

**MAPEAMENTO GENÉTICO DE QTL PARA QUALIDADE DA MADEIRA E
FLORESCIMENTO PRECOCE E, ESTUDOS DE EXPRESSÃO GÊNICA
ALELO ESPECÍFICA EM *Eucalyptus* spp.**

ALEXANDRE ALVES MISSIAGGIA
Engenheiro Florestal

Orientador: Prof. Dr. **DARIO GRATTAPAGLIA**

Tese apresentada à Escola Superior de
Agricultura "Luiz de Queiroz", Universidade
de São Paulo, para obtenção do título de
Doutor em Agronomia, Área de
Concentração: Genética e Melhoramento
de Plantas.

PIRACICABA
Estado de São Paulo - Brasil
Março – 2005

ERRATA

ALEXANDRE ALVES MISSIAGGIA. Mapeamento genético de QTL para qualidade da madeira e florescimento precoce e estudos de expressão gênica alelo específica em *Eucalyptus* spp.

p.	item	linha	onde se lê	leia-se
x	RESUMO	vigésima primeira	...Votorantin Celulose e Papel...	...International Paper...
xiii	SUMMARY	Décima sexta	...Votorantin Celulose e Papel company.	International Paper company.
4	1	décima segunda	...F2 geneticamente não relacionada...	...F2 não endogâmica...
6	2.1	décima primeira	...monóicas...	...hermafroditas...
48	3.2.3	décima primeira, décima segunda e décima terceira	Para cada um de 45 locos microsatélites de <i>Eucalyptus</i> (Brondani et al., 2002). foi resintetizado (Operon Technologies) o primer direto porém adicionando-se uma...	Cada um dos 45 locos microsatélites de <i>Eucalyptus</i> (Brondani et al., 2002) teve o primer direto resintetizado (Operon Technologies), porém adicionando-se uma...
84	4.3.2	Figura 4	Genotipagem dos locos EMBRA 91 e 158...	Genotipagem dos locos EMBRA 91 e 235...
115	5.3.3	sexta	Dos 123 marcadores genotipados, 95 foram mapeados...	Dos 123 marcadores genotipados, 90 foram mapeados...
116	5.3.3	quarta	Deste total, 92 marcadores foram mapeados...	Deste total, 94 marcadores microsatélites foram mapeados...
116	5.3.3	quinta	Para montagem de um mapa genético integrado para os dois parentais foram utilizados 155...	Para montagem de um mapa genético integrado para os dois parentais foram utilizados 146...
117	5.3.3	Figura 5	...parental 221 (direita) e o mapa integrado	...parental 221 (direita) e o mapa integrado. (*) Locos distorcidos. (□) locos que não pertencem ao mapa referência

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

Missiaggia, Alexandre Alves

Mapeamento genético, localização de QTL e estudos de expressão alelo específica em genes de lignificação de eucalipto / Alexandre Alves Missiaggia. - - Piracicaba, 2005. 236 p. : il. + CD-ROM.

Tese (doutorado) - - Escola Superior de Agricultura Luiz de Queiroz, 2005.
Bibliografia.

1. Eucalipto 2. Expressão gênica 3. Genes 4. Hereditariedade 5. Mapeamento genético 6. Marcador molecular 7. Melhoramento genético vegetal I. Título

CDD 634.9734

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

Ao meu pai Sérgio Missiaggia,
À minha mãe Maria Elizabeth Alves Missiaggia,
Ao meu irmão Felipe,
À minha irmã Juliana,
A todos os meus familiares.

DEDICO

AGRADECIMENTOS

Gostaria de agradecer a todas as pessoas que contribuíram para a realização desta tese em especial:

À Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ), particularmente ao Departamento de Genética.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, pela concessão das bolsas de estudo durante o curso.

Ao Prof. Dr. Dario Grattapaglia pelos conselhos, ensinamentos, orientação, críticas, sugestões e apoio durante toda a etapa da realização deste trabalho.

Ao Prof. Dr. Roland Vencovsky pela oportunidade de ter sido seu orientado durante o período em que estive em Piracicaba.

Ao Prof. Dr. Michele Morgante pela confiança e oportunidade de trabalhar em sua equipe no laboratório de Genética Vegetal da Universidade de Udine (Itália).

À Dra. Bobana Radovic, pela amizade e orientação durante o período em que estive desenvolvendo uma parte deste trabalho na Universidade de Udine (Itália).

À empresa ARACRUZ-CELULOSE S.A e, pelo fornecimento de material biológico, e em especial ao Dr. Gabriel Dehon pelas sugestões.

À empresa FERRO GUSA CARAJÁS e ao projeto Genolyptus pelo fornecimento de material biológico.

Ao Prof. Dr. Giancarlo Pasqualli e sua equipe pelas extrações de RNA realizadas.

Ao Prof. Dr. Alexandre Coelho pela ajuda e sugestões na construção dos mapas genéticos e detecção de QTL.

À Nathalia Bueno por ter me ajudado na geração de parte dos dados desta tese.

Aos meu amigos do Laboratório de Genética Vegetal do Cenargen, em especial: Ana Ciampi, Márcio Moretson, Gláucia, Túlio Lins, Isabela Tristan, Rodrigo Tristan e Neide.

Aos meus amigos do Laboratório de Biotecnologia da Universidade Católica de Brasília, em especial: Alessandra Reis, Rosane Colevatti, Georgios Pappas, Lélia, Aiessa e Clarissa.

Aos meus amigos do laboratório de Genética Vegetal da Universidade de Udine: Marco Moroldo, Marco Zantoni, Giusi Zaina, Emanuele de Paulli, Giacomo Prete, Francesco Arzenton, Bobana Radovic e Irena, que me receberam de uma forma muito amigável e me ajudaram durante minha permanência na Itália.

A todos meus amigos de Piracicaba, em especial: Eder Jorge, Tassiano Maxwell, Américo e Adriano de Castro.

Ao Juliano Gomes Pádua pela amizade e pela disposição de me ajudar durante todas minhas idas à Piracicaba.

À Eva Mamani, Marília Pappas e Marco pelas sugestões e revisões desta tese.

À bibliotecária Glória pelas correções de formatação, e ao Berdan pelas impressões da tese.

SUMÁRIO

	Página
RESUMO.....	ix
SUMMARY.....	xii
1 INTRODUÇÃO.....	1
2 REVISÃO DE LITERATURA.....	6
2.1 Histórico e importância do gênero <i>Eucalyptus</i>	6
2.2 Marcadores Microssatélites.....	10
2.2.1 Conservação de locos microssatélites.....	13
2.2.2 Sistemas semi-automatizados de genotipagem baseada em microssatélites.....	15
2.3 Mapeamento genético em Eucalipto.....	17
2.4 Mapeamento de QTL (<i>Quantitative Trait Loci</i>) em <i>Eucalyptus</i>	19
2.5 Florescimento precoce.....	23
2.6 Estratégia de BSA (<i>Bulk Segregant Analysis</i>).....	27
2.7 Genes de lignificação.....	29
2.8 Polimorfismo de Sequência única (<i>Single Nucleotide Polymorphisms</i> – SNP).....	33
2.9 Estudos de Expressão gênica utilizando SNPs.....	35
3 UM MÉTODO PARA GENOTIPAGEM DE MICROSSATÉLITES COM DETECÇÃO FLUORESCENTE EM QUATRO CORES UTILIZANDO MÚLTIPLOS INICIADORES DE PCR COM CAUDA.....	38
Resumo.....	38
Summary.....	40
3.1 Introdução.....	42
3.2 Material e Métodos.....	47
3.2.1 Material genético.....	47
3.2.2 Extração de DNA.....	47
3.2.3 Desenho dos primers.....	48
3.2.4 Amplificação dos locos microssatélites.....	49
3.2.5 Geração dos dados genotípicos.....	50
3.2.6 Análise de segregação dos microssatélites amplificados via primers com cauda.....	50
3.2.7 Montagem do grupo de ligação.....	51
3.3 Resultados.....	52
3.3.1 Amplificação dos locos microssatélites usando a estratégia de múltiplos primers com cauda.....	52
3.3.2 Detecção automatizada dos microssatélites.....	53
3.3.3 Análise de segregação dos microssatélites amplificados via primers com cauda.....	56

3.3.4 Análise de ligação.....	57
3.4 Discussão.....	59
3.5 Conclusão.....	61
4 MAPEAMENTO DO QTL <i>Eef1</i> (<i>Eucalyptus early flowering 1</i>) PARA FLORESCIMENTO PRECOCE EM EUCALIPTO UTILIZANDO “BULK SEGREGANT ANALYSIS” E MARCADORES MICROSSATÉLITES.....	66
Resumo.....	66
Summary.....	69
4.1 Introdução.....	71
4.2 Material e Métodos.....	76
4.2.1 Material genético.....	76
4.2.2 Extração e quantificação de DNA.....	77
4.2.3 Seleção dos indivíduos a serem genotipados utilizando a estratégia de “BSA”.....	78
4.2.4 Amplificação dos locos Microsatélites via PCR.....	78
4.2.5 Geração dos dados genotípicos.....	79
4.2.6 Análise de co-segregação.....	80
4.2.7 Construção do mapa genético.....	80
4.3 Resultados.....	82
4.3.1 Quantificação de DNA e montagem dos Bulks.....	82
4.3.2 Geração dos dados genotípicos.....	83
4.3.3 Análise de co-segregação entre o fenótipo de florescimento precoce e os marcadores microsatélites selecionados.....	85
4.3.4 Construção do mapa genético.....	86
4.3.5 Simulação de seleção assistida por marcadores para o fenótipo de florescimento precoce.....	87
4.4 Discussão.....	89
4.5 Conclusão.....	94
5 MAPEAMENTO DE QTL PARA CARACTERÍSTICAS DA MADEIRA EM CLONES HÍBRIDOS DE <i>Eucalyptus</i> sp. TROPICAL COM BASE EM MICROSSATÉLITES.....	96
Resumo.....	96
Summary.....	99
5.1 Introdução.....	101
5.2 Material e Métodos.....	106
5.2.1 Material genético.....	106
5.2.2 Levantamento dos dados fenotípicos.....	107
5.2.3 Extração e quantificação de DNA.....	109
5.2.4 Amplificação dos locos microsatélites via PCR.....	109
5.2.5 Triagem e seleção de locos microsatélites.....	110
5.2.6 Geração dos dados genotípicos.....	110
5.2.7 Construção do mapa genético.....	111
5.2.8 Detecção dos QTL.....	112
5.3 Resultados.....	114
5.3.1 Levantamento dos dados fenotípicos.....	114
5.3.2 Triagem e seleção dos locos microsatélites.....	115
5.3.3 Geração dos dados genotípicos e construção dos mapas genéticos.....	115

5.3.4 Detecção e mapeamento dos QTL utilizando AMI.....	120
5.3.5 Detecção de QTL utilizando mapeamento de intervalo (MIS e MIC).....	121
5.3.5.1 Detecção de QTL para densidade da madeira.....	121
5.3.5.2 Detecção de QTL para teor de lignina.....	125
5.3.5.3 Detecção de QTL para teor de pentosana.....	125
5.3.5.4 Detecção de QTL para quantidade de fibra/grama.....	130
5.3.5.5 Detecção de QTL para coarseness.....	130
5.3.5.6 Detecção de QTL para rendimento em celulose.....	135
5.3.5.7 Detecção de QTL para consumo específico.....	135
5.3.5.8 Detecção de QTL para crescimento volumétrico (CAP).....	139
5.3.6. Comparação de mapeamento de QTL por AMI e MIC.....	142
5.3.7 Detecção de QTL com base em LOD crítico definido por permutação.....	145
5.4 Discussão.....	147
5.5 Conclusão.....	158
6 REGULAÇÃO ATUANDO EM <i>CIS</i> NA EXPRESSÃO ALELO ESPECÍFICA DE GENES DE LIGNIFICAÇÃO EM <i>EUCALYPTUS</i>	161
Resumo.....	161
Summary.....	164
6.1 Introdução.....	167
6.2 Material e métodos.....	171
6.2.1 Material genético.....	171
6.2.2 Seleção dos genes a serem estudados.....	171
6.2.3 Montagem dos primers.....	174
6.2.4 Extração de DNA.....	174
6.2.5 Quantificação de DNA.....	175
6.2.6 Extração de RNA total.....	175
6.2.7 Extração de mRNA.....	175
6.2.8 Amplificação de cDNA.....	176
6.2.9 Amplificação dos genes estudados via reação em cadeia da Polimerase (PCR).....	176
6.2.10 Purificação da Reação de PCR e Sequenciamento.....	177
6.2.11 Análise das Seqüências e Identificação de SNPs em ESTs.....	179
6.2.12 Detecção dos Níveis de Expressão dos Transcritos com Base em SNPs... ..	179
6.2.13 Curva de normalização de expressão de mRNA usando DNA genômico... ..	182
6.3 Resultados.....	183
6.3.1 Desenho dos primers para amplificação dos genes.....	183
6.3.2 Análise das seqüências e localização de SNPs em regiões transcritas.....	186
6.3.3 Desenho dos primers para a detecção de SNPs via SBE.....	187
6.3.4 Curva de normalização.....	188
6.3.5 Análise de expressão gênica via “ <i>Single Base Extension</i> ” para o gene Cinnamyl Alcohol Dehydrogenase (CAD).....	191
6.3.6 Análise de expressão gênica via “ <i>Single Base Extension</i> ” para o gene Cinnamoyl CoA Redutase (CCR).....	192
6.4. Discussão.....	194
6.5 Conclusões.....	199
7 CONCLUSÕES GERAIS.....	203
ANEXOS.....	210
REFERÊNCIAS BIBLIOGRÁFICAS.....	219

**MAPEAMENTO GENÉTICO DE QTL PARA QUALIDADE DA MADEIRA E
FLORESCIMENTO PRECOCE E, ESTUDOS DE EXPRESSÃO GÊNICA
ALELO ESPECÍFICA EM *Eucalyptus* spp.**

Autor: ALEXANDRE ALVES MISSIAGGIA

Orientador: Prof. DARIO GRATTAPAGLIA

RESUMO

Plantações de eucalipto desempenham um papel de grande destaque na silvicultura intensiva no Brasil e no mundo. Atualmente, estima-se que florestas plantadas de eucalipto ocupem cerca de 3 milhões de hectares no Brasil, sendo 50 % desta produção destinada à produção de celulose. Empresas florestais têm buscado, no melhoramento genético, meios para aumentar a produtividade de seus produtos com o objetivo de expandir a produção sem a necessidade de aumento na extensão de suas plantações. Nos últimos anos tem havido um interesse crescente por parte das empresas florestais quanto à utilização de ferramentas moleculares visando gerar informações que possam ser agregadas aos programas de melhoramento.

Neste trabalho, foram realizados quatro estudos apresentados na forma de artigos. O primeiro trabalho traz a otimização de um método para genotipagem de microssatélites, com detecção fluorescente em quatro cores, utilizando múltiplos iniciadores de PCR (reação em cadeia da polimerase) com cauda, o que gera uma economia em torno de 60 % para rotinas de genotipagem automatizada. O segundo trabalho, realizado em parceria com a empresa FERRO GUSA CARAJÁS, apresenta a detecção e mapeamento de um gene que controla a característica de florescimento precoce denominado *Eef1* (*Eucalyptus early flowering 1*) utilizando marcadores microssatélites com detecção fluorescente em conjunto com a metodologia de BSA (Bulk segregant analysis). Os resultados mostraram que esta técnica pode ser utilizada com sucesso na detecção de genes de herança qualitativa. O terceiro trabalho, realizado juntamente com a empresa ARACRUZ CELULOSE S.A., apresenta a construção de um mapa genético utilizando marcadores microssatélites, marcados com fluorescência, e detecção de QTL para características físico-químicas da madeira para dois híbridos naturais. Os mapas genéticos apresentaram tamanho médio de 70 a 80% do genoma. Nestes mapas foram alocados 15 QTL associados às características de rendimento de celulose, consumo específico *coarseness*, teor de lignina e pentosanas, CAP, densidade básica e fibra/grama. O quarto e último trabalho, realizado na Universidade de Udine (Itália) em parceria com a empresa Votorantin Celulose e Papel, apresenta um estudo de expressão gênica alelo específica de genes envolvidos na via de lignificação utilizando uma metodologia extremamente recente de

extensão de base individual (Single Base Extension-SBE). Foi possível detectar variações gênicas alelo específica para os genes CAD e CCR. Para o gene CCR, os níveis de expressão dos alelos segregou na razão de 1:1, sugerindo um possível efeito de ação combinada de fator em *trans* e fator em *cis* ("*Cis + Trans acting*") afetando a expressão dos alelos para este gene. Estes resultados mostram a possibilidade da utilização destas novas ferramentas e abordagens biotecnológicas em programas tradicionais de genética e melhoramento clássico de eucalipto.

**GENETIC MAPPING OF QTL FOR WOOD PROPERTIES AND EARLY
FLOWERING AND ALLELE SPECIFIC GENE EXPRESSION STUDIES IN
Eucalyptus spp**

Author: ALEXANDRE ALVES MISSIAGGIA

Adviser: Prof. DARIO GRATTAPAGLIA

SUMMARY

Eucalyptus plantations have an important role in the intensive forestry in Brazil and in the world. It is considered that eucalyptus planted forests occupy about 3 million hectares in Brazil, which of them 50% are destined to cellulose production. To increase the productivity of their products, forest companies have been using genetic breeding, in order to improve the production without the necessity of increase the extension of their plantations. In the last years forest companies have been having an increasing in the interest in the use of molecular tools to generate information to be employed in genetic breeding programs. In this work, four studies were accomplished, and were presented in chapters. The first work presents optimization of a microsatellite genotyping method, using four fluorescent color detection and multiple PCR (polymerase

chain reaction) primers tail. This approach generates a cost reduction around 60% for automated genotyping routines. The second work, accomplished with the FERRO GUSA CARAJÁS company, presents the detection and mapping of a gene that controls the early flowering characteristic, named *Eef1* (*Eucalyptus* early flowering 1). It was used microsatellite markers with fluorescent detection associated to BSA (Bulk segregant analysis) methodology. The results showed that this technique can be successfully used to detect the qualitative inheritance genes. The third work was done in association with the ARACRUZ CELULOSE S.A company. It aimed at the construction of a genetic map using microsatellites markers fluorescently labeled. QTL related to physiochemical wood characteristics were mapped for two natural hybrid trees. The genetic maps covered around 70 to 80% of the genome. In these maps 15 QTL associated to the characteristics of cellulose income, specific consumption, coarseness, lignin and *pentosanas* tenor, *CAP*, basic density and fiber/gram were allocated. The fourth and last work, was carried out at University of Udine (Italy), in partnership with the Votorantin Celulose e Papel company. It presented a study of allelic gene expression involved in the lignification pathway using an extremely recent methodology of single base extension (SBE). It was possible to detect allelic variation for CAD and CCR genes. For the CCR gene, the levels of expression of the alleles segregated at the ratio 1:1, suggesting a possible effect of cis + trans acting factor affecting the expression of the alleles for this gene. These results suggest the possibility of the use of these new tools and biotechnological approaches in traditional programs of eucalyptus breeding.

1 INTRODUÇÃO

O gênero *Eucalyptus*, incluindo mais de 600 espécies, a maioria originária da Austrália, desempenha um papel de grande destaque na silvicultura intensiva no Brasil e no mundo. Devido a sua alta adaptabilidade, variabilidade, e rápido crescimento, associadas à aplicabilidade de sua madeira para diversos fins, as florestas plantadas de Eucalipto ocupam cerca de 18 milhões de hectares no mundo (FAO, 2002) com a Índia como maior plantador com 8 milhões de hectares em sua maioria de florestas de baixa produtividade, e o Brasil com 3 milhões de hectares de florestas de altíssima produtividade na faixa de 45 a 60 m³/ha/ano (Mora & Garcia, 2000).

Estudos envolvendo construção de mapas genéticos e detecção de QTL em eucalipto tem demonstrado a possibilidade de encontrar regiões genômicas controlando características de herança complexa e tem gerado as primeiras informações quanto ao número e magnitude de efeitos dos locos controlando características quantitativas na maioria das espécies de *Eucalyptus* utilizadas pelas empresas florestais.

Apesar dos avanços conceituais, estes estudos enfrentaram limitações técnicas que dificultam hoje a transferência das informações geradas

para situações operacionais. Algumas destas limitações envolvem o número limitado de descendentes analisados, a precisão na coleta dos dados fenotípicos e a herança dominante e baixa transferibilidade dos marcadores utilizados para gerar os mapas genéticos, tipicamente marcadores RAPD e AFLP.

A utilização de marcadores moleculares co-dominantes e multialélicos, geneticamente mais informativos, que permitam detectar, comparar e validar a posição e efeito de QTL entre diferentes populações e espécies é hoje um dos aspectos técnicos fundamentais para avançar rumo a um melhor entendimento da arquitetura genética de características quantitativas em *Eucalyptus* e possivelmente para a realização de experimentos integrados com estratégias de mapeamento de associação e identificação de genes responsáveis por QTL. Atualmente, dentre as classes de marcadores moleculares disponíveis, os microssatélites, são aqueles que possuem o maior conteúdo informativo para experimentos de mapeamento de ligação e que podem, portanto suprir esta necessidade.

Marcadores microssatélites para *Eucalyptus* foram desenvolvidos nos últimos anos e utilizados para a construção de um mapa genético referência, composto exclusivamente de marcadores microssatélites cobrindo mais de 95% do genoma de *Eucalyptus* Brondani et al. (1998; 2000; 2002). O fato de marcadores microssatélites serem codominantes, multi-alélicos, baseados em PCR e transportáveis entre diferentes pedigrees representou um avanço

importante para as pesquisas envolvendo genômica de *Eucalyptus*. Informações geradas por esses marcadores permitem a comparação e troca de informações entre diferentes grupos de pesquisa que trabalham com *Eucalyptus*, principalmente no que diz respeito a geração de mapas genéticos, detecção de QTL e procedimentos de seleção assistida por marcadores (Grattapaglia, 2000).

O principal interesse das empresas florestais que plantam eucalipto é compreender as bases genéticas responsáveis pelo controle de características físico químicas envolvidas no processo de formação da madeira. Embora QTL para algumas características importantes já tenham sido mapeados em *Eucalyptus* (Grattapaglia et al., 1995; Grattapaglia et al., 1996; Verhaegen et al., 1997; Marques et al., 1999; Junghans et al., 2003; Kirst et al., 2004; Kirst et al., 2005) as características relacionadas às propriedades físicas e químicas da madeira somente recentemente foram contempladas em estudos de mapeamento genético (Thamarus et al., 2004; Myburg, 2001). Estas características são as de maior impacto no processo industrial e ao mesmo tempo representam o maior desafio para o melhoramento genético de *Eucalyptus* pela sua expressão tardia constituindo portanto, uma excelente oportunidade de desenvolvimento de procedimentos de seleção precoce assistida por marcadores moleculares.

O presente trabalho apresenta resultados de mapeamento genético de QTL para diversas características de importância econômica em *Eucalyptus*,

com padrões contrastantes de herança, com base em marcadores microssatélites. O projeto envolve uma série de experimentos que vão desde a otimização de métodos de alto desempenho para a detecção fluorescente de microssatélites, a aplicação de métodos de genotipagem seletiva de extremos fenotípicos até a análise completa de duas famílias segregantes. O primeiro experimento de mapeamento envolve a localização de um QTL de grande efeito para florescimento precoce denominado *Eef1* (*Eucalyptus* early flowering 1) utilizando a estratégia de Bulk Segregant Analysis em uma família segregante derivada do cruzamento de uma planta de *Eucalyptus grandis* mutante para florescimento precoce. O segundo experimento descreve o mapeamento de diversos QTL para oito características físico-químicas da madeira e crescimento em uma família F2 geneticamente não relacionada derivada do intercruzamento de duas árvores elite híbridas. Neste experimento é realizada ainda uma análise comparativa de localização de QTL e genes candidatos com trabalhos de mapeamento independentes desenvolvidos com outras espécies de *Eucalyptus*.

Fechando a tese, o último capítulo descreve um estudo de expressão de genes envolvidos na via de lignificação de eucalipto. Este trabalho visou avaliar a possibilidade de investigar os níveis de expressão alelo específica de genes selecionados envolvidos na via de lignificação de eucalipto tendo em vista o seu conhecido papel chave na regulação das quantidades e propriedades da lignina produzida. Para o desenvolvimento deste estudo foi utilizado um cruzamento entre *E. grandis* e *E. urophylla* e uma progênie de 40

indivíduos. Este trabalho, realizado na Universidade de Udine, é inovador em espécies arbóreas e demonstrou um grande potencial de aplicação em estudos envolvendo o entendimento do papel da regulação gênica na definição de fenótipos de importância econômica no Eucalipto.

2 REVISÃO DE LITERATURA

2.1 Histórico e importância do gênero *Eucalyptus*

O gênero *Eucalyptus*, nativo da Austrália e ilhas adjacentes, pertence à família *Myrtaceae*, contando hoje com mais de 600 espécies descritas, além de um grande número de variedades e alguns híbridos, sendo a maioria destes descrita no trabalho de S. T. Blake em 1934 (Andrade, 1961; Boland et al., 1994; Lima, 1993). O gênero é originário da Austrália, com exceção de *E. urophylla* e *E. deglupta*, que são naturais do Timor e Papua Nova Guiné, respectivamente (Pryor, 1985). Dentre as principais espécies plantadas para fins comerciais pode-se destacar *E. grandis*, *E. urophylla*, *E. globulus*, *E. camaudulensis*, *E. saligna* e *E. terenticornis* que pertencem ao sub gênero *Symphyomyrtus* (Eldridge, 1994).

Árvores de eucalipto são monóicas e protândricas, apresentando fecundação preferencialmente alógama, mas podendo sofrer autofecundação. Suas flores são agrupadas em corimbos, panículas ou umbelas. Os frutos são capsulares e deiscentes, liberando sementes com tamanho variando de 1mm até 2cm (Cavalcanti et al., 1963). O genoma de *Eucalyptus* tem entre 500 e 650 Mpbs (Grattapaglia et al., 1994a), distribuídos em 22 cromossomos (n=11).

A primeira descrição do gênero *Eucalyptus* ocorreu em 1788 por L'Héritier de Brutelle, sendo que as primeiras mudas e sementes introduzidas na Europa foram levadas por Antônio Guichenot, que participou de uma expedição à Austrália entre 1801 e 1804 (Andrade, 1961). Acredita-se que a introdução deste gênero no Brasil tenha sido feita pelo Sr. Frederico de Albuquerque, no Rio Grande do Sul, em 1868 (Andrade, 1961). Entretanto, o primeiro cientista brasileiro a se interessar em estudar e cultivar plantas de *Eucalyptus* foi Edmundo Navarro de Andrade, que trabalhou na Companhia Paulista de Estradas de Ferro. Seu objetivo principal era utilizar árvores plantadas para alimentar as caldeiras das locomotivas bem como para a produção de moirões e postes. Navarro de Andrade introduziu quase todas as espécies de *Eucalyptus* existentes no Horto de Rio Claro, constituindo desta forma um dos mais completos bancos de germoplasma do país. Graças a este fato e a diversas introduções de germoplasma nas décadas de 70 e 80, o Brasil tornou-se, posteriormente, um grande produtor de papel e celulose de fibra curta de Eucalipto, exportando mais da metade de sua produção. A alta produtividade deve-se em parte ao ótimo clima tropical, condições de solo favoráveis e ausência da maioria dos inimigos naturais. Entretanto o que efetivamente causou aumentos de mais de 300% na produtividade média do país entre o início do século passado e a atualidade foram os ganhos genéticos dos diversos programas de melhoramento existentes seja intra como inter-específicos aliados ao desenvolvimento da clonagem e práticas silviculturais avançadas fornecendo material superior e homogêneo para plantio.

Estima-se que o Brasil possua hoje cerca de 6 milhões de hectares em área reflorestada, sendo 61% ocupada com *Eucalyptus*, 34% com *Pinus* e 5% com outros gêneros (Garcia et al., 1992), Florestas industriais de *Eucalyptus* estão amplamente espalhadas pelo território nacional, principalmente, nos estados de Minas Gerais e São Paulo (SBS, 2002). Atualmente, a cultura de eucalipto gera milhares de empregos, divisas e impostos. O setor florestal brasileiro, principalmente aquele ligado à produção de *Eucalyptus* e seus derivados, tem elevada importância econômica e social, gerando cerca de 500 mil empregos diretos e 2 milhões indiretos tendo sido sendo responsável em 1998, por US\$ 13 bilhões (na época cerca de 2,5% do Produto Interno Bruto brasileiro).

Os principais produtos obtidos a partir da madeira do *Eucalyptus* são o papel e celulose e o carvão vegetal para a indústria siderúrgica, além de outros derivados da madeira, para móveis e construção, papelão, óleos e outros (Mora et al., 2000).

A maior parte dos 3 milhões de hectares cobertos por *Eucalyptus* pertence a empresas produtoras de papel e celulose e siderúrgicas, cultivada, em sua maioria, com clones ou híbridos interespecíficos superiores provenientes de programas de melhoramento genético. Graças a estes programas associados ao manejo florestal, a produtividade do eucalipto foi incrementada, a área de plantação foi triplicada e a qualidade da madeira foi melhorada de forma a atender à demanda das indústrias de papel e de celulose.

Entretanto, a cultura do *Eucalyptus* ainda não supre completamente esta demanda, tanto em qualidade de madeira como em produção, apesar do Brasil possuir alta competitividade. A produtividade e as áreas plantadas anualmente são consideradas insuficientes para suprir a demanda, ameaçando as florestas nativas brasileiras (Scharf, 2003). A melhor estratégia para aumentar a produtividade e qualidade das florestas de *Eucalyptus* envolve principalmente o melhoramento genético. Nesse sentido, novas metodologias e estratégias têm sido utilizadas com o intuito de elevar ainda mais a produção de *Eucalyptus* no Brasil, principalmente envolvendo produtividade, resistência a pragas e doenças e qualidade da madeira. Na área de aplicação das tecnologias de biologia molecular no melhoramento de *Eucalyptus*, diversos trabalhos vêm sendo realizados nos últimos 10 anos, principalmente focando em estudos de variabilidade e distância genética, análises de identidade, certificação e proteção de clones e em escala experimental de seleção assistida por marcadores moleculares. Alguns mapas genéticos já foram gerados e QTL localizados para algumas características.

Atualmente, novos avanços na área de ciências genômicas associados a redução dos custos tecnologias de sequenciamento de DNA, tem favorecido a execução de diversos programas de sequenciamento de genomas para diferentes organismos. Para o eucalipto existem dois projetos brasileiros: o Projeto Forests, que envolveu o seqüenciamento e mineração de 120.000 ESTs (FAPESP, 2003), e o Projeto Genolyptus, que além dos componentes

genômicos de sequenciamento e mapeamento genético e físico, conta com uma ampla rede experimental dedicada especificamente à pesquisa genômica florestal incluindo diversos cruzamentos envolvendo as mais importantes espécies comerciais de eucalipto, bem como híbridos plantados em diversos ambientes. Produtos deste experimento de campo estão servindo de base para pesquisas envolvendo sequenciamento de genes expressos, construção de mapas genéticos, mapeamento de locos controladores de características quantitativas, mapeamento físico, estudos de expressão gênica e desenvolvimento de ferramentas de bioinformática e genética estatística para integrar e disponibilizar toda a informação visando a seleção de árvores superiores (Grattapaglia, 2003).

2.2 Marcadores Microsatélites

Os microsatélites, também conhecidos como SSR (Simple Sequence Repeats), são marcadores moleculares baseados na amplificação via PCR de seqüências repetitivas (1 a 6 pb), abundantemente encontradas no genoma de eucariotos (Hamada et al., 1982).

A alta taxa de mutações encontradas nestas regiões ocorre devido à alta incidência de “escorregamento” ou pareamento incorreto de motivos microsatélites durante a replicação, como foi demonstrado *in vitro* por Schlötterer et al. (1992). Durante a replicação de uma região repetitiva, as fitas de DNA se separam e se reassociam de forma incorreta, resultando em cópias

de trechos de DNA com diferentes tamanhos, gerando locos multialélicos devido ao número variável de repetições. Este polimorfismo gerado, pode ser analisado via PCR utilizando um par de iniciadores específicos que flanqueiam região de DNA repetitivo e definem o loco marcador (Weber et al., 1989).

Devido à abundância e distribuição uniforme no genoma de eucariotos, associadas ao multialelismo e alto nível de polimorfismo com padrão de segregação co-dominante, os microssatélites têm tido ampla aplicação em casos onde há necessidade de discriminação de indivíduos e identificação de parentesco. Estes marcadores têm sido empregados em estudos de genética de populações (Wenburger et al., 1996), na genotipagem de indivíduos para discriminação de genótipos (Cregan et al., 1994), estudos de parentesco (Marklund et al., 1994), e proteção de variedades de plantas (Diwan et al., 1997).

Em *Eucalyptus*, marcadores baseados em microssatélites foram desenvolvidos a partir de bibliotecas genômicas enriquecidas (Brondani et al., 1998). O procedimento utilizado para o seu isolamento não permitiu estimar sua abundância no genoma, bem como frequência de ocorrência de tri e tetra nucleotídeos. Lourenço et al. (2004) a partir de sequenciamento por amostragem de uma biblioteca para sequenciamento por fragmentação randômica de DNA e comparação com bibliotecas de cDNA, identificaram uma ampla bateria de marcadores microssatélites di tri e tetranucleotídeos tanto em regiões não codificadoras como em regiões de ESTs (Expressed Sequence

Tags). Foi verificado que microsatélites com repetições de dinucleotídeos são mais freqüentes que repetições de tri e tetranucleotídeos quando identificados em seqüências de DNA genômico. Porém, quando identificados em seqüências de ESTs verificou-se que microsatélites com repetições de trinucleotídeos são mais freqüentes que microsatélites com repetição de di e tetranucleotídeos, corroborando resultados mostrados por Morgante et al. (2002).

Em um trabalho realizado por Wang et al. (1994) os motivos (TA) n foram descritos como sendo as unidades de repetição mais comum em plantas, porém Lourenço et al. 2004 mostraram que pelo menos para *Eucalyptus* o motivo de (TA) n é quase tão numeroso quanto (TC) n .

Também foi verificado por Lourenço et al. 2004 que 0,75% dos nucleotídeos seqüenciados pela estratégia de sequenciamento randômico estavam contidos em seqüências microsatélites. Com base em um genoma de 640 Mpb e que o tamanho médio de seqüências microsatélites seja de 38 pb, foi estimado que em *Eucalyptus* seriam encontrados mais de 126 mil seqüências microsatélites, somando-se aí seqüências perfeitas e imperfeitas. Isto equivale a 1 seqüência microsatélite a cada 5 kb. Esta proporção de microsatélites observada para o genoma *Eucalyptus* está de acordo com as estimativas de Morgante et al. (2002), onde genomas menores possuem maior porcentagem de seu genoma composta por regiões microsatélites do que genomas maiores, em plantas. Isso pode ser claramente observado quando comparamos com a porcentagem de seqüências microsatélites em

Arabidopsis (0,85% para um genoma de 150Mbp) e para *Z. Mays* (0,37% para um genoma com cerca de 3.000 Mbp) (Toth et al., 2000).

Outro resultado interessante, ainda no trabalho de Lourenço et al. (2004) foi a de que microssatélites pelo menos em *Eucalyptus* parecem ter associação preferencial com regiões codificantes. O número absoluto de microssatélites (1539) identificados em ESTs (4451 seqüências) é muito superior ao encontrado (986 microssatélites) na biblioteca por “seqüenciamento por fragmentação randômica de DNA (shotgun)” de DNA genômico (7395 seqüências). Estes dados corroboram mais uma vez os resultados de Morgante et al. (2002) que propôs como modelo geral para plantas, a ocorrência de microssatélites associados preferencialmente com regiões codificantes.

2.2.1 Conservação de locos microssatélites

O uso amplo de marcadores baseados em microssatélites em análise genômica em plantas ainda é limitado a algumas espécies pelo alto custo e tempo necessário para o desenvolvimento dos iniciadores espécie específicos, que flanqueiam as regiões de DNA repetitivo (Rafalski et al., 1993). Entretanto, a transferibilidade destes marcadores entre espécies pertencentes ao mesmo gênero e até às vezes entre gêneros da mesma família é possível em muitos casos. Essa possibilidade decorre do fato que, embora as regiões microssatélites estejam sujeitas a alta taxa de mutação, as regiões

flanqueadoras complementares aos iniciadores são freqüentemente conservadas em espécies ou gêneros próximos.

Um estudo realizado por White et al. (1997) demonstrou que de um conjunto de 11 locos microssatélites desenvolvidos a partir de DNA genômico de *Swietenia humilis* (Meliaceae), sete amplificaram com sucesso em 11 espécies da mesma família, incluindo *Swietenia macrophylla* e *Cedrela odorata*. Um outro estudo realizado por Collevatti et al. 1999, avaliando um conjunto de 10 locos desenvolvidos a partir de DNA genômico de *Caryocar brasiliense*, observaram 100% de transferibilidade para as espécies *C. coriaseum*, *C. edule*, *C. glabrum*, *C pallidum* e *C. villosum*. Em *Eucalyptus*, microssatélites mostram elevada conservação, entre 80 e 100% entre espécies de um mesmo subgênero, e menor, entre 20 e 60% entre espécies de subgêneros distintos (Brondani et al., 1998).

O desenvolvimento de marcadores microssatélites deve levar em consideração a maximização de sua utilização, ou seja, quanto maior a conservação de locos microssatélite menor será seu custo relativo. Em *Eucalyptus*, locos microssatélites conservados entre espécies, pelo menos entre as mais plantadas, são essenciais, já que programas de melhoramento normalmente lançam mão de várias espécies e híbridos (Byrne et al., 1996). Vale ressaltar que nem todos os locos microssatélites identificados são aproveitáveis para auxiliar programas de melhoramento genético, assim como construção de mapas genéticos. A seleção dos locos informativos é baseada

nas características destes locos e nos resultados que estes microssatélites permitirão obter na identificação de segregação nas progênies resultantes de cruzamentos controlados. Dentre os principais parâmetros utilizados nesta seleção além da robustez de amplificação de detecção estão: o tipo de repetição (di-, tri-, tetra- ou pentanucleotídicos), o número de repetições, a existência ou não de duplicação do loco no genoma, o número e frequência relativa de alelos existentes ao loco determinando a heterozigosidade esperada (Brondani et al., 1998; Edwards et al., 1998; Garner, 2002).

2.2.2 Sistemas semi-automatizados de genotipagem baseada em microssatélites

O uso de marcadores moleculares baseados em microssatélites para fins de genotipagem de indivíduos evoluiu com o desenvolvimento de sistemas de análise semi-automatizada, onde se tornou possível a amplificação e/ou análise eletroforética de vários locos simultaneamente (sistemas multiplex). Estes sistemas apresentam inúmeras vantagens sobre as técnicas convencionais, como rapidez e precisão na geração de dados, análise simultânea de vários locos, minimização da manipulação manual da informação e alto potencial de transferibilidade entre laboratórios.

Desenvolvidos inicialmente para estudos de genética humana (Edwards et al., 1991), estes sistemas baseiam-se na amplificação, via PCR, de um ou mais locos microssatélites, utilizando iniciadores marcados com

fluorocromos. Os fragmentos amplificados são posteriormente separados e visualizados em equipamentos de alta resolução que permitem a detecção de fluorescência e portanto dos alelos, em tempo real.

Recentemente, sistemas multiplex foram descritos para espécies de plantas. Thomas et al. (1994) desenvolveram um sistema para *Vitis sp.*, no qual sete locos microssatélites foram amplificados separadamente via PCR utilizando iniciadores marcados com diferentes fluorocromos e combinados anteriormente à análise em seqüenciador de DNA. O sistema foi utilizado para discriminar 80 cultivares de videira, muitas indistinguíveis em nível de fenótipo, e criar um sistema universal de diferenciação de genótipos do gênero *Vitis*. Diwan & Cregan (1997) propuseram um sistema semelhante, no qual 20 microssatélites, amplificados por meio de iniciadores marcados com fluorescência, eram amplificados via PCR. Os produtos eram posteriormente combinados com produtos de amplificação de microssatélites marcados com diferentes fluorescências e analisados em seqüenciador de DNA, permitindo a discriminação entre cultivares de soja indistinguíveis com base em características morfológicas e 17 locos RFLP.

Em eucalipto foram desenvolvidos três sistemas de genotipagem multiloco (sistemas multiplex) semi-automatizados com detecção fluorescente baseados em 12 marcadores microssatélites que foram utilizados para genotipar indivíduos das espécies *E. camaldulensis*, *E. dunnii*, *E. globulus*, *E. grandis*, *E. saligna* e *E. urophylla*. Os parâmetros de informação genética da

maioria dos locos indicaram que eles são apropriados para genotipagens de indivíduos nessas espécies. Esses sistemas multiplex foram utilizados na avaliação de uma população de melhoramento de *E. urophylla* e uma coleção de clones elite, onde foi possível a detecção de erros de identificação do material genético, além de uma avaliação de diversidade genética e similaridade entre clones (Kirst et al., 1999).

2.3 Mapeamento genético em Eucalipto

A maioria dos trabalhos envolvendo mapeamento genético em plantas como milho é realizada através da análise de meioses geradas a partir de cruzamentos entre linhagens puras, ou seja, indivíduos homocigotos para seus alelos. Em árvores em geral preferencialmente alógamas, com histórico muito recente de domesticação, linhagens não são disponíveis. Tendo em vista o longo tempo necessário para a geração de descendências sob medida para experimentos de mapeamento, têm-se utilizado pedigrees imediatamente disponíveis em programas de melhoramento. Estes pedigrees envolvem muitas vezes somente os dois genitores heterocigotos e sua descendência F1. Nestes casos, a análise de co-segregação de marcadores moleculares é mais complexa do que em situações clássicas de espécies nas quais linhagens são disponíveis como milho, soja ou arroz. Por se tratar de indivíduos heterocigotos, os marcadores segregam com diferentes razões de segregação, desde 1:1 até 1:1:1:1 a depender da constituição alélica dos genitores.

A construção de mapas genéticos em espécies alógamas foi originalmente realizada utilizando marcadores dominantes RAPD combinados com a estratégia de pseudocruzamento (Grattapaglia et al., 1994). O primeiro mapa genético construído para eucalipto foi realizado utilizando 558 marcas geradas por marcadores RAPD (*Random Amplification of Polimorphism DNA*) e envolveu um cruzamento entre *E. grandis* e *E. urophylla* (Grattapaglia et al., 1994). Posteriormente, outros mapas foram gerados utilizando esta mesma estratégia com marcadores RAPD em *E. grandis* e *E. urophylla* (Verhaegen & Plomion, 1996) e *E. tereticornis* e *E. urophylla* (Gan et al., 2003), e com marcadores AFLP em um cruzamento de *E. grandis* e *E. globulus* (Myburg et al., 2003) em um cruzamento de *E. tereticornis* e *E. globulus* (Marques et al., 1997). Mapas integrando diferentes classes de marcadores foram gerados por Byrne et al. (1995) com marcadores RAPD, RFLPs e isoenzimas em *E. nitens*. Por Brondani et al. (2002) com marcadores RAPD e 70 microssatélites em *E. grandis* e *E. urophylla* e finalmente RLFP e microssatélites para *E. globulus* (Moran et al., 2002). Mapas apresentando o mapeamento de genes que participam da via de lignificação em associação com marcadores moleculares também já foram publicados, envolvendo *E. grandis* e *E. urophylla* (Gion et al., 2000) e *E. globulus* (Thamarus et al., 2002; Moran et al., 2002).

O comprimento médio total destes mapas variou entre 1300 a 1400 centiMorgans (cM), com uma media de aproximadamente 120 cM por grupo de ligação com exceção do mapa de AFLP de Marques et al. (1997) que

apresentou um comprimento total abaixo de 1000 cM. Todos esses mapas gerados, com exceção do mapa construído por Thamarus et al. (2002) foram apresentados individualmente para cada parental, e construídos com o programa Mapmaker (Lander et al., 1987). Já Thamarus et al. (2002) optou por construir um mapa integrado para os parentais do cruzamento utilizando o programa JoinMap (Stam, 1993; Stam & Van Ooijen, 1995).

2.4 Mapeamento de QTL (*Quantitative Trait Loci*) em *Eucalyptus*

A tecnologia de marcadores moleculares tem permitido aos geneticistas estudarem variações de seqüências de DNA para diversos genomas de plantas. Cada marcador molecular representa uma pequena seqüência de DNA a qual, uma vez apresentando ligação gênica com um determinado fenótipo ou gene, permite o acompanhamento da herança deste fenótipo ou gene de uma geração para outra (Myburg, 2002). Em espécies arbóreas tem sido demonstrado que coberturas completas do genoma podem ser obtidas quando um grande número de marcadores for testado (Remington et al., 1999). Desta forma, é possível propor que a herança de praticamente todos os genes pode ser seguida de uma geração para a outra e a composição alélica da progênie gerada a partir de um cruzamento pode ser determinada (Myburg, 2002).

A regressão dos valores de mensurações fenotípicas de indivíduos de uma progênie sobre os dados de segregação de marcadores moleculares têm

permitido a identificação de marcadores associados a regiões genômicas que contêm genes ou seqüências regulatórias associadas com o controle das características de interesse. Na maioria dos casos, não se pode determinar se o efeito detectado por um loco marcador corresponde a um ou mais genes ligados afetando a característica. Por este motivo, o termo Quantitative Trait Locus (QTL) é utilizado para descrever uma região do cromossomo que tem um efeito detectável e significativo sobre uma característica quantitativa (Tanksley, 1993). Atualmente, estas regiões genômicas ou QTL são identificadas através de diferentes metodologias estatísticas que associam de forma mais ou menos sofisticada o marcador ou pares de marcadores flanqueantes à característica estudada (Zeng, 1994; Lander & Botstein, 1989). A ocorrência de recombinação entre os marcadores ligados e o QTL nas gerações subseqüentes àquela na qual foi detectada a associação, leva a uma redução na força da associação, pela redução gradativa do desequilíbrio de ligação com o passar das gerações, diminuindo ou eliminando a potencial utilidade deste marcador em programas de melhoramento (Strauss et al., 1992). Entretanto, o desequilíbrio de ligação (LD) entre marcadores e QTL é maximizado em cruzamentos intra e interespecíficos. Desta forma associações entre marcador e gene podem persistir por varias gerações (Dekkers & Hospital, 2002).

A limitação dos experimentos de mapeamento de QTL em árvores tem sido a dificuldade de construção planejada de genótipos contrastantes para maximizar a segregação fenotípica, estratégia clássica para localização mais

eficiente de QTL. Portanto, os únicos QTL que podem ser potencialmente detectados são aqueles que, coincidentemente, encontram-se em heterozigose nos genitores do cruzamento e onde o efeito diferencial entre alelos alternativos aos QTL é relativamente grande para permitir detecção estatística. Devido ao tamanho relativamente pequeno das progênies empregadas, a precisão limitada na avaliação do fenótipo, o número reduzido de “*backgrounds*” genéticos testados e as limitações inerentes aos tipos de marcadores moleculares utilizados até hoje, as características de QTL em árvores ainda são pouco compreendidas (Grattapaglia, 2003).

Em eucalipto, diversos mapas de ligação cobrindo quase a totalidade do genoma já foram relatados (Grattapaglia et al., 1994; Byrne et al., 1995; Verhaegen & Plomion, 1996; Marques et al., 1997; Brondani et al., 2002; Thamarus et al., 2002; Gan et al., 2003; Myburg et al., 2003) e QTL para características de interesse tem sido mapeados em alguns desses mapas.

O primeiro trabalho envolvendo mapeamento de QTL em eucalipto foi realizado por Grattapaglia et al. (1995) onde foram detectados QTL para resposta a micropropagação, capacidade de enraizamento e brotação. Características relacionadas com crescimento volumétrico e densidade da madeira também foram mapeados em uma família de meios irmãos de polinização aberta de *E. grandis* (Grattapaglia et al., 1996). Mapeamento de QTL para altura da plântula, tamanho da área foliar foram mapeadas por Byrne et al. (1997). Características relacionadas com habilidade de propagação

vegetativa como, produção de raízes adventícias e capacidade de brotação *in vitro* também foram mapeadas (Marques et al., 1998). Finalmente, já foram encontrados QTL para características como crescimento volumétrico, comprimento de fibras e forma do fuste (Verhaegen et al., 1997; Squilassi Grattapaglia, 1997; Moran et al., 2002) e recentemente foi identificado um QTL para resistência à *Puccinia psidi* para *E grandis* (Junghans et al. 2003). Embora algumas características importantes tenham sido mapeadas, as características de maior impacto potencial no processo industrial, relacionadas às propriedades físicas e químicas da madeira praticamente não foram contempladas em estudos de mapeamento genético, sendo encontrado na literatura apenas um trabalho envolvendo a detecção de QTL para teor de celulose (Moran et al., 2002).

Atualmente, experimentos de mapeamento de QTL estão sendo aperfeiçoados e expandidos para as principais espécies florestais nos vários projetos de pesquisa genômica no mundo. Busca-se agora delineamentos que permitem não apenas o mapeamento mas a validação de QTL e estudos de interação, fornecendo subsídios fundamentais para a co-localização de QTL e genes candidatos. Ainda há várias questões a serem investigadas quanto à variabilidade existente e o comportamento interativo de QTL específicos com ambientes, backgrounds genéticos distintos e idades. Estes dados são cruciais para que informações de mapeamento possam ser usadas operacionalmente na prática do melhoramento genético (Grattapaglia, 2003). Todas estas

melhorias no planejamento e condução de experimentos de mapeamento visam elucidar a arquitetura genética de características quantitativas importantes em árvores e gerar fundamentos mais sólidos para o estabelecimento de procedimentos de seleção assistida por marcadores (Grattapaglia, 2003).

2.5 Florescimento precoce

Diversas mudanças fisiológicas ocorrem durante o florescimento de uma planta. As principais estão relacionadas ao meristema apical, que, a partir de vegetativo, se torna reprodutivo. Na maioria das plantas, esta transição é fortemente influenciada por mudanças ambientais como o comprimento do dia (fotoperíodo) e exposição prolongada da planta a baixas temperaturas (vernalização).

