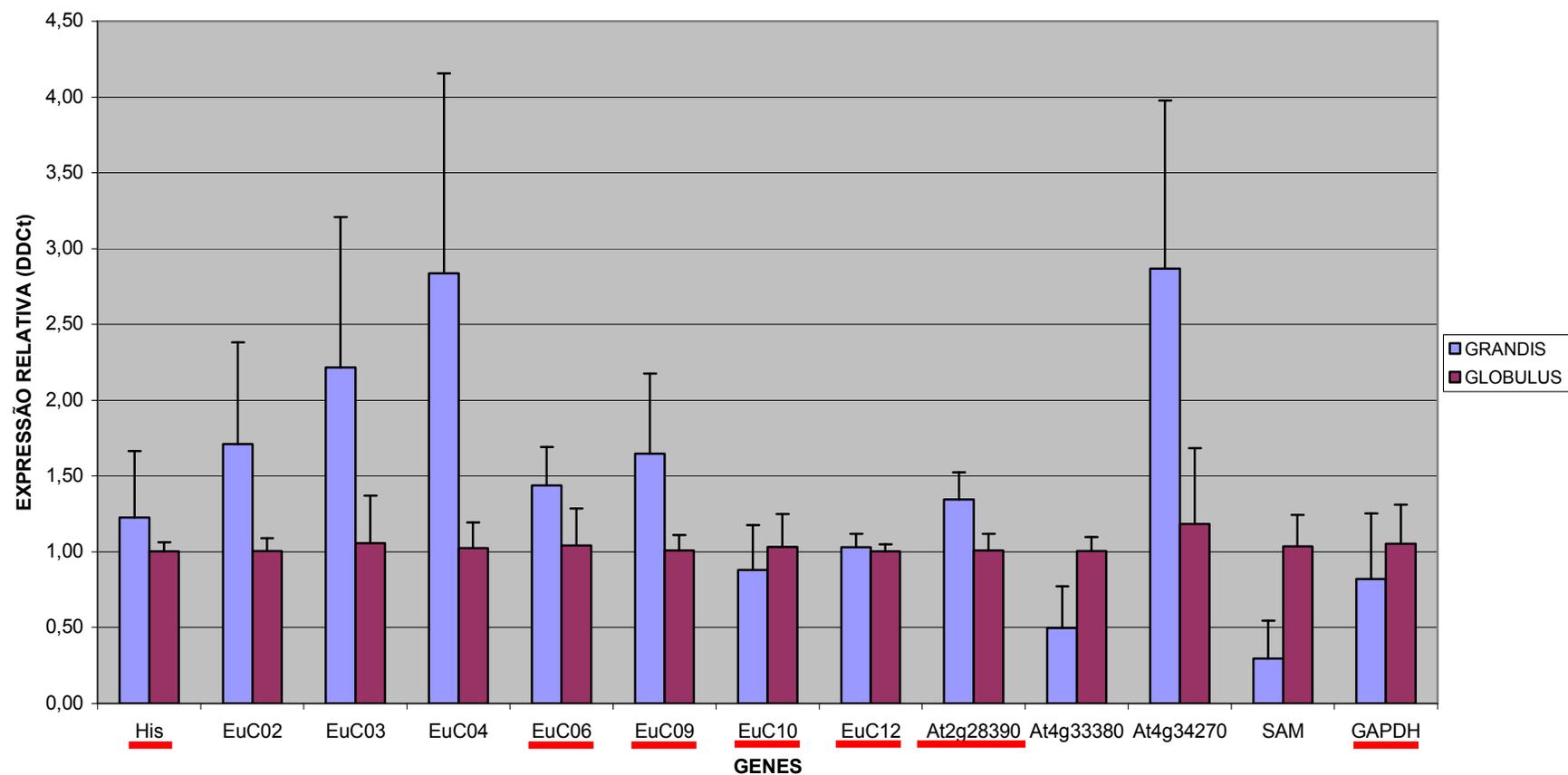


Figura 6. Valores médios de $2^{-\Delta\Delta C_T}$ e o erro padrão para cada gene. Como normalizador para os cálculos foi utilizada a média dos valores de $\Delta\Delta C_T$ dos clones de *E. globulus* para o gene da Ribonucleoproteína L23A. Destacados em vermelho estão os *primers* dos genes como expressão não significativamente variável entre os xilemas de *E. globulus* e folhas e xilemas de *E. grandis*.

possibilidade de ter ocorrido *splicing* diferencial. Não se exclui, também, a possibilidade da ocorrência de falhas na projeção dos *primers* devido à existência de múltiplos alelos ou membros de uma mesma família gênica. Portanto a possibilidade de utilizar outros pares de *primers* para a análise dos mesmos genes é recomendável. Com os resultados obtidos até o momento, não é possível concluir-se a validação dos microarranjos pela qRT-PCR, uma vez que apenas seis genes entre 13 selecionados mostraram comportamento constitutivo entre as duas técnicas.

GENES CONSTITUTIVOS - FOLHAS MADURAS E XILEMA *E. grandis*

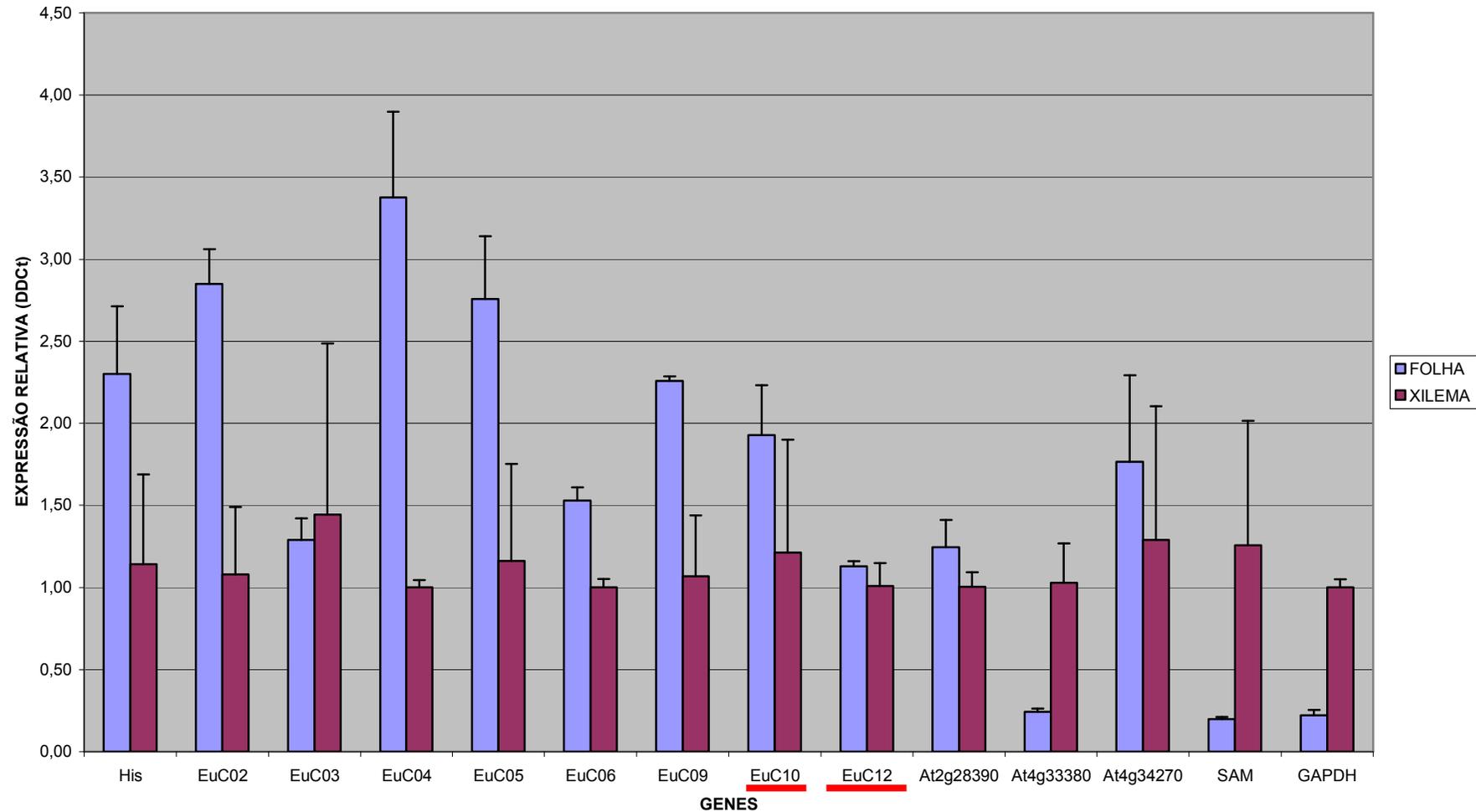


Figura 7. Valores médios de $2^{-\Delta\Delta C_T}$ e o erro padrão para cada gene. Como normalizador para os cálculos foi utilizada a média dos valores de $\Delta\Delta C_T$ dos clones de xilema de *E. grandis* para o gene da Ribonucleoproteína L23A. Destacados em vermelho estão os *primers* dos genes como expressão não significativamente variável entre folhas e xilemas de *E. grandis*.

4.6. Conclusões e perspectivas

Como conclusão da seleção de genes-referência para estudos da expressão gênica em *Eucalyptus*, os genes candidatos EuC12 e At2g28390 são os genes constitutivos mais adequados como controles internos em estudos de expressão gênica para folhas e xilemas de *E. grandis* e *E. globulus*. É importante ressaltar que ainda é necessário investigar a estabilidade de expressão destes genes em outras espécies de *Eucalyptus*, entre diferentes estádios do desenvolvimento, além de outros tecidos e órgãos submetidos, ou não, a estresses bióticos e abióticos. Os genes selecionados, segundo os experimentos realizados, servirão como controles internos em todas as qRT-PCRs subsequentes nas avaliações de dados gerados pelos microarranjos de DNA para todos os demais genes sob análise no Projeto Genolyptus, e demais linhas de investigação de *Eucalyptus*.

5. Referências bibliográficas

Allona I, Quinn M, Shoop E, Swope K, St.Cyr S, Carlis J, Riedl J, Retzel E, Campbell MM, Sederoff R, Whetten RW (1998) Analysis of xylem formation in pine by cDNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America* 95: 9693-9698.

Andersson-Gunneras S, Mellerowicz EJ, Love J, Segerman B, Ohmiya Y, Coutinho PM, Nilsson P, Henrissat B, Moritz T, Sundberg B (2006) Biosynthesis of cellulose-enriched tension wood in *Populus*: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. *The Plant Journal* 45: 144–165.

Aracruz, 2007. Retirado do *site* da empresa na *internet* em julho de 2007 (<http://www.aracruz.com.br/web/pt/ambiente>).

Aroca R, Ferrante A, Vernieri P, Chrispeels MJ (2006) Drought, abscisic acid and transpiration rate effects on the regulation of PIP aquaporin gene expression and abundance in *Phaseolus vulgaris* plants. *Annals of Botany* 98: 1301–1310.

Aspeborg H, Schrader J, Coutinho PM, Stam M, Kallas A, Djerbi S, Nilsson P, Denman S, Amini B, Sterky F, Master E, Sandberg G, Mellerowicz E, Sundberg B, Henrissat B, Teeri TT (2005) Carbohydrate-active enzymes involved in the secondary cell wall biogenesis in hybrid aspen. *Plant Physiology* 137: 983–997.

Atkinson RG, Schröder R, Hallett IC, Cohen D, MacRae EA (2002) Overexpression of polygalacturonase in transgenic apple trees leads to a range of novel phenotypes involving changes in cell adhesion. *Plant Physiology* 129: 122–133.

Ausubel FM, Brent R, Kingston RE, Moore DD, Smith JA, Seidman JG, Stuhl K (1997) *Current Protocols in Molecular Biology*. New York: John Wiley and Sons.

Balestrieri C, Castaldo D, Giovane A, Quagliuolo L, Servillo L (1990) A glycoprotein inhibitor of pectin methylesterase in kiwi fruit (*Actinidia chinensis*). European Journal of Biochemistry 193: 183-187.

Balk PA & de Boer AD (1999) Rapid stalk elongation in tulip (*Tulipa gesneriana* L. cv. Apeldoorn) and the combined action of cold-induced invertase and the water-channel protein gamma TIP. Planta 209: 346–354.

Beritognolo I, Magel E, Abdel-Latif A, Charpentier J, Jay-Allemand C, Breton C (2002) Expression of genes encoding chalcone synthase, flavanone 3-hydroxylase, and dihydroflavonol 4-reductase correlates with flavanol accumulation during heartwood formation in *Juglans nigra* L. Tree Physiology 22: 291–300.

Besseau S, Hoffmann L, Geoffroy P, Lapierre C, Pollet B, Legrand M (2007) Flavonoid accumulation in *Arabidopsis* repressed in lignin synthesis affects auxin transport and plant growth. The Plant Cell 19: 148–162.

Bortolotti C, Cordeiro A, Alcázar R, Borrell A, Culiañez-Macià FA, Tiburcio AF, Altabella T (2004) Localization of arginine decarboxylase in tobacco plants. Physiologia Plantarum 120: 84–92.

Brunner AM, Busov VB, Strauss SH (2004) Poplar genome sequence: functional genomics in an ecologically dominant plant species. Trends in Plant Science 9 (1): 49-56.

Bustin SA, Benes V, Nolan T, Pfaff MW (2005) Quantitative real-time RT-PCR – a perspective. Journal of Molecular Endocrinology 34: 597–601.

BRACELPA (2006) Associação Brasileira de Celulose e Papel: Relatório anual de desempenho (<http://www.bracelpa.org.br/bra/estatisticas/>)

Brito JO, Barrichelo LEG, Seixas F (1993) Análise da Produção Energética e de Carvão Vegetal de Espécies de Eucalipto. Instituto de Pesquisas e Estudos Florestais 23: 53-56.

Chaumont F, Barrieu F, Herman EM, Chrispeels MJ (1998) Characterization of a maize tonoplast aquaporin expressed in zones of cell division and elongation. *Plant Physiology* 117: 1143–1152.

Clarke JD & Zhu T (2006) Microarray analysis of the transcriptome as a stepping stone towards understanding biological systems: practical considerations and perspectives. *The Plant Journal* 45: 630–650.

Colodette JL, Mokfienski A, Gomide A, Oliveira R (2004) Relative importance of wood density and carbohydrate content on pulping yield and product quality. *Journal of Tianjin University of Science and Technology* 19 (4): 71-80.

Cottage A, Mullan L, Portela MDB, Hellen E, Carver T, Patel S, Vavouri T, Elgar G, Edwards YJK (2004) Molecular Characterisation of the sand protein Family: a study based on comparative genomics, structural bioinformatics and phylogeny. *Cellular & Molecular Biology Letters* 9: 739 – 753.

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible W (2005) Genome-Wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiology* 139: 5–17.

Dallas PB, Gottardo NG, Firth MJ, Beesley AH, Hoffmann K, Terry PA, Freitas JR, Boag JM, Cummings AJ, Kees UR (2005) Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR – how well do they correlate? *BMC Genomics* 6: 59-69.

Demura T, Tashiro G, Horiguchi G, Kishimoto N, Kubo M, Matsuoka N, Minami A, Nagata-Hiwatashi M, Nakamura K, Okamura Y, Sassa N, Suzuki S, Yazaki J, Kikuchi S,

Fukuda H (2002) Visualization by comprehensive microarray analysis of gene expression programs during transdifferentiation of mesophyll cells into xylem cells. *Proceedings of the National Academy of Sciences* 99 (24): 15794–15799.

Dambrauskas G, Aves JA, Bryant JA, Francis D, Rogers HJ (2003) Genes encoding two essential DNA replication activation proteins, Cdc6 and Mcm3, exhibit very different patterns of expression in the tobacco BY-2 cell cycle. *Journal of Experimental Botany*: 54 (383): 699-706.

Demura T & Fukuda H (2007) Transcriptional regulation in wood formation. *Trends in Plant Science* 12 (2): 64-70.

Devaux S, Lecordier L, Uzureau P, Walgraffe D, Dierick J, Poelvoorde P, Pays E, Vanhamme L (2006) Characterization of RNA polymerase II subunits of *Trypanosoma brucei*. *Molecular & Biochemical Parasitology* 148: 60–68.

