

**MAPEAMENTO DE QTL PARA CARACTERÍSTICAS DE  
INTERESSE INDUSTRIAL DA MADEIRA DE *Eucalyptus*  
EM PROGÊNIE DE HÍBRIDOS INTERESPECÍFICOS**

**LUCIANO MEDINA MACEDO**

Orientador:

**Prof. Dr. Alexandre Siqueira Guedes Coelho**

Fevereiro - 2009

**TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS NA BIBLIOTECA DIGITAL DA UFG**

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

**1. Identificação do material bibliográfico:**      ☒ **Dissertação**      ☐ **Tese**

**2. Identificação da Tese ou Dissertação**

Nome completo do autor: Luciano Medina Macedo

Título do trabalho:

MAPEAMENTO DE QTL PARA CARACTERÍSTICAS DE INTERESSE INDUSTRIAL DA MADEIRA DE *Eucalyptus* EM PROGÊNIE DE HÍBRIDOS INTERESPECÍFICOS

**3. Informações de acesso ao documento:**

Concorda com a liberação total do documento ☒ SIM      ☐ NÃO<sup>1</sup>

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.



Assinatura do (a) autor (a)

Data: 20 / 09 / 2016

<sup>1</sup> Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

**LUCIANO MEDINA MACEDO**

**MAPEAMENTO DE QTL PARA CARACTERÍSTICAS DE  
INTERESSE INDUSTRIAL DA MADEIRA DE *Eucalyptus* EM  
PROGÊNIE DE HÍBRIDOS INTERESPECÍFICOS**

Dissertação apresentada ao Programa de Pós-graduação em Agronomia, da Universidade Federal de Goiás, como requisito parcial à obtenção do título de Mestre em Agronomia, área de concentração: Genética e Melhoramento de Plantas.

Orientador:

**Prof. Dr. Alexandre Siqueira Guedes Coelho**

Goiânia, GO - Brasil

2009

**Dados Internacionais de Catalogação na Publicação (CIP)**  
**GPT/BC/UFG**

**M141m** Medina-Macedo, Luciano.  
Mapeamento de QTL para características de interesse industrial da madeira de *Eucalyptus* em progênie de híbridos interespecíficos [manuscrito] / Luciano Medina Macedo. – Goiânia, 2009.  
143 f. : figs, tabs.

**Orientador:** Prof. Dr. Alexandre Siqueira Guedes Coelho.

**Dissertação (Mestrado) – Universidade Federal de Goiás, Escola de Agronomia e Engenharia de Alimentos, 2009.**

**Inclui lista de siglas e abreviações, figuras e tabelas.**  
**Bibliografia.**  
**Apêndices.**

**1. *Eucalyptus* 2. Mapeamento genético 3. Marcadores microssatélite 4. mapeamento de QTL 5. SAM. I. Universidade Federal de Goiás. Escola de Agronomia e Engenharia de Alimentos. II. Título.**

**CDU: 631.52:582.776.2**






MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DE GOIÁS  
ESCOLA DE AGRONOMIA E ENGENHARIA DE ALIMENTOS  
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA



**ATA DA REUNIÃO DA BANCA EXAMINADORA DA DEFESA DE DISSERTAÇÃO DE LUCIANO MEDINA MACEDO** - Aos vinte e seis dias do mês de fevereiro do ano de dois mil e nove (26/02/2009), às 08h30min, reuniram-se os componentes da Banca Examinadora: Prof. Alexandre Siqueira Guedes Coelho – Orientador, Prof. Sérgio Tadeu Sibov e Dr. Flávio Breseghello, para, sob a presidência do primeiro, e em sessão pública realizada no auditório do PPGA da EA/UFG, procederem à avaliação da defesa de dissertação intitulada: “Mapeamento de QTL para características de interesse industrial da madeira de *Eucalyptus* em progênie de híbridos interespecíficos”, em nível de **Mestrado**, área de concentração em **Genética e Melhoramento de Plantas**, de autoria de **Luciano Medina Macedo**, discente do Programa de Pós-Graduação em Agronomia da Universidade Federal de Goiás. A sessão foi aberta pelo presidente da Banca Examinadora, Prof. Alexandre Siqueira Guedes Coelho, que fez a apresentação formal dos membros da Banca. A palavra a seguir, foi concedida ao autor da dissertação que, em 30 minutos apresentou o seu trabalho. Terminada a apresentação, cada membro da Banca arguiu o examinando, tendo-se adotado o sistema de diálogo sequencial. Terminada a fase de arguição, procedeu-se à avaliação da defesa. Tendo-se em vista o que consta na Resolução nº. 658/2004 do CEPEC - Conselho de Ensino, Pesquisa, Extensão e Cultura, que regulamenta o Programa de Pós-Graduação em Agronomia e procedidas às correções recomendadas, a dissertação foi **APROVADA** por unanimidade, considerando-se integralmente cumprido este requisito para fins de obtenção do título de **MESTRE EM AGRONOMIA**, na área de concentração em **GENÉTICA E MELHORAMENTO DE PLANTAS** pela Universidade Federal de Goiás. A conclusão do curso dar-se-á quando da entrega, na secretaria do PPGA, da versão definitiva da dissertação, com as devidas correções. A Banca Examinadora recomenda a publicação de artigo científico, oriundo dessa dissertação em periódicos de circulação nacional e, ou, internacional, depois de procedidas as modificações sugeridas. Cumpridas as formalidades de pauta, às 12h00min a presidência da mesa encerrou esta sessão de defesa de dissertação e para constar eu, Welinton Barbosa Mota, secretário do PPGA, lavrei a presente Ata, que depois de lida e aprovada, será assinada pelos membros da Banca Examinadora em quatro vias de igual teor.

  
Prof. Alexandre Siqueira Guedes Coelho  
Presidente – EA/UFG

  
Prof. Sérgio Tadeu Sibov  
ICB I/UFG

  
Dr. Flávio Breseghello  
Embrapa Arroz e Feijão

## **AGRADECIMENTOS**

Aos professores do PPGA da UFG que contribuíram para minha formação acadêmica, em especial ao Prof. Dr. Alexandre Siqueira Guedes Coelho, que aceitou me orientar e confiou a responsabilidade de conduzir este trabalho.

Ao PPGA e a CAPES pela concessão de bolsa durante todo o período de mestrado. Às empresas e pesquisadores participantes do Consórcio Genolyptus, pela oportunidade de participar de um trabalho em rede, troca de informações relevantes sobre o material genético avaliado, fornecimento de reagentes e infra-estrutura para realização de um trabalho de alto nível, além da logística disponibilizada na etapa de coleta de material vegetal.

Aos colegas de laboratório e estagiários pela ajuda nas várias etapas deste trabalho, troca de experiências e amizade. Ao secretário do PPGA Wellington Barbosa, que em muito me ajudou durante o período em que morei em Goiânia.

Ao apoio da minha família e da minha linda e amada esposa Juliana Bittencourt, que apesar das dificuldades impostas pela distância, sempre esteve ao meu lado ajudando a superar as dificuldades que surgiram ao longo desta jornada.

# SUMÁRIO

<b>LISTA DE SIGLAS E ABREVIACÕES .....</b>	<b>06</b>
<b>LISTA DE FIGURAS.....</b>	<b>08</b>
<b>LISTA DE TABELAS .....</b>	<b>10</b>
<b>RESUMO.....</b>	<b>12</b>
<b>ABSTRACT .....</b>	<b>14</b>
<b>1 INTRODUÇÃO .....</b>	<b>16</b>
<b>2 REVISÃO DE LITERATURA .....</b>	<b>18</b>
2.1 O GÊNERO <i>Eucalyptus</i> .....	18
2.2 O <i>Eucalyptus</i> NO BRASIL .....	19
2.3 PROJETOS GENOMA .....	22
2.4 MAPAS GENÉTICOS .....	23
2.5 MAPEAMENTO DE QTL .....	27
2.5.1 Mapeamento de QTL por Análise de Marcas Simples - AMS .....	29
2.5.2 Mapeamento de QTL por Intervalo Simples - MIS .....	30
2.5.3 Mapeamento de QTL por Intervalo Composto - MIC .....	31
2.5.4 Mapeamento de QTL por Intervalo Múltiplo - MIM .....	32
2.5.5 Mapeamento associativo .....	32
2.6 MAPAS GENÉTICOS E MAPEAMENTO DE QTL EM <i>Eucalyptus</i> .....	33
2.6.1 Caracteres quantitativos .....	33
2.6.2 Mapas genéticos .....	35
2.6.3 Distorção de segregação em <i>Eucalyptus</i> .....	39
2.6.4 Mapeamento de QTL .....	40
<b>3 MATERIAL E MÉTODOS .....</b>	<b>45</b>
3.1 MATERIAL VEGETAL .....	45
3.2 ANÁLISE DOS CARACTERES QUANTITATIVOS .....	46
3.3 ESTIMAÇÃO DOS COMPONENTES DE VARIÂNCIA .....	48
3.4 EXTRAÇÃO DE DNA GENÔMICO .....	49
3.5 AMPLIFICAÇÃO DE LOCOS SSR .....	51
3.6 CONSTRUÇÃO DE MAPAS GENÉTICOS .....	54
3.7 ANÁLISE DE QTL .....	55
<b>4 RESULTADOS .....</b>	<b>56</b>
4.1 COMPONENTES DE VARIÂNCIA .....	56
4.1.1 Caracteres de crescimento .....	59

4.1.2	Densidade avaliada por pilodyn .....	61
4.1.3	Caracteres avaliados por NIRS .....	61
4.2	CORRELAÇÕES ENTRE CARACTERES QUANTITATIVOS .....	63
4.3	POLIMORFISMO DE LOCOS SSR .....	64
4.4	CONSTRUÇÃO DO MAPA GENÉTICO .....	66
4.5	DETECÇÃO DE QTL .....	69
4.5.1	Diâmetro na altura do peito - DAP .....	70
4.5.2	Altura .....	72
4.5.3	Volume .....	73
4.5.4	Densidade básica - DB .....	75
4.5.5	Teor de lignina total - LigTotal .....	76
4.5.6	Relação siringil/guaiacil - rSG .....	78
4.5.7	Rendimento depurado de celulose - RD .....	80
4.6	DISTRIBUIÇÃO DOS QTL .....	81
4.7	COMPARAÇÃO ENTRE DIFERENTES PROCESSOS DE SELEÇÃO DE GENÓTIPOS SUPERIORES .....	83
<b>5</b>	<b>DISCUSSÃO .....</b>	<b>88</b>
5.1	CARACTERES QUANTITATIVOS .....	88
5.1.1	Coeficiente de variação do erro experimental - CVe .....	88
5.1.2	Contribuição relativa das fontes de variância na expressão dos caracteres .....	89
5.1.3	Coeficientes de herdabilidade - $h^2$ .....	90
5.1.4	Correlações entre caracteres quantitativos - $r_{xy}$ .....	92
5.2	CONSTRUÇÃO DO MAPA GENÉTICO .....	95
5.2.1	Distorção de segregação .....	95
5.2.2	Mapa genético integrado .....	97
5.3	EFEITO DO TAMANHO DA AMOSTRA DO EXPERIMENTO DE MG NOS QTL E INTERAÇÕES IDENTIFICADAS .....	100
5.4	VALIDAÇÃO DOS QTL MAPEADOS NO CRUZAMENTO DG x GL2 .....	104
5.4.1	Caracteres de crescimento .....	105
5.4.2	Densidade .....	108
5.4.3	Propriedades de lignina .....	109
5.4.4	Rendimento de celulose .....	111
5.5	APLICAÇÃO DA SAM EM PROGRAMAS DE MELHORAMENTO .....	112
<b>6</b>	<b>CONCLUSÕES .....</b>	<b>117</b>
<b>7</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>118</b>
	<b>APÊNDICES .....</b>	<b>128</b>



## LISTA DE SIGLAS E ABREVIACÕES

ABI - Applied Biosystems Inc.

ABRAF - Associação Brasileira de Produtores de Florestas Plantadas

AFLP - Amplified Fragment Length Polymorphism

AMS - Mapeamento de QTL através de Análise por Marcas Simples

BA - Estado da Bahia

BID - Banco Interamericano de Desenvolvimento

BRACELPA - Associação Brasileira de Celulose e Papel

BSA - Bulk Segregant Analysis

CIA - Solução de Clorofórmio e Álcool Isoamílico

cM - Unidade de medida de distância genômica: centi-Morgan

CSIRO - Commonwealth Scientific and Industrial Research Organisation

CTAB - Solução Cetyltrimethylammonium Bromide

CVe - Coeficiente de Variação do Erro Experimental

EUCAGEN - *Eucalyptus* Genome Network

DAP - Diâmetro na Altura do Peito

DB - Densidade Básica

DNA - Ácido Desoxiribonucléico

DOE - US Department of Energy

eBLUP - Empirical Best Linear Unbiased Prediction

EDAS - Electrophoresis Documentation and Analysis System

EDTA - Solução Ethylenediamine Tetraacetic Acid

*Eef1* - *Eucalyptus* Early Flowering Gene 1

EMBRAPA - Empresa Brasileira De Pesquisa Agropecuária

EST - Expressed Sequence Tag

EUCAGEN - *Eucalyptus* Genome Network

FAO - Food and Agriculture Organization of the United Nations

FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo

FDR - False Discovery Ratio

GENOLYPTUS - Rede Brasileira de Pesquisa do Genoma de *Eucalyptus*

IBDF - Instituto Brasileiro de Desenvolvimento Florestal

IDH-M - Índice de Desenvolvimento Humano do Município

JGI - Joint Genome Institute

Kg, g, mg, ng - Unidades de medida de massa: quilograma, grama, miligrama, nanograma

L, mL,  $\mu$ L - Unidades de medida de volume líquido: litro, mililitro, microlitro

LigTotal - Teor de Lignina Total

LOD (LOD-score) - log of the ODDS (logarítmo na base 10 da razão das probabilidades)

m, cm, mm - Unidades de medida de distância: metro, centímetro, milímetro

$m^3$ ,  $cm^3$  - Unidades de medida de volume sólido: metro cúbico, centímetro cúbico

MG - Estado de Minas Gerais

MIS - Mapeamento de QTL por Intervalo Simples

MIC - Mapeamento de QTL por Intervalo Composto

MIM - Mapeamento de QTL por Intervalo Múltiplo

mM - Unidade de medida de concentração: micromolar

NIRS - Near Infra-Red Spectroscopy

PA - Estado do Pará

PCR - Polimerase Chain Reaction

*Ppr1* - *Puccinia psidii* Resistance Gene 1

PRODEPEF - Programa de Pesquisa e Desenvolvimento Florestal

QMres - Quadrado Médio do Resíduo

QTL - Quantitative Trait Loci

$R^2$  - Coeficiente de Determinação

RAPD - Random Amplification of Polymorphic DNA

RD - Rendimento Depurado de Celulose

REMADE - Revista da Madeira

RFLP - Restriction Fragment Length Polymorphism

RS - Estado do Rio de Janeiro

rSG - Relação Siringil/Guaiacil da Lignina

$r_{xy}$  - Coeficiente de Correlação

SAM - Seleção Assistida por Marcadores Moleculares

SBS - Sociedade Brasileira de Silvicultura

SP - Estado de São Paulo

SSR - Simple Sequence Repeats

TBE - Solução Tris-Ácido Bórico e EDTA

## LISTA DE FIGURAS

<b>Figura 1.</b> Diagrama ilustrando a distribuição das amostras nas etapas de trabalho: 1- <i>screening</i> de <i>primers</i> ; 2- testes <i>multiplex</i> ; 3- avaliação da progênie DG x GL2. PB: amostra em branco sem DNA .....	52
<b>Figura 2.</b> Contribuição relativa das fontes de variação na expressão fenotípica dos caracteres de crescimento avaliados nos dois ambientes .....	60
<b>Figura 3.</b> Contribuição relativa das fontes de variação na expressão fenotípica do caráter pilodyn no experimento do RS .....	61
<b>Figura 4.</b> Contribuição relativa das fontes de variação na expressão fenotípica dos caracteres avaliados por NIRS nos dois ambientes .....	62
<b>Figura 5.</b> Mapa genético integrado constituído por 54 marcadores SSR genotipados em uma progênie de 200 irmãos completos de <i>E. globulus</i> x híbrido ( <i>E. dunnii</i> x <i>E. grandis</i> ). Os locos destacados em vermelho foram alocados como marcas acessórias, tendo suas posições aproximadas estimadas com base em informações disponíveis em outros mapas genéticos de <i>Eucalyptus</i> que mapearam estes locos anteriormente .....	68
<b>Figura 6.</b> Identificação dos QTL de acordo com a análise realizada .....	70
<b>Figura 7.</b> Mapa representativo dos QTL detectados para o caráter diâmetro - DAP .....	71
<b>Figura 8.</b> Mapa representativo dos QTL detectados para o caráter altura .....	72
<b>Figura 9.</b> Mapa representativo dos QTL detectados para o caráter volume .....	74
<b>Figura 10.</b> Mapa representativo dos QTL detectados para o caráter densidade básica - DB .....	76
<b>Figura 11.</b> Mapa representativo dos QTL detectados para o caráter teor de lignina - LigTotal .....	77
<b>Figura 12.</b> Mapa representativo dos QTL detectados para o caráter relação siringil/guaiacil - rSG .....	79
<b>Figura 13.</b> Mapa representativo dos QTL detectados para o caráter rendimento depurado de celulose - RD .....	81

- Figura 14.** Mapa representativo de todos os locos onde foram identificados QTL para os caracteres avaliados. Os locos foram identificados com cores de acordo com o caráter associado ao QTL. Os locos que apresentaram QTL para mais de um caráter foram representados com a cor laranja ..... 82
- Figura 15.** Efeito diferencial entre os genótipos dos locos que apresentaram interação entre QTL e ambientes. A, B e C: Efeito diferencial entre as classes genotípicas do loco Embra 0191, que apresentou interação entre QTL e ambientes para o caráter teor de lignina. D, E e F: Efeito diferencial entre as classes genotípicas do loco Embra 1868, que apresentou interação entre QTL e ambientes para o caráter rendimento de celulose ..... 101
- Figura 16.** Fluxograma indicando onde a SAM pode ser aplicada a partir do mapeamento de QTL utilizando marcadores moleculares em programas de melhoramento genético de espécies florestais ..... 116

## LISTA DE TABELAS

<b>Tabela 1.</b>	Modelos utilizados para se estimar os valores eBLUP nos experimentos .....	57
<b>Tabela 2.</b>	Parâmetros experimentais e componentes de variância estimados para os caracteres avaliados no experimento do RS .....	58
<b>Tabela 3.</b>	Parâmetros experimentais e componentes de variância estimados para os caracteres avaliados no experimento de MG .....	58
<b>Tabela 4.</b>	Coefficientes de correlação significativos para valores fenotípicos (quadrante superior) e genotípicos (quadrante inferior) entre os caracteres avaliados .....	63
<b>Tabela 5.</b>	P-valores e coeficientes de determinação dos locos que apresentaram associações a QTL para o caráter diâmetro na altura do peito (DAP) em pelo menos uma das análises realizadas .....	70
<b>Tabela 6.</b>	P-valores e coeficientes de determinação para os locos que apresentaram associações a QTL para o caráter altura em pelo menos uma das análises realizadas .....	73
<b>Tabela 7.</b>	P-valores e coeficientes de determinação dos locos que apresentaram associações a QTL para o caráter volume em pelo menos uma das análises realizadas .....	75
<b>Tabela 8.</b>	P-valores e coeficientes de determinação para os locos que apresentaram associações a QTL para o caráter densidade básica (DB) em pelo menos uma das análises realizadas .....	75
<b>Tabela 9.</b>	P-valores e coeficientes de determinação para os locos que apresentaram associações a QTL para o caráter teor de lignina (LigTotal) em pelo menos uma das análises realizadas .....	78
<b>Tabela 10.</b>	P-valores e coeficientes de determinação para os locos que apresentaram associações a QTL para a relação sirigil/guaiacil (rSG) em pelo menos uma das análises realizadas .....	79
<b>Tabela 11.</b>	P-valores e coeficientes de determinação para os locos que apresentaram associações a QTL para rendimento depurado de celulose (RD) em pelo menos uma das análises realizadas .....	80

<b>Tabela 12.</b> Simulação de seleção e comparação da eficiência da SAM para os caracteres que apresentaram QTL significativos na análise isolada de cada ambiente (A e B) e conjunta das amostras coletadas nos dois experimentos (C) .....	84
<b>Tabela 13.</b> QTL descritos para caracteres de crescimento nos trabalhos citados na revisão bibliográfica .....	106
<b>Tabela 14.</b> QTL descritos para o caráter densidade nos trabalhos citados na revisão bibliográfica .....	108
<b>Tabela 15.</b> QTL descritos para caracteres relacionados com as propriedades da lignina nos trabalhos citados na revisão bibliográfica .....	110
<b>Tabela 16.</b> QTL descritos para o caráter rendimento de celulose nos trabalhos citados na revisão bibliográfica .....	112



## RESUMO

MEDINA-MACEDO, L. **Mapeamento de QTL para características de interesse industrial da madeira de *Eucalyptus* em progênie de híbridos interespecíficos.** 2009. 143 f. Dissertação (Mestrado em Agronomia: Genética e Melhoramento de Plantas) – Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Goiânia, 2009.<sup>1</sup>

Neste trabalho foi construído um mapa genético e realizado o mapeamento de QTL para caracteres de crescimento e qualidade da madeira em uma progênie de híbridos interespecíficos de *Eucalyptus*. Foram isoladas amostras de DNA genômico de 200 indivíduos pertencentes a uma progênie de híbridos interespecíficos. A partir de 300 pares de *primers* microssatélites (SSR), foram selecionados com base na presença de polimorfismo válido para o mapeamento de QTL e na qualidade de amplificação de seus alelos, 76 locos para compor 19 combinações *tetraplex*. Dentre os 76 locos avaliados em sistema multiplex, 14 apresentaram distorção de segregação significativa no teste  $\chi^2$  perante o “Critério de Bonferroni” com nível de significância de 5% e foram excluídos das análises de mapeamento. Utilizando LOD 3 e fração de recombinação de 0,40, foram identificados 12 grupos de ligação compatíveis com os descritos em outros trabalhos de mapeamento em *Eucalyptus*. Foi obtido um mapa genético integrado com comprimento total de 1240 cM e distância média entre marcas de 22 cM, que proporciona uma cobertura genômica de aproximadamente 75% em relação ao tamanho do genoma deste gênero. Com exceção da densidade avaliada por *pilodyn*, todos os outros caracteres avaliados apresentaram associação a QTL por análise de variância (análise de marcas simples). Para os caracteres de crescimento, quatro locos estiveram associados a sete QTL, sendo que o montante da variação genotípica que estes QTL foram capazes de explicar foi de 19% para diâmetro, 18% para altura e 10% para volume. Dentre os caracteres de qualidade da madeira avaliados por NIRS, três QTL foram capazes de explicar 17% da variância genotípica para densidade, enquanto para o rendimento de celulose quatro QTL foram capazes de explicar juntos 7,5% desta variância. Para os caracteres relacionados com as propriedades da lignina também avaliados por NIRS, foram identificados três QTL capazes de explicar 9% da variação genotípica do caráter teor de lignina, e dois QTL capazes de explicar 6% da variação genotípica para a relação siringil/guaiacil. A maioria dos QTL

---

<sup>1</sup>Orientador: Prof. Dr. Alexandre Siqueira Guedes Coelho. EA-UFG.

identificados para os diferentes caracteres avaliados neste trabalho coincidiram com os grupos de ligação e o montante da variância representada com QTL identificados nos trabalhos de outros autores que realizaram mapeamento de QTL para os mesmos caracteres avaliados. Embora comparativamente a seleção assistida por marcadores moleculares tenha apresentado eficiência inferior em relação à seleção fenotípica, ela pode ser muito útil na otimização de futuros experimentos de campo de progênie para as quais já foi realizado o mapeamento de QTL, onde somente os genótipos com desempenho superior poderão ser avaliados em campo ao invés de toda a progênie. A aplicação de SAM em programas de melhoramento genético florestal é possível para a seleção precoce a partir da segunda geração de melhoramento, sempre que forem utilizadas informações de QTL mapeados nos cruzamentos realizados por estes programas de melhoramento. No entanto, estes QTL que serão utilizados para SAM, devem ser robustos nas progênie onde eles foram mapeados, tal como os QTL identificados nesta dissertação no grupo de ligação 9 para o caráter DAP, que além da consistência dos dados da progênie DG x GL2, foram capaz de explicar o mesmo montante da variabilidade com QTL mapeados para este caráter (12%) por Kirst et al. (2007), que utilizaram diferentes marcadores genéticos e técnicas de mapeamento de QTL. Estes dados comprovam que a aplicação da SAM é viável em programas de melhoramento, e que apesar das peculiaridades únicas de cada cruzamento, determinados grupos de ligação apresentam estabilidade de QTL entre diferente progênie para um mesmo caráter.

*Palavras-chave:* *Eucalyptus*, mapeamento genético, marcadores microsatélite, mapeamento de QTL, SAM.

## ABSTRACT

MEDINA-MACEDO, L. **Genetic mapping of QTL for characteristics of industrial interest of the *Eucalyptus*'s wood in a interspecific hybrid progeny.** 2009. 143 f. Dissertation (Master in Agronomy: Genetic and Plant Breeding) – Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Goiânia, 2009.<sup>1</sup>

In this work a genetic map and QTL mapping were built for growth and wood quality characters in an interspecific hybrids progeny of *Eucalyptus*. Samples of genomic DNA were isolated from 200 individuals belonging to a interspecific hybrids progeny. A screening of 300 pairs microsatellite loci (SSR) primers was done and from those 76 loci were selected to compose 19 *tetraplex* combinations, based in presence of valid polymorphism for QTL mapping and in quality of allele's amplification. Among the 76 loci appraised in *multiplex* system, 14 presented significant distortion of segregation in  $\chi^2$  test using "Criterion of Bonferroni" with 5% of significance, and they were excluded of mapping analyses. Using LOD 3 and recombination fraction of 0.40, 12 linkage groups compatible with described in other mapping works in *Eucalyptus* were identified. An integrated genetic map was obtained with total length of 1240 cM and medium distance among marks of 22 cM, It provides a covering genomic of approximately 75% in relation to genome size of *Eucalyptus* gender. A part of density appraised by pilodyn apparatus, all others characters showed QTL association through variance analysis (single marks analysis). For growth characters, four loci were associated to seven QTL and 19% for genotypic variation of diameter, 18% for height and 10% for volume these QTL were capable to explain. Among wood quality characters evaluated by NIRS, three QTL were capable to explain 17% of genotypic variance for density, while for yield cellulose, four QTL together were capable to explain 7,5% of this variance. For the characters related with lignin properties also evaluated by NIRS, were identified three QTL capable to explain 9% of genotypic variation for lignin quantity, and two QTL were capable to explain 6% of genotypic variation for siringil/guaiacil relationship. Most of identified QTL for different characters appraised in this work coincided with linkage groups and as well the amount of variance of QTL detected by other authors, which carried out QTL mapping for these same characters. Although comparatively the SAM showed inferior efficiency in relation to fenotypic selection, it can be useful for optimization of futures field experiments into

---

<sup>1</sup>Advisor: Prof. Dr. Alexandre Siqueira Guedes Coelho. EA-UFG.

progenies which QTL mapping already was accomplished, where only best genotypes can be established in field instead to whole progeny. The application of SAM in forest breeding programs is possible to be used in early selection starting from second breeding generation, using QTL information of crosses mapped in these breeding programs. However, QTL used for SAM should be robust in progenies where they were mapped, like as QTL identified in linkage group 9 for character DAP in this dissertation, that beside consistence data in progeny DG x GL2, were capable to explain the same amount of variability for QTL mapped by Kirst et al. (2007) (12%), which used different genetic markers and techniques for QTL mapping. These results prove that application of SAM is viable in breeding programs, and apart of peculiarities in each crossing, certain linkage groups are stable among different progenies for a same character.

*Key words:* *Eucalyptus*, genetic mapping, microsatellite markers, QTL mapping, SAM.

# 1 INTRODUÇÃO

O setor industrial de base florestal tem sido marcado por um processo de utilização crescente de madeiras provenientes de reflorestamento, colocando o Brasil em sintonia com a ordem mundial, que enfatiza a preservação das florestas naturais e incentiva a implantação de florestas renováveis (Remade, 2005). As indústrias de base florestal promovem e incentivam o desenvolvimento econômico e social em regiões distantes dos grandes centros urbanos onde estão instaladas, e comprovadamente contribuem para a elevação do IDH-M (Índice de Desenvolvimento Humano do Município) nos municípios em que estão presentes. Até o fim de 2007, estudos indicavam que estas indústrias foram responsáveis pela geração de cerca de 4,6 milhões de empregos, dentre empregos diretos, indiretos e resultantes do efeito de renda (Abraf, 2008).

Instituições estaduais de meio ambiente e entidades de classe do setor de base florestal estimam que a área reflorestada no Brasil esteja em torno de 6 milhões de hectares e deste montante, aproximadamente 3,7 milhões de ha são cultivados com *Eucalyptus*. Da área total reflorestada com *Eucalyptus*, cerca de 2 milhões de hectares são destinados à produção de papel e celulose, sendo que 1,4 milhões de ha destas florestas possuem certificação internacional (Abraf, 2008).

O Brasil é um dos três maiores fornecedores de celulose e papel para impressão, sendo um dos maiores exportadores de celulose de fibra curta e o primeiro em fibras produzidas a partir de *Eucalyptus* (Di Ciero & Amaral, 2002). No ano de 2007 estas indústrias produziram 11,9 milhões de toneladas de celulose, contribuindo com um saldo positivo para a balança comercial da ordem de 3,4 bilhões de dólares (Bracelpa, 2008). A elevada produtividade florestal aliada a custos de produção competitivos em relação a outros países do mundo, rendeu ao Brasil no ano de 2007 uma indicação pelo BID (Banco Interamericano de Desenvolvimento), como o país com maior atratividade da América Latina para investimentos florestais (Abraf, 2008).

Os reflorestamentos com *Eucalyptus* no Brasil possuem as mais altas taxas de produtividade do mundo, desempenho alcançado devido ao conjunto de fatores ambientais

favoráveis e elevados investimentos em pesquisa realizados pelas empresas durante os últimos quarenta anos, principalmente em programas de melhoramento genético (Santana et al., 2002). Até meados dos anos 90 era relativamente comum que empresas reflorestadoras obtivessem elevada produtividade devido a altos incrementos volumétricos, porém eram raros os programas de melhoramento genético que iam adiante, chegando a avaliar características mais refinadas (Foelkel, 1997). Com a difusão das técnicas moleculares, instituições de pesquisa e empresas do setor de papel e celulose realizaram um esforço coletivo para criar o Projeto Genolyptus – Rede Brasileira de Pesquisa do Genoma de *Eucalyptus*. Este projeto tem como objetivo identificar e caracterizar genes envolvidos na formação da madeira, para poder utilizar esta informação no melhoramento genético do *Eucalyptus* (Grattapaglia, 2003).

Um dos subprojetos do Projeto Genolyptus é a construção de mapas genéticos e o mapeamento de QTL em diferentes *backgrounds* genéticos, possibilitando assim identificar e validar regiões genômicas envolvidas no controle de caracteres de relevância econômica. Estas ações vêm sendo realizadas através da avaliação molecular de progênies obtidas do cruzamento controlado entre genótipos de diferentes espécies de *Eucalyptus*, avaliadas em experimentos de campo instalados em regiões com condições ambientais contrastantes (Grattapaglia, 2008). A validação de QTL em diferentes progênies em vários ambientes busca o direcionamento dos estudos para regiões genômicas específicas, que por sua vez podem ser caracterizadas com a utilização de técnicas de sequenciamento e análise de expressão gênica, permitindo que sejam identificados genes e regiões reguladoras envolvidas na expressão de caracteres de relevante interesse econômico.

Neste Sentido, o objetivo principal deste trabalho foi construir um mapa genético integrado de *Eucalyptus* utilizando marcadores SSR e identificar QTL associados ao controle genético de características de crescimento e propriedades da madeira de *Eucalyptus*. Posteriormente, foi identificado QTL comuns e específicos para cada um dos ambientes avaliados, bem como a existência de interação entre QTL e ambientes entre os genótipos dos caracteres avaliados nos dois ambientes. Os objetivos secundários consistiram em padronizar uma metodologia para coleta e extração de DNA genômico a partir do tecido do câmbio vascular, e analisar a eficiência das combinações multiplex geradas pelo programa MULTIPLEXER (Coelho, não publicado), utilizadas nas reações de PCR.



## 2 REVISÃO DE LITERATURA

### 2.1 O GÊNERO *Eucalyptus*

O gênero *Eucalyptus* é originário da Oceania, e com exceção das espécies *E. urophylla* e *E. deglupta* que são originários do Timor e Papua Nova Guiné, as demais são originárias da Austrália (Pryor, 1985). Este gênero pertence à família Myrtaceae e possui mais de 600 espécies arbóreas e arbustivas, que são subdivididas em 8 subgêneros. As espécies mais plantadas para fins comerciais pertencem ao subgênero *Symphyomyrtus*, que abrange mais de trezentas diferentes espécies, variedades e híbridos (Eldridge, 1994; Moura, 2004), a maioria delas descrita por S. T. Blake em 1934. De acordo com Grattapaglia & Bradshaw (1994), o tamanho do genoma deste gênero é relativamente pequeno, variando entre 500 e 650 Mpbs e apresentando 22 cromossomos ( $2n = 22$ ).

As espécies deste gênero possuem flores hermafroditas e a fecundação é preferencialmente cruzada (Pryor, 1976), embora possa ocorrer até 30% de autofecundação devido ao sistema reprodutivo misto (Eldridge, 1978). Na maioria das espécies, as flores são agrupadas em estruturas denominadas de corimbos, sendo os frutos capsulares e deiscentes, com sementes que variam em tamanho desde 1 mm até 2 cm de acordo com a espécie (Cavalcanti, 1963).

Pryor (1976) descreveu a existência de um sistema de protandria e também um sistema de autoincompatibilidade controlado geneticamente nas flores muito peculiar ao gênero, que varia em intensidade de acordo com a espécie, favorecendo a fecundação cruzada. Este autor cita que as populações naturais das espécies são compostas principalmente por indivíduos heterozigotos, sendo observada uma depressão no vigor das plantas quando ocorre a autofecundação. Os principais vetores de polinização são insetos, principalmente himenópteros, dípteros, lepidópteros, coleópteros e hemípteros (Pryor, 1985). Nas áreas de ocorrência natural, algumas aves e pequenos marsupiais também atuam como importantes polinizadores (House, 1997, citado por Trigueiro, 2006).

A evolução do gênero ocorreu a partir do período Terciário, e foi associada principalmente a variações climáticas bastante amplas e diferentes taxas de fertilidade do solo existentes na região de origem (Eldridge et al., 1994). Fósseis dos primeiros

*Eucalyptus* datam do Paleoceno, 65 milhões de anos atrás, quando a costa sul da Austrália ainda estava conectada à Antártida. O continente australiano moveu-se lentamente para o norte, formando uma barreira (planície Nullarbor), que isolou um grupo no sudoeste e outro no sudeste australiano pouco antes do Mioceno. As espécies da porção oriental se adaptaram aos verões úmidos e solos pobres em nutrientes, enquanto que as espécies ocidentais se adaptaram aos invernos chuvosos e solos férteis. Desta forma, o gênero ocupa ambientes altamente variáveis em termos de precipitação e de temperaturas, ocorrendo em áreas que variam desde pântanos até regiões semi-áridas, em solos de alta e baixa fertilidade (Williams & Woinarski, 1997). Devido à sua ampla dispersão, o gênero possui rápido crescimento e adaptabilidade natural a diversas áreas geográficas (Sartoretto et al., 2002).

O *Eucalyptus* possui grande importância econômica no mundo devido sua madeira possuir usos múltiplos, podendo ser empregada como fonte de energia renovável (carvão vegetal), matéria-prima para a produção de papel e óleos essenciais, além de sua utilização na fabricação de chapas e móveis de alta qualidade (Bansyopadhyay et al., 1999). Segundo levantamentos da FAO (2004), mais de noventa países possuem programas de reflorestamento com *Eucalyptus*, sendo que 58 destes o fazem em grande escala, utilizando sua madeira como matéria-prima básica para a maioria das suas necessidades. As espécies de *Eucalyptus* mais utilizadas para reflorestamentos em todo o mundo pertencem ao subgênero *Symphyomyrtus* e são: *E. grandis*, *E. saligna*, *E. urophylla*, *E. camaldulensis*, *E. tereticornis*, *E. globulus*, *E. viminalis*, *E. deglupta*, *E. citriodora*, *E. exserta*, *E. paniculata* e *E. robusta* (Remade, 2005).

## 2.2 O *Eucalyptus* NO BRASIL

No Brasil a pesquisa em Silvicultura teve início no começo do Século XX. Apesar de existirem registros de que mais de cinquenta espécies já haviam sido introduzidas antes de Navarro de Andrade em 1903, foi este Agrônomo quem iniciou os primeiros estudos realizados com bases técnicas e científicas que são utilizadas até os dias de hoje. De acordo com Andrade (1939), citado por Missiagia (2005), seu principal objetivo era produzir madeira com propósitos energéticos e estruturais, para suprir as necessidades da ex-Cia. Paulista de Estradas de Ferro.

A utilização do *Eucalyptus* na produção de papel e celulose no Brasil teve início a partir dos anos 40, e desde então este gênero se tornou a principal fonte de matéria-prima para as indústrias do setor (González et al., 2002). A partir de 1960, a pesquisa com *Eucalyptus* se consolidou e expandiu com a realização da “II Conferência Mundial do Eucalipto”, em São Paulo-SP. Nessa mesma época o Governo Federal instituiu o Programa de Incentivos Fiscais ao Reflorestamento e criou o Programa de Pesquisa e Desenvolvimento Florestal (PRODEPEF, ex-IBDF), que motivou a criação de várias cooperativas de pesquisa, dentre as quais podemos destacar o IPEF, o SIF e a FUPEF (Ferreira & Santos, 1997). Em 1967 a área plantada com *Eucalyptus* era de apenas 14 mil hectares, mas devido à política de incentivos fiscais, em 1973 a área cultivada já atingia cerca de um milhão de hectares (Sbs, 2005).

A partir de 1967, as plantações de *Eucalyptus* sofreram perdas sucessivas de produtividade devido à ocorrência do cancro *Cryphonectria cubensis* (Burner) e também da ferrugem causada por *Puccinia psidii* (Winter) no início dos anos 70 (Alfenas, 1989). Em 1975 também ocorreram fortes geadas nas regiões Sul e Sudeste do Brasil, que danificaram grandes extensões de florestas comerciais provocando grandes prejuízos econômicos (Higa, 1998). A sucessiva incidência de vários tipos de injúrias intensificou a seleção de procedências das espécies mais importantes, despertando a atenção para a possibilidade de explorar genótipos híbridos superiores com técnicas de propagação vegetativa.

Na década de 70 diversos projetos de melhoramento para várias espécies do gênero *Eucalyptus* foram desenvolvidos no país, realizando seleção de materiais genéticos adaptados a diversas condições edafo-climáticas, através de testes de procedências e progênies combinados para identificar genótipos superiores (Cruz & Carneiro, 2003). Embora a aplicação dos métodos convencionais de melhoramento na área florestal seja eficiente e tenha garantido ao Brasil os maiores índices de produtividade do mundo, os programas de melhoramento de espécies perenes se caracterizam por ser um processo relativamente lento, devido principalmente ao tempo necessário para se completar cada geração de cruzamentos e obter progênies segregantes (González et al., 2002).

Em 1979 foi estabelecida pela empresa Aracruz a primeira plantação clonal de *Eucalyptus* do Brasil, em uma área de 1.000 ha no município de Aracruz-ES (Ferreira & Santos, 1997). Os reflorestamentos clonais são homogêneos e possuem elevada produtividade, devido principalmente à clonagem permitir captar toda a variância

genotípica presente nos genótipos superiores, ao mesmo tempo em que contribui para minimizar a variância fenotípica dentro dos reflorestamentos (Vencovsky & Barriga, 1992).

Atualmente a maioria dos plantios realizados no Brasil utiliza clones (Sbs, 2005), e as principais espécies cultivadas no Brasil são: *Eucalyptus grandis*, *E. saligna*, *E. urophylla*, *E. viminalis* e híbridos interespecíficos (Remade, 2005). O período de 1986 a 1990 foi caracterizado pelo fim do subsídio para o setor florestal e abertura da economia ao mercado externo, iniciando uma nova fase onde as empresas necessitavam ter elevada eficiência em seu processo produtivo para poderem reduzir custos e se manterem competitivas.

O uso das técnicas de marcadores moleculares aliadas às modernas tecnologias de análise genômica têm proporcionado uma nova perspectiva em termos de avaliação da variabilidade genética para as mais diversas características, ampliando as estratégias disponíveis para aplicação em programas de melhoramento (Cooper, 2003). Um dos objetivos da genômica é o estudo das relações entre variações em características quantitativas e a diversidade molecular (Morgante & Salamini, 2003). Estes estudos podem fornecer o caminho para a próxima geração de pesquisa experimental no melhoramento, abrindo novas áreas antes inacessíveis pelo melhoramento convencional (Martienssem, 2003).

Compreender a arquitetura genética e o funcionamento dos locos envolvidos no processo de formação da madeira auxiliará na utilização da seleção assistida por marcadores (SAM) nos programas de melhoramento de espécies perenes (Cooper, 2003). O conhecimento e possibilidade de manipulação dos genes responsáveis pelo controle de caracteres de crescimento, síntese de celulose, densidade da madeira, teores de lignina, pentosanas entre outros, representam novos desafios para o melhoramento genético de *Eucalyptus* (Grattapaglia, 1999). Segundo este autor, estes caracteres possuem alto impacto no processo industrial por influenciarem diretamente na qualidade da matéria prima utilizada na produção de papel e celulose, reduzindo o consumo de insumos químicos utilizados no processo de beneficiamento da madeira.

## 2.3 PROJETOS GENOMA

Os avanços tecnológicos que ocorreram na década de 90 permitiram aos pesquisadores enfrentar o desafio de seqüenciar genomas inteiros, como o genoma humano e o genoma de algumas espécies agrícolas de relevante interesse econômico. Nos últimos dez anos um número crescente de projetos de seqüenciamento do genoma de espécies florestais de grande interesse econômico foi iniciado, incluindo *Pinus*, *Populus* e *Eucalyptus*. O genoma destes dois últimos gêneros possuem cerca de 600 Mpb, ou seja, cinco vezes o genoma de *Arabidopsis* e não muito mais do que o genoma de arroz. O genoma do gênero *Pinus* por ser cerca de sete vezes maior que o genoma humano, ainda representa uma tarefa difícil de ser alcançada com as tecnologias disponíveis (Grattapaglia, 2003).

O acesso ao genoma completo de várias espécies florestais será extremamente interessante à medida que for possível a condução de análises de mapeamento comparativo e busca de seqüências reguladoras para genes envolvidos na formação da madeira (Grattapaglia, 2003). Atualmente estas análises são muito limitadas ou mesmo impossíveis de serem realizadas com os genomas de *Arabidopsis* e arroz, devido ao fato de que estas espécies possuem crescimento secundário limitado, embora já tenham sido identificados e clonados alguns genes homólogos entre *Eucalyptus* e *Arabidopsis* envolvidos no controle de florescimento (Southerton et al., 1998; Missiaglia, 2005).

No Brasil a partir do Século XXI foram iniciados dois projetos com o objetivo de caracterizar o genoma de *Eucalyptus* – Genolyptus e Forests. Estes projetos utilizaram abordagens diferentes para atingir o mesmo objetivo, que é o melhoramento genético de *Eucalyptus*. O Projeto Forests contou com o apoio da FAPESP e mais quatro empresas com atuação predominante no Estado de São Paulo, enquanto o Projeto Genolyptus – Rede Brasileira de Pesquisa do Genoma de *Eucalyptus*, conta com a participação de empresas privadas do setor de papel e celulose distribuídas pelo país, sete universidades e a Embrapa (Grattapaglia, 2003). Este autor comenta que dificilmente as empresas de papel e celulose teriam individualmente capacidade tecnológica e financeira para sustentar um esforço concentrado na área de ciências genômicas, que se caracteriza pela novidade constante, elevada complexidade técnica, necessidade de equipes especializadas, sem previsão de retorno do investimento realizado em curto ou médio prazo. O Projeto Forests buscou avaliar o genoma expresso de *Eucalyptus* (EST), ou seja, trechos de DNA que comandam

as funções da planta, possibilitando a realização de estudos de estrutura, expressão e função gênicas.