Recentemente, esforços têm sido realizados no intuito de se entender melhor os mecanismos que controlam o florescimento em uma planta. Estudos moleculares realizados com *Arabidopsis* têm revelado que o tempo de florescimento é influenciado por um grande número de genes, a maioria relacionada ao controle do fotoperíodo e vernalização.

Hoje, já é sabido que a regulação do tempo de florescimento se dá através de um conjunto de quatro vias metabólicas, dentre as quais duas respondem a mudanças ambientais (fotoperíodo e vernalização) e duas atuam independentemente de sinais ambientais: a via metabólica autônoma, que

promove o florescimento sob quaisquer condições ambientais, e a via da giberelina, que é necessária para o florescimento em dias curtos (Blázquez et al., 2001).

Em *Arabidopsis*, o florescimento ocorre rapidamente sob fotoperíodos longos. Entretanto, quando expostas a fotoperíodos curtos, as plantas apresentam um período muito maior de crescimento vegetativo antes do florescimento. Genes que controlam a resposta da planta ao fotoperíodo induzindo o florescimento em condições de dias curtos foram identificados, inicialmente, através de mutações que causaram o florescimento tardio em plantas expostas a condições de dias longos. Como exemplos de genes que estão relacionados ao fotoperíodo podem ser citados *Constans*, *Leafy*, *Gigantea* e *Apetala 1*. Um outro grupo de genes, *Luminidependens*, *FCA*, *FVE*, *FY* e *FPA*, forma a via metabólica autônoma, responsável pelo monitoramento do desenvolvimento da planta e que é ativada independentemente do fotoperíodo.

Dos genes que estão envolvidos na resposta a vernalização, podem ser citados *Frigida*, *Agamous* e *Flowering Locus C (FLC)*. Dentre esses genes, o gene *FLC* destaca-se como sendo um dos mais importantes pois a planta se torna apta a florir apenas quando existe um decréscimo no seu nível de transcrição (Rouse et al., 2002). Por outro lado, quando este gene é superexpresso, a planta se torna completamente insensível à vernalização (Sheldon et al., 1999).

Para a formação das estruturas florais, diversos estudos realizados com plantas mutantes demonstraram inicialmente a presença de 3 grupos de genes (modelo ABC) que são expressos simultaneamente (Coen et al., 1991). Em *Arabidopsis*, a atividade dos genes pertencentes ao grupo A, como *Apetala*, *Cauliflower* e *Fruitfull*, é responsável pelo desenvolvimento das sépalas. Estes genes, em conjunto com os genes pertencentes ao grupo B, como *Apetala3* e *Pistilata*, induzem o desenvolvimento das pétalas. Genes pertencentes ao grupo B interagindo com o gene *Agamous*, pertencente ao grupo C, são responsáveis pela produção dos estames. A expressão do gene *Agamous* sozinho induz a produção dos carpelos. Com a descoberta de novos genes atuando de forma isolada ou interagindo com outros genes que atuam na expressão do florescimento, alguns trabalhos propõem a organização dos genes em 5 grupos, organizados em um modelo ABDCE (Leland et al., 2004).

Atualmente, já é sabido que muitos dos genes que participam e interferem no florescimento de uma planta pertencem ao grupo de genes conhecidos como MADS-box, que são fatores de transcrição com seqüências altamente conservadas entre diversos organismos incluindo animais, levedura, plantas e bactéria. Atualmente, sabe-se que *Arabidopsis* possui 107 genes pertencentes à família MADS-box organizados em 5 grupos, sendo que a maioria apresenta função ainda desconhecida (Perenicova et al., 2003). Apesar de serem amplamente estudados em trabalhos envolvendo florescimento, já é

sabido que os genes MADS-box também estão envolvidos na formação do caule, folhas e raízes (Leland et al., 2004).

Em estudos envolvendo florescimento, MADS-box já estão bastante caracterizados, principalmente em herbáceas. Em arroz, Shinozuka et al. (1999) identificaram e mapearam 35 genes MADS Box em 11 espécies de arroz utilizando marcadores RFLP (Restriction Fragment Length Polymorphism). Em cevada, Schmitz et al. (2000) desenvolveram um trabalho envolvendo clonagem, mapeamento e estudo de expressão para 6 genes MADS Box pertencentes às famílias *Agamous* e *Squamosa*. Um trabalho realizado por Tadege et al. (2001), em *Brassica*, estudou como o gene *FLC* influencia o tempo de florescimento e descobriu que pelo menos 5 genes *FLC* estão envolvidos no controle do florescimento para este gênero.

Em plantas perenes, Linden et al. (2002) clonaram e caracterizaram 3 novos genes MADS Box envolvidos no controle do florescimento de maçã como sendo da família *Agamous*. Em *Populus*, onde plantas normais florescem após 4 anos de vida, foi obtida uma planta transgênica expressando o gene *Leafy*, acelerando o florescimento para 7 meses (Smyth, 1996). Ainda em *Populus*, Leland et al. (2004) estudaram diversos grupos de genes MADS-box sendo expresso tanto em florescimento como em estruturas vegetativas sugerindo atuação diferenciada desses genes em árvores quando comparados com herbáceas.

2.6 Estratégia de BSA (*Bulk Segregant Analysis*)

Testes de ligação entre marcadores moleculares e genes controlando fenótipos de interesse se tornaram uma importante ferramenta para entender a arquitetura de características quantitativas e, ainda em pequena escala, para auxiliar programas de melhoramento genético, principalmente em espécies alógamas de longo ciclo de vida como árvores e animais domésticos. Análises envolvendo SAM (Seleção Assistida por Marcadores) para uma característica de interesse necessitam de estudos prévios, como mapeamento de alta resolução para garantir a localização de marcadores co-segregando com a característica. Entretanto, a localização de um marcador molecular co-segregando com um gene de interesse exige a análise de um número considerável de indivíduos da população segregante, o que muitas vezes pode se apresentar como um fator limitante para garantir o sucesso do experimento. Neste sentido, a utilização da estratégia de BSA (*Bulk Segregant Analysis*) se torna interessante para estudos envolvendo a identificação direcionada de segmentos genômicos específicos contendo locos de grande efeito pois permite uma ampla e extensiva triagem de marcadores cobrindo grande parte do genoma a uma alta densidade em um curto espaço de tempo.

A estratégia de BSA foi inicialmente proposta por Michelmore et al. (1991) e se baseia na comparação de dois bulks de DNA formados a partir de uma progênie segregante para determinada característica, onde os bulks são formados utilizando indivíduos com fenótipos contrastantes para a característica

de interesse. Por causa do elevado desequilíbrio de ligação existente dentro de uma família segregante, marcadores moleculares que estão fortemente ligados a loco controlando uma característica de interesse deverão estar fixados em um dos bulks e polimórficos em relação ao outro bulk. Por outro lado, marcadores que não apresentam uma forte ligação ou segregam de forma independente da característica estarão presentes em ambos os bulks. Desta forma, marcadores que apresentarem ligação com o fenótipo poderão ser selecionados de acordo com comparações de frequência dos alelos presentes em um dos bulks.

Esta estratégia tem sido utilizada com sucesso principalmente para a localização de características de herança simples, como resistência a doenças. Em Banana, Lheureux et al. (2003) localizaram 10 marcadores AFLP co-segregando com a presença e/ou ausência de BSV (*Banana Streak Vírus*). Um outro trabalho realizado por Akano et al. (2001), em mandioca, utilizou esta metodologia para localizar um marcador microssatélite e um marcador RFLP que flaqueiam o gene que confere resistência a CMD (*Cassava Mosaic Disease*). Altinkut et al. (2003) utilizaram a estratégia de BSA para identificar marcadores microssatélites associados à tolerância a estresse hídrico em trigo.

Em um trabalho realizado em eucalipto, Junghans et al. (2003) localizaram um marcador RAPD que co-segrega com o loco *Ppr1* que confere resistência a *Puccinia psidii*.

Para características mais complexas, a utilização da estratégia de BSA pela forma como ela é conduzida, consolidando a variação quantitativa em

duas classes extremas, tipicamente resulta na identificação de apenas um QTL ou nenhum. Por se tratar de características mais complexas, a estratégia de BSA não apresenta o mesmo sucesso quando realizada com características qualitativas devido, principalmente, à presença de sobreposições de genótipos dentro das classes fenotípicas analisadas (Pérez-Enciso, 1998). Nesse sentido, alguns trabalhos baseados em modelos estatísticos foram realizados com intuito de estimar com maior acurácia a frequência de alelos que compõem os bulks, garantindo desta forma uma maior confiabilidade na seleção de marcadores supostamente ligados com a característica de interesse (Carleos et al., 2003; Collins et al., 2000; Shaw et al., 1998). Para a localização de QTL envolvendo características multifatoriais, a estratégia de genotipagem seletiva que também utiliza extremos fenotípicos mas realiza a genotipagem individual para os descendentes de uma família segregante, tem sido utilizada com sucesso (Lander et al., 1989; Darvasi et al., 1992).

2.7 Genes de lignificação

A lignina, juntamente com a celulose e hemicelulose são os principais componentes da madeira. Durante a produção da pasta de celulose, para produção de papel, a lignina é quimicamente separada de outros polissacarídeos que compõem a madeira. Esta extração da lignina exige a utilização de uma grande quantidade de substâncias químicas e portanto, requer um investimento de capital significativo por parte das empresas. A separação

da lignina da celulose é de fundamental importância para a obtenção de um papel de boa qualidade, e deve ser realizada de uma forma muito criteriosa, pois resíduos de lignina na pasta de celulose comprometem a cor e o brilho do papel, já tratamentos exaustivos com esses químicos reduzem a qualidade do produto final (Chiang et al., 1988). Por estas razões, diversas pesquisas estão sendo realizadas no sentido de minimizar este custo e facilitar a separação de ambos os polissacarídeos, sem comprometer a qualidade da pasta da celulose.

O monômero de lignina é produzido intracelularmente e depois exportado para a parede celular e subseqüentemente polimerizado na presença dos álcoois *p*-coumaril, coniferil e sinapil. Quando incorporados na lignina, esses álcoois são chamados de *p*-hidroxifenil (H), guaiacil (G) e siringil (S) se transformando em monolignóis. Esses monolignóis diferem entre si pelos graus de metilação e hidroxilação no carbono 3 e carbono 5 de seus anéis aromáticos. Em adição a estes monolignóis, são encontrados também unidades de fenilpropanóides como acetatos, aldeídos, *p*-coumaratos, *p*-hidroxibenzoatos e tiramino ferrulato (Boerjan et al., 2003). A complexidade e a heterogeneidade do polímero de lignina depende da proporção relativa de H, G e S assim como níveis de metilação e hidroxilação desses monolignóis (Baucher et al., 2003). Das enzimas envolvidas na via metabólica da lignina, a fenilalanina amônia-liase (PAL) catalisa o primeiro passo da via do fenilpropanóide. As enzimas ácido cafeico *o*-metiltransferase (COMT) e cafeoil-CoA *o*-metiltransferase (CCOAOMT) estão ligadas à metilação dos carbonos 3 e 5, enquanto a ferulato

hidroxilase (5FH) (também denominada coniferaldeído-5-hidroxilase CAld5H) está ligada à hidroxilação desses carbonos. A enzima 4 coumarato-CoA ligase (4CL) parece estar envolvida na conversão do ácido sinápico em sinapoil-CoA para produção do monolignol S. A cinamoil-CoA redutase (CCR) está ligada à regulação do fluxo de carbono na lignina. Já a álcool cinamil desidrogenase (CAD) catalisa o último passo dos monolignóis, reduzindo os cinamaldeídos em álcool cinamil (Baucher et al., 2003). Diversos estudos envolvendo os genes que codificam para estas enzimas da via de lignificação vem sendo realizados, principalmente em sistemas transgênicos visando principalmente suprimir a expressão desses genes. Um estudo realizado por Korth et al. 2001 diminuindo a expressão em diferentes níveis da PAL, demonstraram que a supressão de 98% na expressão do gene para esta enzima em caules de plantas transgênicas de tabaco resultou na redução de 70% no teor de lignina em este tecido. Porém, estas plantas transgênicas apresentaram atrofiamentos, enrolamento das folhas e susceptibilidade à patógenos (Sewalt et al., 1997). Já a supressão do gene que codifica para a COMT realizada por Marita et al. (2003) em plantas de alfalfa resultou na redução de 20% do teor de lignina e em um aumento do 10% do teor de celulose, no entanto estas plantas não mostraram mudanças estruturais quando comparadas com suas testemunhas. Em álamo, gênero *Populus*, sistemas transgênicos suprimindo a expressão da COMT obtiveram uma redução do teor de lignina de 17% (Jouanin et al., 2000). Em outros trabalhos realizados com esta mesma espécie suprimindo o gene que codifica para COMT, as plantas transgênicas obtidas não sofreram

diferenças de teor de lignina (Van Doorselaere et al., 1995; Pilate et al., 2002). Apesar dessas discrepâncias para o gene COMT, estudos envolvendo supressão destes genes em plantas de *Populus* e tabaco tem demonstrado que o COMT apresenta um importante papel na incorporação da unidade S e G na polimerização da lignina (Li et al., 2000, Guo et al., 2001 e Marita et al., 2003). Em um trabalho realizado por Meyermans et al. 2000 em *Populus* reduzindo os níveis de CCOAOMT em 10% verificou-se uma redução dos níveis de lignina de 12%. Um outro trabalho ainda em *Populus*, níveis da enzima CCOAOMT foram reduzidos a 70% e verificou-se também uma redução do nível de lignina para 60% (Zhong et al., 2000). Em um trabalho realizado por Li et al. (2003) foi realizada a superexpressão da enzima CAld5H (5FH) em conjunto com a subexpressão de 4CL em plantas transgênicas de *Populus*, resultando em um transgênico específico que apresentou uma redução de 52% no teor de lignina e um incremento de 30% no teor de celulose. Em um outro trabalho realizado por Hu et al. (1999), plantas transgênicas de *Populus* subexpressando o gene da enzima 4CL apresentaram uma redução no teor de lignina de 45% e um incremento no teor de celulose de 15%. Em plantas de tabaco subexpressando o gene para a enzima CCR, aberrações fenotípicas como redução no crescimento, deformação foliar e mudanças estruturais na lignina foram observadas. Entretanto ao se obter uma subexpressão de 48% da enzima, o teor de lignina foi reduzido a 58% (O'Connell et al., 2002). Trabalhos realizados em *Populus* subexpressando o gene CAD mostraram que, apesar da pequena diminuição no teor de lignina quando comparados com testemunhas, mudanças

estruturais na lignina facilitaram seu processo de separação da celulose (Lapierre et al., 1999 e Pilate et al., 2003). Em Eucalipto ainda não foram publicados trabalhos envolvendo sub ou superexpressão de genes envolvidos na via de lignificação embora diversos esforços estejam atualmente em andamento em diversos laboratórios públicos e privados no mundo. Trabalhos envolvendo a clonagem e caracterização de genes codificadores de enzimas envolvidas na via de lignificação já foram realizados em eucalipto para o gene CAD (Hawkinns et al., 1994 e Goffner et al., 1998) e para o gene CCR (Lacombe et al., 1997). Em estudos envolvendo mapeamento genético de eucalipto os genes PAL, CCR, COMT, CAD, CCOAOMT e 4CL já foram mapeados (Gion et al., 2000 e Thamarus et al., 2002).

2.8 Polimorfismo de Seqüência única (*Single Nucleotide Polymorphisms – SNP*)

O polimorfismo de nucleotídeo único (*Single nucleotide polymorphisms - SNP*), é a classe de polimorfismo mais comum encontrada no genoma. Resultados do projeto genoma humano mostraram a existência de 1,42 milhões de SNPs presentes no genoma humano, o que representa em média a ocorrência de 1 SNP a cada 1,9 kilobases (The International SNP Map Working Group, 2001). Dependendo da posição de ocorrência de um SNP, podem ser observadas diferenças mensuráveis em nível de fenótipo como reflexo da troca de uma base nucleotídica. SNPs localizados em regiões

codificadoras de genes que alteram a função, estrutura ou até mesmo a expressão de uma proteína são muitas vezes causas suficientes para explicar parte da variação fenotípica entre indivíduos bem como resistência/susceptibilidade a doenças. Em humanos diversos estudos vem sendo realizados no sentido de relacionar a presença de um SNP à susceptibilidade ou resistência a doenças como câncer de pulmão (Zhu et. al., 2001), pressão sangüínea (Koschinsky et. al., 2001) e diabetes (Tsunoda et. al., 2001) por exemplo. Em plantas, estudos de diversidade genética envolvendo SNPs vêm sendo realizados de forma crescente no últimos anos, principalmente em milho e *Arabidopsis*. Em milho, Ching et al. 2002, estudaram a freqüência e distribuição de SNPs em 18 genes em 36 linhagens de milho, e verificaram que a taxa de mutação em milho é bastante elevada, apresentando uma media de 1 SNP a cada 31 pares de bases em regiões não codificantes e uma media de 1 SNP a cada 129 pares de base em seqüências codificantes. Verificaram também que o nível de inserção e deleção de bases (INDELS) é freqüente em regiões não codificantes (1 a cada 85 pares de base), mas raras em regiões codificantes. Um trabalho realizado pela Celeron Genomics em *Arabidopsis* organizou uma coleção de 37.344 SNPs e os disponibilizou em um banco de dados (<http://www.arabidopsis.org/Cereon/>). Um outro trabalho realizado por Cregan et al. (2001) utilizando 22 indivíduos geneticamente contrastantes de soja, identificou que a media de ocorrência de SNP em regiões codificantes é de 1,64 SNPs/Kilobase e de 4,85 SNPs/Kilobase para regiões não codificantes. O GeneBank contém diversos Sequence Tagget Site (STS)

para duas linhagens polimórficas de milho (B73 e Mo17) onde podem ser encontrados diversos SNPs, bem como indels para estes STSs (Rafalski, 2002). Em espécies perenes, Garcia-Gil et.al. (2003) estudaram a variação nucleotídica para os genes PHYP e PHYO em 3 populações de *Pinus sylvestris* e encontraram uma baixa variação nucleotídica com valores de Π (diversidade nucleotídica) iguais a 0,0003 e 0,0002, respectivamente para esses genes. Em *Populus*, um trabalho realizado por Zaina. et al, (2004) envolvendo 29 locos e 12 indivíduos de *Populus nigra* estudaram a frequência e natureza de SNPs nos indivíduos, bem como estudos de desequilíbrio de ligação e inferência de haplótipos. Para um total de 11672 pb analisados foram encontrados 94 SNPs, sendo mais frequentes em seqüências não codificantes (1 SNP a cada 97pb) do que em seqüências codificantes (1 SNP a cada 207pb), apresentando uma diversidade nucleotídica total de 0,0027.

2.9 Estudos de Expressão gênica utilizando SNPs

Uma mutação pontual em regiões codificadoras para um gene pode alterar uma proteína seja alterando sua seqüência de aminoácidos ou alterando sua expressão. Alterações em nível de aminoácidos podem ser caracterizadas diretamente pela análise da seqüência de DNA e já foram bastante descritas. Por outro lado, poucos estudos foram realizados até o momento sobre a consequência de polimorfismos em regiões reguladores da expressão de um gene. Baseado nos trabalhos de Johnson et al. (2000) e Levine (2002) que

mostraram que a diferença entre níveis de expressão de genes é responsável pela maior parte da variação fenotípica dentro e entre espécies, Yan et al. (2002) desenvolveram uma metodologia para quantificar a expressão gênica alelo específica de 13 genes baseada na detecção de SNPs que discriminam alelos via marcação com dideoxi fluorescentes. No trabalho de Van et al. (2002) foram analisadas células linfoblastóides derivadas de 96 indivíduos totalizando uma media de 23 indivíduos informativos para os genes analisados. Como resultado, foi possível verificar que seis dos 13 genes analisados apresentaram diferença significativa na expressão gênica entre alelos, indicando que a expressão gênica em humanos pode ser altamente influenciada por variações genéticas atuando em cis (*cis acting*). Se uma causa da variabilidade fenotípica em humanos, incluindo susceptibilidade a doenças é devida à regulação agindo em cis (*cis acting*), pode-se afirmar que esta regulação deve ser um evento comum no processo de transcrição de um gene. Para verificar a consistência da regulação agindo em cis, Bray et. al. (2003) estudaram a expressão alélica de 15 genes utilizando RNA derivados de 19 amostras de cérebro humano de indivíduos não relacionados. Foi demonstrado que em 7 dos 15 genes estudados ao menos um individuo apresentava variação na expressão alélica superior ou igual a 20%, e que para um determinado gene (DTNBP1), esta variação foi superior a 50%, sugerindo que a variação da expressão gênica devido à regulação agindo em cis pode ser uma importante ferramenta no entendimento da diversidade fenotípica.

Esta abordagem para o estudo e entendimento do papel de variação genética em regiões reguladoras é recente e somente agora começa a ser explorada em plantas. Em milho, Radovic et al. (comunicação pessoal) estudaram a influência do estresse abiótico na expressão alelo específica em dois híbridos recíprocos derivados de um cruzamento entre duas linhas puras. Foi observado que 58% dos genes analisados apresentaram diferenças de segregação alélica significativas quando tecidos de folha e raiz foram analisados.

3 UM MÉTODO PARA GENOTIPAGEM DE MICROSSATÉLITES COM DETECÇÃO FLUORESCENTE EM QUATRO CORES UTILIZANDO MÚLTIPLOS INICIADORES DE PCR COM CAUDA.

Resumo

A técnica considerada como o padrão internacional para a genotipagem de microssatélites se baseia em sistemas que automatizam a captura e análise dos dados através da detecção a laser dos fragmentos amplificados via PCR marcados com diferentes fluorocromos. A discriminação de cada loco é realizada de acordo com o fluorocromo utilizado na marcação do iniciador direto da PCR. Um fator limitante para o uso desta metodologia é o custo da síntese de iniciadores marcados com fluorocromos o que em geral aumenta o custo de genotipagem em cerca de 5 a 10 vezes quando comparado com métodos de detecção colorimétrica. O presente trabalho aprimorou e estendeu o conceito e metodologia de utilização de iniciador de PCR com cauda (*primer tail*) propostos por Oetting et al. (1995) que se baseia na utilização de um iniciador universal de PCR marcado para adicionar marcação fluorescente em locos microssatélites permitindo assim a detecção em seqüenciadores automáticos. Neste trabalho, foi desenvolvido um método de *primer tail* baseado

na utilização de três diferentes caudas e primers universais para a análise de marcadores microssatélites em eucalipto. Para o desenvolvimento das caudas, foram utilizadas 3 seqüências de 20 bases nucleotídicas específicas de microssatélites do genoma humano, não apresentando portanto homologia com o genoma de eucalipto. Este método permitiu a análise e detecção simultânea e eficiente de microssatélites de *Eucalyptus* em um experimento de mapeamento genético. O uso de múltiplos primers com cauda permitiu a multiplexagem de microssatélites seja na mesma PCR bem como no carregamento da eletroforese dos produtos amplificados de cada loco individualmente, aumentando a eficiência de utilização do tempo de um seqüenciador automático e de geração de dados genotípicos. Esta abordagem tem sido amplamente aplicada a várias espécies no nosso laboratório aumentando significativamente a velocidade e qualidade dos dados gerados em comparação com metodologias de detecção baseadas em nitrato de prata e muito mais econômica em comparação à utilização convencional com todos os iniciadores marcados.

A METHOD FOR MICROSATELLITE GENOTYPING WITH FLUORESCENT DETECTION WITH FOUR COLORS USING MULTIPLE TAILED PRIMERS.

Summary

The international standard for microsatellite genotyping is based on real time capture and analysis of data through laser detection of the PCR amplified fragments using primers labelled with different fluorochromes. The discrimination of each marker is carried out according to the different fluorescences. A limiting factor for a more accessible use of this methodology is the cost associated with the synthesis of labeled primers. The cost of fluorescent labeling detection is 5 to 10 times higher when compared with traditional colorimetric methods. The present work extended and improved the concept and method of tailed primer proposed by Oetting et al. (1995). This method is based on the use of a labeled universal PCR primer that is complementary to a tail that is added to the forward primer of the microsatellite marker, thus adding a fluorescence to the microsatellite amplified allele, allowing its detection on an automatic sequencer. In this work, an improved primer tail method was developed and applied to *Eucalyptus* microsatellites, based on the use of three

different tails and universal primers making possible the simultaneous analysis of several microsatellites. For the development of the tails, three different 20-base sequences corresponding to the forward primer sequence of human microsatellites were used, with no homology to eucalyptus or any plant genome. The robustness of the method was tested for the simultaneous detection and genetic analysis of Eucalyptus microsatellites for a genetic mapping experiment. The use of multiple tailed primers allowed the use of multiplexed microsatellites both in the same PCR as well as multiplexed loading of different PCR amplified products to perform the electrophoresis, increasing the efficiency of data generation. This approach has been applied to several species in our laboratory, increasing significantly the speed and quality of the data generated in comparison with colorimetric detection methodologies such as silver nitrate staining, and significantly cheaper when compared to the standard use of all labelled primers especially when applied to multiple species and microsatellite markers.

3.1 Introdução

Genotipagem de marcadores microssatélites tem sido amplamente utilizada, para estudos em diversas áreas desde investigações na genética forense, estudos de estrutura e diversidade genética em populações naturais (Collevatti et al., 1999), diagnose de doenças em humanos (Bouchlaka et al., 2004), construção de mapas genéticos e localização de locos controladores de características quantitativas em plantas (Brondani et al., 2000; Moran et al., 2002), animais domésticos e humanos (Mikawa et al., 2004; Stylianou et al., 2003).

Embora vários sejam os métodos utilizados para a resolução de alelos de microssatélites, atualmente, a técnica de melhor resolução e considerada como padrão internacional se baseia em sistemas que automatizam a análise dos dados através de detecção a laser de fragmentos amplificados via PCR marcados com diferentes fluorocromos. Para tanto, um dos iniciadores (primers) utilizados na reação em cadeia da polimerase (PCR) é marcado com um fluorocromo (ex. 6-FAM, NED, HEX e ROX) que, quando excitado pelo laser, emite uma fluorescência a qual, após passar por filtros virtuais que selecionam o comprimento de onda a ser captado em tempo real

por uma câmara CCD é convertida em um pico no eletroferograma correspondendo ao alelo específico.

A utilização de diferentes fluorocromos cujo máximo de emissão espectral em diferentes faixas de comprimento de onda permite também que a genotipagem de diferentes locos seja realizada simultaneamente. Produtos de PCR de diferentes locos, cada um deles amplificado com primer marcado com um fluorocromo diferente, são misturados e analisados em uma mesma pista de eletroforese. A discriminação de cada loco é realizada de acordo com o fluorocromo utilizado na marcação de cada primer.

Um fator limitante para o uso desta metodologia diz respeito ao custo da síntese de primers marcados com fluorocromos. Enquanto a síntese de um primer com 20 bases custa em torno de 10 dólares, o custo relativo à síntese do mesmo primer, marcado com uma fluorescência, pode aumentar cerca de 5 vezes. O problema se torna ainda maior para estudos que demandam a utilização de uma extensa bateria de locos, como por exemplo a construção de mapas genéticos de ampla cobertura. Para realizar um estudo envolvendo 200 marcadores microssatélites sem marcação fluorescente é necessário um investimento em torno de 4000 dólares. Já se optando por fazer utilizando primers marcados, o mesmo estudo sairia por cerca de 22000 dólares e, na grande maioria dos casos, estes primers são utilizados para um número que gira em torno de algumas centenas de genotipagens e estocados em freezer até uma nova utilização.

Com o objetivo de diminuir os custos da utilização de sistemas de genotipagem automatizada via fluorescência, Oetting et al. (1995) propuseram uma estratégia de PCR que envolve a utilização de um terceiro primer universal, marcado com fluorocromo juntamente com os primers do loco marcador sendo que um destes primers possui uma cauda que permite o anelamento do primer universal marcado. Esta estratégia de três primers baseia-se no acoplamento ao terminal 5' de uma cauda (tail) de uma seqüência não específica à do genoma estudado. Neste trabalho original foi utilizada a seqüência do primer M13 (sigla derivada do bacteriófago M13) comumente utilizado para sequenciamento. Essa seqüência cauda tem a função de gerar, no terceiro ciclo da PCR, uma seqüência complementar àquela do primer universal marcado. Este primer, marcado com um fluorocromo no seu terminal 5', é utilizado na reação de amplificação da PCR e irá anelar na fita de DNA apenas a partir do terceiro ciclo, gerando um produto marcado passível de genotipagem automatizada (Figura 1). Desta maneira, é necessário sintetizar apenas um primer universal marcado com fluorescência e não um para cada loco analisado. Esta estratégia denominada de *primer-tail* foi utilizada por Oetting et al. (1995) para amplificação de 148 marcadores microssatélites em um experimento de mapeamento em seres humanos.

Após o trabalho de Oetting et al. (1995), diversos outros foram realizados utilizando a mesma estratégia. Neilan et al. (1997) desenvolveram um sistema utilizando cauda de primer não específico para o genoma estudado

visando amplificação de marcadores RFLP e microssatélites. A detecção foi feita utilizando um primer específico para anelar na cauda usando o fluorocromo TAMRA (Applied Biosystems). Schuelke (2000) utilizou a mesma estratégia para estudar o loco *MaoA* do cromossomo X em humanos utilizando como cauda a seqüência do primer M13 e marcação fluorescente 6-FAM. Em outro trabalho, Guo et al. (2003) desenvolveram um estudo em humanos também utilizando seqüência do primer M13 para realizar a genotipagem automatizada de 4 locos microssatélites marcados com fluorocromos 6-FAM e HEX. Embora a maioria dos trabalhos usando primer com cauda utilize como cauda a seqüência do primer M13, esta estratégia pode ser realizada utilizando quaisquer outras seqüências que não tenham homologia com o genoma a ser estudado (Neilan et al., 1997). Todos estes trabalhos utilizaram apenas um primer universal marcado com uma fluorescência limitando, portanto a detecção simultânea de locos cujos alelos possuem a mesma faixa de tamanho.

Neste trabalho, foi desenvolvido um método melhorado de *primer tail* baseado na utilização de três diferentes caudas e primers universais para a análise, de 45 marcadores microssatélites de eucalipto. Para o desenvolvimento das caudas, foram utilizadas 3 seqüências de 20 bases nucleotídicas específicas de microssatélites do genoma humano, não apresentando portanto homologia com o genoma de eucalipto. Diferente dos demais trabalhos que utilizaram a metodologia de primer tail único, este método permite a detecção simultânea de microssatélites que possuem alelos na mesma faixa de tamanho

com três fluorocromos, HEX, NED 6-FAM sendo que a quarta fluorescência, ROX, é utilizada para a detecção do padrão interno de fragmentos de tamanho conhecido otimizando assim genotipagens em sistemas multiplex.

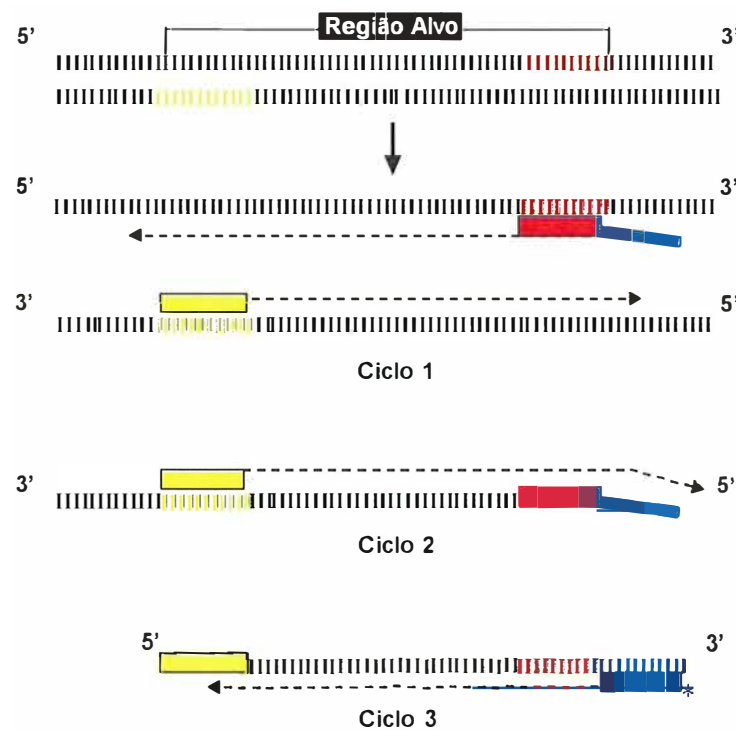


Figura 1 - Desenho esquemático mostrando o princípio da estratégia de primer com cauda para realização de genotipagem automatizada. A barra amarela representa o primer reverso que flanqueia o microsatélite. A barra vermelha representa o primer direto que flanqueia o microsatélite. A barra azul acoplada à barra vermelha representa a seqüência de um primer não específico para o genoma estudado (cauda). A barra azul com um asterisco representa o primer não específico marcado com fluorocromo

3.2 Material e Métodos

3.2.1 Material genético

A amplificação via PCR e detecção de alelos com os primers com caudas foram testados em uma progênie F1 de 88 indivíduos gerados a partir de um cruzamento entre um indivíduo *Eucalyptus grandis* (P1) e um indivíduo *Eucalyptus urophylla* (P2) fornecidos pela empresa FERRO GUSA CARAJÁS S.A.

3.2.2 Extração de DNA

Para a extração do DNA total, foi utilizado tecido foliar dos genitores e de 88 indivíduos da progênie. A extração foi feita pelo método CTAB 2% (Doyle & Doyle, 1987). A maceração das folhas foi realizada utilizando-se nitrogênio líquido. A quantificação do DNA extraído foi feita em gel de agarose 1,5% corado com brometo de etídeo utilizando DNA de fago lambda de concentração conhecida como padrão de comparação.

3.2.3 Desenho dos primers

Três primers todos com 20 bases e já disponíveis no laboratório, utilizados na amplificação de locos microssatélites específicos no genoma humano, e, portanto, sem complementaridade no genoma de *Eucalyptus*, foram utilizados como primers universais para a detecção fluorescente. Estes primers marcados com fluorocromos já estavam disponíveis no laboratório de forma que não foi necessário o investimento inicial na síntese destes primers marcados. Estes primers foram os seguintes: primer direto ("forward") do loco DYS437 marcado com 6-FAM (5' GACTATGGGCGTGAGTGCAT 3'); primer direto ("forward") do loco D8S1132 marcado com HEX (5'GGCTAGGAAAGGTTAGTGGC3'); e o primer direto ("forward") do loco D12S1090 marcado com NED (5' ACCAACCTAGGAAACACAGT 3'). Para cada um de 45 locos microssatélites de *Eucalyptus* (Brondani et al., 2002). foi re-sintetizado (Operon Technologies) o primer direto porém adicionando-se uma cauda no terminal 5', cauda esta correspondendo exatamente à seqüência do primer universal. Visando a multiplexagem de marcadores, um primeiro conjunto de 15 marcadores microssatélites teve incorporado, ao terminal 5' de um dos seus primers iniciadores, a seqüência 5' GGCTAGGAAAGGTTAGTGGC 3', correspondente ao primer D8S1132f (ver Anexo A); um segundo conjunto de 15 marcadores teve incorporado a seqüência 5' ACCAACCTAGGAAACACAGT 3', correspondente ao primer D12S1090f marcado com NED (ver Anexo B), e um terceiro conjunto composto

de 15 marcadores teve incorporado a seqüência 5' GACTATGGGCGTGAGTGCAT 3', correspondente ao primer DYS437f marcado com 6-FAM (ver Anexo C).

3.2.4 Amplificação dos locos microssatélites

A amplificação de cada um dos locos microssatélites foi realizada utilizando os três primers simultaneamente, ou seja, o primer direto com cauda, o primer reverso normal e o primer universal marcado com o fluorocromo. A reação de PCR foi feita utilizando-se 1 X de Tampão de PCR (10 mM Tris-HCl, pH 8.3, 50 mM KCl), 1,5 mM MgCl₂, 0,19 μM de dNTP, 0,24 μg/μl de BSA, 1,0 unidade da enzima *Taq* polimerase, 6,0 ng de DNA, 5,375 μL de água Milli-Q autoclavada. A esta mistura foram adicionados 0,1 μM de cada um dos primers direto + cauda, primer reverso e primer universal marcado conforme as seqüências dos primers + cauda do loco a ser amplificado. A reação da PCR foi realizada utilizando um programa que apresenta uma etapa inicial de desnaturação da fita de DNA a 94° C por 5 minutos, 30 ciclos com etapas de desnaturação (1 minuto a 94° C), anelamento (1 minuto a 57° C) e extensão (1 minuto a 72° C). No final dos ciclos, ocorre uma extensão final de 72° C por mais 30 minutos. A reação de PCR foi realizada separadamente para cada loco e uma mistura dos produtos de PCR de três locos microssatélites foi feita para detecção simultânea dos fragmentos amplificados. Após a mistura das 3 reações de PCR, foi aliqotado 1 μl desta solução e adicionado a 10 μl de Hi-Di

formamida e 1 μ l de marcador de padrão de fragmentos de tamanho conhecido marcado com o fluorocromo ROX, desenvolvido por Brondani et al. (2001).

3.2.5 Geração dos dados genotípicos

As eletroforeses realizadas foram conduzidas na plataforma ABI 3100 da Applied Biosystems. Para a detecção das fluorescências, foi utilizado o filtro virtual D, no qual o fluorocromo 6-FAM emite uma fluorescência azul, o fluorocromo HEX emite uma fluorescência verde, o fluorocromo NED emite uma fluorescência amarela e o fluorocromo ROX emite uma fluorescência vermelha. Os dados foram coletados automaticamente pela detecção das diferentes fluorescências e analisados com os programas GeneScan/Genotyper (Applied Biosystems). Estes programas permitem a genotipagem automatizada dos alelos comparando-os com os valores de um marcador interno.

3.2.6 Análise de segregação dos microssatélites amplificados via primers com cauda

Visando testar a herança e segregação alélica dos marcadores microssatélites amplificados utilizando a estratégia de múltiplos primers com cauda, foram selecionados quatro marcadores microssatélites, dentre os 45 estudados. Estes quatro marcadores selecionados encontravam-se ligados em um mesmo grupo de ligação segundo mapa publicado por Brondani et al.

(2002). Foram selecionados os locos EMBRA 60, EMBRA 164, EMBRA 228 e EMBRA 65.

3.2.7 Montagem do grupo de ligação

Para a montagem do grupo de ligação, utilizou-se o programa Outmap desenvolvido pelo CSIRO especificamente para lidar com misturas de razões de segregação em cruzamentos entre indivíduos heterozigotos (Butcher et al., 2002). Como parâmetros para a montagem do grupo de ligação foram utilizados fração de recombinação de 0,4 e um *Lod Score* mínimo de 3,0. Para ordenamento das marcas foi utilizado o método "Three-opt".

3.3 Resultados

3.3.1 Amplificação dos locos microssatélites usando a estratégia de múltiplos primers com cauda

Foi realizada inicialmente uma análise de 2 indivíduos utilizando-se apenas o primer reverso e o primer direto + cauda do loco EMBRA 2, para verificar se o fato de se ter acoplado uma cauda de 20 nucleotídeos a um dos primers, comprometeria o desempenho da PCR (Figura 2, poços 1 e 2). Após a realização desta análise, foi realizada uma segunda análise para os mesmos indivíduos, utilizando como produto a ser amplificado, uma amostra da PCR anterior. Para a execução desta reação, foi utilizado o primer reverso do loco EMBRA 2 e o direto do loco D8S1132, ou seja o primer universal marcado (Figura 2, poços 3 e 4). Finalmente, realizou-se uma PCR para os mesmos 2 indivíduos utilizando os 3 primers simultaneamente para amplificação do loco EMBRA 2 em uma única reação: o primer reverso, o primer direto + cauda e o primer direto marcado do loco D8S1132 (Figura 2, poços 5 e 6). Foi possível verificar que a amplificação dos alelos ocorreu de forma correta nos três tratamentos. As amplificações utilizando o primer com cauda obviamente geraram alelos com 20 pares de bases adicionais. Esta diferença não é

resolvida e portanto visualizável em gel de agarose 1,5% embora seja possível observar que as bandas nos poços 3 a 6 são de fato ligeiramente maiores em tamanho (Figura 2).

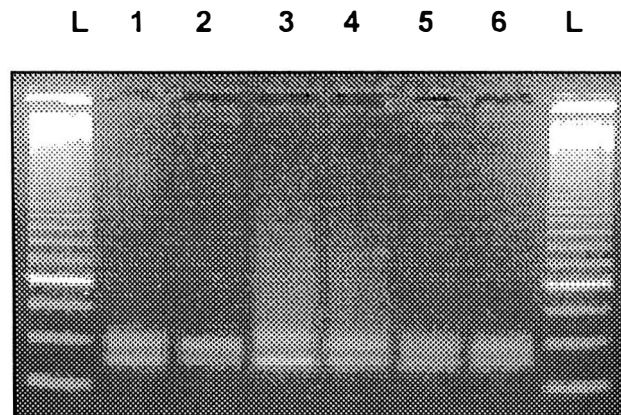


Figura 2 - Foto de um gel de agarose 3,5% corado com brometo de etídeo. L: marcador interno 1Kb; Poços 1 e 2: Resultado de uma PCR realizada com 2 indivíduos da progênie utilizando como iniciadores os primers reverso e direto + cauda para o loco EMBRA 2; Poços 3 e 4: Resultado de uma PCR utilizando como produto a ser amplificado 3ul da PCR anterior, usando como iniciadores os primers reverso do loco EMBRA 2 e o primer direto marcado do loco D8S1132; Poços 5 e 6: Resultado de uma PCR utilizando como iniciadores os primers reverso, direto + cauda para o loco EMBRA 2 e o primer direto do loco D8S1132

3.3.2 Detecção automatizada dos microssatélites

Para confirmar a eficiência e a diferença esperada em pares de bases entre a amplificação do loco de forma convencional e com a metodologia de

primer com cauda, um indivíduo foi genotipado com os locos EMBRA 2, EMBRA 5 e EMBRA 6, utilizando o método convencional de marcação fluorescente onde o primer direto de cada loco era marcado com os fluorocromos HEX, 6-FAM e NED, respectivamente. Este mesmo indivíduo também foi genotipado para os mesmos locos utilizando a estratégia de múltiplos primers com cauda, identificados como EMBRAT 2, detectado com o primer D8S1132; EMBRAT 5, detectado com o primer DYS437 e EMBRAT 6, detectado com o primer OSD12. Os tamanhos dos fragmentos amplificados foram comparados quanto ao tamanho para verificar se os fragmentos amplificados pela nova estratégia realmente eram os desejados. De acordo com o esperado, a diferença de tamanho em pares de bases observada foi de 20 pb para todos os três locos demonstrando que as amplificações com os três primers utilizados simultaneamente ocorreram de forma correta (Figura 3).

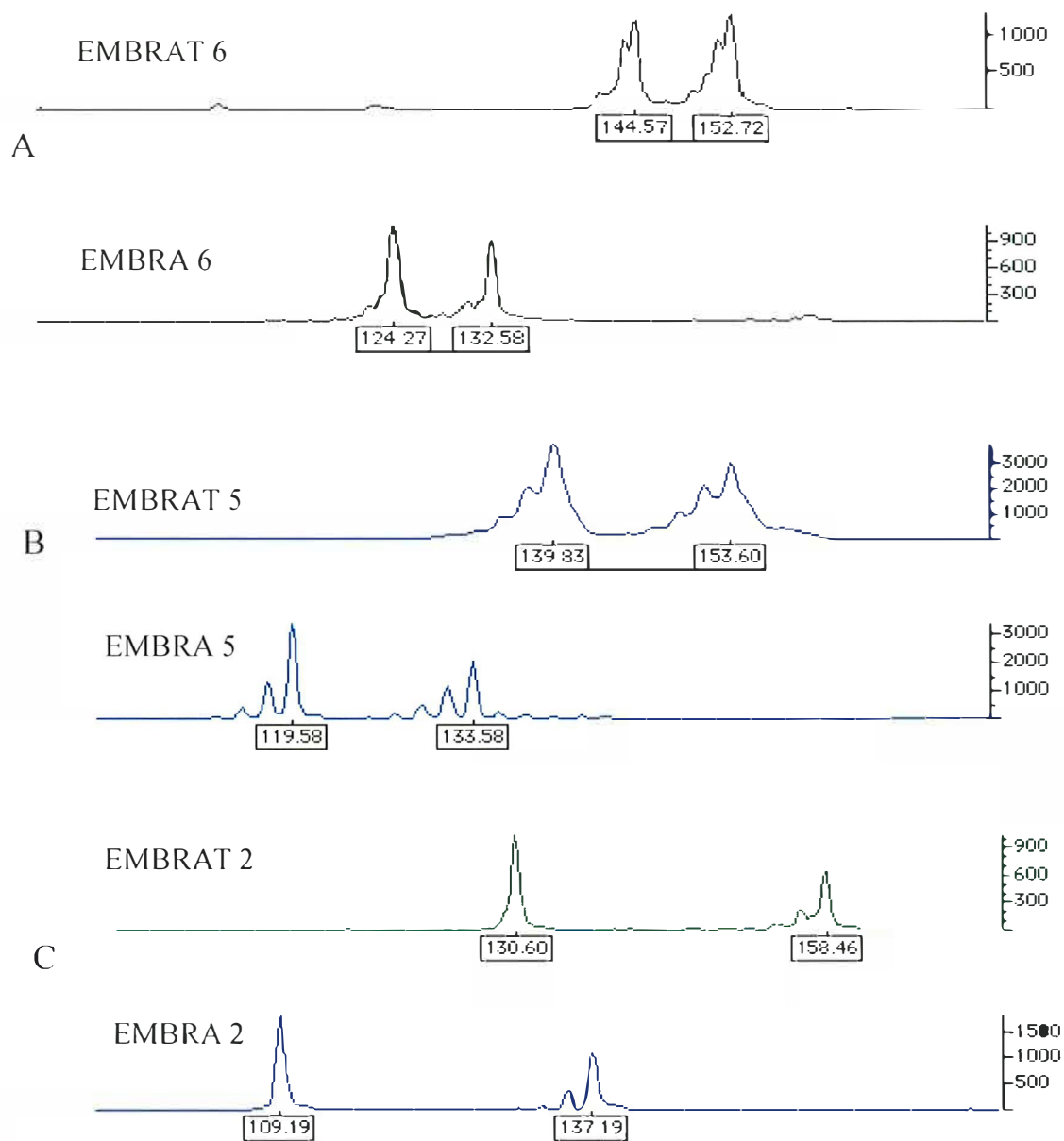


Figura 3 - Comparação do tamanho dos alelos amplificados via PCR utilizando a metodologia convencional de marcação de um dos primers iniciadores e a metodologia utilizando um primer com cauda e primer universal. Comparação das amplificações para os locos EMBRA 6 (A), EMBRA 5 (B) e EMBRA 2 (C)

3.3.3 Análise de segregação dos microssatélites amplificados via primers com cauda

Para verificar o comportamento correto da amplificação dos alelos com a metodologia de múltiplos primers com cauda foi testada a herança e segregação de alguns locos selecionados para genotipagem, locos EMBRA 60, EMBRA 164, EMBRA 228 e EMBRA 65, pertencentes a um mesmo grupo de ligação. Os genótipos para estes quatro locos já haviam sido gerados utilizando o método convencional de marcação. Em dois destes quatro locos um dos genitores apresentava um genótipo heterozigoto para um alelo nulo. A seleção desta configuração visou justamente verificar o comportamento das amplificações em situações de ocorrência de alelo nulo. Estes locos foram utilizados para a genotipagem de 88 indivíduos de uma progênie F1. A tabela 1 mostra o genótipo dos genitores, as classes genotípicas dos indivíduos da progênie e valores de *P* value para teste de distorção de segregação alélica para os locos selecionados. Os locos EMBRA 60, EMBRA 228 e EMBRA 65 apresentaram segregação totalmente informativa do tipo 1:1:1:1 e o loco EMBRA 164 segregou somente a partir de um dos genitores com razão do tipo 1:1. Nenhum loco apresentou distorção na segregação de seus alelos. Os resultados demonstraram que a herança e segregação alélica ocorreu conforme o esperado para estes quatro locos, incluindo a observação correta de genótipos homozigotos aparentes nos casos de segregação envolvendo alelos nulos.

Tabela 1. Resultado da análise de herança e segregação alélica de quatro locos microssatélites analisados utilizando a metodologia de múltiplos primers com cauda. P1 e P2: genótipos dos parentais; classes genotípicas na progênie; Valor p do teste de hipótese nula de segregação mendeliana

Loco	P1	P2	Classes genotípicas				Valor P
EMBRA 60	106/127	106/129	106/106	106/129	106/127	127/129	0,3476
EMBRA 164	133/135	137/137	133/137		135/137		0,6506
EMBRA 228	113/nulo	115/118	113/115	113/118	115/nulo	118/nulo	0,8495
EMBRA 65	209/nulo	219/230	209/219	209/230	219/nulo	230/nulo	0,8810

3.3.4 Análise de ligação

A construção do mapa genético foi realizada utilizando o programa Outmap. O grupo de ligação foi montado utilizando um LOD score 3,0 (OUTMAP). Os quatro locos microssatélites foram mapeados em um mesmo grupo de ligação apresentando um tamanho total de de 4,0 cM. A ordem linear dos marcadores corresponde àquela observada no mapa de ligação gerado por Brondani et al. (2002) (Figura 4).

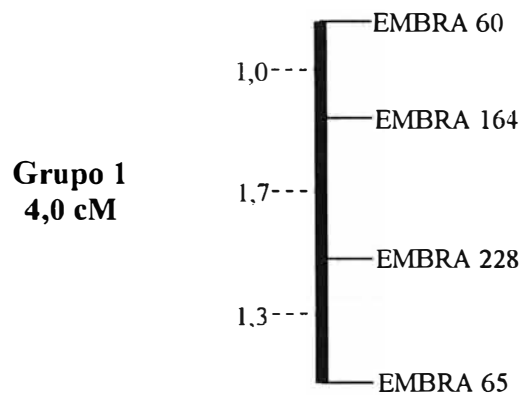


Figura 4 - Grupo de ligação gerado pelo programa OUTMAP para os locos microsatélites EMBRA 60, EMBRA 164, EMBRA 228 e EMBRA 65 analisados com a metodologia de múltiplos primers com cauda

3.4 Discussão

O uso da estratégia de múltiplos primers com cauda para detecção fluorescente foi aplicada em um conjunto de 45 marcadores microssatélites com ampla distribuição no mapa genético de eucalipto. Este trabalho aperfeiçoou a metodologia padrão que se baseia em um único primer com cauda ao utilizar 3 primers específicos marcados com três diferentes cores. Ao utilizar esta metodologia para a amplificação de marcadores em eucalipto, foram utilizadas seqüências específicas de humanos. O oposto possivelmente funcionaria da mesma forma, ou seja, para atuar como caudas podem ser utilizadas seqüências específicas de um determinado genoma sem homologia com o genoma alvo da genotipagem.