Doblin MS, Kurek I, Jacob-Wilk D, Delmer DP (2002) Cellulose biosynthesis in plants: from genes to rosettes. *Plant Cell Physiology* 43 (12): 1407–1420.

Du S & Yamamoto F (2007) An overview of the biology of reaction wood formation. *Journal of Integrative Plant Biology* 49 (2): 131–143.

Eberhard J, Raesecke H, Schmid J, Amrhein N (1993) Cloning and expression in yeast of a higher plant chorismate mutase: Molecular cloning, sequencing of the cDNA and characterization of the *Arabidopsis thaliana* enzyme expressed in yeast. *FEBS* 334 (2): 233-236.

FAO (2000) Global forest resources assessment 2000 – Main report. FAO Forestry paper. 479p (<http://www.fao.org/forestry/fo/fra/main/index.jsp>).

FAO (2006) Food and Agriculture Organization of the United Nations: “Forest Products and Services” (<http://www.fao.org/forestry/index.jsp>).

Foucart C, Paux E, Ladouce N, San-Clemente H, Grima-Pettenati, J, Sivadon P (2006) Transcript profiling of a xylem vs phloem cDNA subtractive library identifies new genes expressed during xylogenesis in *Eucalyptus*. *New Phytologist* 170: 739–752.

Franke R, Humphreys JM, Hemm MR, Denault JW, Ruegger MO, Cusumano JC, Chapple C (2002) The *Arabidopsis* REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. *The Plant Journal* 30 (1): 33-45.

Franklin-Tong VE, Grant H, Hepler PK (1997) Ratio-imaging of Ca²⁺ in the self-incompatibility response in pollen tubes of *Papaver rhoeas*. *The Plant Journal* 12 (6): 1375-1386.

Gachon C, Mingam A, Charrier B (2004) Real-time PCR: what relevance to plant studies? *Journal of Experimental Botany* 55 (402): 1445–1454.

García-Maroto F, Garrido-Cárdenas JA, Rodríguez-Ruiz J, Vilches-Ferrón M, Adam AC, Polainac J, Alonso DL (2002) Cloning and molecular characterization of the $\Delta 6$ -desaturase from two *Echium* plant species: production of GLA by heterologous expression in yeast and tobacco. *Lipids* 37 (4): 417-426.

Giovane A, Servillo L, Balestrieri C, Raiola A, D'Avino R, Tamburrini M, Ciardiello MA, Camardella L (2004) Pectin methylesterase inhibitor. *Biochimica et Biophysica Acta* 1696: 245– 252.

Groover A & Robischon M (2006) Developmental mechanisms regulating secondary growth in woody plants. *Current Opinion in Plant Biology* 9: 55–58.

Ha CH, Jun JH, Nam HG, Fletcher JC (2004) BLADE-ON-PETIOLE1 encodes a BTB/POZ domain protein required for leaf morphogenesis in *Arabidopsis thaliana*. *Plant Cell Physiology* 45 (10): 1361–1370.

Hachez C, Moshelion M, Zelazny E, Cavez D, Chaumont F (2006) Localization and quantification of plasma membrane aquaporin expression in maize primary root: a clue to understanding their role as cellular plumbers. *Plant Molecular Biology* 62: 305–323.

Hadfield KA & Bennett AB (1998) Polygalacturonases: many genes in search of a function. *Plant Physiology* 117: 337–343

Hays DB & Skinner DZ (2001) Development of an expressed sequence tag (EST) library for *Medicago sativa*. *Plant Science* 161: 517–526.

Harper JF, Breton G, Harmon A (2004) Decoding Ca²⁺ signals through plant protein kinases. *Annual Review of Plant Biology* 55: 263–88.

Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalerao R, Uhlén M, Teeri TT, Lundeberg J, Sundberg B, Nilsson P, Sandberg G (2001) A transcriptional roadmap to wood formation. *Proceedings of the National Academy of Sciences* 98 (25): 14732–14737.

Hilaire E, Young SA, Willard LH, McGee JD, Sweat T, Chittoor JM, Guikema JA, Leach JE (2001) Vascular defense responses in rice: peroxidase accumulation in xylem parenchyma cells and xylem wall thickening. *The American Phytopathological Society* 14 (12): 1411–1419.

Hummel I, Bourdais G, Gouesbet G, Couée I, Malmberg RL, Amrani AE (2004) Differential gene expression of Arginine decarboxylase ADC1 and ADC2 in *Arabidopsis thaliana*: characterization of transcriptional regulation during seed germination and seedling development. *New Phytologist* 163: 519–531.

Hrabak EM, Chan CWM, Gribskov M, Harper JF, Choi JH, Halford N, Kudla J, Luan S, Nimmo HG, Sussman MR, Thomas M, Walker-Simmons K, Zhu J, Harmon AC (2003) The *Arabidopsis* CDPK-SnRK superfamily of protein kinases. *Plant Physiology*: 132: 666–680.

Irifune K, Nishida T, Egawa H, Nagatani A (2004) Pectin methylesterase inhibitor cDNA from kiwi fruit. *Plant Cell Reports* 23: 333–338.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4: 249–264.

Iskandar HM, Simpson RS, Casu RE, Bonnett GD, Maclean DJ, Manners JM (2004) comparison of reference genes for quantitative real-time polymerase chain reaction analysis of gene expression in sugarcane. *Plant Molecular Biology Reporter* 22: 325–337.

Israelsson M, Eriksson ME, Hertzberg M, Aspeborg H, Nilsson P, Moritz T (2003) Changes in gene expression in the wood-forming tissue of transgenic hybrid aspen with increased secondary growth. *Plant Molecular Biology* 52: 893–903.

Jacinto E, Guo B, Arndt KT, Schmelzle T, Hall MN (2001) TIP41 Interacts with TAP42 and negatively regulates the TOR signaling pathway. *Molecular Cell* 8: 1017–1026.

Janssens V & Goris J (2001) Protein phosphatase 2A: a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochemical Journal* 353: 417-439.

Jain M, Nijhawan A, Tyagi AK, Khurana JP (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochemical and Biophysical Research Communications* 345: 646–651.

Juge N (2006) Plant protein inhibitors of cell wall degrading enzymes. *Trends in Plant Science* 11 (7): 359-367.

Kaltenbach M, Schröder J, Schmelzer E, Lutz V, Schröder J (1999) Flavonoid hydroxylase from *Catharanthus roseus*: cDNA, heterologous expression, enzyme properties and cell-type specific expression in plants. *The Plant Journal* 19 (2): 183-193.

Kawaura K, Mochida K, Yamazaki Y, Ogihara Y (2006) Transcriptome analysis of salinity stress responses in common wheat using a 22k oligo-DNA microarray. *Functional and Integrative Genomics* 6: 132–142.

Ko J & Han K (2004) Arabidopsis whole-transcriptome profiling defines the features of coordinated regulations that occur during secondary growth. *Plant Molecular Biology* 55: 433–453.

Krishnan HB, Reeves CD, Okita TW (1986) ADPglucose pyrophosphorylase is encoded by different mRNA transcripts in leaf and endosperm of cereals. *Plant Physiology* 81: 642-645.

Kubo M, Udagawa M, Nishikubo N, Horiguchi G, Yamaguchi M, Ito J, Mimura T, Fukuda H, Demura T (2005) Transcription switches for protoxylem and metaxylem vessel formation. *Genes & Development* 19: 1855–1860.

Kurek I, Kawagoe Y, Jacob-Wilk D, Doblin M, Delmer D (2002) Dimerization of cotton fiber cellulose synthase catalytic subunits occurs via oxidation of the zinc-binding domains. *Proceedings of the National Academy of Sciences* 99 (17): 11109–11114.

Lecourieux D, Ranjeva R, Pugin A (2006) Calcium in plant defence-signalling pathways. *New Phytologist* 171 (2): 249–269.

Lennartz K, Plücken H, Seidler A, Westhoff P, Bechtold N, Meierhoff K (2001) HCF164 encodes a thioredoxin-like protein involved in the biogenesis of the cytochrome b6f complex in *Arabidopsis*. *The Plant Cell* 13: 2539–2551.

Li L, Lu S, Chiang V (2006) A Genomic and molecular view of wood formation. *Critical Reviews in Plant Sciences* 25 (3): 215-233.

Li RW & Waldbieser GC (2006) Production and utilization of a high-density oligonucleotide microarray in channel catfish, *Ictalurus punctatus*. *BMC Genomics* 7: 134-140.

Livak KJ & Schmittgen TD (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* 25: 402–408.

Lorenz WW & Dean JFD (2002) SAGE Profiling and demonstration of differential gene expression along the axial developmental gradient of lignifying xylem in loblolly pine (*Pinus taeda*). *Tree Physiology* 22: 301–310.

Luan S, Kudla J, Rodriguez-Concepcion M, Yalovsky S, Gruissem W (2002) Calmodulins and calcineurin B-like proteins: Calcium sensors for specific signal response coupling in plants. *The Plant Cell*: S389–S400.

Ludevid D, Höfte H, Himmelblau E, Chrispeels MJ (1992) The expression pattern of the tonoplast intrinsic protein g-TIP in *Arabidopsis thaliana* is correlated with cell enlargement. *Plant Physiology* 100: 1633–1639.

Mahajan S, Sopory SK, Tuteja N (2006) Cloning and characterization of CBL-CIPK signalling components from a legume (*Pisum sativum*) *FEBS Journal* 273 (5): 907–925.

Ma L, Chen C, Liu X, Jiao Y, Su N, Li L, Wang X, Cao M, Sun N, Zhang X, Bao J, Li J, Pedersen S, Bolund L, Zhao H, Yuan L, Wong GK, Wang J, Deng XW, Wang J (2005) A microarray analysis of the rice transcriptome and its comparison to *Arabidopsis*. *Genome Research* 15: 1274-1283.

Meyer Y, Reichheld JP, Vignols F (2005) Thioredoxins in *Arabidopsis* and other plants. *Photosynthesis Research* 86: 419–433.

MeV Manual (2006) MultiExperiment Viewer, Version 4.0, 276p (<http://mev.tm4.org>).

Mijnsbrugge CV, Meyermans H, Montagu MV, Bauw G, Boerjan W (2000) Wood formation in poplar: identification, characterization, and seasonal variation of xylem proteins. *Planta* 210: 589-598.

Micheli F (2001) Pectin methylesterases: Cell wall enzymes with important roles in plant physiology. *Trends in Plant Science* 6: 414–419.

Milioni D, Sado P, Stacey NJ, Domingo C, Roberts K, McCann MC (2001) Differential expression of cell-wall-related genes during the formation of tracheary elements in the *Zinnia* mesophyll cell system. *Plant Molecular Biology* 47: 221–238.

Mobley EM, Kunkel BN, Keith B (1999) Identification, characterization and comparative analysis of a novel chorismate mutase gene in *Arabidopsis thaliana*. *Gene* 240: 115–123.

Moreau C, Aksenov N, Lorenzo MG, Segerman B, Funk C, Nilsson P, Jansson S, Tuominen H (2005) A genomic approach to investigate developmental cell death in woody tissues of *Populus* trees. *Genome Biology* 6: R34.

Mrusek S, Classen-Linke I, Vloet A, Beier HM, Krusche CA (2005) Estradiol and medroxyprogesterone acetate regulated genes in T47D breast cancer cells. *Molecular and Cellular Endocrinology* 235: 39–50.

Nakashima J, Endo S, Fukuda H (2004) Immunocytochemical localization of polygalacturonase during tracheary element differentiation in *Zinnia elegans*. *Planta* 218: 729–739.

Nicot N, Hausman JF, Hoffmann L, Evers D (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of Experimental Botany* 56 (421): 2907-14.

O'Brien M, Bertrand C, Matton DP (2002) Characterization of a fertilization-induced and developmentally regulated plasma-membrane aquaporin expressed in reproductive tissues, in the wild potato *Solanum chacoense*. *Planta* 215: 485–493.

Oh S, Park S, Han K (2003) Transcriptional regulation of secondary growth in *Arabidopsis thaliana*. *Journal of Experimental Botany* 54 (393): 2709-2722.

Patterson BD & Smillie RM (1971) Developmental changes in ribosomal ribonucleic acid and fraction I protein in wheat leaves. *Plant Physiology* 47: 196-198.

Paux E, Tamasloukht M, Ladouce N, Sivadon P, Grima-Pettenati J (2004) Identification of genes preferentially expressed during wood formation in *Eucalyptus*. *Plant Molecular Biology* 55: 263–280.

Paux E, Carocha V, Marques C, Sousa AM, Borralho N, Sivadon P, Grima-Pettenati J (2005) Transcript profiling of *Eucalyptus* xylem genes during tension wood formation. *New Phytologist* 167: 89-100.

Pereira JCD, Sturion JA, Higa AR, Higa RCV, Shimizu JY (2000) Características da madeira de algumas espécies de eucalipto plantadas no Brasil. *Embrapa Florestas Documentos* 38: 113p

Pérez-Amador MA & Carbonell J (2005) Arginine decarboxylase and putrescine oxidase in ovaries of *Pisum sativum*. *Plant Physiology* 107: 865-872.

Persson S, Wei H, Milne J, Page GP, Somerville CR (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proceedings of the National Academy of Sciences* 102 (24): 8633–8638.

Pischke MS, Huttlin EL, Hegeman AD, Sussman MR (2006) A transcriptome-based characterization of habituation in plant tissue culture. *Plant Physiology* 140: 1255–1278.

Plomion C, Leprovost G, Stokes A (2001) Wood formation in trees. *Plant Physiology* 127: 1513–1523.

Poke FS, Vaillancourt RE, Potts BM, Reid JB (2005) Genomic research in *Eucalyptus*. *Genetica* 125: 79–101.

Potikha TS, Collins CC, Johnson DI, Delmer DP, Levine A (1999) The involvement of hydrogen peroxide in the differentiation of secondary walls in cotton fibers. *Plant Physiology* 119: 849–858.

Prassinis C, Ko J, Yang J, Han K (2005) Transcriptome profiling of vertical stem segments provides insights into the genetic regulation of secondary growth in hybrid aspen trees. *Plant Cell Physiology* 46 (8): 1213–1225.

Rhee SY, Dickerson J, Xu D (2006) Bioinformatics and its applications in plant biology. *Annual Review of Plant Biology*. 57: 335–360.

Rochon A, Boyle P, Wignes T, Fobert PR, Després C (2006) The coactivator function of *Arabidopsis* NPR1 requires the core of its BTB/POZ domain and the oxidation of C-terminal cysteines. *The Plant Cell* 18: 3670–3685.

Ruelland E & Miginiac-Maslow M (1999) Regulation of chloroplast enzyme activities by thioredoxins: activation or relief from inhibition? *Trends in Plant Science* 4 (4): 136-141.

Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J (2003) TM4: a

free, open-source system for microarray data management and analysis. *Biotechniques* 34(2): 374-378.

Saslowsky D & Winkel-Shirley B (2001) Localization of flavonoid enzymes in *Arabidopsis* roots. *The Plant Journal* 27(1): 37-48.

Schrader J, Nilsson J, Mellerowicz E, Berglund A, Nilsson P, Hertzberg M, Sandberga G (2004) A high-resolution transcript profile across the wood-forming meristem of poplar identifies potential regulators of cambial stem cell identity. *The Plant Cell* 16: 2278–2292.

Scognamiglio MA, Ciardiello MA, Tamburrini M, Carratore V, Rausch T, Camardella L (2003) The plant invertase inhibitor shares structural properties and disulfide bridges arrangement with the pectin methylesterase inhibitor. *Journal of Protein Chemistry* 22 (3): 363-369.

Siefritz F, Otto B, Bienert GP, van der Krol A, Kaldenhoff R (2004) The plasma membrane aquaporin NtAQP1 is a key component of the leaf unfolding mechanism in tobacco. *Plant Journal* 37: 147–155.

Schäffner AR (1998) Aquaporin function, structure, and expression: are there more surprises to surface in water relations? *Planta* 204: 131-139.

Smart LB, Nall NM, Bennett AB (1999) Isolation of RNA and protein from guard cells of *Nicotiana glauca*. *Plant Molecular Biology Reporter* 17: 371–383.

Sociedade Brasileira de Silvicultura (2006) Fatos e Números do Brasil Florestal. 108p. (<http://www.sbs.org.br/FatoseNumerosdoBrasilFlorestal.pdf>).