O Projeto Genolyptus teve como objetivo central o descobrimento, seqüenciamento, mapeamento e determinação de função de genes de importância econômica de espécies do gênero *Eucalyptus*, através do mapeamento genético de progênies obtidas de cruzamentos controlados avaliadas em diferentes ambientes, simultaneamente ao seqüenciamento de DNA, EST, mapeamento físico e estudos de sintenia entre os genomas destas espécies (Grattapaglia, 2003). Segundo este autor, a estratégia de ação adotada pelo Projeto Genolyptus permitiu a investigação da relação complexa entre genótipo e fenótipo, integrando diversas tecnologias genômicas aos programas de melhoramento genéticos clássicos de características envolvidas no processo de formação da madeira de *Eucalyptus*.

Mais recentemente, o Departamento de Energia dos Estados Unidos (DOE) aprovou a proposta da rede internacional Eucagen (*Eucalyptus* Genome Network), liderada por três países, entre eles o Brasil, para o seqüenciamento completo do genoma do eucalipto. A espécie escolhida para seqüenciamento por este instituto é um clone de *E. grandis*, desenvolvido no programa de melhoramento genético por uma das empresas participantes do consórcio Genolyptus. Este projeto viabilizado em parceria com o Projeto Genolyptus teve início em 2008, e concorreu com 120 outros projetos submetidos por diversos países, que atenderam à chamada anual realizada pelo Joint Genome Institute (JGI), ligado ao DOE (US Department of Energy), que tem como objetivo o seqüenciamento de genomas de organismos que sirvam como fontes renováveis de energia.

## 2.4 MAPAS GENÉTICOS

O desenvolvimento de mapas genéticos é uma das aplicações de maior impacto da tecnologia de marcadores moleculares em programas de melhoramento de espécies vegetais, principalmente para as culturas perenes (Rocha et al., 2003). Os mapas genéticos possibilitam a cobertura e análise completa de genomas, a decomposição de características genéticas complexas nos seus componentes mendelianos, a localização e quantificação do efeito das regiões do DNA que controlam caracteres de importância econômica e, por fim, a canalização de toda esta informação para uso em programas de melhoramento (Ferreira & Grattapaglia, 1998).



O mapa genético de uma espécie pode ser definido com um arranjo linear de um grupo de genes ou marcadores adjacentes. Dois locos podem ser considerados ligados quando uma progênie apresentar genótipos recombinantes para estas marcas (Stutervant, 1913, citado por Rocha et al., 2003). Com exceção de alguns poucos organismos modelo, a esmagadora maioria das espécies não possuía qualquer informação de mapa até o fim dos anos 80 (Lee, 1995). Este cenário era decorrente da dificuldade em identificar características que tivessem pouca influência do ambiente na sua expressão, permitindo a utilização destas características como marcas para a construção de mapas genéticos (Ferreira & Grattapaglia, 1998). Embora o desenvolvimento da construção de mapas genéticos tenha suas bases desenvolvidas no início do século passado, historicamente os mapas de ligação foram pouco utilizados no auxílio à seleção em programas de melhoramento até o fim do Século XX.

Nos anos 80, Kary Mullis desenvolveu uma técnica denominada de Reação de Polimerização em Cadeia, conhecida como PCR – *Polymerase Chain Reaction* (Mullis & Faloona, 1987). Esta técnica proporcionou um grande avanço na análise genética utilizando marcadores moleculares, pois permitiu a análise do polimorfismo em regiões cromossômicas específicas, a partir de quantidades de DNA muito reduzidas. A difusão dos marcadores moleculares baseados em PCR nos anos 90 permitiu que um grande número de locos, das mais variadas espécies, pudesse ter seus alelos estudados em populações segregantes originadas a partir de cruzamentos controlados, facilitando o desenvolvimento de mapas genéticos (Ferreira & Grattapaglia, 1998). O nível de polimorfismo nos marcadores moleculares é bastante elevado quando comparado ao de marcadores morfológicos, além destes serem seletivamente neutros e em alguns casos codominantes (Ferreira, 2003).

O primeiro passo na construção de um mapa de ligação está relacionado com a escolha dos genitores a serem cruzados, de forma que maximize o polimorfismo genético e o desequilíbrio de ligação presente na progênie (Tanksley, 1993). Após selecionar os genitores para que a progênie obtida maximize os polimorfismos mapeáveis no cruzamento realizado, é necessário o desenvolvimento desta progênie, que deverá ser composta pelo maior número de possível de indivíduos. O desequilíbrio de ligação é atribuído à ligação física entre os locos segregantes, sendo detectado pela redução da frequência de recombinação entre os locos situados em regiões próximas ao longo de um determinado cromossomo (Coelho, 2000). O desequilíbrio de ligação decorrente da ligação física entre

os locos, não ocasionado pelos efeitos de seleção ou deriva genética, atinge sua maior extensão na primeira geração de populações derivadas de cruzamentos controlados, e como consequência, a capacidade de detectar a ligação é maximizada (Tanksley, 1993; Liu, 1998). Desta forma, nas gerações seguintes a extensão do desequilíbrio de ligação será cada vez mais reduzida, devido à recombinação que ocorre entre regiões cada vez mais próximas ao longo das gerações.

O conceito do desequilíbrio de ligação está relacionado com a associação presente entre as regiões genômicas nos pares de homólogos, e tem múltiplas implicações para a aplicação da seleção assistida por marcadores (SAM) nos programas de melhoramento genético. O desequilíbrio de ligação permite estimar a distância entre os marcadores utilizados e também na possibilidade de associar uma região genômica com um genótipo superior de um caráter relevante, apresentando implicações com relação à capacidade de poder utilizar a informação dos marcadores moleculares na SAM.

No reino vegetal a análise de ligação usando marcadores moleculares foi inicialmente empregada na construção de mapas de plantas autógamas, utilizando populações oriundas de gerações F<sub>2</sub>, provenientes na maioria das vezes do cruzamento de plantas diplóides, homozigotas e contrastantes (Malliepaard et al., 1997). Em espécies perenes e preferencialmente alógamas como as do gênero *Eucalyptus*, que apresentam alta depressão endogâmica e mecanismos de autoincompatibilidade, a obtenção de genitores em completa homozigose é inviável (Grattapaglia & Sederoff, 1994). A construção de mapas de ligação utilizando dados moleculares é invariavelmente realizada com dados vindos de algum tipo de delineamento genético, o qual permite esperar um padrão de segregação tipicamente mendeliano, de acordo com a natureza do marcador molecular utilizado e o polimorfismo dos genitores. A ausência desse padrão mendeliano é chamada de distorção de segregação, e os locos que apresentarem o fenômeno de distorção mendeliana devem ser inicialmente descartados da análise (Kao et al., 1999).

O critério mais utilizado para verificar o padrão de segregação mendeliano é o teste de aderência do  $\chi^2$  (Bearzoti, 2000). A avaliação da significância dos p-valores obtidos pela aplicação deste teste deve ser realizada utilizando a correção para múltiplos testes de Bonferroni (Sokal & Rohlf, 1995), ou o Critério FDR – False Discovery Ratio (Benjamini & Hochberg, 1995). Em ambos os casos o que se procura é controlar o aumento da probabilidade de se cometer um erro Tipo I, através da realização de múltiplas comparações. Isto significa que um novo ponto de corte será definido, evitando que alguns

locos com distorção de segregação aparentemente significativa em relação ao intervalo de confiança pré-estabelecido sejam descartados, por não levar em consideração a quantidade de testes realizados.

As marcas que se mostram adequadas ao modelo de segregação são submetidas à análise de ligação aos pares para compor os grupos de ligação. Teoricamente, marcadores com frequência de recombinação inferior a 0,5 podem estar localizados em um mesmo cromossomo, ou seja, no mesmo grupo de ligação. Depois de estabelecidos os grupos de ligação, a ordem dos marcadores deve ser determinada dentro de cada grupo com a respectiva distância entre eles, utilizando procedimentos de análise multiponto (Lopes, 2003).

A construção de mapas de ligação usando marcadores moleculares em espécies de fecundação cruzada é mais complicada do que em progênies derivadas de pais homozigotos, sendo realizada a partir de populações segregantes oriundas do cruzamento entre genitores heterozigotos (Grattapaglia & Sederoff, 1994). Segundo Malliepard et al. (1997), as principais diferenças entre as progênies derivadas de pais homozigotos e as derivadas de pais heterozigotos estão no número de alelos que podem segregar em cada um dos locos dos parentais e na fase de ligação entre estes locos. Na avaliação dos cruzamentos realizados entre plantas heterozigotas com marcadores codominantes, um único loco pode segregar para até quatro genótipos diferentes, além da fase de ligação entre os locos adjacentes serem desconhecidas. A caracterização das fases de ligação é realizada posteriormente, com base na probabilidade de distribuição dos alelos alternativos. Para os cromossomos homólogos de cada parental é calculada a fração de recombinação juntamente com a fase de ligação entre diferentes marcas moleculares (Wu et al., 2002).

Grattapaglia & Sederoff (1994) estabeleceram a estratégia de mapeamento conhecida como pseudo-cruzamento-teste, que basicamente é uma extensão da abordagem de mapeamento de haplóides utilizada em coníferas, onde as meioses de cada um dos pais são analisadas separadamente e são construídos mapas de ligação individuais para cada um dos parentais avaliados. O termo pseudo-cruzamento-teste foi originalmente introduzido para definir as estratégias de mapeamento que usavam marcadores dominantes, onde um dos pais é heterozigoto e o outro homozigoto para um loco qualquer, e a progênie segrega na proporção de 1:1 como em um cruzamento-teste. Quando os dois parentais envolvidos no cruzamento são heterozigotos, o termo duplo-pseudo-cruzamento-teste pode ser mais apropriado (Agrama et al., 2002).

Até o início dos anos 90 não era conhecido nas espécies perenes a frequência e a estrutura de seqüências microssatélites, genes, regiões promotoras e outros elementos nucleotídicos do genoma em geral. Estas informações são fundamentais para o empreendimento de estudos de experimentação genômica envolvendo mapeamento genético, físico e clonagem posicional de genes (Lourenço, 2004). Na segunda metade dos anos 90, o desenvolvimento de marcadores microssatélites amplamente distribuídos ao longo do genoma, possibilitou a saturação, o refinamento e a integração dos mapas genéticos construídos utilizando esta classe de marcador molecular (Moreau et al., 2000). Marcadores microssatélites são codominantes, altamente variáveis, multialélicos e possuem grande conteúdo informativo, e as regiões genômicas adjacentes aos motivos de repetição detectadas por este tipo de marcador, em geral são altamente conservadas entre espécies relacionadas. No gênero *Eucalyptus* esta classe de marcadores apresenta elevada conservação entre espécies de um mesmo subgênero, principalmente no subgênero *Symphyomyrtus*, embora a transferibilidade dos *primers* desenhados para amplificar estes locos entre outros subgêneros não tenha se apresentado tão notável (Brondani et al., 1998).

Em *Eucalyptus* a existência e utilização de locos conservados entre espécies é essencial, visto que os programas de melhoramento normalmente utilizam várias espécies e híbridos interespecíficos (Byrne et al., 1995). De acordo com Brondani et al. (1998), o desenvolvimento de *primers* microssatélites deve levar em consideração a maximização de sua utilização, ou seja, quanto maior a conservação de locos microssatélites menor será seu custo fixo relativo. Estes autores relatam que apesar do grande esforço envolvido no desenvolvimento deste tipo de marcador, nem todos os locos microssatélites disponíveis podem ser aproveitados para o mapeamento genético, pois o grau de informação dos locos depende do polimorfismo existente nos alelos dos parentais utilizados para se obter a progênie de cruzamento controlado.

## 2.5 MAPEAMENTO DE QTL

A maioria dos caracteres economicamente relevantes herdáveis são resultantes da ação conjunta de vários genes, apresentando uma variação contínua ao invés de classes fenotípicas discretas. Um dos problemas que geneticistas e melhoristas têm enfrentado é identificar os locos responsáveis pelo controle destas características de herança complexa, pois existem poucas informações sobre o número, posição cromossômica, magnitude do

efeito e interações dos locos que controlam sua expressão (Grattapaglia, 1999). Estas características são denominadas de poligênicas, quantitativas ou de herança complexa e os locos responsáveis por seu controle são denominados QTL – *Quantitative Trait Loci* (Ferreira & Grattapaglia, 1998).

A maior limitação dos experimentos de mapeamento de QTL em árvores tem sido a dificuldade de construção planejada de genótipos heterozigóticos portadores de alelos contrastantes para maximizar a segregação fenotípica, pressuposto básico para a detecção e localização mais eficiente de QTL (Grattapaglia, 2003). Em parentais heterozigotos, os únicos QTL que podem ser potencialmente detectados são aqueles que se encontram em heterozigose em pelo menos um dos genitores do cruzamento, sendo necessário que o efeito diferencial entre alelos alternativos do loco que apresenta associação significativa a QTL seja grande o suficiente para permitir sua detecção estatística (Missiaggia, 2005).

Com a construção de mapas de ligação, diversos trabalhos que serão descritos a seguir no item 3.6 deste trabalho, relataram a identificação de regiões genômicas que têm efeito significativo na expressão de características economicamente importantes em *Eucalyptus*. Estes trabalhos identificaram regiões genômicas associadas a caracteres de interesse, tais como resistência a pragas e doenças, tolerância ao frio, crescimento, qualidade da madeira, propagação vegetativa e teor de lignina. No início os trabalhos de mapeamento genético em *Eucalyptus* foram realizados utilizando marcadores dominantes (RAPD e AFLP), o que dificultava o compartilhamento interexperimental dos dados, tornando o uso das informações de QTL restritas à progênie avaliada. A utilização de marcadores dominantes também não permitia uma detecção mais fina do efeito dos diferentes alelos nos QTL em cruzamentos de indivíduos heterozigóticos, como é o caso das matrizes utilizadas nos programas de melhoramento do gênero *Eucalyptus*. Posteriormente com a difusão dos marcadores moleculares codominantes, foi possível integrar os mapas genéticos construídos e validar QTL identificados em diferentes trabalhos.

É importante salientar que a dificuldade em se implementar a seleção assistida por marcadores (SAM) nos programas de melhoramento de espécies alógamas se deve à perda de associação entre as marcas moleculares e as características de interesse ao longo das gerações de melhoramento, ocasionada pela redução do desequilíbrio de ligação. Atualmente estão disponíveis diversos métodos para localizar regiões QTL. Estes métodos

são classificados de acordo com o número de marcas utilizadas para se detectar uma região de QTL, que pode ser realizada utilizando uma, duas ou mais marcas simultaneamente (Bearzoti, 2000). Estas metodologias serão descritas a seguir.

### 2.5.1 Mapeamento de QTL por Análise de Marcas Simples - AMS

A análise de marcas simples consiste em associar um marcador ao QTL, pressupondo a existência de diferenças nos valores médios do caráter quantitativo entre os indivíduos com diferentes genótipos no loco marcador. Na análise de marcas simples a distribuição dos valores do caráter quantitativo é examinada separadamente para cada *loco* marcador, independente da informação dos outros locos (Cruz & Pessoni, 2006).

A comparação entre as médias das classes genóticas dos marcadores pode ser realizada pelo Teste t, Regressão Linear Simples, Análise de Variância e Razão de Verossimilhança (*LOD-score*) (Silva, 2001). No mapeamento de QTL, é muito comum o uso do *LOD-score* (*log of the ODDS*), que emprega o logaritmo na base 10 da probabilidade de ocorrência dos dados em relação à hipótese de nulidade. A interpretação do *LOD-score* é baseada no conceito de *odds of ratio* (razão das chances), não sendo necessária nenhuma distribuição teórica para sua interpretação, uma vez que um *LOD-score* igual a 3 indica que a hipótese alternativa é  $10^3 = 1000$  vezes mais provável que a hipótese nula (Coelho & Silva, 2002).

Segundo Silva (2001), diferenças significativas entre as médias dos caracteres quantitativos agrupados nas classes genóticas dos marcadores indicam que pelo menos um QTL pode estar associado à marca. As diferenças observadas entre as médias das classes genóticas resultam tanto dos efeitos diferenciais dos genótipos do QTL quanto da proximidade entre este loco marcador o polimorfismo associado ao QTL (Falconer & Mackay, 1996).

De acordo com Liu (1998), a principal vantagem da análise de marcas simples está em não ser necessário o conhecimento da posição dos marcadores no genoma, mas há algumas desvantagens devido à posição do QTL não poder ser precisamente determinada. Isto ocorre devido a não independência entre os testes de hipótese para ligação, o que confunde o efeito do QTL e sua posição nos procedimentos de análise baseados na comparação de médias, não permitindo diferenciar claramente um QTL de efeito pequeno próximo ao marcador ou de um QTL de efeito grande, porém longe do marcador. Esse tipo de análise é geralmente uma boa alternativa quando o objetivo é a simples detecção de



QTL ligado à marca em detrimento da estimação da posição e do efeito do mesmo (Coelho & Silva, 2002). De acordo com estes autores, apesar das limitações a análise por marcas simples é muito utilizada, principalmente como um estudo preliminar para posterior uso de técnicas mais complexas de mapeamento por intervalo.

### **2.5.2 Mapeamento de QTL por Intervalo Simples - MIS**

O mapeamento por intervalo foi proposto por Lander & Botstein (1989), e é uma extensão do método de máxima verossimilhança para marcas simples. O princípio desse método consiste em avaliar a possibilidade de ocorrência do loco quantitativo em uma série de posições dentro de um intervalo definido por duas marcas adjacentes, através da análise das frequências conjuntas ao par de marcas e o QTL flanqueado por estas duas marcas (Coelho, 2000).

As análises são realizadas considerando cada par de marcas adjacentes, resultando em  $t-1$  testes de associação entre QTL e marcadores em um cromossomo com  $t$  marcas, já que se realiza um teste para cada intervalo (Coelho & Silva, 2002). De acordo com estes autores esta metodologia de mapeamento oferece um incremento no poder de detecção de QTL e estimativas mais precisas da posição e dos efeitos do mesmo. Para utilização desta estratégia é necessário conhecer as frequências de recombinação entre os dois marcadores antes da análise de associação com o QTL, ou seja, é necessário previamente construir um mapa de ligação (Cruz & Pessoni, 2006). Assim, como a distância entre cada par de marcadores é conhecida, o método estuda o intervalo que delimita uma região cromossômica usando a função de mapeamento (Coelho, 2000), aumentando a robustez da análise por utilizar a informação completa da distribuição da associação entre os marcadores e a característica mapeada (Cruz & Pessoni, 2006).

Os valores estimados através da derivação da função de verossimilhança ou pela análise de regressão, são utilizados para cálculos na forma de *LOD-score* para cada ponto de análise ao longo do genoma. Uma alternativa ao uso da máxima verossimilhança é o mapeamento por intervalo baseado na utilização de regressão linear múltipla, conforme proposto por Haley & Knott (1992). Este modelo assume que um QTL está localizado entre duas marcas, e que a variância dentro das classes genóticas do QTL é a mesma, e apresenta distribuição normal, sendo que cada possível genótipo para o QTL possui uma probabilidade de ocorrência em função dos genótipos dos marcadores flanqueadores (Schuster & Cruz, 2004). O mapeamento por intervalo é indiscutivelmente mais eficiente

na localização de QTL do que o método por marcas simples, porém existe a possibilidade de localização de “QTL Fantasmas”, ou seja, é possível identificar um QTL inexistente devido à ocorrência de ligação entre dois QTL adjacentes (Silva, 2001).

### **2.5.3 Mapeamento de QTL por Intervalo Composto - MIC**

No ano de 1993, dois diferentes pesquisadores (Jansen, 1993; Zeng, 1993), propuseram de maneira independente que os QTL fora do intervalo em questão fossem levados em conta através de um modelo de regressão múltipla, eliminando assim o efeito dos “QTL fantasmas” (Silva, 2001). A denominação proposta por Zeng (1993) foi mapeamento por intervalo composto, enquanto a proposta de Jansen foi de mapeamento múltiplo de QTL, embora computacionalmente o produto resultante seja equivalente ao intervalo composto.

De acordo com Coelho (comunicação pessoal), os modelos utilizados para mapear QTL possuem maior influência na redução da ocorrência de “QTL fantasmas” do que o aumento da densidade do mapa. Segundo esta fonte, o mapeamento por intervalo composto pode ser realizado a partir da estimação da razão de verossimilhança, empregando metodologia publicada preconizada por Zeng (1993), ou utilizando recursos computacionais que possibilitem a realização de reamostragens *bootstraps* mais eficientes, empregando a metodologia descrita por Haley & Knott (1992).

O método de mapeamento por intervalo composto combina o mapeamento por intervalo com a regressão linear múltipla, com o objetivo de eliminar a interferência de QTL adjacentes ao intervalo que está sendo mapeado (Carneiro & Vieira, 2002). Com base no método de mapeamento por intervalo composto proposto por Jiang & Zeng (1995), foi desenvolvido um método para mapeamento de QTL para mais de uma característica simultaneamente. Considerando a existência de correlação entre múltiplos caracteres (devido à ligação ou à pleiotropia), estes autores afirmam haver um aumento no poder de detecção de QTL e na precisão do mapeamento, quando comparadas às análises das características isoladamente, possibilitando o estudo de caracteres biológicos mais abrangentes (Bearzoti, 2000).

#### 2.5.4 Mapeamento de QTL por Intervalo Múltiplo – MIM

O método de mapeamento por intervalo múltiplo (MIM) foi proposto por Kao et al. (1999) e consiste em utilizar os conceitos de Cocherham (1954) para interpretar os parâmetros genéticos estimados pelo método de máxima verossimilhança. Desta maneira, múltiplos intervalos são utilizados para mapear múltiplos QTLs simultaneamente para um mesmo caráter (Coelho & Silva, 2002). A metodologia de mapeamento por intervalo múltiplo pode levar a um aumento na precisão e no poder de detecção de QTL, além de permitir estimativas e análise dos efeitos de interações entre QTL (epistasia).

Um problema encontrado nessa técnica de mapeamento é que o número de distribuições envolvidas na verossimilhança e de parâmetros a serem estimados aumenta de maneira drástica com o aumento do número de QTLs considerados (Silva, 2001). Desta maneira, dois fatores devem ser avaliados para utilizar o método. O primeiro diz respeito à estimação de parâmetros, onde é necessário um número limitado de marcas, pois à medida que elas aumentam, a derivação dos estimadores de verossimilhança para os efeitos dos QTL torna-se impraticável. O segundo fator diz respeito ao procedimento de mapeamento utilizado para localizar o QTL, de acordo com o modelo avaliado.

Kao et al. (1999) propõem o uso do procedimento *stepwise*, porém se tratando de análises estatísticas, estes procedimento estão sujeitos aos erros do tipo I e II, ou seja, assumir a existência de QTL inexistente ou eliminar um QTL válido da análise. Desta forma, uma das vantagens do método é que os efeitos dos QTL são estimados sem viés na presença de epistasia, podendo aumentar a eficiência da seleção assistida por marcadores (Matta, 2005).

De acordo com reflexão realizada por Coelho & Silva (2002), são necessários estudos mais aprofundados sobre as metodologias de determinação das marcas e dos QTL que serão inclusos no procedimento de MIM. O ideal seria balancear o poder de detecção de QTL com o número de falsos QTLs detectados, para não causar grandes perdas na detecção de QTL pelo procedimento MIM, e também não tornar este método de mapeamento computacionalmente inviável.

#### 2.5.5 Mapeamento associativo

Embora o mapeamento de associação utilize a mesma metodologia empregada na AMS, a principal diferença entre o mapeamento de associação e o de ligação se

encontra no tipo de população utilizada para a realização do mapeamento de QTL. O mapeamento baseado no desequilíbrio de ligação necessariamente requer a avaliação de uma progênie originada a partir de um cruzamento controlado, enquanto que o mapeamento associativo pode ser realizado utilizando populações naturais ou indivíduos não relacionados, desde que estes indivíduos apresentem fenótipos contrastantes para algum caráter de interesse (Breseghello & Sorrells, 2006).

Desta forma, o mapeamento de associação se refere à associação significativa de um marcador molecular com uma característica fenotípica. Com o desenvolvimento de novas plataformas de genotipagem de polimorfismos em larga escala, tais como dados de *microarranjos* de sondas de DNA, pirosequenciamento ou DarT, este método de mapeamento vem sendo cada vez mais utilizado para o mapeamento de QTL empregando a abordagem de genes candidatos (Myburg et al., 2007). A abordagem de genes candidatos vem sendo cada vez mais difundida devido à constante redução dos custos de sequenciamento, e também ao aumento das informações dos bancos de dados de públicos de sequências, o que possibilita a comparação de sequências e funções gênicas para as espécies cujos genomas já foram ou desejem ser estudados.

## 2.6 MAPAS GENÉTICOS E MAPEAMENTO DE QTL EM *Eucalyptus*

### 2.6.1 Caracteres quantitativos

Os alvos do melhoramento florestal até a década de 1990 eram basicamente os caracteres silviculturais de crescimento, forma do fuste, volume de madeira, e também os caracteres de resistência a fatores bióticos (doenças e pragas) e abióticos (climas e tipos de solos adversos). Em meados da década de 90, novos métodos não destrutivos e o avanço das técnicas moleculares possibilitaram a análise dos caracteres relacionados à qualidade da madeira em uma grande quantidade de árvores a um custo relativamente baixo, permitindo que esses caracteres fossem incorporados aos programas de melhoramento (Raymond, 2002).

Há alguns anos as indústrias reflorestadoras têm aplicado técnicas alternativas na investigação de algumas propriedades da madeira. A espectroscopia no infravermelho próximo (Near Infrared Spectroscopy - NIRS) é uma técnica acurada, rápida, não-destrutiva, e aplicável em um grande número de amostras cujas moléculas contenham ligações C-H, O-H, N-H, S-H ou C=O. A técnica NIRS é fundamentada na absorção

seletiva da luz pelos compostos químicos presentes nos materiais analisados, medindo a interação da luz com o material pela vibração das ligações químicas dos constituintes da amostra (Hein, 2008).

O método é considerado como não-destrutivo e não-invasivo por não exigir preparação prévia da amostra. Contudo, para algumas aplicações com madeiras é necessário a produção de corpos-de-prova para que seja realizada a calibração do equipamento, sem a qual as estimativas geradas para os parâmetros avaliados podem ser imprecisas. O conhecimento da composição química da madeira e seus componentes principais, como celulose, lignina e todo material extraído pelos solventes, é importante para interpretar o comportamento da madeira em processos industriais. Para produção de papel e celulose, madeiras com baixo conteúdo de lignina e alto conteúdo de carboidratos requerem condições menos severas de cozimento, conduzindo a um maior rendimento e eficiência na produção da polpa celulósica. Os outros componentes e extrativos têm também um papel chave na deslignificação e produção. Para a produção de polpa KRAFT, madeiras com menor conteúdo em lignina e solubilidade em álcali abaixo de 1% promovem fácil deslignificação e maior rendimento (Valente et al., 1992).

O teor de lignina é o maior responsável pela carga de álcali necessária para que as fibras atinjam o grau de deslignificação desejado pela indústria de celulose. A qualidade da polpa produzida é monitorada industrialmente, dentre outros fatores, pela solubilidade e estabilidade da lignina sob as condições da polpação, e o conhecimento da relação siringil/guaiacil também apresenta implicações diretas nos processos de polpação e branqueamento (Santos, 2005). Estruturalmente as ligninas são classificadas nos tipos siringil e guaiacil, sendo que a primeira é removida com maior facilidade na etapa inicial de cozimento por possuir menor peso molecular (Pimenta et al., 2007). Quanto maior a relação siringil/guaiacil, mais reativa será a lignina com os produtos utilizados no processo de polpação, consumindo menor carga de álcali. Neste sentido, é de interesse para a indústria de celulose que as árvores acumulem menor teor de lignina, e/ou produzam lignina de fácil remoção durante o processo de polpação (maior relação siringil/guaiacil).

Sabe-se que embora a densidade básica seja um caráter de fácil avaliação e apresente boa herdabilidade, esta é uma propriedade da madeira extremamente complexa por resultar da combinação de elementos anatômicos e compostos químicos presentes na madeira (Foelkel et al., 1974; Barrichelo et al., 1983; Shimoyama, 1990 e Valente et al., 1992). Segundo Foelkel et al. (1992), a densidade básica para *Eucalyptus* pode variar de

0,300 g.cm<sup>-3</sup> a 0,800 g.cm<sup>-3</sup>, porém para a produção de celulose recomenda-se o uso de densidades entre 0,450 g.cm<sup>-3</sup> e 0,550 g.cm<sup>-3</sup> (Wehr, 1991; Fonseca et al., 1996). O uso de madeiras muito leves pode reduzir o rendimento do processo, enquanto que o de madeiras muito densas podem dificultar a picagem das toras e a impregnação dos cavacos pelo licor branco, dificultando a deslignificação e aumentando os teores de rejeitos (Wehr, 1991). A densidade da madeira pode ser avaliada por testes mecânicos, NIRS ou por pilodyn, que é um aparelho que mede a penetração de uma haste de ferro na madeira e assim, espera-se que quanto maior a densidade básica de uma amostra menor seja a penetração da haste do pilodyn.

### 2.6.2 Mapas genéticos

Utilizando marcadores RAPD, Grattapaglia & Sederoff (1994) construíram os primeiros mapas genéticos para este gênero, avaliando a progênie de um cruzamento interespecífico com a estratégia de pseudo-cruzamento-teste, utilizando LOD-score 5 e fração de recombinação de 25% para construir os grupos de ligação destes parentais. Estes autores estimaram que o comprimento dos genomas de *E. grandis* e *E. urophylla* sejam respectivamente de 1620 cM e 1156 cM, e os mapas produzidos apresentaram cobertura genômica acima de 95%. Para o parental feminino de *E. grandis*, foram obtidos 14 grupos de ligação a partir de 142 locos, com uma distância média entre marcas de  $10.2 \pm 6.6$  cM. Para o parental masculino de *E. urophylla*, 119 marcas formaram os 11 grupos de ligação, proporcionando uma distância média entre marcas de  $12.2 \pm 6.3$  cM. Grattapaglia et al. (1995) utilizaram estes mapas para mapear QTL envolvidos no controle de características de enraizamento adventício e multiplicação de plântulas *in vitro*.

Verhaegen & Plomion (1996) também construíram mapas genéticos utilizando marcadores RAPD, avaliando a progênie de um cruzamento de *E. grandis* com *E. urophylla*. Os mapas produzidos por estes autores apresentaram 11 grupos de ligação para as duas espécies, cobrindo uma distância de 1331 cM para *E. urophylla* e 1415 cM para *E. grandis*, com uma distância média entre marcas de 14 cM. Ainda utilizando marcadores RAPD, Gan et al. (2003) construíram mapas genéticos avaliando 82 indivíduos de uma progênie obtida do cruzamento entre *E. urophylla* x *E. tereticornis*. Os mapas gerados por estes autores apresentaram 23 grupos de ligação para cada um dos dois parentais, com cobertura genômica de 95% para o parental feminino de *E. urophylla* e 69% para o parental masculino de *E. tereticornis*, cujos tamanhos dos genomas foram estimados

respectivamente em 1586 cM e 1508 cM. A distância média entre marcas obtidas nestes mapas foram de 11 cM e 10.1 cM respectivamente.

Utilizando marcadores AFLP dominantes, Marques et al. (1998) construíram mapas genéticos com cobertura genômica acima de 80%, avaliando uma progênie de 91 indivíduos obtida do cruzamento entre *E. globulus* com *E. tereticornis*. Para o parental feminino de *E. tereticornis*, 128 marcas foram utilizadas para compor 14 grupos de ligação, com distância média de 8 cM entre marcas e cobertura genômica de 919 cM. O mapa construído para o parental masculino de *E. globulus* utilizou 120 marcas para obter uma cobertura genômica de 967 cM, com distância média entre marcas de 9 cM. Marques et al. (1999) utilizaram estes mapas para mapear QTL envolvidos no controle de características relacionadas à propagação vegetativa. Marques et al. (2005), ainda utilizando marcadores AFLP, construíram três novos mapas genéticos, avaliando o mesmo parental de *E. globulus* em outros dois cruzamentos com *E. globulus* e *E. tereticornis*. Posteriormente, Marques et al. (2002) verificaram a sintenia dos QTL relacionados com características ligadas à propagação vegetativa identificados em seu trabalho, com os identificados no trabalho de Grattapaglia et al. (1995), permitindo validar QTL para características de propagação vegetativa em diferentes *backgrounds* genéticos, e também verificar homologias entre os grupos de ligação obtidos nestes mapas.

Embora os mapas construídos com marcadores dominantes fossem bem saturados e possuísem elevada cobertura genômica, somente uma pequena parte dos polimorfismos avaliados com este tipo de marcador eram informativos para ambos os parentais avaliados, dificultando a integração dos mapas obtidos para cada um dos parentais. Para aumentar a capacidade de integração dos mapas produzidos e validação de regiões QTL entre diferentes *backgrounds* genéticos, alguns autores fizeram uso de mais de um tipo de marcador molecular para a construção de seus mapas genéticos. Agrama et al. (2002) combinaram 73 marcadores RAPD, 81 RFLP e catorze SSR, para construir um mapa genético integrado em uma família de irmãos completos de *E. camaldulensis*, abrangendo uma distância de 1236 cM. Byrne et al. (1995) utilizaram uma combinação de quatro marcadores isoenzimáticos, 125 RAPD e 210 RFLP para construir um mapa genético integrado para uma progênie de *E. nitens*, onde cerca de 20% dos locos avaliados foram informativos para os dois parentais e o tamanho do genoma estimado foi de 1462 cM. Posteriormente, Byrne et al. (1997a; 1997b) utilizaram os mapas construídos em 1995

para identificar QTL envolvidos no controle da altura de plântulas, área foliar e tolerância ao frio.

Myburg et al. (2003) construíram mapas genéticos saturados com mais de oitocentos marcadores AFLP, onde foram avaliados os retrocruzamentos de um híbrido superior obtido do cruzamento de *E. globulus* e *E. grandis* com seus parentais, possibilitando o estudo da colinearidade entre o genoma dessas duas espécies. Os mapas produzidos neste experimento apresentaram uma cobertura genômica acima de 95%, e distância média entre marcas próxima a 10 cM. Thamarus et al. (2002) fizeram uso de cinco marcadores isoenzimáticos, 204 RFLP e quarenta SSR para construir um mapa genético consenso em uma progênie de *E. globulus*, onde foram identificados QTL envolvidos no controle de características da madeira, fibra e floração. Neste mapa foram obtidos 12 grupos de ligação, com comprimento final de 1375 cM e uma distância média de 6 cM entre marcas.

Esta dificuldade de integrar os mapas genéticos construídos fez com que os grupos de ligação identificados nos mapas genéticos construídos nos trabalhos de alguns autores não fossem necessariamente os mesmos grupos de ligação identificados no trabalho de Grattapaglia & Sederoff (1994). Com o avanço das técnicas moleculares e a difusão de marcadores codominantes, os grupos de ligação formados em diferentes trabalhos de mapeamento puderam ser correlacionados com maior precisão. Atualmente a maioria dos mapas genéticos construídos adota a nomenclatura de Grattapaglia & Sederoff (1994) e Brondani et al. (1998, 2002; 2006), devido a estes autores terem sido os pioneiros na construção de mapas genéticos para o gênero *Eucalyptus* e também na utilização de marcadores codominantes desenvolvidos especificamente para este gênero, o que permite relacionar os grupos de ligação obtidos no mapeamento de diferentes espécies de *Eucalyptus*.

Os primeiros vinte marcadores SSR desenvolvidos para o gênero *Eucalyptus* foram publicados por Brondani et al. (1998). Estes marcadores foram ancorados aos mapas genéticos construídos por Grattapaglia & Sederoff (1994) utilizando marcadores RAPD, e permitiram verificar a sintonia de seis grupos de ligação entre *E. grandis* e *E. urophylla*. Brondani et al. (2002) publicaram mais cinquenta novos marcadores SSR, que também foram ancorados nestes mapas genéticos construídos, e assim foi possível verificar a homologia entre os onze grupos de ligação destas espécies. Brondani et al. (2006) publicaram um mapa genético consenso para o gênero *Eucalyptus* utilizando 230 novos



primers SSR juntamente com os outros setenta descritos em seus trabalhos anteriores. Este mapa apresentou uma cobertura de 1568 cM, que representa aproximadamente 90% do tamanho do genoma estimado para este gênero, com uma distância entre marcas de 8.4 cM.

Os marcadores microssatélites publicados por Brondani et al. (1998; 2002; 2006) foram desenvolvidos com o apoio da EMBRAPA, e assim foram identificados com o acrônimo Embra. O CSIRO (Commonwealth Scientific and Industrial Research Organisation, Austrália) também apoiou nos anos 90 o desenvolvimento de marcadores SSR para algumas espécies de *Eucalyptus*, que foram identificados com os acrônimos EG, EN, e ES, que correspondem respectivamente, às espécies *E. grandis*, *E. nitens* e *E. saligna*. Muitos destes marcadores SSR foram apresentados em caráter inédito no trabalho de Byrne et al. (1996), que verificou a transferibilidade dos primers desenvolvidos pelo CSIRO em diferentes espécies de *Eucalyptus*. O trabalho destes autores permitiu que marcadores SSR segregantes e altamente polimórficos fossem selecionados por outros pesquisadores para compor seus mapas genéticos desenvolvidos com espécies e híbridos deste gênero.

Dentro do Projeto Genolyptus também foram desenvolvidas amplas baterias de marcadores SSR para *Eucalyptus*, com o objetivo de consolidar um mapa genético de referência, altamente saturado e composto exclusivamente por esta classe de marcadores moleculares, também identificados pelo acrônimo Embra. No trabalho desenvolvido por Lourenço (2004), foi estudado a complexidade e a abundância de regiões microssatélites no genoma de *Eucalyptus*, identificando cerca de mil novos SSR e mais de 1500 regiões com baixa complexidade. Este autor estimou em seu trabalho que cerca de 6% do genoma de *Eucalyptus* seja composto por regiões repetitivas.

Utilizando exclusivamente marcadores SSR, Missiaglia (2005) fez uso de 169 pares de *primers* para construir mapas genéticos individuais, e também um mapa genético integrado, para os parentais híbridos avaliados em seu trabalho. Este mapa genético integrado foi construído com informações de 96 indivíduos, avaliados com 155 marcadores SSR, que utilizando LOD 10 e fração de recombinação de 40%, foram agrupados em 11 grupos de ligação. O mapa genético integrado construído por este autor apresentou cobertura genômica superior à verificada nos mapas de cada parental individualmente (1033 cM contra 820 cM e 536 cM), mas a distância média entre marcas foi superior (8,08 cM contra 6,6 cM e 4,6 cM). Novaes (2006) também avaliou 276 locos SSR em uma progênie de híbridos interespecíficos composta por 188 indivíduos, obtidos do cruzamento

entre *E. grandis* e *E. urophylla*, no seu trabalho de mapeamento de QTL de interesse para a indústria de celulose. Este autor utilizando LOD 4 e fração de recombinação de 35% construiu um mapa genético integrado composto por 225 marcas, com comprimento total de 951 cM e distância média entre marcas de 4,2 cM.

### **2.6.3 Distorção de segregação em *Eucalyptus***

Lorieux et al. (1995) afirmam que a proporção e a extensão da distorção de segregação pode variar significativamente de acordo com o tipo de marcador molecular utilizado e também com a espécie e o tipo de população ou cruzamento avaliado. De acordo com Missiaglia (2005), a distorção de segregação é maior em cruzamentos interespecíficos do que em cruzamentos entre indivíduos da mesma espécie, e isto ocorre devido à alta taxa de formação de heteroduplex e pareamento errôneo entre os cromossomos homólogos. Liang et al. (2006) também citam que a distorção de segregação em plantas pode ser causada pelo efeito de seleção, transmissão diferencial de alelos ou viabilidade diferencial dos gametas.

Utilizando marcadores RAPD, o trabalho de Grattapaglia & Sederoff (1994) apresentou cerca de 5% dos locos avaliados com distorção de segregação, enquanto no trabalho de Verhaegen & Plomion (1996) este valor foi de 8%, 13% no trabalho de Rocha et al. (2007), e próximo a 25% no trabalho de Gan et al. (2002). Com marcadores AFLP, o percentual de locos que apresentaram distorção de segregação foi de 15% no trabalho de Marques et al. (1998), e próximo a 30% nos mapas genéticos produzidos no trabalho de Myburg et al. (2003). No trabalho de Byrne et al. (1995), onde foi utilizado várias classes de marcadores moleculares, o percentual de locos que apresentaram distorção de segregação foi de 6%.

Utilizando marcadores SSR, Missiaglia (2005) identificou mais de 40% dos locos com distorção de segregação nos parentais avaliados em seu trabalho de mapeamento de QTL para características da madeira de *Eucalyptus*, enquanto no seu trabalho de mapeamento do gene *Eef1*, responsável pelo florescimento precoce neste gênero, este índice foi de 6%. No trabalho de mapeamento de QTL para características de interesse para a indústria de celulose utilizando marcadores SSR realizado por Novaes (2006), 4% dos locos avaliados apresentaram distorção de segregação.

#### 2.6.4 Mapeamento de QTL

Os programas de melhoramento genético basicamente são responsáveis por selecionar dentre os genótipos disponíveis, aqueles de maior interesse para o setor a que o programa se destina, que no contexto desse trabalho é a indústria de celulose. Assim, antes da seleção de caracteres quantitativos, muitos programas de melhoramento necessitam realizar a seleção de caracteres cruciais para o desenvolvimento de uma espécie, tais como resistência a doenças ou condições ambientais limitantes.

Junghans et al. (2003), utilizando marcadores RAPD, avaliaram dez famílias de irmãos completos de *E. grandis*, para mapear genes responsáveis pela resistência à ferrugem causada por *Puccinia psidii* W. Estes autores utilizando a estratégia de BSA (*Bulked Segregant Analysis*) identificaram seis marcadores fortemente ligados ao gene *Ppr1* (*P. psidii* resistance gene 1), que confere resistência à ferrugem. Duas destas marcas distantes 5 cM uma da outra apresentaram um número muito pequeno de recombinantes em relação ao gene *Ppr1* em quase mil meioses avaliadas, permitindo que estes marcadores moleculares sejam utilizados com elevada eficiência em programas de seleção assistida. Avaliando locos SSR e também utilizando a estratégia de BSA, Missiaglia (2005) identificou QTL para florescimento precoce (Eef1), localizado no grupo de ligação 1 de *Eucalyptus*. Rosado (2007) também utilizou marcadores SSR para mapear QTL que confere resistência à ferrugem causada por *Puccinia psidii* em *Eucalyptus* no grupo de ligação 3.

Regiões QTL para características semelhantes às avaliadas neste trabalho tais como crescimento, densidade e qualidade da madeira foram inicialmente identificadas utilizando marcadores RAPD nos trabalhos de Grattapaglia et al. (1996), Verhaegen et al. (1997) e Rocha et al. (2007). Bundock et al. (2008) também mapearam estas características utilizando informações de marcadores RAPD e SSR. Utilizando exclusivamente marcadores microssatélites, Missiaglia (2005) e Novaes (2006) mapearam em *Eucalyptus* QTL relacionados a características de crescimento, densidade, composição química, consumo de álcali efetivo e rendimento de celulose, muitas delas mensuradas a partir de calibrações NIRS.

Verhaegen et al. (1997), ao utilizar técnicas de mapeamento por intervalo, avaliaram a consistência de QTL ligados a características de crescimento e densidade aos 18, 26 e 38 meses de idade. Embora este autor não tenha identificado QTL consistentes ao

longo das três avaliações realizadas, ele verificou QTL significativos em pelo menos dois períodos sucessivos e seu trabalho também demonstrou que a melhor idade para realizar a seleção seria aos 38 meses, período próximo à metade da rotação comercial desta cultura. Neste trabalho foram identificados, aos 38 meses, três QTL para DAP, capazes de representar cerca de 20% da variação fenotípica, localizados no trabalho deste autor nos grupos de ligação 1 e 2. Na avaliação da densidade estimada por pilodyn, este autor identificou aos 38 meses, três QTL capazes de explicar cerca de 20% da variação fenotípica. Estes autores assumindo a herdabilidade desta característica próxima a 40% (Bouvet, 1995), também estimam que os QTL identificados possam representar quase 50% da variação genética.