Dos 45 locos microssatélites utilizados para o desenvolvimento deste trabalho, todos foram amplificados com sucesso utilizando esta nova metodologia. Apesar das reações em cadeia da polimerase (PCR) terem sido realizadas separadamente para cada loco e reunidas para a realização da eletroforese simultaneamente, protocolos podem ser otimizados no sentido de se tentar amplificar três locos conjuntamente em um sistema multiplex de PCR. Alguns testes foram realizados neste sentido, porém, em geral foi observada a perda de amplificação de um dos três locos ao se comparar a intensidade do

sinal de fluorescência com a amplificação utilizando a marcação convencional. Isso pode ser explicado pelo fato de, ao tentar amplificar simultaneamente três locos utilizando primer com cauda, são efetivamente utilizados nove primers que deverão atuar em conjunto para amplificação destes locos. Neste sentido, qualquer complementaridade entre primers poderia comprometer a amplificação de um dos locos. Otimizações no sentido de se obter amplificações simultâneas para três locos podem ser realizadas.

O teste de herança e segregação realizado com 4 marcadores microssatélites demonstrou que a metodologia é robusta mesmo em situações nas quais alelos nulos ocorrem na configuração. Caso algum artefato de amplificação fosse formado, este possivelmente apareceria no lugar de alelos nulos complicando a análise ao gerar alelos inexistentes. Isso não foi observado. Todos os quatro marcadores apresentaram segregação mendeliana evidenciada pela não rejeição da hipótese nula de segregação (valor $p > 0,05$). Os quatro marcadores testados pertencem ao mesmo grupo de ligação do mapa genético de eucalipto. A ligação e ordem dos mesmos foi obtida de forma consistente ao se realizar a genotipagem com a metodologia de múltiplos primers aqui proposta.

3.5 Conclusão

O presente trabalho aprimorou e estendeu o conceito e metodologia propostos por Oetting et al. (1995) que utiliza um primer universal marcado para adicionar marcação fluorescente em marcadores microssatélites permitindo assim a detecção em seqüenciadores automáticos. Embora alguns outros trabalhos tenham sido publicados utilizando ou propondo melhorias nesta abordagem, o trabalho desenvolvido inova no sentido de utilizar três primers com cauda e os correspondentes primers diretos marcados explorando a inexistência de homologia genômica entre o genoma de plantas e de humanos para marcadores microssatélites.

Embora qualquer seqüência de 20 bases de humanos em princípio poderia ser usada para servir como cauda para amplificação em eucalipto, seqüências em genes ou em regiões repetitivas poderiam apresentar homologia entre plantas e humanos e gerar produtos artefatuais de PCR. A utilização de seqüências de iniciadores de PCR de microssatélites de humanos para amplificação em plantas via primer com cauda apresenta uma série de vantagens, a saber: (1) marcadores microssatélites são em geral altamente selecionados para ser espécie específicos e portanto com mínima probabilidade de permitir amplificação cruzada entre plantas e animais; (2) os primers de

marcadores microssatélites são em geral muito robustos com conteúdo GC acima de 50% e temperatura de anelamento otimizada o que permite a amplificação precisa; (3) em geral diversos primers de microssatélites humanos são disponíveis na literatura e laboratórios ou podem ser adquiridos a preços mais econômicos no mercado por serem amplamente usados em diversos laboratórios do mundo; (4) a ampla disponibilidade de primers humanos permite ainda selecionar seqüências com conteúdo GC ou temperatura de anelamento mais adequada aos microssatélites que se pretende amplificar na espécie de planta alvo. Da mesma forma esta metodologia poderia perfeitamente ser aplicada de forma inversa ao se utilizar seqüências específicas de plantas para amplificação em humanos ou outros animais. Mesmo entre diferentes espécies de plantas, tendo em vista que marcadores microssatélites raramente são transferíveis entre espécies de famílias diferentes, em princípio seria possível utilizar esta metodologia utilizando, por exemplo, seqüências de microssatélites de milho como cauda para amplificação de microssatélites de soja.

Apesar de terem sido encontradas limitações para a utilização de reações conjuntas de três locos, ou seja, multiplex de PCR, isso não é uma surpresa tendo em vista que mesmo a multiplexagem de microssatélites com marcação convencional no 5' de cada primer direto é complexa e demanda otimizações específicas de concentrações relativas de primers. A multiplexagem de produtos amplificados de cada loco individualmente, entretanto, não apresentou qualquer problema e pode ser imediatamente utilizada aumentando a eficiência de utilização do tempo de um seqüenciador

automático e de geração de dados genotípicos. Esta abordagem tem se tornado rotineira no nosso laboratório aumentando significativamente a velocidade e qualidade dos dados gerados em comparação com metodologias de detecção baseadas em nitrato de prata e muito mais econômica em comparação à utilização de todos os primers marcados.

Freqüentemente, laboratórios que trabalham com análise de populações ou mapeamento de diversas espécies de plantas e animais utilizando marcadores microssatélites optam por realizar as genotipagens com base em sistemas tradicionais como eletroforese com visualização em nitrato de prata e/ou brometo de etídeo, mesmo tendo o acesso a seqüenciadores automáticos, justamente devido à limitação financeira. A possibilidade de utilizar todas as fluorescências e com isso multiplexar os produtos de PCR de diferentes microssatélites com alelos em faixas sobrepostas gera um ganho significativo não apenas na quantidade de dados gerados por unidade de tempo, mas também na qualidade do dado genético (precisão em pares de bases) em comparação aos dados gerados com detecção por coloração. Para isso basta ao laboratório solicitar a síntese em grande escala de três primers marcados com diferentes fluorescências e que não apresentem homologia com os genomas trabalhados. Ao mesmo tempo, os primers diretos dos marcadores microssatélites utilizados naquele laboratório deverão ser re-sintetizados adicionado exatamente a mesma seqüência do primer universal no terminal 5'. Visando permitir a multiplexagem de produtos amplificados com a máxima eficiência, chegando a 6 ou mesmo 9 marcadores analisados simultaneamente,

a adição de cauda específica que corresponde a uma determinada cor de fluorescência deverá ser planejada de forma que locos que apresentam alelos na mesma faixa de tamanho deverão ter caudas diferentes enquanto que locos que amplificam alelos em faixas não sobrepostas poderão ter a mesma cauda. Em termos de economia, supondo, por exemplo, uma situação de necessidade de genotipar 9 marcadores microssatélites, a opção de sintetizar primers marcados para todos os nove envolveria um custo aproximado de US\$ 170,00 por loco e um total de US\$ 1530,00. Com a metodologia aqui proposta este custo seria de US\$ 30,00 por loco (considerando um total de 60 bases sintetizadas, sendo 40 para o primer direto e 20 para o reverso a um preço de 0,50/base) e US\$ 120,00 para cada primer marcado, totalizando assim US\$ 630,00, ou seja, uma economia de 59%. A vantagem da metodologia entretanto, é tanto maior quanto maior o número de marcadores a serem genotipados. Supondo por exemplo um experimento que envolve o mapeamento de 200 marcadores microssatélites os custos seriam de US\$ 34000,00 para a metodologia convencional ao se adquirir 200 pares de primers sendo um deles marcado ou US\$ 6360,00 ao se adquirir 200 pares de primers um deles com cauda e mais três primers marcados, uma economia de 81%.

Em conclusão, este trabalho apresenta uma alternativa tecnicamente fácil de ser implementada e economicamente acessível para a genotipagem de marcadores microssatélites com detecção fluorescente ao propor a utilização de múltiplos primers com cauda e os correspondentes primers universais

marcados com diferentes fluorescências, explorando assim todo o poder de geração de dados de um seqüenciador automático.

4 MAPEAMENTO DO QTL *Eef1* (*Eucalyptus early flowering 1*) PARA FLORESCIMENTO PRECOCE EM EUCALIPTO UTILIZANDO “BULK SEGREGANT ANALYSIS” E MARCADORES MICROSSATÉLITES

Resumo

Florescimento precoce em espécies perenes é uma característica interessante para acelerar ciclos de programas de seleção recorrente. Uma planta de *Eucalyptus grandis* com crescimento vegetativo normal e fenótipo de florescimento precoce, aproximadamente aos 90 dias de idade, foi observado nos viveiros na empresa FERRO GUSA CARAJÁS S.A. Esta planta mutante quando cruzada com uma planta do tipo selvagem (florescimento aos 2 a 3 anos de idade) gerou uma progênie de 88 indivíduos que segregou na proporção aproximada de 1:1 para o fenótipo de florescimento precoce. Para a busca de marcadores ligados a este fenótipo, uma genotipagem de cerca de 100 microssatélites amplamente distribuídos pelo genoma, foi realizada utilizando inicialmente a estratégia de BSA (*Bulk Segregant Analysis*) utilizando 15 indivíduos expressando o fenótipo de florescimento precoce e 15 indivíduos do tipo selvagem. Após a genotipagem dos *bulks*, foi realizada a comparação das intensidades de fluorescência dos alelos parentais entre os *bulks*

contrastantes para o fenótipo de florescimento. Foram detectadas distorções de intensidade de fluorescência entre os alelos dos dois bulks para os marcadores microssatélites EMBRA 27, EMBRA 60, EMBRA 164, EMBRA 65, EMBRA 158 e EMBRA 91, indicando possível ligação com o fenótipo de florescimento precoce. Estes locos foram utilizados na genotipagem dos indivíduos da progênie visando confirmar o padrão de co-segregação dos marcadores e o fenótipo de florescimento precoce, apresentando valores de $p < 0,0001$, respectivamente $9,4 \times 10^{-8}$, $6,23 \times 10^{-15}$, $8,9 \times 10^{-15}$, $1,8 \times 10^{-13}$, e $2,3 \times 10^{-14}$, confirmando assim os resultados indicados pela estratégia de BSA. Conforme esperado, esses marcadores, quando localizados no mapa referência desenvolvido por Brondani et al. 2001, apresentaram-se no mesmo grupo de ligação. Um grupo de ligação envolvendo os 5 marcadores e o QTL para a característica de florescimento precoce *Eef1* (*Eucalyptus early flowering 1*) foi construído utilizando o programa JoinMap. O gene *Eef1* foi mapeado entre os marcadores EMBRA 27 e EMBRA 164 apresentando distâncias de 17,5cM e 5,8 cM respectivamente. Para programas de introgressão assistida do segmento genômico contendo o QTL *Eef1* para outras árvores, os marcadores EMBRA 27 e EMBRA 164 podem ser utilizados com uma expectativa de sucesso de 95,9% para a seleção de plantas carregando o loco *Eef1* assumindo ausência de interferência na recombinação no segmento genômico alvo. Estes mesmos marcadores microssatélites co-dominantes e amplamente transferíveis intra e interespecificamente, vão permitir ainda o teste de alelismo entre o

segmento contendo *Eef1* e o segmento homólogo em outros mutantes para florescimento precoce que venham a ser identificados em eucalipto.

GENETIC MAPPING OF THE EARLY FLOWERING LOCUS *Eef1* (*Eucalyptus* *early flowering 1*) IN EUCALYPTUS USING BULK SEGREGANT ANALYSIS AND MICROSATELLITE MARKERS

Summary

Early flowering phenotype in perennial species is usually seen as a potentially powerful tool to accelerate breeding cycles in recurrent selection programs. An early flowering mutant plant of *Eucalyptus grandis* with normal vegetative growth was found at FERRO GUSA CARAJÁS S.A in Brasil. Flowering starts at approximately 90 days from germination. A cross between a wild type tree and the mutant was carried out, generating a progeny of 88 individuals where the trait segregated in a 1:1 ratio. A genome scan with 100 microsatellite markers distributed across the genome was carried out with fluorescent detection in an ABI 3100 automatic sequencer. A preliminary genotyping survey was done using Bulk Segragant Analysis. Two bulks were created, one with DNA of 15 early flowering individual sibs and 15 wild type sibs. After the bulk genotyping, fluorescence intensities comparison of the parental

EMBRA 27, EMBRA 60, EMBRA 164, EMBRA 65, EMBRA 158 and EMBRA 91, indicating possible linkage with early flowering phenotype. These loci were used for individual genotyping to confirm the co-segregation of the markers and the early flowering phenotype, presenting $p < 0,0001$ values of $9,4 \times 10^{-8}$, $6,23 \times 10^{-15}$, $8,9 \times 10^{-15}$, $1,8 \times 10^{-13}$, and $2,3 \times 10^{-14}$, respectively, confirming the results for the BSA strategy. Not surprisingly, in the reference linkage map developed by Brondani et.al 2001 these loci were located on the same linkage group. A linkage map involving the five loci and the early flowering QTL named *Eef1* (*Eucalyptus early flowering 1*) was constructed using the Outmap software that deals with mixtures of segregation ratios in crosses between heterozygous individuals. The *Eef1* locus was mapped between markers EMBRA 27 and EMBRA 164, with distances of 17,5 cM and 5,8 cM respectively, In introgression experiments, markers EMBRA 27 and EMBRA 164 could be successfully used with an expected precision of 95,9% to select plants carrying the *Eef1* QTL assuming lack of recombination interference in the genomic segment. These same linked and transferable microsatellites will allow preliminary allelism tests between the segment containing *Eef1* and the homologous segment in other early flowering mutants that could be identified in *Eucalyptus*.

4.1 Introdução

O florescimento em plantas superiores envolve a transição do meristema vegetativo, que produz folhas e troncos, em meristema floral gerando a inflorescência. Os dois principais fatores ambientais que promovem o florescimento, são dias longos (fotoperíodo) e exposição a baixas temperaturas (vernalização). Estudos envolvendo análises genéticas e fisiológicas relacionadas com tempo de florescimento têm demonstrado, em *Arabidopsis*, a presença de mais de 80 genes envolvidos na regulação do tempo de florescimento, respondendo a sinais endógenos e mudanças ambientais (Simpson et al., 1999). Esta regulação ocorre através de um conjunto de vias metabólicas, dentre as quais duas respondem a mudanças ambientais (fotoperíodo e vernalização) e outras duas atuam independentemente de sinais ambientais: uma via metabólica autônoma, que promove o florescimento sob quaisquer condições ambientais; e a via da giberelina, um hormônio necessário para o florescimento em dias curtos (Blázquez et al., 2001).

Estudos genéticos realizados com plantas mutantes têm demonstrado que 3 grupos de genes atuam simultaneamente para a formação de estruturas florais e estão organizados em um modelo conhecido como ABC (Coen et al., 1991). Em *Arabidopsis*, a atividade dos genes pertencentes ao grupo A, como

Apetala (AP1), Cauliflower (CAL), e Fruitful (FUL), é responsável pelo desenvolvimento das sépalas. Estes genes, em conjunto com os genes pertencentes ao grupo 2, como Apetala3 e Pistilata (PI), estão envolvidos no desenvolvimento das pétalas. Genes do grupo 2, interagindo com o gene Agamous (AG), pertencente ao grupo 3, são responsáveis pela produção dos estames. Já a expressão do gene AG sozinho induz a produção dos carpelos. Atualmente, com a descoberta de novos genes interagindo com genes já descritos, foi proposto um modelo ABCDE para explicar a expressão do florescimento (Leland et al., 2004). Diversos estudos comprovaram que mutações em genes específicos atuando no sistema reprodutivo da planta podem induzir à expressão precoce do florescimento.

O interesse das empresas plantadoras de eucalipto é o desenvolvimento de um programa de melhoramento que maximize a produtividade e qualidade da madeira para fins industriais. Árvores que florescem precocemente podem, na verdade, vir a apresentar menor produção do que árvores que florescem no tempo normal, tendo em vista a canalização de energia para a produção das flores em uma idade precoce. Nesse sentido, o florescimento precoce pode ser uma característica, em princípio, indesejável em populações de produção embora não existam dados que comprovem esta hipótese. O florescimento precoce, entretanto, pode ser extremamente útil como uma ferramenta para acelerar ciclos de recombinação e seleção em um programa de melhoramento, reduzindo o tempo necessário para a geração de

novas combinações genótípicas e, conseqüentemente, maximizando os ganhos genéticos por unidade de tempo.

Em eucalipto, alguns genes já foram caracterizados como sendo responsáveis pela indução do florescimento. Em um estudo realizado por Decroocq et al. (1998), o gene *TM3-like* foi isolado e caracterizado a partir de bibliotecas de cDNA de botões florais de *Eucalyptus grandis*. Foi observado que este gene é expresso não só em tecidos reprodutivos mas também em tecidos vegetativos. Em um outro trabalho, dois genes de eucalipto (EAP1 e EAP2) foram isolados e caracterizados como homólogos ao gene *Apetala1* (AP1) encontrado em *Arabidopsis*. Foi mostrado, em sistema transgênico, que ambos os genes, EAP1 e EAP2, apresentam efeitos similares ao AP1, induzindo o florescimento precoce (Kyozyuka et al., 1997). Seguindo a mesma linha de pesquisa, os genes (ELF1 e ELF2) foram isolados e caracterizados como homólogos ao gene *Leafy* encontrado em *Arabidopsis* (Southerton et al., 1997). Em um outro estudo, Southerton et al. (1998) isolaram, através de uma biblioteca de cDNA de flores jovens de *Eucalyptus grandis*, 3 genes MADS-box que são expressos exclusivamente em tecidos florais.

Apesar da existência de estudos que identificaram alguns genes relacionados com florescimento em eucalipto, nada foi relatado até o momento sobre a ocorrência de mutantes naturais do eucalipto para florescimento. Conhecimentos anedóticos sobre a observação de mutantes para florescimento em viveiros são comuns entre melhoristas e viveiristas no Brasil, embora

nenhum estudo sistemático tenha sido feito até o momento. Isso não é surpresa uma vez que mutantes para fenótipos de florescimento são comuns em várias outras espécies de plantas. A dificuldade, no caso do eucalipto, se deve possivelmente pela dificuldade de observação deste fenótipo em espécies perenes. A exploração da variabilidade natural do gênero *Eucalyptus* para os genes envolvidos no controle de florescimento representa assim uma estratégia interessante para a identificação de fenótipos potencialmente úteis de alteração de florescimento.

No presente trabalho é relatado o mapeamento de um QTL de maior efeito para florescimento precoce em *Eucalyptus grandis* utilizando uma árvore mutante para florescimento precoce identificada no Brasil e utilizada em programa de melhoramento. A caracterização genética realizada neste estudo envolveu a verificação da herança e segregação e o mapeamento utilizando uma família gerada pelo cruzamento entre o indivíduo mutante de *E. grandis* apresentando a característica de florescimento precoce (AO085) e uma árvore de *E. urophylla* do tipo selvagem para esta característica (UO243). O fenótipo apresenta uma herança dominante típica de uma mutação de ganho de função. O mutante apresentou florescimento ainda no estágio de mudas, na idade de 3 a 4 meses a partir da germinação, enquanto o florescimento normal do eucalipto ocorre de uma forma consistente entre 2 e 4 anos de idade em condições tropicais brasileiras. O mapeamento foi realizado utilizando marcadores microssatélites com detecção fluorescente com base na estratégia

de BSA (*Bulk Segregant Analysis*) (Michelmore et al., 1991) visando, de uma forma rápida e eficiente, localizar marcadores microssatélites ço-segregando com o fenótipo de interesse. A estratégia de BSA tem se mostrado eficiente para estudos envolvendo detecção de QTL, principalmente para genes de herança monogênica como o caso do loco de resistência a *Puccinia* recentemente mapeado em *Eucalyptus grandis* (Junghans et al., 2003). A informação de mapeamento deste gene poderá ser utilizada para a introgressão direcionada deste segmento genômico contendo o alelo mutante em populações de melhoramento visando acelerar ciclos de recombinação, bem como possibilitar o desenvolvimento de linhagens endogâmicas de eucalipto.

4.2 Material e Métodos

4.2.1 Material genético

O experimento de mapeamento foi conduzido com uma progênie F1 de um cruzamento entre dois indivíduos heterozigotos, sendo um deles o mutante AO085 (*Eucalyptus grandis*) e o outro a árvore UO243 (*Eucalyptus urophylla*) de propriedade da empresa CELMAR S.A., atualmente denominada de FERRO GUSA CARAJÁS. A árvore mutante foi detectada em 1998 florescendo aproximadamente aos 3 meses de idade em viveiro (Anderson Piacezzi, comunicação pessoal) e em seguida utilizada como genitor de cruzamentos por apresentar um bom desempenho de produtividade e adaptabilidade nas condições do Maranhão central na região de Imperatriz. O fenótipo de florescimento precoce observado no mutante segregou de forma mendeliana na sua descendência de 88 plantas em proporção aproximada de 1:1 na idade de 3 a 4 meses, indicando, portanto, se tratar de fato de uma mutação geneticamente herdável em heterozigose e não um efeito ambiental. (Figura 1).



Figura 1 - (A) Foto da inflorescência de um descendente da progênie AO085 x UO243 que apresenta florescimento precoce; (B): Foto de um descendente da mesma progênie que não apresenta florescimento precoce

4.2.2 Extração e quantificação de DNA

Para a extração do DNA total, foi utilizado tecido foliar dos genitores e dos 88 indivíduos disponíveis da progênie. A extração foi feita pelo método CTAB 2 % (Doyle & Doyle, 1987). A maceração das folhas foi realizada utilizando-se nitrogênio líquido. A quantificação do DNA extraído foi feita em gel de agarose 1,5% corado com brometo de etídeo utilizando DNA de fago lambda de concentração conhecida como padrão de comparação.

4.2.3 Seleção dos indivíduos a serem genotipados utilizando a estratégia de “BSA”

Para realização deste experimento, 30 descendentes do cruzamento foram selecionados aleatoriamente, sendo 15 expressando a característica de florescimento precoce e 15 não expressando a característica. Após o ajuste das concentrações individuais, dois *bulks* foram montados, um para indivíduos que expressavam a característica de florescimento precoce (*Bulk 1*) e outro para indivíduos que não apresentavam a característica (*Bulk 2*). Para isso, misturou-se quantidades equimolares de 20 μ l de DNA de cada indivíduo, totalizando 300 μ l de DNA a uma concentração final de 1 ng/ μ l para cada *bulk*.

4.2.4 Amplificação dos locos Microssatélites via PCR

Para a realização das genotipagens dos genitores, dos bulks e dos descendentes individualmente, foram selecionados 100 marcadores microssatélites amplamente distribuídos com base em posicionamento no mapa genético referência de *Eucalyptus* (Brondani, 2000; Brondani et al., 2002). Para detecção automatizada dos alelos via fluorescência, um dos primers de cada par dos 100 microssatélites analisados foi marcado, no seu terminal 5' com um fluorocromo 6-FAM, HEX ou NED (Applied Biosystems). O fluorocromo 6-FAM quando excitado por laser, emite uma fluorescência azul, HEX emite fluorescência verde e NED, fluorescência amarela, detectadas com o filtro

virtual D da Applied Biosystems. A amplificação dos microssatélites via PCR foi feita utilizando-se 1X de Tampão de PCR (10 mM Tris-HCl, pH 8.3, 50 mM KCl), 1,5 mM MgCl₂, 0,19 μM de dNTP, 0,24 μg/μl de BSA, 1,0 unidade da enzima *Taq* polimerase, 0,3 μM de cada primer (direto e reverso), 6,0 ng de DNA genômico total. O volume final da reação foi completado para 13 μL com água Milli-Q autoclavada. A PCR foi realizada utilizando um programa que apresenta uma etapa inicial de desnaturação a 94 °C por 5 minutos, 29 ciclos com etapas de desnaturação (1 minuto a 94 °C), anelamento (1 minuto utilizando temperaturas variando de 48 °C a 65 °C de acordo com o loco marcador) e extensão (1 minuto a 72 °C). Ao fim dos ciclos, foi realizada uma etapa de extensão final por mais 30 minutos a 72 °C. A amplificação dos microssatélites via PCR foi feita separadamente para cada par de primer e uma mistura de reações amplificadas de três marcadores, cada um marcado com um fluorocromo diferente, foi feita para detecção simultânea dos fragmentos amplificados. Após a realização da PCR, 1 μl de cada produto de PCR foi aliqotado e adicionado a 1 μl de marcador molecular interno desenvolvido por Brondani et al. (2001) e 10 μl de Hi-Di formamida (Applied Biosystems).

4.2.5 Geração dos dados genotípicos

As eletroforeses realizadas foram conduzidas em analisador automático ABIPrism 3100 da Applied Biosystems. Os dados foram coletados automaticamente pela detecção das diferentes fluorescências e analisados com

os programas GeneScan/Genotyper (Applied Biosystems). Estes programas permitem a genotipagem automatizada dos alelos comparando-os com os valores de um marcador interno.

4.2.6 Análise de co-segregação

Após a genotipagem dos *bulks*, foi realizada a comparação das intensidades de fluorescência dos alelos parentais entre os *bulks* contrastantes para o fenótipo de florescimento. A observação de uma variação de intensidade de fluorescência de alelos específicos entre os *bulks* indicou uma possível co-segregação entre o marcador microssatélite e um QTL de maior efeito envolvido no controle do florescimento precoce. Desta forma, aqueles marcadores que apresentaram diferença de intensidade de fluorescência para seus alelos entre os *bulks* foram selecionados para genotipagem individual de toda a progênie visando a verificação da possível ligação marcador-fenótipo.

4.2.7 Construção do mapa genético

Os dados genotípicos dos marcadores microssatélites selecionados na etapa de BSA foram analisados inicialmente por um teste de Qui-quadrado visando testar a hipótese nula de segregação 1:1 em cada um dos grupos de fenótipo extremo. Para a criação de um grupo de ligação, foi utilizado o programa JoinMap (Stam, 1993), que permite a construção de mapas genéticos com base em marcadores com diferentes razões de segregação em

cruzamentos entre indivíduos heterozigotos. Foi utilizado um LOD 8,0 e fração de recombinação 0,4 como parâmetros de estringência. Para ordenação das marcas, o programa utiliza automaticamente o comando *ripple* após a adição de cada marca. As distâncias entre os marcadores foram calculados em centimorgans (cM) utilizando a função Kosambi.

4.3 Resultados

4.3.1 Quantificação de DNA e montagem dos *Bulks*

Após a extração de DNA dos indivíduos selecionados para compor os *bulks*, o DNA foi diluído a uma concentração final de 1 ng/ μ l e quantificado antes da montagem dos *bulks* (Figura 2A e 2B). Após o ajuste das concentrações individuais, os dois *bulks* foram montados. Para certificar que os dois *bulks* continham a mesma concentração de DNA, uma nova quantificação foi realizada com o material pertencente a cada *bulk* (Figura 3).

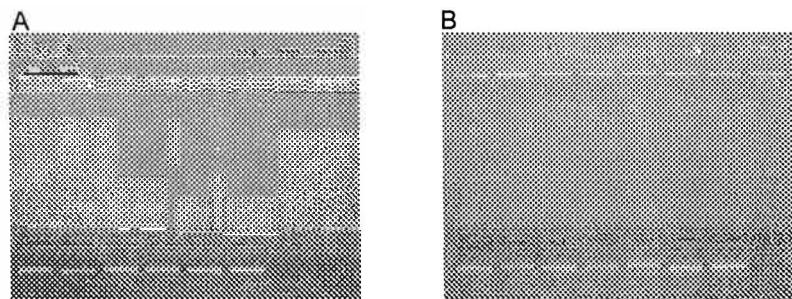


Figura 2 - Eletroforese em gel de agarose 1,5% visualizada com brometo de etídeo para quantificação de DNA de: (A) 14 indivíduos que foram utilizados para montagem do *bulk* 1 (plantas expressando a característica de florescimento precoce e (B) para 15 indivíduos utilizados para montagem do *bulk* 2 (plantas que não expressavam esta característica)

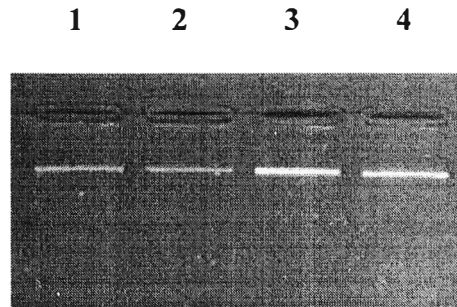
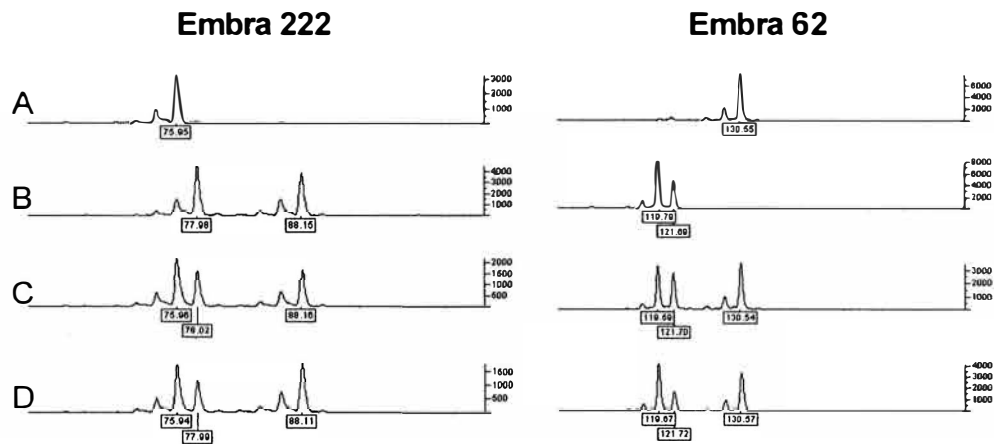


Figura 3 - Eletroforese em gel de agarose 1,5% visualizada com brometo de etídeo para quantificação do DNA dos *bulks*. Poços 1 e 3: *Bulk 1*; poços 2 e 4: *Bulk 2* (10 ng e 20 ng, respectivamente)

4.3.2 Geração dos dados genotípicos

Dos 100 marcadores microsatélites analisados, 6 apresentaram diferenças significativas quando comparada a intensidade de fluorescência dos picos dos alelos observados nos *bulks*. A figura 4 apresenta a genotipagem de dois destes locos (EMBRA 91 e EMBRA 158), assim como a genotipagem de dois locos que não apresentaram evidências de ligação com a característica (EMBRA 222 e EMBRA 62).

1



2

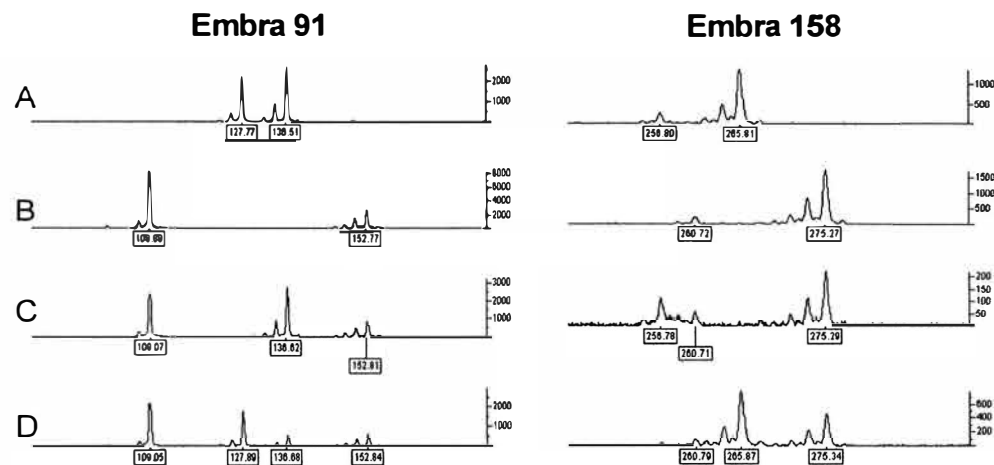


Figura 4 - Genotipagem com detecção fluorescente realizada utilizando o programa Genotyper (Applied Biosystems) para os indivíduos (A) AO085 (mutante); (B) UO243 (tipo selvagem); (C) para o *bulk* 1 (indivíduos que apresentam florescimento precoce); e (D) para o *bulk* 2 (indivíduos que não apresentam florescimento). (1) Genotipagem dos locos EMBRA 222 e 62 que não apresentaram diferenças de intensidade de fluorescência dos alelos entre os *bulks*. (2) Genotipagem dos locos EMBRA 91 e 158 e que mostraram diferenças de intensidade de fluorescência para os alelos entre os *bulks*

4.3.3 Análise de co-segregação entre o fenótipo de florescimento precoce e os marcadores microssatélites selecionados.

Dos 100 locos analisados, os locos EMBRA 91, EMBRA 60, EMBRA 65, EMBRA 164, EMBRA 185 e EMBRA 27 apresentaram uma forte distorção na segregação de seus alelos quando os indivíduos dos dois *bulks* foram analisados. Estes locos foram utilizados na genotipagem dos indivíduos da progênie visando confirmar o padrão de co-segregação dos marcadores e o fenótipo de florescimento precoce. A tabela 1 apresenta o número de indivíduos para cada alelo de cada marcador microssatélite para cada classe fenotípica, valores p de níveis de significância do teste χ^2 visando testar a hipótese alternativa de distorção de segregação, e o alelo do genitor mutante associado ao gene. Os dados da genotipagem individual estão no CD em anexo.

Tabela 1. Número de indivíduos para cada alelo de cada marcador microssatélite para cada classe fenotípica, valores p de níveis de significância do teste χ^2 , e o alelo do genitor mutante associado ao gene para os marcadores EMBRA 27, 60, 164, 158, 91 e 65

	Marcador Microssatélite											
	EMBRA 27		EMBRA 60		EMBRA 164		EMBRA 65		EMBRA 158		EMBRA 91	
Alelos (pb)	136	153	106	127	133	135	209	Nulo	141	151	128	137
Flores. Precoce (n ₁)	4	28	1	29	32	0	0	32	1	25	1	33
Sem florec. (n ₂)	34	12	46	4	5	41	43	9	24	4	41	5
Valores p de significância	9,4x10 ⁻⁸		6,2x10 ⁻¹⁵		8,9x10 ⁻¹⁵		1,8x10 ⁻¹³		1,7x10 ⁻⁹		2,3x10 ⁻¹⁴	
Alelo em acoplamento com o QTL (pb)	153		127		133		Nulo		151		137	

4.3.4 Construção do mapa genético

A construção do mapa genético foi realizada utilizando o programa JoinMap utilizando um *LOD score* de 8,0. Os dados fenotípicos foram organizados de forma binária (florescimento=1 e não florescimento=2) de maneira que pudesse ser analisada a co-segregação do fenótipo de florescimento precoce com os marcadores mapeados no grupo de ligação (Figura 5). Dos seis marcadores microssatélites, apenas o EMBRA 158 apresentou distorção de segregação de seus alelos, apresentando valor de χ^2 de 7,2.

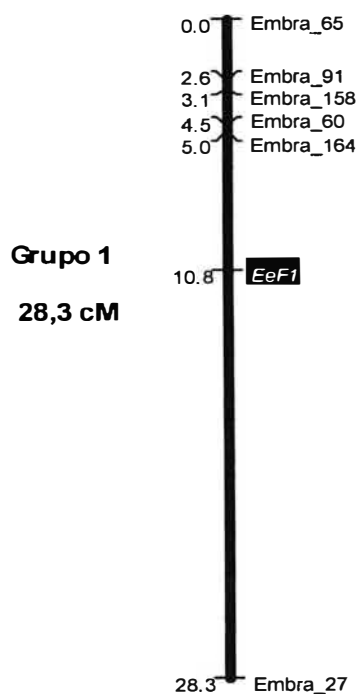


Figura 5 - Grupo de ligação gerado pelo programa JoinMap para 6 locos microssatélites e a mais provável localização do QTL denominado *Eef1* (*Eucalyptus early flowering 1*) envolvido na expressão do fenótipo de florescimento precoce

4.3.5 Simulação de seleção assistida por marcadores para o fenótipo de florescimento precoce

Os dados genéticos permitem inferir que no mutante, os locos EMBRA 27 e EMBRA 60 apresentam os alelos A1 (para o loco EMBRA 27) e B1 (para o loco EMBRA 60) em acoplamento flanqueando o alelo mutante responsável pelo fenótipo de florescimento precoce localizado em um dos cromossomos homólogos (identificado pelo número 1) e os alelos A2 (para o loco EMBRA 27) e B2 (para o EMBRA 60) flanqueando o alelo selvagem no cromossomo homólogo (2) (Figura 6). Estão representados ainda os dois cromossomos homólogos e os gametas formados na meiose com suas respectivas frequências, estimadas a partir das distâncias de mapa, assumindo a ausência de interferência (Figura 6). Para verificar a posição de mapa obtida, foi simulada uma seleção na progênie daqueles indivíduos que apresentavam os alelos A1/B1. Estes indivíduos deveriam apresentar o fenótipo para florescimento precoce. De um total de 76 indivíduos, 28 foram selecionados com base no haplótipo A1/B1. Destes 28 indivíduos, 25 apresentaram o fenótipo de florescimento precoce e 3 indivíduos mesmo apresentando os alelos A1/B1, não expressavam esta característica, possivelmente oriundos de gametas duplos recombinantes herdados a partir do mutante. Por outro lado, quatro indivíduos da progênie com haplótipos A2/B1, A1/B2 ou A2/B2 também apresentaram o fenótipo de florescimento precoce.

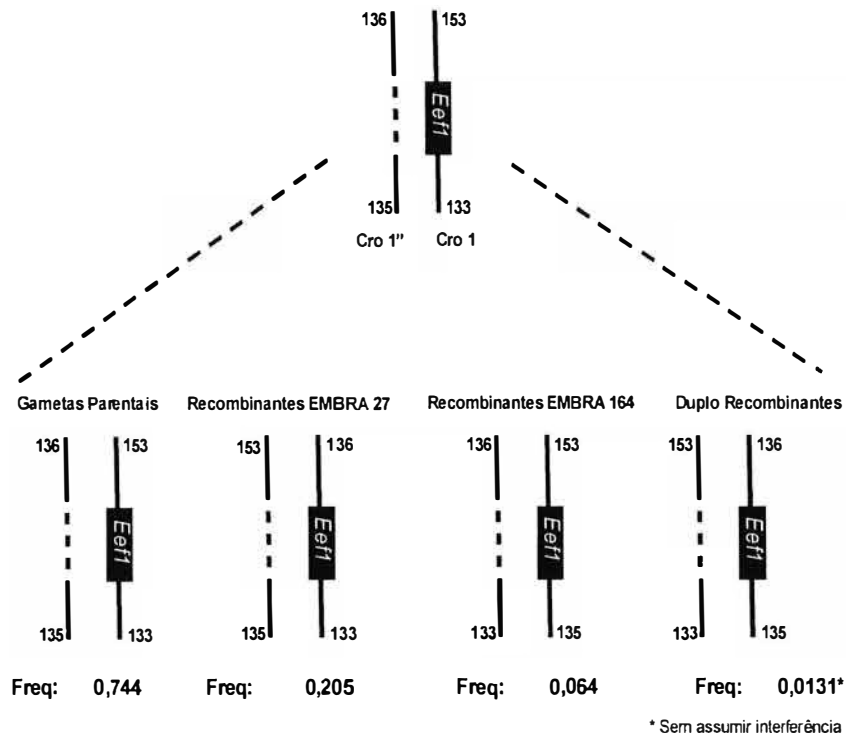


Figura 6 - Representação esquemática dos cromossomos homólogos (1 e 1'), gametas parentais e recombinantes, com suas respectivas frequências, formados na meiose do indivíduo mutante para florescimento precoce heterozigoto para o gene *Eef1*

4.4 Discussão

Dados genotípicos de 30 indivíduos da progênie (15 indivíduos apresentando o fenótipo de florescimento precoce e 15 indivíduos tipo selvagem) foram utilizados para uma triagem preliminar de 100 locos microssatélites cobrindo aproximadamente 90% do genoma a uma distância média de 15 centiMorgans. Os dados genotípicos de marcadores microssatélites foram analisados por um teste de Qui-quadrado visando testar a hipótese de distorção de segregação entre os dois grupos de fenótipo extremo. Dos 100 locos analisados, os locos EMBRA 65, EMBRA 91, EMBRA 158, EMBRA 60, EMBRA 164 e EMBRA 27 apresentaram uma forte distorção na segregação de seus alelos quando os dois grupos de indivíduos foram analisados, apresentando valores p altamente significativos, variando de $9,4 \times 10^{-8}$ a $8,9 \times 10^{-15}$ (Tabela 1). Estes locos, quando posicionados no mapa de ligação desenvolvido por Brondani (2000) e Brondani et al. (2002) encontram-se localizados em um mesmo grupo de ligação, grupo 2. Estes dados indicam que no grupo de ligação 2 está localizado um QTL para florescimento precoce em *Eucalyptus*, denominado de *Eef1* (*Eucalyptus early flowering locus 1*). É importante ressaltar que nenhum outro dos 100 marcadores microssatélites testados, cobrindo grande parte do genoma, cerca de 90%, apresentou

evidência de ligação com o fenótipo de florescimento na análise de BSA. Isto sugere que *Eef1* deve ser responsável pela maior parte da variação no fenótipo de florescimento precoce observado. A identificação de um QTL de maior efeito não descarta a possibilidade de existirem outros locos de menor efeito contribuindo para a penetrância deste fenótipo mutante ou para a uma possível variação na expressividade. A avaliação fenotípica do florescimento precoce foi realizada de forma simples e direta com base em padrões operacionais de melhoramento genético e em uma mesma idade de aproximadamente 4 meses. Não foram conduzidas avaliações da variação na intensidade do florescimento como, por exemplo, número de botões florais nem da variação na precocidade da expressão do florescimento. Para estes outros fenótipos componentes, do que foi chamado globalmente de florescimento precoce, é possível que outros QTL possam ser identificados.

A combinação de marcadores microssatélites e a abordagem de BSA se mostrou factível e potente para a localização do loco *Eef1*. BSA e marcadores dominantes RAPD tem sido amplamente usados para a identificação rápida de alelos de alta penetrância a QTL de maior efeito principalmente para resistência a patógenos. Em árvores, esta abordagem tem sido usada para o mapeamento do primeiro loco de resistência a doenças em eucalipto, o loco de resistência à ferrugem de *Puccinia psiidi Ppr1* em *Eucalyptus grandis* (Junghans et al., 2003). Assim como para locos de resistência em *Populus deltoides* (Cervera et al., 1996; Tabor et al., 2000), *P.*

lambertiana (Devey et al., 1995), *P. taeda* (Wilcox et al., 1996) *Ulmus parvifolia* (Benet et al., 1995).

Neste trabalho, entretanto, o fato que o loco *Eef1* foi mapeado com marcadores microssatélites co-dominantes transferíveis, permite o teste de alelismo entre o segmento contendo *Eef1* e o segmento homólogo em outros mutantes para florescimento precoce que venham a ser identificados futuramente. É importante ressaltar a este respeito que o controle da idade de florescimento, por envolver várias dezenas de genes em plantas, poderá ser afetado por mutações em outros genes que não aquele(s) que estão por trás do QTL *Eef1*.

Para programas de introgressão assistida do QTL *Eef1* para outras árvores de um programa de melhoramento, os marcadores EMBRA 27 e EMBRA 164 podem ser utilizados com sucesso apesar deles estarem localizados a distâncias relativamente grandes de 17,5 e 5,8 cM respectivamente. Na seleção simulada realizada com os indivíduos que apresentaram o haplótipo 153/133, três indivíduos não apresentaram o fenótipo de florescimento precoce (4,10%). Uma parte destes “falsos positivos” (1,3 %) é esperada pela ocorrência de duplos recombinantes no segmento entre EMBRA 27 e EMBRA 164 assumindo interferência zero. O restante (cerca de 2,8 %) pode ser devido à penetrância incompleta ou expressividade variável do QTL *Eef1* na idade de avaliação utilizada no experimento. A penetrância incompleta ou expressividade variável pode ser resultado da interação com o ambiente

e/ou da influência de outros genes modificadores de menor efeito não controlados neste estudo.

Com o objetivo de se obter um marcador mais próximo do loco *Eef1* co-segregando de forma perfeita, uma ampla bateria de primers RAPD (*Random Amplified Polimorphism DNA*), foi utilizada para uma triagem de polimorfismos utilizando ambos os parentais e os *bulks* na estratégia de BSA. Utilizando marcadores RAPD em acoplamento com o alelo mutante e, de preferência, a 0 % de recombinação (estimado em algumas centenas de meioses) seria possível otimizar a o procedimento de introgressão assistida. Além disso, um marcador com essas características seria um importante ponto de partida para a busca do gene em uma biblioteca de BAC (*Bacterial Artificial Chromosome*) previamente desenvolvida para o indivíduo mutante. Porém, mesmo após uma triagem utilizando cerca de 450 primers RAPD e cerca de 2000 marcadores amplificados, não foi possível detectar e confirmar nenhum marcador RAPD apresentando uma forte ligação com o gene que controla este fenótipo (Bueno, comunicação pessoal).

A informação de mapeamento gerada, merece um maior refinamento. Para os marcadores já mapeados, seria interessante a análise de um maior número de descendentes de um cruzamento envolvendo o mutante AO085 para melhorar a estimativa de recombinação entre os marcadores e *Eef1*. Uma saturação com maior número de microssatélites permitiria ainda a detecção de marcadores mais próximos melhorando a eficiência de um futuro exercício de

introgressão assistida. Finalmente, seria interessante cruzar o mutante A0O85 com diversas outras árvores do tipo selvagem de diferentes espécies de *Eucalyptus* para verificar a penetrância e expressividade desta mutação em distintos backgrounds genéticos.

Neste trabalho, nenhum esforço foi dedicado à co-localização de genes candidatos, já conhecidos por seu papel no controle da idade de florescimento com o QTL *Eef1*. Na continuidade desta investigação, seria oportuno este exercício de co-localização visando explorar a possível identificação do gene ou genes efetivamente envolvidos na expressão deste fenótipo de florescimento precoce. A definição de quais genes candidatos selecionar entre as dezenas de genes conhecidos envolvidos no controle do florescimento não é, entretanto, tarefa trivial, pois outros genes, ainda não conhecidos, agindo em trans podem estar envolvidos, os quais dificilmente seriam selecionados com base no conhecimento atual da regulação da idade de florescimento.

4.5 Conclusão

O trabalho teve como objetivo o mapeamento de um fenótipo de alta penetrância para florescimento precoce segregando de forma mendeliana em *Eucalyptus grandis* utilizando estratégia de BSA. Dado um fenótipo de herança simples e fácil avaliação, esta estratégia, permite a rápida identificação de QTL para características de interesse no melhoramento. Isso tem sido comprovado em diversas espécies arbóreas para locos de resistência a patógenos e neste trabalho, de forma pioneira, para um QTL controlando o fenótipo de florescimento. A estratégia de BSA para detecção de QTL para características mais complexas, como teor de lignina e densidade básica da madeira, embora possível (Grattapaglia et al., 1996) resulta na identificação de apenas um dos diversos QTL tipicamente envolvidos em características efetivamente multifatoriais. Um ponto crucial para o sucesso da estratégia de BSA está na qualidade e na concentração do DNA dos indivíduos envolvidos na montagem dos bulks. Isso possivelmente se torna mais crítico ainda ao se utilizar marcadores microssatélites e inferência de ligação marcador-fenótipo utilizando a análise comparativa da intensidade de pico de fluorescência entre os bulks contrastantes. As amostras de DNA dos indivíduos selecionados foram rigorosamente quantificadas antes de serem agrupadas e, após a montagem

dos bulks, houve uma nova quantificação para a certificação de que os *bulks* estavam na mesma concentração. Caso o DNA de algum indivíduo previamente selecionado para compor o *bulk* estivesse a uma concentração diferente dos demais, a intensidade relativa da fluorescência dos picos dos seus alelos amplificados via PCR também seria diferente, podendo levar a falsos positivos ou negativos, dificultando a interpretação dos dados.

Para a execução deste trabalho a estratégia de BSA, seguida de genotipagem com detecção fluorescente apresentou-se bastante satisfatória, detectando de forma rápida e eficiente região do genoma onde a mutação responsável pelo florescimento precoce estava localizada. Desta forma, esta estratégia pode ser utilizada com sucesso para a realização de trabalhos envolvendo a localização de genes de herança qualitativa. Para investigações mais complexas, envolvendo genes de herança quantitativa, as incorporações de ferramentas estatísticas podem ser utilizadas para um melhor refinamento das análises.