Sterky F, Regan S, Karlsson J, Hertzberg M, Rohde A, Holmberg A, Amini B, Bhalerao R, Larsson M, Villarreal R, Montagu MV, Sandberg G, Teeri, Olson O, Terry TT, Boerjan W, Gustafsson P, Uhlén M, Sundberg B, Lundeberg J (1998) Gene discovery

in the wood-forming tissues of poplar: Analysis of 5,692 expressed sequence tags. *Proceedings of the National Academy of Sciences of the United States of America* 95: 13330–13335.

Tattersall EAR, Ergul A, AlKayal F, DeLuc L, Cushman JC, Cramer JR (2005) Comparison of methods for isolating high-quality RNA from leaves of grapevine. *American Journal of Enology and Viticulture* (.56: 400-406.

Torki M, Mandaron P, Thomas F, Quigley F, Mache R, Falconet D (1999) Differential expression of a polygalacturonase gene family in *Arabidopsis thaliana*. *Molecular and General Genetics* 261: 948–952.

Trewavas A (1999) How plants learn. *Proceedings of the National Academy of Sciences of the United States of America* 96: 4216–4218.

Turnbull JW (1999) Eucalypt plantations. *New Forests* 17: 37–52.

Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences* 98 (9): 5116–5121.

Vadesompele J, Preter KD, Pattyn F, Poppe B, Roy NV, Paepe AD, Speleman F (2003) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* 3 (7): RESEARCH0034.1–0034.11

van Hoof NALM, Hassinen VH, Hakvoort HWJ, Ballintijn KF, Schat H, Verkleij JAC, Ernst WHO, Karenlampi SO, Tervahauta AI (2001) Enhanced copper tolerance in *Silene vulgaris* (Moench) garcke populations from copper mines is associated with increased transcript levels of a 2b-type metallothionein gene. *Plant Physiology* 126: 1519–1526.

Veracel, 2007. Retirado do *site* da empresa na *internet* em julho de 2007 (<http://www.veracel.com.br/pt/index.html>).

de Vetten N, ter Horst J, van Schaik HP, de Boer A, Mol J, Koes, R (1999) A cytochrome b5 is required for full activity of flavonoid 3' 5-hydroxylase, a cytochrome P450 involved in the formation of blue flowers. *Proceedings of the National Academy of Sciences of the United States of America* 96: 778–783.

Vuosku J, Jokela A, Läärä E, Sääskilahti M, Muilu R, Sutela S, Altabella T, Sarjala T, Häggman H (2006) Consistency of polyamine profiles and expression of arginine decarboxylase in mitosis during zygotic embryogenesis of *Scots Pine*. *Plant Physiology* 142: 1027–1038.

Wang T, Zhang N, Du L (2005) Isolation of RNA of high quality and yield from *Ginkgo biloba* leaves. *Biotechnology Letters* 27: 629–633.

Weber H, Bernhardt A, Dieterle M, Hano P, Mutlu A, Estelle M, Genschik P, Hellmann H (2005) *Arabidopsis* AtCUL3a and AtCUL3b form complexes with members of the BTB/POZ-MATH protein family. *Plant Physiology* 137: 83–93.

Weber W, Trampczynska A, Clemens S (2006) Comparative transcriptome analysis of toxic metal responses in *Arabidopsis thaliana* and the Cd²⁺-hypertolerant facultative metallophyte *Arabidopsis halleri*. *Plant, Cell and Environment* 29: 950–963.

Werck-Reichhart D, Bak S, Paquette S (2002) *Cytochromes P450. The Arabidopsis Book*, American Society of Plant Biologists. 28p.

Whetten R, Sun Y, Zhang Y, Sederoff R (2001) Functional genomics and cell wall biosynthesis in loblolly pine. *Plant Molecular Biology* 47: 275–291.

Winkel-Shirley B (1999) Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways. *Physiologia Plantarum* 107: 142–149.

Wong JH, Kim Y, Ren P, Cai N, Cho M, Hedden P, Lemaux PG, Buchanan BB (2002) Transgenic barley grain overexpressing thioredoxin shows evidence that the starchy endosperm communicates with the embryo and the aleurone. *Proceedings of the National Academy of Sciences* 99 (25): 16325–16330.

Woodger FJ, Jacobsen JV, Gubler F (2004) GMPOZ, a BTB/POZ domain **nuclear** protein, is a regulator of hormone responsive gene expression in barley aleurone. *Plant Cell Physiology* 45(7): 945–950.

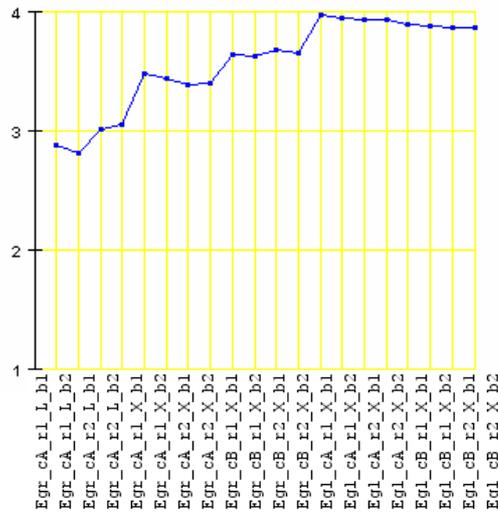
Yang J, Park S, Kamdem DP, Keathley DE, Retzel E, Paule C, Kapur V, Han K (2003) Novel gene expression profiles define the metabolic and physiological processes characteristic of wood and its extractive formation in a hardwood tree species, *Robinia pseudoacacia*. *Plant Molecular Biology* 52: 935–956.

Yang S, Zyl L, No E, Loopstra CA (2004) Microarray analysis of genes preferentially expressed in differentiating xylem of loblolly pine (*Pinus taeda*). *Plant Science* 166: 1185–1195.

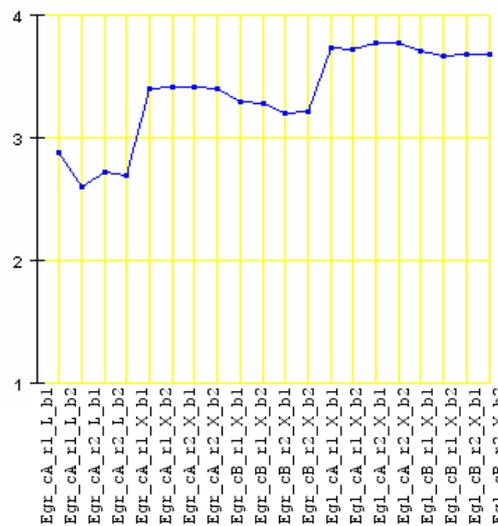
Yokoyama R & Nishitani K (2006) Identification and characterization of *Arabidopsis thaliana* genes involved in xylem secondary cell walls. *Journal of Plant Research* 119: 189–194.

Zhao C, Craig JC, Petzold HE, Dickerman AW, Beers EP (2005) The xylem and phloem transcriptomes from secondary tissues of the *Arabidopsis* root-hypocotyl. *Plant Physiology* 138: 803–818.

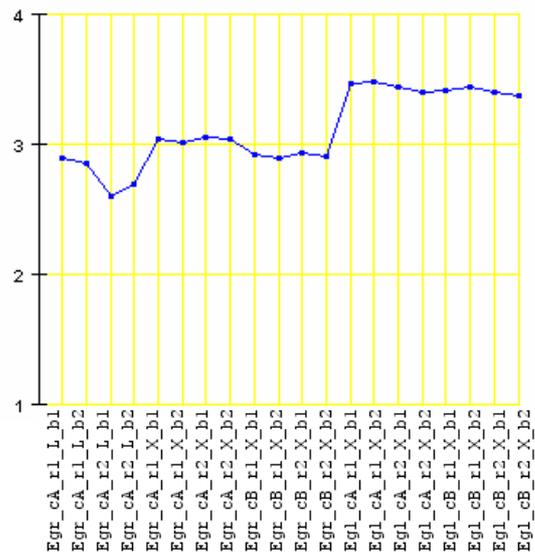
Anexo 1. Gráficos resultantes da análise *in silico* do perfil de expressão relativa dos genes diferencialmente expressos entre xilemas de *E. globulus* e *E. grandis* de acordo com o programa SAM. Os nomes das proteínas potencialmente codificadas pelos genes estão indicados abaixo dos gráficos. O eixo "y" representa a expressão relativa em escala log2. O eixo "x" representa as amostras analisadas. Destacados em vermelho estão os valores de expressão para *E. globulus* e em verde para *E. grandis*.



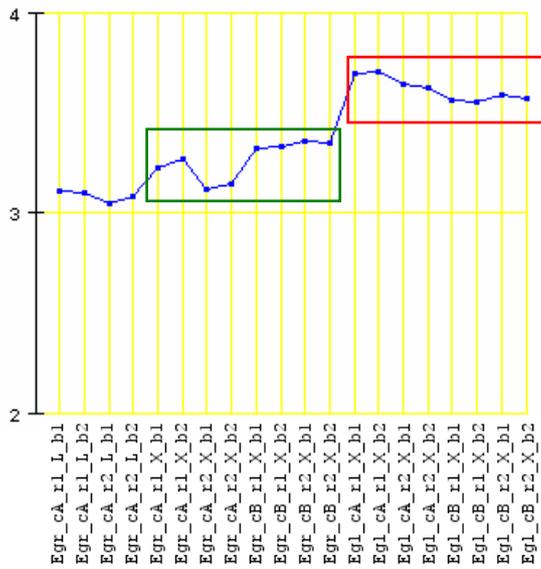
Corismato mutase
[*Arabidopsis thaliana*]



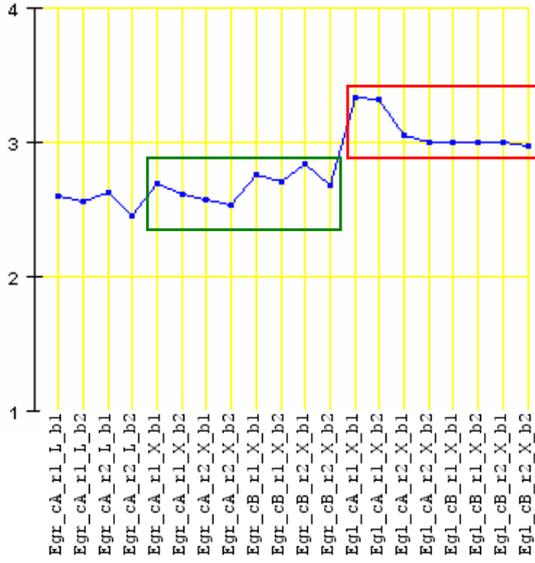
Citocromo P450
[*Petunia x hybrida*]



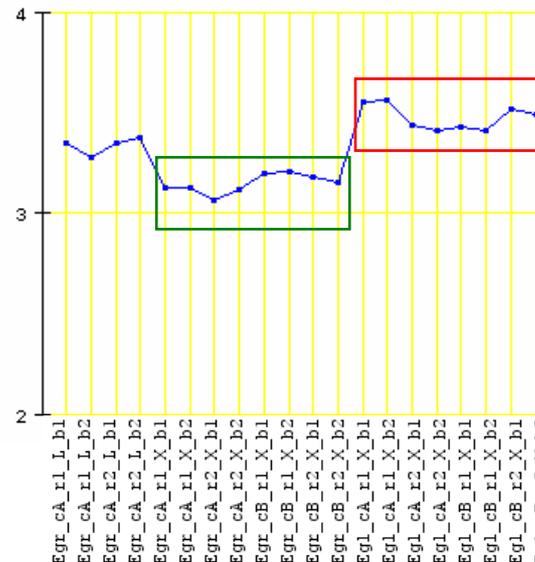
Proteína de função desconhecida
[*Arabidopsis thaliana*]



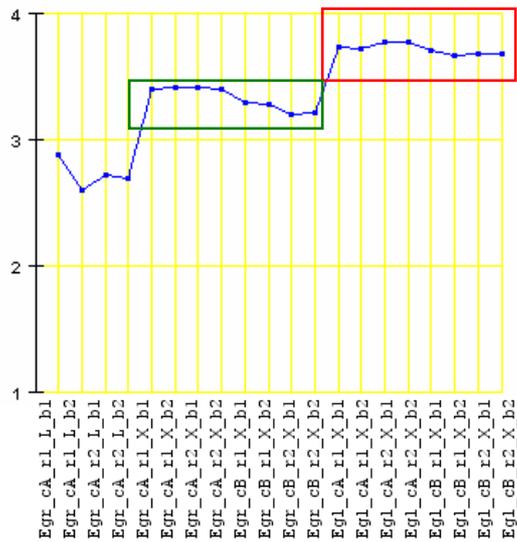
Proteína tipo nodulina
[*Arabidopsis thaliana*]



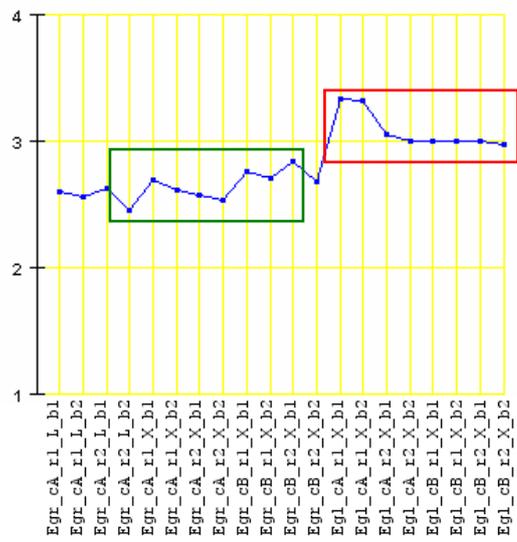
Proteína expressa em *Arabidopsis thaliana*



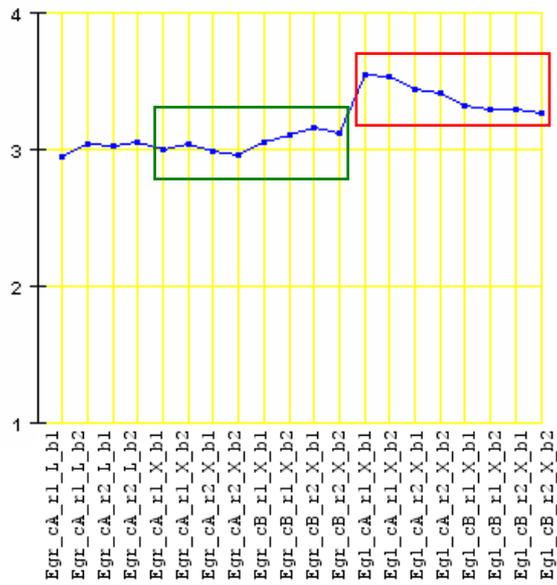
Aquaporina PIP1-2
[*Vitis berlandieri* x *Vitis rupestris*]



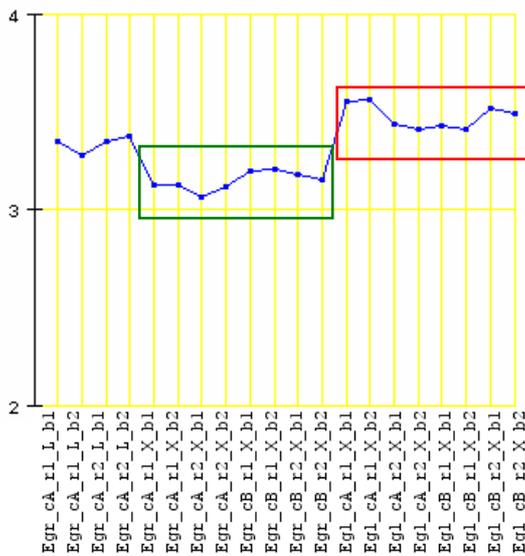
Flavonóide 3',5'-hidroxilase
[*Vitis vinifera*]