Bundock et al. (2008), ao mapear QTL testando quatro diferentes modelos para retirar o efeito ambiental dos valores fenotípicos analisados, validaram QTL em diferentes modelos para características de crescimento e pilodyn. Para crescimento foi identificado através de mapeamento por intervalo um QTL com associação significativa em todos os quatro modelos, situado no grupo de ligação 3. Estes autores também identificaram dois QTL com associações significativas para densidade avaliada por pilodyn em todos os quatro modelos. Estes QTL foram identificados nos grupos de ligação 6 e 11, sendo capazes de representar de 6% a 14% da variação fenotípica observada.

Vários autores compararam a robustez dos QTL identificados na AMS através da comparação dos QTL identificados utilizando técnicas de mapeamento por intervalo. De maneira geral, nestes trabalhos o número de QTL identificados através da AMS foi maior que os identificados via mapeamento por intervalo, devido aos marcadores próximos uns dos outros apresentarem significância para o mesmo caráter na AMS, enquanto no mapeamento por intervalo estes locos situados próximos uns dos outros representam somente um único QTL (Breseguello, comunicação pessoal). Grattapaglia et al. (1996) foram os primeiros a comparar em *Eucalyptus* QTL detectados utilizando diferentes métodos de mapeamento (AMS e MIS). Estes autores validaram QTL para caracteres de crescimento nos grupos de ligação 2 e 5, e ao assumir que a herdabilidade ao nível de indivíduos é próxima a 30% (Rezende & Bertolucci, 1993), estimaram que estes QTL sejam capazes de explicar aproximadamente 44% da variação genética aditiva. Para o caráter densidade foram validados quatro QTL, sendo dois deles presentes no grupo de ligação 5 e os outros dois nos grupos de ligação 2, 7, que juntos são responsáveis pelo controle de quase 25% da variação fenotípica observada. Ao assumir que a herdabilidade

deste caráter é próximo a 60%, estes autores estimaram que os QTL significativos para densidade são capazes de explicar quase 50% da variação genética predita dos genótipos avaliados.

Também utilizando marcadores RAPD, Rocha et al. (2007) compararam QTL significativos para os caracteres DAP, densidade e rendimento de celulose utilizando diferentes metodologias de análises de mapeamento. Para cada um destes caracteres foi validado um QTL, identificados respectivamente nos grupos de ligação 4, 2 e 3. Estes QTL são capazes de representar 20%, 30% e 13,5% da variação observada nos caracteres DAP, densidade e rendimento de celulose. Utilizando informações de marcadores SSR, os trabalhos de Missiaglia (2005) e Novaes (2006) são os mais relevantes para efeito de comparação dos resultados obtidos, pois neles foram mapeados QTL para caracteres similares aos avaliados neste trabalho, utilizando muitas sequências de *primers* em comum.

Missiaglia (2005) validou dois QTL para DAP, um no grupo de ligação 1 e outro no grupo de ligação 3, capazes de representar juntos pouco mais de 18% da variação fenotípica observada para este caráter. Este autor também validou dois QTL para densidade nos grupo de ligação 4 e 6, capazes de controlar juntos pouco mais de 15% da variação fenotípica observada. Para o teor de lignina foram validados outros dois QTL nos grupos de ligação 9 e 2, capazes de explicar respectivamente cerca de 40% e 20% da variação fenotípica. Este autor ainda validou utilizando diferentes técnicas de mapeamento por intervalo um QTL para rendimento de celulose no grupo de ligação 3, capaz de explicar quase 25% da variação fenotípica deste caráter.

Novaes (2006), além de validar QTL empregando diferentes técnicas de mapeamento, também validou QTL para caracteres correlacionados, tendo identificado na AMS 11 QTL significativos comuns para os caracteres DAP e altura. Este autor validou um QTL para diâmetro no grupo de ligação 6, e dois para altura nos grupos de ligação 2 e 11, capazes de explicar, cerca de 24% e 15% da variação fenotípica destes dois caracteres. Para os caracteres densidade e pilodyn, este autor identificou na AMS dois locos significativos para as duas características, e validou QTL nos grupos de ligação 6, 8, 9 e 10 para o caráter pilodyn, capazes de explicar aproximadamente 50% da variação fenotípica, enquanto para o caráter densidade foram validados QTL nos grupos de ligação 1, 6 e 8, capazes de explicar cerca de 20% da variação fenotípica. Para a relação siringil/guaiacil e o teor de lignina, este autor identificou na AMS 12 QTL comuns para estes dois caracteres, localizados nos grupo de ligação 9 e 10, além de outros significativos para os caracteres

isoladamente. Através da análise de MIC, este autor validou três QTL para relação siringil/guaiacil situados nos grupos de ligação 1 e 10, sendo estas regiões também significativas dentre os seis QTL detectados para o teor de lignina, localizados nos grupos de ligação 1, 4, 8 e 10, capazes de explicar respectivamente quase 20% e 50% da variação fenotípica destas duas características. Para o caráter rendimento de celulose foi validado somente um QTL localizado no grupo de ligação 10, capaz de explicar aproximadamente 13% da variação fenotípica deste caráter na progênie avaliada.

O acúmulo de conhecimento que ocorreu com o estudo direcionado de regiões genômicas relevantes, paralelo ao aumento da capacidade de geração de informações moleculares decorrente da evolução das técnicas de sequenciamento, permitiu o desenvolvimento de novas linhas de estudo utilizando ferramentas de bioinformática (Lemsk, 2008). Utilizando a abordagem de genes candidatos, Henery et al. (2007) mapearam em *E. nitens* QTL relacionados com o teor de compostos secundários, avaliados através de NIRS. Gion et al. (2000) também realizaram em *E. grandis* o mapeamento de QTL para genes envolvidos na síntese de lignina utilizando a abordagem de genes candidatos, identificando quatro QTL robustos nos grupos de ligação 2, 6 e 7.

Kirst et al. (2004), analisando dados de expressão gênica de mais de 2600 genes utilizando sondas de *microarrays*, mapearam QTL para caracteres de crescimento e teor de lignina em *Eucalyptus*. Neste trabalho foram identificados via mapeamento por intervalo composto, dois QTL altamente significativos e de efeitos opostos para o caráter DAP. Uma das combinações alélicas do QTL situado no grupo de ligação quatro apresentou efeito aditivo positivo de 1,5 cm no diâmetro em relação à média, enquanto o QTL do grupo de ligação nove apresentou efeito aditivo negativo de 1,25 cm no diâmetro em relação à média, sendo que estes dois QTL juntos são capazes de representar quase 30% da variação fenotípica deste caráter. Embora estes autores não tenham identificado QTL para o caráter teor de lignina, eles verificaram que as unidades siringil da lignina são 38% mais abundantes em árvores de crescimento lento, e o teor de lignina é 10% menor em árvores de rápido crescimento, concluindo que os genes responsáveis pelo crescimento também podem influenciar na composição e no teor de lignina.

Dentro deste contexto, o mapeamento de QTL realizado nesse trabalho, juntamente com informações de QTL mapeados em outros seis cruzamentos realizados pelo do Projeto Genolyptus, permitirá a identificação e validação de QTL robustos para caracteres importantes para a indústria de papel e celulose. A construção destes mapas

genéticos foi realizada utilizando softwares desenvolvidos dentro do Projeto Genolyptus, que se apresentam mais adequados para o mapeamento em espécies perenes do que os softwares comumente utilizados, que foram originalmente desenvolvidos para a construção de mapas genéticos em espécies anuais.

Dentro do Projeto Genolyptus foram realizados mais de trinta cruzamentos controlados entre diferentes espécies e procedências de *Eucalyptus*. Como não seria possível mapear QTL em todos os cruzamentos idealizados, a escolha dos cruzamentos que seriam mapeados com marcadores SSR foi realizada utilizando como critério de seleção os cruzamentos que apresentassem maior variabilidade para os caracteres avaliados. Não por acaso os cruzamentos selecionados foram aqueles que envolviam genitores híbridos, evidenciando o polimorfismo superior presente nas progênies obtidas a partir destes tipos de parentais. De acordo com Cervera et al. (2000), a seleção de genitores geneticamente divergentes ou não relacionados maximiza o nível de heterozigosidade da progênie e a quantidade de locos capazes de segregar para caracteres de interesse.

Um dos genitores avaliados neste trabalho é um híbrido interespecífico, que reúne características de rápido crescimento e resistência a doenças, enquanto o outro é uma matriz pura de *E. globulus*, a espécie mais indicada para a indústria de papel e celulose, mas que no entanto é mais adaptada a climas temperados. Em diferentes cruzamentos realizados dentro do Projeto Genolyptus envolvendo duas matrizes desta espécie, foram verificados altos índices de mortalidade nos experimentos de campo nas progênies obtidas a partir deste tipo de genitor, principalmente nos experimentos instalados em regiões com condições climáticas mais quentes.

Desta forma, a estratégia de mapeamento de QTL em diferentes cruzamentos para um mesmo conjunto de caracteres adotada no Projeto Genolyptus, permite a identificação de regiões que se apresentem significativas para os caracteres avaliados nos diferentes cruzamentos mapeados, possibilitando o direcionamento de estudos de sequenciamento e identificação de genes candidatos nas regiões mais relevantes, aumentando a eficiência e reduzindo os custos envolvidos nestes estudos. Os dados coletados neste trabalho permitem investigar a interação entre QTL e ambientes, através da detecção de genótipos com médias estatisticamente diferentes entre os ambientes avaliados, proporcionando o direcionamento da seleção assistida por marcadores moleculares de maneira independente, de acordo com o ambiente onde estas progênies foram avaliadas.

### 3 MATERIAL E MÉTODOS

#### 3.1 MATERIAL VEGETAL

Para o mapeamento de QTL em diferentes ambientes foram avaliados 203 indivíduos de uma progênie correspondente a uma geração  $F_1$ , proveniente de um cruzamento controlado entre genótipos de diferentes espécies de *Eucalyptus*, utilizados em plantios clonais pelas empresas participantes do Projeto Genolyptus. A progênie avaliada neste estudo foi obtida do cruzamento de um parental de *E. globulus* com um híbrido inter-específico, obtido por sua vez do cruzamento entre matrizes de *E. dunnii* e *E. grandis*, sendo identificado dentro do Projeto Genolyptus como o cruzamento DG x GL2. O parental de *E. globulus* teve seu pólen coletado para fecundar o híbrido inter-específico utilizado como porta-semente, seguindo metodologia para polinização controlada descrita por Assis (1993).

O cruzamento entre indivíduos superiores de espécies diferentes tem como objetivo a captação de características desejáveis presentes nestas espécies, tal qual o processo de piramidização gênica e seleção recorrente utilizado nas espécies agrícolas. De acordo com Duarte (2007), *E. globulus* possui elevado rendimento no processo de produção de celulose por possuir densidade superior às demais espécies, sendo a espécie mais utilizada para reflorestamentos em Portugal e na Espanha. Já a espécie *E. grandis* possui rápido crescimento mas não é tolerante à seca e ao frio, enquanto *E. dunnii* apresenta melhor adaptabilidade a regiões mais frias e resistência a alguns patógenos (Rocha, 2000).

O cruzamento avaliado neste experimento (DG x GL2) produziu mais de mil sementes de irmãos completos, que foram germinadas e plantadas em cinco estados com condições edafo-climáticas diferentes (RS, SP, MG, BA e PA). O delineamento experimental utilizado foi de blocos incompletos, com duas testemunhas por bloco, provenientes de outras duas famílias de polinização controlada onde genitores eram de uma mesma espécie. Estas duas famílias tiveram seus indivíduos clonados e plantados nas cinco regiões onde os experimentos foram implantados, junto com mudas sem réplicas clonais



das outras 29 famílias de polinização controlada idealizadas pelo Projeto Genolyptus, que foram analisadas em conjunto para determinar o melhor modelo estatístico que representaria cada variável analisada, bem como determinar os componentes de variância para a progênie DG x GL2.

Três anos e meio após a instalação dos experimentos, período que corresponde à metade da rotação comercial dos povoamentos desta cultura, foi realizada mensuração das características silviculturais e coleta de material lenhoso de todas as árvores presentes nos experimentos de campo, para posterior análise das propriedades da madeira utilizando Espectroscopia do Infra-Vermelho Próximo - NIRS. De acordo com dados publicados por Resende & Higa (1990), mensurações realizadas a partir da metade do período da rotação de povoamentos florestais possuem elevada correlação com os dados de produtividade no período final da rotação, podendo ser realizadas prognoses a partir deste período com elevada confiabilidade. Atualmente as florestas clonais de alta produtividade possuem uma rotação comercial final entre sete e oito anos de idade, principalmente quando a matéria-prima produzida não será utilizada como fonte de madeira sólida, mas sim como fonte de energia ou celulose (Abraf, 2008).

Ao quarto ano após o plantio foi realizada dentro do Projeto Genolyptus a coleta de material vegetal de cinco dentre os 29 cruzamentos controlados disponíveis, que foram selecionados com base na presença de maior variabilidade para os caracteres avaliados com NIRS. Não por acaso os cruzamentos selecionados possuíam um híbrido interespecífico como um de seus parentais. Foram coletadas para extração de DNA, amostras de material vegetal de 203 indivíduos oriundos do cruzamento DG x GL 2, sendo que 154 destas amostras foram coletadas no experimento instalado no estado do Rio Grande do Sul (RS) e 49 amostras no experimento situado em Minas Gerais (MG).

### 3.2 ANÁLISE DOS CARACTERES QUANTITATIVOS

Os caracteres silviculturais foram avaliados seguindo metodologia amplamente difundida no setor florestal, enquanto as estimativas dos quatro caracteres avaliados utilizando NIRS foram realizadas utilizando curvas de calibrações definidas dentro do âmbito do Projeto Genolyptus. Todos os oito caracteres avaliados apresentaram distribuição similar à distribuição normal, e não precisaram ser submetidos a transformações estatísticas para se adequar aos critérios quantitativos. Na apresentação dos

resultados do mapeamento de QTL, foi inicialmente apresentado o valor médio dos caracteres quantitativos avaliados, seguidos de sua variância, representada pela unidade de medida de cada caractere elevada ao quadrado.

Os três caracteres silviculturais foram avaliados em campo utilizando fita métrica para a mensuração do diâmetro na altura do peito (DAP), hipsômetro para a avaliação da altura e fator de forma para a estimativa do volume, obtido pela multiplicação da área basal pela altura e por um fator de forma de 60%. As unidades de medidas utilizadas na avaliação destes três caracteres foram, respectivamente, centímetros (cm), metro (m) e metros cúbicos (m<sup>3</sup>), conforme apresentado a seguir na fórmula utilizada para a estimativa do volume.

$$V = \left( \frac{\pi \cdot \text{DAP}^2}{40000} \right) \cdot \text{Altura} \cdot 0,6, \text{ em que:}$$

V = volume da árvore (m<sup>3</sup>);

$\pi$  = constante matemática (3,14159...);

DAP = diâmetro da árvore na altura do peito (cm);

Altura = altura da árvore (m).

A massa específica ou densidade básica (DB) foi estimada por NIRS nas duas localidades, mas foi aferida com dados de pilodyn somente no experimento do RS. A densidade avaliada por pilodyn é mensurada em milímetros (mm), enquanto que a densidade estimada por NIRS é dada em quilograma por metro cúbico (Kg.m<sup>-3</sup>).

As propriedades da madeira ligadas à química úmida, tais como teor de lignina, relação siringil/guaiacil da lignina e rendimento depurado de celulose também foram estimadas por NIRS. Estes caracteres são relevantes para o processo de produção de polpa celulósica com propriedades comerciais, que consiste em solubilizar e remover a lignina que confere cor escura à polpa e ao papel (Santos, 2005). O teor de lignina total (LigTotal) e o rendimento depurado de celulose (RD) são expressos em %, enquanto a relação siringil/guaiacil (rSG) é uma medida adimensional.

Inicialmente foi verificada a qualidade dos experimentos de campo através da análise do coeficiente de variação do erro experimental (CVe). De acordo com Vencovsky (1992), em experimentos envolvendo espécies perenes, valores do CVe abaixo de 0,15 indicam que as estruturas experimentais foram bem definidas. O CVe é obtido dividindo-se a raiz quadrada do quadrado médio do resíduo experimental pela média do caráter, conforme apresentado na fórmula geral a seguir:

$$CVe\% = \left( \frac{\sqrt{QMres}}{m\acute{e}dia} \right) \cdot 100, \text{ em que:}$$

QMres e media sao parametros de cada carater validado.

Posteriormente, a associao entre os caracteres dois a dois foi avaliada atraves do coeficiente de correlao de Pearson (R). A significncia destas correlaes foi verificada utilizando o critrio FDR (False Discovery Ratio - Benjamini & Hochberg, 1995), considerando-se um nvel de significncia de 5%. Os valores das correlaes envolvendo a caracterstica pilodyn so referentes somente s correlaes com outras caractersticas avaliadas no experimento do RS, devido a este carater ter sido avaliado somente neste local. A frmula geral utilizada pelo critrio FDR  dada por:

$$FDR_{(\alpha_{i-j})} = \left( \frac{n_i}{n_j} \right) \cdot \alpha, \text{ em que:}$$

$\alpha$  = nvel de confiana utilizado no teste;

$\alpha_{i-j}$  = valor crtico do teste para as variveis ordenadas de  $i$  a  $j$ ;

$n_i$  = nmero de ordenamento da varivel  $i$ ;

$n_j$  = nmero de ordenamento da ltima varivel testada.

### 3.3 ESTIMAO DOS COMPONENTES DE VARINCIA

Para reduzir o efeito da variao ambiental presente nos valores fenotpicos observados nas oito caractersticas, foram obtidos os valores eBLUP (Melhor Preditor Linear no Viciado Emprico), que correspondem aos valores genotpicos. Estes valores genotpicos foram obtidos aps a anlise de todas as famlias de mapeamento presentes nos experimentos de campo, dentre as quais duas destas famlias possuam rplicas clonais e foram utilizadas como testemunhas nos experimentos, permitindo-se estimar a variao dentro das estruturas de blocos e bloquinhos para cada local analisado.

Os componentes de varincia genotpica dos oito caracteres avaliados foram estimados para cada ambiente separadamente. A varincia fenotpica total em cada ambiente foi obtida somando-se a varincia genotpica, a varincia residual, a varincia entre superblocos e a varincia entre bloquinhos dentro de superblocos, quando o melhor modelo definido incluiu todos estes componentes de varincia. As linhas de comandos utilizadas para a estimativa dos componentes de varincia no Programa R para os caracteres avaliados nos experimentos de campo do Projeto Genolyptus, bem como comparar os diferentes modelos testados para se obter os valores eBLUP destes caracteres,

foram desenvolvidas em conjunto com Coelho (comunicação pessoal), e se encontram disponíveis para consulta no Apêndice D.

No modelo descrito o valor eBLUP corresponde à soma dos componentes  $\mu$ ,  $\alpha_i$  e  $\beta_{j(i)}$ , referentes aos componentes genéticos dos indivíduos avaliados. Para cada variável em cada ambiente foram utilizados os modelos mais simples possíveis, escolhidos com base no critério de menor valor AIC – Akaike Information Criterion (Pinheiro et al., 2007). Os cofatores de correção para as falhas do experimento foram utilizados somente nos modelos avaliados no ambiente RS. As estimativas dos eBLUP foram obtidas no Programa R (The R Foundation for Statistical Computing, 2008), utilizando-se o modelo descrito a seguir:

$$Y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + \delta_{l(k)} + \kappa_{L(ijkl)} + \kappa_{C(ijkl)} + \kappa_{D(ijkl)} + e_{ijkl}, \text{ em que:}$$

$Y_{ijkl}$  = valor Genotípico do indivíduo  $ijkl$ ;

$\mu$  = média geral do caráter avaliado em cada ambiente – efeito fixo;

$\alpha_i$  = efeito da família  $i$  – efeito fixo;

$\beta_{j(i)}$  = efeito dos indivíduos  $j$  pertencentes à família  $i$  – efeito aleatório;

$\gamma_k$  = efeito dos superblocos  $k$  em cada ambiente – efeito aleatório;

$\delta_{l(k)}$  = efeito dos bloquinhos  $l$  dentro dos superblocos  $k$  – efeito aleatório;

$\kappa_{L(ijkl)}$  = cofator para a presença de falhas na linha;

$\kappa_{C(ijkl)}$  = cofator para a presença de falhas na coluna;

$\kappa_{D(ijkl)}$  = cofator para a presença de falhas na diagonal;

$e_{ijkl}$  = efeito residual do indivíduo  $ijkl$  – efeito aleatório.

### 3.4 EXTRAÇÃO DE DNA GENÔMICO

Para realizar a extração de DNA genômico foram coletadas amostras do tecido do câmbio vascular, que foram acondicionadas em tubos tipo Falcon de 15 mL contendo solução tamponante CTAB 2% para evitar a oxidação do tecido lenhoso e consequentemente, a degradação do DNA presente nas amostras. O isolamento do DNA das amostras foi realizado com o auxílio do equipamento *TissueLyser*, que permite que até 48 amostras sejam maceradas simultaneamente, além de possibilitar a agitação das amostras em outras fases do protocolo de extração quando necessário (Qiagen, 2004).

A extração de DNA dos parentais DG e GL2 foi realizada a partir de tecido foliar, seguindo o protocolo CTAB 2% descrito para o gênero *Eucalyptus* por Ferreira &

Grattapaglia (1998), adaptado de Doyle & Doyle (1991). Com o objetivo de homogeneizar e otimizar a quantidade de DNA isolada a partir do tecido do câmbio vascular com o auxílio do equipamento *TissueLyser*, foram testadas alterações no protocolo descrito por Ferreira & Grattapaglia (1998). Estes testes indicaram que a quantidade de DNA obtida é inversamente proporcional à quantidade de sal utilizada durante o processo de extração, e por este motivo o protocolo original de Doyle & Doyle (1991) permitiu obter maior quantidade de DNA a partir do tecido do câmbio vascular. A principal diferença entre estes dois protocolos está na ausência do componente CTAB 10% na segunda lavagem com a solução de CTAB e álcool isoamílico (CIA), componente adotado por Ferreira & Grattapaglia (1998) para melhorar a qualidade do isolado a partir do tecido foliar, devido à quantidade de DNA obtida neste tipo de tecido ser superior. O protocolo de extração otimizado para isolamento de DNA genômico a partir do câmbio vascular utilizando o equipamento *TissueLyser* se encontra no Apêndice A.

Para quantificar o DNA, 5  $\mu\text{L}$  do DNA ressuspendido de cada amostra foram coletados e misturados a 7  $\mu\text{L}$  de tampão de carregamento para aplicação em gel de agarose 1%, juntamente com alíquotas de 12  $\mu\text{L}$  de DNA de fago Lambda nas quantidades de 20 ng, 50 ng e 100 ng (Invitrogen, Inc.). Após meia hora de corrida em solução tampão TBE 0,5x (Tris-Ácido Bórico, 90 mM; EDTA, 2 mM), os géis foram corados com brometo de etídeo e visualizados em luz ultravioleta. As amostras foram fotografadas e quantificadas utilizando o sistema de foto-documentação EDAS 290 (Kodak, 2000). A concentração do DNA foi estimada pelo programa EDAS, através da comparação do DNA das amostras com padrões de DNA de fago Lambda de concentração conhecida. A quantidade de DNA estimada pelo programa foi dividida por cinco, para se obter a concentração de DNA das amostras em nanogramas por microlitro (ng/ $\mu\text{L}$ ).

As amostras quantificadas foram diluídas para a concentração de 1 ng/ $\mu\text{L}$  e novamente requantificadas para verificar a necessidade de realizar novos ajustes nas diluições quando necessário. Após realizar o último ajuste na diluição das amostras de DNA para a concentração de 1 ng/ $\mu\text{L}$  com base na visualização em gel de agarose 1%, as diluições foram testadas em uma reação de PCR utilizando um dos pares de *primers* marcados que foram disponibilizados, para verificar a qualidade dos fragmentos amplificados. As amostras que não tiveram seus alelos amplificados satisfatoriamente sofreram novos ajustes na diluição, para homogeneizar a qualidade dos fragmentos

amplificados na reação de PCR e detectados na plataforma de sequenciamento de DNA semi-automática ABI3100 (Applied Biosystems, Inc.).

### 3.5 AMPLIFICAÇÃO DE LOCOS SSR

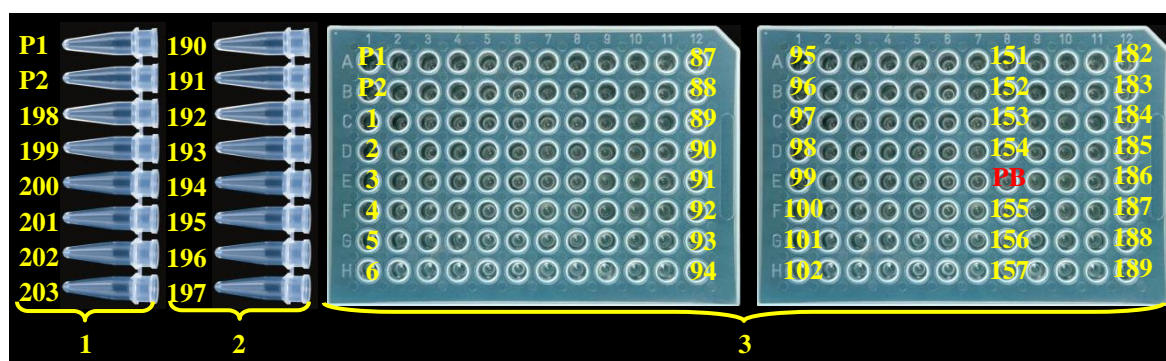
Foram disponibilizados pelo Projeto Genolyptus cerca de trezentos marcadores SSR identificados pelos acrônimos EG e EMBRA, os primeiros desenvolvidos pelo CSIRO e os demais por Brondani et al. (1998; 2002; 2006), Lourenço et al. (2004) e ainda outros marcadores SSR desenvolvidos no âmbito do Projeto Genolyptus, também identificados pelo acrônimo EMBRA, não publicados anteriormente. Todos os pares de *primers* disponibilizados passaram por uma fase de *screening* para identificar os locos polimórficos no cruzamento DG x GL2, utilizando o DNA dos dois parentais juntamente com outros seis indivíduos da progênie.

Os pares de *primers* que não amplificaram satisfatoriamente na primeira fase de *screening* foram submetidos à amplificação em novas reações PCR, agora conduzidas no modo *touch-down*. Neste modo de PCR, a temperatura de anelamento decresce gradativamente conforme programado, permitindo que pares de *primers* com elevada sensibilidade em relação à temperatura de anelamento sejam amplificados com maior segurança devido à ampla variação de temperaturas na qual a reação é amplificada, reduzindo o efeito do fator temperatura como condição limitante para que não ocorra a amplificação dos locos. Nestas reações os ciclos de PCR tiveram início com temperatura de anelamento inicial de 63°C, decrescendo 0,2°C por ciclo até atingir a temperatura de 55°C após quarenta ciclos. Após esta etapa, os locos que novamente não amplificaram alelos nítidos foram definitivamente descartados.

Todas as reações de PCR foram conduzidas no volume de 5 µL, utilizando termociclador GeneAmp PCR System 9700 (Applied Biosystems, Inc.) e o KIT para PCR *Multiplex* da Qiagen para amplificação dos locos SSR. Este *kit* para PCR utiliza uma enzima de qualidade superior (HotStarTaq DNA Polimerase), que diminui as restrições impostas por *primers* com diferentes temperaturas de anelamento. Embora o custo efetivo deste *kit* por reação seja superior, ele apresenta como vantagem a eliminação do tempo gasto com o ajuste das temperaturas de anelamento para os *primers*, além de permitir que até 16 locos SSR possam ser amplificados simultaneamente em uma mesma reação de PCR no volume de 50 µL (Qiagen Inc., 2004).

Os locos que apresentaram bom padrão de amplificação tiveram seus alelos identificados, a cor do fluorocromo confirmada e a faixa de variação dos alelos (*range*) determinada. Após identificar os locos polimórficos, foram testadas combinações *multiplex* de *primers* em uma mesma reação de PCR, fazendo uso do DNA de outras oito amostras da progênie nesta etapa. Foram testadas combinações *multiplex* utilizando de três a sete pares de *primers* SSR simultaneamente, com o objetivo de determinar quantos pares de *primers* SSR poderiam ser amplificados com segurança em uma mesma reação de PCR com volume de 5 µl. As combinações *multiplex* foram definidas com o auxílio do programa MULTIPLEXER (Coelho, não publicado), que define as combinações utilizando como critério de seleção a mínima complementariedade entre as sequências dos *primers* e a similaridade entre suas temperaturas de anelamento. Com base nos resultados de amplificação dos pares de *primers* utilizados isoladamente e em sistema *multiplex*, foram selecionados 76 locos SSR para compor 19 combinações *tetraplex* utilizadas para avaliar o restante da progênie.

O DNA diluído das 189 amostras restantes da progênie foi distribuído em duas microplacas com capacidade para 96 amostras. Na primeira microplaca estava presente o DNA dos dois parentais e de mais 94 indivíduos, enquanto na segunda placa estava presente o DNA de 95 indivíduos mais uma prova em branco, alocada no ponto da placa que separava os dois locais onde foram realizadas as coletas. A distribuição do DNA das amostras em cada uma das etapas do trabalho pode ser visualizada na Figura 1. Informações detalhadas sobre as sequências, fluorocromos e referências dos 76 pares de *primers* SSR amplificados em sistema *multiplex* podem ser consultadas no Apêndice C. A visualização gráfica das combinações *multiplex* se encontra disponível no Apêndice B.



**Figura 1.** Diagrama ilustrando a distribuição das amostras nas etapas de trabalho: 1- *screening* de *primers*; 2- testes *multiplex*; 3- avaliação da progênie DG x GL2. PB: amostra em branco sem DNA.

As reações de PCR conduzidas no volume de 5 µL foram compostas por 2,5 µL de Master Mix 1x; 0,5 µL de Q-Solution 0,2 µM; 0,1 µL de cada par de *primers* na concentração de 10 µM; 1,0 ng de DNA genômico e água ultra-pura qsp. Os ciclos de PCR tinham início com uma etapa de desnaturação a 95°C durante 15 minutos, seguida por 35 ciclos de desnaturação a 94°C por 30 segundos, anelamento a 57°C por 90 segundos e extensão a 72°C por 90 segundos, mais uma etapa final de extensão na temperatura de 72°C durante 30 minutos.

Ao término dos ciclos de PCR a reação era diluída uma vez acrescentando-se 5 µL de água ultra-pura. Para a detecção dos alelos das amostras era coletado 1 µL da reação de PCR diluída, à qual eram adicionados 0,5 µL de solução contendo padrões de peso molecular para análise no sequenciador (*size standard*), e mais 8,5 µL de formamida, constituindo uma alíquota final de 10 µL. Esta mistura era submetida à desnaturação em termociclador a 95°C durante 5 minutos e, em seguida, era imersa em gelo picado por mais 10 minutos, para então ser encaminhada para análise de fragmentos no equipamento ABI3100 (Applied Biosystems).

As reações *tetraplex* de algumas amostras que não tiveram seus alelos identificados na primeira injeção no sequenciador foram novamente reinjetadas. Os indivíduos que ainda não tiveram seus alelos identificados após a reinjeção da primeira reação de PCR tiveram a reação repetida para os locos que não amplificaram. Da mesma maneira que realizado anteriormente, a nova reação de PCR foi encaminhada para análise de fragmentos no equipamento ABI3100 por no máximo mais duas vezes quando necessário. Este procedimento propiciou que o percentual médio de dados moleculares perdidos fosse reduzido de 40% para 10%, sendo ainda verificadas falhas entre 20% e 2% ao término do processo de amplificação de locos SSR, dependendo do loco analisado.

O padrão de fragmentos utilizado no sequenciador foi produzido de acordo com metodologia desenvolvida por Brondani & Grattapaglia (2001), realizando combinações entre *primers* específicos para DNA de cloroplasto, amplificando fragmentos com tamanhos conhecidos e marcados com a fluorescência vermelha (ROX). Os pares de *primers* SSR disponibilizados possuíam uma de suas extremidades 5' marcada com os fluorocromos 6-FAM, NED ou HEX, identificados na plataforma de sequenciamento de DNA semi-automatizada ABI3100 através da detecção das cores azul, amarelo e verde respectivamente. A leitura dos fragmentos e geração das tabelas de genotipagem foi realizada no programa GENEMAPER (Applied Biosystems).



### 3.6 CONSTRUÇÃO DE MAPAS GENÉTICOS

Inicialmente foi realizada a análise de distorção de segregação para todos os locos avaliados pela utilização do teste  $\chi^2$  (Liu, 1998). Esta análise foi realizada no programa PRELIM (Coelho, não publicado). Os locos que apresentaram distorção de segregação significativa perante o “Critério de Bonferroni” ao nível de significância de 5% (Sokal, 1996), foram excluídos das análises por violarem os pressupostos estatísticos utilizados pelos programas que realizam a construção de mapas genéticos.

Os locos que não apresentaram distorção de segregação significativa foram submetidos à análise de ligação para construir um único mapa genético integrado, conforme sugerido por Ritter et al. (1990). De acordo com este autor, a construção de um mapa genético integrado é realizada reunindo a informação dos polimorfismos mapeáveis somente em um dos parentais isoladamente ao polimorfismo mapeável nos dois parentais em conjunto, proporcionando obter um mapa com cobertura genômica superior em relação à construção de mapas individuais para cada parental. Outra vantagem da construção dos mapas integrados é que eles representam as regiões genômicas presentes na progênie onde efetivamente será realizada a seleção, e não as regiões genômicas dos parentais que originaram esta progênie segregante (Ritter et al., 1990).

O mapa integrado foi gerado utilizando o *software* TreeMap (Coelho, não publicado), e a representação gráfica dos grupos de ligação foi realizada com o auxílio do programa MapChart (Voorrips, 2002). Os parâmetros utilizados para declaração da ligação entre marcadores foram: frequência máxima de recombinação de 40% e valor mínimo de LOD (logaritmo da razão de verossimilhança) igual a 3,0. As distâncias entre locos foram calculadas através da função de mapeamento de Kosambi. A identificação dos grupos de ligação do mapa genético integrado foi realizada de acordo a nomenclatura adotada no mapa genético de referência descrito para o gênero *Eucalyptus* por Brondani et al. (2006).

Alguns dos locos, que permaneceram desconectados nas análises de mapeamento mas já haviam sido anteriormente mapeados em outros trabalhos, tiveram suas posições estimadas no mapa genético integrado construído neste trabalho a partir de referências disponíveis, sendo alocados como marcas acessórias. Os locos inéditos avaliados neste trabalho, mas que não ofereceram suporte estatístico para ligação aos grupos formados, não foram alocados na representação gráfica dos grupos de ligação do cruzamento DG x GL2.

### 3.7 ANÁLISE DE QTL

Para realizar o mapeamento de QTL foram utilizados os valores genotípicos preditos pelos eBLUP. Os resultados obtidos foram interpretados fazendo uso mapa de ligação integrado dos parentais. O modelo utilizado para a análise de QTL foi o modelo de análise de marcas simples, devido a nem todas as marcas avaliadas neste trabalho terem proporcionado o suporte estatístico necessário sua devida alocação nos grupos de ligação. A realização de mapeamento por intervalo utilizando as informações aproximadas de alguns locos poderia introduzir um forte *viés* nas análises de mapeamento por intervalo.

Na análise de marcas simples, os marcadores foram testados individualmente quanto à ligação a QTL em diferentes ambientes através de uma Análise de Variância, onde os efeitos avaliados foram os diferentes genótipos encontrados na população para cada marcador. Estas análises foram realizadas utilizando o Programa R (The R Foundation for Statistical Computing, 2008), sendo consideradas válidas somente as associações com p-valores menores que 0,01. As linhas de comandos utilizadas no Programa R para mapeamento de QTL através da AMS e verificação da existência de interação entre QTL e ambientes para algum dos locos avaliados, foram desenvolvidas com a colaboração de Coelho e Pot (comunicação pessoal, 2008), e encontram-se disponíveis no Apêndice E.

Os locos que apresentaram interações significativas entre QTL e ambientes foram avaliados separadamente em cada um dos ambientes. O percentual da variação genotípica explicada pelos QTL declarados como válidos, foi obtido pelo coeficiente de determinação ( $R^2$ ), associado ao efeito dos genótipos no modelo (Beartozzi, 2000). A existência de interação entre QTL e ambientes foi avaliada de acordo com o modelo:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}, \text{ em que:}$$

$\mu$  = média geral do caráter – efeito fixo;

$\alpha_i$  = efeito do genótipo  $i$  – efeito aleatório;

$\beta_j$  = efeito do ambiente  $j$  – efeito aleatório;

$\gamma_{ij}$  = efeito da interação do genótipo  $i$  com o ambiente  $j$  – efeito aleatório;

$e_{ijk}$  = efeito residual do indivíduo  $ijk$  – efeito aleatório.

## **4 RESULTADOS**

Os resultados serão apresentados seguindo a sequência de trabalho descrita na seção de Material e Métodos. Inicialmente serão descritos os resultados das análises dos caracteres quantitativos, tais como componentes de variância e correlações entre os caracteres analisados, para posteriormente serem abordados os aspectos referentes à construção do mapa enético e identificação de QTL para estes caracteres. Para encerrar a seção de Resultados foi realizada uma comparação entre os ganhos genéticos capazes de serem obtidos utilizando diferentes critérios de seleção de genótipos superiores. A comparação dos resultados obtidos com os descritos em por outros autores, bem como a validação e considerações sobre os resultados obtidos, será realizada na seção da Discussão.

### **4.1 COMPONENTES DE VARIÂNCIA**

Inicialmente será apresentado quais caracteres consideraram as estruturas de variação de bloquinhos dentro de superblocos na estimativa dos valores eBLUP para cada ambiente, agrupados de acordo com o modelo selecionado para cada caráter, do modelo mais completo para o mais simples (Tabela 1). Com exceção do caráter pilodyn que foi mensurado somente no RS, todos os outros sete caracteres foram avaliados em ambos os experimentos. Os valores dos componentes de variância e coeficientes de herdabilidade dos oito caracteres avaliados foram estimados para cada ambiente utilizando-se os valores genotípicos preditos dos indivíduos (eBLUP), separados de acordo com o local do experimento onde estes caracteres foram avaliados (Tabelas 2 e 3).

Conforme apresentado na Tabela 1, no experimento do RS os modelos mistos selecionados para estimar os valores eBLUP dos caracteres diâmetro (DAP), volume, pilodyn e teor de lignina (LigTotal), não incluíram a variação presente entre os bloquinhos dentro dos superblocos. Somente o caráter rendimento de celulose (RD) teve seu valor eBLUP estimado utilizando o modelo completo, enquanto nos modelos selecionados para

estimar os eBLUP dos caracteres altura, densidade (DB) e relaçãoiringil/guaiacil (rSG), não foram inclusos os efeitos das falhas identificadas neste local.

**Tabela 1.** Modelos utilizados para se estimar os valores eBLUP nos experimentos.

Experimento	Caracteres	Modelo selecionado
RS	RD	$Y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + \delta_{l(k)} + \kappa_{L(ijkl)} + \kappa_{C(ijkl)} + \kappa_{D(ijkl)} + e_{ijkl}$
	Diâmetro, Volume, Pilodyn e LigTotal	$Y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + \kappa_{L(ijkl)} + \kappa_{C(ijkl)} + \kappa_{D(ijkl)} + e_{ijkl}$
	Altura, DB e rSG	$Y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + \delta_{l(k)} + e_{ijkl}$
MG	Altura, DB, Pilodyn, LigTotal, rSG e RD	$Y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + \delta_{l(k)} + e_{ijkl}$
	Diâmetro e Volume	$Y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + e_{ijkl}$

Em que:  $Y_{ijkl}$  = valor Genotípico do indivíduo  $ijkl$ ;

$\mu$  = média geral do caráter avaliado em cada ambiente – efeito fixo;

$\alpha_i$  = efeito da família  $i$  – efeito fixo;

$\beta_{j(i)}$  = efeito dos indivíduos  $j$  pertencentes à família  $i$  – efeito aleatório;

$\gamma_k$  = efeito dos superblocos  $k$  em cada ambiente – efeito aleatório;

$\delta_{l(k)}$  = efeito dos bloquinhos  $l$  dentro dos superblocos  $k$  – efeito aleatório;

$\kappa_{L(ijkl)}$  = cofator para a presença de falhas na linha;

$\kappa_{C(ijkl)}$  = cofator para a presença de falhas na coluna;

$\kappa_{D(ijkl)}$  = cofator para a presença de falhas na diagonal;

$e_{ijkl}$  = efeito residual do indivíduo  $ijkl$  – efeito aleatório.

Nos modelos mistos utilizados no experimento de MG para se estimar os valores genotípicos eBLUP e componentes de variância, não foram avaliados os fatores de correção para as eventuais falhas presentes no experimento de campo. Da mesma forma que no experimento do RS, dentre os vários modelos testados para cada caráter, os selecionados para estimar os valores eBLUP dos caracteres DAP e volume no experimento de MG não incluíram a variação presente nos bloquinhos dentro dos superblocos (Tabela 1). No entanto, esta fonte de variação esteve presente nos modelos utilizados para se estimar os valores eBLUP dos outros cinco caracteres avaliados neste ambiente. As estimativas dos componentes de variância, coeficientes de herdabilidade e de variação experimental, dos caracteres avaliados nos experimentos de RS e MG, encontram-se respectivamente nas Tabelas 2 e 3.

**Tabela 2.** Parâmetros experimentais e componentes de variância estimados para os caracteres avaliados no experimento do RS.

RS	DAP (cm)	Altura (m)	Volume (m <sup>3</sup> )	Pilodyn (mm)	DB (Kg.m <sup>-3</sup> )	LigTotal (%)	rSG (-)	RD (%)
Média	6,81	8,64	0,027	18,50	453	23,17	3,32	54,79
$\hat{\sigma}^2_{(k+l)}$	0,0509	0,4150	2,9E-05	0,1938	127,4946	0,0256	0,0201	0,1996
$\hat{\sigma}^2_e$	2,0767	1,2683	0,0004	1,6281	126,2984	0,4339	0,0309	0,7236
$\hat{\sigma}^2_g$	4,1755	3,2406	0,0003	1,1444	24,1898	0,3841	0,0372	0,4944
$\hat{\sigma}^2_f$	6,3031	4,9239	0,0007	2,9663	277,9828	0,8436	0,0881	1,4176
$h^2$	0,6625	0,6581	0,4355	0,3858	0,0870	0,4553	0,4223	0,3488
CVe	0,212	0,13	0,074	0,069	0,025	0,028	0,053	0,016

Em que:  $\hat{\sigma}^2_{(k+l)}$ : é o componente de variância dos efeitos dos superblocos e bloquinhos;

$\hat{\sigma}^2_e$ : é o componente de variância residual;

$\hat{\sigma}^2_g$ : é o componente de variância genética;

$\hat{\sigma}^2_f$ : é o componente de variância fenotípica;

$h^2$ : é o coeficiente de herdabilidade no sentido amplo;

CVe: é o coeficiente de variação do erro experimental.

**Tabela 3.** Parâmetros experimentais e componentes de variância estimados para os caracteres avaliados no experimento de MG.

MG	DAP (cm)	Altura (m)	Volume (m <sup>3</sup> )	DB (Kg.m <sup>-3</sup> )	LigTotal (%)	rSG (-)	RD (%)
Média	7,48	10,50	0,031	472	24,15	3,03	56,53
$\hat{\sigma}^2_{(k+l)}$	0,0022	0,1180	1,9E-06	17,1526	0,2564	0,0240	0,2085
$\hat{\sigma}^2_e$	1,1509	0,6909	0,0002	213,1869	0,6013	0,0182	1,7610
$\hat{\sigma}^2_g$	1,6897	1,2033	0,0002	21,7695	0,6744	0,0163	0,3830
$\hat{\sigma}^2_f$	2,8427	2,0122	0,0004	252,1090	1,5322	0,0585	2,3526
$h^2$	0,5944	0,5980	0,5508	0,0863	0,4402	0,2778	0,1628
CVe	0,143	0,079	0,046	0,031	0,032	0,045	0,023

Em que:  $\hat{\sigma}^2_{(k+l)}$ : é o componente de variância dos efeitos dos superblocos e bloquinhos;

$\hat{\sigma}^2_e$ : é o componente de variância residual;

$\hat{\sigma}^2_g$ : é o componente de variância genética;

$\hat{\sigma}^2_f$ : é o componente de variância fenotípica;

$h^2$ : é o coeficiente de herdabilidade no sentido amplo;

CVe: é o coeficiente de variação do erro experimental.