5 MAPEAMENTO DE QTL PARA CARACTERÍSTICAS DA MADEIRA EM CLONES HÍBRIDOS DE *Eucalyptus* SP. TROPICAL COM BASE EM MICROSSATÉLITES

Resumo

A construção de mapas genéticos e localização de QTL (*Quantitative Trait Loci*) economicamente relevantes tem sido realizada em eucalipto com base em marcadores dominantes RAPD e AFLP. Estes marcadores apresentam limitações quanto ao conteúdo informativo e, principalmente, quanto à transferibilidade de marcas entre diferentes mapas e pedigrees não permitindo estudos comparativos e validações de posicionamento de QTL. Marcadores microssatélites, por outro lado, apresentam uma ampla distribuição ao longo do genoma e maior fidelidade na transferência de locos entre indivíduos e pedigrees permitindo, assim, a validação e estudos detalhados de variabilidade alélica a QTL. Este trabalho teve por objetivo a construção de um mapa genético utilizando marcadores microssatélites, a localização de QTL para características físico-químicas da madeira e a demonstração da utilidade de microssatélites para análises comparativas de mapeamento entre experimentos independentes, particularmente, quanto à localização de QTL. A

construção do mapa genético foi realizada a partir da genotipagem de uma progênie de 188 indivíduos de um cruzamento entre híbridos naturais superiores de *E. grandis* com *E. urophylla* oriundos do Horto de Rio Claro. Foram genotipados 169 microssatélites, em sistemas multiplex com detecção automatizada, que, posteriormente, foram mapeados em 11 grupos de ligação. Mapas genéticos individuais foram construídos para cada genitor, utilizando o programa Mapmaker/EXP. Para a construção de um mapa integrado, foi utilizado o programa JoinMap. Para o genitor materno clone 235, foi construído o mapa genético com 102 marcadores microssatélites e 819,6 cM, cobriu estimados 74,6% do genoma. Para o genitor paterno clone 221, o mapa incluiu 92 marcadores 536,4 cM e cobriu 66,67% do genoma. O mapa integrado, com 155 marcas, com 851 cM cobriu 78% do genoma. A análise de detecção de QTL foi realizada utilizando diferentes análises estatísticas com os programas Mapmaker/QTL e QTL Cartographer. Para o genitor materno 235, foram detectados 10 QTL com valores de LOD variando de 2,9 (teor de lignina) a 4,2 (consumo específico). Para o genitor paterno 221, foram detectados 5 QTL com valores de LOD variando de 2,9 (rendimento de celulose) a 4,8 (densidade básica). Mapeamento comparativo de QTL detectados neste trabalho com outros QTL e genes candidatos da via de lignificação mapeados em estudos independentes, revelaram diversos casos de sintonia de QTL para rendimento de celulose, teor de lignina e propriedades de fibra. Estes resultados são inéditos para *Eucalyptus* e apontam para perspectivas de validação de QTL e detecção direcionada de variação alélica em outros pedigrees visando a

determinação precisa de segmentos genômicos alvo para experimentos de mapeamento de alta resolução, testes de associação e, futuramente, seleção assistida por marcadores.

QTL MAPPING OF WOOD PROPERTIES IN TROPICAL EUCALYPTUS HYBRID CLONES USING MICROSATELLITES

Summary

The construction of genetic maps and detection of QTL (Quantitative Trait Loci) for economically relevant traits has been accomplished in eucalyptus mainly with dominant RAPD and AFLP markers. These markers have limited information content and are useless for comparative mapping studies and QTL validation with unrelated maps and pedigrees. Microsatellite markers, on the other hand can easily be transferred among maps therefore allowing validation of QTL position. This work describes the construction of a genetic map based exclusively on microsatellites and the detection of QTL for several wood properties. It also shows the possibility of performing comparative map analysis between independent experiments, specifically involving QTL location. Genetic maps were constructed using a set of 188 F1 individuals from a cross between two elite Rio Claro natural hybrid trees involving predominantly *E. grandis* and *E. urophylla*. A set of 169 microsatellite markers were genotyped in multiplex systems with automated detection and mapped on 11 linkage groups. Individual genetic maps were built for each parent using Mapmaker/EXP and an integrated

map with JoinMap. For parent clone 235 102 markers were mapped on 819,6 cM covering an estimated 74,6% of the genome. For parent clone 221, the map had 92 microsatellites, 536,4 cM covering 66,7% of the genome. The integrated map involved 155 markers, 851 cM covering 78% fo the genome. QTL detection was accomplished under different models with Mapmaker/QTL and QTL Cartographer. For hybrid parent tree 235 10 QTL were detected, with LODs varying between 2,9 (lignin content) to 4,2 (specific wood consumption). For hybrid parent 221, five QTL were detected with LODs varying from 2,9 (cellulose yield) to 4,8 (basic wood density). Comparative QTL mapping with QTL and candidate genes from other studies revealed a number of syntenic QTL for cellulose yield, lignin content and for different but correlated fiber traits as well as candidate genes fo the lignification pathway. These are novel results for *Eucalyptus* and point to exciting perspectives for QTL validation, directed search for allelic variants at QTL in multiple pedigrees and the precise determination of target genomic regions for association mapping studies, high resolution mapping and eventually marker assisted selection.

5.1 Introdução

Nos últimos anos tem havido um interesse crescente por parte das empresas florestais na utilização de técnicas de marcadores moleculares para auxiliar os programas de melhoramento. Informações de distância genéticas entre genitores têm sido utilizadas operacionalmente por algumas empresas para a organização e gerenciamento da variabilidade genética e otimização de esquemas de cruzamento. Marcadores moleculares de diferentes classes tem sido propostos e utilizados com sucesso para resolver questões de identidade clonal e parentesco em eucalipto (Keil & Griffin, 1994; Grattapaglia et al., 2004 ; Kirst et al., 2004) A construção de mapas de ligação e localização de QTL (Quantitative Trait Loci) economicamente relevantes, por outro lado, ainda constitui tópico de pesquisa e desenvolvimento e ainda não foi convertida em aplicações operacionais em seleção assistida por marcadores (Grattapaglia, 2000).

Os estudos de mapeamento de QTL realizados com Eucalipto têm demonstrado a possibilidade de detectar regiões genômicas controlando características de herança complexa. Diversos grupos relataram a identificação de regiões genômicas que têm efeito significativo na expressão de características economicamente importantes em *Eucalyptus*. QTL para

características juvenis, tais como a altura de plântulas, área foliar e tolerância da muda à geada foram mapeados (Vaillancourt et al., 1995; Byrne et al., 1997). Características relacionadas à habilidade de propagação vegetativa, tais como enraizamento adventício, brotamento de toco e multiplicação de plântulas in vitro também foram detectados (Grattapaglia et al., 1995; Marques et al., 1999). Finalmente, QTL importantes também foram encontrados para características em idade de rotação, tais como crescimento volumétrico, densidade básica da madeira, espessura da casca e forma do caule (Grattapaglia et al., 1996; Verhaegen et al., 1997), características físico químicas da madeira (Myburg 2001; Moran et al., 2002; Thamarus et al., 2004), QTL para resistência a insetos e características responsáveis pela produção de óleos essenciais foram mapeados (Shepherd et al., 1995) e recentemente um QTL de maior efeito para resistência à ferrugem de *Puccinia psidii* foi encontrado e mapeado em *E. grandis* (Junghans et al., 2003). Uma característica comum observada é a presença de um número relativamente pequeno de locos sendo responsáveis por 10 a 40% de toda a variação fenotípica de características de herança quantitativa.

Recentemente QTL de expressão gênica (eQTL) foram mapeados para características de crescimento em *Eucalyptus* utilizando microarranjos contendo cerca de 2600 genes e um mapa construído com marcadores AFLP (Kirst et al. 2004). Dois genes foram detectados, co-localizados com QTL para crescimento sugerindo que as mesmas regiões genômicas regulam crescimento

e teor de lignina Em uma análise global de QTL, Kirst et al. (2005) identificaram ainda eQTL para 1067 dos 2608 genes sendo que para 821 genes um único eQTL foi identificado explicando 70% da variação em expressão gênica. Estes trabalhos demonstram o poder da integração de análise de expressão gênica e mapeamento de ligação para entender a complexidade de características multifatoriais.

Apesar dos avanços conceituais gerados por estes estudos, todos eles enfrentaram limitações técnicas que dificultam a transferência das informações geradas para situações operacionais. Algumas destas limitações são: o número limitado de descendentes analisados, a precisão na coleta dos dados fenotípicos e a herança dominante e baixa transferibilidade dos marcadores utilizados para gerar os mapas genéticos, tipicamente marcadores RAPD e AFLP sobre os quais os QTL são localizados. Com isso, a posição exata e particularmente a variabilidade alélica aos QTL encontrados e a magnitude relativa do efeito do QTL detectado ainda demandam estudos mais detalhados antes que esta tecnologia possa efetivamente ser utilizada para seleção assistida.

Um dos aspectos técnicos fundamentais para avançar rumo ao um entendimento mais detalhado das características complexas em *Eucalyptus* é a utilização de marcadores moleculares geneticamente mais informativos que permitam detectar e validar QTL através de comparações entre experimentos independentes, bem como acompanhar os vários alelos existentes aos QTL nas

populações de melhoramento. Dentre as classes dos marcadores moleculares disponíveis hoje, os microssatélites, são, indubitavelmente, os mais informativos e talvez os únicos que atendam à esta demanda. Em um trabalho pioneiro, Brondani et al. (1998, 2000; 2002) desenvolveram amplas baterias de marcadores microssatélites em *Eucalyptus* e finalizaram um mapa genético referência composto exclusivamente de marcadores microssatélites cobrindo mais de 95% do genoma de *Eucalyptus*. Este mapa “referência” útil para a grande maioria das espécies comercialmente importantes de *Eucalyptus* permite a rápida disseminação do uso destes marcadores em análises genéticas de *Eucalyptus*, facilitando a troca e comparação de informação entre laboratórios, particularmente quanto ao mapeamento e validação de QTL entre experimentos independentes (Grattapaglia, 2000). Estas possibilidades técnicas, entretanto, ainda não foram testadas em eucalipto.

Este trabalho apresenta a detecção e mapeamento de QTL para características de interesse econômico como densidade básica da madeira, consumo específico, rendimento de celulose, *coarseness*, fibra/grama, teor de pentosana e teor de lignina. Diferentemente dos trabalhos publicados até hoje, os QTL mapeados neste estudo foram localizados utilizando marcadores microssatélites (di e trinucleotídeos) de ampla transferibilidade entre espécies de Eucalipto. Os objetivos específicos deste trabalho foram: (1) construir um mapa genético *de novo* utilizando uma bateria de marcadores microssatélites previamente mapeados por Brondani (2000) e comparar a ordem relativa de

marcadores nos dois mapas; (2) mapear QTL para oito características quantitativas de importância econômica, incluindo seis derivadas de mensurações de propriedade física e químicas da madeira geradas por espectrometria de infravermelho próximo; (3) comparar o poder de detecção de QTL com tamanhos variáveis de amostras fenotipadas da progênie segregante; (4) comparar as localizações dos QTL mapeados neste estudo com aquelas dos QTL e genes candidatos mapeados em estudos anteriores em eucalipto.

5.2 Material e Métodos

5.2.1 Material genético

A população de mapeamento foi derivada de um cruzamento entre dois híbridos naturais de *Eucalyptus*, sendo o clone 221 utilizado como genitor paterno e o clone 235 como genitor materno. Ambos são clones elite de altíssima produtividade e adaptação utilizados operacionalmente em plantios comerciais em Espírito Santo, Brasil. Estes dois clones são derivados de sementes colhidas no Horto de Rio Claro (SP) e apresentam características fenotípicas sugestivas de possuir base genética predominante de *E. grandis* e *E. urophylla* respectivamente para os genitores 221 e 235. O cruzamento foi realizado na ARACRUZ-CELULOSE S.A em abril de 1998 resultando em uma descendência de 270 indivíduos levados ao campo em delineamento experimental em junho de 1999, visando à maximização da segregação e potencial recuperação de híbridos transgressivos. Para o experimento de mapeamento foram tomados aleatoriamente 188 descendentes. (Figura 1).



Figura 1 - Foto do experimento de campo mostrando a população de descendentes do cruzamento entre híbridos naturais de *E. grandis* x *E. urophylla* utilizada para o mapeamento de QTL

5.2.2 Levantamento dos dados fenotípicos

Dados fenotípicos de CAP (Circunferência na Altura do Peito) e de densidade básica da madeira foram levantados em campo aos dois anos de idade, durante o mês de Julho de 2001, para os 188 indivíduos da população de mapeamento. Para densidade básica da madeira foram obtidas medidas de penetrância em milímetros com o equipamento Pilodyn, um dinamômetro que estima a densidade básica da madeira com base na penetração de um elemento metálico previamente calibrado (Figuras 2A e 2B). Dados de teor de lignina (%), teor de pentosana (%), Número de fibras/grama, *Coarseness*, correspondendo ao peso em mg de 100 metros de fibra, (característica relacionada a espessura de fibra, densidade e número de fibras por grama), rendimento em celulose (%) e consumo específico (m^3 de madeira/ton de celulose) foram coletados para 80 dos 188 indivíduos da progênie com base em

espectros de Infravermelho próximo (NIRS – *Near Infrared Spectroscopy*). Para isso, foram coletadas baguetas de madeira retiradas com trado na altura do peito (cerca de 1,30 m do chão) (Figura 3). Para a análise NIRS foi utilizado o procedimento previamente desenvolvido e otimizado com base na calibração existente no equipamento da Aracruz Celulose S.A (G. Dehon comunicação pessoal).

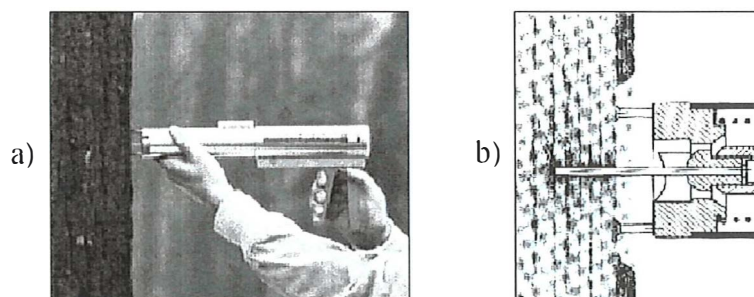


Figura 2 - a: Foto do aparelho pilodyn utilizado para fazer as mensurações de densidade da madeira; b: Ilustração do princípio de funcionamento do aparelho. que estima a densidade básica da madeira com base na penetração de um elemento metálico previamente calibrado



Figura 3 - Coleta de baguetas de madeira com um trado para análises físico químicas da madeira

5.2.3 Extração e quantificação de DNA

A extração foi feita a partir de folhas expandidas dos 188 descendentes e dos genitores utilizando o método padrão baseado no detergente CTAB a 2% (Doyle & Doyle 1987). A maceração das folhas foi realizada utilizando-se nitrogênio líquido. A quantificação do DNA extraído foi feita em gel de agarose 1% em tampão TBE corado com brometo de etídeo utilizando DNA de fago lambda de concentração conhecida como padrão de comparação.

5.2.4 Amplificação dos locos microssatélites via PCR

A amplificação dos microssatélites via PCR (Reação em Cadeia da Polimerase) foi feita utilizando-se 1 X de Tampão de PCR (10 mM Tris-HCl, pH 8.3, 50 mM KCl), 1,5 mM MgCl₂, 0,19 μM de dNTP, 0,24 μg/μl de BSA, 1,0 unidade da enzima *Taq* polimerase (Invitrogen), 0,3 μM de cada iniciador (primer) (direto e reverso), 6,0 ng de DNA, e 5,375 μL de água Milli-Q autoclavada para um total de 13 μl de reação final. A PCR foi realizada utilizando um programa que apresenta uma etapa inicial de desnaturação a 94 °C por 5 minutos, seguida de 29 ciclos envolvendo uma etapa de desnaturação (1 minuto a 94 °C), uma de anelamento (1 minuto utilizando temperaturas variando de 48 °C a 65 °C) e uma de extensão (1 minuto a 72 °C) e finalizando com uma etapa única de extensão final por mais 30 minutos a 72 °C.

5.2.5 Triagem e seleção de locos microssatélites

Foram utilizados 326 marcadores microssatélites, sendo que 196 desenvolvidos por Brondani (2000) e Brondani et al. (2002), 40 marcadores desenvolvidos pela CSIRO (<http://www.ffp.csiro.au/tigr/molecular/eucmsps.html>), 87 microssatélites de distribuição aleatória gerados a partir de clones genômicos “shot-gun” e três microssatélites derivados de clones ESTs (Lourenço et al., 2004). Esses microssatélites foram submetidos a um processo de triagem na busca de polimorfismo entre os dois genitores. Para isso foram utilizados, além dos genitores, dois indivíduos da progênie para verificar a herança e confirmar a presença de polimorfismo.

5.2.6 Geração dos dados genotípicos

Para detecção automatizada dos genótipos aos microssatélites, um dos iniciadores de cada um dos marcadores microssatélites utilizados foi marcado, no seu terminal 5' com um dos fluorocromos 6-FAM, HEX ou NED (Applied Biosystems). O fluorocromo 6-FAM emite uma fluorescência azul, o fluorocromo HEX emite uma fluorescência verde e o fluorocromo NED emite uma fluorescência amarela no filtro virtual D da Applied Biosystems (Figura.4). A eletroforese dos marcadores foi conduzida em sistemas triplex com um marcador em cada uma das três diferentes cores no analisador automático ABI 3100 da Applied Biosystems em placas de 96. Os dados foram coletados

automaticamente pela detecção das diferentes fluorescências e analisados com os programas GeneScan/Genotyper (Applied Biosystems).

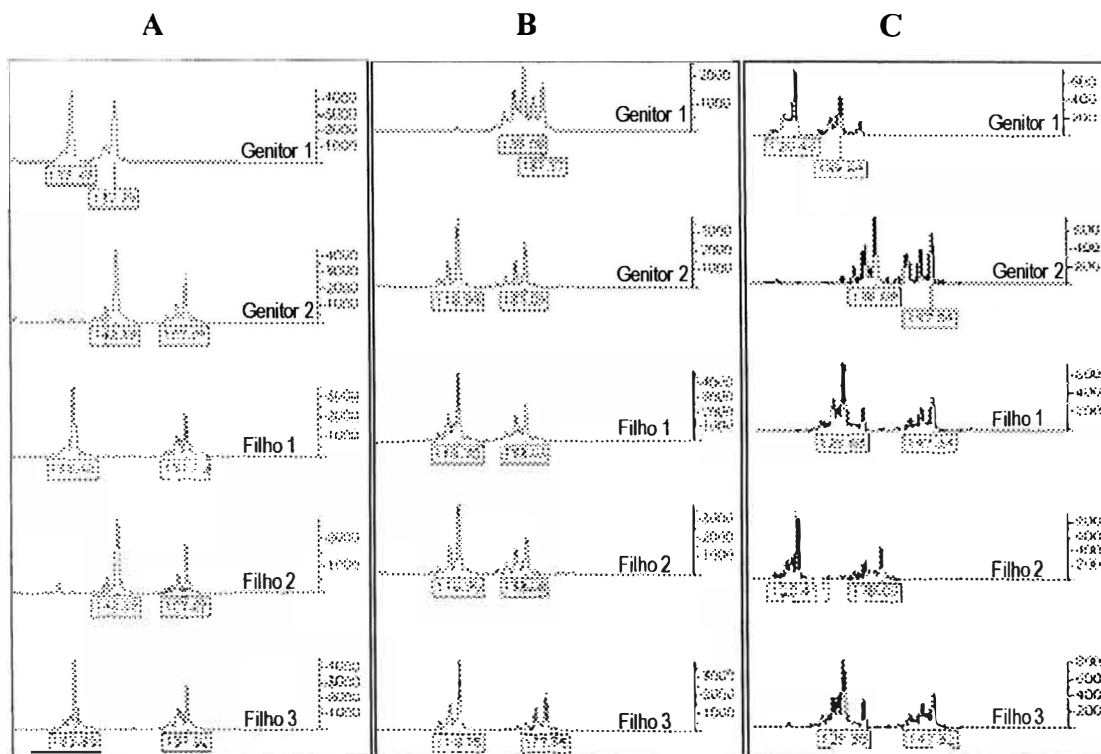


Figura 4 - Genotipagem de três microssatélites nos genitores e três descendentes com os locos EMBRA 37 (A), EMBRA 166 (B) e EMBRA 24 (C) com o programa Genotyper (Applied Biosystems)

5.2.7 Construção do mapa genético

Foi inicialmente utilizado o programa Mapmaker/EXP. 3.0 (Lander et al., 1987) para geração de um mapa genético individual para cada genitor da população segregante. Para montagem dos mapas individuais utilizou-se a

estratégia de pseudocruzamento teste proposto por Grattapaglia & Sederoff (1994). Para o agrupamento dos marcadores foi utilizado um LOD=10 e fração máxima de recombinação de 0,4. Os grupos de ligação suporte (*framework maps*) foram montados utilizando o comando *ripple* para gerar grupos com suporte de verossimilhança maior ou igual a 3,0 seguindo o proposto por Grattapaglia & Sederoff (1994). Para a criação de um mapa integrado, foi utilizado o programa JoinMap (Stam, 1993), que permite a construção de mapas genéticos com base em marcadores com diferentes razões de segregação em cruzamentos entre indivíduos heterozigotos. Os grupos de ligação foram montados com os mesmos parâmetros de estringência (LOD 10 e fração de recombinação 0,4). Para ordenação das marcas, o programa utiliza automaticamente o comando *ripple* após a adição de cada marca. Para ambos os programas as distâncias entre os marcadores foram calculados em centimorgans (cM) utilizando a função Kosambi. Para desenhar os mapas genéticos foi utilizado o programa MapChart (Voorrips, 2002). Os dados de genotipagem utilizados para as construções dos mapas genéticos são apresentados no CD-ROM.

5.2.8 Detecção dos QTL

A detecção de QTL foi realizada com base nos mapas genéticos de cada genitor separadamente, tendo em vista que ainda não existem metodologias de mapeamento de intervalo ou por análise de marcadores

individuais estabelecidas para mapeamento de QTL com base em mapas integrados. Por se tratar de um programa que lida apenas com cruzamentos endogâmicos, os mapas de pseudocruzamento teste gerados com o Mapmaker/EXP 3.0 foram utilizados para a detecção e mapeamento dos QTL, realizados individualmente para cada genitor, utilizando os programas Mapmaker/QTL (Lander et al., 1987), utilizando mapeamento por intervalo simples (MIS) e o programa QTL Cartographer (Wang et al., 2001), utilizando análise de marcadores individuais (AMI) (*single marker analysis*) por mapeamento de intervalo simples e mapeamento por intervalo composto (MIC). Para a análise de AMI, com base na correção de Bonferroni, foi adotado um nível de significância de 0,00055 para a declaração de QTL em cada teste individual resultando em nível de significância para o experimento todo (*experiment wise error rate*) de 0,05. Um LOD score igual a 3,0 foi utilizado para a declaração de um QTL nas análises de MIS utilizando Mapmaker-QTL. Para determinar o LOD limite para declarar significância nas análises de MIS e MIC com o QTL-Cartographer foi realizado um teste envolvendo 500 permutações de acordo com o procedimento descrito por Churchill & Doerge (1994) implementado no programa QTL-Cartographer. Para as análises realizadas via MIC, foi utilizado o modelo padrão de detecção, com uma janela de tamanho 10 cM e precisão de 0,5 cM. O método de regressão utilizado foi o de "*forward e backward*". Os dados utilizados para as análises são apresentados no CD-ROM.

5.3 Resultados

5.3.1 Levantamento dos dados fenotípicos

Dados fenotípicos de CAP (Circunferência na Altura do Peito) e de densidade básica da madeira foram levantados em campo durante o mês de Julho de 2001. Para estas características de campo os dados foram levantados para toda a população segregante de 270 árvores (ver Anexo D) e os valores de média, variância e desvio padrão para medidas de CAP e penetrância do Pilodyn em milímetros foram calculados para toda a progênie, e para os 188 indivíduos utilizados no mapeamento (Tabela 1). As mensurações de teor de lignina, teor de pentosana, fibra/grama, *coarseness*, rendimento em celulose e consumo específico obtidas com o NIRS foram realizadas em 80 indivíduos dos 188 genotipados em função da limitação de tempo de análise e custo (ver anexo D).

Tabela1. Valores de média, variância e desvio padrão para medidas de CAP (cm) e penetrância do Pilodyn (mm) considerando dados de toda a progênie e considerando os 188 indivíduos tomados ao acaso para o experimento de mapeamento

	Toda a Progênie (270 ind.)		Progênie utilizada para mapeamento (188 ind.)	
	CAP (cm)	Pilodyn (mm)	CAP (cm)	Pilodyn (mm)
Média	35,103	14,72	35,25	14,856
Variância	50,83	4,098	53,833	3,920
Desvio Padrão	7,129	2,024	7,337	1,979

5.3.2 Triagem e seleção dos locos microssatélites

Dos 487 locos testados na etapa de triagem para herança e polimorfismo, 169 locos segregaram em configurações informativas na população, apresentando segregações do tipo 1:2:1, 1:1 e 1:1:1:1.

5.3.3 Geração dos dados genotípicos e construção dos mapas genéticos

Dos 169 marcadores microssatélites genotipados na população, 123 foram utilizados para a construção do mapa genético do genitor materno 235, com base em segregação 1:1. Com base em um teste de χ^2 , 53 dos 123 marcadores (43%) apresentaram distorção na segregação alélica, com χ^2 variando de 5,3 (EMBRA 210) a 125,9 (EMBRA 629). Dos 123 marcadores genotipados, 95 foram mapeados em 10 grupos de ligação (Figura 5). Para a montagem do mapa genético do genitor 221, foram utilizados 117 dos 169

marcadores microsatélites genotipados, com base em segregação esperada 1:1. O teste de χ^2 indicou que 64 (54,7%) apresentaram distorção na segregação, com valores de χ^2 variando de 5,10 (EMBRA 72) a 103,5 (EN-14). Deste total, 92 marcadores microsatélites foram mapeados em 10 grupos de ligação (Figura 5). Para a montagem de um mapa genético integrado para os dois parentais, foram utilizados 155 dos 169 marcadores microsatélites genotipados. Para estes marcadores, 119 (70,1%) apresentaram distorção na segregação de seus alelos, em relação a uma das três possíveis proporções Mendelianas esperadas (1:1:1:1; 1:2:1; 1:1), com valores de χ^2 variando de 5,9 (EMBRA 56) a 136,1 (EMBRA 659). O mapa genético integrado apresentou os 155 marcadores microsatélites distribuídos em 11 grupos de ligação utilizando os critérios de LOD e fração de recombinação pré-estabelecidos (Figura 5). O comprimento total do mapa genético realizado para o genitor materno 235 foi de 819,6 cM, a distância média entre marcadores no mapa foi de 6,6 cM e a maior distância entre marcadores foi de 40,2 cM no grupo de ligação 3. Para o genitor paterno 221, o comprimento total do mapa genético foi de 536,4 cM, a distância média entre marcadores foi de 4,58 cM e a maior distância entre marcadores foi de 30,8 cM observada no grupo de ligação 3. Já o mapa integrado apresentou um comprimento total de 1032,9 cM com distância média entre marcadores de 8,08, onde a maior distância encontrada também no grupo de ligação 3 (48,1cM). O mapa genético do genitor materno 235 não apresentou um grupo de ligação referente ao grupo de ligação 7, já o mapa genético do genitor

paterno 221 não apresentou grupo de ligação referente ao grupo 8. O mapa integrado apresentou o número de grupos de ligação equivalente ao número haplóide de cromossomos de Eucalipto ($n = 11$) (Figura 5).

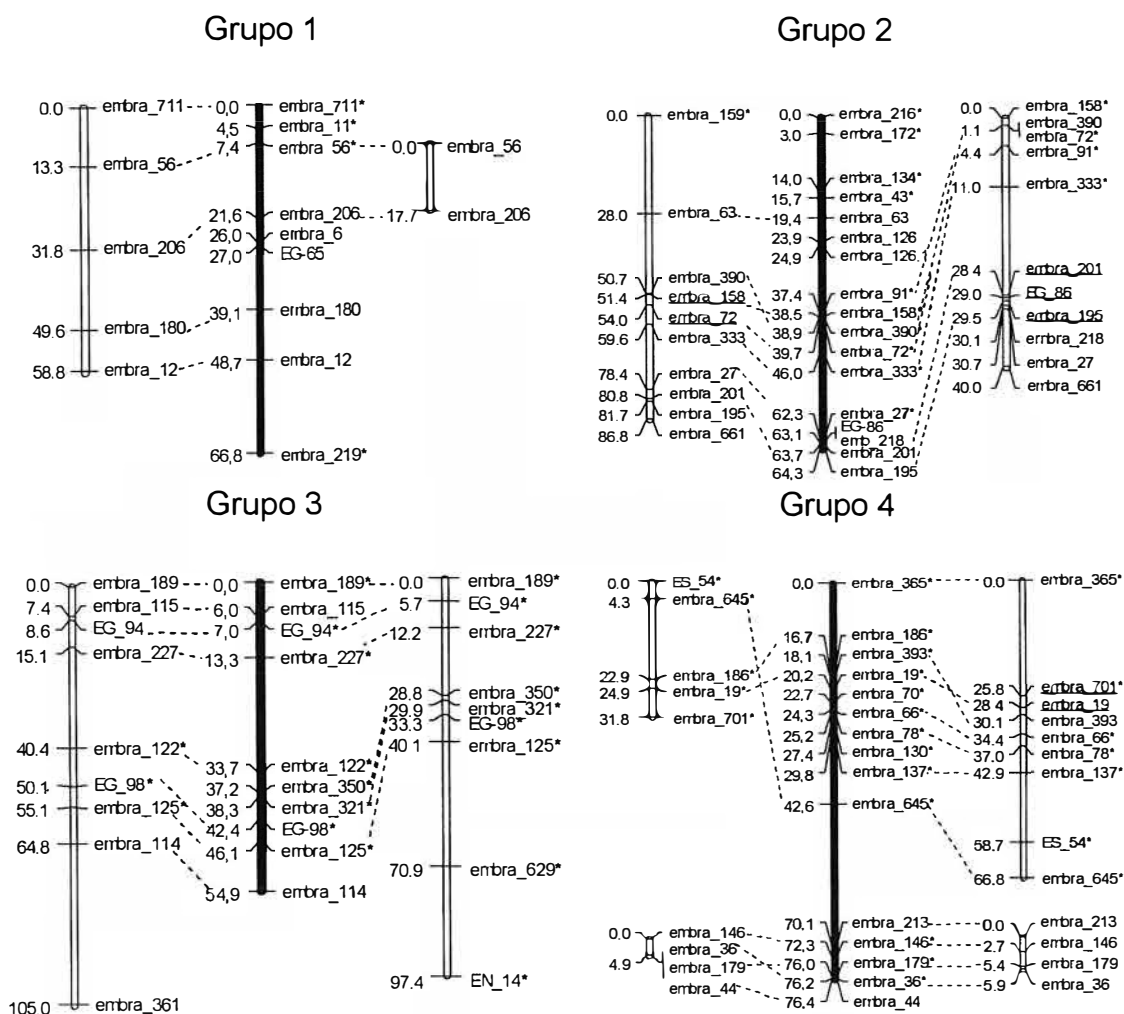


Figura 5 - Mapas genéticos construídos com base em 169 marcadores microssatélites genotipados em 188 descendentes de uma progênie gerada a partir do cruzamento entre híbridos naturais de *E. grandis* e *E. urophylla*. Para cada grupo de ligação, são apresentados os mapas para o genitor 235 (esquerda) para o parental 221 (direita) e o mapa integrado

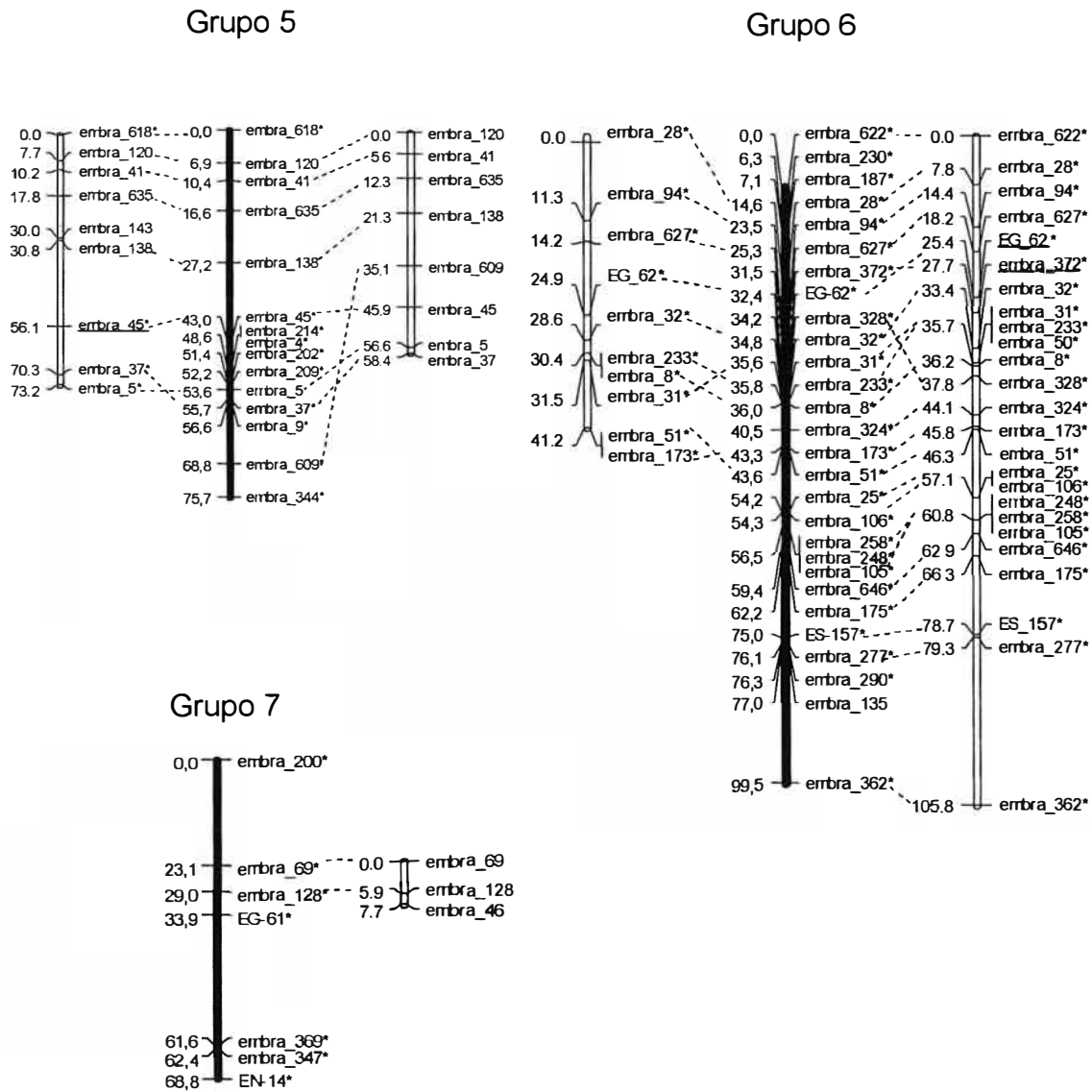


Figura 5 Mapas genéticos construídos com base em 169 marcadores microssatélites genotipados em 188 descendentes de uma progênie gerada a partir do cruzamento entre híbridos naturais de *E. grandis* e *E. urophylla*. Para cada grupo de ligação, são apresentados os mapas para o genitor 235 (esquerda) para o parental 221 (direita) e o mapa integrado

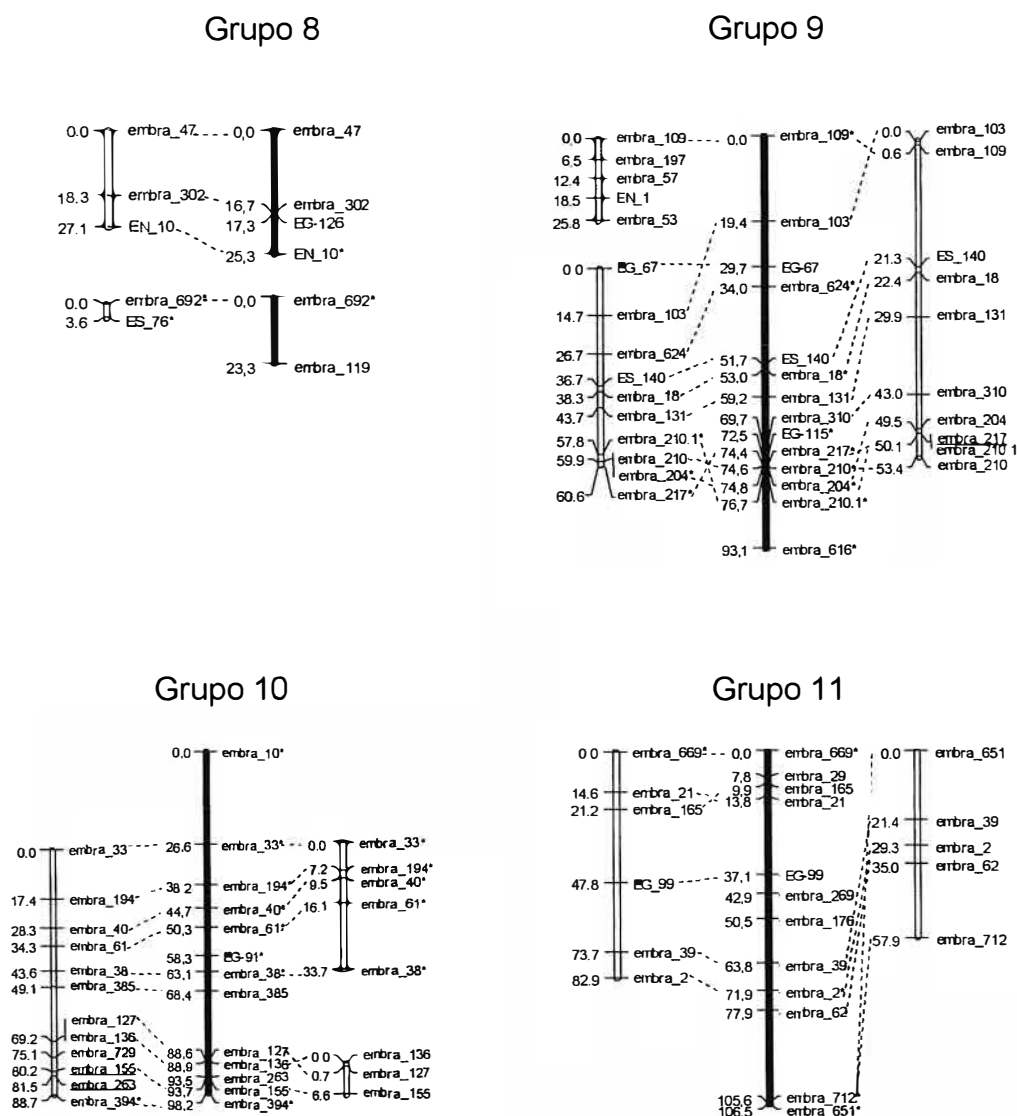


Figura 5 - Mapas genéticos construídos com base em 169 marcadores microssatélites genotipados em 188 descendentes de uma progênie gerada a partir do cruzamento entre híbridos naturais de *E. grandis* e *E. urophylla*. Para cada grupo de ligação, são apresentados os mapas para o genitor 235 (esquerda) para o parental 221 (direita) e o mapa integrado

5.3.4 Detecção e mapeamento dos QTL utilizando AMI

A análise de marcadores individuais foi realizada com um nível de significância para todo o experimento (*experiment-wise*) de 0,05 aplicando-se a correção de Bonferroni em cada teste de marcador individual para os 92 testes realizados com os 92 marcadores em cada característica revelou a localização de QTL com valor $p < 0,001$ para quatro das oito características (Tabela 2). Para as características de teor de pentosanas e rendimento em celulose foram detectados QTL com um nível e significância de 0,01 e para coarseness N° de fibras/grama ao nível de 0,05.

Tabela 2. Análise de QTL por marcadores individuais (AMI) realizada com marcadores microssatélites que compõem os mapas genéticos dos genitores materno 235 e paterno 221. São apresentados os marcadores com valores de níveis de significância iguais ou menores que $\alpha \leq 0,00054$ (*experiment-wise* de 0,05) com exceção de QTL para teor de lignina e consumo específico declarados com $\alpha \leq 0,001$ (*experiment-wise* de 0,1)

Característica	Genitor	Grupo de Ligação	Marcador	Valor p
CAP	221	3	EMBRA 189	0,000
CAP	221	3	EMBRA 227	0,000
CAP	221	6	EMBRA 28	0,000
CAP	235	1	EMBRA 711	0,000
Densidade	221	4	EMBRA 137	0,000
Densidade	221	4	EMBRA 78	0,000
Densidade	221	4	EMBRA 66	0,000

Tabela 2. Análise de QTL por marcadores individuais (AMI) realizada com marcadores microssatélites que compõem os mapas genéticos dos genitores materno 235 e paterno 221. São apresentados os marcadores com valores de níveis de significância iguais ou menores que $\alpha \leq 0,00054$ (*experiment-wise* de 0,05) com exceção de QTL para teor de lignina e consumo específico declarados com $\alpha \leq 0,001$ (*experiment-wise* de 0,1)

Característica	Genitor	Grupo de Ligação	Marcador	Valor p
Densidade	221	4	EMBRA 393	0,000
Densidade	221	6	EMBRA 622	0,000
Densidade	221	6	EMBRA 28	0,000
Densidade	221	6	EMBRA 233	0,000
Densidade	221	6	EMBRA 50	0,000
Densidade	221	6	EMBRA 8	0,000
Densidade	221	6	EMBRA 627	0,000
Densidade	221	6	EMBRA 328	0,000
Densidade	221	6	EMBRA 324	0,000
Densidade	221	6	EMBRA 173	0,001
Densidade	235	4	EMBRA 186	0,000
Lignina	221	2	EMBRA 333	0,001
Cons.Específico	235	10	EMBRA 385	0,001

5.3.5. Detecção de QTL utilizando mapeamento de intervalo (MIS e MIC)

5.3.5.1 Detecção de QTL para densidade da madeira

Foram detectados dois QTL para densidade da madeira no genitor paterno 221. O primeiro no grupo de ligação 4 (LOD 3,4), localizado entre os marcadores EMBRA 137 e EMBRA 393, explicando 6,39% da variação

fenotípica e o segundo, no grupo de ligação 6 (LOD 4,8), localizado entre os marcadores EMBRA 662 e EMBRA 94, explicando 9,01% da variação fenotípica. A detecção dos dois QTL no genitor 221 foram coincidentes para as análises utilizando MIS realizadas pelo programas Mapmaker/QTL e QTL Cartographer, assim como para MIC utilizando o programa QTL Cartographer (Figura 6 e 7).

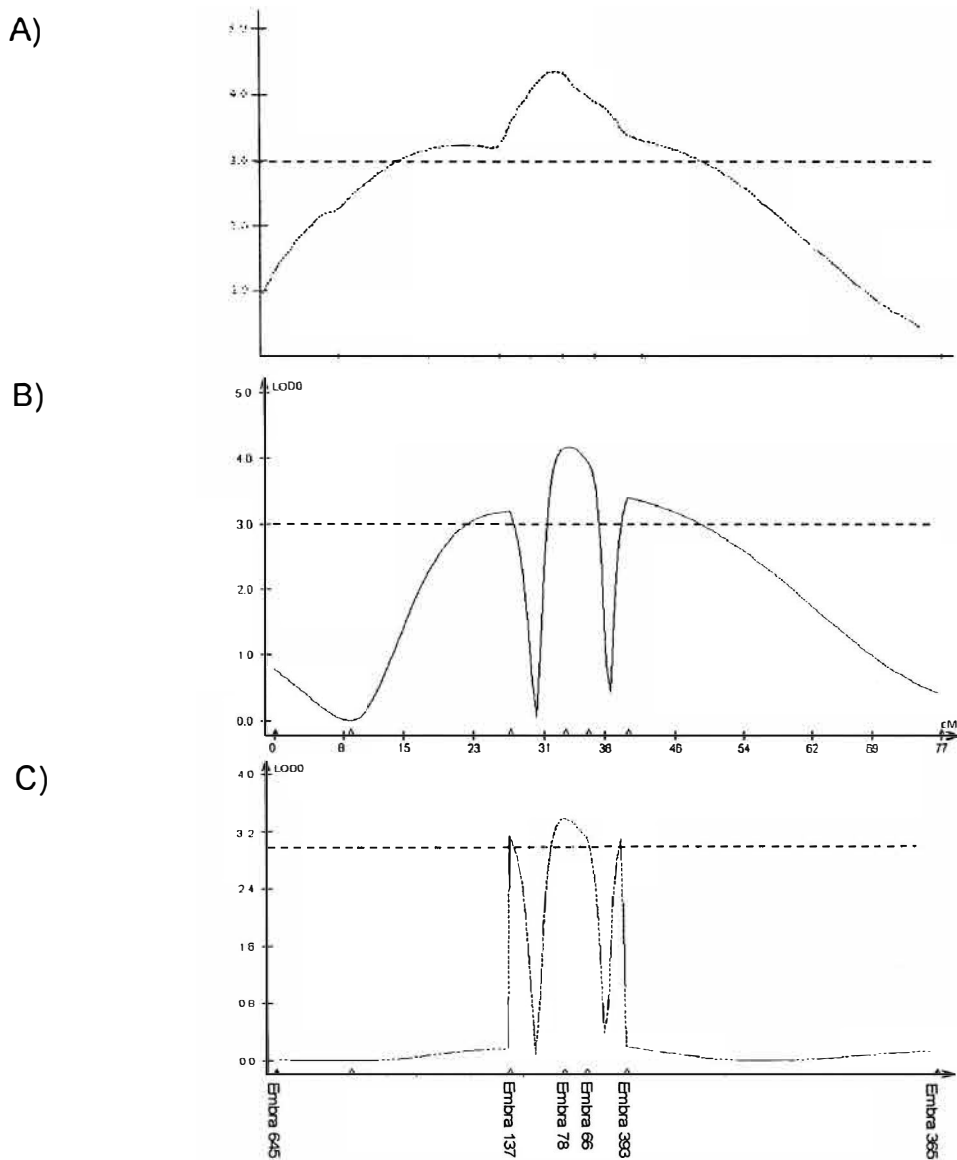


Figura 6 - Mapeamento de um QTL no grupo de ligação 4 do genitor paterno 221 para densidade da madeira. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTL Cartographer

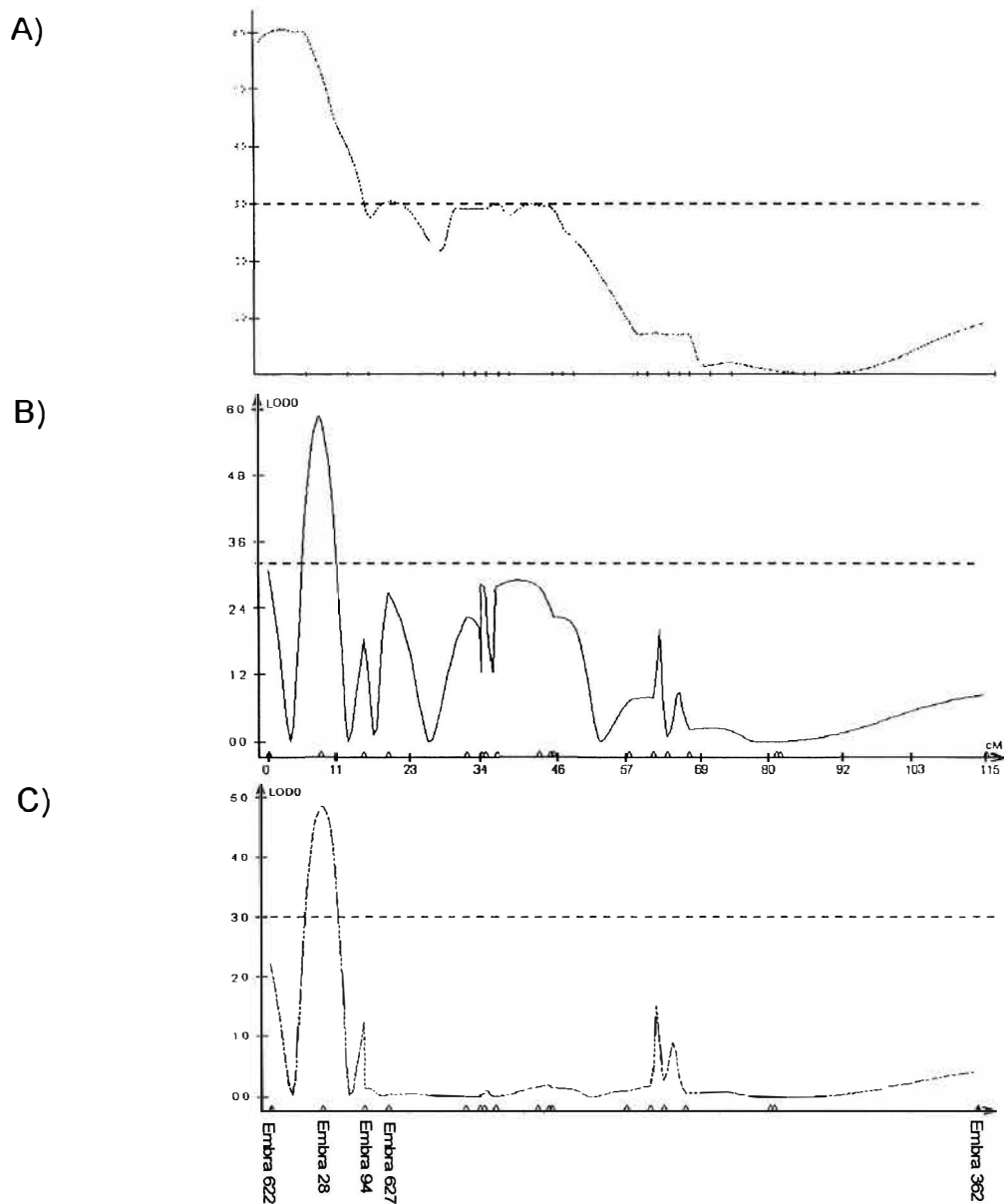


Figura 7 - Mapeamento de um QTL no grupo de ligação 6 do genitor paterno 221 para densidade da madeira. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTL Cartographer

5.3.5.2 Detecção de QTL para teor de lignina

Foram detectados dois QTL para teor de lignina. Para o genitor 235, foi detectado um QTL no grupo de ligação 9. O QTL (LOD 2,92) está localizado entre os marcadores EMBRA 131 e ES-140, explicando 39,13% da variação fenotípica. A detecção deste QTL foi feita utilizando MIC com o programa QTL Cartographer e não foi coincidente quando se utilizou MIS com os programas Mapmaker/QTL e QTL Cartographer (Figura 8). O segundo QTL foi detectado no grupo de ligação 2 do genitor 221. O QTL (LOD 4,09) está localizado entre os marcadores EMBRA 333 e EMBRA 218, e explica 19,84% da variação fenotípica. A detecção deste QTL foi feita utilizando MIC, sendo coincidente para as análises utilizando MIS realizadas pelo programas Mapmaker/QTL e QTL Cartographer, apesar do LOD ter diminuído de 4,09 para 2,4 (Figura 9).