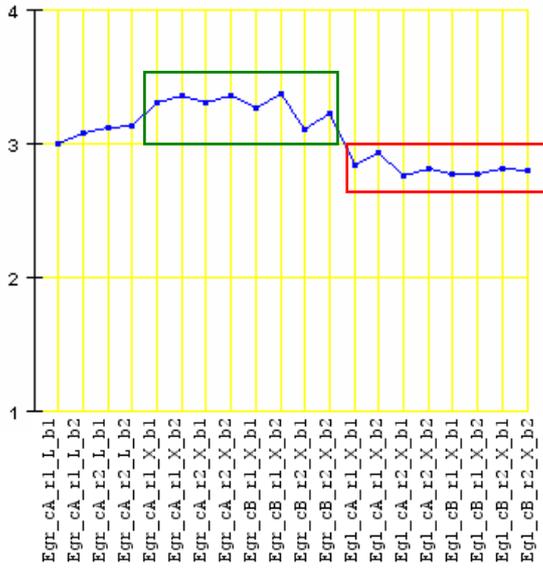
Proteína tipo poligalacturonase
[*Fragaria x ananassa*]



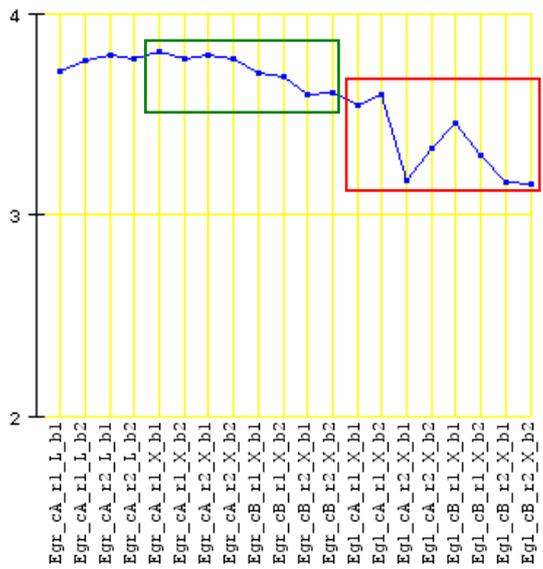
Arginina descarboxilase
[*Prunus persica*]



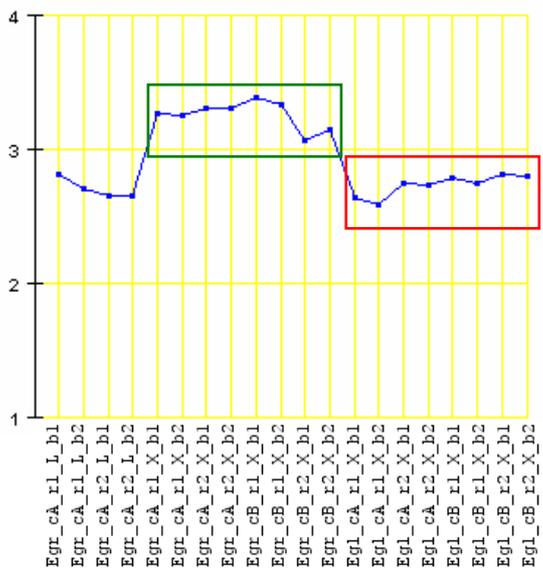
Aquaporin [*Arabidopsis thaliana*]



Proteína tipo calcineurina B
[*Pisum sativum*]



Thioredoxina [*Nicotiana tabacum*]



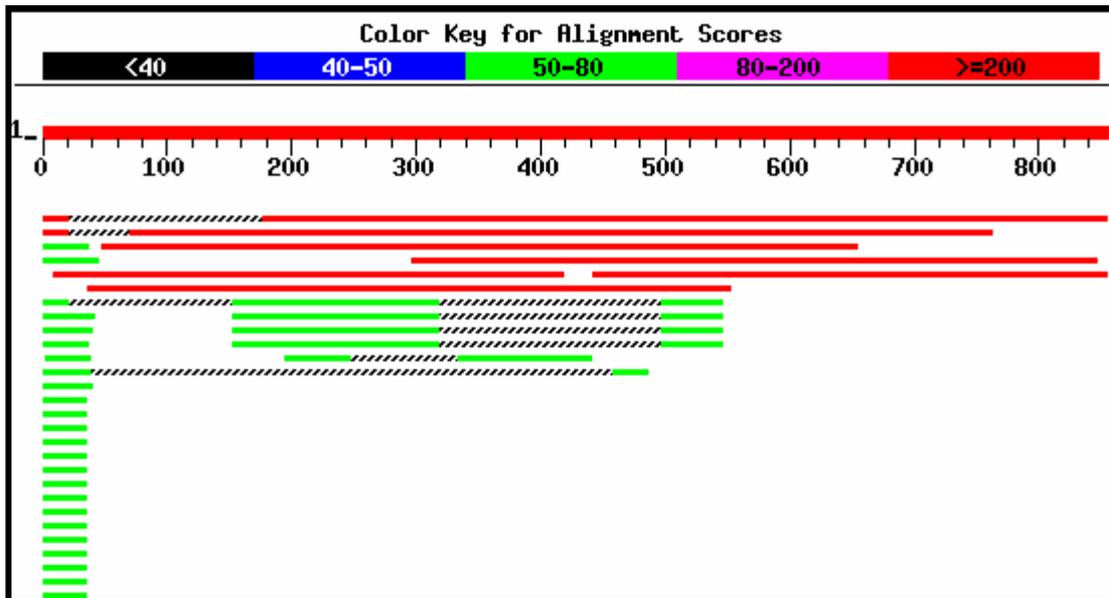
Família de proteínas inibitórias de invertase/pectina metil-esterase [*Arabidopsis thaliana*]

Anexo 2. Alinhamentos realizados com a seqüência do *Contig* utilizado na projeção dos oligonucleotídeos, sintetizados pela NimbleGen para hibridização dos microarranjos. Os alinhamentos foram realizados com a ferramenta BLASTn no banco de ESTs do Genolyptus. Para cada gene está representada a seqüência consenso e os locais de projeção dos *primers* utilizados no qRT-PCR, bem como o gráfico que ilustra o resultado dos alinhamentos.

EuC01

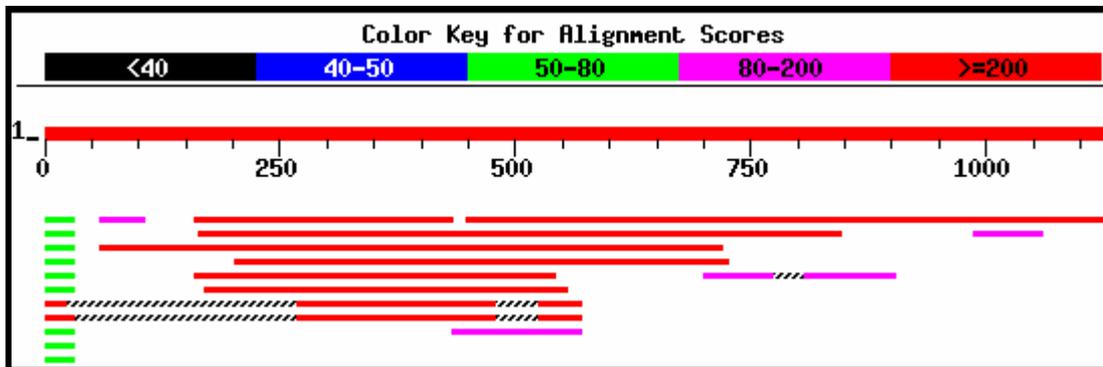
```

CCGGGTCGACCCACGCGTCCGCGGACGCGTGGGCGGCGAACCAGTTGGAGTTTGGCGTACGGTCCGCTAGGGT
TTCCCCGTCACCTCCCTCGCTCCCGCAGGGAAAATGGCACGGGTATATGTTGGTAACCTGGATTCTCGTGTTA
CAGAACGGGATCTTGAGGATGAGTTTCGCACATTTGGAATTATAAGAAGTGTGTGGGTCGCCCGAAGACCACC
GGGATATGCTTTTATTGATTTTGATGACCATAGGGATGCACAAGATGCGATCCGTGAGCTAGATGGTAAAAAT
GGCTGGAGAGTTGAGCTCTCTCATAACTCTAGTGGTGGACGTGGTGGTTCGTGGTGGTTCGATCTGGTGGTTCTG
ATTTGAAGTGCTATGAATGTGGTGAGCCTGGACACTTTGCTCGGGAGTGCCGCCTGCGTGTGGGTTCCAGCAGC
ACGTCGAAGTCGCAGCCGCAAGCCGCAGCCGCAGCCGCAGTCGCAGTCCAAGATACCGCAGGAGCCCAAGCTA
TGGTCGAAGGAGCTACAGCCACGTGGTCGATCTCCAAGACGTCGCAGTCTCTCCTCGTGGTCGCAGCTAT
AGCAGGTCGCCCCATTCCGCGCCCGTGACGAAGTGCCATATGCTAATGGCAATGGTCGTGATCGTCGCCGAA
GCAGGAGCTGAATGGTGCAGAGCTAAAAAATGTGGTTTTGTCTGGATCGTATCTGGATTAAGTGTGTTTTAGTA
GTTTGGACTTGGTACAACCTTTGTTTTAGTCGGGCTCGCCCTCAGCCCCACCTCTTCGATTTGGATAAGTTCT
CATTACTGAGTATGTTTTCTGTCTTTTACTTTGTCTCAGAATTATATGGTTAT
  
```



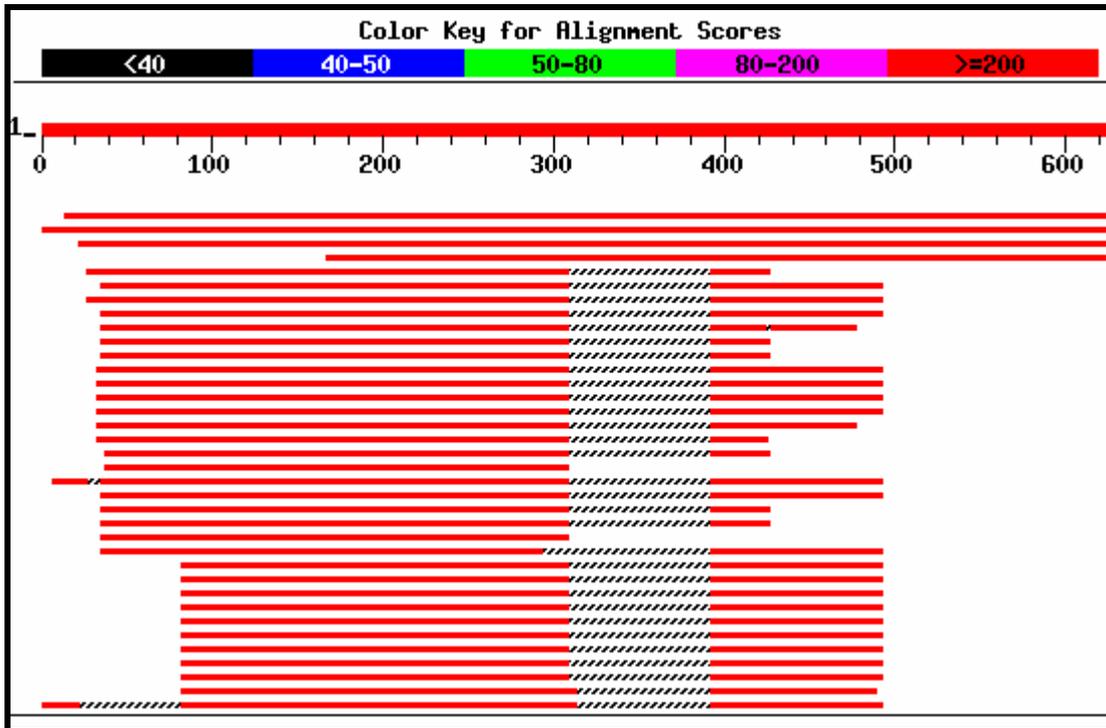
EuC02

GTCCGGAATTCCGGGTCGACCCACGCGTCCGCTCTCTCCCTTCTTTCTCTCTCTTCTTCGTGCGCGATCGCCA
TTGAAGAAGCTCCGCGCGAGACCTCGGAAGGCGGAGGAGGAATCCCCCACGCGATTTTCGCATCCGCCGCCGC
CGCCGGCCGCCGCTTTCCGTTCTTATTGCAATTCTCAAGATAGATCCATGGCATTTCGAGCAGTACTTTGCCCA
GGAGTGAGGCCATCTCTGGACCAGCTATGGATTCTGGAAGTAGCGATGGCTGTTTCGACTGCAACATCTGT
CTAGACTTTGCGATTGAGCCTGTGGTCACTCTCTGTGGTCACCTCTACTGCTGGCCCTGCATCTACAAATGGC
TCCACGTGCAAAGCGCCTCGCTTGCTTCTGATGAGCACCCACAGTGTCCCGTCTGCAAGGCTGAAATATCCCA
CACAGCCATGGTCCCTCTCTATGGCCGTGGCCAAAGCTCCAAAGAGTCTGATCTGCAAGACAAGGCACTCCAA
CTAGGAACAATTGTACCCCGAGACCAGCGGCTTGTGGCATCCAAGCTCTCGCCTCTACAACACCCCGCAGTG
GTCAGCAGCTCCCTACCGTAATCCTTACCAAAATCCGTACTACAGCGCCAATTCGTTTGAAGAAGATTCCACC
ATCACCATTCCCTCAGCATGGGATCCGTAGGCCTCGACCACCTCGATGGTGGGAATGTTTGGGGAAATGTTTTAT
TGCAAGAGTCTTCGGCAATTCGGACAGCCTATATGGTTACCCTGAATTCATATAGGTACATGGGAGGAAGTAG
CCCTGAGGTTGAGGAGGCAGGAAATGCAGGCGAATAAGTCGCTGAGCAGAGTAACTTTCTTCTCTTCTGTTG
CTTCTGCTCTGTCTAATTGTATTCTGAGCTCGGTTGAAGCCGGTGTTCACCCCTCACCCATTTTGTATA
CTTGTATAGCGTAAAAATGAATCTTAATGGTACTAGTAGACGGCACGTCACACGGCTTGCCTAGAGAATGACT
CTTTGTTACATATTGAGAGATGGATTGCTATAGACAAAACATCATGGTCTTCGTGAGCGTGTCTCAGCGACC
ATCTTGTGCTTTAACTTCTAGTCCTCTTTGGCGTTC



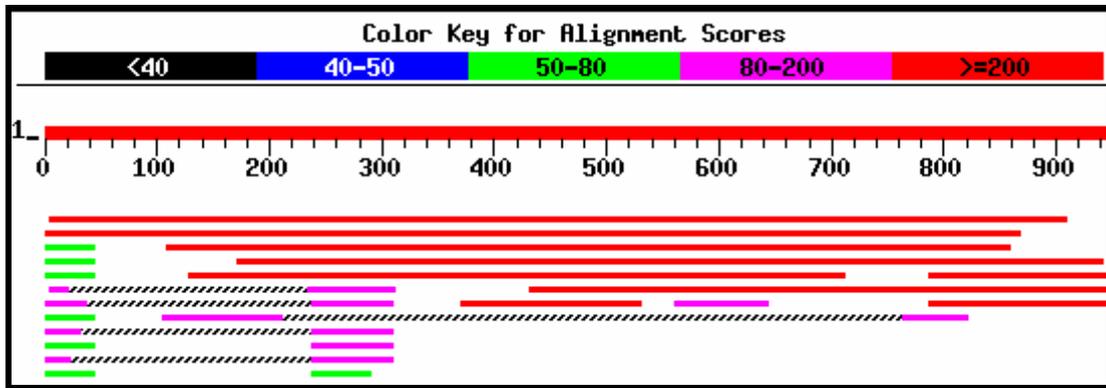
EuC03

CACGGTCCGCGGACGCGTGGGCTGTCTCTCTTCTTCTTCTCGATCGATTCTCGGCCAAGCCCCTTGCGAAC
TCTTTCAAGATGCAGATCTTCGTCAAGACCCTCACCGGAAGACCATCACCCCTCGAGGTGGAGAGCTCTGACA
CAATCGACAACATCAAGGCCAAGATCCAGGACAAGGAGGGCATCCACCA**GACCAGCAGAGGCTCATCTTCGC**
CGGAAAGCAGCTCGAGGACGGCCGCACCCTCGCTGATTACAACATCCAGAAG**GAGTCCACCCTCACCTTGTG**
CTGAGGCTCAGGGGAGGATCTTGAATTTTTTCAAGAACCCAGCGGCAAGATCGAGTCCAATATCCAGGACG
AGGCAGGCATTCCATAACCAGGAGCCCGAGATCCAGGACGGGGAAGGCGTTCCCCAGACCAGGGGCCGCACCC
TGGCGGACCACAACATCCGGAAGGGGGTCAACCCTTACCTCGAGCTCCGTCTTCGTCTCTAGAGTGGTCTATT
GGGGCCATCTTCTTCAAATTTCTTGGCTCTAGGAAGAATAAGTTTTATTCCATTGTCTCTGCATTGT
TATCTGTTATTTGGAGGGTTCAAGTTGAGTGAGTGTGGTC