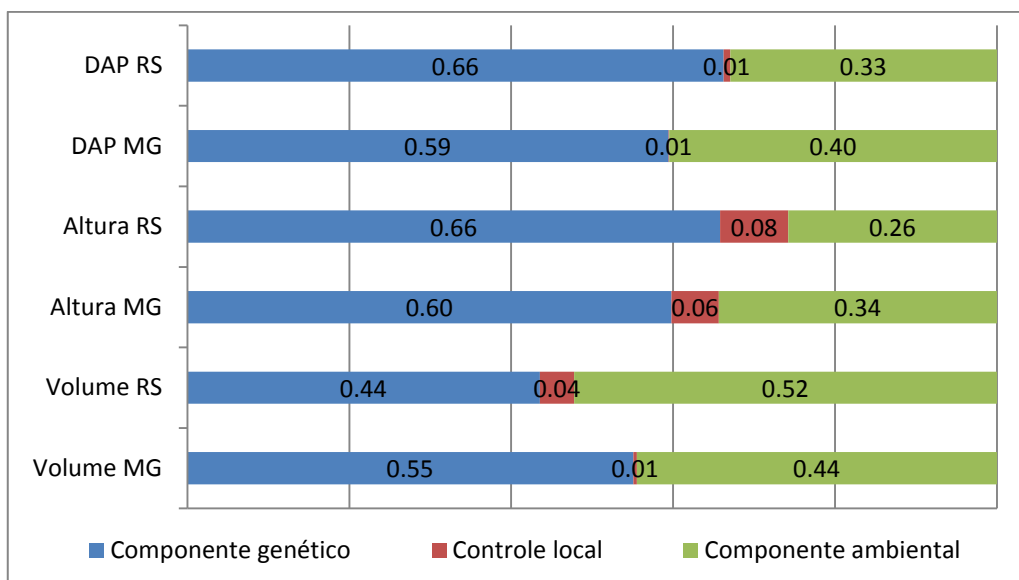
O coeficiente de variação do erro experimental (CVe), que é uma medida da eficiência das estruturas experimentais na captação da variância por estes componentes, somente apresentou valor acima de 0,15 para o caráter diâmetro no experimento do RS. Em ambos os experimentos os maiores valores absolutos do CVe foram verificados para os caracteres de crescimento, sendo que o caráter diâmetro apresentou os maiores valores absolutos de CVe dentre todos os caracteres avaliados. Os menores índices foram verificados para o caráter RD, cujo valor do CVe foi de 0,02 nos dois ambientes. Os coeficientes de variação do erro experimental dos oito caracteres avaliados no experimento do RS variaram de 0,02 a 0,21, enquanto os CVe dos sete caracteres avaliados no experimento de MG variaram entre 0,02 e 0,14.

A comparação dos componentes de variância estimados nos dois experimentos foi realizada reunindo os caracteres avaliados em conjuntos, de acordo com a origem de seus dados, sendo estes agrupados em caracteres de crescimento (seção 5.5.1), densidade estimada por pilodyn (seção 5.5.2) e caracteres avaliados por NIRS (seção 5.5.3). Nestas seções se encontram as Figuras 2, 3 e 4, nas quais é possível verificar a contribuição relativa das fontes de variação consideradas pelos modelos selecionados nos dois experimentos de acordo com os grupos de caracteres avaliados.

Nas Figuras 2, 3 e 4, a proporção explicada pelo componente genético corresponde à herdabilidade no sentido amplo dos caracteres observados, estimados individualmente para cada local onde esta progênie foi avaliada. Devido aos efeitos dos componentes de bloquinhos terem apresentado pequena contribuição para a expressão dos fenótipos, optou-se por representar nas figuras somente o somatório dos efeitos das estruturas de blocos e bloquinhos, identificados pelo componente controle local do experimento.

#### **4.1.1 Caracteres de crescimento**

A influência dos componentes de variância avaliados pelo modelo eBLUP nos caracteres silviculturais podem ser visualizados na Figura 2. Somente para o caráter altura os efeitos dos bloquinhos dentro de superblocos foram significativos nos dois ambientes, e provavelmente por este motivo a fonte de variação representada pelo controle local conseguiu captar para este caráter um montante maior da variância que os demais caracteres de crescimento.



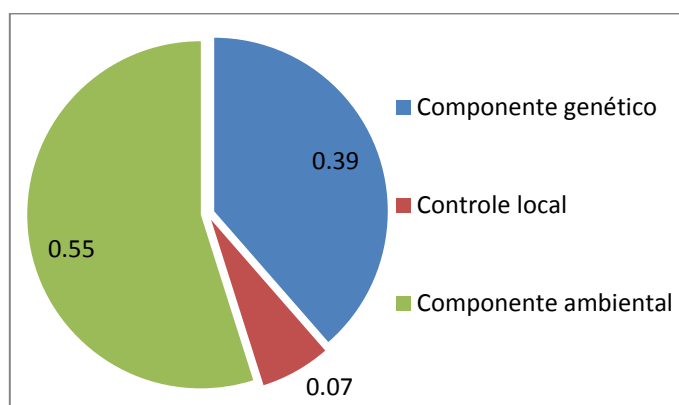
**Figura 2.** Contribuição relativa das fontes de variação na expressão fenotípica dos caracteres de crescimento avaliados nos dois ambientes.

Dentre os caracteres de crescimento, DAP e altura apresentaram coeficientes de herdabilidade similares entre si nos dois experimentos, apresentando valores praticamente idênticos para estes dois caracteres nos dois experimentos, sendo verificados valores ligeiramente superiores no RS em relação ao experimento de MG. Esperava-se que o efeito do componente genético para o caráter volume apresentasse a mesma tendência que os demais caracteres de crescimento, devido ao seu parâmetro ser resultante de uma função dos caracteres diâmetro e altura. No entanto, o coeficiente de herdabilidade do caráter volume no experimento de MG foi maior que o estimado para o experimento do RS, que apresentou o menor valor absoluto dentre todos os caracteres de crescimento avaliados nos dois experimentos.

O coeficiente de herdabilidade do caráter volume em MG (0,55) foi ligeiramente inferior aos dos outros caracteres de crescimento avaliados neste local, que apresentaram valores em torno de 0,60. No experimento do RS a contribuição relativa do componente genético para o volume foi de 0,44, enquanto para os outros dois caracteres foi de 0,66. A herdabilidade do volume no RS foi próxima a  $\frac{2}{3}$  dos coeficientes de herdabilidade estimados para os outros caracteres de crescimento avaliados neste mesmo local, e 20% abaixo do valor estimado para este mesmo componente no experimento de MG.

#### 4.1.2 Densidade avaliada por pilodyn

O modelo misto utilizado para análise deste caráter não incluiu a variação presente nas estruturas de bloquinhos, e seu valor do CVe foi o menor dentre todos os caracteres avaliados em campo. A influência relativa dos componentes de variância do modelo misto, selecionado para a estimativa dos valores eBLUP do caráter densidade mensurado com pilodyn no experimento do RS, pode ser visualizada na Figura 3.



**Figura 3.** Contribuição relativa das fontes de variação na expressão fenotípica do caráter pilodyn no experimento do RS.

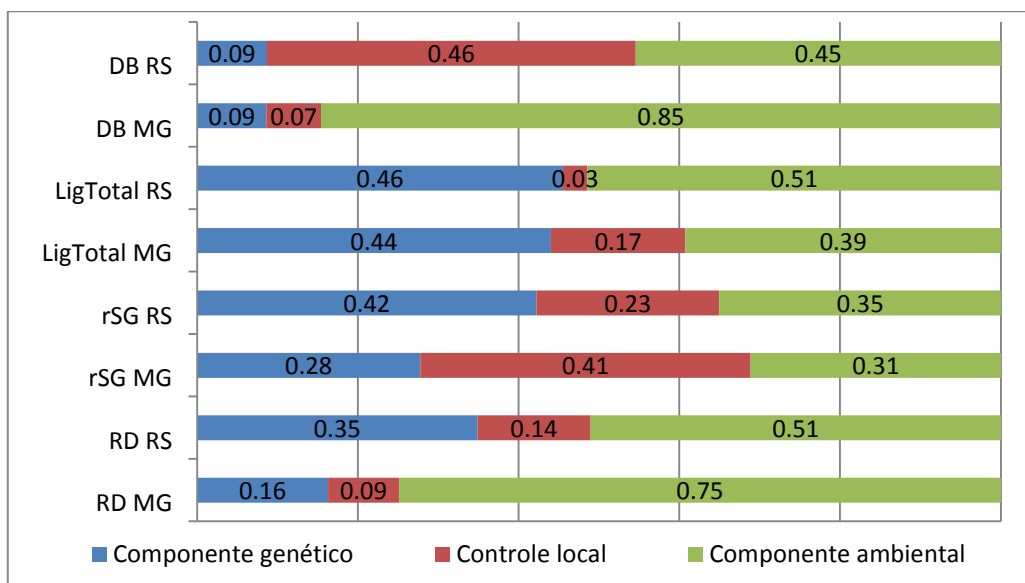
A herdabilidade do caráter densidade avaliado por pilodyn no experimento do RS apresentou valor superior ao verificado na estimativa obtida por NIRS nos dois experimentos. No entanto, a contribuição do componente ambiental na expressão deste caráter neste ambiente foi próximo ao verificado na estimativa obtida por NIRS neste mesmo ambiente, enquanto o montante que o componente de controle local foi capaz de captar apresentou valor idêntico ao verificado para o caráter densidade estimado por NIRS no experimento de MG.

#### 4.1.3 Caracteres avaliados por NIRS

A influência dos componentes de variância estimados pelo modelo eBLUP nos caracteres avaliados por NIRS pode ser verificada na Figura 4. Conforme apresentado nas Tabelas 2 e 3, os valores dos CVe foram muito próximos entre os pares de caracteres avaliados por NIRS nos dois ambientes, se tornando praticamente iguais quando realizado o arredondamento dos valores para somente duas casas decimais. Esta semelhança entre os valores verificados para cada caráter nos dois ambientes é muito provavelmente decorrente das curvas de calibrações utilizadas para estimar estes parâmetros com tecnologia NIRS,



onde foi utilizada a mesma curva de calibração para obter as estimativas destes parâmetros nos dois ambientes.



**Figura 4.** Contribuição relativa das fontes de variação na expressão fenotípica dos caracteres avaliados por NIRS nos dois ambientes.

Os caracteres densidade e teor de lignina apresentaram coeficientes de herdabilidade semelhantes nos dois locais, embora o componente de controle local tenha apresentado diferença de cerca de cinco vezes para estes caracteres entre os experimentos avaliados. Os caracteres rendimento de celulose e relação siringil/guaiacil apresentaram coeficientes de herdabilidades diferentes entre os dois locais, sendo que estes coeficientes apresentaram tendência semelhante à verificada para os caracteres de crescimento diâmetro e altura, sendo verificados valores superiores nos coeficientes de herdabilidades destes caracteres no experimento do RS em relação aos do experimento de MG.

A magnitude dos efeitos associados ao controle local dos experimentos para o caráter rendimento de celulose foi similar ao dos caracteres de crescimento, com valores ligeiramente superiores para o experimento do RS em relação ao de MG. Já a influência deste componente na relação siringil/guaiacil apresentou padrão similar ao do caráter teor de lignina, sendo a magnitude do controle local no experimento de MG superior em relação à verificada no experimento do RS.

## 4.2 CORRELAÇÕES ENTRE CARACTERES QUANTITATIVOS

Os oito caracteres quantitativos tiveram seus dados fenotípicos e genotípicos correlacionados aos pares pelo coeficiente de correlação de Pearson. A Tabela 4 apresenta os coeficientes de correlação significativos que foram estimados entre os diferentes pares de caracteres ( $r_{xy}$ ). No quadrante superior, acima da diagonal principal, encontram-se as correlações entre os valores fenotípicos, enquanto no quadrante inferior encontram-se as correlações entre os valores genotípicos (eBLUP).

**Tabela 4.** Coeficientes de correlação significativos para valores fenotípicos (quadrante superior) e genotípicos (quadrante inferior) entre os caracteres avaliados.

Caracteres	DAP	Altura	Volume	Pilodyn	DB	LigTotal	rSG	RD
DAP	-	0,831	0,967	0,490	-	-	-	-
Altura	0,823	-	0,845	0,428	-	-	-	-
Volume	0,965	0,805	-	0,430	-	-	-	-
Pilodyn	0,524	0,503	0,469	-	-	0,198	-	-
DB	-	0,229	-	-	-	-	0,519	-
LigTotal	-	0,233	-	0,171	0,294	-	-0,567	-0,448
rSG	-	-0,184	-	-	-0,177	-0,713	-	-
RD	-	0,175	-	-	0,572	-	-0,245	-

A utilização do critério FDR para avaliar a significância das correlações foi útil para definir o ponto de corte que estabelecia quais os valores de correlações que seriam significativos, sendo que os pontos de corte determinados se apresentaram próximos para os conjuntos de dados avaliados. O critério FDR realiza uma ponderação a partir do ordenamento decrescente dos p-valores dos testes realizados, definindo um nível crítico específico para cada teste realizado, de acordo com sua posição de ordenamento em relação aos demais. Para as correlações fenotípicas o ponto de corte estimado ficou entre 0,16 e 0,20, enquanto para as genotípicas o ponto de corte ficou entre 0,12 e 0,17.

Somente as correlações fenotípicas entre os caracteres de crescimento apresentaram valores positivos e altamente significativos, acima de 0,80. A correlação entre DAP e volume foi quase linear ( $r_{xy} > 0,96$ ), enquanto a correlação da altura com estes dois caracteres foi muito similar (respectivamente 0,83 e 0,85).

Os caracteres de crescimento também apresentaram correlações genotípicas e fenotípicas positivas e significativas com o caráter pilodyn, embora suas magnitudes

tenham sido moderadas (entre 0,40 e 0,50). O caráter altura ainda apresentou correlações genéticas significativas com todos os caracteres avaliados por NIRS, embora suas amplitudes tenham sido pequenas (próximas a 0,20), e estas correlações não tenham sido significativas para os dados fenotípicos. Estas correlações genéticas do caráter altura se apresentaram positivas com os caracteres densidade, teor de lignina e rendimento de celulose, e negativa para a relação siringil/guaiacil.

Dentre os caracteres de qualidade da madeira estimados por NIRS, algumas correlações não foram significativas nos dois conjuntos de dados avaliados. A correlação entre o rendimento de celulose e o teor de lignina que havia sido significativa nos valores fenotípicos (negativa e próxima a -0,45), não foi verificada para os valores genotípicos, da mesma forma com que a correlação genotípica entre o teor de lignina e a densidade (positiva e próxima a 0,30) não foi verificada para os valores fenotípicos.

A relação siringil/guaiacil também apresentou correlação genética com o rendimento de celulose (negativa e próxima a 0,25), que não foi verificada nos dados fenotípicos. A densidade básica estimada por NIRS ainda apresentou correlação genética significativa acentuada (positiva e acima de 0,50) com o rendimento de celulose, que não foi significativa nos dados fenotípicos.

As demais correlações que haviam se apresentado significativas nos dados fenotípicos também se apresentaram significativas nos dados genotípicos, embora com sentidos contrários em algumas delas. A relação siringil/guaiacil apresentou correlações fenotípicas significativas e superiores a 0,50 com a densidade e o teor de lignina total, sendo a primeira positiva e a segunda negativa. Entretanto as correlações genotípicas da relação siringil/guaiacil se apresentaram negativas tanto para com a densidade, quanto para com o teor de lignina (próximas a -0,20 e -0,70 respectivamente), valores diferenciados com relação aos verificados para os valores fenotípicos.

O teor de lignina ainda apresentou correlação fenotípica e genotípica significativa, positiva e de pequena magnitude com o caráter pilodyn (próximas a 0,20). No entanto, as demais correlações entre os caracteres avaliados por NIRS com a densidade avaliada por pilodyn não se apresentaram significativas.

### 4.3 POLIMORFISMO DE LOCOS SSR

Foram disponibilizados pelo Projeto Genolyptus 292 pares de *primers* SSR. Na fase de screening inicial, 170 locos amplificaram alelos satisfatoriamente com 35 ciclos na temperatura de anelamento de 57°C. Os 122 locos que não amplificaram nestas condições foram novamente submetidos à amplificação em reações de PCR *touch-down* de 40 ciclos. Ao término da fase de *screening*, 53 pares de *primers* não amplificaram alelos detectáveis e foram definitivamente descartados, enquanto dentre os 239 pares de *primers* que apresentaram bom padrão de amplificação, 94 foram descartados por não apresentar polimorfismo mapeável nos genitores que deram origem à progênie avaliada (DG x GL2).

Com base nos resultados obtidos utilizando reações de PCR *multiplex*, ou seja, utilizando conjuntos de três a sete pares de *primers* simultaneamente, foram otimizados conjuntos de quatro pares de *primers* SSR em reações de PCR *tetraplex* com volume final de 5 µL. Embora as combinações tetraplex não tenham funcionado com 100% de eficiência para todas as amostras avaliadas neste trabalho (falha final média de 10%), acredita-se que esta redução na eficiência tenha sido decorrente de variações na qualidade do DNA de algumas amostras, e não a fatores relacionados à competição por sítios de ligação entre os *primers* selecionados para compor as combinações *tetraplex*. O fracionamento do volume da reação original de 50 µL em dez volumes de 5 µL possibilitou que fosse amplificada uma quantidade de locos duas vezes e meia acima da indicada pelo fabricante do *kit* de PCR, sendo amplificado um total de 40 locos em dez reações de 5 µL ao invés de 16 locos em uma única reação de 50 µL.

Uma representação gráfica contendo informações sobre a faixa de amplificação dos alelos (*range*) e a cor do fluorocromo presente nos 76 pares de *primers* SSR selecionados para compor as dezenove combinações *tetraplex*, se encontra disponível para visualização no Apêndice B. Informações detalhadas sobre as sequências dos pares de *primers*, temperatura de anelamento, o grau de polimorfismo dos locos avaliados neste cruzamento, o grupo de ligação onde estes marcadores foram mapeados e o primeiro autor a publicar informações sobre estes locos, se encontram organizadas em forma de tabela no Apêndice C.

Os 76 pares de *primers* SSR selecionados para compor os conjuntos *tetraplex* foram classificados de acordo com seu conteúdo informativo, sendo 29 completamente informativos (em heterozigose nos dois parentais), 39 informativos para o parental DG (em

heterozigose somente no parental DG) e somete oito informativos para o parental GL2 (em heterozigose somente no parental GL2). Após a inspeção da genotipagem realizada pelo programa GENEMAPPER, foram identificadas três árvores como pertencentes a outros cruzamentos (amostras contaminantes), que foram excluídas das análises de mapeamento da progênie DG x GL2. Estes eventos de contaminação foram detectados inicialmente pela presença de alelos diferentes daqueles existentes nos parentais avaliados neste trabalho. Posteriormente, ao conferir os dados dos experimentos constatou-se que duas destas árvores foram coletadas por engano no experimento localizado no RS, enquanto uma árvore do experimento de MG é provavelmente resultante da contaminação de pólen por ocasião do cruzamento.

Dentre os 76 locos avaliados em sistema *multiplex*, catorze apresentaram distorção de segregação significativa em relação ao p-valor obtido no teste de  $\chi^2$ , ao empregar o “Critério de Bonferroni” com intervalo de confiança de 95% (Embra 0078, Embra 0157, Embra 0187, Embra 0197, Embra 0209, Embra 0214, Embra 1320, Embra 1456, Embra 1474, Embra 1688, Embra 1753, Embra 1770, Embra 1934 e Embra 1960). Estes catorze locos foram descartados das análises de construção do mapa genético e de mapeamento de QTL, por violarem os pressupostos estabelecidos por estes modelos, o que poderia interferir nos resultados obtidos nas análises. Dos 62 locos que não apresentaram distorção de segregação significativa, 21 se apresentaram heterozigóticos para ambos os genitores, com a progênie avaliada segregando próximo à proporção de 1:1:1:1. Outros 33 locos foram informativos somente para o parental híbrido interespecífico, enquanto oito foram informativos exclusivamente para o genitor *E. globulus*, todos segregando próximo à proporção de 1:1 na progênie.

#### 4.4 CONSTRUÇÃO DO MAPA GENÉTICO

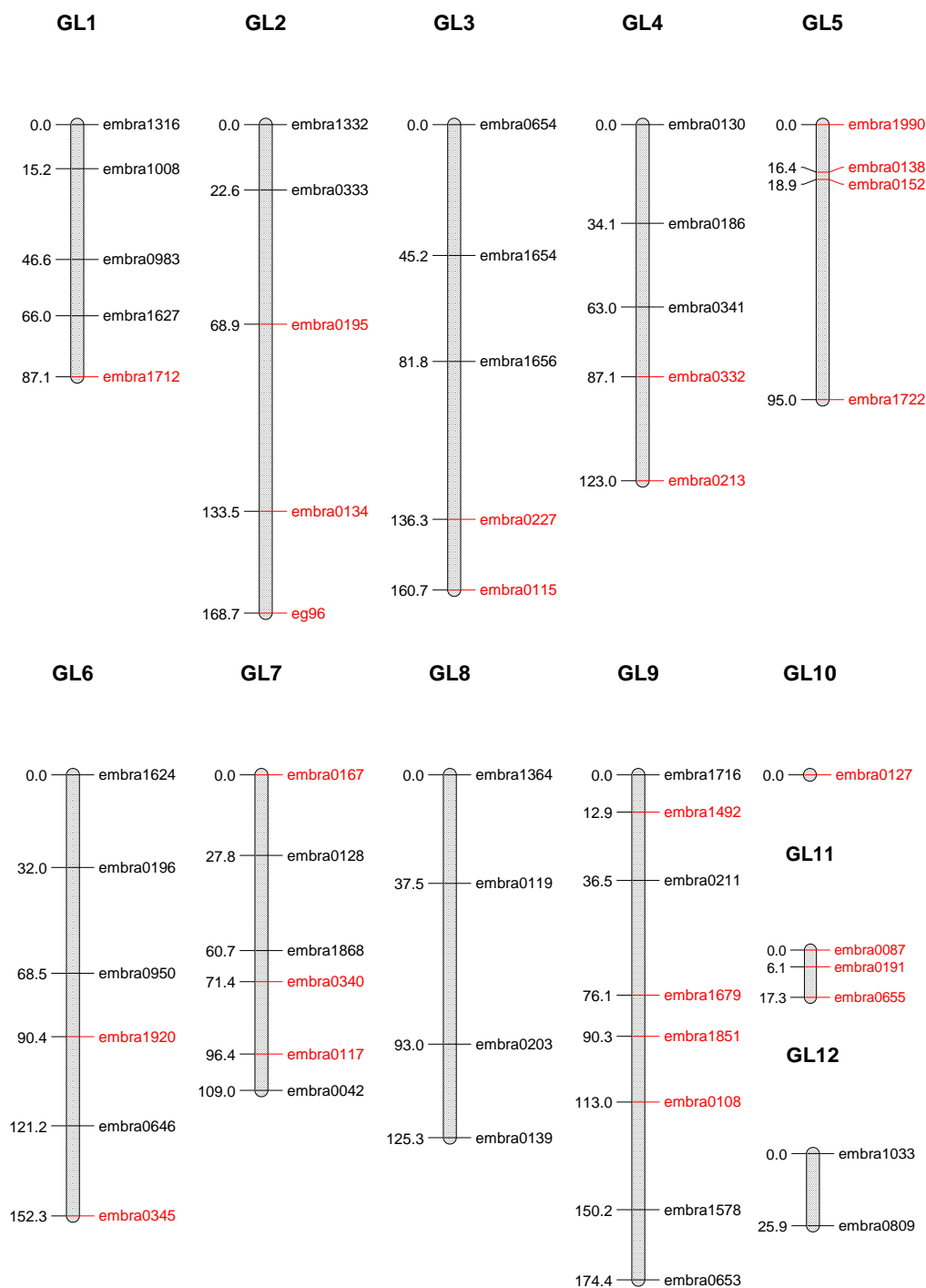
Neste trabalho foi utilizada abordagem proposta por Ritter et al. (1990), segundo a qual se constrói um único mapa genético para os dois parentais avaliados, que representará a distribuição das marcas moleculares analisadas nos indivíduos da progênie avaliada. Os grupos de ligação obtidos foram identificados seguindo nomenclatura definida por Grattapaglia & Sederoff (1994), e adotada na grande maioria dos trabalhos de mapeamento genético em *Eucalyptus*.

Dentre os 62 locos que não apresentaram distorção de segregação significativa, 33 permaneceram desconectados após as análises de mapeamento por não propiciarem o suporte estatístico necessário para definir com segurança a posição destas marcas no mapa genético integrado. Destes 33 locos que não tiveram suas posições definidas, 25 já haviam sido mapeados anteriormente nos trabalhos de Brondani et al. (2002; 2006), Lourenço (2004) e Novaes (2005), e assim, estas marcas puderam ter suas posições estimadas nos grupos de ligação formados neste trabalho como marcas acessórias, utilizando as informações disponíveis em literatura (vide Apêndice C).

Utilizando LOD 3 e fração de recombinação de 0,40, foram obtidos dez grupos de ligação no programa TREEMAP. Ao incluir as informações dos 25 locos que tiveram suas posições estimadas, dois grupos de ligação que inicialmente haviam sido mapeados separadamente foram reunidos e outros três foram identificados, constituindo um total de 12 grupos de ligação. Os grupos de ligação 5 e 11 foram compostos exclusivamente por marcadores que não apresentaram ligação estatística, enquanto o grupo de ligação 9 teve todos seus locos conectados somente ao incluir informações de quatro locos que haviam inicialmente permanecidos desconectados nas análises de mapeamento.

Agregando as informações disponíveis em literatura, o grupo de ligação 10 foi representado somente por uma única marca isolada (Embra 0127), o grupo de ligação 11 foi composto por três locos, os grupos de ligação 5 e 8 por quatro locos, os grupos de ligação 1, 2, 3, 4, e 8 por cinco locos, os grupos de ligação 6 e 7 por seis locos, e no grupo de ligação 9 foram alocados oito locos. O décimo segundo grupo de ligação foi composto somente por duas marcas, que como não haviam sido mapeadas em trabalhos anteriores, permaneceram representadas como um grupo de ligação isolado.

A extensão dos grupos de ligação formados deste trabalho variou de 17,3 cM a 174,4 cM, e a maior distância verificada entre duas marcas (*gap*) foi de 91,8 cM, situada no grupo de ligação 5, entre os marcadores Embra 0152 e Embra 1722. O comprimento total do mapa genético integrado obtido foi próximo a 1240 cM, apresentando uma distância média entre marcas de 22 cM. Este mapa genético abrangeu uma cobertura genômica de aproximadamente 75% em relação ao tamanho do genoma do gênero *Eucalyptus*, estimado por Brondani et al. (2006). O mapa genético integrado dos parentais avaliados nesta população de mapeamento pode ser visualizado na Figura 5.



**Figura 5.** Mapa genético integrado constituído por 54 marcadores SSR genotipados em uma progênie de 200 irmãos completos de *E. globulus* x híbrido (*E. dunnii* x *E. grandis*). Os locos destacados em vermelho foram alocados como marcas acessórias, tendo suas posições aproximadas estimadas com base em informações disponíveis em outros mapas genéticos de *Eucalyptus* que mapearam estes locos anteriormente.

## 4.5 DETECÇÃO DE QTL

Os 62 marcadores analisados foram testados quanto à ligação a QTL através de análise de variância, sendo consideradas significativas as associações com p-valor inferior a 0,01 pelo Teste F, nas análises individuais e conjuntas das árvores amostradas nos experimentos instalados no RS e em MG. A apresentação dos QTL mapeados será realizada inicialmente para cada caráter avaliado, e posteriormente na discussão serão abordados os efeitos do tamanho da amostragem do experimento de MG nos QTL e interações identificadas envolvendo este ambiente. Conforme descrito na metodologia dos caracteres quantitativos, nos resultados do mapeamento de QTL serão apresentados os valores médios de cada um dos caracteres avaliados, seguidos de sua variância, representada pela unidade de medida de cada caractere elevada ao quadrado.

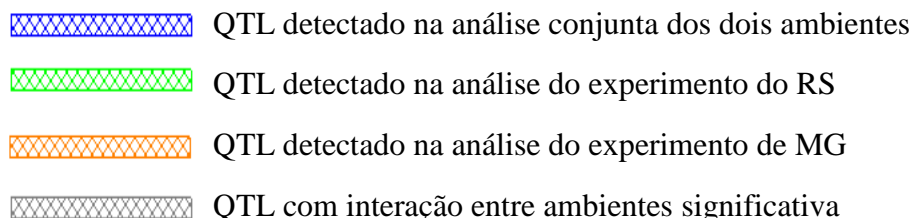
Na análise conjunta dos experimentos foram detectados dois locos que apresentaram interação entre QTL e ambientes, sendo um deles para o caráter teor de lignina e outro para rendimento de celulose. Como o caráter pilodyn foi mensurado em apenas um dos experimentos, não foi possível avaliar a interação entre QTL e ambientes para este caráter, que apresentou no experimento do RS média de 18,77 mm e variância de 0,89 mm<sup>2</sup>. A densidade avaliada por pilodyn foi o único caráter que não apresentou associação significativa a QTL, enquanto para os demais caracteres foram identificados de dois a cinco QTL, que serão apresentados para cada caráter individualmente adiante.

A representação gráfica dos QTL significativos foi realizada seguindo o padrão adotado para a representação de QTL dentro do Projeto Genolyptus, onde os QTL para cada caráter avaliado são destacados em relação a todo o mapa genético construído, e não somente nos grupos de ligação onde estes se encontram, como comumente realizado nos trabalhos de mapeamento de QTL. Os valores das médias genotípicas dos locos associados a QTL podem ser verificados no Apêndice F, enquanto os valores médios das classes genotípicas dos locos que apresentaram interação entre QTL e ambientes encontram-se no Apêndice G.

A Figura 6 representa a legenda que auxilia na interpretação dos mapas genéticos presentes nas Figuras enumeradas de 7 a 13, onde são representados os QTL significativos para os diferentes caracteres avaliados. Os QTL detectados foram representados nos mapas genéticos com uma indicação hachurada próxima ao loco, identificada de acordo com a análise onde o QTL foi significativo. Os QTL válidos na



análise dos dois ambientes em conjunto foram identificados com a cor azul, enquanto os QTL significativos na análise isolada dos experimentos em cada ambiente foram representados com a cor verde no experimento do RS e laranja no experimento de MG. Os locos que apresentaram interação entre QTL e ambientes foram representados com a cor cinza.



**Figura 6.** Identificação dos QTL de acordo com a análise realizada.

#### 4.5.1 Diâmetro na altura do peito - DAP

O caráter DAP apresentou na progênie avaliada valor médio de 7,84 cm com variância de 2,46 cm<sup>2</sup>, não sendo verificada interação entre QTL e ambientes. Na análise do experimento de MG não foram identificados QTL com 99% de confiança, enquanto no experimento do RS foram verificados dois QTL dentro deste intervalo de confiança. Também foram detectados dois QTL na análise conjunta dos dois experimentos, sendo que um destes também foi significativo na análise individual do experimento do RS. Os p-valores e coeficientes de determinação ( $R^2$ ) dos locos com associações significativas a QTL para o caráter DAP, encontram-se na Tabela 5, e o posicionamento destes locos pode ser visualizado na Figura 7.

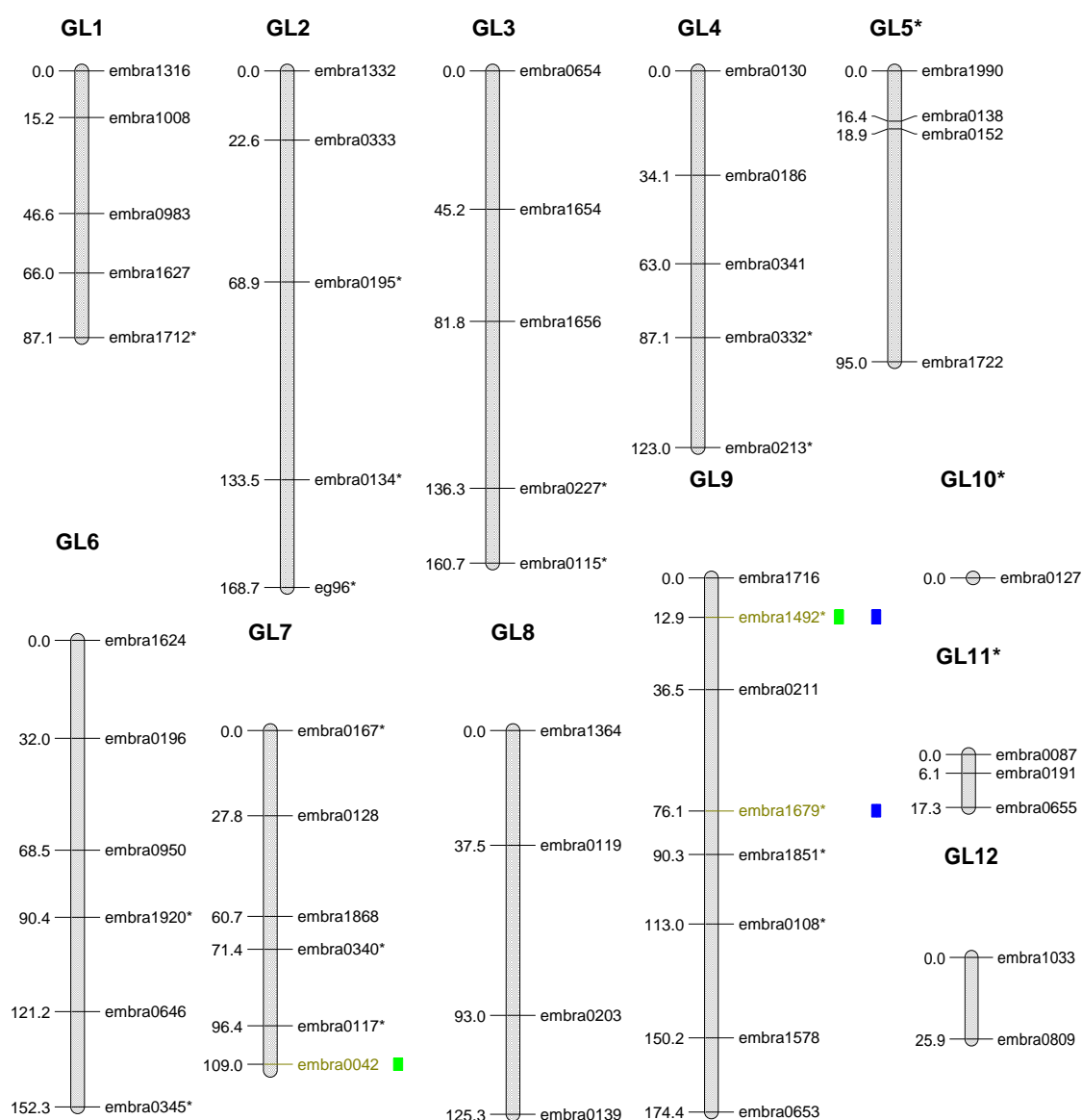
**Tabela 5.** P-valores e coeficientes de determinação dos locos que apresentaram associações a QTL para o caráter diâmetro na altura do peito (DAP) em pelo menos uma das análises realizadas.

Loco	Conjunto		Experimento RS		Experimento MG		Interação
	p-valor	$R^2$ (%)	p-valor	$R^2$ (%)	p-valor	$R^2$ (%)	
Embra 0042	0,015*	6,73	0,007**	8,82	0,386	11,67	0,187
Embra 1492	0,004**	4,74	0,004**	7,42	0,782	0,36	0,339
Embra 1679	0,006**	7,21	0,018*	7,30	0,019*	35,81	0,202

\*: Significativo 0,95; \*\*: Significativo 0,99.

Os locos Embra 1492 e Embra 1679 mapeados no grupo de ligação 9 apresentaram associações a QTL na análise dos experimentos em conjunto, sendo capazes de explicar respectivamente 4,7% e 7,2% da variação genotípica deste caráter na progênie

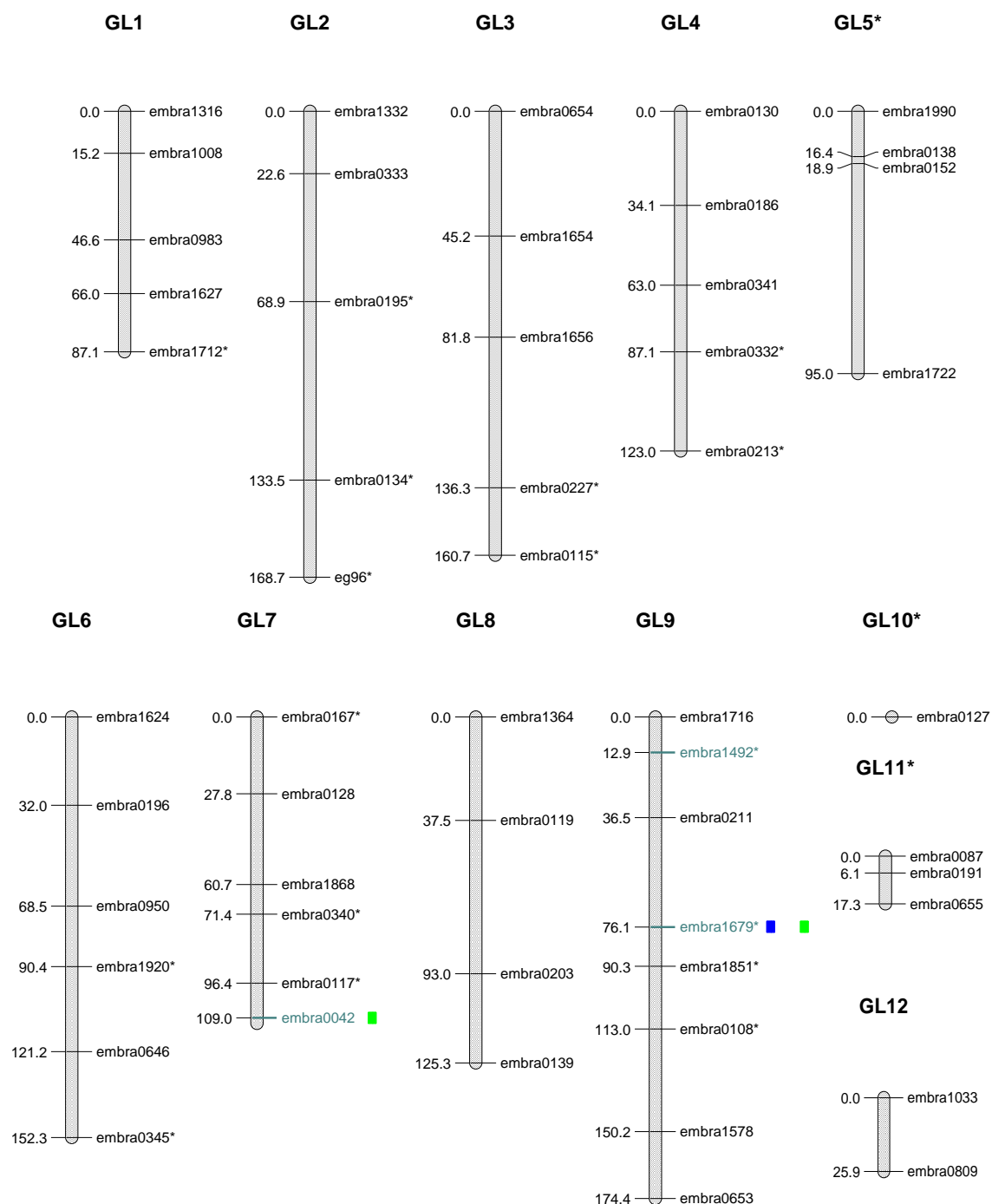
avaliada. O loco Embra 1492 também foi significativo para DAP na análise isolada do experimento do RS, sendo capaz de explicar 7,4% da variação genotípica neste ambiente. O loco Embra 0042 mapeado no grupo de ligação 7 também foi significativo no experimento do RS, sendo capaz de representar quase 9% da variação genotípica neste local. Os locos com associação significativa a QTL na análise conjunta dos dois ambientes foram capazes de controlar juntos aproximadamente 12% da variação genotípica na progênie, e 16% desta variação nos indivíduos avaliados no experimento do RS.



**Figura 7.** Mapa representativo dos QTL detectados para o caráter diâmetro - DAP.

### 4.5.2 Altura

A população de mapeamento avaliada neste trabalho apresentou altura média de 9,76 m e variância de 1,42 m<sup>2</sup>, não sendo observada interação significativa entre QTL e ambientes para este caráter. Os coeficientes de determinação ( $R^2$ ) e p-valores dos QTL identificados nestas análises podem ser verificados na Tabela 6, enquanto a representação dos QTL significativos para este caráter pode ser visualizada na Figura 8.



**Figura 8.** Mapa representativo dos QTL detectados para o caráter altura.

Nos dois ambientes analisados em conjunto foi identificado somente um QTL. Da mesma maneira verificada para o caráter DAP, o QTL detectado na análise conjunta dos dois ambientes também foi significativo na análise individual do experimento do RS, onde foram identificados dois QTL significativos, enquanto no experimento de MG não foram identificados QTL. O loco Embra 1679 mapeado no grupo de ligação 9, apresentou QTL significativo na análise conjunta dos dois ambientes e também na análise individual do experimento do RS, sendo capaz de explicar cerca de 9% da variação genotípica predita em cada uma destas análises. O loco Embra 0042 mapeado no grupo de ligação 7, foi capaz de explicar 8,4% da variação genotípica predita para este caráter na análise do experimento do RS. Os locos associados a QTL para o caráter altura foram capazes de explicar juntos aproximadamente 18% da variação genotípica predita no experimento do RS e 9% desta variação na análise conjunta dos dois experimentos.

**Tabela 6.** P-valores e coeficientes de determinação para os locos que apresentaram associações a QTL para o caráter altura em pelo menos uma das análises realizadas.

Loco	Conjunto		Experimento RS		Experimento MG		Interação
	p-valor	R <sup>2</sup> (%)	p-valor	R <sup>2</sup> (%)	p-valor	R <sup>2</sup> (%)	p-valor
Embra 0042	0,017*	5,28	0,009**	8,38	0,186	17,83	0,113
Embra 1679	0,004**	8,97	0,005**	9,31	0,322	14,37	0,323

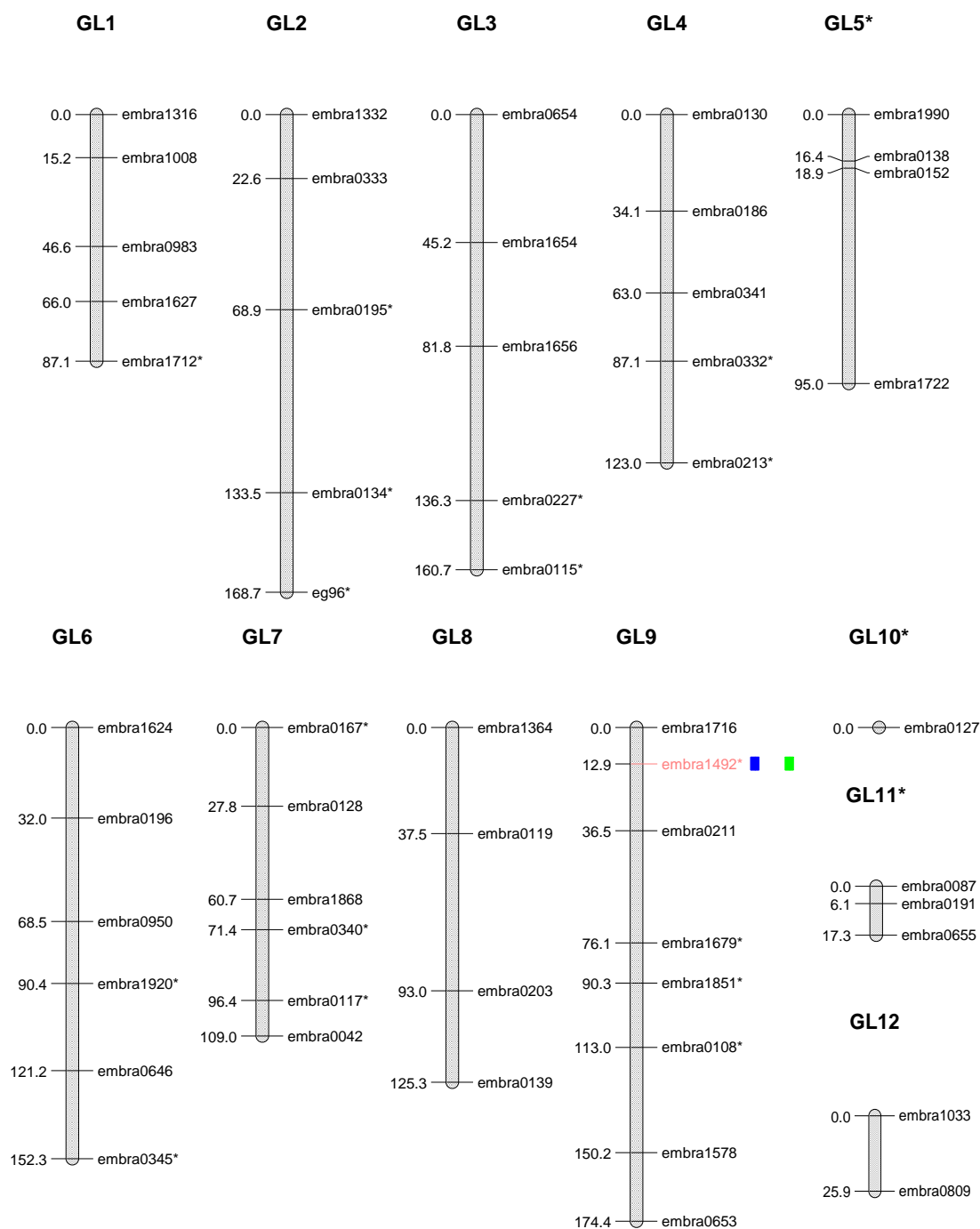
\*: Significativo 0,95; \*\*: Significativo 0,99.