5.3.5.3 Detecção de QTL para teor de pentosana

Foram detectados dois QTL para teor de pentosana no genitor 235. O primeiro QTL (LOD 3,1) está localizado no grupo de ligação 9 entre os marcadores EMBRA 57 e EMBRA 53, explicando 36% da variação fenotípica. O segundo QTL (LOD 3,6) está localizado no grupo de ligação 6 entre os marcadores EMBRA 28 e EMBRA 267, explicando 14,6% da variação fenotípica. A detecção destes QTL só foi possível utilizando MIC, não sendo detectada para MIS (Figuras 10 e 11).

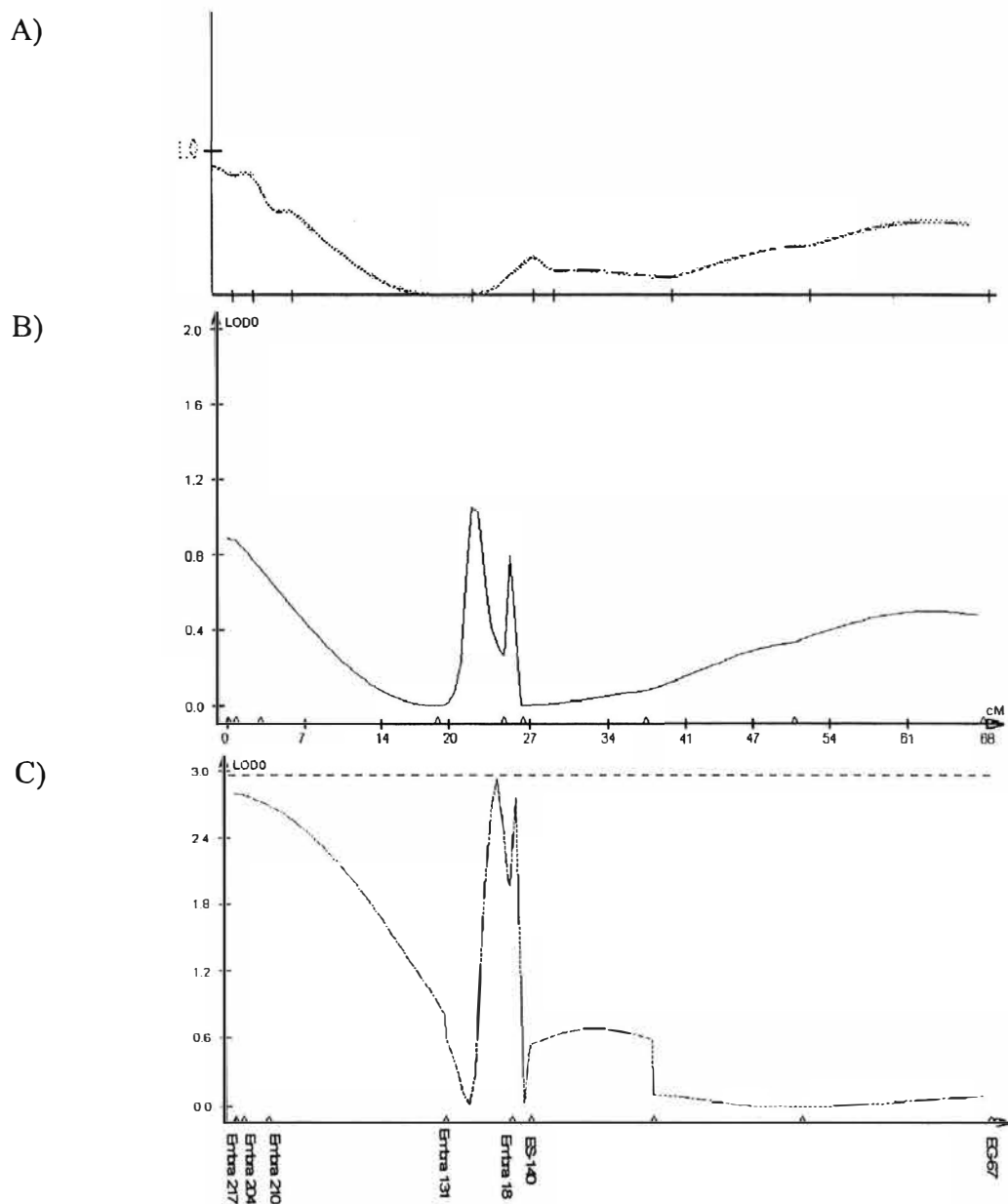


Figura 8 - Mapeamento de um QTL no grupo de ligação 9 do genitor materno 235 para teor de lignina. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTLCartographer

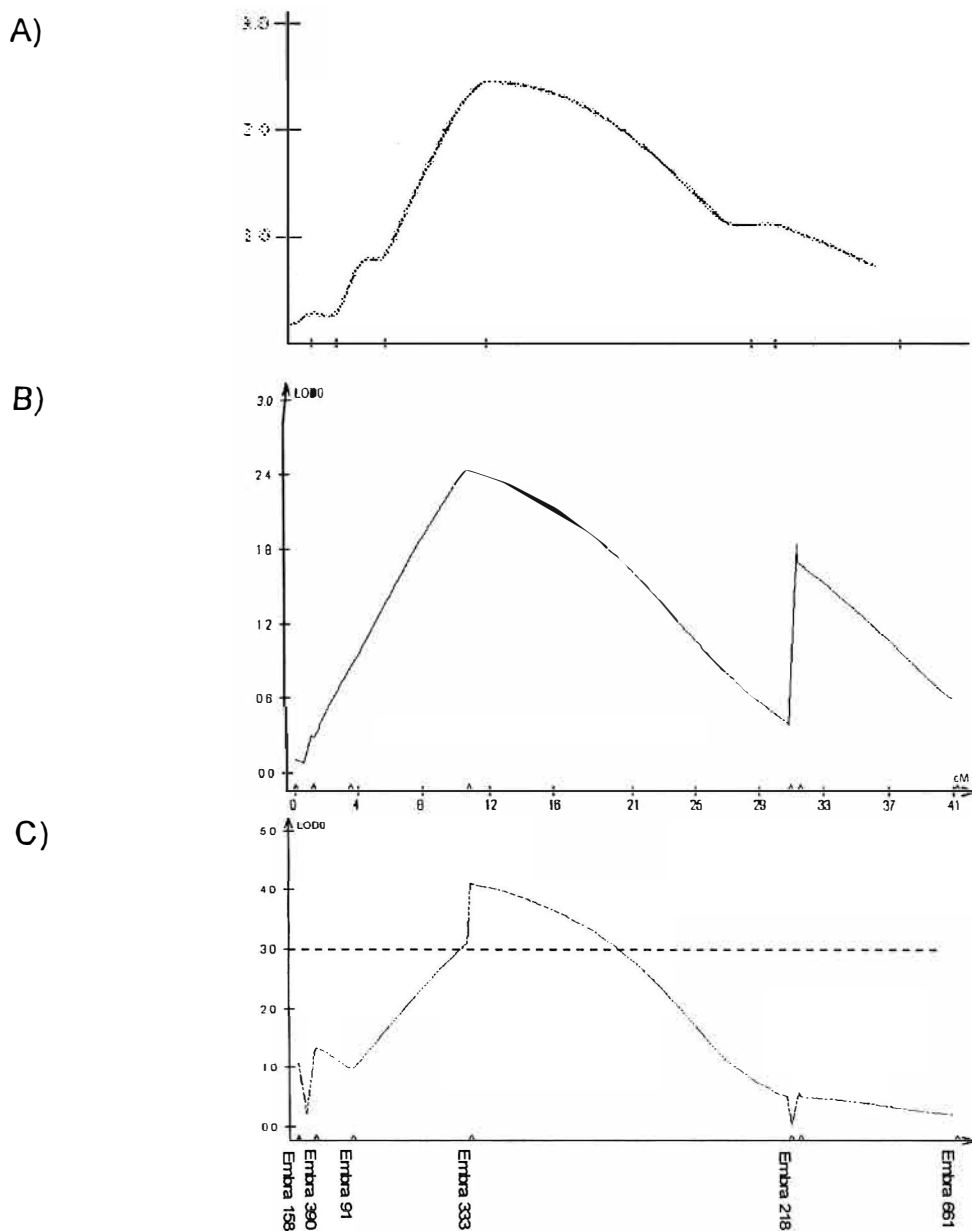


Figura 9 - Mapeamento de um QTL no grupo de ligação 2 do genitor paterno 221 para teor de lignina. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTLCartographer

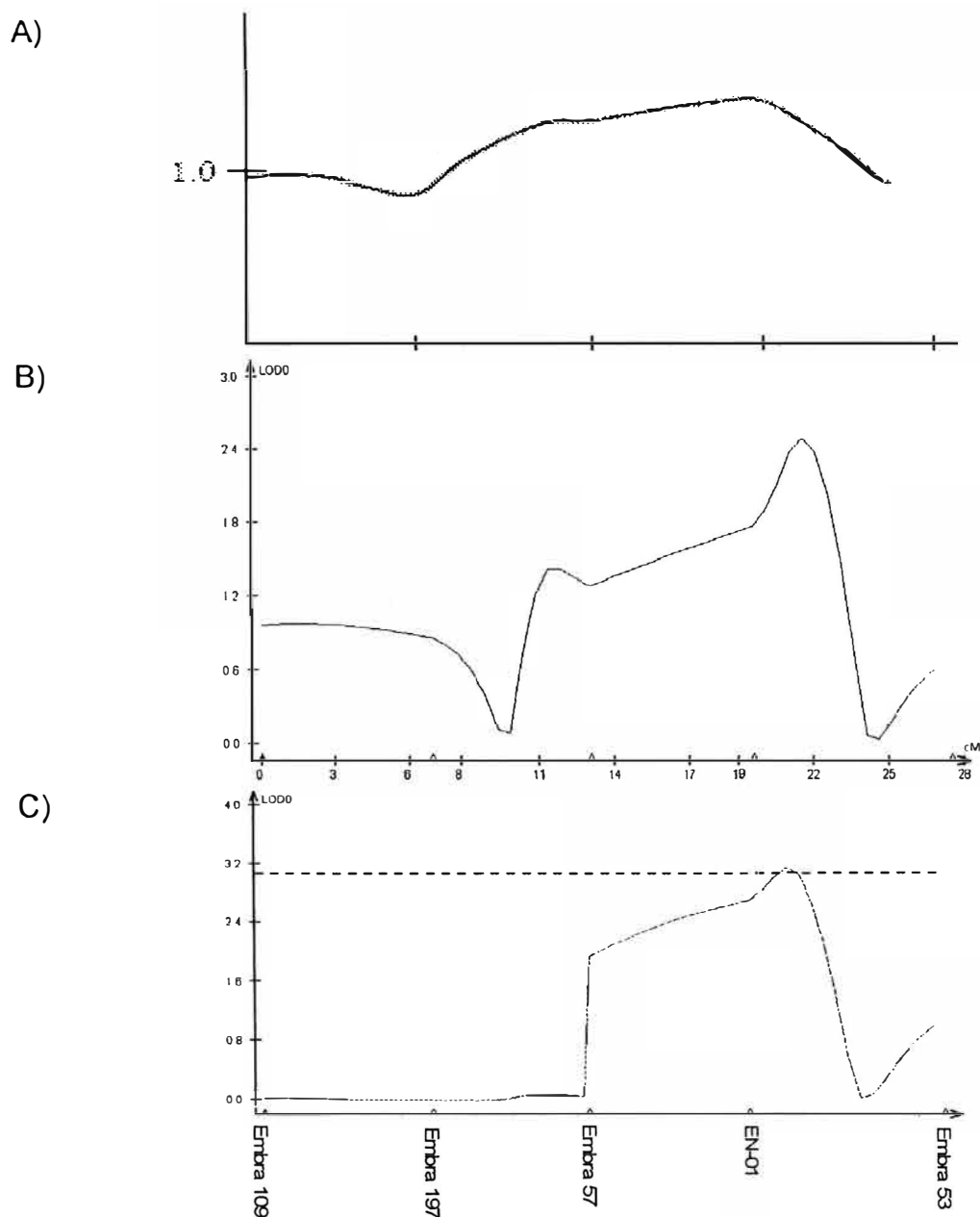


Figura 10 - Mapeamento de um QTL no grupo de ligação 9 do genitor materno 235 para teor de pentosana. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTLCartographer

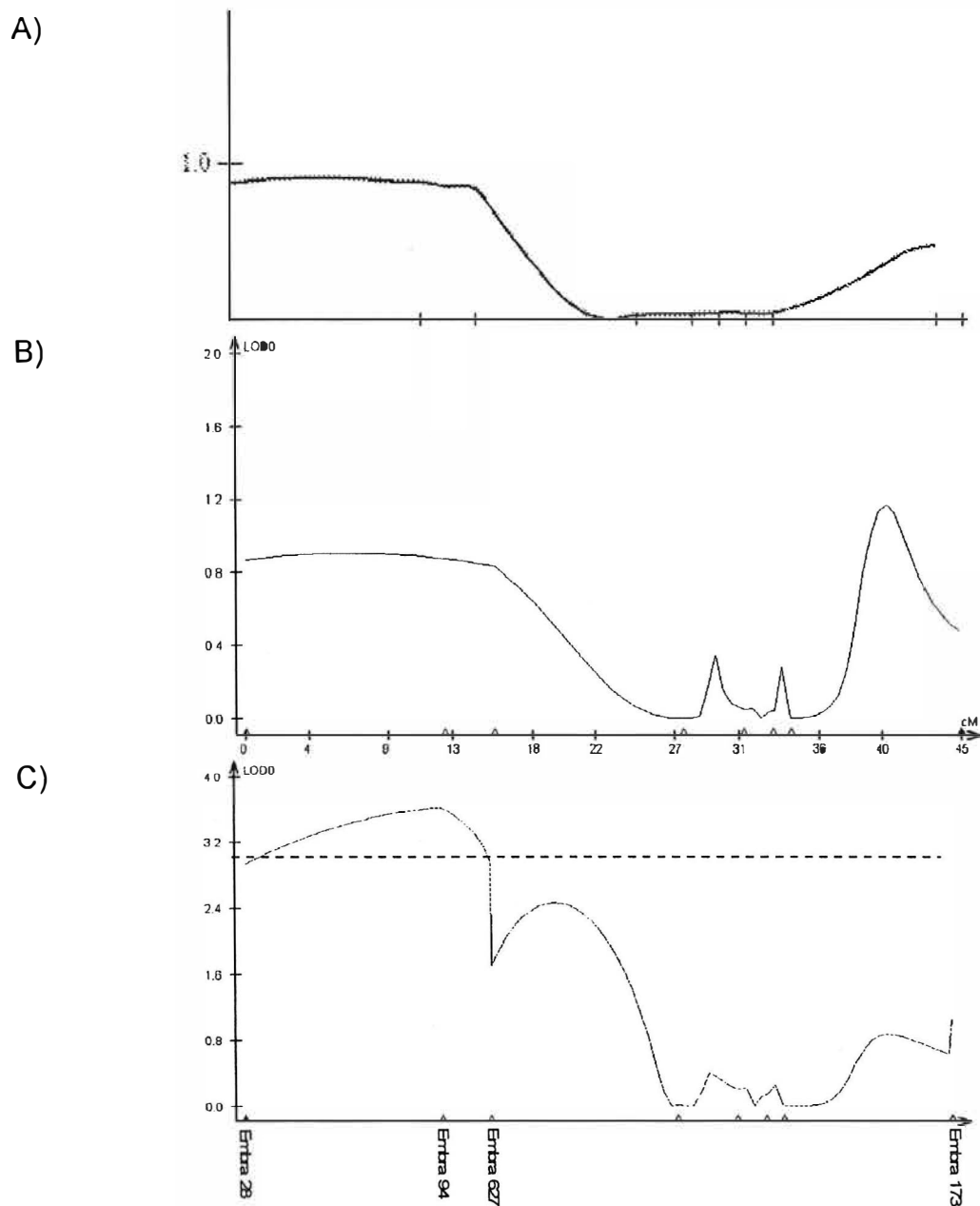


Figura 11 - Mapeamento de um QTL no grupo de ligação 6 do genitor materno 235 para teor de pentosana. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTLCartographer

5.3.5.4 Detecção de QTL para quantidade de fibra/grama

Para esta característica foi detectado um QTL (LOD 3,0) no grupo de ligação 10 do genitor 235. Este QTL está localizado entre os marcadores EMBRA 61 e EMBRA 94, sendo responsável por 22,46% da variação fenotípica. A detecção deste QTL foi realizada utilizando MIC, não sendo detectado pelas análises realizadas com MIS (Figura 12).

5.3.5.5 Detecção de QTL para *coarseness*

Foram detectados três QTL para *coarseness* no genitor 235. O primeiro QTL (LOD 4,1) foi detectado no grupo de ligação 3 entre os marcadores EG-94 e EMBRA 227, explicando 58% da variação fenotípica. O segundo QTL (LOD 3,5) foi detectado no grupo de ligação 9 entre as marcas EG-18 e ES-140, explicando 53% da variação fenotípica. Já o terceiro QTL (LOD 4,0), foi detectado no grupo de ligação 4 entre as marcas EMBRA 186 e EMBRA 19, explicando 53% da variação fenotípica. A detecção destes QTL foi realizada utilizando MIC, e MIS para o programa QTL Cartographer, apesar dos valores de LOD Score terem diminuído. Já nas análises de MIS realizadas com o programa Mapmaker/QTL nenhum desses QTL foram detectados (Figuras 13, 14 e 15).

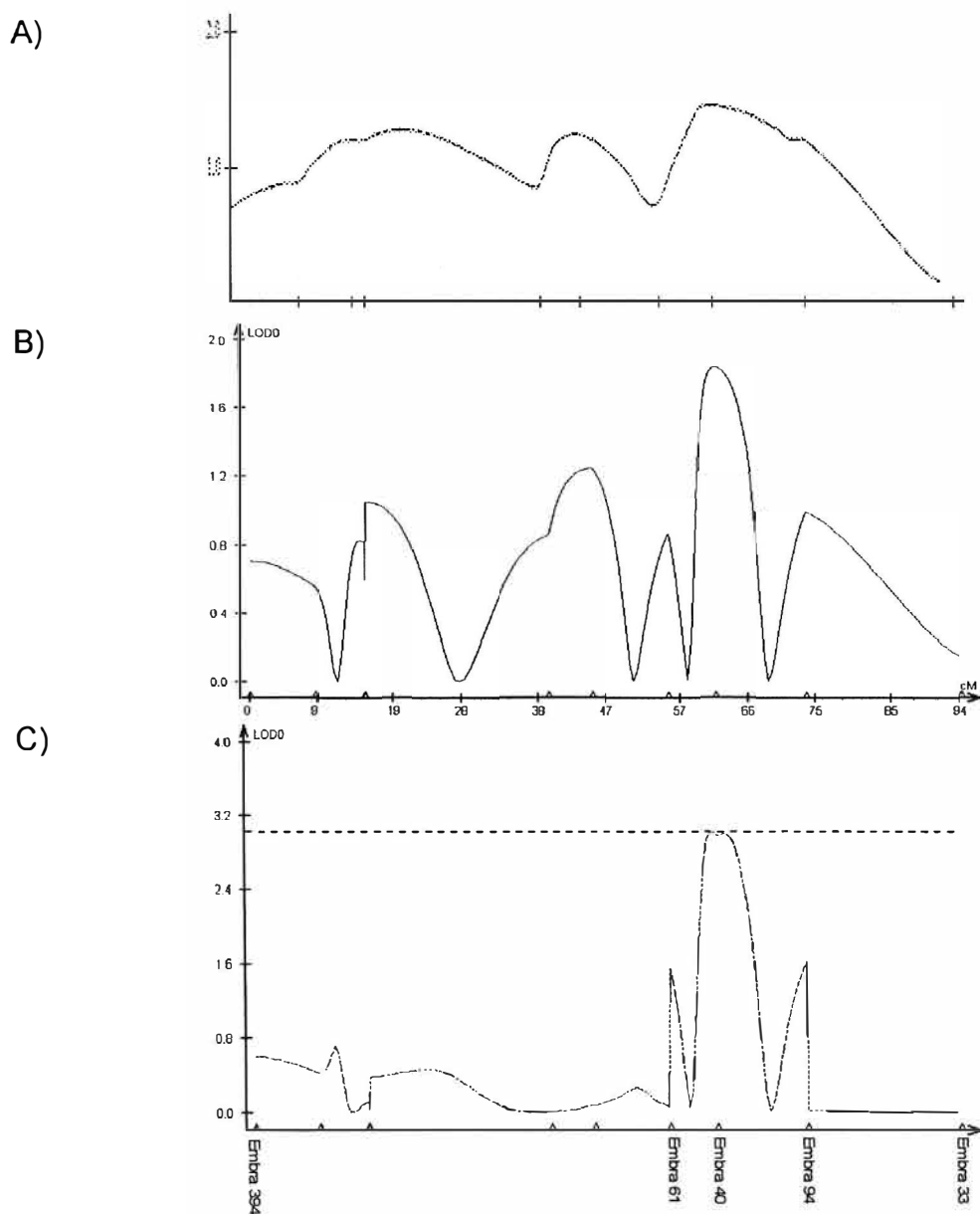


Figura 12 - Mapeamento de um QTL no grupo de ligação 10 do genitor materno 235 para quantidade de fibra/grama. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTLCartographer

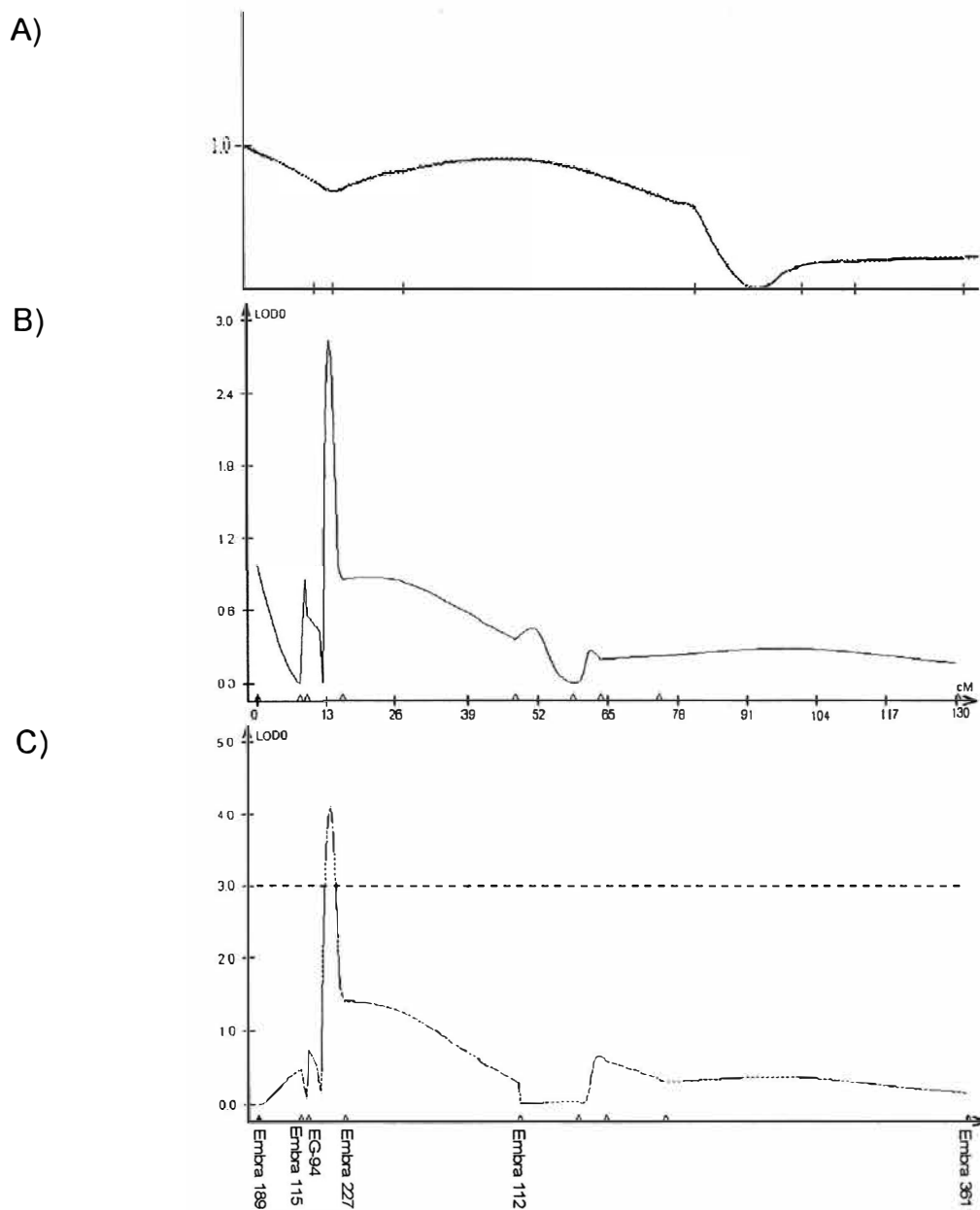


Figura 13 - Mapeamento de um QTL no grupo de ligação 3 do genitor materno 235 para *coarseness*. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTLCartographer

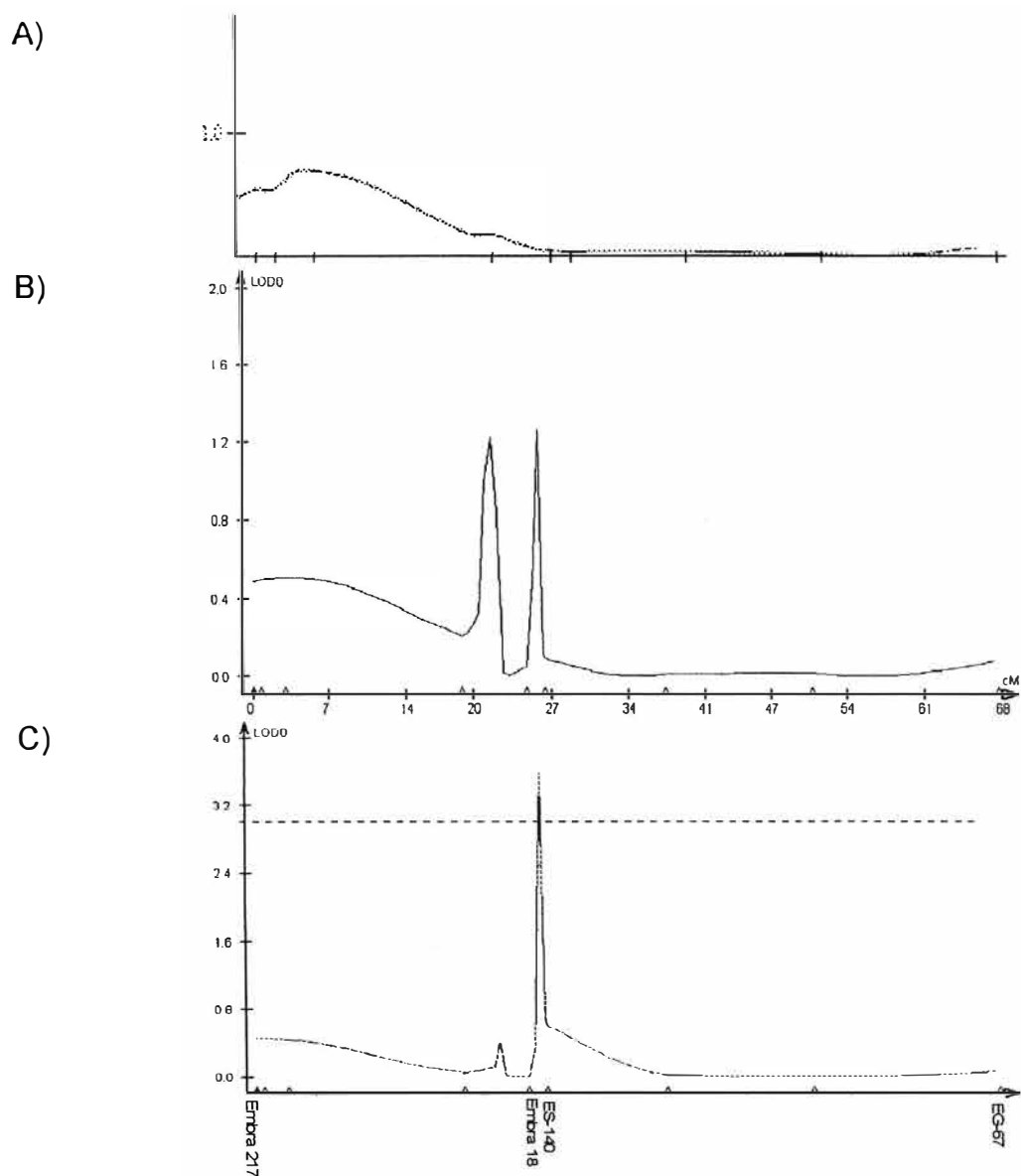


Figura 14 - Mapeamento de um QTL no grupo de ligação 9 do genitor materno 235 para *coarseness*. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTLCartographer

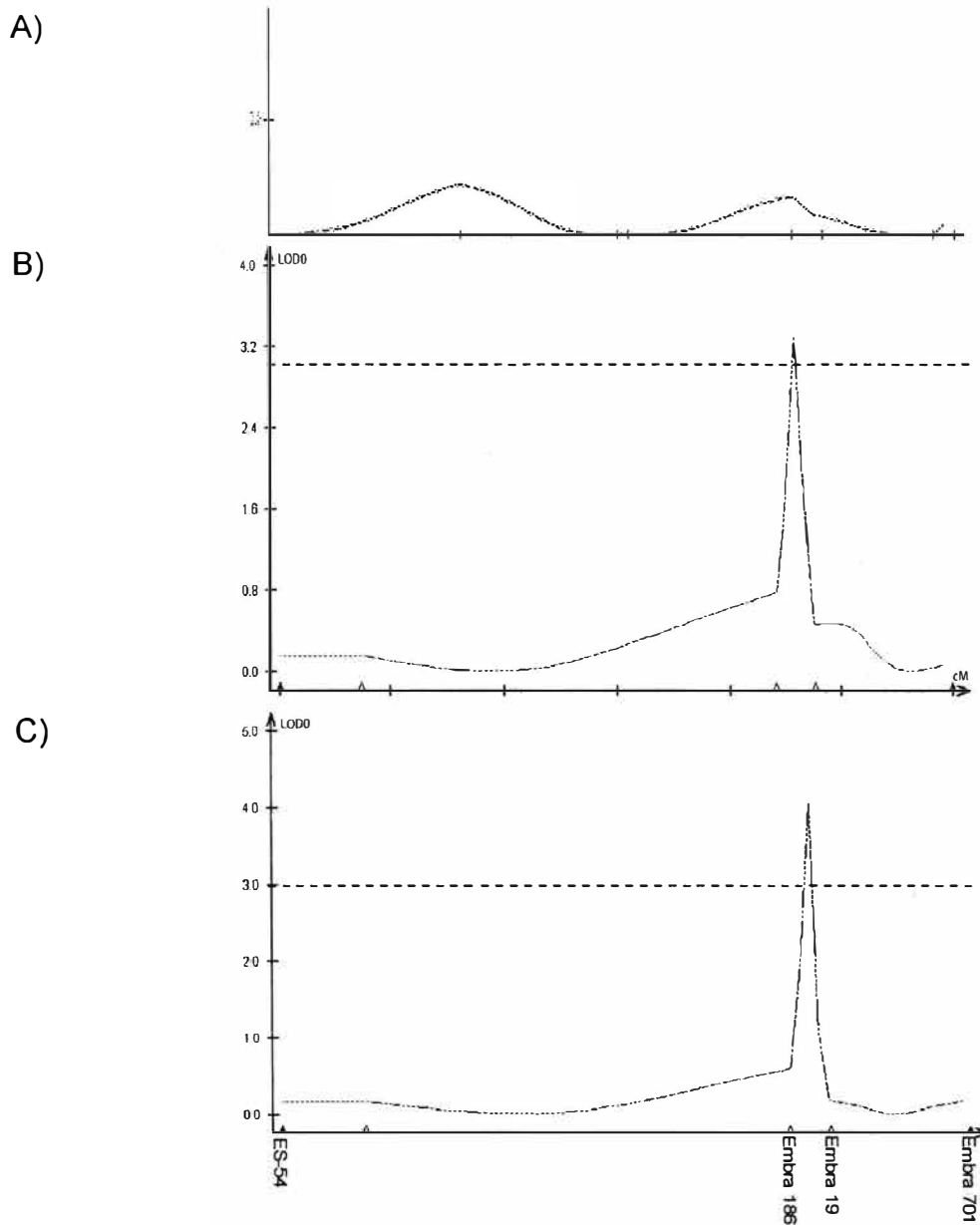


Figura 15 - Mapeamento de um QTL no grupo de ligação 4 do genitor materno 235 para *coarseness*. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTLCartographer

5.3.5.6 Detecção de QTL para rendimento em celulose

Foram detectados dois QTL para rendimento em celulose. Para o genitor 235, foi detectado um QTL (LOD 3,5) localizado no grupo de ligação 9 entre os marcadores EMBRA 131 e ES-140, explicando 47,75% da variação fenotípica. A detecção deste QTL foi realizada utilizando MIC, não sendo detectado pelas análises realizadas com MIS (Figura 16). O segundo QTL (LOD 2,9) foi detectado no grupo de ligação 3 no genitor 221 entre as marcas EMBRA 125 e EMBRA 629, explicando 23,64% da variação fenotípica. A detecção deste QTL foi consistente para as análises de MIC bem como MIS (Figura 17).

5.3.5.7 Detecção de QTL para consumo específico

Para esta característica foi detectado um QTL (LOD 4,2) no grupo de ligação 9. do genitor 235. Este QTL está localizado entre os marcadores EMBRA 217 e EMBRA 131, sendo responsável por 20,60% da variação fenotípica. A detecção deste QTL foi realizada utilizando MIC, não sendo detectado pelas análises realizadas com MIS (Figura 18).

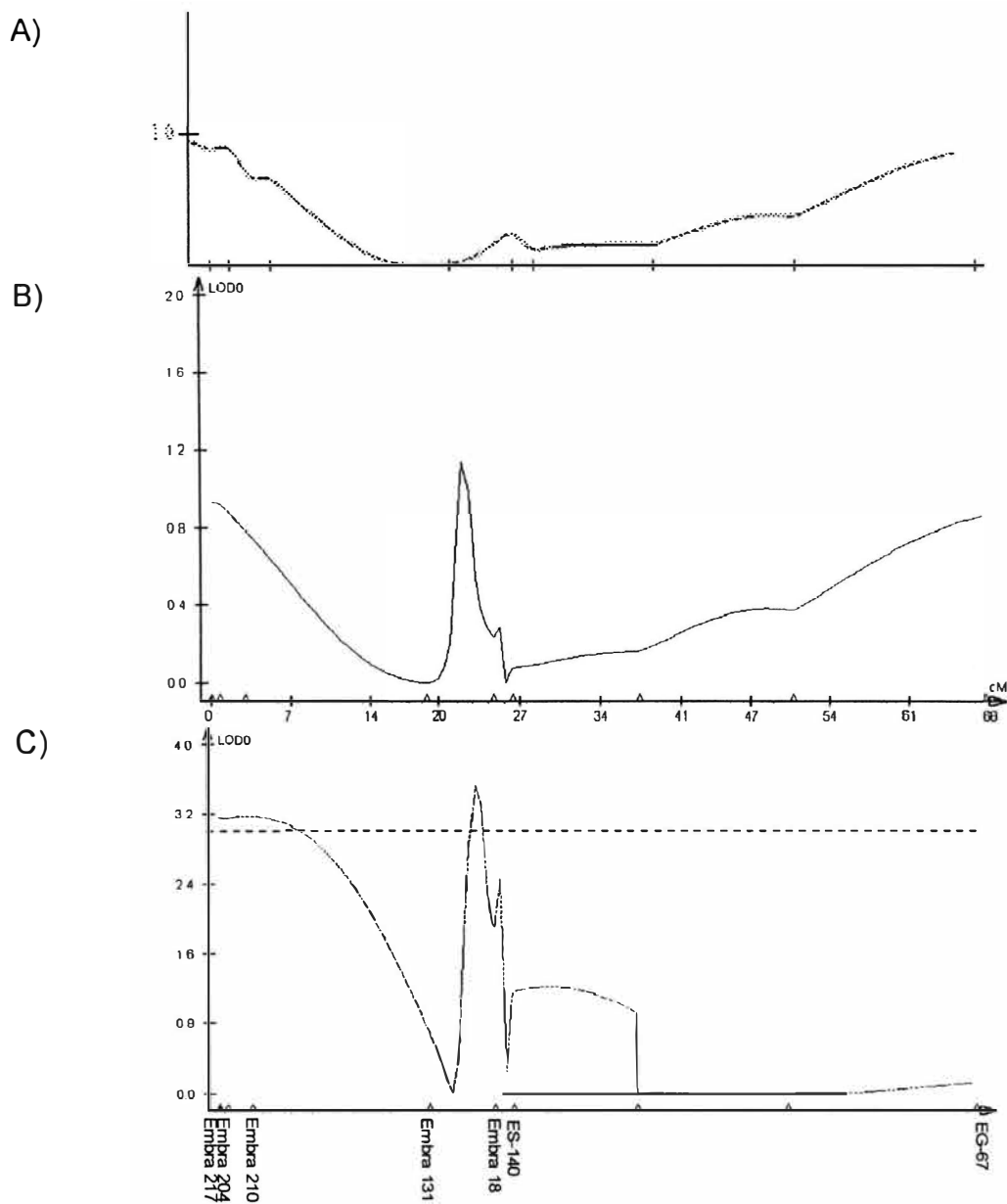


Figura 16 - Mapeamento de um QTL no grupo de ligação 9 do genitor materno 235 para rendimento. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTLCartographer

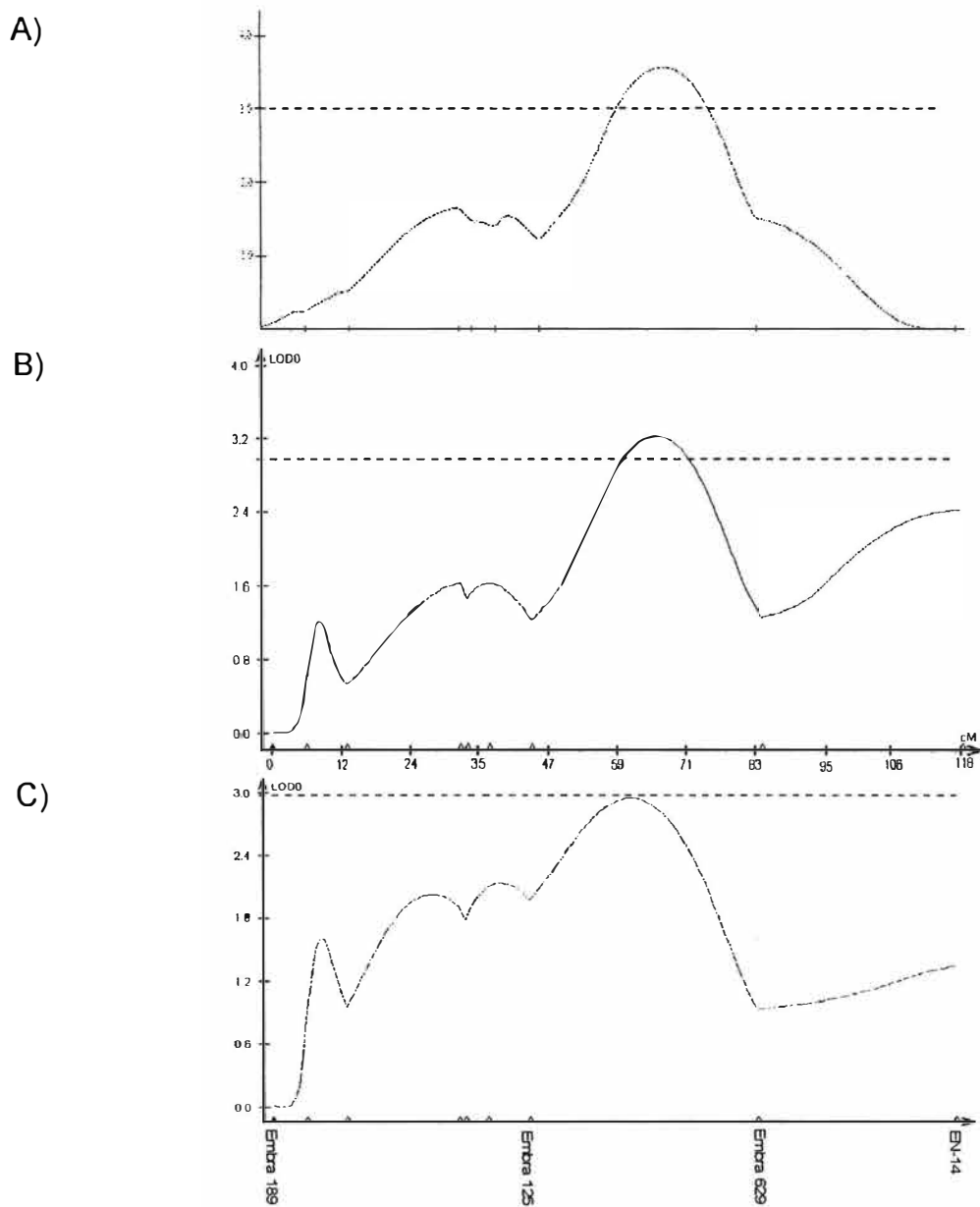


Figura 17 - Mapeamento de um QTL no grupo de ligação 3 do genitor paterno 221 para rendimento de celulose. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTL Cartographer

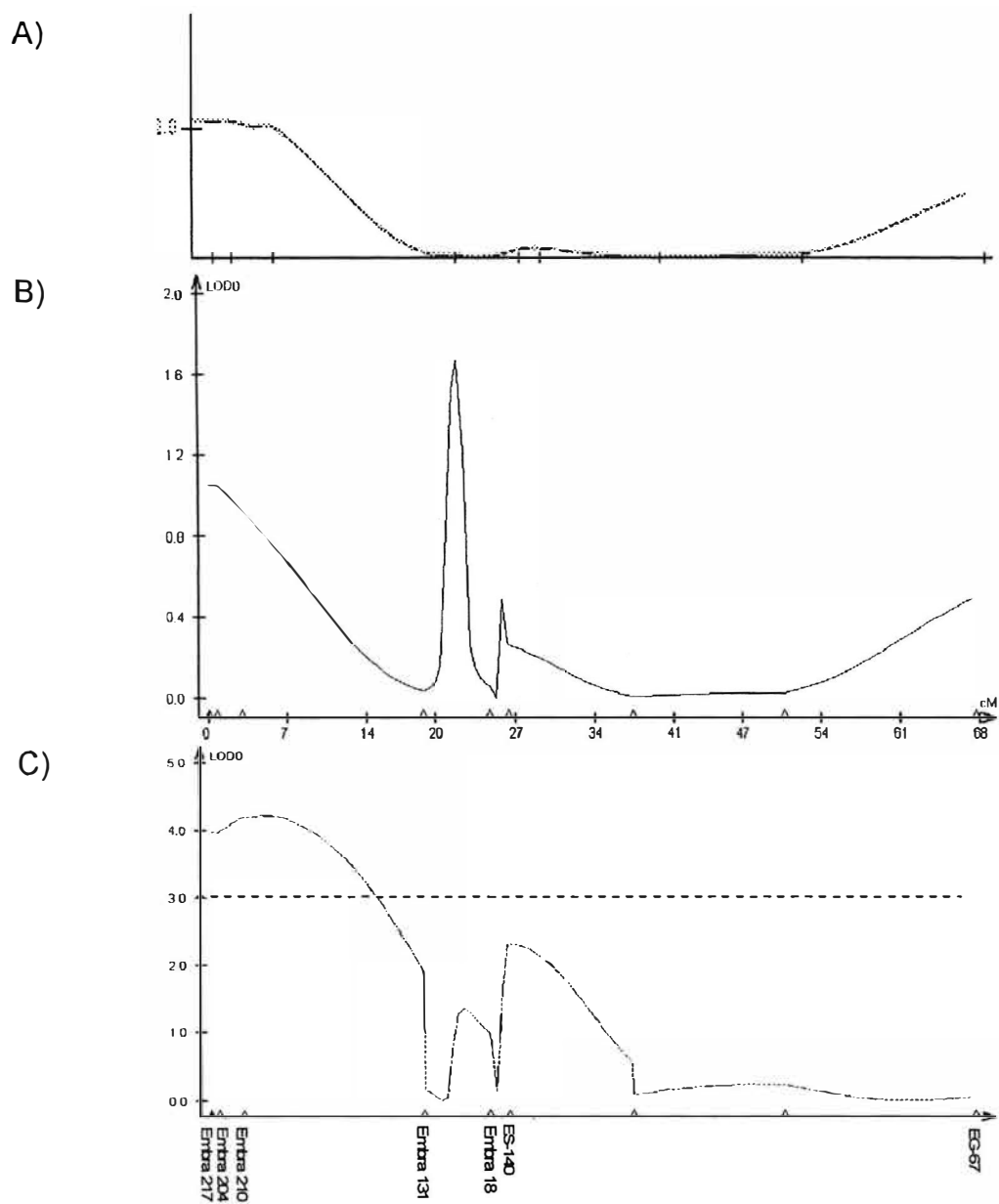


Figura 18 - Mapeamento de um QTL no grupo de ligação 9 do genitor materno 235 para consumo específico. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTLCartographer

5.3.5.8 Detecção de QTL para crescimento volumétrico (CAP)

Para a característica de CAP, não foram detectados QTL utilizando MIC com valores de LOD iguais ou superiores a 3,0. Porém para análises de MIS realizadas com os programas Mapmaker/QTL e QTL Cartographer foram detectados 2 QTL para esta característica. O primeiro QTL (LOD 3,18) está localizado no grupo de ligação 1 do genitor materno 235. Este QTL está flanqueado pelos marcadores EMBRA 711 e EMBRA 56, sendo responsável por 9,03% da variação fenotípica (Figura 19). O segundo QTL (LOD 4,2 pela análise com Mapmaker/QTL e LOD 3,7 pela análise realizada com QTL Cartographer) foi detectado no grupo de ligação 3 do genitor paterno 221. Este QTL está flanqueado pelos marcadores EMBRA 189 e EMBRA 122, sendo responsável por 9,32% da variação fenotípica (Figura 20).