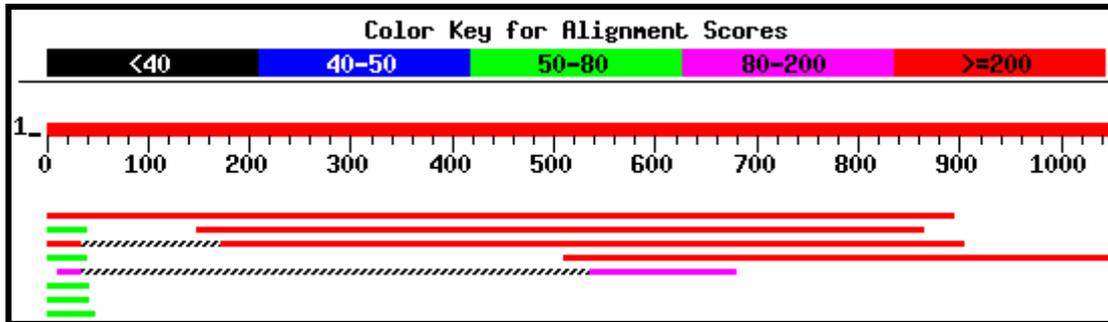
EuC04

CGGTCCGGATTCCCAGGTCGACCCACGCTTCCGCTCTTTTCTCTCACACGCAAAAAGCACCCGTTCTTGTTCT
CTACCTTTGCTGCAAGTCTTTCGGTATTCTCTTTCTCCAGTCTCTGCAAGATGGGTCTGCTCTCTCTCCT
GTTGCTCCTTGCCATCTTCTTGTCTCTGAGCCCAGCACTCTCCACCAACTCGGAAGGAAATGCTTTGCATGC
CTTGAGAACCGGGCTCTCAGATCCGAACAACGTGCTTTCAGAGCTGGGACCCGACTCTGGTCAATCCCTGCACT
TGGTTCCATATCACTTGCAGCTCCAACAATCATGTGATTAGATTGGATTTGGGTAATTCTAATGTCTCTGGGA
CTTTAGGGCCAGAGCTTGGCCAGCTGAAGCATCTCAAATACTTAGAACTTTACATGAATAGCCTGTCGGGGGA
AATTCCGAAGGAGTTGGGTAATTTGAAGAACTGGTTAGCTTGGATATGTATGGCAACAAATTGGAAGGACAA
ATACCA**AAGGCCTTTGCTAAGCTGAA**GTCACTTAGATTTCTCCGTCTCAACAACAATAAGCTCACGGGACCGA
TCCCAGGGGA**ACTCACCAACCCTCTCTGACC**TCAAAGTATTTGATGTTTCTAACAATGATCTCTGTGGAACGAT
TCCCGTGGATGGTCCATTCGCCAATTTCCAAATGGAAAGCTTTGAGAACAACCGGCTCAATGGACCGGAGCTG
CAAGGACTAGTGCCATACGACTTTGGATGCTGATGAAAGCTCGCTTGCCTGGATATCATGTCTAGGACCCTAA
AGTTAAGTTGGTGGTAATTA AAAACTCGTCAAGCTCAGGCGATCTCGGAGACATATAAAAACAACAATAGTCGC
GTTAATAGTCAGAAGTCGGAATGTACTTGCAAGTTAGCTGAATAAAAGTATTGGCACTTGATAATGTGTTCCA



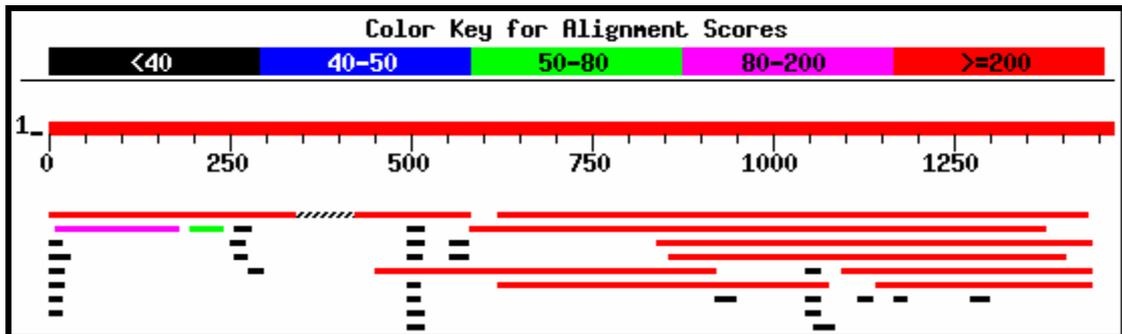
EuC05

CGGTCCGGATTCCC GGGTCGACCCACGCGTCCGGTCTGAAGATCAATTGCACTGTGGGTGTTGTGGTCTCGGA
 AATAGACTGTCCAAGATTACATTCCATACAAGTTCCTGAGTCTGATATTGGTGCACACTTTGGCTTCTTGCTG
 GAGAAAGAAGAAGGCTCCGATGTTACTTTTAATGTGTCTGGGGAAAAGTTTCATGGTCATAAGCTGGTATTGG
 CTGCTCGATCTCCTGTATTTGAGTCTGAATTCCTGAATGCAATGGAGGAAGACAATGATGAAGTAATTATAAG
 TGATATGGAACCTACAGTCTTCCAGGCTTTGCTACATTTTCATCTACAAAGACACTTTAGATGAAGACGAGGAG
 CTTTCAATAACCAAGTGCCTTTTGCAGTCCATCAGAAAATGCCGGTGGCAAAGCTTCTAGCTGCTGCAGATAGAT
 ACAATTTGCCTAGATTGAAACTGATTTGTGAGTCCATTCTCTGTAAGGGTATATCTGTGGATTCTGTTGCCAG
 GATCTTGGCCCTTGCTGAGCAGCATCATGCCATGGATCTAAAATCTATTTGTTTCAAATTTGCTGCTGAAAC
 CTTGTAGCTGTCATGCGCTCAGATGGTTTTGAGTATCTTAAAGAGAACTGCCCATCTCTGCAGTCAGAGCTCC
 TAAAGACAGTTGCAGGATGCGAAGAGCAAGTCAGTGGAGGAGGAAATCGAAGTGTTTTTGCACAATTTTCTGA
 TGGCAATGAAACAAATGACAGGAGTATTAGGCAACAAGACTGGGATAATTTGAATGTGAGAACTCAGAGCTTG
 TGGATGCCCTCTCCGACGGAGGAGATGCAAGTGGTCCGAGCCCTGGCCAAGAAAGCTGAATGGTAAATATGT
 TGCTTTGGATGACCAATTAACATTTTCGTTTGTCTGCCAGATCAAGAATGTGAAAGAAAAGGACTACATTAGAA
 ACAAAGAGTGGGGGAAAATTTTTCCCTCCTCTTCTTTTGTCCGTAAACCATCCATGTCTCTNCAAGATTT
 GGGTTCTGTAGATTATTCTTGGCTTTCCCT



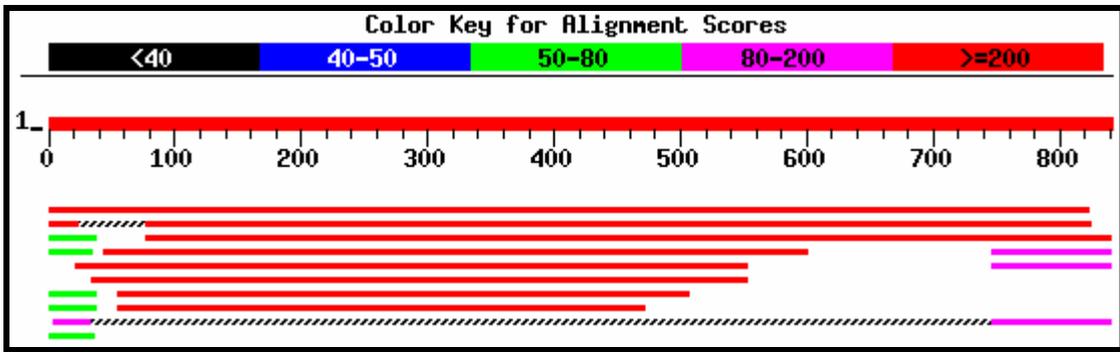
EuC06

CCACGCGTCCGCCAGAGGTTGCTTGGGTGACAAAATCTGGGGAATCTGATTTGGAAGTGCCCATTTGCAATCCG
ACCAACAAGTGAGACAGTGATGTATCCTTACTTTTTCTAAGTGGATAAAGGGGACACAGAGACTTGCCTTTGAGA
CTTAATCAGTGGTGCAATGTTGTTAGATGGGAGTTTAGCAATCCTACACCCTTTATCAGGAGTCGTGAGTTTC
TTTGGCAGGAAGGACATACAGCTTTTGTACAAAAGAAGAAGCAGATGCAGAGGTTCTTGATATACTGGA
AAATTTTATAAGGGGAACGTATATATGAAGAGTACTTAGCTGTTCCAGTAATAAAGGGCAAGAAGAGTGAGT
TGGAGAAGTTTGTGGTGGTTTTTATACAACACTACTGTAGAGGCATTTGTTCCCTAACACTGGGCGTGGAATTCA
AGGGGCAACATCACACTGTTTGGGTCAGAAATTTGCAAAAATGTTGAAATATTCTTTGAGAATGAAAAGAGA
GAGAAAACCATGGTCTGGCAGAACTCGTGGGCCATACCCTAGAACGATTGGGGTTCATGGTGTGGTTCATG
GGGATGATAAGGGGTTGGCTCTCCCTCCCAAAGTGGCATCAGTTCAAGTTATCATTGTACCTGTTCCATATAA
AGATGCTAATACTCAAGGTATTTTTGATGCCTGCACAGCCACTGTTAATACCTTGTGTGAAGCAGGTATTCGT
GCAGAGGCAGACCTCAGAGATAACTATTCGCCTGGTGGAAAGTACTCACATTGGGAAATGAAAGGTGTTCCCTC
TTAGAATTGAAATTGGACCAAAGGACTTGGCAAATAATCAGGTCGAGCTGTTCCGGCGTGATAATTCTGCTAA
AAAAGATATACCGAGGGCTGACTTAGTTGAGCAGGTGAAAGAATTGTTGGCCAACATTCAACAGAGCCTGTTT
GATGCTGCAAAACAAAAACGAGATGCTTGCAATCAAGTTGTGAAAA**CATGGGATGAATTTGTGGAA**GCACTGG
GCCAAAAGAAGATGGTTTTGGCTCCTTGGTGTGATGAGGAAGAGGTGGAGAAAAGAAGTAAAAGCACGGACGAA
GGTGAGATGGGAGCATGTAAGACACTTTGTTCCCCATTTGAAGAGCCCCGAGCTTCCTGAAGGTACTATATGC
TTTGCTTCTGGAAAAGCCAGCAAAAAGGTGGAGCTACTGGGGCAGGAGTTACTAGTTCTACCGTCGGTCCTTCC
TCCACATGAAGATCGGATCGGCCCCAGCTTGGAGCAAGCTCAGTCAAATTCAGACCAGTAATCTTATGGCAGT
GCTGATTTATTTATTGACATCGACTCGAGCCACCCAAGTTTTCTGAATCATAGTTTGGCGTTTTTCCTTTTTTT
CAATACACAGGTGATGTTTGAAAAATGCTTTGAAAACATGGAGCAATACCACCCAAAAAAAATAAAAAATTGA
TATTTAGTA



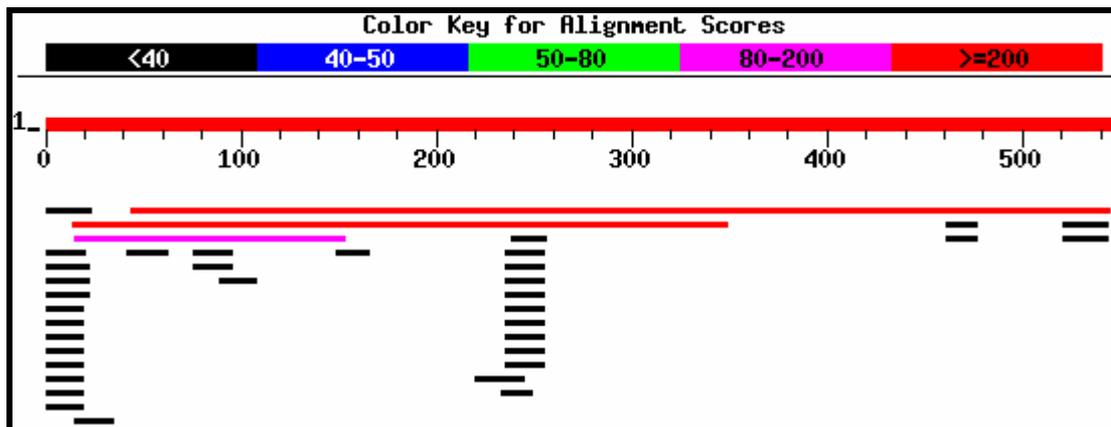
EuC07

CGGTCCGGATTCCCGGGTCGACCCCGCGTCCGATTCCACAATACAGTAGCAGAGAAAAAGCTCTCGAGCAGCT
CGGTTCTCCGACGATGGAAGGGGGCGGCGCGTCCGTACCAGCGGTTCCCGGTGGTGAAGGTCCGGGAGACG
AAGGACGACTACCTCAAGTTCGAGCTCCGGGGGACGGACGCCAGCGTGGCCAACGCCCTCCGCCGCGTCATGA
TCGCGGAGGTCCCCACCATCGCCATCGACCTCGTCGAGATCGAGGTCAACTCCTCCGTCTCAACGACGAGTT
CATCGCCACCAGCTCGGCTCGTCCCCCTACCTCCGACCGCGC**CATGTCCATGCGCTTCTCCCGC**ACTGC
GACGCTGCGACGGCGACGGCCAGTGCAGTTCTGCTCCGTCGAGTTCCACCTCCGCGCCA**AGTGCATGACCG**
ACCAGACCCTCGACGTCACCAGCAAGGACCTCATCAGCTCCGACCACACCGTCGTCCCCGTCGACTTCTCCGA
CGCCCTCGCCGACGATTCCGCCGAGCAGAAGGGAATTACTATTGTGAAGCTTCGCCGAGGGCAAGAGCTAAGG
CTGCGAGCAATAGCAAGAAAAGGAATTGGCAAAGATCATGCAAAGTGGTCCCCCGCAGCCACAGTGACATTCA
TGTATGAACCTGAAATTCACATCAATGAAGATTTAGTGGAGACCTTGTCCTTGAGGAAAAACAAAGTTGGGT
GGAAAGCAGTCCACCAAAGTGTGACATTGATCCAAATACTGAGCAGGTGGTAGTGGTTGATCCTGAGGCT
TACTTATGATGATGAAGTAATAAAAAAAGCTGAAGC