#### 4.5.3 Volume

Para este caráter não foi observada interação entre QTL e ambientes. O volume médio verificado nesta progênie aos 3,5 anos de idade foi de 0,03 m<sup>3</sup>, com variância de 3x10<sup>-4</sup> m<sup>6</sup>. Na análise dos dois ambientes em conjunto foram detectados dois QTL, e da mesma forma verificada anteriormente para os demais caracteres de crescimento, um destes locos foi significativo para o experimento do RS, enquanto no experimento de MG não foram identificados QTL. Os p-valores e coeficientes de determinação observados nestas análises podem ser visualizados na Tabela 7, enquanto a representação gráfica destes QTL pode ser verificada na Figura 9.

O loco Embra 1234 apresentou associação significativa a QTL na análise dos dois ambientes em conjunto, mas foi um dos oito locos avaliados que não tiveram sua posição determinada no mapa genético integrado construído neste trabalho. O loco Embra 1492 mapeado no grupo de ligação 7 apresentou QTL na análise individual do experimento

do RS e na análise conjunta dos dois experimentos, sendo capaz de explicar respectivamente, 7,5% e 4,9% da variação genotípica predita nestas análises. Os QTL identificados na análise conjunta dos ambientes foram capazes de explicar juntos, quase 10% da variação genotípica desta progênie para o caráter volume.



**Figura 9.** Mapa representativo do QTL detectado para o caráter volume.

**Tabela 7.** P-valores e coeficientes de determinação dos locos que apresentaram associações a QTL para o caráter volume em pelo menos uma das análises realizadas.

Loco	Conjunto		Experimento RS		Experimento MG		Interação
	p-valor	R <sup>2</sup> (%)	p-valor	R <sup>2</sup> (%)	p-valor	R <sup>2</sup> (%)	
Embra 1234	0,006**	4,61	0,012*	4,43	0,291	4,11	0,846
Embra 1492	0,005**	4,90	0,003**	7,54	0,762	0,43	0,383

\*: Significativo 0,95; \*\*: Significativo 0,99.

#### 4.5.4 Densidade básica - DB

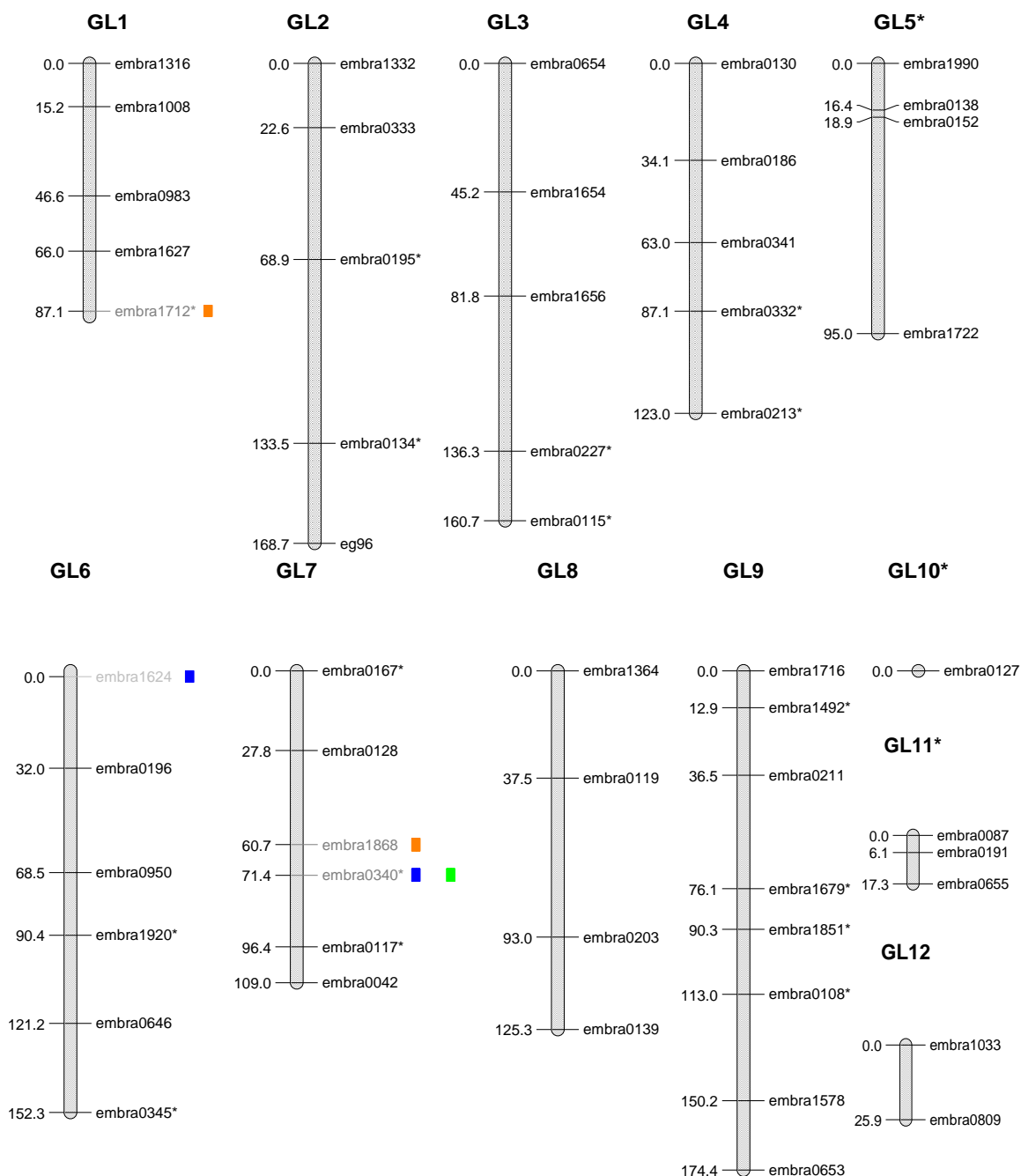
Para o caráter densidade foram identificados QTL somente nas mensurações realizadas por NIRS, uma vez que não foram identificados QTL nas mensurações realizadas utilizando pilodyn. Este caráter não apresentou interações significativas entre QTL e ambientes. O valor médio da densidade estimada por NIRS foi de 457 Kg.m<sup>-3</sup>, com variância de 82,42 (Kg.m<sup>-3</sup>)<sup>2</sup>. O único loco que apresentou QTL para o experimento do RS também foi significativo na análise dos dois ambientes em conjunto, onde foram identificados dois QTL. No experimento de MG também foram identificados dois QTL, conforme dados apresentados na Tabela 8 e na Figura 10.

Os locos Embra 0340 e Embra 1624 apresentaram QTL na análise conjunta dos dois ambientes, e estes locos foram capazes de explicar juntos pouco mais de 17% da variação genotípica predita. O loco Embra 0340 também apresentou QTL na análise individual do experimento do RS, onde foi capaz de controlar quase 5% da variação genotípica neste ambiente. Os locos Embra 1712 e Embra 1868 apresentaram associações significativas a QTL no experimento de MG, e o montante da variação genotípica que estes locos foram capazes de representar foi de quase 100%.

**Tabela 8.** P-valores e coeficientes de determinação para os locos que apresentaram associações a QTL para o caráter densidade básica (DB) em pelo menos uma das análises realizadas.

Loco	Conjunto		Experimento RS		Experimento MG		Interação
	p-valor	R <sup>2</sup> (%)	p-valor	R <sup>2</sup> (%)	p-valor	R <sup>2</sup> (%)	
Embra 0340	0,005**	13,30	0,008**	4,93	0,312	3,65	0,037*
Embra 1624	0,009**	3,80	0,074	2,24	0,018*	15,92	0,193
Embra 1712	0,738	1,50	0,614	4,47	0,004**	62,27	0,025*
Embra 1868	0,403	2,28	0,903	0,58	0,003**	37,65	0,029*

\*: Significativo 0,95; \*\*: Significativo 0,99.

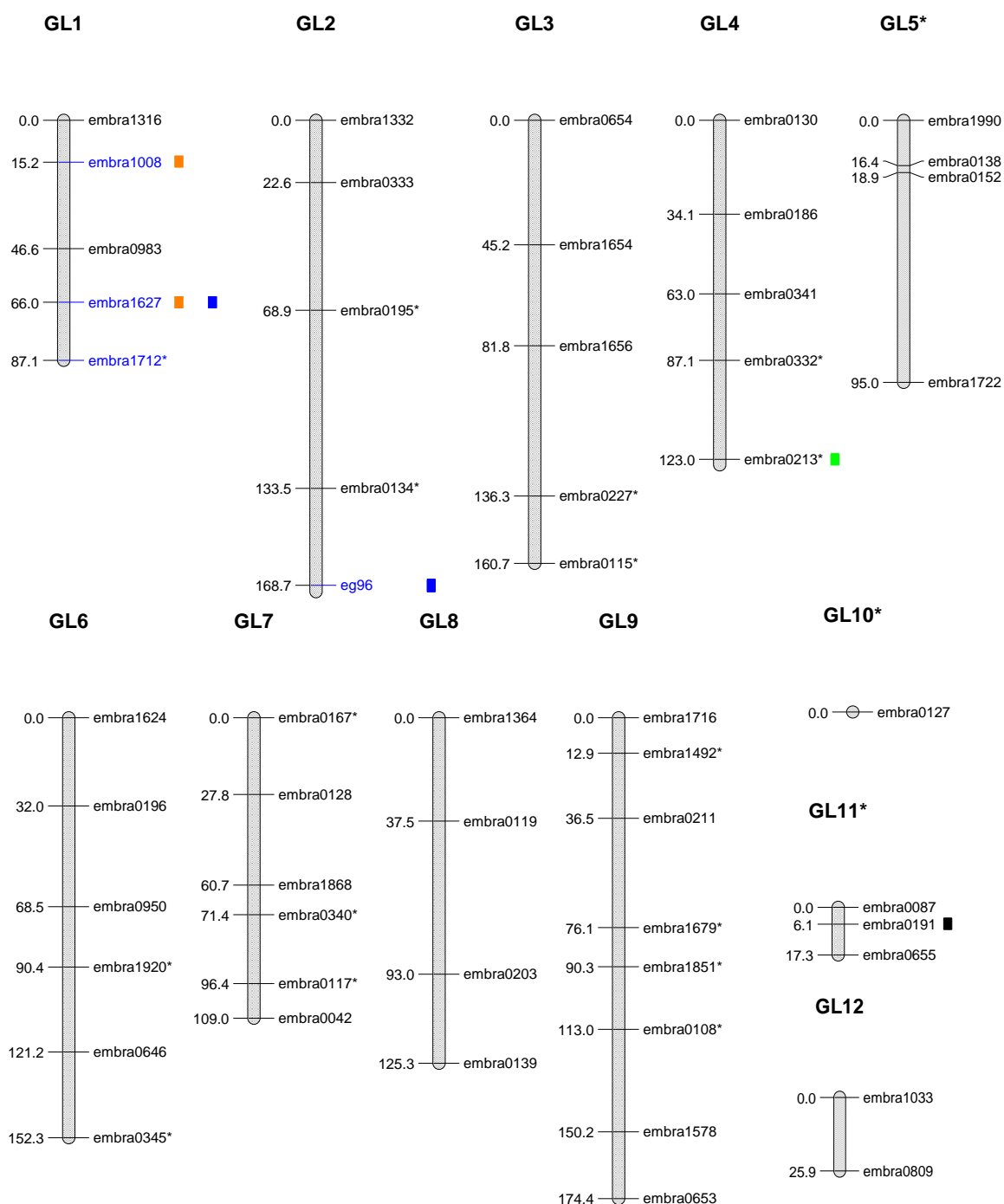


**Figura 10.** Mapa representativo dos QTL detectados para o caráter densidade básica - DB.

#### 4.5.5 Teor de lignina total - LigTotal

A média verificada para o caráter teor de lignina foi de 23,39% com variância de 0,46, sendo identificado um loco que apresentou interação entre os dois ambientes avaliados. A representação dos QTL identificados para este caráter pode ser visualizada na Figura 11, enquanto os p-valores e coeficientes de determinação ( $R^2$ ) estimados para os

locos onde foram identificados QTL para este caráter encontram-se na Tabela 9, logo após a Figura 11.



**Figura 11.** Mapa representativo dos QTL detectados para o caráter teor de lignina - LigTotal.

Embora os genótipos do loco Embra 0191 mapeado no grupo de ligação 11 tenham apresentado interação entre os dois ambientes avaliados, este loco não apresentou QTL significativo em nenhuma das análises realizadas. Na análise do experimento de MG



e dos dois ambientes em conjunto foram detectados dois QTL, capazes de explicar respectivamente 40% e 9% da variação genotípica do teor de lignina nestas análises. No experimento do RS foi detectado somente um QTL, capaz de explicar quase 12% desta variância no experimento do RS.

**Tabela 9.** P-valores e coeficientes de determinação para os locos que apresentaram associações a QTL para o caráter teor de lignina (LigTotal) em pelo menos uma das análises realizadas.

Loco	Conjunto		Experimento RS		Experimento MG		Interação
	p-valor	R <sup>2</sup> (%)	p-valor	R <sup>2</sup> (%)	p-valor	R <sup>2</sup> (%)	p-valor
EG 96	0,008**	1,94	0,053	2,71	0,065	12,52	0,116
Embra 0191	0,443	0,04	0,630	1,61	0,099	37,28	0,004**
Embra 0213	0,037*	5,38	0,009**	11,43	0,183	17,97	0,033*
Embra 1008	0,024*	1,65	0,330	0,75	0,009**	19,55	0,015*
Embra 1627	0,006**	7,12	0,118	1,99	0,009**	19,28	0,038*

\*: Significativo 0,95; \*\*: Significativo 0,99.

O loco Embra 0213 mapeado no grupo de ligação 4, apresentou QTL somente na análise individual do experimento do RS, sendo capaz de explicar pouco mais de 11% da variação genotípica neste local. Os locos Embra 1008 e Embra 1627 situados no grupo de ligação 1, apresentaram associação a QTL no experimento de MG, sendo que cada um deles foi capaz de explicar quase 20% da variação genotípica predita na análise deste ambiente. O loco Embra 1627 também apresentou associação a QTL na análise conjunta dos experimentos, capaz de representar cerca de 7% da variação genotípica predita nesta progênie, enquanto o loco EG 96 mapeado no grupo de ligação 2 foi capaz de explicar somente cerca de 2% desta variação nesta mesma análise.

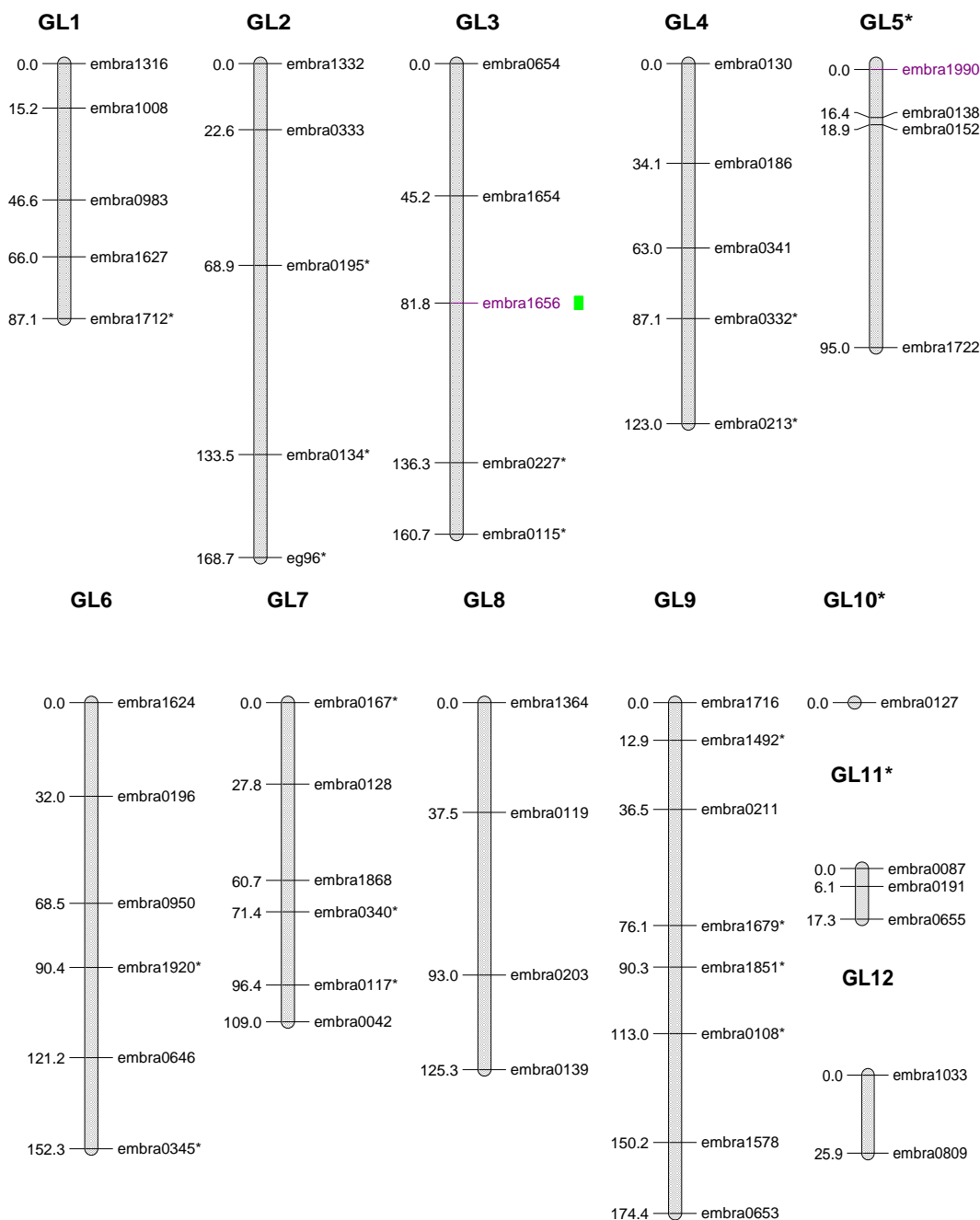
#### 4.5.6 Relação siringil/guaiacil - rSG

O valor médio observado na relação siringil/guaiacil foi 3,26, com variância de 0,04. Este caráter não apresentou interação entre QTL e ambientes ou QTL significativos no experimento de MG. Na análise dos experimentos, um QTL foi significativo no experimento do RS e outro na análise em conjunta dos dois ambientes (Tabela 10). A representação dos QTL identificados para a relação siringil/guaiacil pode ser visualizada na Figura 12.

**Tabela 10.** P-valores e coeficientes de determinação para os locos que apresentaram associações a QTL para a relação sirigil/guaiacil (rSG) em pelo menos uma das análises realizadas.

Loco	Conjunto		Experimento RS		Experimento MG		Interação
	p-valor	R <sup>2</sup> (%)	p-valor	R <sup>2</sup> (%)	p-valor	R <sup>2</sup> (%)	
Embra 1656	0,022*	5,31	0,008**	8,95	0,967	0,85	0,269
Embra 1990	0,007**	5,30	0,020*	6,73	0,173	14,64	0,757

\*: Significativo 0,95; \*\*: Significativo 0,99.



**Figura 12.** Mapa representativo dos QTL detectados para o caráter relação sirigil/guaiacil - rSG.

O loco Embra 1656 mapeado no grupo de ligação 3 apresentou QTL no experimento do RS capaz de representar quase 9% da variação genotípica predita deste caráter neste local. O loco Embra 1990 mapeado no grupo de ligação 5 apresentou QTL na análise conjunta dos experimentos, sendo capaz de explicar pouco mais de 5% da variação genotípica da progênie.

#### 4.5.7 Rendimento depurado de celulose - RD

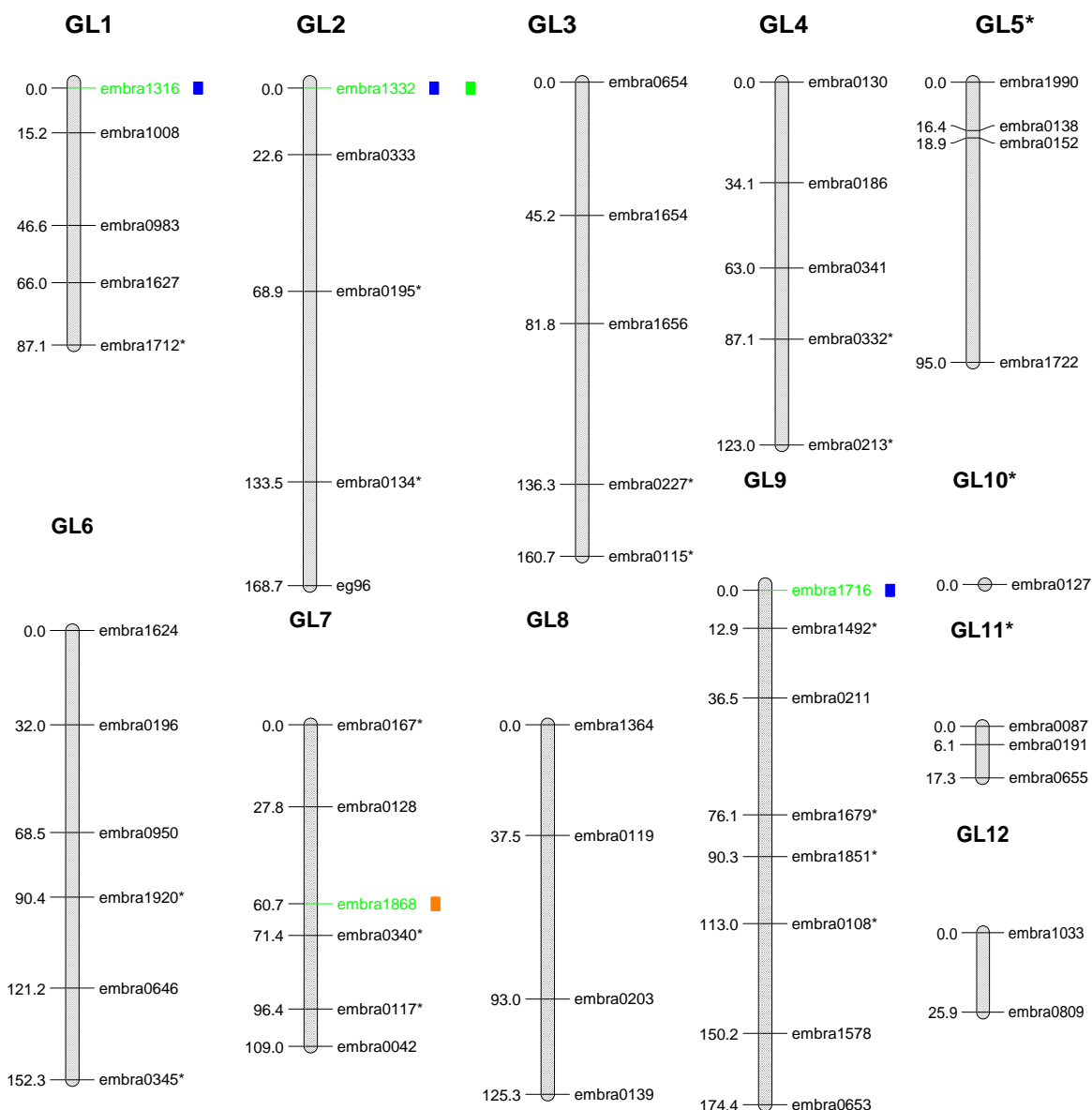
O rendimento de celulose médio nas árvores avaliadas nesta progênie foi 0,55, com variância de 0,96. Para este caráter foi detectado um QTL com interação significativa entre ambientes e três QTL na análise conjunta dos experimentos. Na análise individual dos experimentos foi detectado um QTL em cada ambiente avaliado. O loco associado a QTL no experimento de MG também apresentou interação significativa entre os dois ambientes, que será discutida com maiores detalhes na seção 6.3, durante a análise dos QTL significativos no experimento de MG. Os p-valores estimados para estes QTL se encontram disponíveis na Tabela 11, e a representação gráfica na Figura 13.

**Tabela 11.** P-valores e coeficientes de determinação para os locos que apresentaram associações a QTL para rendimento depurado de celulose (RD) em pelo menos uma das análises realizadas.

Loco	Conjunto		Experimento RS		Experimento MG		Interação
	p-valor	R <sup>2</sup> (%)	p-valor	R <sup>2</sup> (%)	p-valor	R <sup>2</sup> (%)	
Embra 1316	0,007**	1,40	0,014*	4,29	0,298	3,37	0,794
Embra 1332	0,004**	2,22	0,004**	5,89	0,556	1,17	0,435
Embra 1716	0,008**	3,01	0,025*	4,59	0,149	6,61	0,787
Embra 1868	0,523	2,70	0,223	4,39	0,000**	46,90	0,001**

\*: Significativo 0,95; \*\*: Significativo 0,99.

Na análise conjunta dos experimentos os locos Embra 1316, Embra1332 e Embra 1716 apresentaram associações significativas a QTL, embora estes locos juntos sejam capazes de explicar pouco mais de 6% da variação genética desta progênie. O loco Embra 1332 também apresentou associação significativa no experimento do RS, sendo capaz de explicar sozinho cerca de 6% da variação genotípica predita neste local. O loco Embra 1868 que apresentou interação significativa entre ambientes, também apresentou QTL na análise do experimento de MG, sendo capaz de explicar quase 50% da variação genética predita neste local.

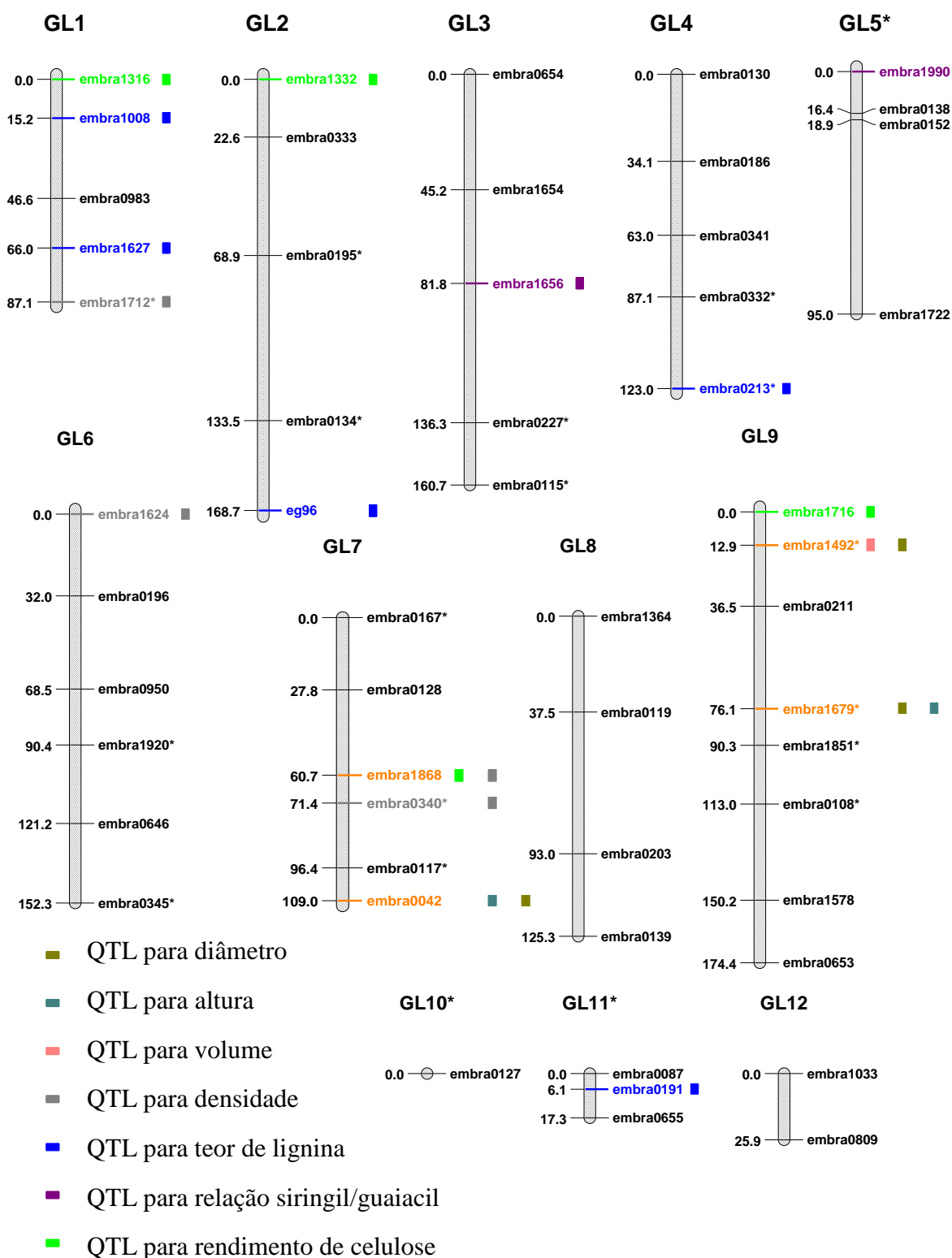


**Figura 13.** Mapa representativo dos QTL detectados para o caráter rendimento depurado de celulose - RD.

#### 4.6 DISTRIBUIÇÃO DOS QTL

A representação dos locos que se apresentaram associados a QTL para todos os caracteres que apresentaram QTL significativos pode ser visualizada na Figura 14. A representação gráfica dos QTL identificados para os diferentes caracteres avaliados na progênie DG x GL2 pode ser visualizada na Figura 14. Os quatro locos que apresentaram associação significativa com QTL para os caracteres de crescimento foram responsáveis pela identificação de dez QTL, indicando a existência de locos atuando no crescimento das árvores em diferentes sentidos. Três dos locos que apresentaram QTL significativos para a

densidade e o rendimento de celulose, foram localizados nos grupos de ligação 7 e 9, juntamente com os locos significativos para os caracteres de crescimento.



**Figura 14.** Mapa representativo de todos os locos onde foram identificados QTL para os caracteres avaliados. Os locos foram identificados com cores de acordo com o caráter associado ao QTL. Os locos que apresentaram QTL para mais de um caráter foram representados com a cor laranja.

Três QTL foram identificados nos grupos de ligação 1 e 2, juntamente com QTL significativos para o teor de lignina. No grupo de ligação 1 foram localizados dois locos associados a QTL para teor de lignina, um para a densidade e outro para rendimento de celulose. No grupo de ligação 2 foram localizados dois QTL, um para rendimento de celulose e outro para teor de lignina, posicionados em regiões extremas neste grupo de ligação. O teor de lignina apresentou cinco locos associados com QTL, sendo dois destes localizados no grupo de ligação 1, enquanto os demais se encontram distribuídos nos grupos de ligação 2, 4 e 11. A relação siringil/guaiacil apresentou QTL significativos nos grupos de ligação 3 e 5, onde não foram verificados QTL para nenhum outro caráter.

O posicionamento dos QTL identificados para os caracteres de crescimento somente em dois grupos de ligação e situados nos mesmos grupos de ligação onde foram identificados QTL para os caracteres rendimento de celulose e densidade da madeira, sugerem que os grupos de ligação 7 e 9 são extremamente relevantes para caracteres ligados ao crescimento e qualidade da madeira.

#### 4.7 COMPARAÇÃO ENTRE DIFERENTES PROCESSOS DE SELEÇÃO DE GENÓTIPOS SUPERIORES

Nesta seção foi realizada a avaliação da eficiência da SAM, através da comparação do processo de seleção de genótipos superiores tradicionalmente realizado em programas de melhoramento genéticos de espécies perenes com base nos valores eBLUP, com o processo de seleção de árvores com base nos genótipos superiores identificados com marcadores moleculares. Esta comparação foi conduzida incluindo todos os locos que apresentaram associação significativa a QTL para seus respectivos caracteres, na análise dos ambientes isolados e em conjunto.

Nos itens A, B e C da Tabela 12, inicialmente são apresentados os valores médios dos caracteres que apresentaram associação a QTL nas diferentes análises realizadas, seguidos do número de indivíduos selecionados para todos os QTL descritos para um determinado caráter na referida análise. Em seguida, encontram-se colunas com o valor médio dos N melhores indivíduos, selecionados tanto pelos valores e-BLUP quanto pela SAM para os QTL identificados, cada qual precedido pelo quanto a média destes indivíduos selecionados é superior em relação à média geral deste caráter na análise em questão ( $\Delta g$  %). Na última coluna é verificada a eficiência da SAM em relação à seleção

com base nos valores eBLUP. Informações detalhadas sobre os alelos dos locos que apresentaram associação a QTL encontram-se disponíveis no Apêndice F.

**Tabela 12.** Simulação de seleção e comparação da eficiência da SAM para os caracteres que apresentaram QTL significativos na análise isolada de cada ambiente (A e B) e conjunta das amostras coletadas nos dois experimentos (C).

A) Experimento RS:

Caráter	Média	Nº Ind. Sel.	Seleção eBLUP		SAM			Relação eBLUP/ SAM
			Média	$\Delta g$ (%)	Nº QTL	Média	$\Delta g$ (%)	
DAP	7,95	26	10,22	28,5	2	8,43	7,5	0,26
Altura	9,66	13	11,52	19,0	2	10,26	6,2	0,33
Volume	0,0353	30	0,0607	71,9	2	0,0526	49,0	0,68
DB	453,29	30	460,95	1,7	1	458,88	1,2	0,71
LigTotal	23,22	30	22,49	-3,3	1	23,11	-0,5	0,15
Rsg	3,32	30	3,55	6,9	1	3,35	0,9	0,25
RD	54,74	30	55,66	1,7	1	55,53	1,4	0,82

B) Experimento MG:

Caráter	Média	Nº Ind. Sel.	Seleção eBLUP		SAM			Relação eBLUP/ SAM
			Média	$\Delta g$ (%)	Nº QTL	Média	$\Delta g$ (%)	
DB	472,18	1	490,03	3,8	2	474,23	0,4	0,11
LigTotal	24,08	12	23,30	-3,3	2	23,61	-2,0	0,61
RD	56,58	18	57,01	0,8	1	56,87	0,5	0,63

C) Análise conjunta dos experimentos:

Caráter	Média	Nº Ind. Sel.	Seleção eBLUP		SAM			Relação eBLUP / SAM
			Média	$\Delta g$ (%)	Nº QTL	Média	$\Delta g$ (%)	
DAP	7,84	23	10,41	32,8	2	7,93	1,2	0,04
Altura	9,76	30	11,34	16,2	1	10,48	7,4	0,46
Volume	0,0342	30	0,0623	82,2	1	0,0468	36,8	0,45
DB	456,98	30	473,20	3,5	2	459,88	0,6	0,17
LigTotal	23,39	30	22,49	-4,0	2	22,88	-2,2	0,55
rSG	3,26	30	3,55	8,9	1	3,50	7,4	0,83
RD	55,09	29	56,84	3,2	3	55,45	0,7	0,22

O número máximo de árvores selecionadas pelos dois métodos de seleção não ultrapassou de trinta indivíduos, valor que corresponde a uma intensidade de seleção de 15% em relação ao tamanho da população de mapeamento avaliada neste trabalho. Com exceção do caráter pilodyn que não apresentou associação a QTL, todos os demais caracteres apresentaram associação a QTL em alguma das análises realizadas (ambientes isolados ou em conjunto). A relação siringil/guaiacil e os caracteres de crescimento não apresentaram associação a QTL no experimento de MG, enquanto os caracteres densidade, teor de lignina e rendimento de celulose apresentaram QTL em todas as análises realizadas.

Para o caráter densidade, a inclusão dos dados do experimento de MG junto aos do RS aumentou de um para dois o número de QTL significativos na análise conjunta dos dados em relação à análise individual do experimento do RS, embora a seleção simultânea para estes dois locos tenha reduzido os possíveis ganhos genéticos pela metade e a eficiência da SAM de 71% para 17% entre estas análises. A eficiência da SAM para densidade em MG foi a menor de todas, devido principalmente à seleção simultânea para os dois locos que apresentaram QTL significativos neste ambiente ter reduzido o número de árvores selecionadas para um único exemplar, em função do pequeno número de indivíduos do cruzamento DG x GL2 que sobreviveram neste ambiente.

Já para os caracteres teor de lignina e rendimento de celulose, que apresentaram interação entre QTL e ambientes significativa, foi verificada maior eficiência da SAM na análise do experimento de MG em relação à análise conjunta dos experimentos, mesmo havendo um menor número de amostras avaliadas neste ambiente. A inclusão das amostras de MG na análise conjunta com as do RS proporcionou um aumento na eficiência da SAM para o caráter teor de lignina e uma redução na eficiência para o caráter rendimento de celulose, ambas na ordem de aproximadamente quatro vezes em relação à análise isolada do experimento do RS.

O caráter rendimento de celulose apresentou a menor eficiência da SAM na análise conjunta dos experimentos devido à identificação de três QTL significativos nesta análise, em detrimento de somente um QTL identificado nas análises isoladas dos experimentos. No entanto, para o caráter teor de lignina foram identificados dois QTL na análise conjunta dos dados, capazes de proporcionar ganhos genéticos superiores em relação à análise isolada do experimento do RS onde foi identificado somente um QTL.

Para os caracteres de crescimento que não apresentaram QTL na análise de MG, foram verificados dois QTL para cada caráter na análise do experimento do RS. O



aumento do número QTL significativos no RS somente não ocasionou uma redução do número de árvores selecionadas para o caráter volume, pois proporcionou a seleção de trinta árvores para os dois locos. Para os demais caracteres de crescimento avaliados no RS, bem como para o caráter DAP avaliado na análise conjunta dos dois ambientes, a seleção simultânea de genótipos superiores para mais de um loco restringiu o número de árvores selecionadas. A inclusão dos dados do experimento de MG junto ao do RS reduziu o número de QTL identificados na análise conjunta dos dados dos caracteres altura e volume. No entanto, esta redução na quantidade de QTL identificados para o caráter altura possibilitou a seleção de uma quantidade maior de indivíduos na análise conjunta, o que possibilitou que os ganhos genéticos nesta análise fossem superiores em relação aos obtidos na análise individual do experimento do RS, onde somente treze árvores seriam selecionadas para os dois QTL identificados neste ambiente.

No caráter volume a redução do número de QTL na análise conjunta dos experimentos em relação à análise individual do RS não proporcionou aumento da eficiência ou dos ganhos genéticos capazes de serem obtidos, devido aos dois QTL identificados para este caráter no experimento do RS não terem restringido o montante de árvores selecionadas. O mesmo padrão foi verificado para o caráter DAP, onde a análise do RS também apresentou uma diferença aproximadamente 20% superior em relação à análise conjunta dos dados. O caráter DAP apresentou dois QTL tanto na análise do RS quanto na análise conjunta dos experimentos, mas a eficiência da SAM foi superior para o experimento do RS devido aos dois QTL identificados neste ambiente terem restringido menos o montante de árvores capazes de serem selecionadas.

Apesar da SAM apresentar melhores índices de eficiência para os caracteres DAP, volume, densidade e rendimento de celulose na análise isolada do RS, a eficiência da SAM verificada para o caráter DAP foi de somente 26%, enquanto para os demais caracteres ela foi superior a 50%. Os caracteres altura e relaçãoiringil/guaiacil apresentaram melhor eficiência na SAM na análise conjunta dos dados, apesar de somente o segundo caráter ter apresentado eficiência acima de 50%. O caráter teor de lignina apresentou maior eficiência na SAM na análise isolada do experimento de MG, também com eficiência superior a 50%.

Os ganhos possíveis de serem obtidos para os caracteres avaliados na análise de MG foram ligeiramente inferiores aos capazes de serem obtidos para estes mesmos caracteres na análise conjunta dos experimentos. No entanto, a eficiência da SAM para este

ambiente só não foi superior para o caráter densidade, que apresentou diferença de 0,06. Para o teor de lignina a eficiência da seleção para os dados de MG foi 0,06 acima à da seleção conjunta. Para o caráter rendimento de celulose, a relação eBLUP/SAM apresentou eficiência superior na análise dos experimentos isolados (0,63 em MG e 0,82 no RS) em relação à análise conjunta (0,22), indicando que existe possibilidade de utilizar a informação destes QTL para realizar seleção direcionada para cada um dos ambientes avaliados, de maneira independente.

## 5 DISCUSSÃO

Nesta seção os resultados obtidos neste trabalho serão comparados com os descritos por outros autores que realizaram trabalhos similares em *Eucalyptus*, citados na revisão bibliográfica. Inicialmente serão discutidos os perfis dos caracteres quantitativos analisados, para em seguida serem comentados o processo de construção do mapa genético integrado os efeitos do tamanho da amostra do experimento de MG nos QTL e interações identificadas envolvendo este local, que foram abordados separadamente. Posteriormente será realizada uma validação dos QTL identificados nas análises de mapeamento através da comparação com QTL descritos por outros autores. Antes de concluir foram realizadas algumas considerações sobre a aplicação da SAM em programas de melhoramento genético florestal.

### 5.1 CARACTERES QUANTITATIVOS

#### 5.1.1 Coeficiente de variação do erro experimental - CVe

O coeficiente de variação do erro experimental é uma medida que indica o quanto o delineamento das estruturas experimentais em campo foi eficiente na captação de um montante da variância total observada (estruturas de superblocos e bloquinhos neste caso). Com exceção do caráter DAP avaliado no experimento do RS que apresentou CVe de 0,21, todos os demais caracteres avaliados nos dois experimentos apresentaram CVe abaixo de 0,15, o que caracteriza que as estruturas experimentais foram bem definidas nos experimentos de campo (Tabelas 2 e 3). Os valores superiores dos CVe verificados para os caracteres de crescimento em relação aos avaliados por NIRS, provavelmente são decorrentes dos erros inerentes ao método de avaliação, que estão mais sujeitos à presença de imprecisões durante o processo de coleta de dados em campo.

Os CVe dos caracteres de crescimento apresentaram uma tendência entre os dois ambientes, sendo observado valores superiores para o experimento do RS em relação ao de MG. Já os caracteres de qualidade da madeira estimados por NIRS, apresentaram CVe de pequena magnitude e idênticos entre os experimentos do RS e de MG para cada

caráter avaliado. Acredita-se que esta coincidência entre os valores dos CVe para estes caracteres sejam decorrentes das curvas de calibrações utilizadas para estimar os parâmetros por NIRS, uma vez que foram utilizadas as mesmas curvas de calibração para avaliar as amostras de madeira coletadas nos dois locais. O baixo valor do CVe para os caracteres avaliados por NIRS é um indicativo de que embora as curvas de calibrações utilizadas não sejam as melhores possíveis de serem ajustadas, elas foram eficientes no sentido de permitir que as estruturas experimentais captassem o máximo possível da variação experimental destes caracteres, contribuindo para a redução do coeficiente de variação do erro experimental (CVe).