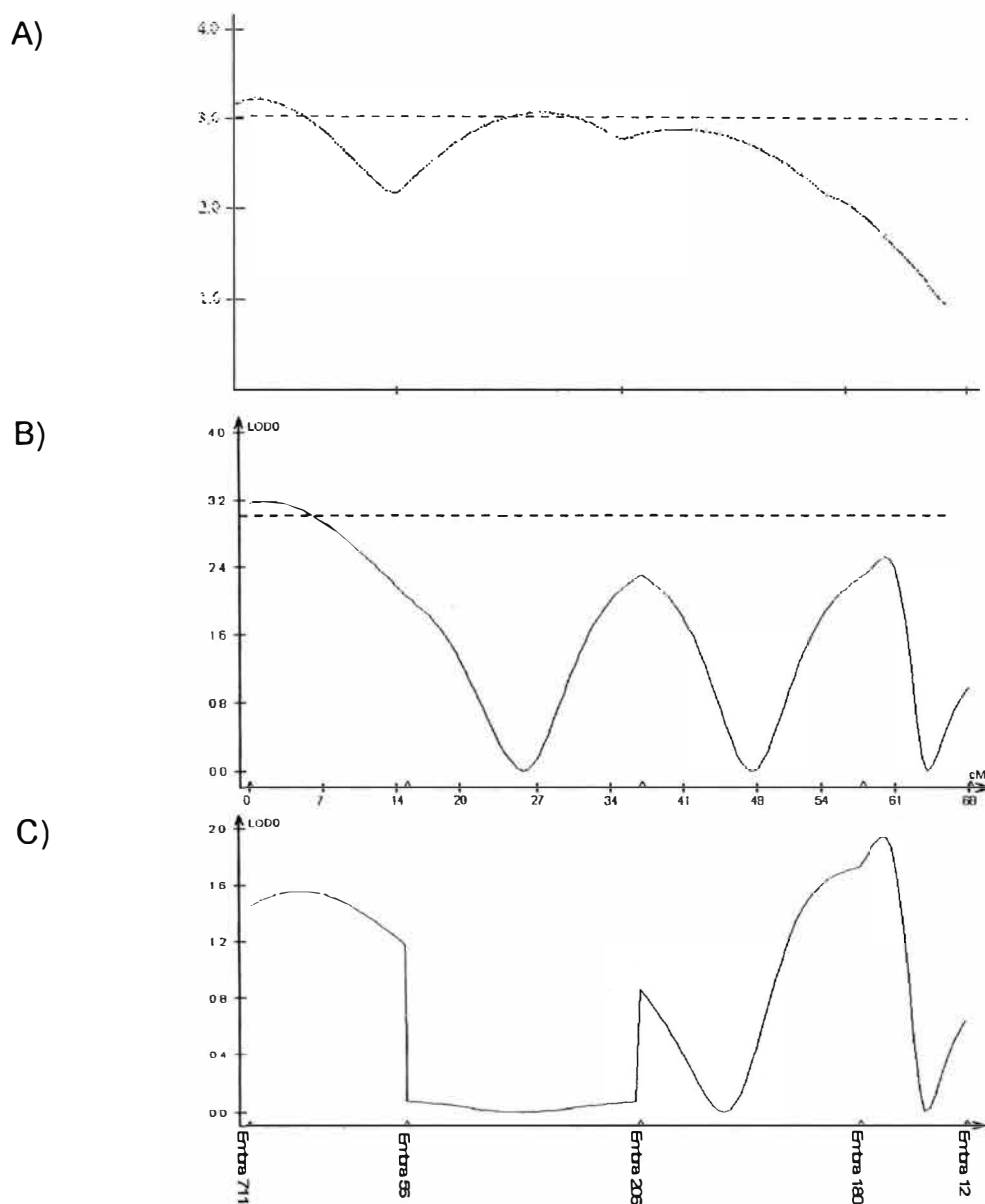


Figura 19 - Mapeamento de um QTL no grupo de ligação 1 do genitor materno 235 para CAP. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTLCartographer

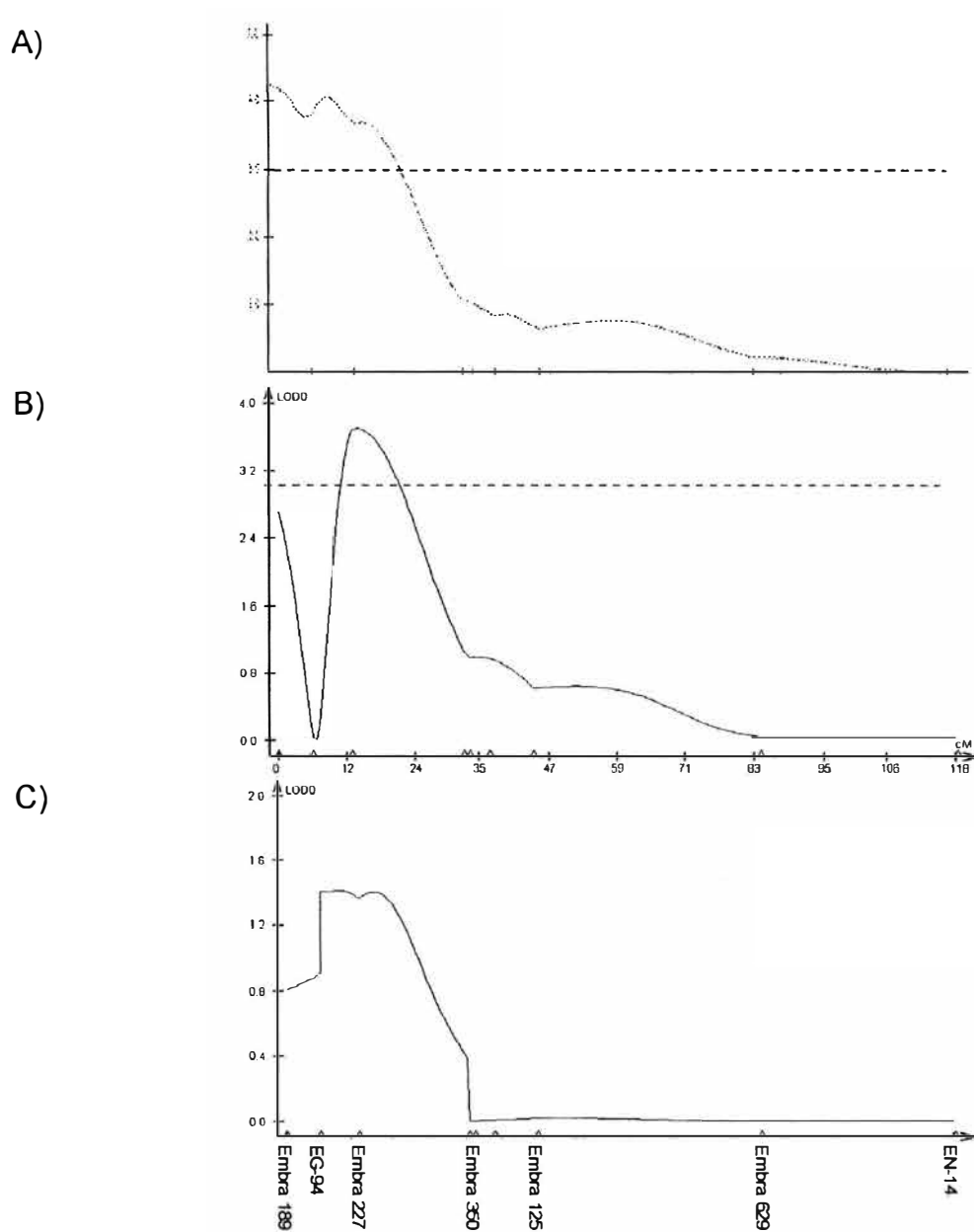


Figura 20 - Mapeamento de um QTL no grupo de ligação 3 do genitor paterno 221 para CAP. Eixo Y: LOD Score; eixo X: Grupo de ligação com indicação de marcadores. (A) Análise de QTL com MIS no programa Mapmaker/QTL; (B) Análise de QTL com MIS no programa QTL Cartographer; (C) Análise de QTL com MIC no programa QTLCartographer

5.3.6. Comparação de mapeamento de QTL por AMI e MIC

Com exceção de CAP, para todas as características fenotípicas avaliadas, foram detectados um ou mais QTL em um ou ambos os genitores pelos procedimentos de mapeamento de intervalo composto utilizando um LOD padrão igual a 3,0 (Tabela 3). Na comparação de resultados de detecção de QTL entre mapeamento por intervalo e mapeamento por marcadores individuais, observou-se que maioria dos QTL detectados por MIC não haviam sido detectados por AMI, particularmente se utilizado o valor crítico após correção de Bonferroni de $\alpha=0,00055$. Isso demonstra o maior poder de detecção de QTL pelo MIC em comparação com AMI.

Uma ilustração resume o posicionamento de todos os QTL detectados para ambos os genitores nos respectivos grupos de ligação permitindo comparar a posição relativa dos QTL nos mapas genéticos dos dois genitores (Figura 21).

Tabela 3. Sumário dos QTL detectados pela análise de MIC nos genitores 221 e 235 para as características fenotípicas estudadas. São descritos o grupo de ligação, marcadores flanqueantes, valor de LOD score, % da variação fenotípica explicada. Para efeito de comparação também é listado o valor p de significância da análise de marcadores individuais (AMI) estimado nos marcadores flanqueantes do QTL

Característica	Genitor	Grupo de ligação	Marcadores flanqueantes	LOD score	% de variação	Valor p
Densidade	221	6	EMBRA 662 EMBRA 94	4,8	9,01	0,000 0,004
Densidade	221	4	EMBRA 137 EMBRA 393	3,4	6,39	0,000 0,000
Lignina	221	2	EMBRA 333 EMBRA 218	4,09	19,84	0,001 0,179
Lignina	235	9	EMBRA 131 ES-140	2,92	39,13	0,984 0,918
Pentosanas	235	6	EMBRA 28 EMBRA 627	3,6	14,6	0,047 0,052
Pentosanas	235	9	EMBRA 57 EMBRA 53	3,1	36	0,016 0,075
Fibra/Grama	235	10	EMBRA 61 EMBRA 194	3,0	22,46	0,047 0,038
Coarseness	235	4	EMBRA 186 EMBRA 19	4,0	53	0,060 0,150
Coarseness	235	9	EMBRA18 ES-140	3,5	53	0,650 0,525
Coarseness	235	3	EG-94 EMBRA 227	4,1	58	0,112 0,048
Rendimento	221	3	EMBRA 125 EMBRA 629	2,9	23,64	0,017 0,014
Rendimento	235	9	EMBRA 131 ES-140	3,5	47,75	0,942 0,561
Cons. Esp.	235	9	Embtra 217 EMBRA 131	4,2	20,6	0,028 0,699
CAP #	221	3	EMBRA 189 EMBRA 350	3,7	9,32	0,000 0,028
CAP#	235	1	EMBRA 711 EMBRA 56	3,18	9,03	0,000 0,002

QTL para CAP foram mapeados exclusivamente com base em mapeamento de intervalo simples e análise de marcadores individuais.

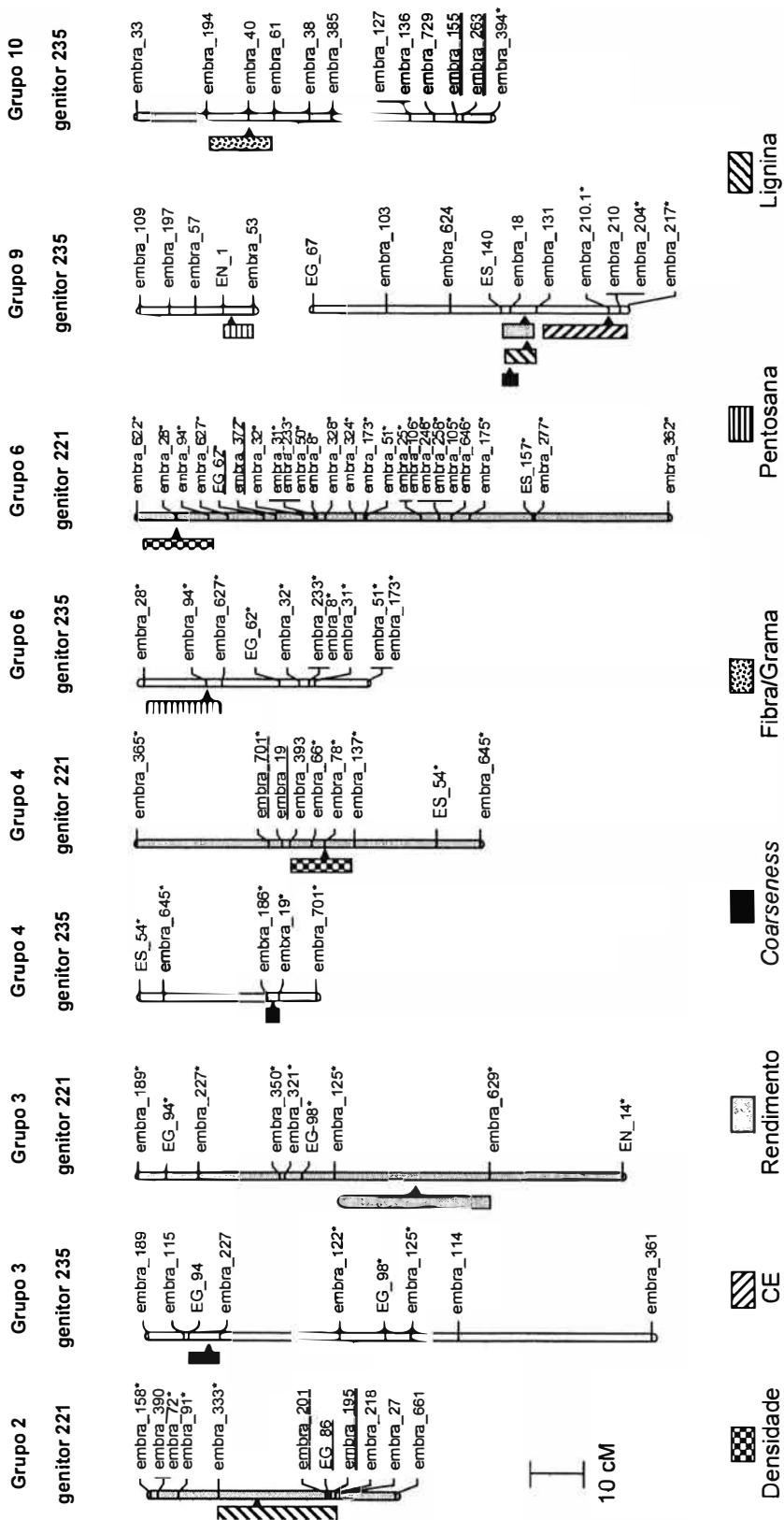


Figura 21 - Localização de QTLs para as várias características fenotípicas nos grupos de ligação dos genitores 235 e 221. São mostrados apenas os grupos de ligação nos quais foram mapeados QTLs com base em mapeamento de intervalo composto utilizando como nível de significância um LOD 3,0

5.3.7 Detecção de QTL com base em LOD crítico definido por permutação

Valores de LOD críticos para as análises de mapeamento por intervalo simples e intervalo composto implementadas no QTL Cartographer foram definidos para cada genitor por permutação para cada uma das oito características (Tabela 4). O objetivo foi verificar a robustez de QTL definidos com base em um LOD padrão de 3,0 em comparação ao LOD definido por permutação que corrige para efeitos de distorção de segregação, variação ambiental, tamanho amostral reduzido e dados incompletos (Doerge 2001).

Tabela 4. Valores de LOD críticos para declaração de significância de QTL, obtidos pela metodologia de permutação (Churchill e Doerge 1994) para as características quantitativas estudadas nos mapas dos dois genitores

Característica	LOD para genitor paterno	LOD para genitor materno
<i>Coarseness</i>	3,8	4,1
CAP	3,0	2,8
Densidade básica	2,8	2,8
Teor de lignina	3,1	3,2
Teor de pentosanas	3,0	2,9
Fibra/Grama	2,9	2,9
Rendimento de celulose	3,5	3,5
Cons. Específico	3,8	3,8

De maneira geral o LOD crítico para declaração de QTL definido por permutação foi maior que o LOD crítico teórico assumido (LOD 3,0). Mesmo

assim, tendo em vista que a maioria dos QTL haviam sido declarados com LOD elevados, o quadro geral de declaração de QTL se manteve relativamente igual ao aplicar a correção por permutação. Utilizando os valores de LOD score calculados conforme descrito por Churchill & Doerge (1994), os QTL para teor de lignina e localizado no grupo de ligação 9 (LOD 2,92) e para coarseness localizados nos grupos de ligação 9 (LOD 3,5) e 4 (LOD 4,0) do genitor materno 235 não seriam selecionados. Para o genitor paterno 221, apenas o QTL (LOD 2,9) para teor de celulose, encontrado no grupo de ligação 3, não seria selecionado.

5.4 Discussão

Dos 326 locos microssatélites testados, 169 foram utilizados para a montagem dos mapas genéticos para os dois genitores e o mapa integrado. Foi observada distorção de segregação para 43,05% dos marcadores mapeados para o parental 235 e 54,71% para o parental 221. Marcadores com distorção de segregação da razão esperada de 1:1 tem sido observados em todos os trabalhos de construção de mapas genéticos para espécies de eucalipto, variando de 5% a 30,9% para cruzamentos interespecíficos envolvendo indivíduos de espécies puras (Grattapaglia et al., 1994; Marques et al., 1998; Myburg et al., 2003) e em menor escala em cruzamentos intraespecíficos (Byrne et al., 1995; Thamarus et al., 2002). Uma parte desta distorção de segregação pode ser explicada por seleção gamética, i.e. genes ligados aos marcadores analisados que reduzem a capacidade de fertilização dos gametas ou seleção pós-zigótica, i.e genes ligados aos marcadores que reduzem a viabilidade dos descendentes, resultando em frequências distorcidas de uma classe genotípica do marcador em relação à outra. Entretanto neste trabalho diferentemente dos demais realizados até hoje com exceção do trabalho de Myburg et al. (2003) os genitores utilizados são híbridos interespecíficos. Neste caso espera-se que a frequência de marcadores com distorção de segregação

seja mais elevado devido a possíveis problemas de pareamento dos cromossomos homólogos de espécies distintas durante a meiose no híbrido. De fato, em um trabalho de construção de um mapa genético envolvendo um híbrido F1 *E. grandis* com *E. globulus* como um dos parentais, Myburg et al. (2003) verificaram uma distorção de segregação em 28,5% dos marcadores quando este híbrido foi utilizado como genitor paterno e de 30,6% quando o híbrido foi utilizado como genitor materno. Neste trabalho, a proporção de marcadores com segregação distorcida foi significativamente mais alto, da ordem de 40 a 50%. Uma explicação para isso pode ser o tipo de marcador utilizado o qual pode influenciar na estimativa da proporção final de marcadores distorcidos. No trabalho de Myburg et al. (2003) assim como a grande maioria dos mapas genéticos construídos, foram utilizados marcadores dominantes RAPD ou AFLP em uma configuração de pseudo cruzamento teste. Nestes casos dos inúmeros marcadores polimórficos gerados e que segregam na população de mapeamento existe uma tendenciosidade (*bias*) para favorecer a interpretação somente daqueles marcadores que mais se aproximam da segregação esperada de 1:1 fazendo com que diversos marcadores que estejam distorcidos sejam descartados antes mesmo de serem considerados. No caso de se utilizar marcadores microssatélites, todos os marcadores que apresentam robustez na amplificação e segregação são interpretados sem qualquer tendenciosidade de seleção daqueles que mais se aproximam da razão de segregação esperada. Com base nesta hipótese, as proporções de marcadores com distorção de segregação estimadas neste trabalho devem se

aproximar com mais precisão das proporções reais a serem observadas em cruzamentos de Eucalipto envolvendo híbridos como genitores de cruzamentos.

Na construção do mapa integrado um total de 70,1% dos marcadores apresentaram distorção de segregação a um nível de 0,05. Esta proporção elevada pode ser facilmente explicada pelo fato de que o conjunto de marcadores com distorção na meiose de um dos genitores não é coincidente com o conjunto de marcadores distorcidos do outro genitor. Uma vez que segregações do tipo 1:1:1:1 são, na verdade o produto das segregações esperadas 1:1 em cada genitor separadamente, espera-se uma proporção aumentada de distorção ao analisar as segregações combinadas pois um marcador não distorcido na análise dos gametas de um parental individualmente, pode se revelar distorcido ao se reunir a sua análise com a análise de segregação distorcida proveniente do outro genitor. Além disso, pode ainda haver seleção pós-zigótica na F1 devida ao encontro específico de alelos deletérios a genes ligados aos marcadores analisados.

O mapa genético desenvolvido para o genitor paterno 221 apresentou um tamanho total de 536 cM, já o mapa para o genitor materno 235 apresentou um tamanho de 819 cM. Dois aspectos devem ser avaliados ao se comparar estes comprimentos totais de mapas com mapas anteriormente publicados. O primeiro é relativo á cobertura genômica efetiva obtida com os cerca de 100 microssatélites mapeados no mapa suporte (*framework map*). Para isso é necessária uma comparação entre os tamanhos observados e os tamanhos

esperados dos mapas. Os tamanhos esperados dos genomas de acordo com as estimativas pelo método de Hulbert et al. (1998) foram respectivamente de 1107 cM e 818 cM para os genitores materno e paterno. A cobertura genômica obtida foi, portanto, de 74,36% para o mapa do genitor materno e 66,67% para o genitor paterno.

O segundo aspecto refere-se ao tamanho esperado em comparação com os tamanhos esperados de mapas publicados. Os comprimentos esperados destes mapas estão abaixo da média dos mapas desenvolvidos para eucalipto até o momento, que apresentam tamanho em torno de 1200 a 1400 cM com exceção dos mapas construídos por Marques et al. (1998) que apresentaram tamanho de cerca de 900 cM. Esta comparação permite fazer inferências quanto à restrição de recombinação em regiões genômicas específicas. Foi observada uma redução significativa nos comprimentos totais esperados dos mapas deste trabalho em comparação com mapas de trabalhos anteriores. Uma redução generalizada de recombinação é frequentemente observada em genomas híbridos devida á supressão de recombinação em regiões divergentes entre os dois genomas parentais (Williams et al., 1995; Kreike & Stiekema, 1997; Chetelat et al., 2000). Entretanto em um híbrido de *E. grandis* x *E. globulus*, Myburg et al. (2003) não observou este fenômeno ao comparar o comprimento total de mapa entre o híbrido e as duas espécies puras embora diferenças significativas tenham sido observadas de forma

localizada em alguns grupos de ligação possivelmente como reflexo de diferentes coberturas de mapa e não de supressão de recombinação.

Apesar de se ter observado altas taxas de distorção de segregação alélica para os locos analisados, os mapas individuais apresentaram, no geral uma consistência da ordem das marcas dentro de cada grupo de ligação. Para o mapa integrado, foi realizada uma análise de sintenia com o mapa gerado por Brondani et al. (2000). Apesar de encontrar algumas divergências quanto ao ordenamento das marcas, os grupos de ligação foram coincidentes (dados não mostrados).

No total, foram detectados 15 QTL para todas as características avaliadas, incluindo CAP para a qual foram detectados somente por MIS e AMI (Tabela 3). QTL para teor de pentosana e densidade foram localizados na mesma região do grupo de ligação 6 em ambos os parentais (Figura 21). Este resultado representa uma validação indireta da declaração de QTL nesta região. Para as outras características analisadas, não foi possível localizar um QTL no mesmo grupo de ligação para ambos os parentais. A detecção de QTL no mesmo intervalo em ambos os parentais depende essencialmente da ocorrência de heterozigiosidade ao QTL em ambos os genitores e ainda dos alelos derivados de cada genitor terem efeito de substituição suficientemente diferente para permitir a detecção do QTL em uma família de 188 plantas. No grupo de ligação 9 do genitor materno 235 foram co-localizados QTL no mesmo intervalo do grupo de ligação para diferentes características, incluindo consumo

específico, *coarseness*, lignina e rendimento (Figura 21). Co-localização de QTL para diferentes características também foram observados por Grattapaglia et al. (1995), Grattapaglia et al. (1996), Marques et al. (1999), Thamarus et al. (2004). É interessante observar que teor de lignina e rendimento em celulose são características, em princípio, negativamente correlacionadas. Um maior rendimento em celulose reflete um menor teor de lignina. Este potencial correlação sugere poder se tratar de um QTL com efeito pleiotrópico.

Foram utilizados diferentes métodos estatísticos para a detecção e localização de QTL visando comparar os resultados tendo em vista as diferentes características de cada método em termos de proteção contra erros do tipo I e tipo II (Doerge, 2001). As análises de detecção de QTL foram primeiramente realizadas com AMI com um nível de significância para todo o experimento de 0,05 com base em correção de Bonferroni para múltiplos testes, com o objetivo de fazer uma primeira triagem de detecção. Este método não permite a localização precisa do QTL por não levar em consideração as relações de distância entre marcadores mapeados. Com AMI foram detectados QTL somente para CAP e densidade. Relaxando o nível de significância para 0,1 para todo o experimento foram também detectados marcadores ligados a QTL para teor de lignina e consumo específico (Tabela 2). Análises de detecção e posicionamento dos QTL por mapeamento de intervalo simples (MIS) foram realizadas utilizando os programas QTL Cartographer e Mapmaker/QTL. Análises de MIC foram realizadas para localização desses QTL, utilizando

também o programa QTL Cartographer. Mapeamento de intervalo tem maior poder estatístico de detecção e permite a localização mais provável do QTL no intervalo por utilizar a informação da ordem linear dos marcadores ao longo do grupo de ligação e com isso levar em consideração os efeitos confundidos da magnitude do efeito do QTL e a sua distância do marcador mais próximo. O mapeamento de intervalo composto possui ainda mais poder de detecção (menor probabilidade de erro do tipo II) por incluir outros marcadores como cofatores. É interessante observar que alguns QTL foram detectados por análises de MIS com o programa QTL Cartographer e não com o programa Mapmaker/QTL (Figuras 13 e 15). Por outro lado todos os QTL localizados com análises de MIS utilizando o programa Mapmaker/QTL, também foram localizados quando a mesma análise foi realizada com o programa QTL Cartographer (Figuras 6, 7, 17, 19 e 20). Estas diferenças de resultados não são esperadas considerando que ambos os programas teoricamente implementam o mesmo algoritmo de detecção e localização. Uma explicação para esta observação demandaria uma análise comparativa mais aprofundada dos detalhes computacionais implementados pelos dois programas. Um outro resultado interessante, foi a detecção de um QTL para CAP utilizando MIS com os dois diferentes programas e AMI mas que não foi detectado pelas análises de MIC (Figuras 19 e 20 e Tabela 3). Estes resultados sugerem ser interessante realizar múltiplas análises de QTL com diferentes métodos estatísticos que possuem diferentes poderes estatísticos e níveis de estrigência antes de se declarar um QTL. Em uma abordagem mais conservadora visando evitar erros

do tipo I seriam declarados somente aqueles QTL detectados com base em AMI com correção de Bonferroni e via MIS. Entretanto esta abordagem pecaria pela possibilidade de erros do tipo II. Os resultados deste trabalho demonstram claramente este fato, ao revelar que com AMI somente QTL para CAP e densidade foram detectados enquanto que com base em MIC foram detectados QTL para sete das oito características estudadas. Curiosamente para CAP somente foram detectados QTL com base em AMI e MIS mas não com MIC.

Foram realizadas análises de sintenia de marcadores e QTL envolvendo os mapas individuais construídos neste trabalho com outros mapas genéticos construídos para eucalipto, que envolveram detecção de QTL ou mapeamento de genes candidatos codificantes de enzimas que participam da via de lignificação. A possibilidade de mapeamento comparativo foi relativamente limitada considerando que poucos microssatélites foram mapeados em outros mapas genéticos construídos para eucalipto. No mapa de Thamarus et al. (2002) foram localizados 40 microssatélites. Mesmo assim, foram observadas co-localizações tentativas interessantes entre QTL neste estudo e genes da via de lignificação e QTL em outros mapas. No mapa genético construído por Thamarus et al. (2002) o marcador microssatélite En-14 está posicionado a 21 cM do gene CCOAOMT, importante gene envolvido na via de lignificação. Um QTL para ângulo de microfibrila foi posicionado em um intervalo de 25 cM envolvendo este marcador (Thamarus et al. 2004). Este marcador, posicionado no grupo de ligação 3 do genitor paterno 221 dista 26,5

cM do marcador EMBRA 629 que flanqueia um QTL para rendimento de celulose (Figura 17). Ainda em relação ao mapa genético gerado por Thamarus et al. (2002), o marcador microssatélite EG-86, está posicionado a 48cM do gene PAL (outro gene envolvido na via de lignificação). Dois QTL para rendimento em celulose e teor de celulose foram mapeados a cerca de 25 cM deste microssatélite (Thamarus et al. 2004). No mapa gerado para o genitor paterno 221, este marcador está posicionado dentro do intervalo de suporte de um QTL para teor de lignina e a cerca de 25 cM da posição mais provável deste QTL (Figura 3.21). Thamarus et al. (2004) localizaram um QTL para comprimento de fibra junto ao microssatélite EG-94 associado a um QTL para comprimento de fibra e um QTL co-localizado para ângulo de microfibrilas. Este marcador, posicionado no grupo de ligação 3 do genitor materno 235 flanqueia um QTL para *coarseness* a 5 cM da posição mais provável (Figura 14). Esta característica está relacionada a espessura de fibra, densidade e número de fibras por grama. Ainda no trabalho de mapeamento de QTL de Thamarus et al. (2004), os marcadores microssatélites EG-67 e ES-140, encontram-se ligados no mesmo grupo de ligação, sendo que o EG-67 flanqueia um QTL para comprimento de fibra. Estes dois marcadores, também posicionados no grupo de ligação 9 do genitor materno 235, entretanto o marcador ES-140 está associado a vários QTL co-localizados para características de rendimento de celulose, teor de lignina e *coarseness* (Figura 21).

É importante ressaltar, entretanto, que todas estas comparações são tentativas e não podem ser realizadas com a precisão necessária para declarar inequivocamente que se tratam de genes candidatos ou de QTL sintênicos entre trabalhos independentes tendo em vista que: (1) o número de marcadores microssatélites presentes nos únicos mapas genéticos que permitem comparação é ainda limitado; (2) a ordem relativa dos microssatélites em relação aos QTL mapeados pode variar de acordo com o conjunto específico de meioses analisadas em cada experimento; (3) a precisão do posicionamento de QTL em todos os trabalhos é limitada por diversos fatores tais como qualidade das avaliações fenotípicas, número de indivíduos analisados na descendência e tipo de marcador utilizado.

Apesar das limitações inerentes a um exercício de mapeamento comparativo entre experimentos independentes, é interessante observar que, mesmo com muito poucos microssatélites em comum entre os mapas de QTL de Thamarus et al. (2004) e os mapas deste trabalho, foi possível detectar diversos casos de sintonia entre QTL e entre QTL e genes candidatos o que sugere não se tratarem de coincidências mas sim de evidências efetivas de validação interespecífica do posicionamento de QTL para características físico-químicas da madeira de grande importância econômica em eucalipto.

A validação da real existência de QTL em determinados segmentos genômicos depende em grande parte de experimentos de mapeamento comparativo entre diferentes backgrounds genéticos. Estes resultados abrem,

portanto, uma perspectiva muito promissora para futuros experimentos de mapeamento comparativo de QTL entre diferentes pedigrees e espécies tendo como base o mapeamento de microssatélites transferíveis e polimórficos, como proposto por Grattapaglia (2000) alguns anos atrás. Estes resultados reforçam ainda a necessidade e significativa vantagem de se utilizar marcadores microssatélites, mesmo que apenas algumas dezenas, e não marcadores dominantes de baixa ou nenhuma transferibilidade tais como RAPD e AFLP, em experimentos de mapeamento genético e localização de QTL. Mapas genéticos altamente saturados construídos com centenas ou milhares de marcadores AFLP ou RAPD tem uma aplicabilidade muito limitada ou praticamente nula para exercícios de análise de sintenia e validação de QTL.

5.5 Conclusão

Este trabalho teve por objetivo a construção de um mapa genético utilizando marcadores microssatélites, a localização de QTL e a demonstração da utilidade de microssatélites para análises comparativas de mapeamento entre experimentos independentes, particularmente quanto à localização de QTL. Embora mapas genéticos baseados em microssatélites tenham sido gerados (Brondani et al., 2000) e alguns marcadores microssatélites tenham sido utilizados para mapeamento de QTL (Thamarus et al., 2004) este é o primeiro trabalho descrevendo a detecção e mapeamento de QTL utilizando exclusivamente microssatélites. A quase totalidade dos trabalhos anteriores de mapeamento de QTL em eucalipto se basearam em marcadores RAPD (Grattapaglia et al., 1996, Verhaegen et al., 1997; Junghans et al., 2003) ou marcadores AFLP (Marques et al. 1999; Myburg, 2001). QTL mapeados com marcadores de baixa ou nula transferibilidade entre indivíduos ou espécies distintas de *Eucalyptus*, representam interessantes demonstrações da arquitetura genética de características quantitativas. Entretanto em nada contribuem para o verdadeiro desafio atual de validação de QTL visando, em seguida, realizar experimentos de mapeamento de associação com genes

candidatos co-localizados com QTL validados ou mesmo identificar e clonar os genes responsáveis pelos QTL.

Marcadores microssatélites, por outro lado, por se basearem em ampliações de regiões específicas do genoma, uma vez polimórfico e transferível entre espécies, pode ser utilizado para gerar informações de segregação alélica em outros cruzamentos mesmo de espécies distintas. Embora a resolução de mapeamento de QTL dos trabalhos utilizados nas comparações não permita declarar de forma certa que foi possível validar QTL, os resultados de mapeamento comparativo de QTL detectados neste trabalho com QTL detectados em trabalhos anteriores com experimentos independentes indicam que esta estratégia é possível. Este trabalho abre assim uma importante perspectiva na investigação detalhada das bases moleculares e arquitetura genética de características quantitativas em diferentes backgrounds genéticos seja de indivíduos bem como espécies de *Eucalyptus*.

A resolução de mapeamento de QTL com marcadores microssatélites obtida neste trabalho é da ordem de dezenas de cM o que não permite ainda utilizar os marcadores microssatélites para iniciar buscas de clonagem posicional. Para isso, a análise de um número muito maior de indivíduos e uma avaliação fenotípica de alta precisão são necessárias a fim de estreitar ao máximo o segmento genômico no qual o QTL se localiza. Além disso, tecnologias de marcadores que permitem uma elevada saturação do segmento genômico alvo são necessárias tais como AFLP ou SNPs de forma a identificar

marcadores que co-segrem perfeitamente com o fenótipo. Enquanto isso já tenha sido demonstrado para diversas características qualitativas ou de herança simples em plantas anuais (Morgante & Salamini 2002) o desafio para características multifatoriais é muito maior e ainda não existem exemplos de identificação e clonagem de genes responsáveis por QTL de características efetivamente quantitativas.

6 REGULAÇÃO ATUANDO EM CIS NA EXPRESSÃO ALELO ESPECÍFICA DE GENES DE LIGNIFICAÇÃO EM *EUCALYPTUS*.

Resumo

Um método para detectar variações do nível de transcrição de alelos, baseado em diferenças nas seqüências atuando em *cis*, foi aplicado a genes envolvidos na via de lignificação do Eucalipto. Genes envolvidos nesta via podem ser utilizados para testar genes candidatos e caracterizar QTL envolvidos no controle da qualidade da madeira em plantas arbóreas. Neste trabalho foi utilizada uma metodologia genômica recente que combina mensuração de níveis de expressão diferencial alelo específica em genes candidatos usando RT-PCR com a discriminação dos transcritos correspondentes aos diferentes alelos com base na variabilidade de bases individuais, SNPs (*Single Nucleotide Polymorphisms*). A discriminação dos SNPs foi feita com a técnica de extensão de base individual (SBE - *Single Base Extension*). Desta forma, é possível distinguir na progênie os transcritos herdados de cada um dos parentais e detectar variações de expressão de mRNA relacionadas a cada alelo. As análises foram realizadas para os parentais e 40 indivíduos da progênie de um cruzamento envolvendo

Eucalyptus grandis (G38) e *Eucalyptus urophylla* (U15). A busca por identificação de SNPs em regiões transcritas foi realizada utilizando DNA genômico dos parentais para os genes: Fenilalanina amônia-liase (PAL) ácido cafeico o-metiltransferase (COMT), cafeoil-CoA o-metiltransferase (CCOAOMT), ferulato hidroxilase (5FH), 4 coumarato-CoA ligase (4CL), cinamoil-CoA redutase (CCR) e o gene álcool cinamil desidrogenase (CAD). De todos os genes estudados, foram localizados SNPs em regiões transcritas para os genes das enzimas CAD, CCR e PAL. Para o gene da CAD, o genitor G38 apresentou o genótipo heterozigoto AC e o genitor U15, o genótipo homozigoto AA. Para o gene da CCR o genitor G38 apresentou o genótipo heterozigoto CT e o genitor U15 o genótipo homozigoto CC. Para o gene da PAL genitor G38 apresentou o genótipo heterozigoto AG e o genitor U15 apresentou o genótipo homozigoto AA. Estudos de expressão alélica foram realizados para os indivíduos heterozigotos da progênie apresentando genótipos AC para o gene CAD, CT para o gene CCR e AG para o gene PAL, utilizando mRNAs extraídos de xilema. Para o gene da CAD, foram analisados 18 indivíduos, dos quais 4 apresentaram expressão do alelo C superior a 60%, representando uma expressão 1,5 vez maior que a expressão do alelo A. Já para o gene da CCR, foram analisados 22 indivíduos, dos quais 11 apresentaram valores de expressão para o alelo T superior a 60%, representando uma expressão 1,5 vez maior que a expressão do alelo C. Para este gene, a segregação dos níveis de expressão dos alelos segregou na razão de 1:1, sugerindo um possível efeito de ação combinada de fator em trans e fator em cis ("*Cis + Trans acting*")

afetando a expressão dos alelos para este gene. Para o gene da PAL, não foi possível amplificar seqüências transcritas. Esta metodologia apresenta um amplo potencial de aplicação para aprofundar conhecimentos rumo a um melhor entendimento do comportamento dos genes envolvidos na via de lignificação.

DETECTION OF CIS-ACTING REGULATORY VARIATIONS THAT AFFECTS LIGNINE GENES IN *EUCALYPTUS*.

Summary

The study of candidate genes that putatively affect trait expression, has been an important strategy to characterize a QTL and understand the biochemical or developmental pathways affecting the trait of interest. Biosynthetic (structural) genes of the lignification pathway constitute a unique opportunity for testing the candidate gene approach and potentially characterizing the QTL involved in the control of wood quality in woody plants. Variation in non-coding cis-regulatory DNA sequences, which affects gene expression levels, has also been proposed as a major component of the genetic basis for phenotypic evolution, but it is mostly unknown. A method to detect changes in transcript levels, due to cis acting sequence differences, without the necessity to recognize specific regulatory variants that can be hundreds or even thousands of bases upstream from the transcription unit, has been developed in *Eucalyptus*, based on studies about methods for measurement of allelic levels of gene expression. The assay involves the study of two alleles of a gene under

identical conditions and the comparison of the transcript expression level associated with each. To distinguish between the transcripts derived from each of the two parental alleles we have used SNP markers (Single Nucleotide Polymorphisms) in the transcript itself. The assay involves RT-PCR amplification in the region surrounding the SNP marker, followed by a SBE (Single Base Extension) using a primer adjacent to the variant base in the presence of fluorescently labeled nucleotides and detection on a capillary DNA sequencer. The ratio of the levels of the two alleles is inferred by comparison with known mixtures of both parental genomic DNA used as reference standard, which after PCR amplification, are subjected to the same SNP assay. Eucalyptus species have a high level of heterozygosity so that a high degree of nucleotide diversity affecting both coding sequence and expression level of genes could be detected. For this study we set a survey of polymorphisms among a list of candidate genes involved in the lignine biosynthetic pathway including Phenylalanine ammonia lyase (PAL), Caffeate/5 hydroxyferulate o-Methyltransferase (COMT), CCOAOMT Caffeiol o-Methyltransferase (CCOAOMT), Ferulate 5-hydroxylase (5FH), 4 coumarate CoA ligase (4CL), Cinnamoyl CoA Redutase (CCR), Cinnamyl Alcohol dehydrogenase (CAD). For all these genes, SNPs were localised in the transcription region of the genes coding for PAL, CCR and CAD. Studies of allelic expression for the CAD gene was carried out using stem mRNA extraction for 18 individuals. We analysed the selected genes in the parents and in 16 offsprings from the a cross between *Eucalyptus grandis* X *Eucalyptus urophylla*. By using the SNP markers detected in the transcribed

regions we could distinguish between transcripts derived from each of the two parental alleles in order to reveal abundance variation in the mRNA related to each single allele. For CCR the analysis was carried out using stem mRNA extracted from 22 individuals. For the CAD gene, 4 of 18 individuals analysed showed expression of C allele higher than 60%, representing an expression profile 1,5 times more than A allele. For the CCR gene, 11 of 22 individuals analysed showed expression of the T allele higher than 60%, representing an expression profile 1,5 times more than C allele. For this gene, the segregation of expression levels for its alleles was in a 1:1 ratio, suggesting a possible combining effect of a cis acting and trans acting ("cis + trans acting") affecting the alleles expression for this gene. For the PAL gene it was not possible to amplify transcribed sequence. An hypothesis to explain the non amplification for this gene could be the presence of different isoforms for PAL in different cellular tissues. This methodology present a powerfull application to better understand the behavior of the genes involving on the lignin pathway.

6.1 Introdução

Tecnologias genômicas de alto desempenho, integradas nos programas de melhoramento florestal, estão abrindo novas perspectivas para o estudo de genes e genomas, facilitando a compreensão das relações complexas entre variabilidade genética e diversidade fenotípica. Avanços relacionados à construção de mapas genéticos e localização de QTL, apesar de serem ferramentas importantes para a focalização de esforços de investigação em regiões específicas do genoma e possivelmente permitir a seleção direcional de indivíduos superiores, são estratégias que não permitem a inferência direta dos elementos responsáveis pela variação fenotípica, tais como quais e quantos genes estão envolvidos na diferenciação de um fenótipo, quais são os padrões de expressão desses genes ou até mesmo a ação de elementos em *cis*, promotores, envolvidos no processo de regulação dos genes. Um dos aspectos fundamentais para avançar rumo ao entendimento mais detalhado de características multifatoriais e economicamente relevantes em eucalipto é a aplicação de metodologias que permitam o conhecimento do comportamento de genes de maior efeito que participam das vias metabólicas relacionadas às características de importância econômica tais como formação da madeira e resistência a doenças.

Nos últimos anos, um esforço crescente tem sido feito no sentido de investigar as bases moleculares da formação, estrutura e função da madeira. As propriedades físicas de interesse incluem, por exemplo: densidade básica, ângulo das microfibrilas, espessura da parede celular, diâmetro do lúmen e comprimento das fibras. Propriedades químicas de interesse incluem, por sua vez, a abundância e composição dos três principais componentes da madeira, celulose, hemicelulose e lignina. A via de biossíntese dos precursores da lignina é a mais bem conhecida no processo de formação da madeira e tem sido o foco da experimentação em biologia molecular. A maioria dos genes que codificam para as enzimas conhecidas desta via, bem como fatores de transcrição e proteínas de parede, foram clonados e caracterizados particularmente em *Pinus taeda* (Whetten et al., 1998) e álamo (Sterky et al., 1998). Além disso, árvores transgênicas de álamo superexpressando ou subexpressando algumas destas enzimas apresentaram alterações importantes na quantidade e qualidade final de lignina (Hu et al., 1999; Pilate et al., 2002; Li et al., 2003). Embora genes já tenham sido identificados e expressados em sistemas transgênicos, nada se sabe sobre a relação entre a diversidade nucleotídica dos variantes alélicos destas enzimas presentes na natureza e as características finais da lignina. Potenciais limitações para a utilização em larga escala de árvores transgênicas e a ampla variabilidade existente no gênero *Eucalyptus* para qualidade da madeira, sugerem que o entendimento preciso e a exploração refinada da diversidade gênica natural para estas enzimas poderá ser uma estratégia de grande impacto no melhoramento florestal.

Este trabalho visou avaliar a possibilidade de investigar os níveis de expressão alelo específica de genes envolvidos na via de lignificação de eucalipto selecionados tendo em vista o seu conhecido papel chave na regulação das quantidades e propriedades da lignina produzida. Para isso, foi utilizada uma metodologia genômica recente que combina mensuração de níveis de expressão diferencial alelo específica em genes candidatos usando RT-PCR com a discriminação dos transcritos correspondentes aos diferentes alelos com base na variabilidade de bases individuais, SNPs (*Single Nucleotide Polymorphisms*). A discriminação dos SNPs detectados via sequenciamento de DNA genômico é feita com a técnica de extensão de base individual (SBE - *Single Base Extension*) utilizando primers adjacentes à base polimórfica na presença de dideoxis marcados com fluorocromos e detecção em sequenciador automático (Yan et al., 2002).

Indo além das demais metodologias utilizadas em estudos de expressão de genes, que tipicamente comparam a expressão entre dois ou mais indivíduos, esta estratégia possibilita a discriminação de diferentes níveis de expressão entre os alelos de um mesmo gene para um mesmo indivíduo na mesma condição de background genômico. Esta abordagem fornece, portanto, uma nova possibilidade de conectar genótipo com variabilidade fenotípica determinada por diferenças em níveis de expressão de alelos a genes específicos e não mudanças na estrutura da proteína codificada. Ou seja, permite a investigação de variabilidade herdável que atua em *cis* alterando, em

última análise, a definição final dos fenótipos. Uma questão sempre complexa na investigação da relação entre genótipo e fenótipo é a seleção adequada dos genes candidatos a serem testados. Este estudo permite a análise direta e precisa da variação na expressão de um grande número de genes candidatos. Uma vez identificados genes com padrões diferenciados de expressão interessantes, estudos detalhados podem então ser conduzidos principalmente no que se refere a regiões reguladoras da expressão do gene.

6.2 Material e métodos

6.2.1 Material genético

O experimento de análise de expressão gênica foi realizado utilizando uma progênie de um cruzamento entre o clone G38 (*Eucalyptus grandis*) e o clone U15 (*Eucalyptus urophylla*) plantados na empresa International Paper do Brasil. Esta família faz parte do grupo de famílias segregantes estudadas no âmbito do projeto Genolyptus. Tecido foliar e de xilema de 48 indivíduos tomados ao acaso foram coletados e armazenados em microtubos eppendorfs de 2,0 mL. Para evitar a degradação do RNA, os tubos contendo o material vegetal foram armazenados em nitrogênio líquido até o processamento.

6.2.2 Seleção dos genes a serem estudados

A seleção de genes candidatos ainda é uma questão complexa para a maioria dos fenótipos de relevância em espécies florestais. Esta definição demanda conhecimentos de bioquímica, fisiologia e desenvolvimento muitas vezes ainda não disponíveis, mesmo para fenótipos bem caracterizados ou rotas metabólicas conhecidas. Sabendo que para a realização de estudos de expressão gênica utilizando esta metodologia necessita-se da localização de

SNPs em regiões transcritas nos genes nos genitores, foram desenhados primers para os seguintes genes: Fenilalanina amônia-liase (PAL), cafeico o-metiltransferase (COMT), cafeoil-CoA o-metiltransferase (CCOAOMT), ferulato hidroxilase (5FH), 4 coumarato-CoA ligase (4CL), cinamoil-CoA redutase (CCR), álcool cinamil dehidrogenase (CAD), visando maximizar a cobertura de regiões codificadoras. Estes genes foram selecionados em função do seu conhecido envolvimento na biossíntese de lignina, já bastante caracterizado para *Arabidopsis* (Figura 1). Para obtenção de dados de seqüências para estes genes, foi utilizado o banco de dados do NCBI (*National Center of Biotechnology and Information*).

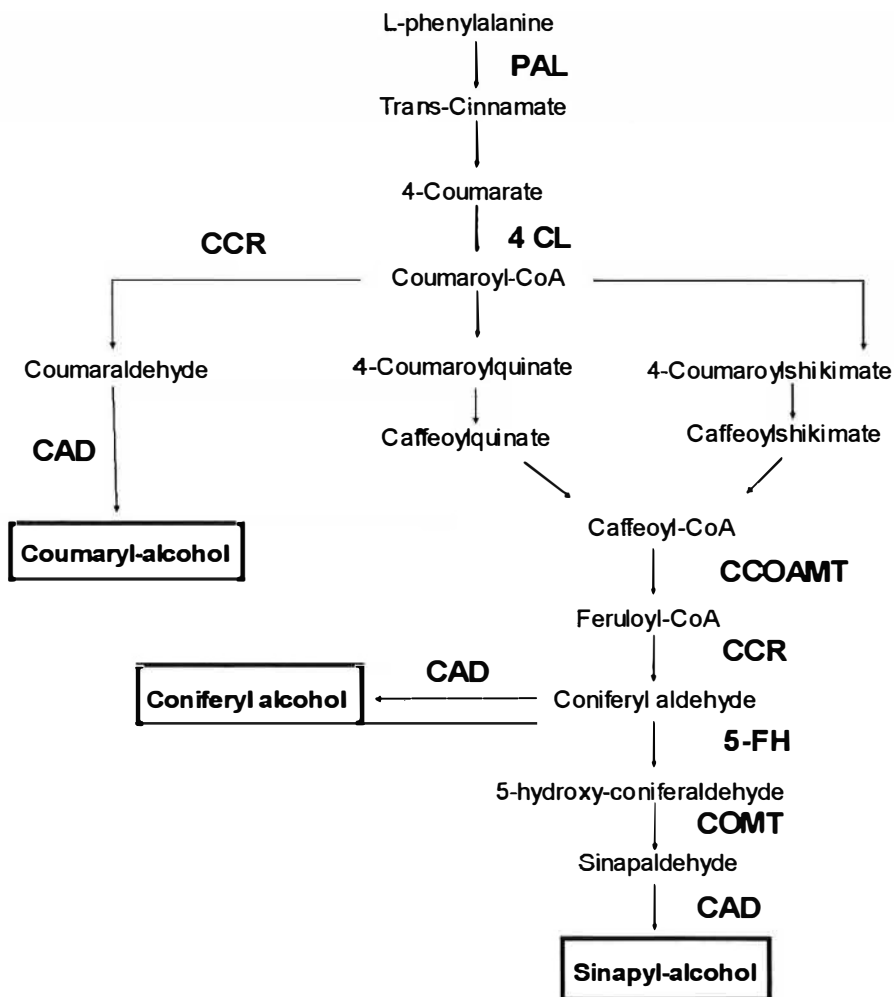


Figura 1 - Via de lignificação (modelo proposto para Arabidopsis) e os genes mais importantes envolvidos na sua biossíntese. PAL: Fenilalanina amônia-liase; 4CL: 4 coumarato-CoA ligase; CCR: cinamoil-CoA redutase; CCOAOMT: cafeoil-CoA o-metiltransferease; 5FH: ferulato hidroxilase; COMT: cafeico o-metiltransferease; CAD: álcool cinamil dehidrogenase

6.2.3 Montagem dos primers

Para a identificação dos SNPs em seqüências transcritas, foi desenhado pelo menos um par de primers para cada um dos genes selecionados e muitas vezes mais de um par. A montagem dos primers foi realizada utilizando o programa Primer3 (<http://www.genome.wi.mit.edu>), Rozen & Skaletsky, 1998). Foram utilizados para a construção dos primers os parâmetros pré-selecionados pelo próprio programa. Os primers foram desenhados em pelo menos duas regiões transcritas do DNA genômico a fim de maximizar a probabilidade de se encontrar um SNP entre os genitores.

6.2.4 Extração de DNA

DNA genômico total foi extraído de tecido foliar de 24 indivíduos. A extração foi feita pelo método CTAB 2% (Doyle & Doyle, 1987). A maceração das folhas foi realizada utilizando-se nitrogênio líquido. O DNA foi utilizado para selecionar os indivíduos heterozigotos via sequenciamento, para os SNPs detectados, bem como para montagem da curva de tratamento visando transformar os valores de altura de pico em expressão alélica.

6.2.5 Quantificação de DNA

Visando obter uma quantificação precisa do DNA extraído tendo em vista a aplicação deste DNA para fins de calibração do sistema de detecção de expressão diferencial em RNA, a quantificação de DNA foi realizada em uma plataforma VICTOR da Applied Biosystems mediante o kit PicoGreen dsDNA Quantification da Molecular probes (www.probes.com).

6.2.6 Extração de RNA total

Para extração de RNA total foi utilizado o kit *Concert™ Plant RNA Reagent* da Invitrogen. Este kit apresenta uma metodologia simples de extração e resulta em uma boa quantidade de RNA total. Para realização da extração de RNA, tecidos de xilema armazenados em microtubos eppendorf de 2,0 mL foram macerados utilizando nitrogênio líquido. As amostras de RNA foram visualizadas em sistema de eletroforese de gel de agarose 1% corado com brometo de etídeo.

6.2.7 Extração de mRNA

Após a extração de RNA total, foi realizada uma extração de mRNA com o intuito de garantir um material livre de DNA. Para isso, foi utilizado o kit *Purification of poly A+ RNA from total RNA* da Qiagen.

6.2.8 Amplificação de cDNA

A metodologia de RT-PCR foi aplicada para a conversão de mRNA em cDNA e amplificação do mesmo visando disponibilizar cópias de DNA dos transcritos alvo. A reação de transcriptase reversa foi realizada seguindo três passos. Como primeiro passo foi feito um aquecimento à 70⁰C do mRNA por 10 minutos e em seguida o material foi armazenado em gelo. A geração de cDNA foi feita utilizando-se o kit *Reverse Transcription System* da Promega. Para realização da reação de cDNA foram utilizados 1% de um tampão da enzima Reverse Transcriptase (100mM Tris-HCL, 500mM KCL, 1% Triton X-100), 5mM de Cloreto de Magnésio, 1 mM de dNTP, 20 unidades de inibidor de ribonuclease (RNAsin), 40 unidades da enzima AMV Reverse transcriptase, 25µg/ml de random primers e 6,5 µL de mRNA, para um total de 20 µL de reação final. Antes de iniciar a reação de cDNA, o mix foi mantido a temperatura ambiente por 10 minutos. A reação de geração da segunda fita de cDNA foi realizada utilizando um programa que apresenta uma temperatura inicial de 42⁰C por 45 minutos, seguido de 95⁰C por 5 minutos e 4⁰C por 5 minutos.