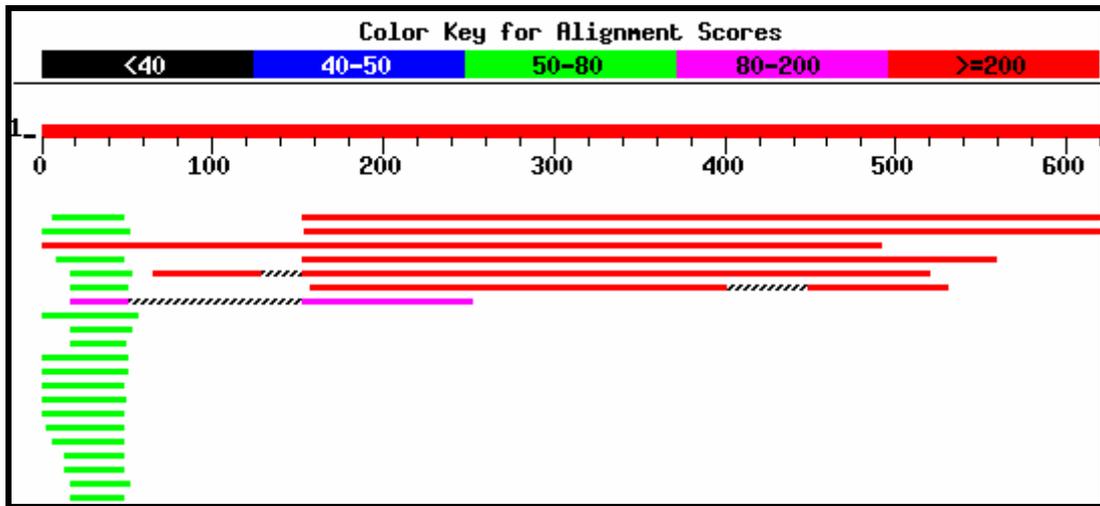
EuC08

GACCCACGCGTCCTATTGAATGAGGGTGAAAATGAGGTTGCGATTGCTAATCCAAAAAGACTTGTCTCCTCTG
GAGTCTGTGCTTCTGATATTGGAATCATAACACCTTATGCTGCACAGGTGGTTTTATTGAAGATGCTGAGAGG
CAATGAAAACAAACTCAAGGATATGAAAATCTCAACGGTGGATGGTTTTCAAGGTCGGGAGAAGGAAGCGATC
GTCATTTCAATGGTTCGTTCAAATTCAAAGAAAGAGGTGGGTTTCCTGAGTGACCGCAGGCGAATGAATGTGG
CTGTTACACGGGCAAGAAGACAGTGTGTCTCATCTGTGATACCGAAACCGTGAGTAGTGATGGATTCTTAAA
GAGACTAATTGAGTACTTTGAGGAGCATGGCGAGTATCAAAGTGCCTCAGAGTATTGCAACGAATAAGCAAAC
CGTCTGGTTAGGTGAAGCCTTTGTTATTGCTTTCCTTTGGGGTACTTGTTCATACAGGTTTCGGACTAATTGTT
AGGGCCTTGAATTCAACACCCAAAAAGAAAGAAATG



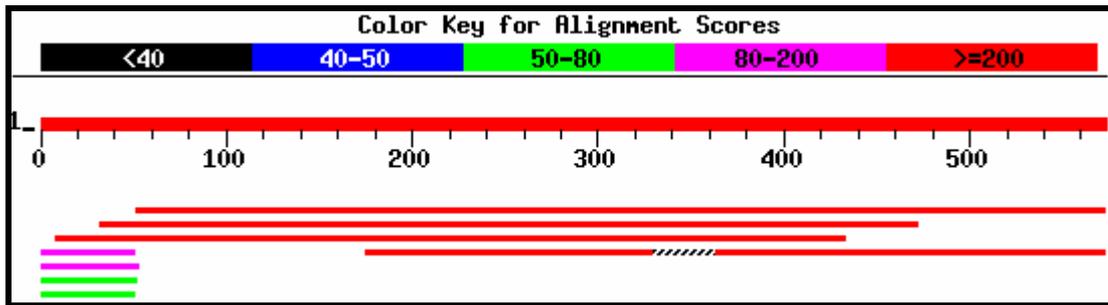
EuC09

AATGGTACCTGCAGGTCACGGTCCGGAATCCGGGTCGACCACGCGTCCGCTCCGGGGATCAGGAAGAAGA
TAAGGGAGAGCGTCCGTCCGCCCTCAGCCTCGGCGGCTCAGTCCCACCTGAGACAGCGAGAGAGAGAGAGAG
AGAGAGAAAGGAAAATGGAGAAGTACTTCGGGAACGCGTACAGGGGCGACCCGGGAGTGCCGCACGCCGACCC
GGACAGGTTTCGTGAACATATGGATCGGGTCCGCCGCTTCTCCGCCCTCACCTGGTCCAACCCTTACATTTGG
CAGCTCTCCAATCAATTCAACTGGCATGACAAAGCCATGATGTTTCGAGCACTATCACTGGAAGAAAGCCAAGG
CTAAA AATCAGCCATATAATTTCAAGTGGAAATGAGTACATGGACAGGGACCACCGGACTCATACTATTTCAA
TTGGCCTGTTTACTTCCCATAGGAAGTATCATTGTATTGGGTTGCCTTTGGGTGCTGACCTTACTCAAATAAC
ATGCGCTTGTTTCAGGAGTTTGATAGTATTTCTCATGGTTCCTTCTTTTACCTTGTCTATTCTGTTTGAA
TGTTTCAACAAAAACAAAAACAGTTGGCTTCTTATGCTGCT



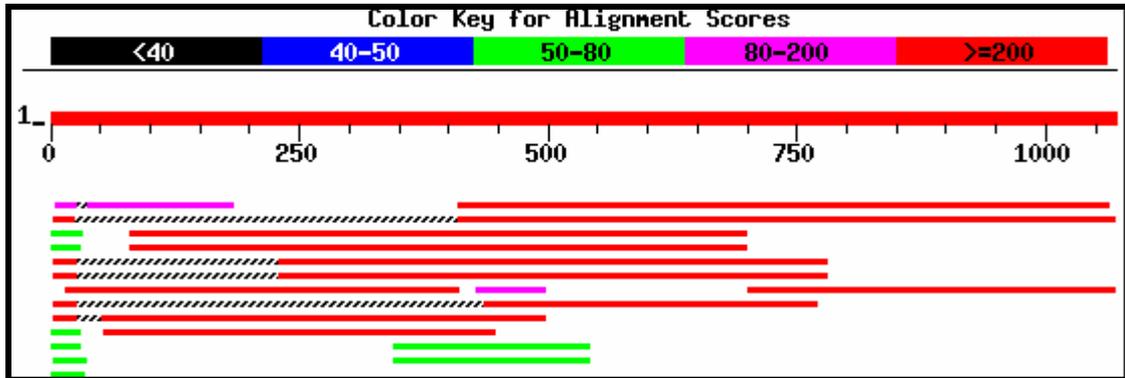
EuC10

NAATGGTCACCTGCAGTCACGGTCCGGAATCTCGGGTCGACCCCGCGTCCGGCGGAGGAACGCGGAGGCAGGC
GGAGGAAGGAGGAGACGAGGCCGAGCGTAGAGCTGCACAATACATCCGATCGAGAGACATGAAAGTCGCTCCG
AAGGTGGTTTTCTGTCAAGGACCAGGACGGGCTCGGTCCGGCGTTCGCCGCCGCTCTCCGGCCGACCCGA
GCTCCGACCTTCGCCGGCTGGAGGAGTCCTTCGAGCTTCCGTTGGAGCGCTACGGCATCGGGGACGGCGGAGC
TCGCCGGGAAGTCGTTCACTTCGTCGATGGCCATGGGGTTTATCAGGTGTCCGTGCTGCAAATGGAAAGTTAC
AAGCCACCTCTATTGGCTTGCGCCCTTAATGAAGTTTTGGCACAATTAATGGATGAAACATCAGAAACGATGC
CTACTCTTCTAGTTCCCTTTATTTTCATCATCAACAAACCTCAAGTCTGAGAAAAGAAGTTTACCCATGGACAG
CAGAAGAGTATCTCTTTACGGCCTGCATATAGGGAGAAGAACAGACGTCACCTTGTGCTATCA



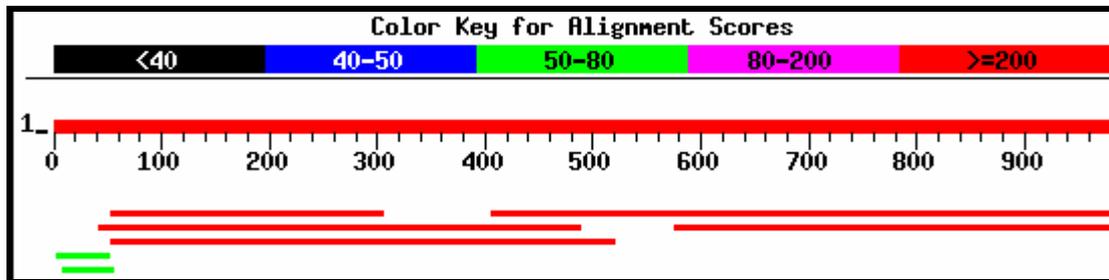
EuC12

GAATCCCGGGTCGACCCACGCGTCCGGCAACTTGCAGCTTTATTTGTGGCTGTGGGCAGGTTGTGGACTGTC
GTATTTGTGGTGACCCCAATTCTGTACTTCGCTTTGCCTTTATTGAGTTCACGGATGAAGAAGGTGCAATGGC
TGCCCTCAACTTGTGAGGACAATGCTTGGATTCTACCCTGTGAGAGTGCTGCCCTCTAAAACAGCTATCGCA
CCAGTTAACCCGACATTTTTGCCAAGGAATGAGGATGAGCGGGAGATGTGCACGAGGACCATATATTGTACAA
ACATTGATAAGAAGGTTACTCAAGCAGACGTTAAACTATTTTTTGTGAGTCTGTCTGTGGAGAGGTTTATCGTTT
GAGGCTACTAGGAGACTATCACCATTCTACTCGAATTGCTTTTTGTTGAGTTCGTAATGGCTGAAAGTGCTATT
GCTGCCCTCAACTGTAGTG**GCGTGGTTCTTGGATCACTA**CCCATAGGGTGAGCCCATCGAAAACGCCTGTTC
GCCACGTGCCCTCGCCTGCCGATGCATTGACCGTTCT**AGCACCTGACTTTGTCACCA**TCTGGATATATAT
TCGAACGCATGATTATCCCTTCTTCAATGTAAAAATTGCTCTTTGGCCATCTGGCTTATCTCTGTCTTATAGA
TACGGTCATCCACTTCGATATGAAAGCTTACCATAGGTGCCATGTTTTGTTTATTCTCACTTATAGAGTTGT
TGGATGGAACTTTGAGGCTCTAAGAGATTGATGTTACTATAGTTGGTTCATCACTCTTCTGGACATGGGATGC
TTTTTTAAGTATAGAATGCTCCCTGAGCACTAGCAGATTCTGACATGGTCTTTTGAACCTGACCCCGACAAATC
GGTTAGTTTTGATGCTTATGAGTCTTGTTCATGCTCGGTTTTGTGCGAGTTTAACTTATCATGAAAATATTAG
CTTTGAAATCTGTGTCCAGTTGTCCACATGGTGGTCTGATAATCTATGTCCACATTGTGGTTTTGTTTGTGCA
TANGGTATTTTTGCATGAAGTGGTGTGGTAGTGTATGTTGGCTTTTG



EuC13

AAGTGGTAGCCTGCAGGTACCGGTCCGGATTCCCTGGGTGCACCCACGCGTCCGATCCGGCCGGTTATGAGCCT
CTATCTCAATCCTTACTCCTGGAATCTAGTTGCCAACTTGTGTTTATTTGATTTCACCCGTTGGTGTGGTTTT
TCCTACTCCAATACGTCTCAGATTTGCTGAACAATGGAGACAAAAGGACTGTTGAGGATTCACTGGCATTTC
TACTCAAGTGGTTTGAGCGCTTCCACAGTTTAAGGGGAGGGAATTTTATATCACTGGAGAAAAGCTATGGAGG
TCACTATGTTCCCTCAATTGAGTAAAGCCATTGTAAAGTACAACAAAGCATCCGGTGAAAATGTAATCAATCTG
AAGGGCTATATGGTTGGAAAATGCTCTAACTGATGATTACCATGATCATCTTGGGGTCTTCCAATTTATGTGGT
CAGCTGGTTTTGATTTCTGATCAGACATTC AAGCTGCTGAATGTTCTCTGTGATTTCCAGTCTTTTATAACAC
CTCAAGTGCTTGTGATAAAAGTTCTAGATATAGCTAGCGACGAGCTCGGAGATATTGATCCATACAGTATATAT
ACTCCTCTTTGCCCTGTCAATGGTAGCCAGTCAAACCGCCTTCGCAAAAGAATGAAAGTCATTGGCCATGTCA
ATGAGAAAATATGATCCCTGCACGGAGGAGCACTCAGTTAAATATTTCAACCGTGCTGAGGTACAGACAGCACT
TCATGTAAACAAGGCATTTGCGCCAGCTAAATGGGACACTTGCAGTGACTTAGTTAGTAAAATTTGGAAGGAT
GCTCCTCGTACGGTGCTGGACATCTACCGTGAGCTTATGCGGAAGGACTTCGCATCTGGATGTTTCAGCGGTG
ACACGGATGCTATCATCCCTGTCACATCCACCGGTACACGATAGATTCTTTAAAGCTTCCAACGGTGAAGCC
TTTTTCGCGCTTGGTATGATGATGGGACGGTGGTGGAT



Anexo 3. Gráficos resultantes da análise *in silico* do perfil de expressão relativa dos genes constitutivos de acordo com o programa SAM. Os nomes das proteínas potencialmente codificadas pelos genes estão indicados abaixo dos gráficos. O eixo "y" representa a expressão relativa em escala log2. O eixo "x" representa as amostras analisadas onde

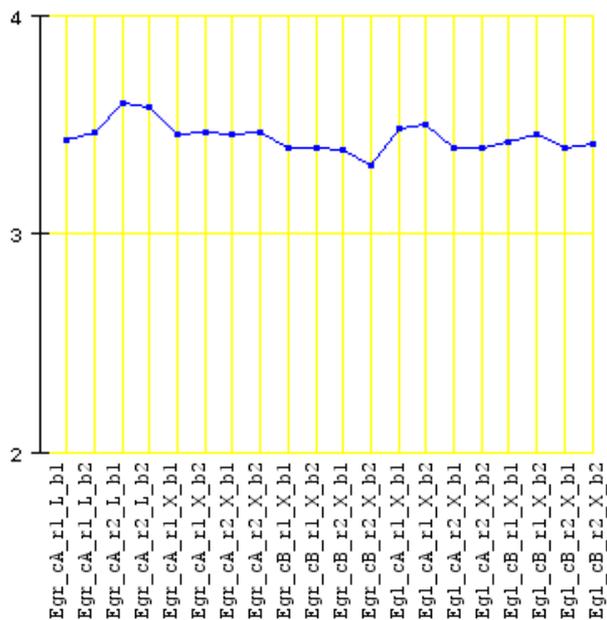
Egr - *Eucalyptus grandis*
Egl - *Eucalyptus globulus*

cA - matriz A
cB - matriz B

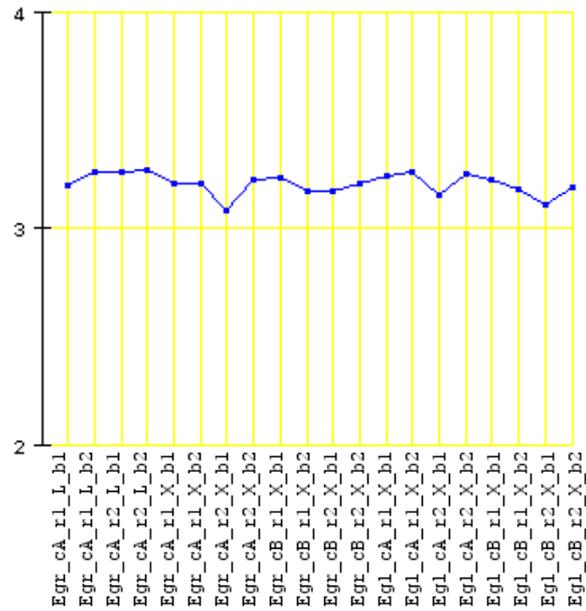
r1 - clone 1
r2 - clone 2

L - folha
X - xilema

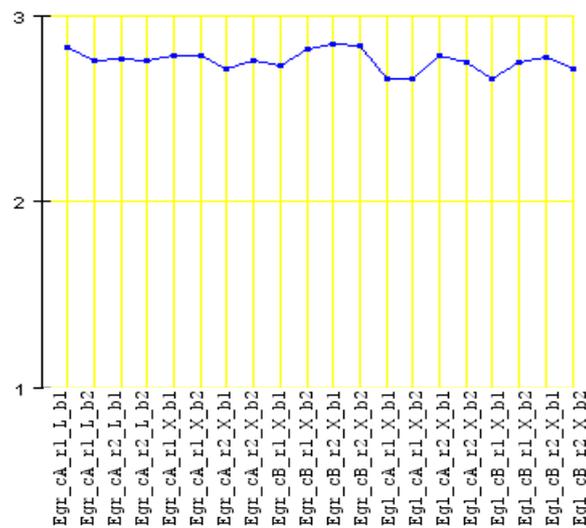
b1 - bloco 1 (do slide)
b2 - bloco 2 (do slide)