### **5.1.2 Contribuição relativa das fontes de variância na expressão dos caracteres**

Ao analisar a contribuição relativa das fontes de variância na expressão dos caracteres de crescimento na Figura 2, verificam-se similaridades entre os diferentes componentes para cada um dos três caracteres de crescimento avaliados nos dois ambientes. Já os caracteres estimados por NIRS apresentados na Figura 4 não apresentaram padrões similares para os componentes ambientais e de controle local entre os dois experimentos.

A influência relativa do componente de variação ambiental na expressão dos caracteres de crescimento apresentou diferença entre os dois ambientes de no máximo 0,08, enquanto para os caracteres avaliados por NIRS esta diferença chegou a até 0,40 para um mesmo caráter entre os dois ambientes. Dentre os caracteres avaliados por NIRS somente a relação siringil/guaiacil apresentou valores semelhantes para o componente de variação ambiental entre os dois experimentos, com diferença de apenas 0,04. A influência relativa do componente de variação ambiental na expressão dos caracteres avaliados no experimento do RS variou de 0,26 a 0,55, enquanto no experimento de MG os valores deste componente variaram entre 0,31 e 0,85.

Conforme apresentado nas Figuras 2, 3 e 4, mesmo para os caracteres que consideraram as estruturas de bloquinhos dentro de superblocos significativas (Tabela 1), o componente de controle local apresentou de maneira geral pequeno efeito na captação de uma parcela da variância total. Somente para o caráter densidade no experimento do RS e para a relação siringil/guaiacil no experimento de MG, o controle local conseguiu captar uma parte considerável da variação observada (acima de 0,40). Para os caracteres rendimento de celulose e relação siringil/guaiacil no experimento do RS, e para o caráter

teor de lignina em MG, os fatores de controle local ainda conseguiram captar uma boa parte da variância observada (entre 0,10 e 0,25). Para os demais caracteres avaliados por NIRS e todos os caracteres de crescimento, independente de considerarem as estruturas de bloquinhos dentro de superblocos significativas ou não, o montante da variância total que estas estruturas juntas conseguiram captar não chegou a 0,10.

O desejável seria que os componentes de controle local (estruturas de blocos e bloquinhos) captassem o máximo possível da variância presente nos caracteres avaliados, reduzindo a influência do componente ambiental devido à existência de fontes de variância associadas à estrutura experimental. Este cenário levaria tanto a um aumento da precisão experimental (C<sub>Ve</sub>), quanto das estimativas da contribuição do componente genético na expressão dos caracteres avaliados, incluindo os efeitos que os QTL identificados seriam capazes de explicar. No entanto, embora os fatores de controle local tenham captado somente uma pequena parte da variância na maioria dos caracteres avaliados nos dois ambientes, estes fatores foram significativos para todos os caracteres avaliados, pois pelo menos a estrutura de superblocos esteve presente nos modelos mistos selecionados para prever os valores genotípicos (eBLUP) destes caracteres (Tabela 1).

### **5.1.3 Coeficientes de herdabilidade - $h^2$**

Os valores dos coeficientes de herdabilidade estimados para os caracteres de crescimento foram similares entre os dois locais e compatíveis com valores existentes em literatura, conforme serão apresentados a seguir. Dentre os caracteres avaliados por NIRS, o teor de lignina e a densidade apresentaram coeficientes de herdabilidade similares entre os dois ambientes, enquanto os coeficientes de herdabilidade dos caracteres rendimento de celulose e relação siringil/guaiacil apresentaram valores diferenciados, sendo verificados valores superiores no experimento do RS em relação aos de MG para estes caracteres. Segundo Paula (1995), as estimativas de herdabilidade obtidas sempre estão associadas aos erros presentes nas estimativas das variáveis, e por isso, estas somente fornecem uma indicação relativa do controle genético do caráter, não devendo ser interpretada como valor absoluto ou invariante.

De acordo com Cornelius (1993), em espécies florestais a herdabilidade de caracteres ligados ao crescimento como DAP, altura e volume variam entre 0,1 e 0,3. Segundo Resende (2001), os coeficientes de herdabilidade no sentido restrito de características ligadas ao crescimento apresentam valores próximos a 0,30, enquanto Tolfo

et al. (2005) estimaram os coeficientes de herdabilidade no sentido amplo em 0,69, 0,48 e 0,70, respectivamente para os caracteres DAP, altura e volume. Os valores dos coeficientes de herdabilidade no sentido amplo obtidos nas análises dos experimentos avaliados neste trabalho foram semelhantes aos descritos por Tolfo et al. (2005), sendo que os valores médios entre os dois ambientes foram de 0,63 para os caracteres DAP e altura, e 0,50 para o volume.

O caráter densidade avaliado por pilodyn apresentou coeficiente de herdabilidade praticamente idêntico ao descrito em um trabalho onde este caráter foi avaliado em um experimento contendo várias populações de *Eucalyptus* (Bouvet, 1995). O autor estimou a herdabilidade do caráter densidade mensurado com pilodyn em torno de 0,40, enquanto o valor verificado para este caráter neste trabalho foi de 0,39. No entanto, a densidade que é considerada um caráter de alta herdabilidade (Ribeiro & Filho, 1993), apresentou neste trabalho coeficientes de herdabilidade próximos a 0,10 para os dois locais onde a progênie foi avaliada com NIRS, valores diferenciados em relação aos descritos em literatura. Grattapaglia et al. (1996) consideraram em seu trabalho de mapeamento de QTL a herdabilidade do caráter densidade entre 0,5 e 0,7, enquanto Tolfo et al. (2005) estimaram a herdabilidade no sentido amplo do caráter densidade em 0,56.

Os coeficientes de herdabilidade do caráter teor de lignina apresentaram valores muito próximos entre os dois ambientes avaliados neste trabalho, sendo estimado em 0,46 no experimento do RS e em 0,44 no experimento de MG. Garcia (1998) e Tolfo et al. (2005) ao avaliar experimentos independentes onde foram realizadas análises físico-químicas da madeira, estimaram a herdabilidade no sentido amplo do caráter teor de lignina em 0,84. Para a relação sirigil/guaiacil ainda não haviam sido descritos estimativas de herdabilidade, e no presente trabalho foram verificados valores distintos entre os dois experimentos (0,42 no RS e 0,28 em MG).

Embora o caráter rendimento de celulose também tenha apresentado valores diferenciados nas estimativas de herdabilidade entre os dois experimentos (0,16 em MG e 0,35 no RS), estes valores foram aproximados em relação a outros já descritos em literatura. Demuner & Bertolucci (1992), citados por Grattapaglia et al. (1996), estimaram a partir de avaliações físico-químicas a herdabilidade do caráter rendimento depurado de celulose em 0,20, valor próximo ao valor médio verificado entre os dois ambientes neste experimento (0,25). No entanto, Tolfo et al. (2005) ao realizar avaliações físico-químicas de diferentes progênies, estimaram a herdabilidade no sentido amplo deste caráter em 0,86.

A coincidência dos valores dos coeficientes de herdabilidade verificada para os caracteres densidade e teor de lignina entre os dois ambientes, é um indicativo da boa qualidade das estimativas obtidas para os caracteres avaliados com NIRS. No entanto, acredita-se que as diferenças entre os coeficientes de herdabilidade verificados neste trabalho para estes caracteres com os descritos por outros autores, também tenham origem nas curvas de calibrações utilizadas na obtenção de parâmetros utilizando esta tecnologia.

As curvas de calibrações utilizadas para estimar os caracteres relacionados com a qualidade da madeira no Projeto Genolyptus não se encontram perfeitamente otimizadas para cada condição experimental onde estes caracteres foram avaliados. Este desajuste pode levar a uma superestimativa do efeito do componente ambiental na expressão dos caracteres avaliados com esta tecnologia, reduzindo o efeito do componente genético na expressão dos caracteres avaliados. As curvas de calibrações utilizadas no Projeto Genolyptus se encontram em fase de ajuste, para que futuramente possam ser obtidas estimativas mais precisas para parâmetros mensurados em diferentes ambientes com tecnologia NIRS (Coelho, comunicação pessoal).

#### **5.1.4 Correlações entre caracteres quantitativos - $r_{xy}$**

Observou-se uma tendência para os caracteres que apresentaram correlações significativas nos dois conjuntos de dados, de que os valores observados nas correlações fenotípicas foram superiores em relação aos verificados nas correlações genotípicas, tanto para os caracteres de crescimento quanto para os avaliados com NIRS. Para as correlações recíprocas envolvendo o caráter pilodyn foi verificado o inverso, sendo observados valores superiores nas correlações genotípicas em relação às fenotípicas.

De acordo com Rocha et al. (2008), entre os caracteres de crescimento são esperadas correlações elevadas e positivas. Neste trabalho os caracteres de crescimento (DAP, altura e volume) apresentaram correlações fenotípicas e genotípicas positivas e acentuadas, acima de 0,8. Novaes (2006) verificou correlação fenotípica de 0,72 entre DAP e altura. No entanto, estas correlações elevadas não necessariamente indicam a existência de genes de efeito pleiotrópico controlando estes caracteres, pois estes caracteres são extremamente relacionados entre si. De acordo com Breseguello (comunicação pessoal), o conceito de pleiotropia é melhor aplicado para caracteres não diretamente relacionados entre si, tal qual são os caracteres de crescimento.

Dentre os caracteres de crescimento somente o caráter altura apresentou correlação genotípica com as propriedades da madeira estimadas por NIRS, sendo todas elas de intensidade fraca e em torno de 20%. No entanto, este resultado indica que é possível se obter pequenos progressos genéticos para as propriedades da madeira através da seleção indireta para a variável de crescimento altura.

Para os caracteres avaliados por NIRS, Rocha et al. (2008) descrevem que correlações acentuadas e negativas entre o rendimento de celulose e o conteúdo de extrativos ou o teor de lignina insolúvel, indicam boa qualidade nos dados e curvas de calibrações utilizadas para mensurar estes caracteres. Barrichelo et al. (1983), Perez (2002), Novaes (2006) e Tolfo et al. (2005) reportaram em seus estudos correlações fenotípicas negativas entre rendimento de celulose e teor de lignina. No presente estudo a correlação fenotípica entre o rendimento de celulose e o teor de lignina foi negativa e de intensidade moderada (próxima a -0,50), embora esta correlação não tenha sido verificada para o conjunto de dados genotípicos. De maneira contrária, a correlação entre o rendimento de celulose e a relação siringil/guaiacil foi significativa somente para os dados genotípicos, também se apresentando negativa e de pequena magnitude (próxima a -0,25). Conforme esperado foram verificadas correlações robustas e negativas entre o teor de lignina e a relação siringil/guaiacil, sendo verificados valores negativos e superiores a 0,50 nos dois conjuntos de dados. O teor de lignina ainda apresentou no experimento do RS correlações positivas e de pequena magnitude com a densidade avaliada por pilodyn, sendo verificados valores genotípicos e fenotípicos inferiores a 0,20.

Verhaegen et al. (1997) verificaram em seu trabalho no qual foi avaliado uma progênie de híbridos interespecíficos, correlações significativas e positivas entre os caracteres de crescimento e a penetração da haste do pilodyn, e negativas entre os caracteres de crescimento e a densidade estimada com ensaios mecânicos. Dean et al. (1990) observaram em *E. globulus* correlação de 0,93 entre pilodyn e densidade avaliada com testes mecânicos. O caráter densidade avaliado por pilodyn neste trabalho apresentou correlações fenotípicas e genotípicas positivas significativas e de intensidade moderada com os caracteres de crescimento (todas próximas a 0,50), embora não se tenha verificado correlações significativas com a densidade estimada por NIRS.

Estes resultados fundamentam a hipótese de que quanto maior a taxa de crescimento de uma árvore, menor será a densidade de sua madeira, e quanto menor a densidade, maior será a penetração da haste do pilodyn. No entanto, os resultados obtidos



neste trabalho concordam com os apresentados no trabalho de Novaes (2006), que ao avaliar uma progênie de híbridos interespecíficos também não verificou correlações significativas entre pilodyn e densidade.

Existem duas hipóteses, não mutuamente excludentes, que podem explicar os resultados verificados nas correlações envolvendo o caráter densidade avaliado por pilodyn. A primeira é de que a região do tronco avaliada por este equipamento não é representativa em relação à densidade da tora, uma vez que a densidade apresenta variação ao longo do tronco de uma mesma árvore, o que possibilita que sejam obtidos valores diferenciados de acordo com a região do tronco na qual o operador do equipamento realizou a medida. A segunda é de que este equipamento não é eficiente na mensuração da densidade de todas as espécies do gênero *Eucalyptus*, sendo mais preciso na estimativa da densidade das espécies que possuem casca delgada em relação às espécies que possuem casca espessa.

Estas particularidades entre as diferentes espécies de *Eucalyptus* podem explicar os efeitos das mensurações realizadas com este aparelho serem significativas para algumas espécies e independentes para outras, e também a dificuldade em se obter correlações consistentes em progênies de híbridos interespecíficos, onde uma única progênie pode segregar para diferentes fenótipos. Progênies híbridas podem apresentar variabilidade superior para o caráter espessura da casca em relação às progênies obtidas a partir do cruzamento de indivíduos de uma mesma espécie, que aliado com a imprecisão inerente às dificuldades operacionais deste equipamento, dificulta a obtenção de dados consistentes nas mensurações realizadas utilizando o equipamento pilodyn.

A densidade avaliada por NIRS apresentou correlações fenotípicas e genotípicas com a relação siringil/guaiacil, embora a fenotípica tenha se apresentado positiva e acentuada ( $r_{xy} = 0,50$ ), enquanto a genotípica foi negativa e com o valor mais próximo do limite de significância dentre todas as correlações genotípicas significativas verificadas neste trabalho ( $r_{xy} = -0,17$ ). Novaes (2006) também verificou nos caracteres avaliados por NIRS em seu trabalho correlações fenotípicas positivas entre a densidade e a relação siringil/guaiacil ( $r_{xy} = 0,16$ ) e entre a densidade e o diâmetro ( $r_{xy} = 0,26$ ). Ainda foram verificadas correlações genotípicas negativas e de pequena magnitude entre a relação siringil/guaiacil com o rendimento de celulose e a densidade (próximas a  $-0,20$ ), e também correlações positivas entre a densidade com os caracteres teor de lignina e rendimento de celulose.

Conforme comentado anteriormente, embora as curvas de calibração NIRS não estejam perfeitamente ajustadas para cada ambiente onde os experimentos foram avaliados, algumas das correlações verificadas entre os caracteres avaliados, juntamente com os resultados verificados nas análises dos experimentos, indicam que as estimativas obtidas por NIRS no Projeto Genolyptus foram confiáveis.

## 5.2 CONSTRUÇÃO DO MAPA GENÉTICO

### 5.2.1 Distorção de segregação

Dentre os 76 locos avaliados, 18% (14 locos) apresentaram distorção de segregação significativa com um nível de significância de 5% para o Critério de Bonferroni, e estes locos foram excluídos das análises de mapeamento. Vários autores que construíram mapas genéticos descrevem em seus trabalhos que a taxa de locos que apresentam distorção de segregação significativa pode variar de acordo com a natureza do polimorfismo detectado pelos marcadores moleculares utilizados (dominante ou codominante), e também de acordo com o tipo de cruzamento avaliado.

Em *Eucalyptus* são verificados percentuais de distorção superiores nos cruzamentos interespecíficos em relação aos cruzamentos intraespecíficos. No entanto, ao analisar os dados deste e de outros trabalhos, verifica-se que a taxa de locos que apresentam distorção de segregação significativa também é afetada pelo processo de *screening* de *primers*, onde os mais informativos são selecionados e os que amplificam regiões inespecíficas ou não informativas são descartados, conforme será discutido a seguir.

Marques et al. (1998), ao analisarem uma progênie de híbridos interespecíficos obtida do cruzamento entre as espécies *E. tereticornis* e *E. globulus* com marcadores AFLP, observaram 15% dos locos avaliados com distorção de segregação significativa, percentual semelhante ao verificado neste trabalho. Os mapas genéticos construídos por Myburg et al. (2003), utilizando marcadores AFLP, apresentaram cerca de 30% e 26% dos locos com distorção de segregação para os parentais *E. globulus* e *E. grandis*, valores acima do observado neste trabalho.

Os mapas genéticos de Gan et al. (2003) construídos com base em marcadores RAPD avaliando uma progênie híbrida entre as espécies *E. urophylla* e *E. tereticornis*, apresentaram cerca de 15% e 35% das marcas avaliadas para estes parentais com distorção

de segregação significativa, valores com média superior ao percentual verificado neste trabalho. Nos trabalhos destes autores não existem relatos sobre a realização de processos de seleção de marcas informativas.

Grattapaglia & Sederoff (1994) selecionaram dentre 305 *primers* RAPD disponíveis, 151 para amplificar as marcas utilizadas na construção de seus mapas genéticos, observando distorção de segregação em somente 10 marcas polimórficas para cada parental, equivalente a menos de 4% das marcas avaliadas. Thamarus et al. (2002) ao avaliarem uma progênie pura de *E. globulus*, identificaram apenas duas marcas com distorção de segregação significativa, dentre as 249 avaliadas com diferentes tipos de marcadores moleculares. Nos trabalhos descritos acima e também nos de Brondani et al. (1998; 2002; 2006) e Agrama et al. (2002), foi realizado um intenso processo de seleção de marcas informativas, e por este motivo não são relatados os percentuais de locos que apresentaram distorção de segregação.

Assim, embora o percentual de locos que apresentaram distorção de segregação verificado neste trabalho (18%) tenha sido superior ao de alguns outros citados na literatura, podemos considerar que ele foi dentro do esperado em um trabalho onde foi realizado um processo rigoroso de seleção de locos SSR informativos, para avaliar uma progênie obtida do cruzamento de um genitor híbrido interespecífico (*E. dunnii* x *E. grandis*), com uma outra espécie pura (*E. globulus*). Conforme os dados de polimorfismo dos 62 locos que não apresentaram distorção de segregação significativa na progênie DG x GL2 (Apêndice C), verifica-se que o genitor híbrido foi polimórfico para 87% das marcas avaliadas, apresentando polimorfismo quase duas vezes superior em relação ao genitor puro de *E. globulus*, que foi polimórfico somente para 47% dos locos avaliados.

Estes resultados concordam com os resultados obtidos por Cervera et al. (2000) e Bison et al. (2007), que indicam a existência de polimorfismo superior em populações híbridas em relação às populações originadas do cruzamento de parentais da mesma espécie. Como a estratégia de pseudo-cruzamento teste utiliza o polimorfismo presente dentro de cada genitor para mapear QTL, o polimorfismo presente no genitor híbrido é capaz de representar os QTL existentes nas espécies *E. dunnii* e *E. grandis*, que deram origem ao parental híbrido avaliado neste cruzamento.

### 5.2.2 Mapa genético integrado

A natureza codominante e multialélica dos marcadores SSR aumenta o poder de detecção de polimorfismos relevantes para a construção de mapas genéticos em espécies florestais (Grattapaglia & Sederoff, 1996). Além disso, os marcadores SSR apresentam altas taxas de transferibilidade entre espécies do mesmo gênero (Brondani et al., 2002), permitindo que grupos de ligação mapeados em diferentes cruzamentos possam ser facilmente relacionados. Os grupos de ligação formados apresentaram consistência com os verificados nos mapas genéticos construídos também utilizando marcadores SSR nos trabalhos de Brondani et al. (2002; 2006), Missiaglia (2005), Novaes (2006), e outros três ainda não publicados, construídos no âmbito do Projeto Genolyptus.

Devido aos trabalhos de Brondani et al. (1998; 2002; 2006) estarem avaliando pela primeira vez pares de *primers* SSR com alto conteúdo informativo, da mesma forma que não foram descritos os percentuais de locos que não amplificaram alelos adequadamente, não são relatados nestes trabalhos os percentuais de locos que permaneceram desconectados após as análises de mapeamento. Aplicando os critérios de LOD 3 e fração de recombinação de 0,40 foram obtidos grupos de ligação consistentes com os trabalhos descritos por outros autores, embora quase metade dos locos avaliados (33) não ter apresentado suporte estatístico para ligação na análise de mapeamento. Utilizando informações disponíveis em outros mapas onde alguns destes locos foram avaliados, 75% destes locos (25) tiveram suas posições estimadas no mapa genético integrado construído neste trabalho. Este mapa integrado foi composto por 12 grupos de ligação, com comprimento total de 1240 cM e distância média entre marcas de 22 cM, não sendo inclusos os marcadores que apresentaram distorção de segregação significativa.

Nos trabalhos de Missiaglia (2005) e Novaes (2006) foram utilizados exclusivamente marcadores SSR para a construção de mapas genéticos e mapeamento de QTL para características de crescimento e qualidade da madeira em *Eucalyptus*. Os trabalhos destes autores apresentaram grande diferença com relação o percentual de locos que apresentaram distorção de segregação significativa, e a inclusão destes locos nas análises de mapeamento, tiveram influência nos valores de LOD que precisaram de ser utilizados nas análises de mapeamento, para se obter uma quantidade de grupos de ligação compatível com os existentes no gênero *Eucalyptus*.

Missiaglia (2005) identificou 70% dos locos com distorção de segregação significativa e não os excluiu das análises de mapeamento. Este autor teve de utilizar LOD score 10 e fração de recombinação de 0,40 para separar os 169 locos SSR avaliados em 11 grupos de ligação, permanecendo cerca de 8% dos locos avaliados desconectados após as análises de mapeamento. Novaes (2006), aplicando a Proteção de Bonferroni com um nível de significância de 5%, verificou que entre 235 locos SSR avaliados em seu trabalho, somente 4% apresentaram distorção de segregação significativa. Este autor excluiu estes locos das análises de mapeamento e utilizou LOD 4 com fração de recombinação de 0,35 para separar os locos polimórficos para os parentais de sua progênie em 11 grupos de ligação, onde somente um marcador permaneceu desconectado em relação aos demais.

Myburg et al. (2003), utilizando LOD 10 e 12, obtiveram 11 grupos de ligação nos mapas construídos para cada um dos parentais analisados em seu trabalho, com cobertura genômica acima de 95%, e distância média entre marcas próxima a 10 cM. Estes autores não retiraram os marcadores com distorção de segregação significativa das análises de mapeamento, e o número de marcas AFLP que permaneceram desconectadas após as análises de mapeamento foi 4. Dentre os 227 marcadores RAPD e RFLP avaliados no trabalho de Agrama et al. (2002), 26% permaneceram desconectados após as análise de mapeamento e foram descartados das análises de mapeamento, que utilizou LOD 3 para separar os 11 grupos de ligação de um cruzamento puro de *E. camaldulensis*.

Os mapas genéticos de *Eucalyptus* construídos por Grattapaglia & Sederoff (1994) utilizaram LOD 5 e fração de recombinação de 0,25 para agrupar mais de 500 marcas RAPD, obtendo 14 grupos de ligação no genitor *E. grandis* e 11 grupos de ligação no genitor *E. urophylla*. Após as análises de mapeamento, 4,5% das marcas avaliadas permaneceram desconectadas em relação aos grupos de ligação obtidos, valor inferior ao verificado neste trabalho. No trabalho de Gan et al. (2003), onde também foram utilizados em média 250 marcadores RAPD na construção dos mapas genéticos para os parentais avaliados em seu cruzamento, cerca de 16% dos locos permaneceram desconectados após as análises de mapeamento, valor próximo ao verificado neste trabalho.

Marques et al. (1998), utilizando LOD 5 e fração de recombinação de 0,25, identificaram 14 e 16 grupos de ligação para os parentais *E. tereticornis* e *E. globulus* avaliados em seu trabalho. Os mapas genéticos construídos por estes autores apresentaram comprimento médio próximo a 950 cM, enquanto a taxa de locos que permaneceram desconectados após as análises de mapeamento foi de aproximadamente 5%. A cobertura

genômica verificada no presente trabalho foi superior em relação à do trabalho citado anteriormente neste parágrafo, mesmo tendo sido avaliado um menor número de marcas que estes autores citados. Este fato indica que os locos SSR avaliados neste trabalho encontram-se bem distribuídos ao longo do genoma das espécies envolvidas neste cruzamento.

Thamarus et al. (2002), utilizando diferentes classes de marcadores, obtiveram um mapa genético integrado para uma progênie de *E. globulus* com características similares às observadas neste trabalho de mapeamento. No mapa destes autores foram obtidos 12 grupos de ligação apresentando uma cobertura de 1375 cM, valores próximos aos verificados no presente trabalho. No entanto, ao comparar o número de locos utilizados nas análises de mapeamento destes autores com o presente trabalho, verifica-se que a distância média entre marcas é relativamente proporcional ao número de locos analisados. Estes autores avaliaram 249 marcas moleculares, obtendo distância média entre marcas de 6 cM, enquanto que os 62 locos mapeados neste trabalho apresentaram distância média entre marcas próxima a 22 cM, grandezas inversamente proporcionais na ordem de proximadamente quatro vezes.

Embora o número de grupos de ligação descritos nos trabalhos de diferentes autores apresentem pequenas variações, a obtenção de 12 grupos de ligação no programa TREEMAP empregando LOD 3 e fração de recombinação de 0,40 foi confiável e representativa, em um trabalho onde foram avaliados poucos locos e os marcadores que apresentaram distorção de segregação significativa perante a Proteção de Bonferroni foram descartados. A alocação das marcas que permaneceram desconectadas nas análises de mapeamento, utilizando dados de outros mapas genéticos, se fez necessária devido à quantidade pequena de locos avaliados, e ao fato de que alguns dos locos inéditos se conectaram a outros locos já mapeados anteriormente, que também haviam permanecido desconectados inicialmente.

Desta forma, acredita-se que se tivesse sido possível avaliar um maior número de locos, menos marcas iriam permanecer desconectadas nas análises de mapeamento, devido ao aumento da saturação do mapa genético integrado construído reduzir a distância média entre as marcas avaliadas. Mesmo assim, a construção de um mapa genético integrado para a progênie segregante proporcionou obter uma cobertura genômica superior em relação à construção de mapas genéticos individuais para cada um dos parentais avaliados.

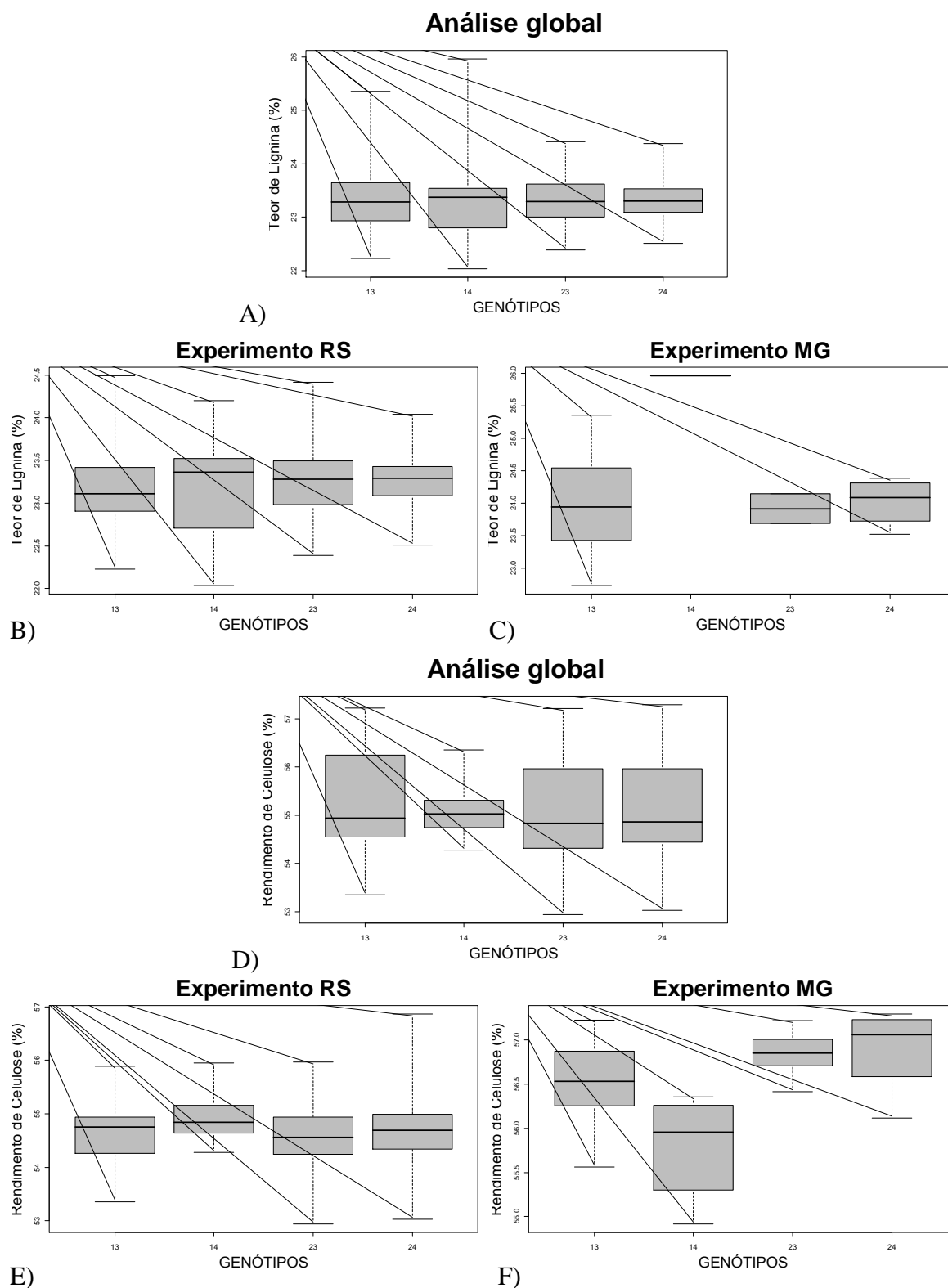
### 5.3 EFEITO DO TAMANHO DA AMOSTRA DO EXPERIMENTO DE MG NOS QTL E INTERAÇÕES IDENTIFICADAS

Os QTL identificados no experimento de MG e os locos que apresentaram interação entre QTL e ambientes serão discutidos à parte nesta seção, pois o alto índice de mortalidade que ocorreu neste ambiente teve influência direta na magnitude dos QTL identificados e na significância das interações envolvendo este local. O efeito diferencial entre os alelos mapeados na análise conjunta e individual dos ambientes para os locos que apresentaram interação entre ambientes significativa pode ser visualizado na Figura 15, enquanto informações detalhadas destes locos encontram-se no Apêndice G.

De acordo com Resende & Duarte (2007), o erro amostral reduz na medida em que se aumenta o tamanho da amostra. Desta forma, devido ao reduzido número de amostras avaliadas no experimento de MG, os QTL identificados nas análises isoladas deste local possuem maior erro em suas estimativas. Os erros ocasionados em virtude do pequeno número de amostras avaliadas neste ambiente podem ser verificados tanto no processo de declaração de um QTL como significativo, quanto na estimativa do montante da variabilidade que um QTL qualquer pode ser capaz de explicar ( $R^2$ ).

O efeito da amostra reduzida no experimento de MG foi mais acentuado de acordo com o aumento do grau de polimorfismo dos locos SSR avaliados. Para um tipo de cruzamento onde os dois parentais podem ser heterozigotos em alguns locos, é possível que a progênie segregue para até quatro classes genotípicas diferentes. Devido à progênie DG x GL 2 ter apresentado somente 48 árvores sobreviventes até os quatro anos de idade no experimento de MG, o número médio de indivíduos que seriam esperados para representar um loco que apresente quatro classes genotípicas neste ambiente seria de 12 indivíduos. Este tamanho de amostra inicial sem dúvida já seria pequeno para se realizar mapeamento de QTL com elevada confiabilidade, e o número final de amostras válidas eventualmente ainda pode ser reduzido devido a falhas podem ocorrer durante o processo de produção de dados fenotípicos e moleculares.

Para os locos onde apenas um dos parentais se apresente polimórfico, é esperado que uma progênie de 48 indivíduos segregue para duas classes genotípicas, com 24 indivíduos em média em cada uma delas. Este montante de indivíduos nas classes genotípicas já permite que o melhorista considere a magnitude dos efeitos destes QTL ( $R^2$ ) mais factíveis, mesmo com eventuais falhas que possam ocorrer durante o processo de obtenção de dados.



**Figura 15.** Efeito diferencial entre os genótipos dos locos que apresentaram interação entre QTL e ambientes. A, B e C: Efeito diferencial entre as classes genotípicas do loco Embra 0191, que apresentou interação entre QTL e ambientes para o caráter teor de lignina. D, E e F: Efeito diferencial entre as classes genotípicas do loco Embra 1868, que apresentou interação entre QTL e ambientes para o caráter rendimento de celulose.



Conforme apresentado nos resultados e disponibilizado de maneira resumida no Apêndice F, na análise individual do experimento de MG foram identificados QTL para os caracteres densidade, teor de lignina e rendimento de celulose. Os locos Embra 1008 e Embra 1627, que apresentaram associação a QTL para o caráter teor de lignina, foram capazes de explicar cada um quase 20% da variação genotípica ( $R^2$ ). Estes valores apesar de serem acima dos verificados na maioria das análises realizadas para outros caracteres, não foram muito superiores em relação aos verificados nas demais análises para este mesmo caráter (Tabela 9). Isto se deve ao fato destes dois locos apresentarem somente duas classes genotípicas, o que permitiu que a grande maioria destas classes genotípicas fossem representadas por pelo menos vinte amostras válidas no experimento de MG.

No Apêndice F também pode ser verificado que o único QTL identificado para o caráter teor de lignina na análise do experimento do RS apresentou quatro classes genotípicas, compostas por uma quantidade de amostras similar ao verificados nas classes genotípicas dos locos identificados na análise do experimento de MG. O montante que este loco foi capaz de explicar foi pouco inferior a 12%. Para os dois locos significativos na análise do experimento de MG, as médias dos genótipos com menor valor foram cerca de 3% inferiores às médias dos genótipos de maior valor.

No entanto, os dois locos que apresentaram QTL significativos para o caráter densidade juntamente com o QTL para o caráter rendimento de celulose segregaram para quatro diferentes classes genotípicas. Devido ao pequeno número de amostras válidas em algumas classes genotípicas destas análises, o montante que estes QTL foram capazes de explicar ( $R^2$ ) foram superestimados em relação aos valores verificados nas outras análises destes mesmos caracteres, onde foi avaliada uma amostra mais representativa.

O loco Embra 1868 juntamente com o loco Embra 1712 apresentaram QTL significativos para o caráter densidade no experimento de MG, e de acordo com a variabilidade avaliada na amostra, estes locos capazes de explicar juntos cerca de 100% da variação genética predita para este caráter neste ambiente. O loco Embra 1868 apresentou cerca de dez indivíduos por classe genotípica, e foi capaz de explicar pouco menos de 40% desta variabilidade. O loco Embra 1712 apresentou cerca de cinco indivíduos por classe genotípica, e foi capaz de explicar mais de 60% da variabilidade genotípica neste ambiente. Conforme pode ser verificado no Apêndice F, os outros três locos que apresentaram QTL significativos nas demais análises segregaram para apenas duas classes genotípicas, sendo todas estas representadas por um montante considerável de indivíduos.

Dentre estes três QTL significativos nas outras análises, o loco que explicou individualmente a maior parcela da variância genotípica predita foi capaz de explicar somente pouco mais de 13% desta variabilidade.

O baixo índice de plantas sobreviventes no experimento de MG também influenciou os resultados de interação entre QTL e ambientes, onde dois locos que apresentaram interação significativa entre ambientes eram polimórficos para quatro diferentes genótipos cada. Embora estas interações tenham sido significativas, é importante atentar para o número de amostras efetivamente utilizadas para validar esta interação em cada um dos ambientes avaliados antes de utilizar esta informação para SAM. A condição necessária para que um loco qualquer apresente interação entre QTL e ambientes, é de que os valores médios de suas classes genotípicas sejam estatisticamente distintas entre os ambientes avaliados.

Apesar do loco Embra 0191 situado no grupo de ligação 11 ter apresentado interação entre QTL e ambientes significativa para o caráter teor de lignina, não foi verificada associação a QTL para este loco na análise isolada ou conjunta dos experimentos. Este fato já tornaria esta interação pouco útil do ponto de vista prático, pois embora existam diferenças entre os valores médios dos genótipos entre os dois ambientes, a estruturação da variabilidade não permite que sejam obtidos ganhos de seleção utilizando este loco para SAM. No entanto, ainda existe o fato de que uma das classes genotípicas do experimento de MG apresentou somente uma amostra com dados moleculares e quantitativos válidos, onde foi verificado o maior contraste entre os valores médios dos dois ambientes. Assim, acredita-se que esta interação significativa seja um artefato decorrente do pequeno número de amostras sobreviventes no experimento de MG, conforme apresentado na Figura 15 – A, B e C.

O loco Embra 1868 além de apresentar interação entre os ambientes significativa também apresentou QTL significativo para o caráter rendimento de celulose no experimento de MG, conforme apresentado na Figura 15 – D, E e F. No entanto, esta interação não traz vantagem para ser empregada para seleção diferencial entre os dois ambientes em um programa de melhoramento, uma vez que o genótipo capaz de proporcionar um maior rendimento de celulose é o mesmo nos dois ambientes analisados. Conforme comentado anteriormente, o montante que este QTL é capaz de explicar provavelmente se encontra distorcido e superestimado, devido ao pequeno número de amostras avaliadas no experimento de MG.

Estes resultados demonstram claramente a forma como a qual uma amostra de tamanho reduzido pode influenciar no montante da variação que um QTL identificado é capaz de explicar. Os montantes que os QTL significativos para o experimento de MG foram capazes de explicar apresentaram valores elevados, muito superiores em relação aos verificados nos QTL identificados para o experimento do RS e também na análise conjunta dos dados. Foi verificada uma relação de inversa entre o aumento do montante que o QTL é capaz de explicar em função da redução do tamanho da amostra avaliada. Desta maneira, quanto maior for o tamanho da amostra avaliada mais confiáveis serão os resultados obtidos, proporcionando utilizar esta informação na SAM com maior segurança e estimar resultados confiáveis, compatíveis com os que poderão ser obtidos futuramente por um programa de melhoramento genético.

Desta forma, as magnitudes dos QTL identificados exclusivamente para o experimento de MG e os locos que apresentaram interação entre ambientes devem ser considerados com ressalvas, levando em conta o tamanho da amostra avaliada no experimento de MG em função do número de genótipos que estejam segregando nas classes genóticas destes locos. Assim, devido ao elevado *viès* presente nas análises envolvendo o experimento de MG, os QTL relacionados com este ambiente não serão discutidos nas demais análises dos caracteres avaliados a seguir.

#### 5.4 VALIDAÇÃO DOS QTL MAPEADOS NO CRUZAMENTO DG x GL2

Os resultados obtidos no mapeamento de QTL pela análise de marcas simples na análise global da progênie DG x GL2 e isolada do experimento do RS foram similares aos demais trabalhos encontrados na literatura onde foram mapeados QTL para caracteres de crescimento e qualidade da madeira. A comparação e localização dos QTL identificados neste trabalho com os descritos em outros foi realizada seguindo a nomenclatura dos grupos de ligação inicialmente estabelecida no trabalho de Grattapaglia & Sederoff (1994). No entanto, as nomenclaturas adotadas por alguns autores na identificação dos grupos de ligação construídos em seus trabalhos nem sempre são correspondentes e por este motivo devem ser feitas com ressalvas.

Esta dificuldade em se relacionar os grupos de ligação era inerente à baixa capacidade que existia anteriormente em se transferir marcas genéticas entre diferentes genótipos, o que diminuía reprodutibilidade e o relacionamento dos polimorfismos obtidos

a partir de diferentes marcadores moleculares, principalmente os dominantes. Assim, a nomenclatura adotada para realizar a comparação de QTL foi a mesma empregada nos trabalhos de Brondani et al. (1998; 2002; 2006), Missiaglia (2005), Novaes (2006), que seguiram a nomenclatura descrita pela primeira vez para o gênero *Eucalyptus* por Grattapaglia & Sederoff (1994), permitindo validar e verificar a robustez dos QTL identificados em diferentes cruzamentos para um mesmo caráter.

A discussão foi realizada comparando os resultados obtidos neste trabalho com outros, de maneira semelhante a qual os resultados dos componentes de variância foram apresentados, reunindo os caracteres avaliados de acordo com suas afinidades. Os caracteres diâmetro, altura e volume foram agrupados como caracteres de crescimento, enquanto os caracteres teor de lignina e relação siringil/guaiacil foram agrupados como propriedades da lignina. Somente os caracteres densidade e rendimento de celulose avaliados por NIRS foram discutidos separadamente, visto que o caráter pilodyn não apresentou associação significativa a QTL.

Para cada grupo de caracteres discutidos foram construídas tabelas, para facilitar a comparação dos dados obtidos neste trabalho com outros. Nas Tabelas de 11 a 14, pode ser verificado para os trabalhos citados informações tais como o tipo de marcador molecular empregado, o número de QTL identificados de acordo com o tipo de análise realizada, os grupos de ligação onde estes QTL foram mapeados, e a maior parcela possível da variância que os QTL identificados seriam capazes de explicar para os diferentes caracteres. Nos trabalhos de outros autores onde foram realizadas análises de mapeamento por diferentes metodologias, é descrito a metodologia de mapeamento utilizada e o total da variância que os QTL identificados para o caráter em discussão foram capazes de representar nestas análises.

#### **5.4.1 Caracteres de crescimento**

Foram identificados neste trabalho quatro locos responsáveis pela detecção de sete QTL para os caracteres DAP, altura e volume. Três destes locos foram responsáveis pela detecção de QTL para mais de um caráter, e as classes genóticas superiores e inferiores para estes locos foram praticamente os mesmos para os diferentes caracteres. A única exceção foi verificada no loco Embra 1679, onde ocorreu mudança do genótipo superior entre os caracteres diâmetro e altura, embora o genótipo com o menor valor médio tenha permanecido o mesmo. Estes resultados concordam com os resultados verificados

nas correlações, onde os caracteres de crescimento apresentaram correlações robustas tanto para os conjuntos de dados genotípicos quanto para os fenotípicos.

Conforme apresentado na Tabela 13, os locos Embra 1492 e Embra 1679 mapeados no grupo de ligação 9, coincidiram com QTL mapeados para caracteres de crescimento por Kirst et al. (2007). O QTL identificado no trabalho deste autor no grupo de ligação 9, utilizando técnicas de mapeamento por intervalo, foi capaz de explicar 12% da variação fenotípica na progênie avaliada em seu trabalho. Neste trabalho, os dois locos identificados no grupo de ligação 9 também foram capazes de representar cerca de 12% da variação genotípica na análise conjunta do caráter DAP, o único caráter para o qual estes dois locos situados no mesmo grupo de ligação foram significativos simultaneamente.

**Tabela 13.** QTL descritos para caracteres de crescimento nos trabalhos citados na revisão bibliográfica.

Autor	Marcador Molecular	Método de Mapeamento	Nº de QTL	Grupos de Ligação	R <sup>2</sup>
Grattapaglia et al. (1996)	RAPD	AMS	3	2 e 5	11%
		Intervalo	2	2 e 5	14%
Verhaegen et al. (1997)	RAPD	Intervalo	3	1 e 2	20%
Missiaglia (2005)	SSR	AMS	4	1, 3 e 6	N.I.
		Intervalo	2	1 e 3	18%
Novaes (2006)	SSR	AMS	52	2, 4, 6, 8 e 11 - DAP e altura	N.I.
		Intervalo	3	2, 6 e 11 – DAP	24%
			2	2 e 11 – altura	15%
Kirst et al. (2007)	genes <i>microarrays</i>	Intervalo	2	4 e 9	30%
Rocha et al. (2007)	RAPD	AMS	11	1, 2, 4 e 5	N.I.
		Intervalo	1	4	20%
Bundock et al. (2008)	RAPD	Intervalo	1	3	8%
	SSR				
presente trabalho	SSR	AMS	3	7 e 9 – DAP	23%
			2	7 e 9 – altura	18%
			2	9 – volume	12%

O loco Embra 1492 apresentou associação a QTL para os caracteres diâmetro e volume nas análises conjunta e isolada do experimento do RS. Para o caráter DAP, o genótipo de maior valor foi em média 10% superior em relação ao de menor valor nas duas análises onde foram identificados QTL. Para o caráter volume, o genótipo superior foi mais que 25% superior em relação ao de menor valor nos dois tipos de análises realizadas. O loco Embra 1679 esteve associado a QTL para os caracteres altura e diâmetro, sendo que os genótipos superiores não foram os mesmos para estes dois caracteres, embora o genótipo de menor valor não tenha apresentado diferença.