6.2.9 Amplificação dos genes estudados via reação em cadeia da Polimerase (PCR)

Para a amplificação dos genes selecionados usando DNA como molde, utilizou-se 1% de Tampão de PCR, 1,5 mM MgCl₂, 250 µM de dNTP,

0,24 $\mu\text{g}/\mu\text{l}$ de BSA, 5 unidades da enzima Taq polimerase, 6,0 ng de DNA, 1 μM de cada primer (forward e reverse) e 9,75 μL de água Milli-Q autoclavada. Para amplificação dos segmentos gênicos selecionados utilizando cDNA como molde foi utilizado o mesmo protocolo, porém utilizando 0,5 μl de cDNA e 12,25 μL de água Milli-Q autoclavada, para um volume final de 25 μL . A reação de PCR foi realizada em uma placa de 96 poços da Applied Biosystems. A amplificação via PCR foi feita utilizando um programa que apresenta uma etapa inicial de desnaturação a 94^oC por 10 minutos, 35 ciclos com etapas de desnaturação (1 minuto a 94^o C), anelamento (1 minuto utilizando temperaturas variando de 57^o C a 61^o C) e extensão (1 minuto a 72^o C). No final dos ciclos, ocorreu uma etapa de extensão final por mais 8 minutos a 72^o C. Ao termino da reação de PCR, os fragmentos amplificados foram visualizados em sistema de eletroforese de gel de agarose 1% corado com brometo de etídeo. Para verificar se os tamanhos dos fragmentos eram realmente aqueles esperados, estes foram carregados em gel de agarose, juntamente com um padrão de fragmentos de DNA de tamanho conhecido (100 pb Invitrogen).

6.2.10 Purificação da Reação de PCR e Sequenciamento

Para a realização da reação de sequenciamento visando a identificação de SNPs utilizou-se o produto de PCR proveniente da amplificação do DNA dos genitores. Para verificar a consistência da herança e segregação de um possível SNP, a amplificação e sequenciamento do segmento foi

realizado para os genitores e dois descendentes da progênie. Após a confirmação do sucesso da amplificação via PCR, as amostras foram purificadas adicionando inicialmente água Milli-Q autoclavada, para um volume final de 100 μL . Esta diluição foi transferida para uma placa “*PCR 96 Clean up plate*” da Millipore (Montáge). Foi realizada uma filtragem a 10mmHg por 10 minutos no equipamento *Vacuum System* da *Millipore*, e a seguir adicionados 50 μL de água Milli-Q autoclavada. Foi realizada uma agitação a 700 rotações por minuto (RPM) no miniagitador da IKA. O produto de PCR purificado foi transferido para uma outra placa para ser armazenado. A reação de sequenciamento foi realizada utilizando 1,85 μl de tampão 5 X, 0,32 μM de primer, 0,1 μL de BigDye, 2,0 μL de produto de PCR purificado e 5,91 μl de água Milli-Q autoclavada, para um volume final de 10 μL de reação. A reação de sequenciamento foi realizada em uma placa de 96 poços da Applied Biosystems. Para garantir uma maior amplitude de seqüências, a reação de sequenciamento foi realizada nos dois sentidos (5'e 3') para cada gene. Após a realização da reação de sequenciamento, o produto foi purificado adicionando 2,5 μL de EDTA 125 mM e 25 μL de etanol absoluto. A placa foi coberta com adesivo de alumínio e após uma homogeneização por inversão, a mistura foi mantida por 15 minutos a temperatura ambiente. A seguir foi realizada uma centrifugação da placa por 55 minutos a 3000 RPM a 4 $^{\circ}\text{C}$ e posteriormente, o sobrenadante foi descartado. Para acelerar o processo de secagem do DNA sequenciado, foi realizada uma centrifugação a 185 RPM por 1 minuto a 4 $^{\circ}\text{C}$,

com a placa colocada em posição invertida. As amostras foram ressuspendidas em 8,0 μ L de HIDI formamida e submetidas a uma desnaturação por 2 minutos a 94 °C. A eletroforese da reação de sequenciamento foi realizada na plataforma 3730 de 48 capilares da Applied Biosystems.

6.2.11 Análise das Seqüências e Identificação de SNPs em ESTs

Após a realização do sequenciamento, os eletroferogramas das seqüências 5'e 3'dos genitores e dois descendentes da progênie foram alinhadas para cada gene utilizando o programa Phred Phrap Consed (Ewing et al., 1998) em sistema operacional Linux. Visando maximizar o grau de informação das análises, preferência foi dada para SNPs em homozigose em cada genitor mas polimórficos entre eles. Neste caso, todos os indivíduos da progênie poderiam ser utilizados, uma vez que todos eles seriam heterozigotos para o SNP identificado.

6.2.12 Detecção dos Níveis de Expressão dos Transcritos com Base em SNPs

Para a detecção dos níveis de expressão dos transcritos foi utilizada a metodologia de extensão de base individual (Single Base Extension-SBE) com base no kit *Multiplex SnaPshot* da Applied Biosystems (Figura 2). Foram sintetizados primers no sentido 3' e 5' adjacentes ao SNP. Para amplificação

dos SNPs via SBE, o produto da reação de PCR foi purificado utilizando 2,0 μL da enzima *ExoSAP-IT*, 1,0 μL do produto de PCR, e 4,0 μL de água Milli-Q autoclavada, para um total de 7,0 μL . A reação de purificação ocorreu utilizando um programa de 37 $^{\circ}\text{C}$ a 30 minutos, seguido de 15 minutos a 80 $^{\circ}\text{C}$, realizada em termociclador da Applied Biosystems. Para a reação de SBE utilizou-se o Kit *SnapShot* da Applied Biosystems que contém dideoxi nucleotídeos trifosfatos (ddNTPs) marcados com fluorocromos. O ddATP é marcado com o fluorocromo dR6G e fluoresce na cor verde, o ddCTP é marcado com o fluorocromo dTAMRA e fluoresce na cor amarela, o ddGTP é marcado com o fluorocromo dR110 e fluoresce na cor azul, e o ddTTP (uracila) é marcado com o fluorocromo dROX e fluoresce na cor vermelha. As cores referentes às fluorescências são para um filtro virtual específico. Para a reação, foram utilizados 1,5 μL do kit *SnapShot* da Applied Biosystems, 3,0 μL do produto de PCR purificado, 0,4 μM de primer e 4,5 μL de água Milli-Q autoclavada, para um volume total de 10 μL . A amplificação da SBE ocorreu em termociclador da Applied Biosystems utilizando um programa de 25 ciclos seguindo os passos de 96 $^{\circ}\text{C}$ por 10 segundos, 55 $^{\circ}\text{C}$ por 5 segundos e 60 $^{\circ}\text{C}$ por 30 segundos. Ao término da reação de SBE, foi realizada uma outra purificação utilizando a enzima SAP. A purificação foi realizada adicionando à reação de SBE, 1 unidade da enzima SAP. A reação de purificação foi realizada em termociclador da Applied Biosystems utilizando um programa de 37 $^{\circ}\text{C}$ a 60 minutos seguido de 75 $^{\circ}\text{C}$ por 15 minutos. Após a reação de purificação foram adicionados 9,85

μL de formamida e $0,15 \mu\text{L}$ de padrão de fragmentos de tamanho conhecido marcados com fluorescência LIZ (ABI LIZ 50 pares de base). A eletroforese da reação de SBE foi realizada na plataforma 3730 de 48 capilares da Applied Biosystems. Os dados de intensidade de fluorescência dos picos de cada alelo foram convertidos em níveis relativos de expressão de cada alelo utilizando o programa Genemapper da Applied Biosystems.

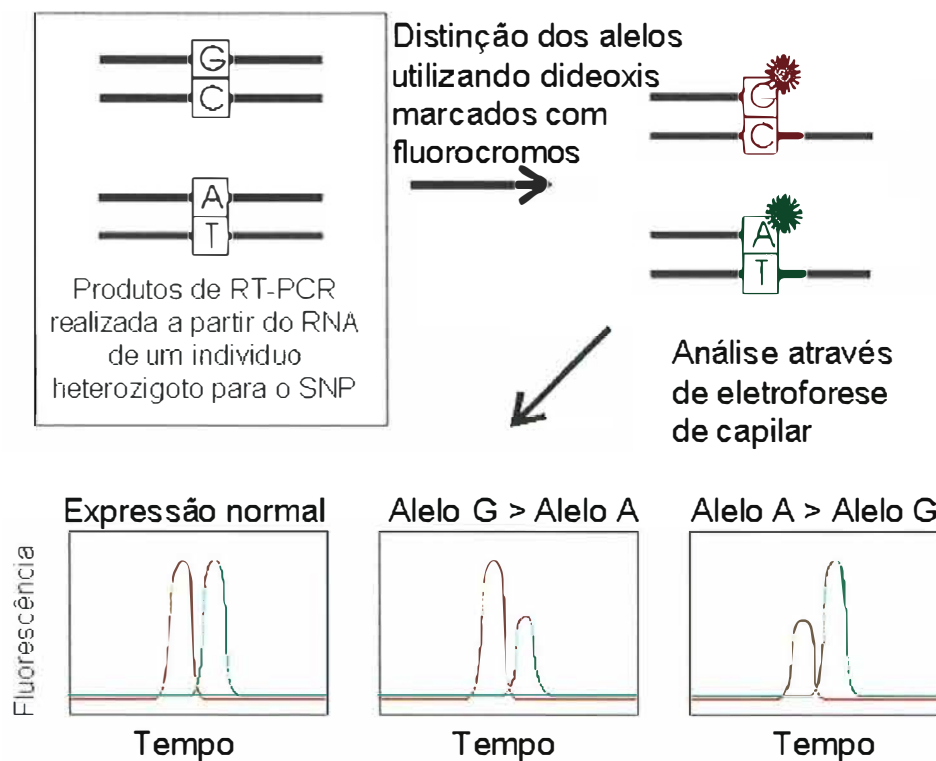


Figura 2 - Estratégia utilizada para mensurar variação na expressão de alelos de genes utilizando uma combinação de RT-PCR e posterior detecção e discriminação dos transcritos com base em extensão de base individual e marcação com terminadores marcados (adaptada de YAN et al., 2002)

6.2.13 Curva de normalização de expressão de mRNA usando DNA genômico

Os níveis de expressão alelo específica dos mRNAs analisados foram inferidos após uma normalização da intensidade de sinal de fluorescência mensurada em rfu (*relative fluorescence units*). Esta normalização foi realizada a partir de uma comparação da intensidade de fluorescência do pico observado na detecção do SNP em cDNA, com a intensidade de fluorescência do pico observado na detecção do mesmo SNP a partir de quantidades conhecidas de DNA molde dos genitores do cruzamento, que após amplificação via PCR foram igualmente submetidos a SBE juntamente com as amostras de mRNA. Foram realizadas misturas de DNAs nas proporções 75:25, 80:20 e 90:10 e nas proporções 75:25, 87,5:12,5 e 62,5:37,5. Os valores referentes aos picos de cada alelo para cada uma das misturas realizadas com os DNAs dos genitores foram utilizados para criar uma equação de regressão de modo a ajustar os valores dos picos observados a partir da SBE conduzida nos mRNAs dos alelos. Após a normalização dos dados, o valor observado da intensidade do pico de fluorescência foi transformado em valor de expressão do alelo específico do gene. Este procedimento foi realizado para cada um dos genes testados.

6.3 Resultados

6.3.1 Desenho dos primers para amplificação dos genes

Foram desenhados e sintetizados pares de primers para amplificar segmentos específicos de exons dos genes CAD, CCOAMT, COMT, CCR, PAL, 5FH e 4CL. As seqüências dos primers desenhados, as temperaturas de anelamento para cada par de primer, o tamanho total dos produtos amplificados bem como o tamanho dos produtos amplificados em seqüências exônicas são descritos (Tabela 1). Os primers foram desenhados objetivando amplificar a maior extensão de seqüência transcrita possível. A figura 3A mostra a posição de amplificação dos primers CADF e CADR no gene CAD e a figura 3B mostra a foto de uma eletroforese em gel de agarose (1%) corado com brometo de etídeo para amplificações via PCR de 4 amostras de DNA genômico e 4 amostras de RNA para eucalipto.

Tabela 1. Seqüência dos primers desenhados para os genes CAD, CCOAMT, COMT, CCR, PAL, 5FH e 4CL indicando temperatura de anelamento dos primers (Ta), extensão em pares de bases da amplificação em seqüências exônicas (Exon pb) bem como o tamanho total do fragmento amplificado para cada par de primer (F+R) em pares de bases (Total pb)

Primer	Seqüência (5' – 3')	Ta °C	Exons (pb)	Total (pb)
CADF	ATGAAGTGGTGGGTGAGGTT	61	225	1190
CADR	AGCTTCAACAAGGCCAGGTA		368	
CCOAMTF	GGACATTAACAGAGAGAACTATGAAC	57	120	920
CCOAMTR	ACGGGGAGCATAACAGATCTCA		253	
COMTF	GAACCTCTTCGCCATGCAG	61	361	914
COMTR	TCAATCACGTGAGGCAAGTC		304	
CCRF	GGTGATCGGGACGAAAAAT	57	165	663
CCRR	GGGAAGAACTTGCCAAGGAT		401	
CCR2F	CTTCATCGCCTCCTGGATT	61	73	319
CCR2R	AGAGGCGGTGTGGACGAC		136	
PALF	CTAGGGCCCCAGATTGAGGT	57	994	994
PALR	GCAGTGAAGACCTTGTCGAA		994	
5FHF	TAAGTATGGCGGGATCTTCC	61	189	189
5FHR	CAGAACGGGCCGTAGTGC		189	
4CLF	GGAATGACCGAGGCTGGCCC	57	215	215
4CLR	AATGCGACAGGAACTTCACC		215	
4CLF2	ATGCCAAGCTCGGACAGG	61	619	619
4CLR2	CCCGGTTGATCCTCTTGTA		619	

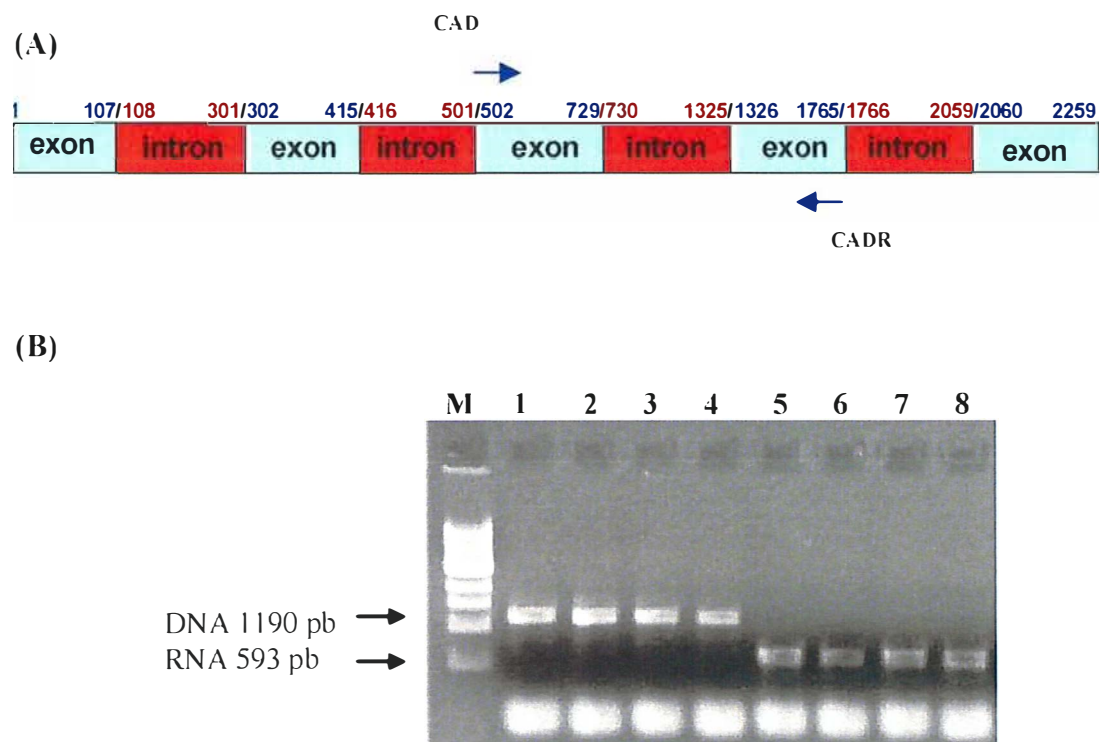


Figura 3 - (A) Desenho esquemático mostrando a organização da seqüência de DNA genômico para o gene CAD em *Eucalyptus saligna* (Dados obtidos do Gene Bank NCBI acesso N° AF294793) e posição de anelamento dos primers CADF e CADR. (B) Imagem de um gel de agarose 1% corado com brometo de etídeo mostrando a eletroforese de produtos de PCR realizados com DNA genômico de eucalipto e RNA para o gene CAD. M; marcador 1kb ladder. 1-4; amplificações de DNA genômico para os genitores G38, U15 e dois indivíduos da progênie, respectivamente. 5-8; amplificações de RNA para os genitores G38, U15 e dois indivíduos da progênie, respectivamente. Conforme esperado, o produto amplificado via DNA é de 1190 pb, enquanto para RNA é de 595 pb

6.3.2 Análise das seqüências e localização de SNPs em regiões transcritas

Uma vez realizado o sequenciamento dos produtos de PCR para os genes selecionados, as seqüências dos genitores e dos dois indivíduos da progênie foram alinhadas utilizando o programa Phred Phrap Consed, que permite o alinhamento de eletroferogramas indicando as possíveis posições de um SNP (Figura 4). Foram localizados SNPs em regiões transcritas para os genes CAD, PAL e CCR. Para o gene PAL foi selecionado um SNP na posição consenso 112 após o alinhamento das seqüências geradas utilizando como iniciador o primer PALF. O genitor G38 apresentou o genótipo heterozigoto AG e o genitor U15 apresentou o genótipo homozigoto AA. Para o gene CAD, foi selecionado um SNP na posição consenso 135 após o alinhamento das seqüências geradas utilizando como iniciador o primer CADR. O genitor G38 apresentou o genótipo heterozigoto AC e o genitor U15 apresentou o genótipo homozigoto AA. Para o gene CCR foi selecionado um SNP na posição consenso 257 após o alinhamento das seqüências geradas utilizando o primer CCRR. O genitor G38 apresentou o genótipo heterozigoto CT e o genitor U15 apresentou o genótipo homozigoto CC. Não foi possível detectar SNPs em regiões transcritas para os genes 5FH, 4CL, COMT e CCOAMT utilizando os primers selecionados impossibilitando desta maneira a continuidade dos estudos para estes genes especificamente no cruzamento selecionado para o estudo.



Figura 4 - Tela de saída de resultado do programa Phred Phrad Consed para a análise das seqüências geradas para o gene CCR utilizando como iniciador o primer CCRR. A parte superior da figura mostra o alinhamento das seqüências para os genitores e dois indivíduos da progênie bem como a indicação de um possível SNP. A parte inferior da imagem mostra o alinhamento dos eletroferogramas

6.3.3 Desenho dos primers para a detecção de SNPs via SBE

Foram desenhados primers para a detecção dos SNPs descobertos nos genes, PAL, CCR e CAD. Uma vez que por restrição da técnica de SBE, os primers obrigatoriamente devem começar exatamente antes ou exatamente depois do SNP, não foi necessária a utilização de um programa específico para desenhar os primers. Para a análise da expressão gênica no gene PAL foi desenhado o primer PAL2B cuja seqüência é 5'TGAGCTTCCCTATGGACGCAA3' para amplificação do SNP selecionado.

Para a análise da expressão gênica no gene CCR foi desenhado o primer CCR4A cuja seqüência é 5'AGGTCAACCCCTCTCTCCTT3'. Para a análise da expressão gênica no gene CAD foi desenhado o primer CAD3A cuja seqüência é 5'ATGTGACTGTGATAAGCTCTTC3'.

6.3.4 Curva de normalização

Curvas de normalização de dados foram criadas para corrigir os valores de intensidade dos picos de fluorescência de expressão gênica, gerados com base em mRNA para os genes estudados com base na comparação da intensidade de sinal gerado a partir de misturas conhecidas de DNA genômico dos genitores (Figura 5). Para os SNPs encontrados no gene PAL (AG para o genitor G38 e AA para o genitor U15) e para o gene CAD (AC para o genitor G38 e AA para o genitor U15), as curvas de normalização foram geradas utilizando as relações de misturas de 75:25, 80:20 e 90:10 para os alelos AG, e AC, respectivamente. Para o SNP encontrado no gene CCR, os genitores apresentaram os genótipos TC (G38) e CC (U15). A curva de normalização foi realizada utilizando as relações de misturas de 75:25, 87,5:12,5 e 62,5:37,5 para os alelos TC. Para os SNPs encontrados nos genes PAL e CAD, a obtenção de uma relação de 75:25 entre os alelos, foi utilizado 3,0 ng do DNA genômico do genitor G38 com 3,0 ng do DNA genômico do genitor U15. Para a obtenção de uma relação 80:20, foram misturados 4,5ng do DNA genômico do genitor U15 com 3,0 ng do DNA genômico do genitor G38, e

para a obtenção de uma relação 90:10, foram misturados 12,0 ng do DNA genômico do genitor U15 com 3,0 ng do DNA genômico do genitor G38. Para o SNP encontrado no gene CCR, a relação 62,5:37,5 entre os alelos foi obtida utilizando 3,0 ng do DNA genômico do genitor U15 com 9,0 ng do DNA genômico do genitor G38. Para obtenção de uma relação 75:25, foram misturados 3,0 ng do DNA genômico do genitor U15 com 3,0 ng do DNA genômico do genitor G38, e para obtenção de uma relação 87,5:12,5, foram misturados 9,0 ng do DNA genômico do genitor U15 com 3,0 ng do DNA genômico do genitor G38.

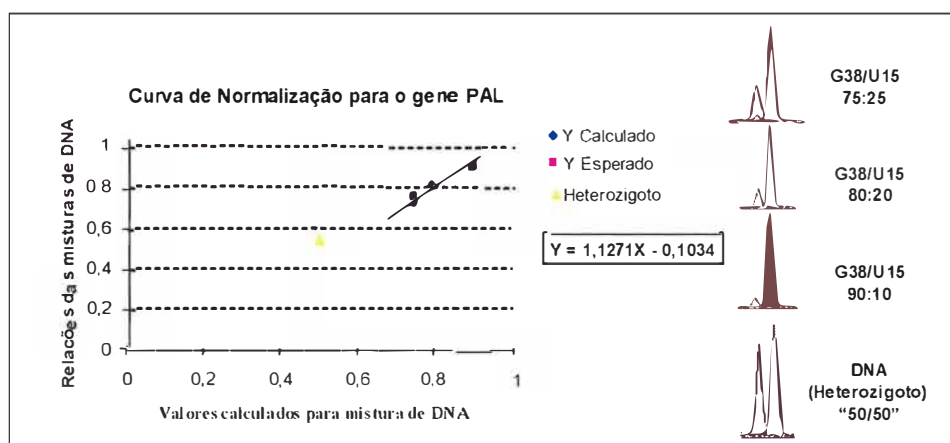


Figura 5 - Genotipagem via SBE do SNP localizado na posição consensus 112 do exon1 para o gene PAL utilizando misturas de DNA genômico dos genitores G38 e U15. No gráfico, é possível observar uma boa correlação entre as razões das misturas esperadas e observadas. Foi calculada uma equação para a linha de tendência a qual foi em seguida utilizada para corrigir os valores de expressão alélica dos indivíduos

A conversão dos valores de altura de pico de fluorescência em estimativas de níveis de expressão gênica para os indivíduos analisados foi realizada utilizando uma equação gerada a partir dos dados referentes às misturas de DNA genômico de proporções conhecidas para cada gene. Em seguida foi realizada a normalização dos dados utilizando informação de um DNA genômico para um indivíduo heterozigoto, representando o que seria uma expressão de 50:50 para os alelos de um gene (Figura 6).

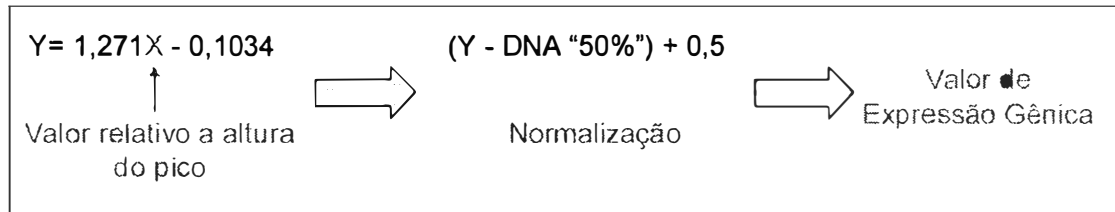


Figura 6 - Etapas realizadas para a conversão de dados referentes à altura do pico em eletroferograma em estimativa da intensidade da expressão gênica para um alelo específico. O valor referente à altura do pico para um alelo é transformado em valor de expressão gênica utilizando a equação gerada com dados da genotipagem da SBE para misturas de proporções conhecidas de DNA genômico dos genitores. Em seguida, a normalização dos dados foi realizada utilizando os dados da SBE de um DNA genômico de um indivíduo heterozigoto da progênie previamente selecionado

6.3.5 Análise de expressão gênica via “Single Base Extension” para o gene Cinnamyl Alcohol Dehydrogenase (CAD).

Na análise de expressão alélica diferencial para o gene CAD foram utilizados 18 indivíduos da progênie que apresentaram o genótipo heterozigoto AC ao SNP 135. A reação de SBE foi realizada a partir de mRNA extraído do xilema. Os valores referentes às expressões dos alelos A e C estão apresentados na figura 7.

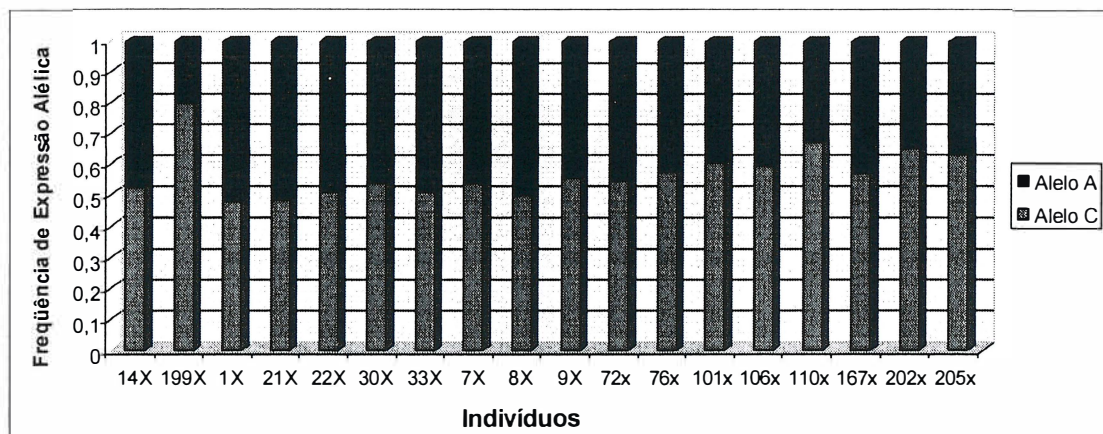


Figura 7 - Distribuição dos níveis de expressão gênica alelo específica para o gene CAD em 18 indivíduos da progênie do cruzamento entre os indivíduos G38 e U15. As barras vermelhas representam a expressão do alelo C e as barras azuis representam a expressão do alelo A

6.3.6 Análise de expressão gênica via “Single Base Extension” para o gene Cinnamoyl CoA Redutase (CCR)

Na análise de expressão alélica diferencial para o gene CCR foram utilizados 22 indivíduos da progênie que apresentaram o genótipo heterozigoto TC. A reação de SBE foi realizada a partir de mRNA extraído do xilema. Os valores referentes às expressões dos alelos T e C estão apresentados na figura 8.

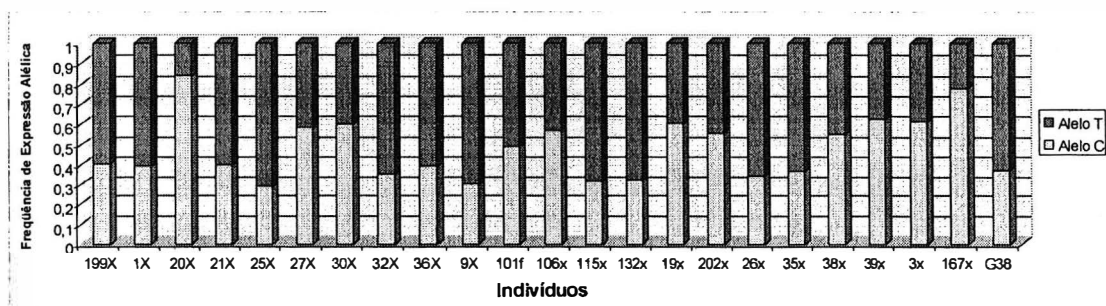


Figura 8 - Distribuição dos níveis de expressão gênica alelo específica para o gene CCR de 22 indivíduos da progênie do cruzamento entre os indivíduos G38 e U15. As barras azuis representam a expressão do alelo C e as barras vermelhas representam a expressão do alelo T

Os resultados obtidos para os genes CAD e CCR mostram a presença de expressão diferencial em seus alelos. Para o gene CAD, os indivíduos 199, 110, 202 e 205 apresentaram expressão do alelo C superior a 60% representando uma expressão 1,5 vez maior que a expressão do alelo A.

Já para o gene CCR, os indivíduos 199, 1, 21, 25, 32, 36, 9, 115, 132, 26 e 35 apresentaram valores de expressão para o alelo T superior a 60%, representando uma expressão 1,5 vez maior que a expressão do alelo C.

6.4. Discussão

Este trabalho teve como objetivo estudar variações de expressão gênica alelo específica para todos os genes envolvidos na via de lignificação do eucalipto. Uma vez que por exigência da metodologia de SBE, só é possível realizar este estudo para aqueles genes que apresentam SNPs em seqüências transcritas nos genitores do cruzamento, somente os genes PAL, CAD e CCR puderam ser estudados especificamente para este cruzamento. Outros cruzamentos, possivelmente envolvendo espécies de *Eucalyptus* mais divergentes poderão fornecer configurações genotípicas aos SNPs identificados, bem como outras posições polimórficas que permitam realizar esta análise para outros genes da via de lignificação.

Para o gene PAL, não foi possível obter amplificações via PCR para os mRNAs extraídos de tecido de xilema. Uma hipótese para explicar a não amplificação deste gene é a existência de diversas isoformas do gene responsáveis pela formação de precursores de fenilpropanóides em diversos tecidos celulares (Fiona et al., 2004). Desta forma, o par de primers desenhado para amplificação deste gene pode ser específico para uma outra isoforma atuando em algum outro tecido que não seja xilema. Estudos adicionais devem ser realizados no intuito de resgatar seqüências específicas de isoformas que

atuam na via de lignificação do eucalipto especificamente em xilema para que este gene possa ser enquadrado em estudos futuros envolvendo sua participação na via de lignificação. Recentemente sequências de EST geradas a partir de diferentes bibliotecas de xilema bem como sequências geradas a partir de mRNA de folhas foram geradas no âmbito do projeto Genolyptus (Pasquali, comunicação pessoal). Uma análise comparativa dos transcritos gerados nos dois tipos de tecidos poderá fornecer pistas para a identificação das isoformas específicas de xilema.

Outra limitação encontrada no estudo e passível de ocorrência com outros genes foi a discriminação de polimorfismos entre os membros da família gênica do gene COMT. Apesar dos primers terem sido desenhados para permitir a amplificação específica do membro da família COMT1, foram observados produtos amplificados de mais de um gene para esta família. Resultados de sequenciamento dos produtos de PCR utilizando o par de primer desenhado mostraram que as seqüências desses genes, para esta região, apresentam uma alta similaridade, impossibilitando a discriminação de um polimorfismo entre alelos ao mesmo loco com um polimorfismo entre alelos a locos distintos (dados não mostrados).

Para o gene CAD, foram analisados 18 indivíduos, dos quais 4 apresentaram expressão do alelo C superior a 60% representando uma expressão 1,5 vez maior que a expressão do alelo A (199, 110, 202, 205). Para o gene CCR, foram analisados 22 indivíduos, dos quais 11 (199, 1, 21, 25, 32,

36, 9, 115, 132, 26 e 35) apresentaram valores de expressão para o alelo T superior a 60%, representando uma expressão 1,5 vez maior que a expressão do alelo C. Para este gene, a segregação dos níveis de expressão dos alelos segregou na razão de 1:1, sugerindo um possível efeito de ação combinada de fator em trans e fator em cis ("*Cis + Trans acting*") afetando a expressão dos alelos para este gene (Figura 9). Como o genitor G38 apresentou uma baixa expressão para o alelo C (menor que 40%), e foram observados na progênie indivíduos com uma elevada expressão para este alelo, o genitor U15 pode ser heterozigoto para um componente em *Trans*, ou seja um fator de transcrição, que se ligaria à região *Cis* que controla a expressão do alelo C. Desta forma, indivíduos da progênie que herdaram o alelo C do genitor G38 e o alelo responsável pelo efeito *Trans* do genitor U15, apresentariam uma elevada expressão do alelo C, consequência esta do efeito "*Cis + Trans acting*". Para investigar esta hipótese em detalhe, novas análises serão realizadas com maior número de descendentes para uma melhor compreensão do modelo genético que controla a expressão dos alelos deste gene. Uma vez realizada a genotipagem para mais indivíduos, os dados referentes à expressão gênica poderão ser utilizados para realizar o mapeamento do efeito como uma característica quantitativa, ou seja mapeamento de um eQTL (*expression-QTL*) para possível fator de transcrição envolvido para este gene.

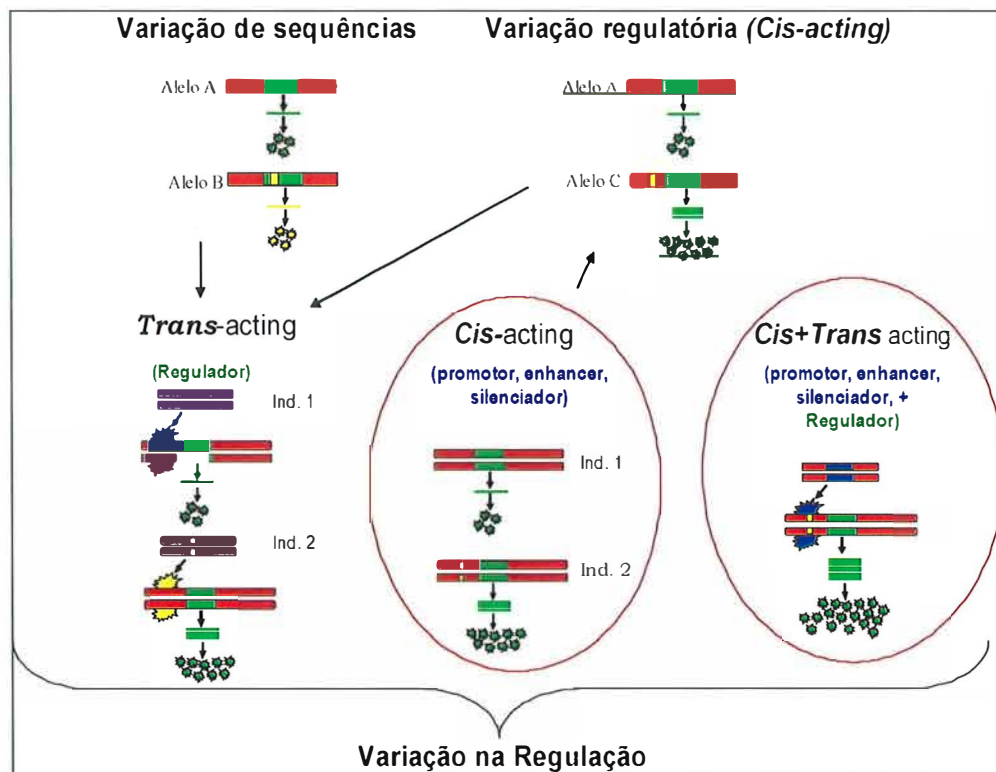


Figura 9 - Possíveis modelos de polimorfismos funcionais que podem atuar na regulação da expressão de um gene. O modelo *trans-acting* é explicado pela ação de uma proteína que se liga ao promotor de um gene alterando sua expressão final. Variações *cis-acting* envolvem mutações no promotor de um gene, alterando sua expressão final. O modelo *cis + trans acting* pode ser explicado pela união dos fatores *cis* e *trans acting* alterando a expressão final de um gene

Para verificar a consistência das análises de variação da expressão gênica alelo específico utilizando SBE, novas análises foram realizadas para os genes CAD e CCR partindo de novas extrações de mRNA para alguns indivíduos. O resultado das diferentes reações de SBE não apresentou diferenças significativas, confirmando a robustez desta metodologia (Figura 10).

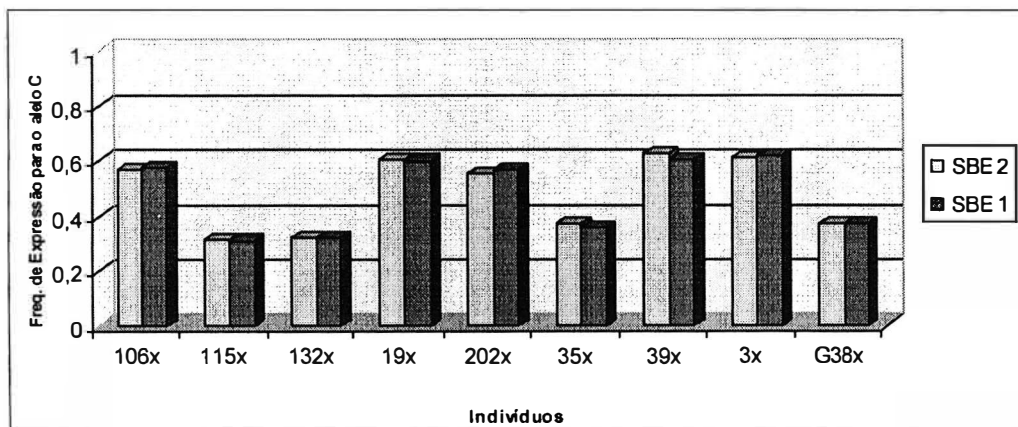


Figura 10 - Distribuição dos níveis de expressão gênica alelo específica para o gene CCR em dois experimentos independentes (SBE 1 e SBE 2). Foram realizadas 2 reações de SBE partindo de diferentes extrações de mRNA de tecidos de xilema para 9 indivíduos. A frequência de expressão alélica apresentada é relativa ao alelo C

6.5 Conclusões

Este estudo se baseou na otimização e implementação de uma metodologia recente para estudos de expressão gênica que permite a análise diferencial de expressão de alelos ao mesmo gene. Foi demonstrado que a metodologia permite de fato a análise de expressão gênica de alelos específicos em genes específicos no eucalipto. Para isso são necessárias famílias segregantes e informações prévias de seqüência de genes candidatos para permitir a amplificação de segmentos exônicos, o sequenciamento, descobrimento de SNPs desenho de primer para a detecção do SNP que diferencia os alelos e uma metodologia de normalização da intensidade de sinal.

Os poucos trabalhos que se tem publicado até o momento utilizando esta metodologia foram realizados em humanos. Desta forma, este trabalho é pioneiro se tratando de espécies florestais e um dos primeiros envolvendo plantas em geral. (Um outro trabalho ainda não publicado vem sendo realizado na universidade de Udine envolvendo estudos de variação genética atuando em *cis* em linhagens de milho).

Estudos de expressão alelo específica em espécies alógamas e geneticamente heterogêneas como o eucalipto, apresentam algumas limitações

quando comparados a estudos conduzidos em plantas autógamas ou plantas com disponibilidade de linhagens. Os cruzamentos disponíveis são entre indivíduos heterozigotos o que gera na descendência diversas possíveis configurações envolvendo genótipos homozigotos e heterozigotos. Nos casos nos quais a configuração envolve um SNP em heterozigose para um genitor, e homozigose ou heterozigose no outro genitor para um alelo em comum com o primeiro genitor, a análise não é totalmente informativa, uma vez que somente os indivíduos que são heterozigotos para o SNP poderão ser analisados. Para os demais descendentes homozigotos, a análise não é possível para o SNP alvo. Por outro lado, o eucalipto apresenta uma vantagem em comparação a espécies endogâmicas ao se utilizar esta metodologia para estudar variações de expressão gênica. O possível efeito combinado *trans* e *cis* observado neste estudo para o gene CAD, por exemplo, não poderia ser caracterizado em milho apesar de possivelmente também existir naquela espécie.

Esta metodologia apresenta um enorme potencial de aplicação para aprofundar conhecimentos rumo a um melhor entendimento do comportamento dos genes na via de lignificação. A utilização de cruzamentos interespecíficos envolvendo genitores contrastantes e possivelmente altamente polimórficos para SNPs nos genes responsáveis pela produção de lignina certamente deverá gerar oportunidades interessantes de análise de múltiplos genes com esta metodologia. Neste sentido, análises de sequenciamento destes genes da via de lignificação foram realizadas para outros genitores de diversas espécies

de eucalipto envolvidos na geração de famílias segregantes utilizadas para mapeamento de QTL no âmbito do projeto Genolyptus. Foi observado que, em diversos cruzamentos, principalmente envolvendo os eucaliptos temperados *Eucalyptus dunni* e *E. globulus* como um dos genitores, ocorre um elevado nível de polimorfismo de seqüência para todos os genes da via de lignificação aqui estudados. O fato de se trabalhar com um cruzamento envolvendo como um dos genitores um indivíduo da espécie *Eucalyptus dunni* ou *E. globulus*, tornar-se-á extremamente interessante não somente pela ampla possibilidade de detecção de múltiplos SNPs em regiões exônicas, mas também pelo fato do *Eucalyptus dunni* e principalmente o *E. globulus* apresentar madeira com características físico químicas altamente diferenciadas em relação às espécies tropicais *E. grandis* e *E. urophylla* que resultam em um rendimento significativamente superior na produção de celulose. Neste sentido, dados fenotípicos para as diversas características da madeira em conjunto com os dados genotípicos gerados utilizando a metodologia de detecção de expressão alelo específica, poderão abrir novas e importantes perspectivas para a compreensão não somente de quais os genes candidatos que são responsáveis pela formação da madeira mas também de que forma que a variação de expressão dos alelos desses genes ocorre quando indivíduos contrastantes para um fenótipo são combinados. Finalmente, a identificação de padrões de diferenciação de expressão gênica alelo específica em genes servirá como mais uma ferramenta para a seleção de genes candidatos com elevado potencial de sucesso em estudos de mapeamento de associação nos quais polimorfismos

específicos em forte desequilíbrio de ligação por estarem dentro do gene são identificados e correlacionados com diferenças fenotípicas e posteriormente utilizados para a seleção assistida.

7 CONCLUSÕES GERAIS

A seguir são resumidas as principais conclusões e observações de cada um dos quatro trabalhos desenvolvidos nesta tese:

Capítulo 1: Metodologia de múltiplos iniciadores com cauda:

(1) Foi desenvolvido um método de múltiplos iniciadores com cauda (*primer tail*) baseado na utilização de três diferentes primers universais com seqüências do genoma humano não apresentando portanto homologia com o genoma de eucalipto. Este método foi utilizado com sucesso para a detecção simultânea e eficiente de microssatélites de *Eucalyptus* em um experimento de mapeamento genético aumentando a eficiência de utilização do tempo de um seqüenciador automático e de geração de dados genotípicos.

(2) A metodologia desenvolvida tem uma ampla aplicação para todas as espécies de plantas aumentando significativamente a velocidade e qualidade dos dados genéticos de microssatélites em comparação com metodologias de detecção baseadas em nitrato de prata e muito mais econômica em comparação à utilização convencional com todos os iniciadores marcados.

Capítulo 2: Mapeamento do loco controlador de florescimento precoce *Eef1*:

(1) Foi identificada uma planta mutante de *E. grandis* com um fenótipo de florescimento precoce aos 90 dias a partir da germinação. Em um cruzamento deste mutante com uma planta normal, marcadores microssatélites com detecção fluorescente associados com a estratégia de BSA, foram utilizados com sucesso para a detecção rápida e eficiente de um QTL de grande efeito (*major effect locus*) controlando florescimento precoce em eucalipto, loco este batizado *Eef1* (*Eucalyptus early flowering 1*).

(2) O fato do loco *Eef1* ter sido mapeado com marcadores microssatélites co-dominantes e transferíveis, permite o teste de alelismo entre o segmento contendo o loco *Eef1* e o segmento homólogo em outros mutantes para florescimento precoce que venham a ser identificados futuramente.

(3) O florescimento precoce, é uma característica extremamente útil como uma ferramenta experimental em trabalhos de transgenia ou desenvolvimento de linhagens puras de eucalipto. Do ponto de vista operacional em programas de melhoramento, não havendo perda de poder de crescimento vegetativo da árvore, o florescimento precoce pode permitir acelerar ciclos de recombinação e seleção reduzindo o tempo necessário para a geração de novas combinações genótípicas e, conseqüentemente, maximizando os ganhos genéticos por unidade de tempo.

Capítulo 3: Mapeamento genético de QTL para propriedades de madeira

(1) Mapas genéticos de microssatélites com cobertura da ordem de 70 a 80% foram construídos para duas árvores híbridas e utilizados para a localização de um total de 15 QTL para oito características.

(2) O comprimento total esperado reduzido dos mapas em cM, juntamente com a observação de proporções da ordem de 40 a 50% de marcadores com distorção de segregação apontam para a ocorrência de restrição de recombinação.

(3) Embora parte da distorção de segregação possa ser explicada por seleção gamética e/ou seleção pós-zigótica, o fato dos genitores utilizados serem híbridos interespecíficos sugere que sejam as distorções bem como a redução do tamanho dos mapas em unidades de recombinação seja uma consequência de problemas de pareamento dos cromossomos homólogos de espécies distintas durante a meiose nos híbridos.

(4) O tipo de marcador molecular utilizado na construção de mapas genéticos pode também influenciar na observação das proporções reais de marcadores distorcidos. Foi sugerido que as proporções de marcadores com distorção de segregação estimadas neste trabalho devem se aproximar com mais precisão

das proporções reais a serem observadas em cruzamentos de Eucalipto envolvendo híbridos como genitores de cruzamentos.

(5) A localização de QTL para teor de pentosana e densidade na mesma região do grupo de ligação 6 nos mapas de ambos os genitores representa uma evidência de validação da declaração de QTL nesta região.

(6) A co-localização de QTL para consumo específico, *coarseness*, lignina e rendimento de celulose no grupo de ligação 9 do genitor materno 235 indica a possibilidade de haver nesta região um gene de maior efeito envolvido no controle da expressão dessas características. Para teor de lignina e rendimento de celulose este QTL pode apresentar efeito pleiotrópico, tendo em vista que essas características são negativamente correlacionadas.

(7) Foram utilizados diferentes métodos estatísticos para a detecção e localização de QTL. Na comparação de resultados, as análises de MIC apresentaram maior poder de detecção do que as metodologias de AMI e MIS indicando ser recomendada a utilização de diferentes métodos de detecção de QTL para evitar erros do tipo I e tipo II.

(8) Diferenças referentes à detecção de QTL encontradas pelas mesmas análises de MIS utilizando os programas QTL Cartographer e Mapmaker QTL

sugerem a necessidade de uma análise comparativa mais aprofundada dos detalhes computacionais implementados pelos dois programas.

(9) Marcadores microssatélites apresentam uma grande vantagem perante as outras classes de marcadores utilizadas para construção de mapas genéticos em eucalipto por possibilitar estudos de sintenia e transferibilidade de marcas entre mapas gerados a partir de diferentes cruzamentos.

(10) Mapeamento comparativo de QTL detectados neste trabalho com outros QTL e genes candidatos da via de lignificação mapeados em estudos independentes, revelaram diversos casos de sintenia de QTL para rendimento de celulose, teor de lignina e propriedades de fibra. Estes resultados são inéditos para *Eucalyptus* e apontam para perspectivas de validação de QTL e detecção direcionada de variação alélica em outros pedigrees visando a determinação precisa de segmentos genômicos alvo para experimentos de mapeamento de alta resolução, testes de associação e, futuramente, seleção assistida por marcadores.

Capítulo 4: Expressão alelo específica de genes de lignificação em *Eucalyptus*.

(1) A metodologia genômica que combina mensuração de níveis de expressão diferencial alelo específica em genes candidatos usando RT-PCR com a discriminação dos transcritos correspondentes aos diferentes alelos com base

na variabilidade de bases individuais, SNPs (*Single Nucleotide Polymorphisms*) permitiu detectar diferenças significativas níveis de expressão gênica entre alelos de um mesmo gene. Esta observação foi realizada com sucesso para alguns genes que codificam para enzimas chave na via de biossíntese de lignina em Eucalipto.

(2) Apesar de estudos de expressão gênica alelo específica em espécies alógamas e geneticamente heterogêneas como o eucalipto apresentarem algumas limitações quando comparados a estudos conduzidos em plantas autógamas com disponibilidade de linhagens, em plantas geneticamente heterogêneas é possível a detecção de variações atuando em *cis* agindo em conjunto com variações atuando em *trans*, que apesar de existirem, não são facilmente detectadas em plantas autógamas.

(3) Para o gene CCR, os níveis de expressão dos alelos segregou na razão de 1:1, sugerindo um possível efeito de ação combinada de fator em *trans* e fator em *cis* ("*Cis + Trans acting*") afetando a expressão dos alelos para este gene. Futuramente, estudos mais aprofundados serão realizados visando mapear este fator atuando em *trans*.