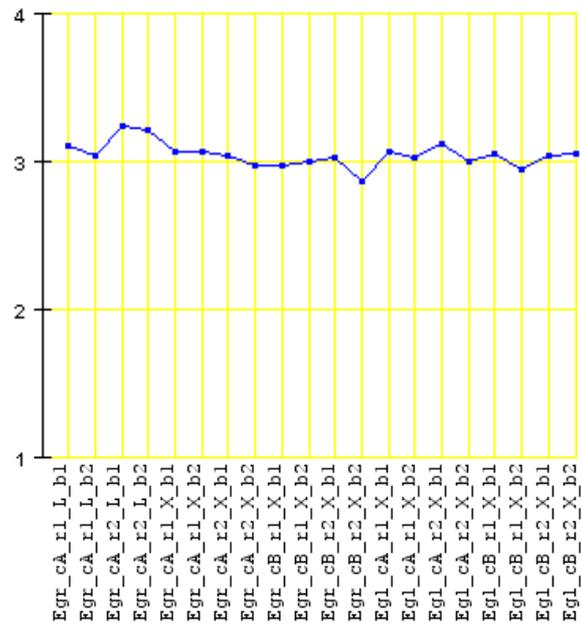
EuC01: Fator de *splicing* 7, rico em arginina/serina
[*Oryza sativa* (japonica)]



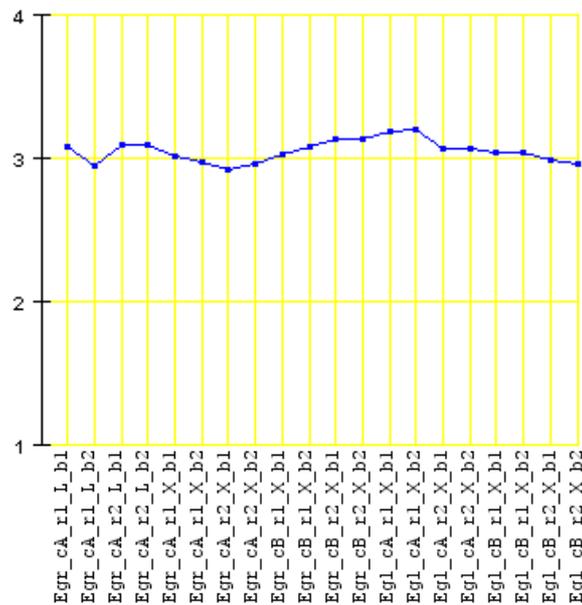
EuC02: Proteína da família RING *finger*
[*Arabidopsis thaliana*]



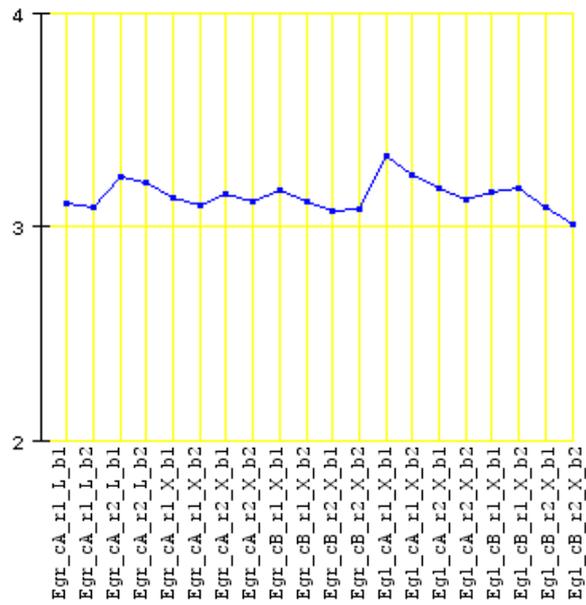
EuC03: Poliubiquitina [*Pinus sylvestris*]



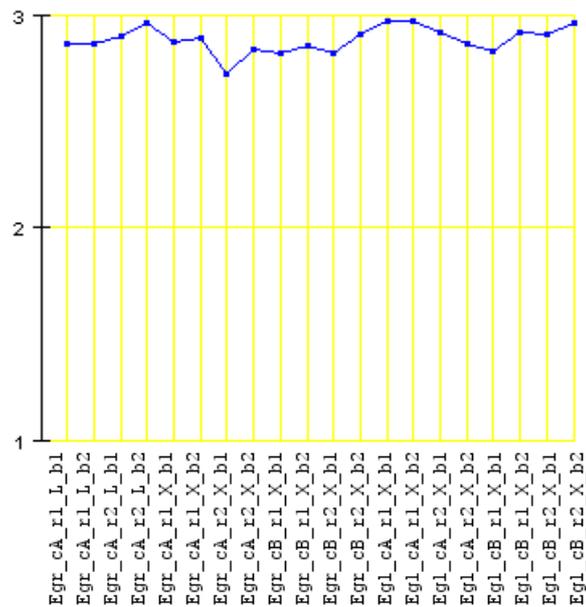
EuC04: Proteína rica em leucinas
[*Triticum aestivum*]



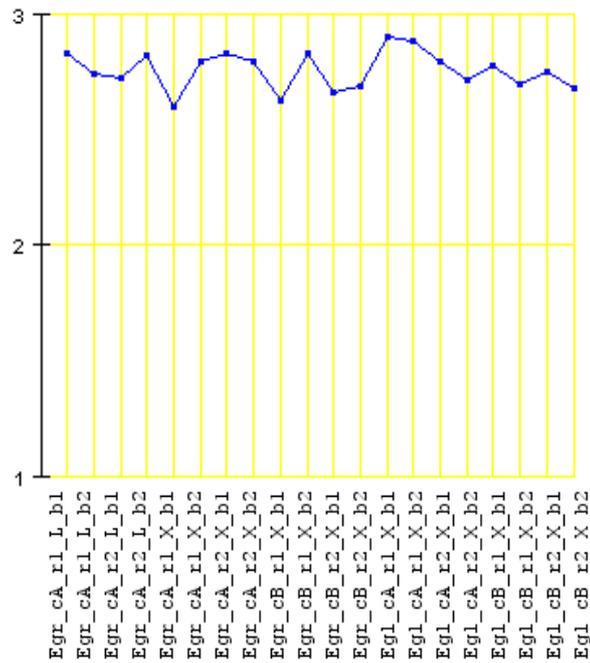
EuC05: BTB/POZ; MATH
[*Medicago truncatula*]



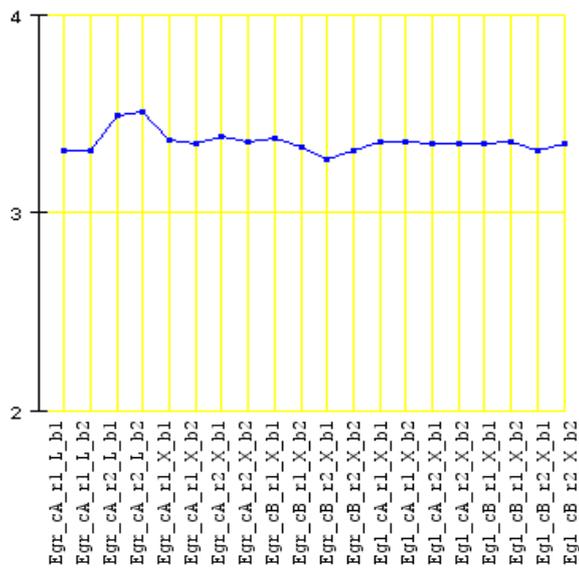
EuC06: Proteína de ligação a aminoacil-tRNAs
[*Arabidopsis thaliana*]



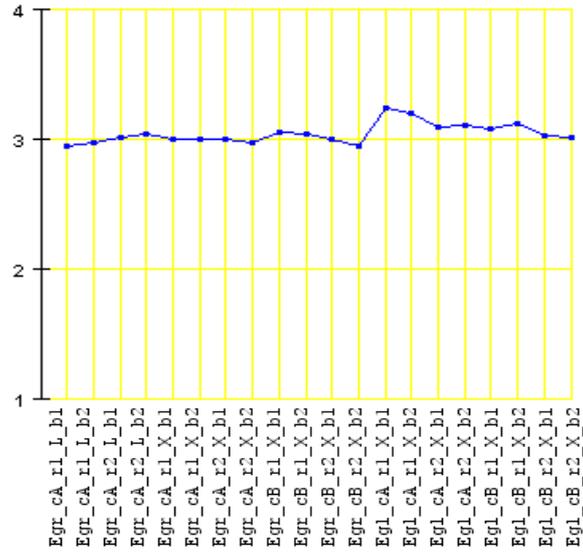
EuC07: Terceira subunidade da RNA Polimerase II
[*Arabidopsis thaliana*]



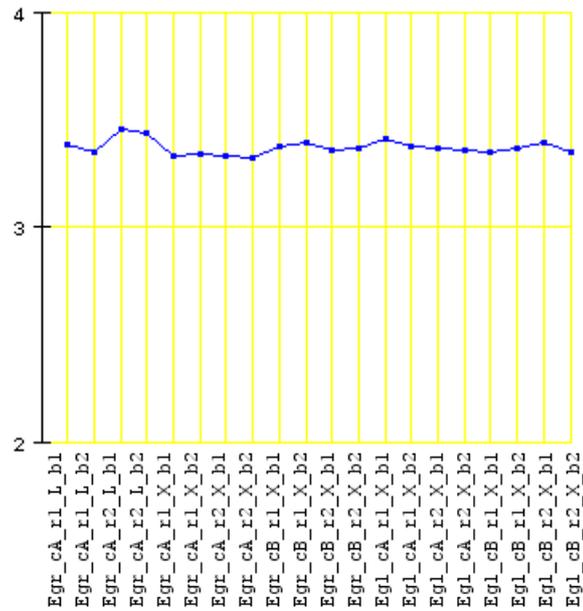
EuC08: Helicase ligante aos módulos DEAD/DEAH [*Medicago truncatula*]



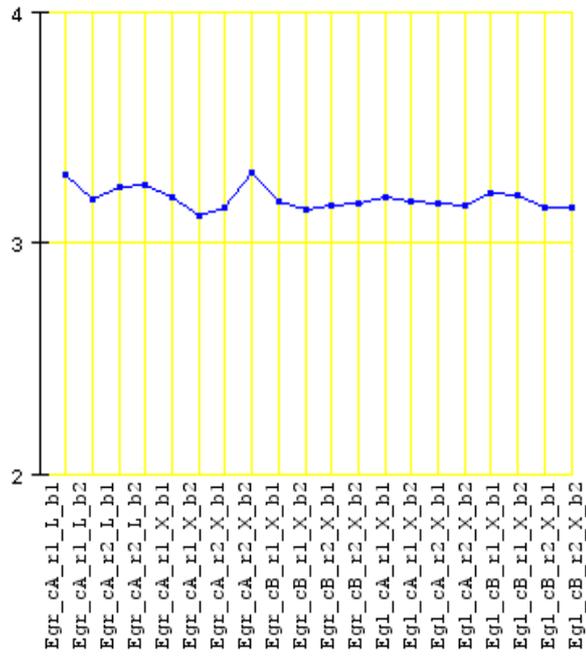
EuC09: Proteína expressa em *Arabidopsis thaliana*



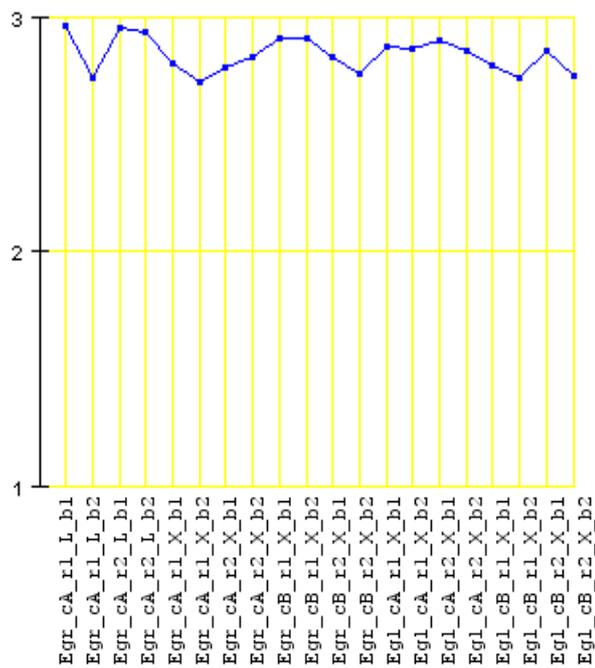
EuC10: Proteína hipotética de *Arabidopsis thaliana*



EuC11: Fator de transcrição [*Arabidopsis thaliana*]

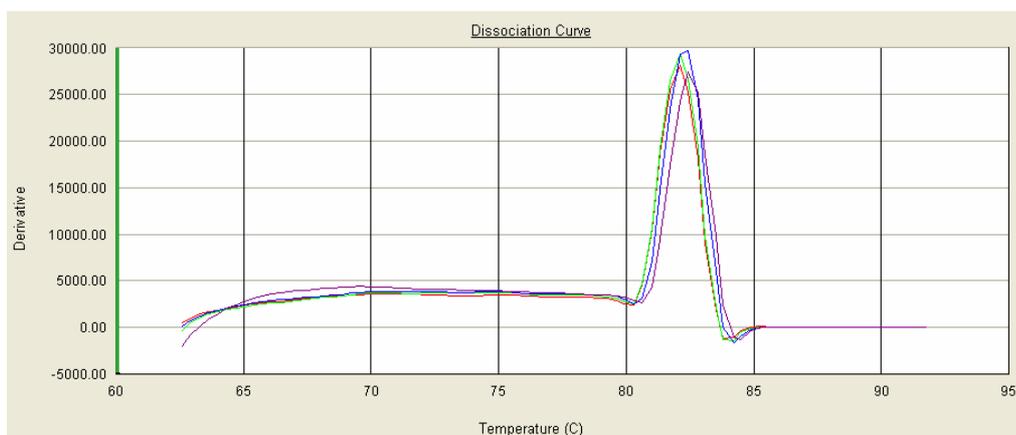


EuC12: Proteína de ligação a RNAs
[Arabidopsis thaliana]

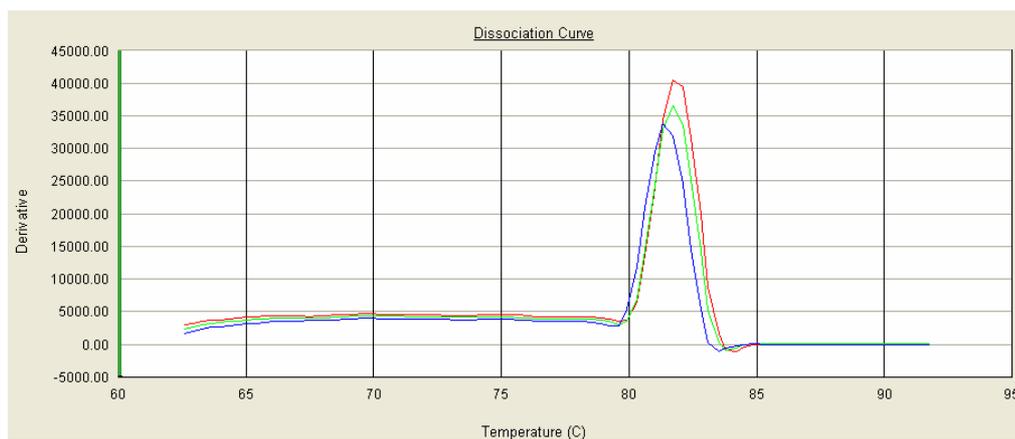


EuC13: Serina carboxilase
[Arabidopsis thaliana]

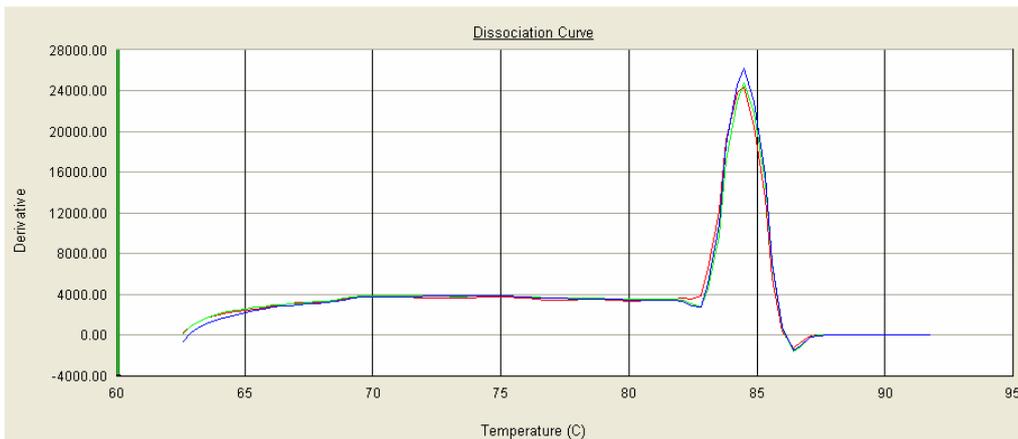
Anexo 4. Curvas de dissociação dos *primers* geradas durante qRT-PCRs segundo o programa *7500 System SDS Software* (Applied Biosystems). No eixo "y" está a derivada que indica a taxa de mudança na fluorescência do SYBR Green em função da temperatura. No eixo "x" está representada a temperatura.