Foi verificado no caráter altura que um alelo específico do parental híbrido esteve associado aos genótipos de maior valor nas duas condições onde foram detectados QTL significativos, indicando a existência de efeito aditivo neste alelo para este caráter. Os genótipos de maior valor se apresentaram aproximadamente 10% superiores em relação ao de menor valor para a altura, enquanto para o diâmetro esta diferença foi de 15%.

A coincidência dos QTL situados no grupo de ligação 9 com os identificados por Kirst et al. (2007), mesmo avaliando genótipos não relacionados com técnicas moleculares e de mapeamento de QTL diferentes, permite validar estes locos para utilização em SAM com elevada confiança. Já a utilização do QTL identificado no grupo de ligação 7 para fins de SAM deve ser realizada com mais cautela apesar do tamanho considerável da amostra analisada, pois nenhum dos trabalhos descritos anteriormente havia identificado QTL para caracteres de crescimento no grupo de ligação 7.

O loco Embra 0042 mapeado no grupo de ligação 7 apresentou associação com os caracteres altura e diâmetro na análise individual do experimento do RS. O genótipo de maior valor foi quase 18% superior em relação ao genótipo de menor efeito para o caráter diâmetro e quase 11% superior em relação ao genótipo de menor valor para o caráter altura.

O loco Embra 1234 foi o único que foi associado a QTL e não teve sua posição definida no mapa genético construído neste trabalho. Este loco apresentou associação significativa a QTL para o caráter volume, e o genótipo de maior valor foi 20% superior em relação ao de menor valor. Os valores médios das classes genotípicas dos locos que apresentaram QTL significativos para o caráter altura podem ser visualizados no Apêndice F.

### 5.4.2 Densidade

Para o loco Embra 0340 mapeado no grupo de ligação 7, o genótipo de maior valor foi o mesmo nos dois tipos de análises onde foram detectados QTL significativos, apresentando médias 5% acima das médias do genótipo de menor valor. O loco Embra 1624 mapeado no grupo de ligação 6 apresentou associação a QTL somente na análise conjunta dos experimentos, sendo que a média do genótipo superior 2% acima da média do genótipo com menor densidade. Embora os QTL identificados na análise do experimento de MG devam ser considerados com ressalvas devido ao efeito da pequena amostra analisada neste local, os QTL mapeados nos grupos de ligação 1 e 7 (Embra 1712 e Embra 1868), também coincidiram com QTL mapeados no trabalho de outros autores, conforme apresentado na Tabela 14 e no Apêndice F.

**Tabela 14.** QTL descritos para o caráter densidade nos trabalhos citados na revisão bibliográfica.

Caráter	Autor	Marcador Molecular	Método de Mapeamento	Nº de QTL	Grupos de Ligação	R <sup>2</sup>
Pilodyn	Verhaegen et al. (1997)	RAPD	Intervalo	4	1, 2, 8 e 11	25%
	Novaes (2006)	SSR	AMS	27	2, 6, 8, 9 e 10	N.I.
			Intervalo	5	1, 3, 6 e 8	50%
	Bundock et al. (2008)	RAPD SSR	Intervalo	2	6 e 11	20%
Densidade	Grattapaglia et al. (1996)	RAPD	AMS	11	2, 5 e 7	N.I.
			Intervalo	5	2, 5, 7 e 6/9	25%
	Missiaglia (2005)	SSR	AMS	14	4 e 6	N.I.
			Intervalo	2	4 e 6	15%
	Novaes (2006)	SSR	AMS	5	1, 6 e 8	N.I.
			Intervalo	3	6, 8, 9 e 10	20%
	Rocha et al. (2007)	RAPD	AMS	5	2, 4 e 5	N.I.
			Intervalo	1	2	30%
	presente trabalho	SSR	AMS	4	1*, 6 e 7	17%

\*: GL onde foi identificado QTL exclusivamente na análise do experimento de MG.

A localização dos grupos de ligação que apresentaram locos significativos para este caráter coincidiram com grupos de ligação onde foram identificados QTL para densidade nos trabalhos de Grattapaglia et al. (1996), Missiaglia (2005) e Novaes (2006), que também verificaram a significância destes QTL validando-os em diferentes técnicas de mapeamento. A densidade estimada com pilodyn nos trabalhos de Novaes (2006) e Bundock et al. (2008) também apresentaram QTL significativos no grupo de ligação 6.

O montante da variação genotípica que estes locos foram capazes de explicar na análise global experimento foi de aproximadamente 17%, valor próximo ao verificado por Missiaglia (2005), que foi de 15%. No entanto, na análise individual do experimento do RS este montante não tenha chegou a 5%, provavelmente devido ao fato de que somente um único QTL ter sido significativo nesta análise. Estes QTL situados nos mesmos grupos de ligação onde foram mapeados QTL nos trabalhos de outros autores, permite que os mesmos sejam validados para utilização na SAM com elevada confiabilidade, embora os percentuais da variação genotípica que eles possam explicar no experimento de MG não sejam reais.

### **5.4.3 Propriedades da lignina**

Nas análises onde o número de amostras nas classes genotípicas proporcionou a detecção de QTL com efeitos confiáveis, foram identificados três QTL para o caráter teor de lignina e outros dois para a relação siringil/guaiacil, conforme apresentado a seguir na Tabela 15. Nenhum dos locos que foram associados a estes cinco QTL foram comuns entre estes dois caracteres ou mesmo estiveram presentes nos mesmos grupos de ligação, indo de encontro com os dados verificados nas correlações, onde a correlação estes caracteres se apresentou negativa e acentuada, embora esta não seja uma condição determinante que indique a existência de independência entre dois caracteres.

Para o caráter teor de lignina, os locos EG 96 e Embra 1627 apresentaram associação a QTL na análise conjunta dos experimentos, e foram mapeados respectivamente nos grupos de ligação 2 e 1, enquanto na análise do experimento do RS o loco Embra 0213 foi mapeado no grupo de ligação 4. De maneira contrária à verificada para a relação siringil/guaiacil, os grupos de ligação onde os QTL para o teor de lignina foram mapeados coincidiram com os grupos de ligação de todos os outros trabalhos que já haviam mapeado QTL para este caráter, inclusive nos QTL mapeados no experimento de MG, também estiveram localizados no grupo de ligação 1. Os QTL identificados para o



caráter teor de lignina nos grupos de ligação 1 e 4 coincidiram com os identificados por Novaes (2006), enquanto o QTL identificado no grupo de ligação 2 coincidiu com os detectados nos trabalhos de Gion et al. (2000) e Missiaglia (2005).

**Tabela 15.** QTL descritos para caracteres relacionados com as propriedades da lignina nos trabalhos citados na revisão bibliográfica.

Caráter	Autor	Marcador Molecular	Método de Mapeamento	Nº de QTL	Grupos de Ligação	R <sup>2</sup>
Teor de Lignina	Gion et al. (2000)	genes candidatos	AMS	4	2, 6 e 7	N.I.
	Missiaglia (2005)	SSR	AMS	1	2	N.I.
			Intervalo	2	2 e 9	60%
	Novaes (2006)	SSR	AMS	27	4, 7 e 11	N.I.
			Intervalo	6	1, 4, 8 e 10	50%
	presente trabalho	SSR	AMS	4	1, 2 e 4	20%
Relação Siringil / Guaiacil	Novaes (2006)	SSR	AMS	12	9 e 10	N.I.
			Intervalo	3	1 e 10	20%
	presente trabalho	SSR	AMS	2	3 e 5	14%

Na análise do caráter teor de lignina no experimento do RS, o genótipo de menor valor do loco Embra 0213 foi 2% inferior em relação ao genótipo de maior valor. Ainda foi verificado que um dos alelos deste loco, herdado do parental híbrido, conferiu média superior nos genótipos da progênie onde esteve presente, indicando a existência de efeito de aditividade para este alelo em relação aos demais. Na análise conjunta dos experimentos, os genótipos de menor valor dos locos EG 96 e Embra 1627 foram os mesmos, e apresentaram valores médios respectivamente 1% e 2% inferiores em relação aos genótipos de maior valor. Estes dois QTL significativos na análise conjunta dos dados apresentaram somente duas classes genotípicas, que foram representadas por um grande número de amostras, foram capazes de explicar juntos menos de 10% da variância genotípica predita. Os dois QTL identificados na análise do experimento de MG situam-se no grupo de ligação 1, e também apresentaram somente duas classes genotípicas que foram representadas cada uma delas por cerca de vinte amostras ou mais, o que nos permite

considerar estes QTL para utilização em SAM, embora a magnitude de seus efeitos se encontrem super estimadas (próximas a 20% para cada loco).

Somente um autor já havia realizado mapeamento de QTL para o caráter relaçãoiringil/guaiacil anteriormente. Os locos Embra 1656 e Embra 1990 mapeados nos grupos de ligação 3 e 5, apresentaram QTL na análise individual e conjunta do experimento do RS. Estes grupos de ligação não coincidiram com os identificados no trabalho de Novaes (2006). Nestes dois locos que segregaram para quatro genótipos diferentes e foram muito bem amostrados, as duas classes genotípicas superiores de cada um destes QTL apresentaram valores exatamente iguais, conforme pode ser verificado no Apêndice F. No entanto, na análise conjunta dos experimentos a média das classes genotípicas com maior valor para o loco Embra 1990 foram somente cerca de 1% superiores em relação ao genótipo de menor valor, sendo este QTL capaz de explicar aproximadamente 5% da variação genotípica na progênie DG x GL2. Na análise individual do experimento do RS, os genótipos com maiores valores do loco Embra 1656 foram 4% superiores em relação ao genótipo de menor valor, sendo este loco capaz de explicar quase 9% da variação genotípica predita neste ambiente.

A coincidência dos QTL significativos para o caráter teor de lignina com QTL identificados no trabalho de outros autores possibilita que estes locos sejam utilizados na SAM com relativa confiabilidade para a seleção precoce de genótipos superiores. Já os QTL identificados para a relaçãoiringil/guaiacil essa utilização deve ser realizada com mais cautela, devido principalmente à pequena quantidade de trabalhos que avaliaram este caráter, o que dificulta a validação destes QTL, que não coincidiram com os identificados por Novaes (2006).

#### **5.4.4 Rendimento de celulose**

O caráter rendimento de celulose apresentou três QTL significativos nas análises onde todos os genótipos foram representativamente amostrados. Os locos Embra 1316, Embra 1332 e Embra 1716 apresentaram associações significativas a QTL nos grupos de ligação 1, 2 e 9, respectivamente. Os QTL mapeados nos grupos de ligação 1 e 9 coincidiram com QTL identificados nos trabalhos de Missiaglia (2005) e Rocha et al. (2008). O QTL identificado no grupo de ligação 2 não apresentou semelhança com nenhum outro mapeado anteriormente, conforme apresentado na Tabela 16.

O loco Embra 1716 apresentou diferença de 0,07% entre a média de seus genótipos contrastantes, sendo capaz de explicar cerca de 3% da variação genotípica predita na progênie DG x GL2. Nos demais locos que apresentaram QTL significativos na análise global e no experimento do RS, a diferença verificada entre a média dos genótipos superiores e inferiores não foi superior a 0,5%. O montante que estes locos foram capazes de explicar foi de quase 7% na análise conjunta dos experimentos, enquanto na análise individual do experimento do RS, o loco Embra 1332 sozinho foi capaz de explicar quase 6% da variação genotípica predita para este ambiente.

**Tabela 16.** QTL descritos para o caráter rendimento de celulose nos trabalhos citados na revisão bibliográfica.

Autor	Marcador Molecular	Método de Mapeamento	Nº de QTL	Grupos de Ligação	R <sup>2</sup>
Missiaglia (2005)	SSR	MAS	0	-	-
		Intervalo	2	3 e 9	70%
Novaes (2006)	SSR	MAS	7	4 e 10	N.I.
		Intervalo	1	10	13%
Rocha et al. (2007)	RAPD	MAS	8	1 e 3	N.I.
		Intervalo	1	3	13%
presente trabalho	SSR	MAS	4	1, 2, 7* e 9	10%

\*: GL onde foi identificado QTL exclusivamente na análise do experimento de MG.

## 5.5 APLICAÇÃO DA SAM EM PROGRAMAS DE MELHORAMENTO

A seleção de genótipos superiores em programas de melhoramento genético de espécies florestais é um processo lento, e que geralmente é conduzido através da avaliação de experimentos que abrangem grandes áreas. Como as espécies florestais apresentam forte interação entre ambientes, estes experimentos precisam ser representativos com relação ao local para o qual está sendo realizada a seleção dos genótipos, não existindo garantias de que o melhor clone selecionado para um determinado local tenha o mesmo desempenho em um ambiente diferente do qual este clone foi selecionado (Resende & Duarte, 2007). Neste sentido, a seleção precoce de caracteres de relevância econômica e a clonagem de genótipos superiores, ajudam a reduzir o tempo necessário para se poder usufruir dos benefícios que os programas de melhoramento de espécies florestais podem oferecer.

A seleção assistida por marcadores (SAM) em programas de melhoramento de espécies florestais permite que caracteres de expressão tardia sejam avaliados e selecionados ainda na fase juvenil (Grattapaglia, 2003). No entanto, a prática desta atividade ainda não é uma rotina nestes programas de melhoramento, devido às informações produzidas serem restritas aos *pedigrees* avaliados, aos elevados custos dispendidos na avaliação de genótipos, e principalmente à escassez de informações confiáveis que possam ser úteis a estes programas de melhoramento (Tuscan et al., 2001). Por estes motivos, com exceção de alguns caracteres qualitativos tais como resistência a doenças ou florescimento precoce, a seleção de caracteres quantitativos ainda é realizada através de experimentos de campo pela maioria dos programas de melhoramento destas empresas reflorestadoras.

De acordo com dados apresentados por Aguiar et al. (2007), marcadores microsatélites não são bons preditores do comportamento de clones de *Eucalyptus*, principalmente quando comparados com predições realizadas a partir da avaliação de cruzamentos dialélicos. A comparação de diferentes métodos de seleção de genótipos superiores realizada neste trabalho, também indicou que a seleção realizada pelo valor eBLUP apresentou eficiência superior em relação à SAM. Este resultado é esperado, uma vez que a seleção realizada pelo valor eBLUP avalia o produto final de todos os genes envolvidos na expressão de um determinado caráter, enquanto a SAM utilizando marcadores SSR avalia somente uma pequena porção das regiões genômicas mais relevantes envolvidas na expressão dos caracteres analisados.

No entanto, o conhecimento da qualidade dos processos envolvidos no estudo permite que o melhorista tome uma decisão confiável para viabilizar a aplicação da SAM em um programa de melhoramento genético. A decisão de utilizar uma marca molecular no processo de seleção assistida, bem como o momento de se realizar esta seleção, deve ser consequência de vários fatores, que vão desde a coleta de dados fenotípicos e moleculares até a interação do melhorista com os resultados obtidos durante as diversas etapas de análise de dados quantitativos e moleculares. Por estes motivos, acredita-se que devem ser estabelecidos critérios para viabilizar a utilização da SAM dentro de um programa de melhoramento genético. A seguir serão descritos cinco critérios para viabilizar a SAM em programas de melhoramento, sendo que três deles estão relacionados com a consistência e validação dos QTL identificados, e outros dois com o momento e a maneira com a qual a SAM será implementada dentro do programa de melhoramento genético em andamento.

O primeiro destes critérios deve levar em conta a robustez e confiabilidade das análises realizadas. O presente trabalho apesar de adotar os melhores critérios estatísticos possíveis, identificou associações significativas a QTL no experimento de MG com base em uma pequena amostra avaliada neste ambiente. Estes QTL identificados no experimento de MG apresentaram duas interações entre ambientes significativas porém duvidosas, além do montante que estes QTL seriam capazes de explicar terem se apresentado super estimados em relação aos QTL identificados no experimento do RS, onde foi avaliada uma amostra aproximadamente três vezes maior. O montante da variação que os QTL foram capazes de representar foi altamente influenciado pelo número de observações válidas durante o mapeamento de QTL. Desta forma, os QTL identificados no experimento de MG devem ser considerados com ressalvas, devido principalmente ao pequeno número de amostras avaliadas neste local.

Um segundo critério que também poderia ser considerado para viabilização da SAM em programas de melhoramento é a validação dos QTL identificados com os descritos por outros autores. Com exceção dos QTL identificados no grupo de ligação 7 para os caracteres de crescimento, nos grupos de ligação 3 e 5 para a relação siringil/guaiacil, e nos grupos de ligação 2 e 7 para o rendimento de celulose, todos os demais QTL identificados situaram-se em grupos de ligação onde já haviam sido descritos QTL para estes mesmos caracteres. A análise de marcas simples utilizada para identificar QTL neste trabalho, permitiu identificar QTL capazes de explicar uma porção da variabilidade compatível com QTL identificados para estes mesmos caracteres que já haviam sido mapeados por outros autores, que empregaram marcadores moleculares e análises de mapeamento distintas em seus trabalhos. Desta maneira o critério de coincidência não deve ser considerado como fator principal para a exclusão de uma marca molecular do processo de SAM, pois mesmo que um QTL identificado não coincida com os descritos por outros autores, ele pode ser utilizado para SAM desde que a detecção deste QTL tenha sido realizada com base em uma amostra representativa para o ambiente onde a progênie foi avaliada.

Um terceiro critério ainda pode ser considerado no que diz respeito à seleção para múltiplos QTL. Os resultados obtidos neste trabalho indicaram que para um determinado caráter, quando a seleção de genótipos superiores para mais de um QTL simultaneamente não restringiu o montante de árvores selecionadas, ocorreu um aumento na eficiência e nos ganhos possíveis de serem obtidos através da utilização da SAM.

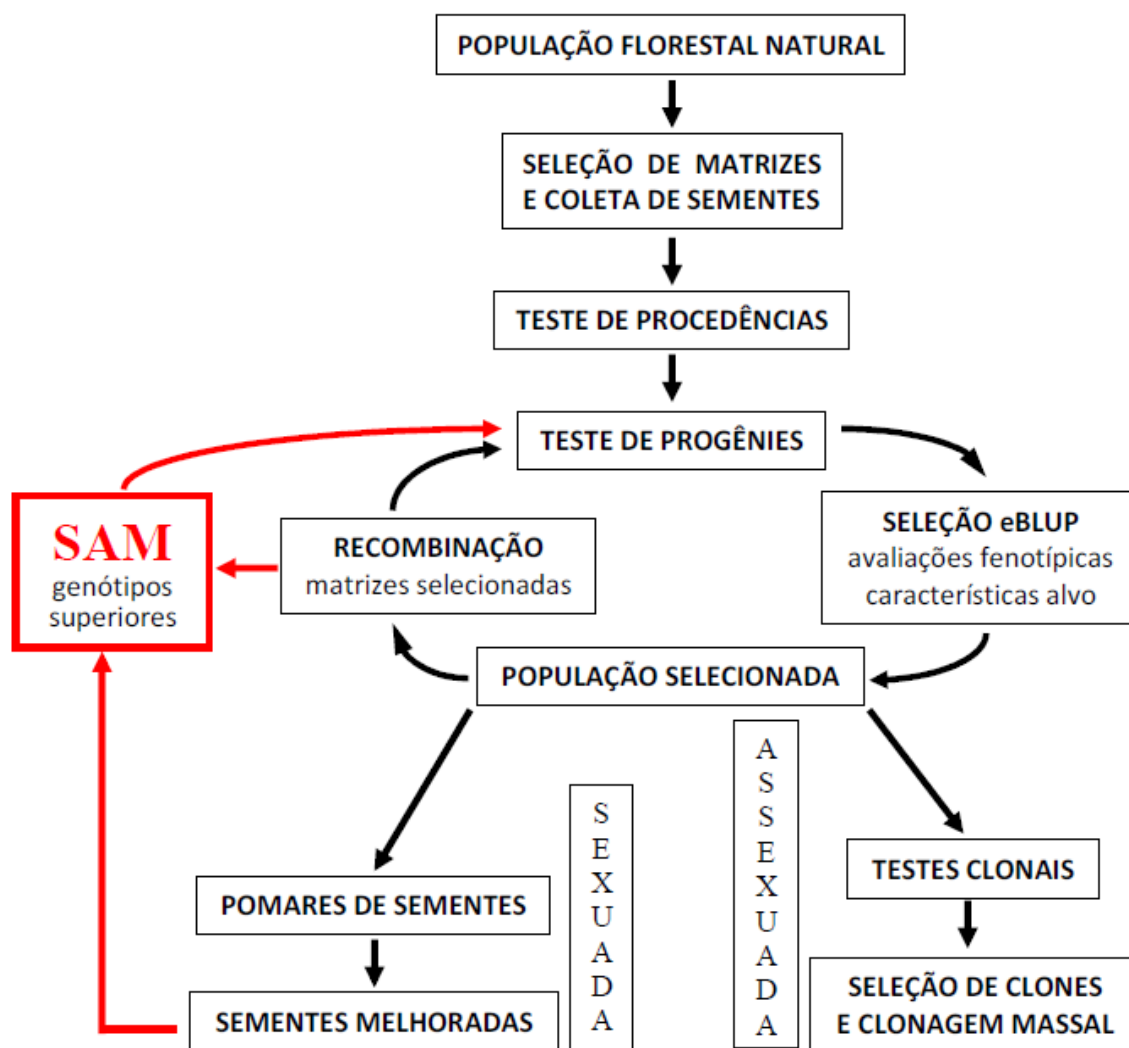
Portanto, a seleção para vários locos simultaneamente pode ser realizada com segurança, desde que o montante de árvores selecionadas não seja reduzido drasticamente.

O quarto critério diz respeito à maneira com a qual a SAM será realizada para os QTL identificados nos diferentes ambientes avaliados. A eficiência da SAM se apresentou superior na maioria das vezes quando sua aplicação foi direcionada para cada ambiente avaliado individualmente, indicando que sempre que possível a aplicação da SAM em programas de melhoramento genético deve ser direcionada para um ambiente específico, devido à acentuada interação entre genótipos e ambientes presente nas espécies perenes. No entanto este fato não pode ser tomado como regra, pois para os caracteres altura e relação siringil/guaiacil foi verificado maior eficiência da SAM na análise conjunta dos dois experimentos.

O quinto e último critério está relacionado com a fase do programa de melhoramento na qual a SAM pode ser aplicada a partir de marcadores microssatélites. Os resultados obtidos no mapeamento de QTL podem ser utilizados para selecionar ainda na fase de viveiro as mudas que irão para os testes de campo, permitindo que o melhorista direcione o processo seletivo para uma progênie originada a partir dos parentais que já tiveram QTL avaliados. As classes genotípicas superiores e que ainda apresentem um montante de indivíduos suficiente para comportar variabilidade, permitem reduzir o tamanho e os custos dos experimentos de campo, bem como o número de genótipos avaliados em futuros experimentos, conforme demonstrado na Figura 16. Para fins de seleção direcionada para caracteres de interesse, a SAM e programas de melhoramento de espécies florestais pode ser aplicada com eleada precisão a partir da segunda geração de melhoramento, ou a partir de cruzamentos cujos parentais já foram mapeados anteriormente.

Desta maneira, a eficiência na aplicação da SAM em programas de melhoramento genético depende de fatores tais como o número de QTL associados ao caráter, o tamanho amostral das classes genotípicas selecionadas nestes locos e da distribuição da variabilidade genética nos alelos destes locos. Estes resultados indicam que é preciso estabelecer critérios para fazer uso da SAM em programas de melhoramento genético florestal com segurança e eficiência, e estes critérios devem ser estabelecidos de acordo com o perfil genético da população alvo e o comportamento dos caracteres avaliados. Dentro do contexto de espécies perenes e de ciclo longo, a SAM se apresenta como uma ferramenta viável no sentido de reduzir o número de árvores necessárias para se

avaliar em um experimento de campo, uma vez que as classes genótípicas com médias superiores ainda apresentem variabilidade distribuída entre as amostras selecionadas precocemente.



**Figura 16.** Fluxograma indicando onde a SAM pode ser aplicada a partir do mapeamento de QTL utilizando marcadores moleculares em programas de melhoramento genético de espécies florestais.

## 6 CONCLUSÕES

A extração de DNA do tecido do câmbio vascular é uma alternativa viável para espécies florestais de grande porte, reduzindo o tempo e os custos envolvidos na fase de coleta de material vegetal para estudos genéticos.

A amplificação de pares de *primers* SSR em sistema *tetraplex* geradas pelo programa MULTIPLEXER foi satisfatória, e as eventuais falhas verificadas provavelmente foram decorrentes de concentrações diferenciadas do DNA genômico das amostras ou falhas de pipetagem.

A construção de mapas genéticos utilizando marcadores SSR permitiu relacionar os grupos de ligação construídos no trabalho de diferentes autores, devido à elevada transferibilidade desta classe de marcadores entre espécies do mesmo gênero.

A análise de marcas simples utilizando critérios estatísticos rigorosos permitiu validar os QTL identificados neste trabalho com QTL detectados por outros autores, que avaliaram diferentes *back-grounds* genéticos e utilizaram técnicas de mapeamento mais elaboradas.

A identificação de QTL significativos para diferentes análises de ambientes permite que sejam adotadas estratégias de SAM diferenciadas pelos programas de melhoramento de espécies perenes para cada local avaliado.

A eficiência da SAM embora tenha se apresentado inferior à seleção realizada com base nos valores genotípicos, é uma alternativa viável para ser empregada na seleção recorrente em programas de melhoramento de espécies perenes, permitindo direcionar a seleção de genótipos superiores ainda na fase de viveiro, reduzindo a quantidade de árvores avaliadas e o tamanho dos futuros experimentos de campo.

A aplicação da SAM em programas de melhoramento de espécies perenes deve ser realizada seguindo critérios, que devem ser adotados com o objetivo de garantir a obtenção de ganhos genéticos os mais próximos possíveis ou até mesmo superiores em relação aos preditos.



## 7 REFERÊNCIAS BIBLIOGRÁFICAS

AGUIAR, M. S.; FERREIRA, D. F.; AGUIAR, A. M.; BISON, O.; REZENDE, G. D. S. P.; GRATTAPAGLIA, D. Potencial de híbridos entre clones-elite de eucalipto por meio de marcadores microssatélites. **Pesquisa Agropecuária Brasileira**, Brasília, v. 42, n. 7, p. 1007-1012, 2007.

ALFENAS, A. C.; DEMUNER, N. L.; BARBOSA, M. M. Eucalipto: a ferrugem e as opções de controle. **Correio Agrícola**, São Paulo, v. 1, n. 1, p. 18-20, 1989.

ASSOCIAÇÃO BRASILEIRA DE CELULOSE E PAPEL. **Bracelpa News**. São Paulo, 2008. 31 p.

ASSOCIAÇÃO BRASILEIRA DE PRODUTORES DE FLORESTAS PLANTADAS. **Anuário estatístico da ABRAF 2008**: ano base 2007. Brasília, 2008. 90 p.

BANDYOPADHYAY, S.; CANE, K.; RASMUSSEN, G.; HAMILL, J. D. Efficient plant regeneration from seedling explants of two commercially important temperate eucalypt species- *Eucalyptus nitens* and *E. globulus*. **Plant Production Science**, Tokyo, v. 1, n. 140, p. 189-198, 1999.

BARRICHELO, L. E. G.; BRITO, J. O.; COUTO, H. T. Z.; CAMPINHOS JÚNIOR, E. Densidade básica, teor de holocelulose e rendimento em celulose de madeira de *Eucalyptus grandis*. **Silvicultura em São Paulo**, São Paulo, v. 8, n. 32, p. 802-808, 1983.

BEARZOTI, E. M. Mapeamento de QTL. In: PINHEIRO, J. B.; CARNEIRO, I. F. (Org.). **Análise de QTL no melhoramento de plantas**: 2ª Jornada em Genética e Melhoramento de Plantas realizada na Escola de Agronomia da Universidade Federal de Goiás. Goiânia: Funape, 2000. p. 63-224.

BENJAMINI, Y.; HOCHBERG, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of the Royal Statistical Society**, London, v. 2, n. 57, p. 289-300, 1995.

BONFERRONI, C. E. Teoria statistica delle classi e calcolo delle probabilità. **Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze**. Firenze, v. 1, n. 8, p. 3-62, 1936.

BOTSTEIN, D.; WHITE, R. L.; SKOLNICK, M.; DAVIS, R. W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**, Houston, v. 1, n. 32, p. 314-331, 1980.

BOUVET, J. M. **Evolution de la variabilité avec l'âge et corrélation juvénile-adulte dans des populations d'*Eucalyptus***. 1995. 175 f. Tese (Doutorado) – Institut National Agronomique Paris-Grignon, França, 1995.

BRESEGHELLO, F.; SORRELLS, M. E. Association analysis as a strategy for improvement of quantitative traits in plants. **Crop Breeding and Applied Biotechnology**, Viçosa, v. 1, n. 46, p. 1323-1330, 2006.

BRONDANI, R. P. V.; BRONDANI, C.; TARCHINI, R. & GRATTAPAGLIA, D. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. **Theoretical and Applied Genetics**, Berlin, v. 1, n. 97, p. 816-827, 1998.

BRONDANI, R. P. V.; GRATTAPAGLIA, D. Cost-effective method to synthesize a fluorescent internal DNA standard for automated fragment sizing. **Biotechniques**, New York, v. 1, n. 31, p. 793-795, 2001.

BRONDANI, R. P. V.; BRONDANI, C.; GRATTAPAGLIA, D. Towards the construction of a genus wide reference linkage map for *Eucalyptus* based on microsatellite markers. **Molecular and General Genomics**, Berlin, v. 1, n. 267, p. 338-347, 2002.

BRONDANI, R. P. V.; WILLIAMS, E. R.; BRONDANI, C.; GRATTAPAGLIA, D. A microsatellite-based consensus linkage map for species of *Eucalyptus* and a novel set of 230 microsatellite markers for the genus. **BMC Plant Biology**, London, v. 1, n. 6, p. 20-36, 2006.

BYRNE M.; MURRELL J. C.; ALLEN B.; MORAN G. An integrated genetic linkage map for eucalypts using RFLP, RAPD and isozyme markers. **Theoretical and Applied Genetics**, Berlin, v. 1, n. 91, p. 869-875, 1995.

BYRNE, M.; MARQUES-GARCIA, M. I.; UREN, T.; SMITH, D. S.; MORAN, G. F. Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus*. **Australian Journal of Botany**. Collingwood, v. 44, n. 3, p. 331-341, 1996.

BYRNE M.; MURRELL J. C.; OWEN J. V.; KRIEDEMANN, P.; WILLIAMS E. R.; MORAN, G. F. Identification and mode of action of quantitative trait loci affecting seedling height and leaf area in *Eucalyptus nitens*. **Theoretical and Applied Genetics**, Berlin, v. 1, n. 94, p. 674-681, 1997a.

BYRNE M.; MURRELL J. C.; OWEN J. V.; WILLIAMS E. R.; MORAN, G. F. Mapping of quantitative trait loci influencing frost tolerance in *Eucalyptus nitens*. **Theoretical and Applied Genetics**, Berlin, v. 1, n. 95, p. 975-979, 1997b.

CARNEIRO, M. S.; VIEIRA, M. L. C. Mapas genéticos em plantas. **Bragantia**, Campinas, v. 61, n. 2, p. 89-100, 2002.

CAVALCANTI, G. Identificação das principais espécies de *Eucalyptus* existentes no Brasil. **Silvicultura em São Paulo**, São Paulo, v. 1, n. 1, p. 149-169, 1963.

COCHERHAM, C. C. An extension of the concept of partitioning hereditary variance for analysis of covariance among relatives when epistasis is present. **Genetics**, Pittsburgh, v. 39, p.859-882, 1954.

COELHO, A. S. G. Considerações gerais sobre a análise de QTL. In: PINHEIRO, J. B.; CARNEIRO, I. F. (Org.). **Análise de QTL no melhoramento de plantas**: 2ª Jornada em Genética e Melhoramento de Plantas realizada na Escola de Agronomia da Universidade Federal de Goiás. Goiânia: Funape, 2000. p. 1-15.

COELHO, A. S. G.; SILVA, H. D. **Construção de Mapas Genéticos & Mapeamento de QTL**. Piracicaba: ed. USP, 2002. 73 p.

CORNELIUS, J. Heritabilities and additive genetic coefficients of variation in forest trees. **Cambridge Journal for Researchers**, Cambridge, v. 24, p. 372-379, 1993.

DEAN, G. H.; FRENCH, J.; TIBBITS, W. N. Variation in pulp making characteristics in a field trial of *Eucalyptus globulus*. In: ANNUAL APPITA GENERAL CONFERENCE, 44th., 1990, Appita. **Proceedings...** New Zealand, Austrália: 1990. p. 1-33.

DI CIERO, L.; AMARAL, W. Árvores Geneticamente Modificadas na Silvicultura Intensiva. **Biotecnologia Ciência e Desenvolvimento**, Brasília, v. 29, n. 1, p. 92-98, 2002.

DOYLE, J. J.; DOYLE J. L. Isolation of plant DNA from fresh tissue. **Focus**, Madison, v. 12, n. 1, p. 13-15, 1991.

ELDRIDGE, K. G. Genetic Improvement of *Eucalyptus*. **Silvae Genetica**, Frankfurt, v. 27, n. 5, p. 173-216, 1978.

ELDRIDGE, K. G.; DAVIDSON, J.; HARWOOD, C.; WYK, G. ***Eucalyptus domestication and breeding***. Oxford: Oxford University Press, 1994. 288 p.

FALCONER, D. S.; MACKAY, T. F. C. **Introduction to quantitative genetics**. 4. ed. New York: Longman Scientific & Technical, 1996. 464 p.

FAO - **Preliminary review of biotechnology in forestry, including genetic modification**. Rome – Italy: Forest Resources Development Service, Forest Resources Division, 2004. 117 p.

FERREIRA, M.; SANTOS, P. E. T. Melhoramento genético forestal dos *Eucalyptus* no Brasil – Breve histórico e perspectivas. In: IUFRO CONFERENCE ON SILVICULTURE AND IMPROVEMENT OF EUCALYPTUS, 1997, Salvador. **Proceedings...** Salvador: Embrapa, 1997. v. 1, p. 14-34.

FERREIRA, M.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3. ed. Brasília: Embrapa-Cenargen, 1998. 220 p.

FERREIRA, M. Melhoramento Genético do Arroz: impactos da genômica. In: BORÉM, A.; GIUDICE, M.; SEDYIAMA, T. (Eds). **Melhoramento Genômico**. Viçosa: UFV, 2003. p. 73-127.

FOELKEL, C. E. B. Rendimentos em celulose sulfato de *Eucalyptus spp* em função do grau de deslignificação e da densidade da madeira. **IPEF**, Piracicaba, v. 9, n. 1, p. 61-77, 1974.

FOELKEL, C. E. B.; MORA, E.; MENOCELLI, S. Densidade Básica: sua verdadeira utilidade como índice de qualidade da madeira de eucalipto para produção de celulose. **O Papel**, São Paulo, v. 53, n. 5, p. 35-40, 1992.

FOELKEL, C. Qualidade da madeira de eucalipto para atendimento das exigências do mercado de celulose e papel. In: IUFRO CONFERENCE ON SILVICULTURE AND IMPROVEMENT OF EUCALYPTS, 1997, Salvador. **Proceedings...** Salvador: Embrapa, 1997. v. 3, p. 15-22.

FONSECA, S. M.; OLIVEIRA, R. C. & SILVEIRA, P. N. Seleção da árvore industrial. **Revista Árvore**, Viçosa, v. 20, n. 1, p. 69-85, 1996.

FULKER, D. W.; CHERNY, S. S.; SHAM, P. C.; HEWITT, J. K. Combined Linkage and Association Sib-Pair Analysis for Quantitative Traits. **American Journal of Human Genetics**, Bethesda, v. 64, n. 1, p. 259-267, 1999.

GAIOTTO, F. A.; BRAMUCCI, M.; GRATTAPAGLIA, D. Estimation of outcrossing rate in a breeding population of *Eucalyptus urophylla* with dominant RAPD and AFLP markers. **Theoretical and Applied Genetics**, Berlin, v. 1, n. 95, p. 842-849, 1997.

GAN, S.; SHI, J.; LI, M. Moderate-density molecular maps of *Eucalyptus urophylla* S. T. Blake and *E. tereticornis* Smith genomes based on RAPD markers. **Genetics**, Pittsburgh, v. 118, n. 1, p. 59-67, 2003.

GONZÁLEZ, R. E.; LABATE, M. T. V.; LABATE, C. A. Transformação genética do eucalipto. **Biotecnologia, Ciência e Desenvolvimento**, Brasília, v. 1, n. 26, p. 18-22, mai./jun. 2002.

GRATTAPAGLIA, D.; BRADSHAW, H. J. Nuclear DNA content of commercially important *Eucalyptus* species and híbridos. **Canadian Journal of Forest Research**, Ottawa, v. 24, n. 1, p. 1074-1078, 1994.

GRATTAPAGLIA, D.; SEDEROFF, R. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. **Genetics**, Pittsburgh, v. 137, n. 1, p. 1121-1137, 1994.

GRATTAPAGLIA, D.; BERTOLUCCI, F. L.; PENCHEL, R. 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**, Berlin, v. 90, n. 1, p. 933-947, 1995.

GRATTAPAGLIA, D.; BERTOLUCCI, F. L.; PENCHEL, 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**, Pittsburgh, v. 144, n. 1, p. 1202-1214, 1996.

GRATTAPAGLIA, D. Molecular Breeding of Eucalyptus: State of the Art, Applications and Technical Challenges. In: JAIN, S. M. **Molecular Markers and Genome Mapping in Wood Plants**. Norwell: Kluwer Academic Publishers, 1999. v. 1, p. 451-474.

GRATTAPAGLIA, D. Além dos QTL's: Desafios para a incorporação de informações de QTL's no melhoramento florestal. In: PINHEIRO, J. B.; CARNEIRO, I. F. (Org.). **Análise de QTL no melhoramento de plantas: 2ª Jornada em Genética e Melhoramento de Plantas** realizada na Escola de Agronomia da Universidade Federal de Goiás. Goiânia: Funape, 2000. p. 21-35.

GRATTAPAGLIA, D. Genolyptus. In: BORÉM, A.; GIUDICE, M.; SEDYIAMA, T. (Eds). **Melhoramento Genômico**. Viçosa: UFV, 2003. p. 51-71.

HALEY, C. S.; KNOTT, S. A. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. **Heredity**, Sheffield, v. 69, n. 1, p. 315-324, 1992.

HAYASHI, T.; AWATA, T. Efficient method for analysis of QTL using F<sub>1</sub> progenies in an outcrossing species. **Genetics**, Pittsburgh, v. 122, n. 1, p. 173-183, 2004.

HEIN, P. R. G. **Avaliação das propriedades da madeira de Eucalyptus urophylla por meio da espectroscopia no infravermelho próximo**. 2008. 75 f. Dissertação (Mestrado em Ciência e Tecnologia da Madeira) - Universidade Federal de Lavras, Lavras, 2008.

HIGA, R. C. V. **Avaliação e Recuperação de Eucaliptus dunnii Atingidos por Geada em Campo do Tenente - PR**. 1998. 100 f. Tese (Doutorado em Engenharia Florestal) Universidade Federal do Paraná, Curitiba, 1998.

HOUSE, S. M. Reproductive biology of *Eucalyptus*. In: WILLIAMS, J.; WOINARSKI, J. (Ed.) **Eucalypt Ecology**. Cambridge: Cambridge University Press, 1997. p. 56-92.

JANSEN, R. C. Interval mapping of multiple quantitative trait loci. **Genetics**, Pittsburgh, v. 135, n. 1, p. 205-211, 1993.

JIANG, C.; ZENG, Z. B. Multiple trait analysis and genetic mapping for quantitative trait loci. **Genetics**, Pittsburgh, v. 140, n. 1, p. 1111-1127, 1995.

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

KAO, C. H.; ZENG, Z. B.; TEASDALE, R. D. Multiple interval mapping for quantitative traits loci. **Genetics**, Pittsburgh, v. 152, n. 1, p. 1203-1216, 1999.

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 & Cell Physiology**, Kyoto, v. 135, n. 1, p. 2368-2378, 2004.

KODAK COMPANY. **Kodak Electrophoresis, Documentation and 1D Analysis System (EDAS) 290**: Windows User's Guide. Rochester: Kodak, 2000. 68 p.

LANDER, E.; GREEN, P.; ABRAHAMSON, J.; BARLOW, A.; DALY, M. J.; LINCOLN, S. E.; NEWBURG, L. Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. **Genomics, Society and Policy**, Exeter, v. 1, n. 114, p. 174-181, 1987.

LANDER, E.; BOTSTEIN, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, Pittsburgh, v. 1, n. 121, p. 185-199, 1989.

LANDER, E.; THOMPSON, R. Efficiency of marker-assisted selection in the improvement of quantitative traits. **Genetics**, Pittsburgh, v. 1, n. 124, p. 743-756, 1990.

LESK, A. M. **Introdução à Bioinformática**. 2ª ed. Porto Alegre: ArtMed, 2008. 385 p.

LIU, B. H. **Statistical Genomics – Linkage, Mapping and QTL Analysis**. Boca Raton - Florida: CRC Press, 1998. 611 p.

LOPES, R. **Mapas de ligação AFLP e identificação de genes de resistência à *Xanthomonas campestris* pv. *Passiflorae* em maracujá-amarelo**. 2003. 125 f. Tese (Doutorado em Agronomia: Genética e Melhoramento de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2003.

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**. 2004. 164f. Dissertação (Mestrado em Biologia Molecular) – Universidade de Campinas, Campinas, 2004.

MALIEPAARD, C.; JANSEN J.; VAN OOIJEN, J. W. Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences for applications. **Genetic Research**, Cambridge, v. 1, n. 70, p. 237-250, jul. 1997.

MARGARIDO, G. R. A.; SOUZA, A. P.; GARCIA, A. A. F. OneMap: software for genetic mapping in outcrossing species. **Hereditas**, Hoboken, v. 1, n. 144, p. 78-79, 2007.

MARQUES, C. M.; ARAUJO, J. A.; FERREIRA, J. G.; WHETTEN, R.; O'MALLEY, D. M.; LIU, B. H.; SEDEROFF, R. AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. **Theoretical and Applied Genetics**, Berlin, v. 1, n. 96, p. 727-737, 1998.

MARQUES, C. M.; VASQUEZ-KOOL, J.; CAROCHA, V. J.; FERREIRA, J. G.; O'MALLEY, D. M.; LIU, B. H.; SEDEROFF, R. Genetic dissection of vegetative propagation traits in *Eucalyptus tereticornis* and *E. globulus*. **Theoretical and Applied Genetics**, Berlin, v. 1, n. 99, p. 936-946, 1999.

MARQUES, C. M.; BRONDANI, R. P. V.; GRATTAPAGLIA, D.; SEDEROFF, R. Conservation and syntenic of SSR loci and QTLs for vegetative propagation in four *Eucalyptus* species. **Theoretical and Applied Genetics**, Berlin, v. 1, n. 105, p. 474-478, 2002.

MISSIAGGIA, A. A. **Mapeamento genético de QTL para qualidade da madeira e florescimento precoce e, estudos de expressão gênica alelo específica em *Eucalyptus* spp.** 2005. 236 f. Tese (Doutorado em Agronomia: Genética e Melhoramento de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2005.

MOREAU, V.; FRISCHKNECHT, F.; RECKMANN, I.; VINCENTELLI, R.; RABUT, G.; STEWART, D.; WAY, M. A complex of N-WASP and WIP integrates signalling cascades that lead to actin polymerization. **Nature Cell Biology**, London, v. 2, n. 1, p. 441–466, 2000.

MORGANTE, M.; SALAMINI, F. From plant genomics to breeding practice. **Current Opinion in Biotechnology**, Amsterdam, v. 1, n. 14, p. 214-219, 2003.

MOURA, V. P. G. **O germoplasma de *Eucalyptus urophylla* S. T. Blake no Brasil.** Comunicado técnico n. 111. Brasília: Embrapa – Cenargen, 2004. 12 p.

MULLIS, K.; FALOONA, F. Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction. **Methods Enzymology**, Pasadena, v. 1, n. 55, p. 335-350, 1987.