(4) A metodologia utilizada apresenta um amplo potencial de aplicação para aprofundar os conhecimentos rumo a um melhor entendimento do comportamento dos genes envolvidos na via de lignificação em híbridos

superiores de eucalipto, principalmente aqueles gerados entre espécies com características contrastantes para qualidade da madeira. Genes cuja expressão apresente diferenças importantes entre alelos de espécies contrastantes, são candidatos em potencial para estudos de genética de associação visando desenvolver metodologias de seleção assistida baseadas em marcadores diretamente dentro de genes ou regiões reguladoras.

Em conclusão, os resultados obtidos nesta tese contribuem para o avanço contínuo do conhecimento genético e genômico de características de importância econômica do eucalipto, além de fornecer novas ferramentas e abordagens experimentais rumo à aplicação destes conhecimentos na prática diária dos programas de melhoramento genético de espécies do gênero.

ANEXOS

Anexo A. Seqüência dos primers microssatélites selecionados para genotipagem automatizada utilizando como cauda a seqüência do primer D8S1132, bem como a temperatura de anelamento para realização da PCR

Nome do Loco	Seqüência dos primers direto, reverso e direto + cauda	Temp. PCR °C
EMBRA 2	5'CGTGACACCAGGACATTAC3' 5'ACAAATGCAAATTC AAAATGA3' 5'GGCTAGGAAAGGTTAGTGGCC GTGACACCAGGACATTAC3'	56
EMBRA 36	5'TTATCGTCAATTCTTGCTTG3' 5'AATTCAGCTCAAGATTGGT3' 5' GGCTAGGAAAGGTTAGTGGCT TTATCGTCAATTCTTGCTTG3'	56
EMBRA 102	5'GGCTAGTGATCCAAACATG3' 5'GTCTCTCTCGCGTATGACA3' 5' GGCTAGGAAAGGTTAGTGGC GGCTAGTGATCCAAACATG3'	54
EMBRA 37	5'CACCTCTCCAACTACACAA3' 5'CTCCTCTCTTCCACCATT3' 5' GGCTAGGAAAGGTTAGTGGC CACCTCTCCAACTACACAA3'	56
EMBRA 38	5'GGTTCTCTAGTGAAAATGTCG3' 5'ATACATCCATCAAAGCACAA3' 5' GGCTAGGAAAGGTTAGTGGC GGTTCTCTAGTGAAAATGTCG3'	56
EMBRA 41	5'ATGATTTTGTGCGTGGAC3' 5'TCAGGTGAAAGGATGGAG3' 5' GGCTAGGAAAGGTTAGTGGC ATGATTTTGTGCGTGGAC3'	56
EMBRA 105	5'GGAACGCATGTTCTTTTC3' 5'TGTCTGCTTTGGAATGTG3' 5' GGCTAGGAAAGGTTAGTGGC GGAACGCATGTTCTTTTC3'	58
EMBRA 42	5'GAGTAAAAATTGGTTTTGAGTG3' 5'CCCTCTTTTCATTTTGTCTT3' 5' GGCTAGGAAAGGTTAGTGGC GAGTAAAAATTGGTTTTGAGTG3'	56
EMBRA 109	5'AAAGAAACCGTAAGTTCGC3' 5'CTCACCAATGGACAATAGC3' 5' GGCTAGGAAAGGTTAGTGGC AAAGAAACCGTAAGTTCGC3'	58
EMBRA 45	5'GTCATTTGCACACAGTTTTTC3' 5'AGTTCATAGAATGCAGAAAATG3' 5' GGCTAGGAAAGGTTAGTGGC GTCAATTTGCACACAGTTTTTC3'	56
EMBRA 12	5'AGGATTTGTGGGGCAAGT3' 5'GTTCCCCATTTTCATGTCC3' 5' GGCTAGGAAAGGTTAGTGGC AGGATTTGTGGGGCAAGT3'	56
EMBRA 18	5'CAGCTAGGATGTTAGACTTGG3' 5'GCACACGAATTTTCAAATA3' 5' GGCTAGGAAAGGTTAGTGGC CAGCTAGGATGTTAGACTTGG3'	56
EMBRA 122	5'TTGCTCCATCTTTCTTG3' 5'AAAACGATTAGAGGGTCATG3' 5' GGCTAGGAAAGGTTAGTGGC TTGCTCCATCTTTCTTG3'	54
EMBRA 19	5'GACGGTTGATTTCTGATT3' 5'GTGGTGCTCCTCTCCTCT3' 5' GGCTAGGAAAGGTTAGTGGC GACGGTTGATTTCTGATT3'	56
EMBRA 124	5'TGTGGACATGATAGATGAAGT3' 5'AATTATTTTCGATAGAAACGGA3' 5' GGCTAGGAAAGGTTAGTGGC TGTGGACATGATAGATGAAGT3'	58

Anexo B. Seqüência dos primers microssatélites selecionados para genotipagem automatizada utilizando como cauda a seqüência do primer D12S1090, bem como a temperatura de anelamento para realização da PCR

Nome do Loco	Seqüência dos primers direto, reverso e direto + cauda	Temp. PCR (°C)
EMBRA 6	5'AGAGAATTGCTCTTCATGGA3' 5'GAAAAGTCTGCAAAGTCTGC3' 5' ACCAACCTAGGAAACACAGT AGAGAATTGCTCTTCATGGA3'	56
EMBRA 126	5'GTAGTCCCTGACGCAAAC3' 5'AGTACGATATTTTCGCGAGT3' 5' ACCAACCTAGGAAACACAGT GTAGTCCCTGACGCAAAC3'	58
EMBRA 140	5'GGTTTGATAGGTGGATTGGG3' 5'GGTTGATGACTGAGAGATGAAGG3' 5' ACCAACCTAGGAAACACAGT GGTTTGATAGGTGGATTGGG3'	64
EMBRA 145	5'AAGCCACACCGTGTCAGTTAGC3' 5'AACTCACAGGCGCAAGGCA3' 5' ACCAACCTAGGAAACACAGT AAGCCACACCGTGTCAGTTAGC3'	64
EMBRA 65	5'GACATCTCCTCCTCAAGC3' 5'CGATATGCTACGTCTTCC3' 5' ACCAACCTAGGAAACACAGT GACATCTCCTCCTCAAGC3'	58
EMBRA 66	5'ACTTCTTAGGCTACAGCA3' 5'GAAGGATCACGAGACATA3' 5' ACCAACCTAGGAAACACAGT ACTTCTTAGGCTACAGCA3'	56
EMBRA 67	5'GGAAGAATCTAAGCGTCA3' 5'GAAGAAGATGAATGTAGGTG3' 5' ACCAACCTAGGAAACACAGT GGAAGAATCTAAGCGTCA3'	58
EMBRA 164	5'CCTTGTTGAGCTCCTGTCT3' 5'ACTATCAGCGTCCTGCAA3' 5' ACCAACCTAGGAAACACAGT CCTTGTTGAGCTCCTGTCT3'	60
EMBRA 165	5'AAGAACATGCAGCGGAGA3' 5'TTGTGATGGACCACTCAATG3' 5' ACCAACCTAGGAAACACAGT AAGAACATGCAGCGGAGA3'	56
EMBRA 177	5'CCTCACTTGCCGAATAC3' 5'CTTGCATGTCCTCCAACAC3' 5' ACCAACCTAGGAAACACAGT CCTCACTTGCCGAATAC3'	60
EMBRA 60	5'AACAGCAGTTGCTACACCAC3' 5'GAGCGAAAAGGAGAACC3' 5' ACCAACCTAGGAAACACAGT AACAGCAGTTGCTACACCAC3'	60
EMBRA 224	5'TAAGATAGAGACGGGCCAGC3' 5'TGCAATACAAGTCTGAGAGCCA3' 5' ACCAACCTAGGAAACACAGT TAAGATAGAGACGGGCCAGC3'	58
EMBRA 227	5'CGAATGCCATAGATTGTCAG3' 5'CAGGCATCTCGTACGTGGA3' 5' ACCAACCTAGGAAACACAGT CGAATGCCATAGATTGTCAG3'	56
EMBRA 228	5'TTGCCCATGAAGTTAACAC3' 5'GCATTCCGATTCTTGTGG3' 5' ACCAACCTAGGAAACACAGT TTGCCCATGAAGTTAACAC3'	56
EMBRA 230	5'GTACTTAATCTGTGGCAATC3' 5'GATCCTTGCATTATCGAC3' 5' ACCAACCTAGGAAACACAGT GTACTTAATCTGTGGCAATC3'	56

Anexo C. Sequência dos primers microssatélites selecionados para genotipagem automatizada utilizando como cauda a seqüência do primer DYS437, bem como a temperatura de anelamento para a realização da PCR

Nome do Loco	Seqüência dos primers direto, reverso e direto + cauda	Temp. PCR (°C)
EMBRA 3	5'GGTTGTTTCATCTTTTCCATG3' 5'AGCGAAGGCAATGTGTTT3' 5' GACTATGGGCGTGAGTGCAT GGTTGTTTCATCTTTTCCATG3'	56
EMBRA 5	5'ATGCTGGTCCAACCTAAGATT3' 5'TGAGCCTAAAAGCCCAAC3' 5' GACTATGGGCGTGAGTGCAT ATGCTGGTCCAACCTAAGATT3'	56
EMBRA 27	5'ATAACCCACACCAATCTGCA3' 5'TATAGCTCGAACGCTCAAC3' 5' GACTATGGGCGTGAGTGCAT ATAACCCACACCAATCTGCA3'	56
EMBRA 10	5'GTAAAGACATAGTGAAGACATTCC3' 5'AGACAGTACGTTCTCTAGCTCA3' 5' GACTATGGGCGTGAGTGCAT GTAAAGACATAGTGAAGACATTCC3'	56
EMBRA 45	5'GTCATTTGCACACAGTTTTTC3' 5'AGTTCATAGAATGCAGAAAATG3' 5' GACTATGGGCGTGAGTGCAT GTCATTTGCACACAGTTTTTC3'	56
EMBRA 118	5'GTGCCCTCAGCATCTCAG3' 5'CAGGAGGAAGGGAGTGGGA3' 5' GACTATGGGCGTGAGTGCAT GTGCCCTCAGCATCTCAG3'	58
EMBRA 133	5'CGGTCGTTGTCGGAATCTC3' 5'AGTTGGGTAACGCCAGGTTT3' 5' GACTATGGGCGTGAGTGCAT CGGTCGTTGTCGGAATCTC3'	54
EMBRA 77	5'GAACACTATTTGGGAAGCA3' 5'CTCCTCTCTCCTCTCAACTC3' 5' GACTATGGGCGTGAGTGCAT GAACACTATTTGGGAAGCA3'	60
EMBRA 20	5'GTGAGTGGGTATCCATCG3' 5'GCTGGAAGTGGTCTTGAG3' 5' GACTATGGGCGTGAGTGCAT GTGAGTGGGTATCCATCG3'	56
EMBRA 78	5'TATTATCCAAAACGTCATTAGC3' 5'TAGAGGCAGGTGTTAGCTG3' 5' GACTATGGGCGTGAGTGCAT TATTATCCAAAACGTCATTAGC3'	58
EMBRA 82	5'GGAGACAGACATTTTGTTCG3' 5'TAAGATTCTTGAATGCACA3' 5' GACTATGGGCGTGAGTGCAT GGAGACAGACATTTTGTTCG3'	54
EMBRA 27	5'ATAACCCACACCAATCTGCA3' 5'TATAGCTCGAACGCTCAAC3' 5' GACTATGGGCGTGAGTGCAT ATAACCCACACCAATCTGCA3'	56
EMBRA 89	5'AACAGTTTCGATTTGAAGGT3' 5'GAGATTGTGGAGGAATCAAT3' 5' GACTATGGGCGTGAGTGCAT AACAGTTTCGATTTGAAGGT3'	54
EMBRA 97	5'GAGATTGTGGAGGAATCAAT3' 5'TGCAAGAAAAGCATAAAGAA3' 5' GACTATGGGCGTGAGTGCAT GAGATTGTGGAGGAATCAAT3'	58
EMBRA 33	5'CAATTTGCATGTCCAGTTTG3' 5'GCAGAAGTTGATTGAAAGCA3' 5' GACTATGGGCGTGAGTGCAT CAATTTGCATGTCCAGTTTG3'	58

Anexo D. Dados fenotípicos para as características avaliadas no capítulo 3 que foram utilizadas para mapeamento de QTL

Ind. Sel.	CAP	Pilodyn	Rendimento	CE	Lignina	Pentosanas	Fibra/Grama	Coarseness
119	28.1	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
123	34.6	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
124	26.8	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
125	47.3	14	49.6616	4.4689	30.4727	17.2586	25.7245	7.9816
126	41.6	15.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
129	40.8	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
132	40.1	17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
136	39.9	18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
138	41.9	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
142	39.6	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
144	37.3	14.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
146	31.2	16.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
149	45.7	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
151	31.1	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
156	41	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
158	38.7	15.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
159	43.5	12	51.5139	4.3009	29.1935	18.2931	27.5097	7.8950
161	36.6	14	51.3870	4.1782	29.9494	16.8310	23.7591	8.4335
162	30.2	18	50.9421	4.5114	30.1402	16.5077	26.4139	7.5750
164	38.1	14.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
165	34	16.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
167	42.3	11.5	50.6944	4.3292	30.4523	17.3773	24.9613	8.3502
169	33	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
176	29.8	12.5	51.4577	4.0646	30.1617	16.6140	23.1036	8.5412
177	41.2	14	50.9916	4.3196	30.7685	16.6048	24.7238	8.4415
178	28.7	16.5	51.8519	4.0284	29.3473	17.6206	23.6147	8.5166
179	39.5	16	52.4824	4.1297	29.5626	16.4775	25.9079	8.0726
181	20.2	16	51.5298	4.0525	28.8027	17.7053	26.9963	8.0503
182	30.5	14.5	52.2705	3.9388	29.2804	16.7866	25.5755	8.0850
183	33.5	14.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
184	29.9	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
185	34	18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
186	24.6	12	53.5198	3.6245	27.8626	18.1948	23.4449	8.6971
187	40.5	16.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
192	31.5	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
198	38	15.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
199	36.8	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
202	47.3	14.5	51.9931	3.9076	29.1867	16.8482	27.3103	7.8003
204	40.8	15.5	54.0410	3.7653	28.5734	16.5827	22.0819	8.5482
205	28.7	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
206	42.9	11.5	53.0348	4.0585	30.0089	16.0534	24.2753	8.5183
209	28.7	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
213	36.8	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
215	27.3	17	53.9853	4.2579	28.4218	17.9662	26.1185	8.0982
216	46.5	13	53.0624	4.2887	29.6882	17.1069	27.0412	7.8748

Anexo D. Dados fenotípicos para as características avaliadas no capítulo 3 que foram utilizadas para mapeamento de QTL

Ind. Sel.	CAP	Pilodyn	Rendimento	CE	Lignina	Pentosanas	Fibra/Grama	Coarseness
306	40.1	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
308	39.3	17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
309	32.7	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
311	43.9	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
312	39.3	14.5	54.2765	3.7859	28.3559	16.7214	25.4774	8.1993
313	29.5	13.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
319	30.8	13.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
320	26.7	15	54.2904	4.0793	28.3884	17.3254	25.0709	8.1284
321	37.2	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
323	36.4	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
326	28.3	13.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
328	25.2	14.5	54.7490	3.6869	28.2505	17.1498	22.4372	8.5783
330	21.3	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
332	30	13.5	53.0966	4.3172	29.3215	16.9145	28.1792	7.3374
333	39.7	13.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
335	41.7	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
340	41.7	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
342	42.3	15.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
343	37.2	14.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
345	38	12	53.0438	4.0247	29.5819	16.0275	20.2396	8.2247
351	37.9	15.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
353	19.8	18	53.8398	4.0044	28.5180	17.1781	20.2788	8.5012
357	42.2	11.5	53.3270	3.8247	29.3685	15.9956	22.5958	8.1639
361	28.2	12.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
362	30.1	15.5	53.7160	4.1037	28.5553	17.5454	24.2374	7.6679
363	34.5	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
365	29.4	13.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
366	36	16.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
368	43.8	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
373	43.8	13.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
374	38.1	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
375	35.8	19	53.6592	3.9611	28.7037	16.7987	23.7118	7.9805
376	20.1	14	54.1398	3.7126	27.9351	17.1373	24.9160	7.8261
378	38	12.5	52.7635	4.0956	29.3540	17.3765	25.9424	7.5942
379	22.6	12.5	53.5813	3.8623	28.3438	17.2837	24.6539	7.8469
380	28.5	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
383	35.4	12	53.3997	4.0125	29.2797	16.5710	24.9235	7.7685
385	33.3	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
386	35.4	14.5	53.5091	3.9865	29.2413	16.2750	23.8391	7.7280
390	40.1	13.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
393	19.6	13	51.3284	4.5908	29.3258	17.5965	29.2907	7.1462
394	40.1	18	53.0365	4.2005	29.4898	16.4871	24.5937	7.6154
396	44.1	13.5	53.9405	3.8348	28.7193	16.4053	24.8466	7.6012
398	42	7.5	52.8110	3.8548	29.2745	17.2413	25.6255	8.0878

Anexo D. Dados fenotípicos para as características avaliadas no capítulo 3 que foram utilizadas para mapeamento de QTL

Ind. Sel.	CAP	Pilodyn	Rendimento	CE	Lignina	Pentosanas	Fibra/Grama	Coarseness
399	27.9	13	53.9557	3.7311	28.6500	17.1970	26.6583	7.5244
400	33.7	16.5	54.0256	3.8801	28.6862	16.8826	24.2007	7.7885
401	27.8	13.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
403	35.7	13.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
405	38.7	11.5	53.6574	3.9545	29.0272	16.2197	25.9748	7.5230
406	46.5	15	54.9963	3.6735	28.7845	15.6676	23.3565	7.9793
407	36	15	53.0847	4.0159	29.1639	16.8029	25.6110	7.6092
411	36.3	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
412	36.2	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
413	17.7	15	53.3879	4.0568	28.7389	17.3559	25.4943	7.6087

REFERÊNCIAS BIBLIOGRÁFICAS

- AKANO, A. O.; DIXON, A. G. O.; BARRERA, E. et al. Genetic mapping of a dominant gene conferring resistance to cassava mosaic disease. **Theoretical and Applied Genetics**, v.105, p.521-525, 2002.
- ALTINKUT, A.; GOZUKIRMIZI, N. Search for microsatellite markers associated with water-stress tolerance in wheat through bulked segregant analysis. **Molecular Biotechnology**, v.23, p.97-106, Feb. 2003.
- ANDRADE, E. N. **O Eucalipto**. Jundiaí: Companhia Paulista de Estradas de Ferro, 1961. 660p.
- BAUCHER, M.; HALPIN, C.; PETIT-CONIL, M. et al. Lignin: Genetic Engineering and Impact on Pulping. **Critical Reviews in Biochemistry and Molecular Biology**, v.38, p.305-350, 2003.
- BLAZQUEZ, M. A.; GREEN, R.; NILSSON, O. et al. Gibberellins promote flowering of Arabidopsis by activating the LEAFY promoter. **The Plant Cell**, v.10, p.791-800, 1998.
- BOERJAN, W.; RALPH, J.; BAUCHER, M. Lignin biosynthesis. **Annual Review of Plant Biology**, v.54, p.519-546, 2003.
- BOLAND, D. J.; BROOKER, M. I. H.; CHIPPENDALE, G. M. et al. **Forest trees of Australia**. Melbourne, Austrália: CSIRO Publications, 1994. 687p.
- BOUHLAKA, C.; OTHMAN, T. B.; AISSAOUI, L. et al. Fanconi anemia: contribution of molecular analyses to the identification of bone marrow graft donors and the study of chimerism in grafted patients. **Genetic Testing**, v.8, p.268-275, 2004.

- BRAY, N. J.; BUCKLAND, P. R.; OWEN, M. J. et al. Cis-acting variation in the expression of a high proportion of genes in human brain. **Human Genetics**, v.113, p.149-153, 2003.
- BRONDANI R.P.V.; BRONDANI C.; GRATTAPAGLIA D. Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. **Molecular Genetics and Genomics**, v.267, p.338-347, 2002.
- BRONDANI, R. P. V. Desenvolvimento, mapeamento e caracterização de marcadores microssatélites em *Eucalyptus grandis* e *E. urophylla*. Brasília, 2000. 145p. Tese (Doutorado) - Universidade de Brasília.
- BRONDANI, R. P. V.; BRONDANI, C.; TARCHINI, R. et al. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. **Theoretical and Applied Genetetics**, v.97, p.816-827, 1998.
- BRONDANI, R. P.; BRONDANI, C.; GRATTAPAGLIA, D. Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. **Molecular Genetics and Genomics**, v.267, p.338-347, 2002.
- BRONDANI, R. P.; D. GRATTAPAGLIA. Cost-effective method to synthesize a fluorescent internal DNA standard for automated fragment sizing. **Biotechniques**, v.31, p.793-795, 2001.
- BUTCHER, A.; WILLIAMS, R.; WHITAKER, D. et al. Improving linkage analysis in outcrossed forest trees - an example from *Acacia mangium*. **Theoretical and Applied Genetics**, v.104, p.1185-1191, 2002.
- BYRNE M.; MURRELL J. C.; OWEN J. V. et al. Identification and mode of action of quantitative trait loci affecting seedling height and leaf area in *Eucalyptus nitens*. **Theoretical and Applied Genetics**, v.94, p.674-681, 1997.

- BYRNE, N.; MARQUEZ-GARCIA M. Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus*. **Australian Journal of Botany**, v.44, p.331-341, 1996.
- BYRNE M.; MURRELL JC.; ALLEN B.; MORAN G. et al. An integrated genetic linkage map for eucalypts using RFLP, RAPD and isozyme markers. **Theoretical and Applied Genetics**, v. 91, p.869-875, 1995.
- CARLEOS, C.; BARO, J. A.; CANON, J. et al. Asymptotic variances of QTL estimators with selective DNA pooling. **Journal of Heredity**, v.94, p.175-179, 2003.
- CAVALCANTI, G. Identificação das Principais Espécies de Eucalyptus Existentes no Brasil. **Silvicultura em São Paulo**, v.1, p.149-169, 1963.
- CHABBERT, B.; TOLLIER, M. T.; MONTIES, B. et al. Biological variability in lignification of maize: expression of the brown midrib *bm3* mutation in three maize cultivars. **Journal of Science Food Agriculture**, v.64, p.349-355, 1994.
- CHETELAT, R. T.; MEGLIC. V.; CISNEROS, P. A genetic map of tomato based on BC1 *Lycopersicon esculentum* - *Solanum lycopersicoides* reveals overall synteny but suppressed recombination between these homeologous genomes. **Genetics**, v.154, p.857-867, 2000.
- CHIANG, V. L.; PUMALA, R.J.; TAKEUCHI, H. et al. Comparison of softwood and hardwood kraft pulping. **Tapping Journal**, v.71, p.173-176, 1988.
- CHING, A.; CALDWELL, K. S.; JUNG, M. et al. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. **BMC Genetics**, v.3, p.19, 2002.
- CHURCHILL, G. A.; DOERGE, R. W. Empirical threshold values for quantitative trait mapping. **Genetics**, v.138, p.963-971, 1994.

- COEN, E. S.; MEYEROWITZ, E. M. The war of the whorls: genetic interactions controlling flower development. **Nature**, v.353, p.31-37, 1991.
- COLLEVATTI, R. G.; BRONDANI, R. V.; GRATTAPAGLIA, D. Development and characterization of microsatellite markers for genetic analysis of a Brazilian endangered tree species *Caryocar brasiliense*. **Heredity**, v.83, p.748-756, 1999.
- COLLINS, H. E.; LI, H.; INDA, S. E. et al. A simple and accurate method for determination of microsatellite total allele content differences between DNA pools. **Human Genetics**, v.106, p.218-226, 2000.
- CREGAN, P.; QUIGLEY, C.; ZHU, Y. et al. Detection of Single Nucleotide Polymorphisms (SNPs) in Soybean. /Apresentado em Plant and Animal Genome IX Conference, San Die.go, 2001/
http://www.intlpag.org/pag/9/abstracts/P3e_11.html. (4 Feb. 2005).
- CREGAN, P. B.; BHAGWAT, A. A.; AKKAYA, M. S. Microsatellite fingerprinting and mapping of soybean. **Methods in Molecular and Cell Biology**, v.5, p.49-61, 1994.
- CSEKE, L. J.; PODILA, G. K. MADS-box genes in dioecious aspen II: A review of MADSbox genes from trees and their potential in forest biotechnology. **Physiology and Molecular Biology of Plants**, v.10, p.7-28, 2004.
- DARVASI, A.; SOLLER, M. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. **Theoretical and Applied Genetics**, v.85, p.353-359, 1992.
- DECROOCQ, V.; ZHU, X.; KAUFFMAN, M. et al. A TM3-like MADS-box gene from *Eucalyptus* expressed in both vegetative and reproductive tissues. **Gene**, v.228, p.155-160, 1999.
- DEKKERS, J. C. M.; HOSPITAL F. The use of molecular genetics in the improvement of agricultural populations. **Nature Reviews Genetics**, v. 2, p.22-32, 2002.

- DIWAN, N.; CREGAN, P. B. Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. **Theoretical and Applied Genetics**, v.95, p.723-733, 1997.
- DOERGE, R. W. Mapping and analysis of quantitative trait loci in experimental populations. **Nature Reviews Genetics**, v.3, p.43-52, 2003
- DOYLE, J. J.; J. L. DOYLE. Isolation of plant DNA from fresh tissue. **Focus**, v.12, p.13-15, 1987.
- EDWARDS, A.; CIVITELLO, A.; HAMMOND, H. A. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. **American Journal of Human genetics**, v.49, p.746-756, 1991
- EDWARDS, Y. J.; ELGAR, G.; CLARK, M. S. et al. The identification and characterization of microsatellites in the compact genome of the Japanese pufferfish, *Fugu rubripes*: perspectives in functional and comparative genomic analyses. **Journal of Molecular Biology**, v.278, p.843-854, 1998.
- ELDRIDGE, K.; DAVIDSON, D.; HARDWOOD, C. et al. **Eucalypt Domestication and Breeding**. Oxford, New York: Clarendon Press, 1994. 288p.
- EWING, B.; GREEN, P. Basecalling of automated sequencer traces using phred. II. Error probabilities. **Genome Research**, v.8, p.186-194, 1998.
- FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO. **Programa Genoma do Eucalipto**. [http://www.fapesp.br/materia.php?data\[id_materia\]=468](http://www.fapesp.br/materia.php?data[id_materia]=468) (28 Feb. 2005)
- GAN, S.; SHI, J.; LI, M. et al. Moderate-density molecular maps of *Eucalyptus urophylla* S. T. Blake and *E. tereticornis* Smith genomes based on RAPD markers. **Genetica**, v.118, p.59-67, 2003.
- GARCIA, C. H.; F. PIMENTEL-GOMES. Forest outlines of Brazil. **Revista de Agricultura**, v.67, p.105-107, 1992.

- GARCIA-GIL, M. R.; MIKKONEN, M.; SAVOLAINEN, O. Nucleotide diversity at two phytochrome loci along a latitudinal cline in *Pinus sylvestris*. **Molecular Ecology**, v.12, p.1195-1206, 2003.
- GARNER, T. W. Genome size and microsatellites: the effect of nuclear size on amplification potential. **Genome**, v.45, p.212-215, 2002.
- GION, J. M.; RECH, P.; GRIMA-PETTENATI, J. et al. Mapping candidate genes in *Eucalyptus* with emphasis on lignification genes. **Molecular Breeding**, v.6, p.441-449, 2000.
- GOFFNER, D.; VAN DOORSSELAERE, J.; YAHIAOUI, N. et al. A novel aromatic alcohol dehydrogenase in higher plants: molecular cloning and expression. **Plant Molecular Biology**, v.36, p.755-765, 1998.
- GOUJON, T.; SIBOUT, R.; POLLET, B. et al. A new *Arabidopsis thaliana* mutant deficient in the expression of o-methyltransferase impacts lignins and sinapoyl esters. **Plant Molecular Biology**, v.51, p.973-989, 2003.
- GRATTAPAGLIA, D. Genômica Florestal In: MIR, L. (Ed.) **Genômica**. Editora Atheneu, 2004a. p.917-934.
- GRATTAPAGLIA, D. Molecular breeding of *Eucalyptus* - State of the art, operational applications and technical challenges. **Molecular Biology Of Woody Plants**, v.1, p.451-474, 2000.
- GRATTAPAGLIA, D. Genolyptus. In: BORÉM, A.; GIUDICE, M.; SEDYIAMA, T. (Eds). **Melhoramento Genômico**. Viçosa: UFV, 2003. cap 3, p. 51-71.
- GRATTAPAGLIA, D. RIBEIRO, V.J. REZENDE, G.D.S.P. Retrospective selection of elite parent trees using paternity testing with microsatellite markers: an alternative short term breeding tactic for *Eucalyptus*. **Theoretical and Applied Genetics**, v.109, p.192-199, 2004b.

- GRATTAPAGLIA, D.; BERTOLUCCI, F. L.; PENCHEL, R. et al. Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers. **Genetics**, v.144, p.1205-1214, 1996b.
- GRATTAPAGLIA, D.; BERTOLUCCI F. L.; SEDEROFF R. R. Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers. **Genetics**, v. 144, p.1205-1214, 1996a.
- GRATTAPAGLIA, D.; BERTOLUCCI, F. L.; PENCHEL, R. et al. Genetic mapping of quantitative trait loci controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. **Theoretical and Applied Genetics**, v.90, p.933-947, 1995.
- GRATTAPAGLIA, D.; H. J. BRADSHAW. Nuclear DNA content of commercially important *Eucalyptus* species and hybrids. **Canadian Journal of Forest Research**, v.24, p.1074-1078, 1994a.
- GRATTAPAGLIA, D.; SEDEROFF, R. Genetic linkage maps of *eucalyptus grandis* and *eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. **Genetics**, v. 137, p.1121-1137, 1994b.
- GUO, D.; CHEN, F.; INOUE, K. et al. Downregulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-omethyltransferase in transgenic alfalfa: impacts on lignin structure and implications for the biosynthesis of G and S lignin. **Plant Cell**, v.13, p.73-88, 2001.
- GUO, D.C.; MILEWICZ, D.M. Methodology for using a universal primer to label amplified DNA segments for molecular analysis. **Biotechnology Letters**, v.25, p.2079-2083, 2003.

- HAWKINS, S. W.; BOUDET, A. M. Purification and characterization of cinnamyl alcohol dehydrogenase isoforms from the periderm of *Eucalyptus dunnii* Hook. **Plant Physiology**, v.104, p.75-84, 1994.
- HU, W. J.; HARDING, S. A.; LUNG, J. et al. Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. **Nature Biotechnology**, v.17, p.808-12, 1999.
- HULBERT, S. H.; ILOTT, T. W.; LEGG, E. J.; LINCOLN, S. E.; LANDER, E. S.; MICHELMORE, R. W. Genetic analysis of the fungus, *Bremia lactucae*, using restriction fragment length polymorphisms. **Genetics**, v.120, p.947-958, 1988.
- JOHNSON, N. A.; PORTER, A. H. Rapid speciation via parallel, directional selection on regulatory genetic pathways. **Journal of Theoretical Biology**, v.205, p.527-542, 2000.
- JOUANIN, L.; GOUJON, T.; NADAI, V. et al. Lignification in transgenic poplars with extremely reduced caffeic acid o-methyltransferase activity. **Plant Physiology**, v.123, p.1363-1374, 2000.
- JUNGHANS, D. T.; ALFENAS A. C.; BROMMONSCHENKEL S. H. et al. Resistance to rust (*Puccinia psidii* Winter) in Eucalyptus: mode of inheritance and mapping of a major gene with RAPD markers. **Theoretical and Applied Genetics**, v.108, p.175-180, 2003.
- KIRST, M.; BASTEN, C. J.; MYBURG, A. A.; ZENG, Z. B.; SEDEROFF, R. R. Genetic architecture of transcript level variation in differentiating xylem of an *Eucalyptus* hybrid. **Genetics**, Jan. 2005. /no prelo/
- KIRST, M.; CORDEIRO, C. M.; REZENDE, G. D. S. P.; GRATTAPAGLIA, D. Power of microsatellite markers for fingerprinting and parentage analysis in *Eucalyptus grandis* breeding populations. **Journal of Heredity**, v.96, p.161-166, 2005.

- KIRST, M.; MYBURG, A. A.; DE LEON, J. P.; KIRST, M. E.; SCOTT, J.; SEDEROFF, R. Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of eucalyptus. **Plant Physiology**, v.135, p.2368-2378, 2004.
- KIRST, M. Desenvolvimento e aplicação de sistemas de genotipagem multiloco, semi-automatizados, baseados em marcadores microssatélite, para espécies do gênero *Eucalyptus*. Viçosa, 1999. 146p. Dissertação (Mestrado) - Universidade Federal de Viçosa.
- KORTH, K. L.; BLOUNT, J. W.; CHEN, F. et al. Changes in phenylpropanoid metabolites associated with homology dependent silencing of phenylalanine ammonialyase and its somatic reversion in tobacco. **Physiologia Plantarum**, v.111, p.137-143, 2001.
- KOSCHINSKY, M. L.; BOFFA, M. B.; NESHEIM, M. E. et al. Association of a single nucleotide polymorphism in CPB2 encoding the thrombin-activable fibrinolysis inhibitor (TAF1) with blood pressure. **Clinical Genetics**, v.60, p.345-349, 2001.
- KREIKE, C. M.; STIEKEMA, W. J. Reduced recombination and distorted segregation in a *Solanum tuberosum* (2x) - *S. spegazzinii* (2x) hybrid. **Genome**, v.40, p.180-187, 1997.
- KYOZUKA, J.; HARCOURT, R.; PEACOCK, W. J. et al. *Eucalyptus* has functional equivalents of the *Arabidopsis AP1* gene. **Plant Molecular Biology**, v.35, p.573-584, 1997.
- LACOMBE, E.; HAWKINS, S.; VAN DOORSSELAERE, J. et al. Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. **Plant Journal**, v.11, p.429-441, 1997.

- LANDER ES.; GREEN P.; ABRAHAMSON J. et al. Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. **Genomics**, v.1, p.174 -181, 1987.
- LANDER, E.; BOTSTEIN, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, v.121, p.185-199, 1989.
- LAPIERRE, C.; POLLET, B.; PETIT-CONIL, M. et al. Structural alterations of lignins in transgenic poplars with depressed cinnamyl alcohol dehydrogenase or caffeic acid o-methyltransferase activity have an opposite impact on the efficiency of industrial kraft pulping. **Plant Physiology**, v.119, p.153-164, 1999.
- LEVINE, M. How insects lose their limbs. **Nature**, v.415, p.848, 2002.
- LHEUREUX, F.; CARREEL, F.; JENNY, C. et al. Identification of genetic markers linked to banana streak disease expression in inter-specific *Musa* hybrids. **Theoretical and Applied Genetics**, v.106, p.594-598, 2003.
- LI, L.; POPKO, J. L.; UMEZAWA, T. et al. 5-hydroxyconiferyl aldehyde modulates enzymatic methylation for syringyl monolignol formation, a new view of monolignol biosynthesis in angiosperms. **Journal of Biological Chemistry**, v.275, p.6537-6545, 2000.
- LI, L.; ZHOU, Y.; CHENG, X. et al. Combinatorial modification of multiple lignin traits in trees through multigene cotransformation. **Proceedings of the National Academy of Sciences of the United States of America**, v.100, p.4939-4944, 2003.
- LIMA, W. P. **Impacto ambiental do Eucalipto**. São Paulo: Editora da Universidade de São Paulo, 1993. 302 p.
- LITT, M.; LUTY, J.A. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. **American Journal of Human genetics**, v.44, p.397-401, 1989.

- LOURENÇO R.T. Estrutura genômica de três megabases de DNA genômico e seqüenciamento por fragmentação randômica de DNA ("*shotgun*") de *Eucalyptus*: conteúdo nucleotídico, seqüências repetitivas e genes. Campinas, SP, 2004. 164p. Dissertação (Mestrado) - Universidade de Campinas.
- MARITA, J. M.; RALPH, J.; HATFIELD, R. D. et al. Structural and compositional modifications in lignin of transgenic alfalfa down-regulated in caffeic acid 3-o methyltransferase and caffeoyl coenzyme A 3-o methyltransferase. **Phytochemistry**, v.62, p.53-65, 2003.
- MARQUES C. M.; ARAÚJO J. A.; FERREIRA J. G. et al. AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. **Theoretical and Applied Genetics**, v.96, p.727-737, 1998.
- MARQUES C.M.; VASQUEZ-KOOL J.; CAROCHA V.J. et al. Genetic dissection of vegetative propagation traits in *Eucalyptus tereticornis* and *E. globulus*. **Theoretical and Applied Genetics**, v.99, p.936-946, 1999.
- MEYERMANS, H.; MORREEL, K.; LAPIERRE, C. et al. Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A o-methyltransferase, an enzyme involved in lignin biosynthesis. **Journal of Biological Chemistry**, v.275, p.36899-36909, 2000.
- MICHELMORE, R. W.; PARAN, I.; KESSELI, R. V. et al. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. **Proceedings of the National Academy of Sciences of the United States of America**, v.88, p.9828-9832, 1991.
- MIKAWA, S.; SHIMANUKI, S.; MOROZUMI, T. et al. Comparative analysis and development of microsatellite markers on swine (*Sus scrofa*) chromosome 1qter. **Animal Genetics**, v.35, p.445-450, 2004.

- MORA, A.; GARCIA, C. **A Cultura do Eucalipto no Brasil**. 1.ed. São Paulo: Sociedade Brasileira de Silvicultura, 2000. 112p
- MORAN, G. F.; THAMARUS, K. A.; RAYMOND, C. A. et al. Genomics of *Eucalyptus* wood traits. **Annals of Forest Science**, v.59, p.645-650, 2002.
- MORGANTE, M.; HANAFEY, M.; POWELL, W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. **Nature Genetics**, v.30, p.194-200, 2002.
- MORGANTE, M.; SALAMINI, F. From plant genomics to breeding practice. **Current Opinion in Biotechnology**, v.14, p.214-219, 2003
- MYBURG, A. A. **Molecular genetic analysis of wood quality traits in eucalyptus**. http://tappsa.co.za/archive/APPW2002/Title/Molecular_genetic_analysis_of_/molecular_genetic_analysis_of_.html (28 Feb. 2004)
- MYBURG A. A.; GRIFFIN, A. R.; SEDEROFF, R. R. et al. Comparative genetic linkage maps of *Eucalyptus grandis*, *Eucalyptus globulus* and their F1 hybrid based on a double pseudo-backcross mapping approach. **Theoretical and Applied Genetics**, v.107, p.1028-1042, 2003.
- MYBURG, A. A. Genetic architecture of hybrid fitness and wood quality traits in a wide interspecific cross of *Eucalyptus* tree species. Raleigh, 2001. 228p. Thesis (PhD) - North Carolina State University
- NEILAN, B. A.; WILTON, A. N.; JACOBS, D. A universal procedure for primer labelling of amplicons. **Nucleic Acids Research**, v.25, p.2938-2939, 1997.
- O'CONNELL, A.; HOLT, K.; PIQUEMAL, J. et al. Improved paper pulp from plants with suppressed cinnamoyl-CoA reductase or cinnamyl alcohol dehydrogenase. **Transgenic Research**, v.11, p.495-503, 2002.
- OETTING, W. S.; LEE, H. K.; FLANDERS, D. J. et al. Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. **Genomics**, v.30, p.450-458, 1995.

- PARENICOVA, L.; FOLTER, S.; KIEFFER, M. et al. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: New openings to the MADS world. **Plant Cell**, v.15, p.1538-1551, 2003.
- PÉREZ-ENCISO, M. Sequential bulked typing: a rapid approach for detecting QTL. **Theoretical and Applied Genetics**, v.96, p.551-557, 1998.
- PILATE, G.; GUINEY, E.; HOLT, K. et al. Field and pulping performances of transgenic trees with altered lignification. **Nature Biotechnology**, v.20, p.607-612, 2002.
- PILATE, G.; GUINEY, E.; HOLT, K. et al. Field and pulping performances of transgenic trees with altered lignification. **Nature Biotechnology**, v.20, p.607-612, 2002.
- PINCON, G.; MAURY, S.; HOFFMANN, L. et al. Repression of o-methyltransferase genes in transgenic tobacco affects lignin synthesis and plant growth. **Phytochemistry**, v.57, p.1167-1176, 2001.
- PRYOR, L. *Eucalyptus*. In:HALEVY A. H. (Ed) **Handbook of flowering**. 1.ed. Boca Raton: Franklin Book Company Inc./CRC Press,1985. cap.2, p. 476-482.
- RAFALSKI, A. Applications of single nucleotide polymorphisms in crop genetics. **Current Opinion in Plant Biology** , v.5, p.94-100, 2002.
- RAFALSKI, J. A.; TINGEY S. V. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. **Trends in Genetics**, v.9, p.275-280, 1993.
- RATCLIFFE, O. J.; KUMIMOTO, R. W.; WONG, B. J. et al. Analysis of the *Arabidopsis* MADS AFFECTING FLOWERING gene family: *MAF2* prevents vernalization by short periods of cold. **Plant Cell**, v.15, p.1159-1169, 2003.
- REMYNGTON, D. L.; WHETTEN, R. W.; LIU, B. H. et al. Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. **Theoretical and Applied Genetics**, v. 101, p.1279-1292, 1999.

- ROUSE, D. T.; SHELDON, C. C.; BAGNALL, D. J. et al. FLC, a repressor of flowering, is regulated by genes in different inductive pathways. **Plant Journal**, v.29, p.183-191, 2002.
- SOCIEDADE BRASILEIRA DE SILVICULTURA <http://www.sbs.org.br>. (4 Mar. 2005)
- SCHARF, R. Falta madeira na terra do Pau-Brasil. **Revista Galileu**, maio 2003. p.52-60.
- SCHLOTTERER, C.; TAUTZ, D. Slippage synthesis of simple sequence DNA. **Nucleic Acids Research**, v.20, p.211-215, 1992.
- SCHMITZ, J.; FRANZEN, R.; NGYUEN, T. H. et al. Cloning, mapping and expression analysis of barley MADS-box genes. **Plant Molecular Biology**, v.42, p.899-913, 2000.
- SCHUELKE, M. An economic method for the fluorescent labeling of PCR fragments. **Nature Biotechnology**, v.18, p.233-234, 2000.
- SEWALT, V.; NI, W.; BLOUNT, J. W. et al. Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of l-phenylalanine ammonia-lyase or cinnamate 4-hydroxylase. **Plant Physiology**, v.115, p.41-50, 1997.
- SHAW, S. H.; CARRASQUILLO, M. M.; KASHUK, C. et al. Allele frequency distributions in pooled DNA samples: applications to mapping complex disease genes. **Genome Research**, v.8, p.111-123, 1998.
- SHELDON, C. C.; BURN, J. E.; PEREZ, P. P. et al. The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. **Plant Cell**, v.11, p.445-458, 1999.
- SHINOZUKA, Y.; KOJIMA, S.; SHOMURA, A. et al. Isolation and characterization of rice MADS box gene homologues and their RFLP mapping. **DNA Research**, v.6, p.123-129, 1999.

- SIMPSON, G. G.; GENDALL, A. R.; DEAN, C. When to switch to flowering. **Annual Review of Cell and Developmental Biology**, v.15, p.519-550, 1999.
- SMYTH, D. R. Plant genetics: Fast flowering. **Current Biology**, v.6, p.122-124, 1996.
- SOUTHERTON, S. G.; MARSHALL, H.; MOURADOV, A. et al. Eucalypt MADS-box genes expressed in developing flowers. **Plant Physiology**, v.118, p.365-372, 1998a.
- SOUTHERTON, S. G.; STRAUSS, S. H.; OLIVE, M. R. et al. *Eucalyptus* has a functional equivalent of the Arabidopsis floral meristem identity gene LEAFY. **Plant Molecular Biology**, v.37, p.897-910. 1998b.
- SQUILASSI, M. G. Mapping QTL using linkage disequilibrium and efficiency of early marker assisted selection in *Eucalyptus*. Botucatu, 1997. 84p. Dissertação (M.S.) - Universidade Estadual Paulista.
- STAM, P. Construction of integrated genetic linkage maps by means of a new computer package: Joinmap. **Plant Journal**, v.3, p.739-744, 1993.
- STAM, P.; VAN OOIJEN, J. W. **Joinmap version 2.0**: software for the calculation of genetic linkage maps. Wageningen: CPRO-DLO, 1995.
- STERKY, F.; REGAN, S.; KARLSSON, J. et al. 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**, v.95, p.13330-13335, 1998.
- STRAUSS, S. H.; LANDE R.; NAMKOONG G. et al. Limitations of molecular marker-aided selection in forest tree breeding. **Canadian Journal of Forest Research**, v.32, p.1050-1061, 1992.
- STYLIANOU, I. M.; CHRISTIANS, J. K.; KEIGHTLEY, P. D. et al. Genetic complexity of an obesity QTL (*Fob3*) revealed by detailed genetic mapping. **Mammalian Genome**, v.15, p.472-481, 2004.

- TADEGE, M.; SHELDON, C. C.; HELLIWELL, C. A. et al. Control of flowering time by FLC orthologues in *Brassica napus*. **Plant Journal**, v.28, p.545-553, 2001.
- TANKSLEY, S. D. Mapping polygenes. **Annual Review of Genetics**, v.27, p.205-233, 1993.
- THAMARUS, K.; GROOM, K.; BRADLEY, A.; RAYMOND, C. A.; SCHIMLECK, L. R.; WILLIAMS, E. R.; MORAN, G. F. Identification of quantitative trait loci for wood and fibre properties in two full-sib properties of *Eucalyptus globulus*. **Theoretical and Applied Genetics**, v.109, p.856-864, 2004
- THAMARUS, K. A.; GROOM, K.; MURRELL, J. et al. A genetic linkage map for *Eucalyptus globulus* with candidate loci for wood, fibre, and floral traits. **Theoretical and Applied Genetics**, v.104, p.379-387, 2002.
- THOMAS, M. R.; CAIN, P.; SCOTT, N. S. DNA typing of grapevines: a universal methodology and database for describing cultivars and evaluating genetic relatedness. **Plant Molecular Biology**, v.25, p.939-949, 1994.
- TOTH, G.; GASPARI, Z.; JURKA, J. Microsatellites in different eukaryotic genomes: survey and analysis. **Genome Research**, v.10, p.967-981, 2000.
- TSAI, C. J.; POPKO, J. L.; MIELKE, M. R. et al. Suppression of o-methyltransferase gene by homologous sense transgene in quaking aspen causes red-brown wood phenotypes. **Plant Physiology**, v.117, p.101-112, 1998.
- VAN DER LINDEN, C. G.; VOSMAN, B.; SMULDERS, M. J. Cloning and characterization of four apple MADS box genes isolated from vegetative tissue. **Journal of Experimental Botany**, v.53, p.1025-1036, 2002.
- VAN DOORSSELAERE, J.; BAUCHER, M.; CHOENOT, E. et al. A novel lignin in poplar trees with a reduced caffeic acid/5-hydroxyferulic acid O-methyltransferase activity. **Plant Journal**, v.8, p.855-864, 1995.

- VERHAEGEN D.; PLOMION, C. Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus grandis* using RAPD markers. **Genome**, v.39, p.1051-1061, 1996.
- VERHAEGEN, D.; PLOMION, C.; GION, J. M. et al. Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers.1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. **Theoretical and Applied Genetics**, v.96, p.597-608, 1997.
- WANG, S.; BASTEN, C. J.; ZENG, Z. B. **Windows QTL Cartographer 2.0**. Raleigh: Department of Statistics, North Carolina State University, 2001. <http://statgen.ncsu.edu/qtlcart/WQTLCart.htm> (26 Feb. 2004)
- WANG, Z.; WEBER, J. L.; ZHONG, G.; TANKSLEY, S. D. Survey of plant short tandem DNA repeats. **Theoretical and Applied Genetics**, v.88, p.1-6, 1994.
- WEBER, J. L.; MAY, P. E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. **American Journal of Human genetics**, v.44, p.388-396, 1989.
- WENBURG, J. K.; OLSEN, J. B.; BENTZEN, P. Multiplexed systems of microsatellites for genetic analysis in coastal cutthroat trout (*Oncorhynchus clarki clarki*) and steelhead (*Oncorhynchus mykiss*). **Molecular Marine Biology and Biotechnology**, v.5, p.273-283, 1996.
- WHETTEN, R. W.; MACKAY, J. J.; SEDEROFF, R. R. Recent Advances In Understanding Lignin Biosynthesis. **Annual Review of Plant Physiology and Plant Molecular Biology**, v.49, p.585-609, 1998.
- WHITE, G.; W. POWELL. Isolation and characterization of microsatellite loci in *Swietenia humilis* (Meliaceae): an endangered tropical hardwood species. **Molecular Ecology**, v.6, p.851-860, 1997.
- WILLIAMS, C. G.; GOODMAN, M. M.; STUBER, C. W. Comparative recombination distances among *Zea mays* L. inbreds, wide crosses and interspecific hybrids. **Genetics**, v.141, p.1573-1581, 1995.

- YAN, H.; YUAN, W.; VELCULESCU, V. E. et al. Allelic variation in human gene expression. **Science**, v.297, p.1143, 2002.
- ZAINA, G.; MORGANTE, M. Nucleotide diversity and Linkage Disequilibrium in *Populus nigra*. **New Phytologist**. /no prelo/
- ZENG, Z. B. Precision mapping of quantitative trait loci. **Genetics**, v.139, p.1457-1468, 1994.
- ZHONG, R.; MORRISON, W. H.; HIMMELSBACH, D. S. et al. Essential role of caffeoyl coenzyme o-methyltransferase in lignin biosynthesis in woody poplar plants. **Plant Physiology**, v.124, p.563-578, 2000.
- ZHU, Y. L.; SONG, Q. J.; HYTEN, D. L. et al. Single-nucleotide polymorphisms in soybean. **Genetics**, v.163, p.1123-1134, 2003.
- ZHU, Y.; SPITZ, M. R.; LEI, L. et al. A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter enhances lung cancer susceptibility. **Cancer Research**, v.61, p.7825-7829, 2001.