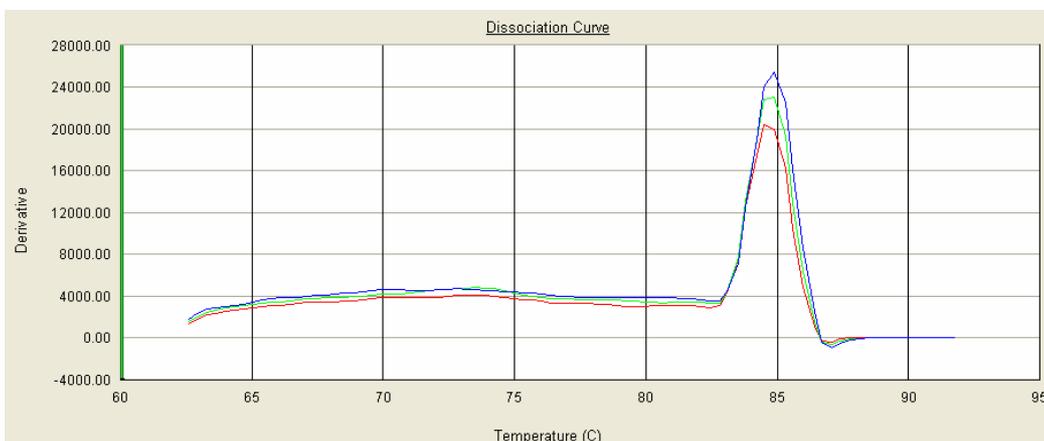
At2g28390



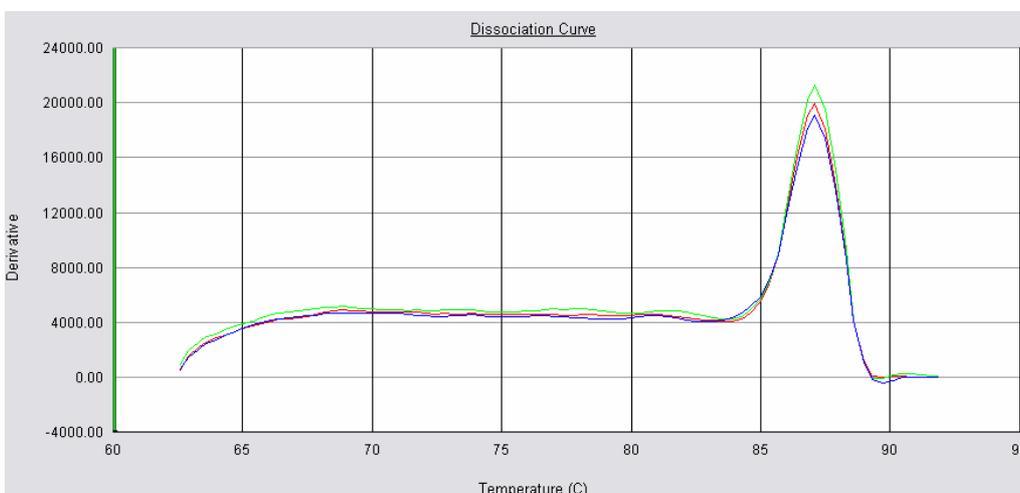
At4g33380



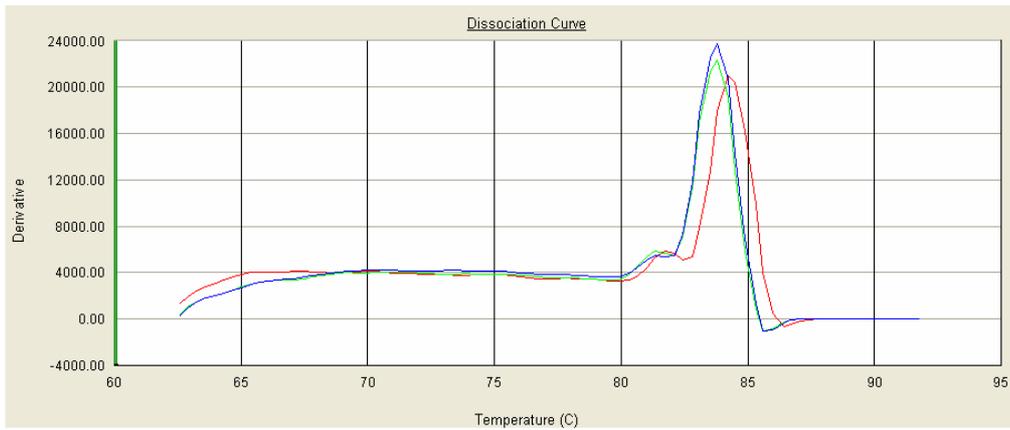
At4g34270



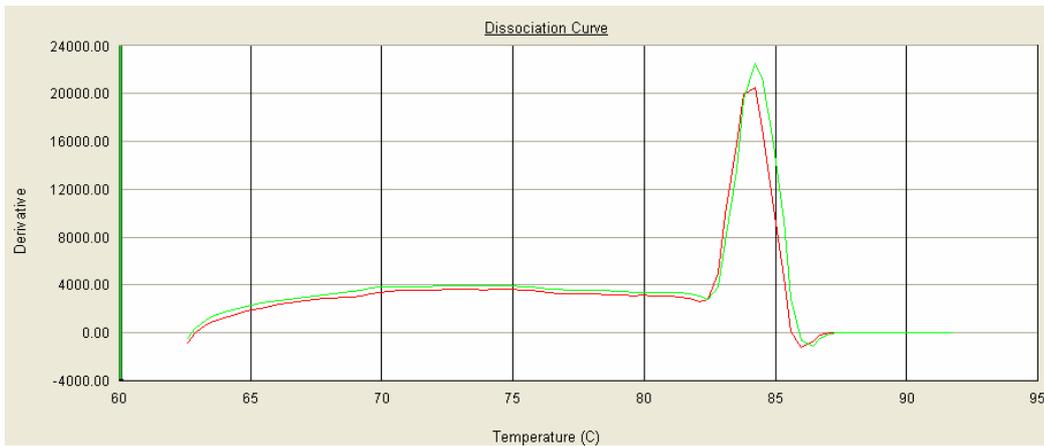
EuC02



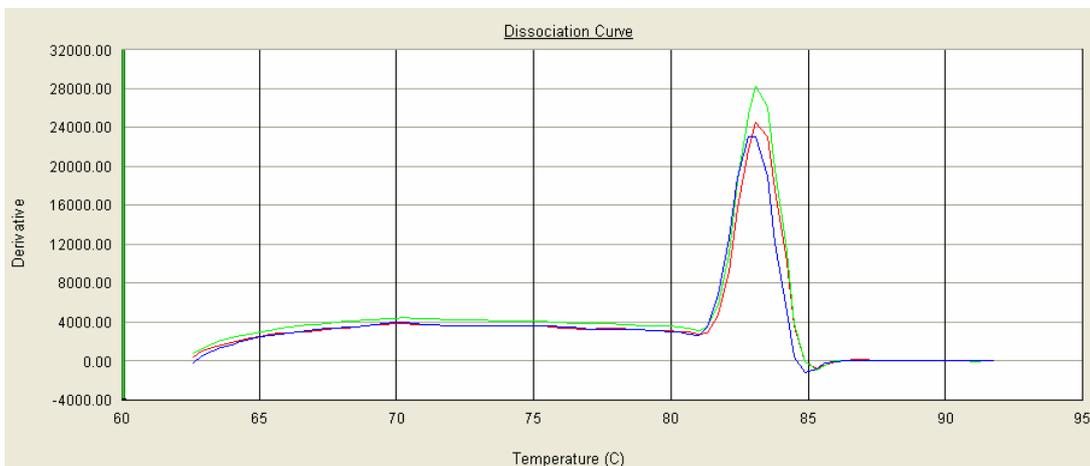
EuC03



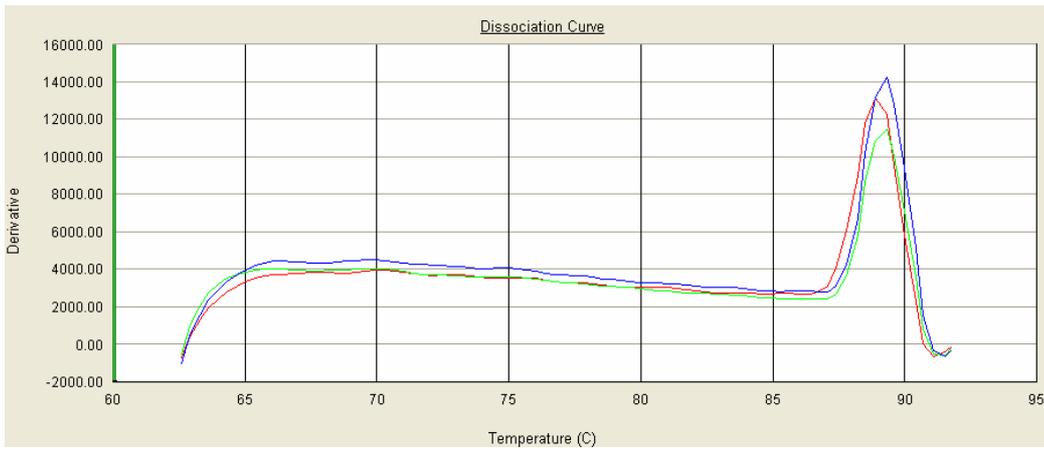
EuC04



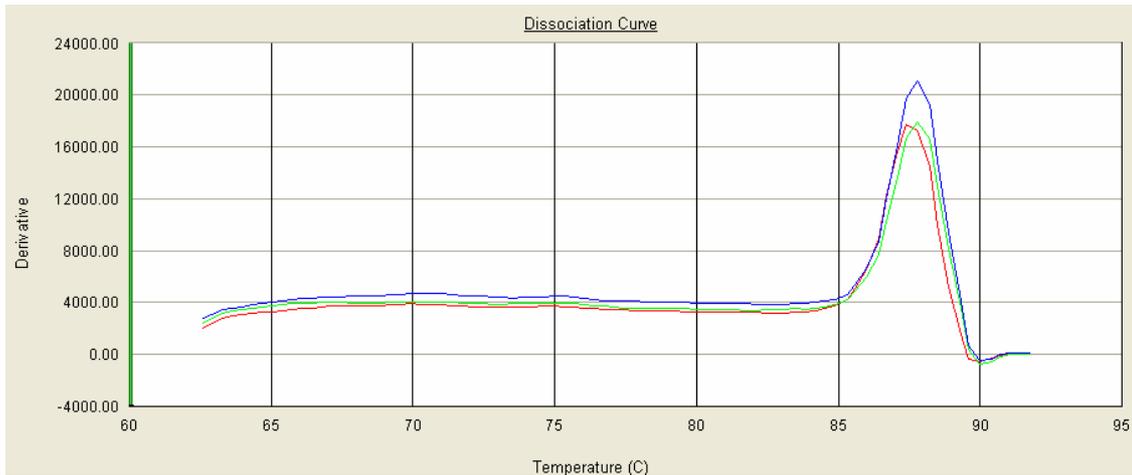
EuC05



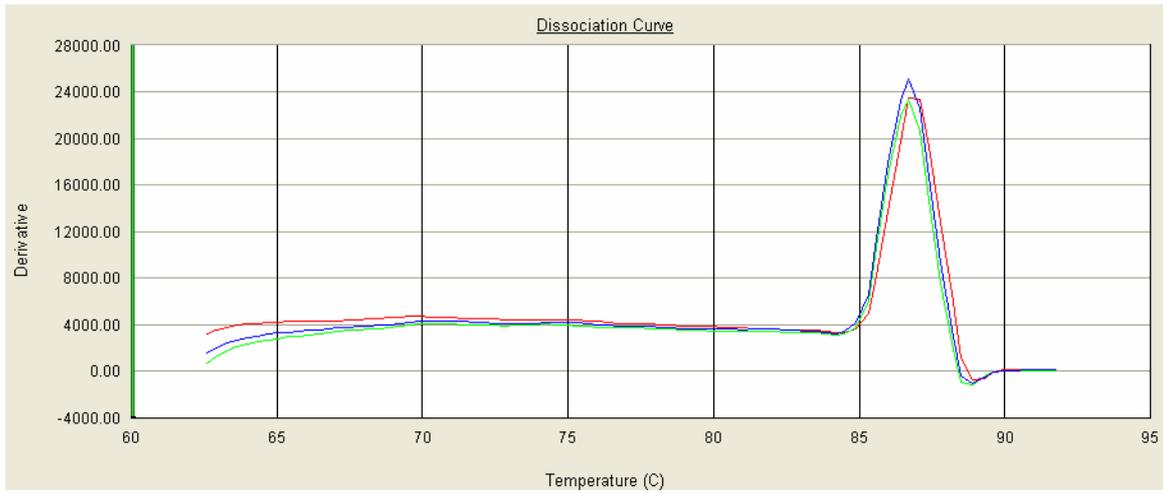
EuC09



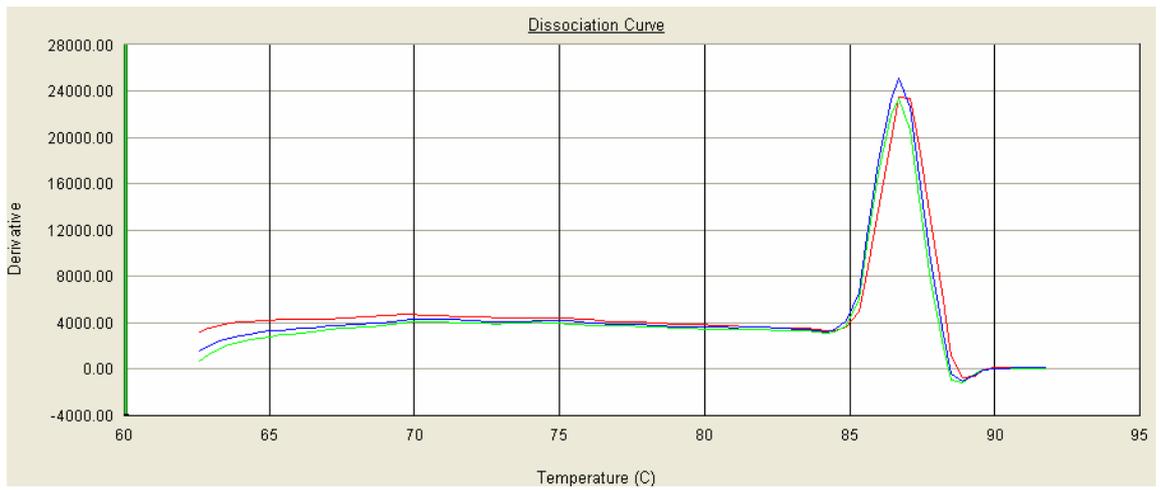
EuC10



EuC12



GAPDH



SAM