MYRBURG, A. A.; GRIFFIN, A. R.; SEDEROFF, R. R. 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**, Berlin, v. 1, n. 107, p. 1028-1042, 2003.

MYBURG, A. A.; POTTS, B.; MARQUES, C.; KIRST, M.; GION, J.; GRATTAPAGLIA, D.; GRIMA-PETTENATI, J. *Eucalypts*. In: KOLE, C. R. (Org.). **Genome Mapping and Molecular Breeding in Plants** 1. ed. Heidelberg, Tokyo, New York: Springer, 2007. v. 7: Forest Trees, p. 115-160.

NOVAES, E. **Mapeamento de QTLs em uma Família de Irmãos Completos de *Eucalyptus Grandis* x *E. Urophylla* Utilizando Microssatélites e Triagem dos Marcadores em Biblioteca BAC.** 2006. 150 p. Tese (Doutorado em Agronomia: Genética e Melhoramento de Plantas) – Universidade Federal de Viçosa, Viçosa, 2006.

PIMENTA, L. R.; GOMIDE, J. L.; COLODETTE, J. L.; SHIN, N. H. **Comportamento dos constituintes químicos da madeira de eucalipto na polpação lo-solids.** Portal Celulose Online, Disponível em: <[www.celuloseonline.com.br/imagembank/Docs/DocBank/Eventos/430/2Pimenta/Oral.pdf](http://www.celuloseonline.com.br/imagembank/Docs/DocBank/Eventos/430/2Pimenta/Oral.pdf)>. Acesso em: 5 mai. 2008

PRYOR, L. **The Biology of *Eucalyptus*.** London: Edward Arnold Publ., 1976. 83 p.

PRYOR, L. *Eucalyptus*. In HALEVY, A. H. (Ed) **Handbook of flowering.** Boca Raton: Franklin Book Company Inc./CRC Press, 1985. p. 476-482.

QIAGEN INC. **TissueLyser Handbook**. Austin: QIAGEN, 2004. 31 p.

RAYMOND, C. A. Genetics of *Eucalyptus* wood properties. **Annals of Forest Science**, Champenoux, v. 1, n. 59, p. 522-531, 2002.

RESENDE, M. D. V.; HIGA, A. R. Estratégias de melhoramento para eucaliptos visando a seleção de híbridos. **Boletim de Pesquisa Florestal**, Colombo, v. 1, n. 21, p. 49-60, 1990.

RESENDE, M. D. V. Melhoramento de espécies perenes. In: NASS, L. L.; VALOIS, A. C. C.; MELO, I. S.; VALADARES-INGLIS, M. C. **Recursos Genéticos & Melhoramento**. Rondonópolis: Fundação MT, 2001. v. 1, p. 357-422.

RESENDE, M. D. V.; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, Goiânia, v. 37, n. 1, p. 182-194, 2007.

REMADE. Dados do setor florestal ano base 2005. **Revista da madeira**: online. Disponível em: <<http://www.remade.com.br/artigos/materias.php?artigo=Setor+Florestal>>. Acesso em: 13/09/2006.

RIBEIRO, F. A.; FILHO, J. Z. Variação da densidade básica da madeira em espécies/procedências de *Eucalyptus spp.* **IPEF**, Piracicaba, v. 1, n. 46, p. 76-85, 1993.

RITTER, E.; GEBHARDT, C.; SALAMINI, F. Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. **Genetics**, Pittsburgh, v. 1, n. 125, p. 645-654, 1990.

ROCHA, R. B.; PEREIRA, J. F.; CRUZ, C. D.; QUEIROZ, M. V.; ARAÚJO, E. F. O Mapeamento Genético no Melhoramento de Plantas. **Biotecnologia, Ciência e Desenvolvimento**, Brasília, v. 1, n. 30, p. 27-31, 2003.

ROCHA, R. B.; BARROS, E. G.; CRUZ, C. D.; ROSADO, A. M.; ARAÚJO, E. F. Mapping of QTLs Related With Wood Quality and Developmental Characteristics in Hybrids (*Eucalyptus Grandis* x *Eucalyptus Urophylla*). **Revista Árvore**, Viçosa, v. 31, n. 1, p. 13-24, 2007.

ROSADO, T. B. **Mapeamento de gene letal, responsável pela distorção de segregação e detecção de QTL para resistência à ferrugem (*Puccinia psidii*) em *Eucalyptus spp.*** 2007. 123 f. Tese (Doutorado em Agronomia: Genética e Melhoramento de Plantas) – Universidade Federal de Viçosa, Viçosa, 2007.

SANTANA, R.; BARROS N. F.; NEVES J. C. L. Nutrient use: efficiency and production sustainability of *Eucalyptus grandis* and *Eucalyptus saligna* provenances at different sites of São Paulo state. **Revista Árvore**, Viçosa, v. 26, n. 4, p. 16-21, abr. 2002.

SANTOS, S. N. **Genes de lignificação em *Eucalyptus*: estrutura e diversidade genética dos genes 4cl e ccoaomt**. 2005. 208 f. Dissertação (Mestrado em Ciências Genômicas e Biotecnologia) – Universidade Católica de Brasília, Brasília, 2005.



SANTOS, S. R. **Influência da qualidade da madeira de híbridos de *Eucalyptus grandis* x *Eucalyptus urophylla* e do processo KRAFT de polpação na qualidade da polpa branqueada.** 2005. 160 f. Dissertação (Mestrado em Recursos Florestais: Tecnologia de Produtos Florestais) – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2005.

SARTORETTO M. L.; BARRUETO, CID L. P.; BRASILEIRO, A. C. M. Biolistic transformation of *Eucalyptus grandis* x *E. urophylla* callus. **Functional Plant Biology**, Collingwood, v. 1, n. 29, n. 1, p. 917-924, jan. 2002.

SBS. Estatísticas do setor florestal. **Sociedade Brasileira de Silvicultura**: online. Disponível em: < <http://www.sbs.org.br/estatisticas.htm>>. Acesso em: 02 out. 2006.

SHIMOYAMA, V. R. S. **Variações da densidade básica e características anatômicas e químicas da madeira de *Eucalyptus* spp.** 1990. 101 f. Dissertação (Mestrado) – Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 1990.

SILVA, H. D. **Aspectos biométricos da detecção de QTLs (“Quantitative Traits Loci”) em espécies cultivadas.** 2001. 166f. Tese (Estatística e Experimentação Agrônômica) – Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 2001.

TANKSLEY, S. D. Mapping polygenes. **Annual Review of Genetics**, Palo Alto, v. 27, n. 1, p. 205-233, 1993.

THAMARUS, K. A.; GROOM, K.; MURRELL, J.; BYRNE, M.; MORAN, G. F. A genetic linkage map for *Eucalyptus globulus* with candidate loci for wood, fibre, and floral traits. **Theoretical and Applied Genetics**, Berlin, v. 1, n. 104, p. 379-387, 2002.

**The R Foundation for Statistical Computing.** v. 2.7.2, 2008.

TUSKAN, G.; WEST, D.; DAVID, M.; ELAM, C.; EVANS, R.; BRADSHAW, H.; NEALE, D.; SEWELL, M.; WHEELER, N.; MEGRAW, B.; JECH, K.; WEIRMAN, C.; DINUS, R. **Development and validation of marker-aided selection methods for wood property traits in loblolly pine and hybrid poplar**, Oak Ridge: Oak Ridge National Laboratory, 2001. 13 p.

VALENTE, C. A.; SOUSA, A. P. M.; FURTADO, F. P.; CARVALHO, A. P. Improvement program for *Eucalyptus globulus* at Portucel. **Technological Components Appita**, Appita, v. 45, n. 6, p. 403-407, 1992.

VAN OOIJEN, J. W. Lod significance thresholds for QTL analysis in experimental population of diploid species. **Heredity**, Sheffield, v. 1, n. 83, p. 613-624, 1999.

VENCOVSKY, R.; BARRIGA, P. **Genética biométrica no fitomelhoramento.** Ribeirão Preto: Sociedade Brasileira de Genética, 1992. 486 p.

VERHAEGEN, D.; PLOMION, C. Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus grandis* using RAPD markers, **Genome Research**, Cold Spring Harbor, n. 39, v. 1, p. 1051-1061, 1996.

VERHAEGEN D.; PLOMION, C.; GION, J. M.; POITEL, M.; COSTA, P.; KREMER, A. Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers: Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. **Theoretical and Applied Genetics**, Berlin, v. 1, n. 95, p. 597-608, 1997.

VOORIPS, R. E. MapChart: Software for the graphical presentation of linkage maps and QTLs. **The Journal of Heredity**, Oxford, v. 1, n. 93, p. 77-78, 2002.

WEHR, T. R. **Variações nas características da madeira de *Eucalyptus grandis* W. Hill ex –Maiden e suas influências na qualidade de cavacos em cozimentos Kraft**. 1991. 84 f. Dissertação (Mestrado) – Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 1991.

WILLIAMS, J.; WOINARSKI, J. **Eucalypt Ecology**. Cambridge: Cambridge University Press, 1997. 442 p.

ZENG, Z. B. Theoretical basis for separation of multiple linked gene effects in mapping of quantitative trait loci. **Proceedings of the Natural Academic Sciences of USA**, Washington, v. 1, n. 90, p. 10972-10976, 1993.

## APÊNDICES

<b>Apêndice A.</b> Protocolo de extração para isolamento de DNA genômico a partir do tecido do câmbio vascular utilizando o equipamento <i>TissueLyser</i> .....	129
<b>Apêndice B.</b> Combinações tetraplex geradas pelo programa MULTIPLEXER .....	131
<b>Apêndice C.</b> Informações de sequências, fluorescência, conteúdo informativo para o cruzamento analisado (C.I.), grupo de ligação onde foi mapeado (G.L.) e autoria dos 76 primers microsatélites selecionados para compor as combinações <i>tetraplex</i> .....	132
<b>Apêndice D.</b> Linhas de comandos utilizadas no Programa R para a estimativa dos componentes de variância dos caracteres avaliados nos experimentos do Projeto Genolyptus .....	138
<b>Apêndice E.</b> Linhas de comandos utilizadas no Programa R para o mapeamento de QTL através de Análise de Marcas Simples (AMS) .....	139
<b>Apêndice F.</b> Caracteres para os quais foram identificados QTL com $p < 0,01$ , com indicação do quanto os locos com associações válidas influenciam na variação fenotípica observada ( $R^2$ ), o grupo de ligação onde este QTL foi mapeado (G.L.), os valores médios e o número de indivíduos mapeados nas respectivas classes genotípicas .....	141
<b>Apêndice G.</b> Caracteres que apresentaram interação entre QTL e ambientes, com indicação do loco e grupo de ligação onde este QTL foi mapeado (G.L.), o quanto os locos com associações válidas influenciam na variação fenotípica observada ( $R^2$ ), e os valores médios das classes genotípicas nos ambientes separados e em conjunto.....	143

**Apêndice A.** Protocolo otimizado para extração de DNA genômico a partir de tecido do câmbio vascular utilizando o equipamento de maceração *TissueLyser*.

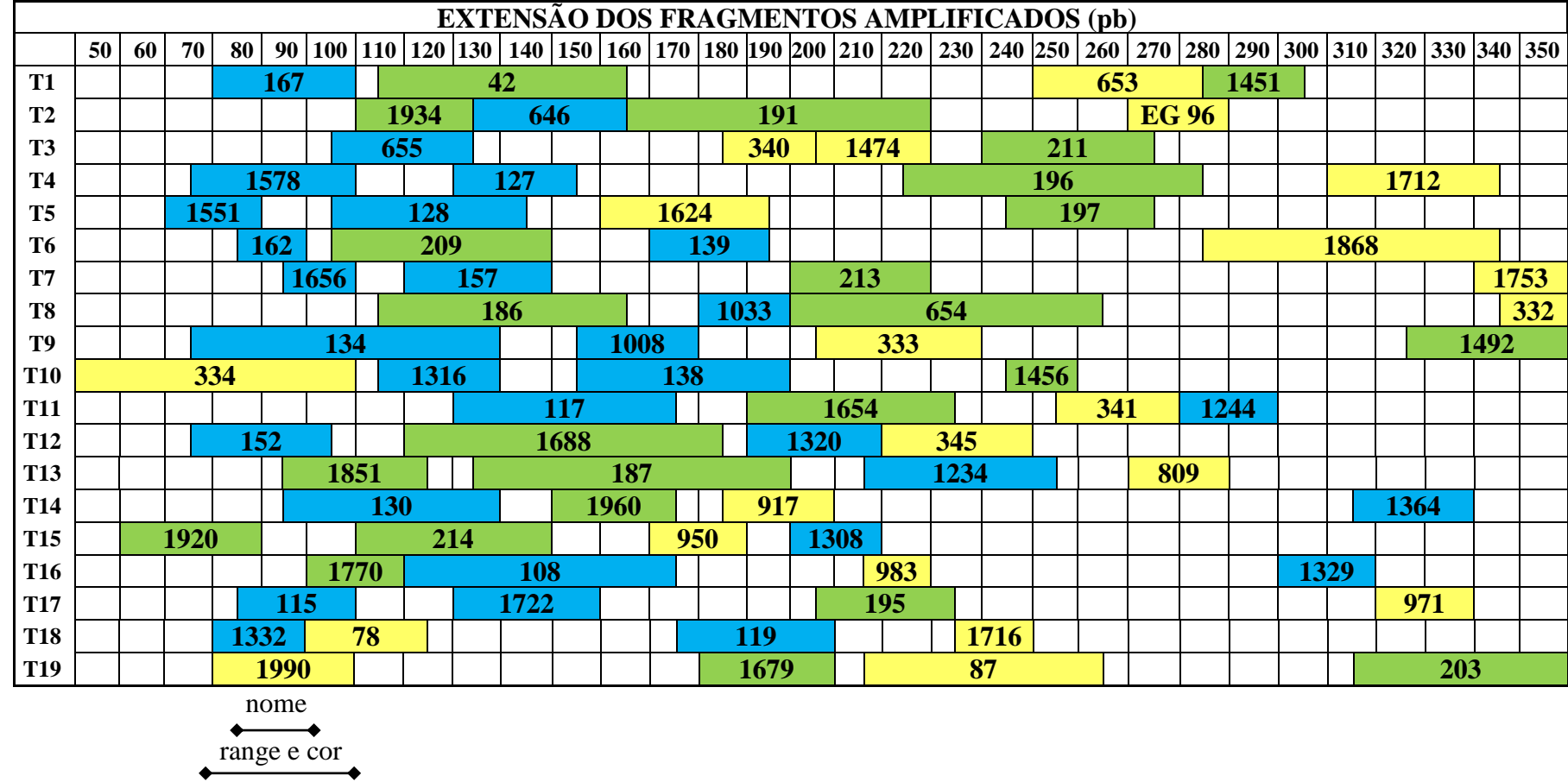
- 1- Prepare três baterias de microtubos *Eppendorf*, sendo que a primeira deve ser feita com microtubos de fundo circular e as demais com microtubos convencionais. Coloque o tampão CTAB2% com  $\beta$ -Mercaptoetanol (2  $\mu$ L/mL) para aquecer em banho-maria na temperatura de 65°C.
- 2- Forre o fundo do microtubo de fundo circular com pequenos pedaços do tecido do câmbio cortados com tesoura (aproximadamente 100 mg). Coloque nos microtubos as *beads* esterilizadas e 700  $\mu$ L de tampão CTAB2% pré aquecido.
- 3- Insira os microtubos no *TissueLyser* e agite as amostras por 1 minuto na potência máxima. Leve os tubos para o banho-maria por 30 minutos, dando uma leve agitada no material a cada 10 minutos (agitação manual).
- 4- Após 30 minutos, macere as amostras no equipamento na potência máxima por 15 segundos. Centrifugue as amostras em uma microcentrífuga, o suficiente para abaixar a espuma formada (spin), e em seguida adicione 600  $\mu$ L de CIA (Clorofórmio com Álcool Isoamílico na proporção 24:1).
- 5- Agite as amostras no equipamento por 5 minutos na potência intermediária e em seguida centrifugue os microtubos por 10 minutos a 12.000 rpm.
- 6- Colete o sobrenadante e transfira esta fase coletada para uma nova bateria de microtubos. Adicione novamente 600  $\mu$ L de CIA e coloque as amostras para agitar no equipamento por um período de 5 minutos na potência intermediária.
- 7- Centrifugue as amostras por 5 minutos a 12.000 rpm. Em seguida colete a fase sobrenadante (cerca de 400  $\mu$ L) e transfira para a última bateria de microtubos.
- 8- Adicione 500  $\mu$ L de Isopropanol e deixe no freezer por no mínimo duas horas. O ideal é deixar no freezer de um dia para o outro.
- 9- Retire as amostras do freezer e centrifugue por 3 minutos a 7.000 rpm.
- 10- Descarte o álcool com cuidado para não perder o pellet formado e adicione 1 mL de Álcool 70%, colocando em seguida as amostras no freezer por pelo menos uma hora.

Continua...

**Apêndice A.** Continuação.

- 11- Descarte o Álcool 70% com cuidado para não perder o pellet e adicione 1 mL de Álcool Absoluto. Leve novamente as amostras ao freezer por pelo menos uma hora.
- 12- Retire as amostras do freezer, centrifugue por 3 minutos a 5.000 rpm e em seguida descarte o Álcool Absoluto.
- 13- Deixe as amostras secarem em temperatura ambiente por cerca de uma hora e dilua os pellets em TE com RNase, na proporção de 2  $\mu$ L/mL de TE.

**Apêndice B.** Combinações tetraplex\* geradas pelo programa MULTIPLEXER (Coelho, não publicado).



\*: Com excessão do primer EG 96 (T2), todos os demais são identificados pelo acrônimo EMBRA.

**Apêndice C.** Informações de seqüências, fluorescência, conteúdo informativo para o cruzamento analisado (C.I.), grupo de ligação onde foi mapeado (G.L.) e autoria dos 76 primers microssatélites selecionados para compor as combinações *tetraplex*.

C.I.: 1-completamente informativo; 2-informativo para o parental DG; 3-informativo para o parental GL2.

Identificação	Sequências	Temp. Anelamento	Motivo de repetição	Fluorescência	C.I.	G.L.	Autoria
Eg 96	F: ccagggaacaattcaagc R: gagcgacaacccaagtctc	58 °C	(CA)20	Ned	2	2	Byrne et al. (1996)
Embra0042	F: gagtaaaaattggtttgagtg R: ccctcttttcatttgtctt	56 °C	(AG)15	Hex	1	7	Brondani et al. (2002)
Embra0078	F: tattatccaaaacgtcattagc R: tagaggcaggtgtagctg	58 °C	(TC)24	Ned	1	4	Brondani et al. (2006)
Embra0087	F: cttegtcacatcagtc R: agtcaataacattcaagactgc	56 °C	(CT)17AGA...(CT)18	Ned	1	11	Brondani et al. (2006)
Embra0108	F: aaagaaaccgtaagtcgc R: ctaccaatggacaatagc	58 °C	(CT)17	6-Fam	2	9	Brondani et al. (2006)
Embra0115	F: ctctactactggttctatcact R: tctgtcttttctctctcta	56 °C	(GA)13	6-Fam	2	3	Brondani et al. (2006)
Embra0117	F: ggaattcccgtcaaac R: ctgaagcttgacaatcagg	58 °C	(CT)26	6-Fam	3	7	Brondani et al. (2006)
Embra0119	F: ctaccttccactaccaaag R: caacaaaagtatgtgggtga	56 °C	(AC)4G(CA)2CTTTT CT(TC)2CA(TC)15	6-Fam	1	8	Brondani et al. (2006)
Embra0127	F: aaatgctcgtcctatttgat R: agatataggtcaaatccctcc	56 °C	(CT)16CGG...(TC)15	6-Fam	2	10	Brondani et al. (2006)
Embra0128	F: aatttggttgatctgatcc R: gctccaaagttatcgaaagt	58 °C	(TC)25	6-Fam	1	7	Brondani et al. (2006)
Embra0130	F: catgattacgccagctcg R: gaacgctattcctcggatga	58 °C	(AG)9AGG(AG)15	6-Fam	2	4	Brondani et al. (2006)
Embra0134	F: agacaccatgtccccagg R: gttatgtttggtgaaaggag	56 °C	(AG)22	6-Fam	2	2	Brondani et al. (2006)
Embra0138	F: gccgtatgcctactgtta R: taccaccatgtgcgagt	56 °C	(AG)n	6-Fam	3	5	Brondani et al. (2006)
							Continua...

**Apêndice C.** Continuação.

Identificação	Sequências	Temp. Anelamento	Motivo de repetição	Fluorescência	C. I.	G.L.	Autoria
Embra0139	F: gaagccgtatgcctactgtt R: atgtgcgagtgggagatg	58 °C	(GA)32	6-Fam	2	8	Brondani et al. (2006)
Embra0152	F: tgccagaatgtatcgccc R: tctggcttcttctgttg	56 °C	(AC)24	6-Fam	2	5	Brondani et al. (2006)
Embra0157	F: ctgccttcaatatgtgt R: gaattgtgagcggataac	56 °C	(GT)16	6-Fam	2	8	Brondani et al. (2006)
Embra0167	F: gggttgataggtggattggg R: gggttgatgactgagagatgaagg	56 °C	(TC)19	6-Fam	2	7	Brondani et al. (2006)
Embra0186	F: accgagaccagttgagag R: ggaaggctttgagataac	56 °C	(GA)27	Hex	2	4	Brondani et al. (2006)
Embra0187	F: gcagctcagtgatattgg R: ctcatgcatagctgtactc	56 °C	(GA)9CAGG(GA)20	Hex	1	6	Brondani et al. (2006)
Embra0191	F: aggtgtgtctctctgttg R: gatccagctgtgtacatc	56 °C	(CT)23(CATA)5	Hex	1	11	Brondani et al. (2006)
Embra0195	F: gaatcactaatcgaactagg R: ggagttacgttctttcactt	56 °C	(GA)24	Hex	2	2	Brondani et al. (2006)
Embra0196	F: gtgaagctcaacctgttgtct R: gtgaccgatcatgtgtggact	58 °C	(GA)46	Hex	1	6	Brondani et al. (2006)
Embra0197	F: gcagcaacgtctacctcttc R: ctactgtacgttgagcgaagg	57 °C	(TC)21	Hex	1	8	Brondani et al. (2006)
Embra0203	F: cagcaacgtctacctcttc R: tgatccagttccgactagt	57 °C	(TC)21	Hex	1	8	Brondani et al. (2006)
Embra0209	F: tgcttccacttcttcag R: catcagataatcaagcaagg	56 °C	(TC)22	Hex	2	5	Brondani et al. (2006)
Embra0211	F: ttctttcccttcggccacc R: gagatggagggttatgtgtcgg	52 °C	(TC)18	Hex	1	9	Brondani et al. (2006)
Embra0213	F: agtgaatgctgtggagaacc R: ggagaagcaagaccaggagt	56 °C	(CT)17	Hex	1	4	Brondani et al. (2006)

Continua...



**Apêndice C.** Continuação.

Identificação	Sequências	Temp. Anelamento	Motivo de repetição	Fluorescência	C. I.	G.L.	Autoria
Embra0214	F: ccttcactttcttcaggt R: catcagataatcaagcaagc	56 °C	(TC)22	Hex	2	5	Brondani et al. (2006)
Embra0332	F: tggccatggtgtgttgag R: cttcatgcagacataccgtg	54 °C	(GA)11	Ned	1	4	Novaes (2005)
Embra0333	F: ctattagcctgcagttgacc R: ggatgttcattgtgacctcca	56 °C	(TC)7AG(TC)5	Ned	1	2	Brondani et al. (2006)
Embra0334	F: ttgattgggatcgaagtaga R: ggacacatgtcgtgttctc	56 °C	(TC)4(GA)7	Ned	2	N.I.	Brondani et al. (2006)
Embra0340	F: gttgagtatgcaggcaagtc R: atccataccgtatagcttagg	58 °C	(GA)35	Ned	3	7	Novaes (2005)
Embra0341	F: cctcacactccacgaacacg R: ctgtgtttcgtggtgatgc	56 °C	(TC)22	Ned	2	4	Novaes (2005)
Embra0345	F: ctgatccgataccatgagag R: cagctgtgaccacagagaga	56 °C	(TC)18	Ned	1	6	Novaes (2005)
Embra0646	F: aaagcgttacgtgcgactct R: ggtacagaagaggcggtcaa	58 °C	(GA)17	6-Fam	1	6	Lourenço (2004)
Embra0653	F: ttcccccgaagcagaaacta R: agaagcaggtgcagaggttg	60 °C	(CT)12	Ned	2	9	Lourenço (2004)
Embra0654	F: gggggcaaaatacaatccaaa R: caaaattggcaaaatcacga	56 °C	(AG)14	Hex	2	3	Lourenço (2004)
Embra0655	F: atgtggctaaccgcaaaac R: caaataggtcgtcgattgtca	56 °C	(TA)18	6-Fam	2	11	Lourenço (2004)
Embra0809	F: cctcacgccccaaagaagaag R: gggaatcgaagaaacgatga	60 °C	(CT)7	Ned	2	N.I.	Genolyptus
Embra0917	F: cttttgggaacttttggcac R: gtacgtcaagcagtcgcgaca	60 °C	(CCT)6	Ned	2	N.I.	Genolyptus
Embra0950	F: cgtccgcctaaatcctctct R: gaatcgagctggacgaagac	60 °C	(CTCG)4	Ned	2	6	Genolyptus

Continua...

**Apêndice C.** Continuação.

Identificação	Sequências	Temp. Anelamento	Motivo de repetição	Fluorescência	C. I.	G.L.	Autoria
Embra0971	F: gaaattcatcgacgagagcc R: ttgtagaactcggcgtcctt	60 °C	(AGG)6	Ned	2	N.I.	Genolyptus
Embra0983	F: gttcgcaaagcttcctca R: ggccatcatgagggagtaga	60 °C	(CT)8	Ned	2	1	Genolyptus
Embra1008	F: aagctcgcagctcagaaaaa R: gtacttgtcctccgccatgt	60 °C	(ATCG)4	6-Fam	2	1	Genolyptus
Embra1033	F: tctcttcgtccttccttga R: atcagtcgcagcggcataaac	60 °C	(CTT)8	6-Fam	2	N.I.	Genolyptus
Embra1234	F: attcaccaagccgatcagag R: atggaccggtagagggtctt	60 °C	(GAG)7	6-Fam	3	N.I.	Genolyptus
Embra1244	F: cgtgactctgtctgagctgg R: tcttctcacatgccttgac	60 °C	(AG)14	6-Fam	3	N.I.	Genolyptus
Embra1308	F: cgcttcattcaaagcctctc R: gcagatgagttcccattggt	60 °C	(AG)14	6-Fam	3	N.I.	Genolyptus
Embra1316	F: cgcacaaacagaaactccag R: gcgttcggatgtacctttgt	60 °C	(TC)12	6-Fam	2	1	Genolyptus
Embra1320	F: gccatgaatcaaaatttgcc R: cttcctttttagcggcag	60 °C	(TCA)6	6-Fam	1	N.I.	Genolyptus
Embra1329	F: agaactggcttcgcactact R: taatgctacggagggttggg	60 °C	(GGA)4	6-Fam	1	N.I.	Genolyptus
Embra1332	F: tgagggtgctggttgatcttg R: gcgatctgtcgtcttcac	60 °C	(GAA)7	6-Fam	2	2	Genolyptus
Embra1364	F: cgttttcgctcctctctctc R: tgtagagatcggggtccttg	60 °C	(CTCC)3	6-Fam	1	8	Genolyptus
Embra1451	F: gcgtacttgaagatccgctc R: agacgcactacgaggaag	60 °C	(TCC)5	Hex	2	N.I.	Genolyptus
Embra1456	F: ttccgacgggtatttgaagg R: gaaacgatttcttggttgc	60 °C	(CCGCCT)3	Hex	1	1	Genolyptus

Continua...

**Apêndice C.** Continuação.

Identificação	Sequências	Temp. Anelamento	Motivo de repetição	Fluorescência	C. I.	G.L.	Autoria
Embra1474	F: atgtctgctgctcttcat R: tccctctccagtttcacacc	60 °C	(CT)8	Hex	1	6	Genolyptus
Embra1492	F: catgtcttcgcagtgaataa R: agtcttcgcctgacttcaa	60 °C	(AG)11	Hex	2	9	Genolyptus
Embra1551	F: gtccgctcggttactcgtc R: ctcaagtcgaacggaggagag	60 °C	(CT)6	6-Fam	2	7	Genolyptus
Embra1578	F: gggcgagcgagagaagac R: gtcaggaaggtggacatcgt	60 °C	(AG)13	6-Fam	2	9	Genolyptus
Embra1624	F: caccgcgaaagttgtaagc R: ggggaaggggaagatgaagaa	60 °C	(ACC)5	Ned	2	6	Genolyptus
Embra1627	F: acgacgttgactttgatcc R: tccgcttcacaaatttctc	60 °C	(AG)6	6-Fam	2	1	Genolyptus
Embra1654	F: cgacgtcatcggtcaaacta R: atacgtacgcctcgagaaa	60 °C	(CGG)7	Hex	1	3	Genolyptus
Embra1656	F: tccgctcacatctcactgtt R: gaagaggaggaggatcaggg	60 °C	(TC)8	6-Fam	1	3	Genolyptus
Embra1679	F: gtccgccttccccttcac R: atcctcaagctctggaccac	60 °C	(TC)8	Hex	1	9	Genolyptus
Embra1688	F: ctttctcccccttctccatc R: gtgatctcctcctccgacag	60 °C	(CCG)4	Hex	1	3	Genolyptus
Embra1712	F: tccgagagaaggctcgttcat R: cgtccggagagacagaagag	60 °C	(GA)9	Ned	1	1	Genolyptus
Embra1716	F: ttggttcttccttgcttget R: gggtcgtgttgagttgagt	60 °C	(CTTGC)5	Ned	2	9	Genolyptus
Embra1722	F: ttaccgcgaatcagatcgac R: ccgacttcgaaaaagacgag	60 °C	(GA)6	6-Fam	3	5	Genolyptus
Embra1753	F: gagagcttgacatgaaggg R: ctctctctctctctcactc	60 °C	(CT)6	Ned	2	N.I.	Genolyptus

Continua...

**Apêndice C.** Continuação.

Identificação	Sequências	Temp. Anelamento	Motivo de repetição	Fluorescência	C. I.	G.L.	Autoria
Embra1770	F: aattttggtctgctggaac R: tgatccgaaaaggtgaatcc	60 °C	(TC)8	Hex	2	5	Genolyptus
Embra1851	F: gtcgtcgccattgaagttct R: cgatcctatcaggctcagtg	60 °C	(ACGG)7	Hex	3	9	Genolyptus
Embra1868	F: tgtggagcatggagtagcag R: caaatctcagagacgccaca	59 °C	(TC)18	Ned	1	7	Genolyptus
Embra1920	F: agccaaaaggaatattgggc R: gttcgccctctccctctc	60 °C	(AG)11	Hex	2	6	Genolyptus
Embra1934	F: cagagcagaagcaagcaciaa R: tctgtgcgtagaaccactgc	60 °C	(CT)6	Hex	1	N.I.	Genolyptus
Embra1960	F: gtcgaggcaggtggagtaga R: tctcatcaatggcttctcc	60 °C	(GAG)7	Hex	2	N.I.	Genolyptus
Embra1990	F: ccgctcacttcagacaagc R: cgattatcaccatcccatc	60 °C	(AG)9	Ned	1	5	Genolyptus

**Apêndice D.** Linhas de comandos utilizadas no Programa R para a estimativa dos componentes de variância dos caracteres avaliados nos experimentos do Projeto Genolyptus (Coelho - comunicação pessoal, 2005).

```
# CARREGAR TABELA DE DADOS PARA A FUNÇÃO UTILIZADA NA
ESTIMATIVA DOS VALORES E-BLUP
```

```
EXPERIMENTO<- read.table(file='EXPERIMENTO.csv', header=TRUE, sep=';')
```

```
# GERAR OS FATORES PARA A ANÁLISE
```

```
EXPERIMENTO$superbloco <- as.factor(EXPERIMENTO$superbloco)
```

```
EXPERIMENTO$bloquinho <- as.factor(EXPERIMENTO$bloquinho)
```

```
EXPERIMENTO$parcela <- as.factor(EXPERIMENTO$parcela)
```

```
EXPERIMENTO$familia <- as.factor(EXPERIMENTO$familia)
```

```
EXPERIMENTO$genotipo <- as.factor(EXPERIMENTO$genotipo)
```

```
# CRIAR CONJUNTOS DE DADOS ESPECÍFICOS PARA CADA VARIÁVEL
```

```
# VARIÁVEL_EXPERIMENTO <- na.exclude(EXPERIMENTO[, c(1:CI, CF)])
```

```
# CI e CF: LIMITES DAS COLUNAS DOS CARACTERES NA TABELA DE
ENTRADA DE DADOS)
```

```
dap_EXPERIMENTO <- na.exclude(EXPERIMENTO[, c(1:11, 12)])
```

```
altura_EXPERIMENTO <- na.exclude(EXPERIMENTO[, c(1:11, 13)])
```

```
pilodyn_EXPERIMENTO <- na.exclude(EXPERIMENTO[, c(1:11, 14)])
```

```
ligTotal_EXPERIMENTO <- na.exclude(EXPERIMENTO[, c(1:11, 15)])
```

```
rSG_EXPERIMENTO <- na.exclude(EXPERIMENTO[, c(1:11, 16)])
```

```
RD_EXPERIMENTO <- na.exclude(EXPERIMENTO[, c(1:11, 17)])
```

```
DB_EXPERIMENTO <- na.exclude(EXPERIMENTO[, c(1:11, 18)])
```

```
Vol_EXPERIMENTO <- na.exclude(EXPERIMENTO[, c(1:11, 19)])
```

```
# salvar conjuntos de dados
```

```
save.image(file='c:\\ EXPERIMENTO.Rdata')
```

```
# EXEMPLO CARÁTER DAP NO EXPERIMENTO DO RS
```

```
dap_m1 <- lmer(dap ~ familia-1 + (FL+FC+FD) + (1|familia/genotipo) + (1|superbloco/
bloquinho), data=dap_EXPERIMENTO)
```

```
dap_m2 <- lmer(dap ~ familia-1 + (FL+FC+FD) + (1|familia/genotipo) + (1|superbloco),
data=dap_EXPERIMENTO)
```

```
dap_m3 <- lmer(dap ~ familia-1 + (1|familia/genotipo) + (1|superbloco),
data=dap_EXPERIMENTO)
```

```
dap_m4 <- lmer(dap ~ familia-1 + (FL+FC+FD) + (1|familia/genotipo),
data=dap_EXPERIMENTO)
```

```
# COMPARAR OS MODELOS PARA SELECIONAR O MELHOR
```

```
aov (dap_m1, dap_m2, dist norm)
```

```
# REALIZAR INTERAÇÕES ENTRE TODOS OS MODELOS AVALIADOS
```

```
# REPETIR PARA CADA CARÁTER AVALIADO
```

**Apêndice E.** Linhas de comandos utilizadas no Programa R para o mapeamento de QTL através de Análise de Marcas Simples (AMS).

**#CARREGAR TABELA DE DADOS PARA A FUNÇÃO UTILIZADA PARA AMS**

```
test<-read.table("TABELA_DE_DADOS.csv",header=T,sep=";")
```

**#FUNÇÃO DE MAPEAMENTO**

```
single3<-function(test,marker1,marker2,trait1,trait2)
#test=tabela; marker1e2=col_prim e ult_marc;trait1 e 2prim e ult_caractquant
{
y<-0
markers<-marker2-marker1+1
traits<-trait2-trait1+1
results<-matrix(nrow=markers*traits,ncol=12)
for (a in (trait1:trait2))
{
for (b in (marker1:marker2))
{
y<-y+1
test[,b]<-factor(test[,b])
test1<-aov(test[,a]~test[,b],test)
pvalue<-summary(test1)[[1]][[5]][1]
cat(dimnames(test)[[2]][a],dimnames(test)[[2]][b],pvalue,"\n")
r<-
summary(test1)[[1]][[2]][1]/(summary(test1)[[1]][[2]][1]+summary(test1)[[1]][[2]][2])*10
0
numniveis<-length(levels(test[,b]))
if (numniveis<3)
{
lev1<-levels(test[,b])[1]
lev2<-levels(test[,b])[2]
medialev1<-mean(test[test[,b]==lev1,a],na.rm=T)
medialev2<-mean(test[test[,b]==lev2,a],na.rm=T)
results[y,4]<-lev1
results[y,5]<-lev2
results[y,8]<-medialev1
results[y,9]<-medialev2
}
if (numniveis>2)
{
lev1<-levels(test[,b])[1]
lev2<-levels(test[,b])[2]
lev3<-levels(test[,b])[3]
lev4<-levels(test[,b])[4]
medialev1<-mean(test[test[,b]==lev1,a],na.rm=T)
medialev2<-mean(test[test[,b]==lev2,a],na.rm=T)
medialev3<-mean(test[test[,b]==lev3,a],na.rm=T)
medialev4<-mean(test[test[,b]==lev4,a],na.rm=T)
```

Continua...

## Apêndice E. Continuação.

```

results[y,4]<-lev1
results[y,5]<-lev2
results[y,6]<-lev3
results[y,7]<-lev4
results[y,8]<-medialev1
results[y,9]<-medialev2
results[y,10]<-medialev3
results[y,11]<-medialev4
}
results[y,2]<-dimnames(test)[[2]][b]
results[y,1]<-dimnames(test)[[2]][a]
results[y,3]<-pvalue
results[y,12]<-r
b<-b+1
cat("FIM!!!!!!", "\n")
}
a<-a+1
}
dimnames(results)[[2]]<-
c("trait", "marker", "pvalue", "G1", "G2", "G3", "G4", "mediaG1", "mediaG2",
"mediaG3", "mediaG4", "r2")
write.table(results, "results.txt")
}

```

# ANÁLISE DOS DADOS

single3(test, CI, CF, CI, CF)

# CI e CF: LIMITES DAS COLUNAS DOS GENÓTIPOS E CARACTERES NA TABELA DE ENTRADA DE DADOS)

**Apêndice F.** Caracteres para os quais foram identificados QTL com  $p < 0,01$ , com indicação do quanto os locos com associações válidas influenciam na variação fenotípica observada ( $R^2$ ), o grupo de ligação onde este QTL foi mapeado (G.L.), os valores médios e o número de indivíduos mapeados nas respectivas classes genotípicas.

Caráter	QTL	Loco	$R^2$	G.L.	Classes Genotípicas				
Diâmetro (cm)	RS	Embra 0042	8,82	7	Genótipo	13	14	23	24
					Média	7,21	7,78	8,49	7,73
					Nº Ind.	24	26	54	32
		Embra 1492	7,42	9	Genótipo	13	23	-	-
					Média	8,28	7,45	-	-
					Nº Ind.	62	50	-	-
	Global	Embra 1492	4,74	9	Genótipo	13	23	-	-
					Média	8,07	7,41	-	-
					Nº Ind.	94	59	-	-
		Embra 1679	7,21	9	Genótipo	13	14	23	24
					Média	7,77	8,35	8,09	7,26
					Nº Ind.	51	63	29	35
Altura (m)	RS	Embra 0042	8,38	7	Genótipo	13	14	23	24
					Média	9,08	9,56	10,05	9,50
					Nº Ind.	24	26	54	32
		Embra 1679	9,31	9	Genótipo	13	14	23	24
					Média	9,99	9,90	9,87	9,09
					Nº Ind.	28	49	28	31
	Global	Embra 1679	8,97	9	Genótipo	13	14	23	24
					Média	10,06	10,03	9,84	9,17
					Nº Ind.	51	63	29	35
Volume (m <sup>3</sup> )	RS	Embra 1492	7,54	9	Genótipo	13	23	-	-
					Média	3,83e <sup>-2</sup>	2,96e <sup>-2</sup>	-	-
					Nº Ind.	62	50	-	-
	Global	Embra 1234	4,61	-	Genótipo	11	12	-	-
					Média	3,75e <sup>-2</sup>	3,04e <sup>-2</sup>	-	-
					Nº Ind.	101	88	-	-
		Embra 1492	4,90	9	Genótipo	13	23	-	-
					Média	3,64e <sup>-2</sup>	2,93e <sup>-2</sup>	-	-
					Nº Ind.	94	59	-	-
Densidade (Kg.m <sup>-3</sup> )	MG	Embra 1712	62,27	1	Genótipo	13	14	23	24
					Média	469,23	469,45	477,84	470,77
					Nº Ind.	5	5	5	4
		Embra 1868	37,65	7	Genótipo	13	14	23	24
					Média	472,83	478,21	469,41	470,07
					Nº Ind.	13	7	14	11
	RS	Embra 0340	4,93	7	Genótipo	12	13	-	-
					Média	452,20	454,46	-	-
					Nº Ind.	81	67	-	-

Continua...



**Apêndice F.** Continuação.

Caráter	QTL	Loco	R <sup>2</sup>	G.L.	Classes Genotípicas				
Densidade (Kg.m <sup>-3</sup> )	Global	Embra 0340	13,30	7	Genótipo	12	13	-	-
					Média	455,66	457,72	-	-
					Nº Ind.	101	88	-	-
		Embra 1624	3,80	6	Genótipo	11	12	-	-
					Média	457,62	456,50	-	-
					Nº Ind.	89	111	-	-
Relação Siringil / Guaiacil	RS	Embra 1656	8,95	3	Genótipo	12	13	22	24
					Média	3,24	3,28	3,37	3,37
					Nº Ind.	33	39	42	23
	Global	Embra 1990	5,30	5	Genótipo	11	12	13	23
					Média	3,28	3,26	3,28	3,25
					Nº Ind.	42	50	41	67
Teor de lignina (%)	MG	Embra 1008	19,55	1	Genótipo	13	23	-	-
					Média	23,79	24,45	-	-
					Nº Ind.	28	18	-	-
		Embra 1627	19,28	1	Genótipo	13	23	-	-
					Média	23,74	24,38	-	-
					Nº Ind.	20	27	-	-
	RS	Embra 0213	11,43	4	Genótipo	13	14	23	24
					Média	22,98	23,17	23,36	23,40
					Nº Ind.	27	19	31	27
	Global	EG 0096	1,94	2	Genótipo	13	23	-	-
					Média	23,30	23,48	-	-
					Nº Ind.	108	78	-	-
		Embra 1627	7,12	1	Genótipo	13	23	-	-
					Média	23,26	23,64	-	-
					Nº Ind.	98	79	-	-
Rendimento de Celulose (%)	MG	Embra 1868	46,90	7	Genótipo	13	14	23	24
					Média	56,48	55,76	56,86	56,89
					Nº Ind.	13	7	14	11
	RS	Embra 1332	5,89	2	Genótipo	12	22	-	-
					Média	54,87	54,55	-	-
					Nº Ind.	81	63	-	-
	Global	Embra 1316	1,40	1	Genótipo	13	23	-	-
					Média	55,19	54,95	-	-
					Nº Ind.	114	82	-	-
		Embra 1332	2,22	2	Genótipo	12	22	-	-
					Média	55,20	54,91	-	-
					Nº Ind.	107	80	-	-
		Embra 1716	3,01	9	Genótipo	13	23	-	-
					Média	55,17	55,13	-	-
					Nº Ind.	100	60	-	-

**Apêndice G.** Caracteres que apresentaram interação entre QTL e ambientes, com indicação do loco e grupo de ligação onde este QTL foi mapeado (G.L.), o quanto os locos com associações válidas influenciam na variação fenotípica observada ( $R^2$ ), e os valores médios das classes genotípicas nos ambientes separados e em conjunto.

Caráter	Loco	G.L.	Análise	Classes Genotípicas				$R^2$
Teor de Lignina (%)	Embra 0191	11	Genótipos	13	14	23	24	
			Média global	23,35	23,34	23,32	23,35	0,04
			Nº Ind.	39	33	36	40	
			Média RS	23,13	23,22	23,28	23,26	1,61
			Nº Ind.	29	32	30	33	
			Média MG	23,95	25,96	23,92	24,02	37,28
			Nº Ind.	10	1	6	7	
Rendimento de Celulose (%)	Embra 1868	7	Genótipos	13	14	23	24	
			Média global	55,25	55,12	55,12	55,19	2,70
			Nº Ind.	43	37	46	43	
			Média RS	54,64	54,94	54,56	54,77	4,39
			Nº Ind.	30	30	32	32	
			Média MG	56,48	55,76	56,86	56,89	46,90
			Nº Ind.	13	7	14	11	