

MATIAS KIRST

DESENVOLVIMENTO DE SISTEMAS DE GENOTIPAGEM MULTILÓCOS
SEMI-AUTOMATIZADOS, BASEADOS EM MARCADORES
MICROSSATÉLITES, E SUA APLICAÇÃO EM

DESENVOLVIMENTO DE SISTEMAS DE GENOTIPAGEM MULTILÓCOS
SEMI-AUTOMATIZADOS, BASEADOS EM MARCADORES
MICROSSATÉLITES, E SUA APLICAÇÃO EM
ESPÉCIES DO GÊNERO *Eucalyptus*

Tese apresentada à Universidade
Federal de Viçosa, como parte das
exigências do Curso de Genética e
Melhoramento, para obtenção do
título de *Magister Scientiae*.

VIÇOSA

MINAS GERAIS – BRASIL

JULHO – 1999

Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV

T Kirst, Matias, 1973-
K61d Desenvolvimento de sistemas de genotipagem multilo-
1999 cos semi-automatizados, baseados em marcadores micros-
satélites, e sua aplicação em espécies do gênero
Eucalyptus / Matias Kirst. – Viçosa : UFV, 1999.
150p. : il.

Orientador: Acelino Couto Alfenas.

Dissertação (mestrado) – Universidade Federal de Vi-
çosa.

1. Eucalipto – Melhoramento genético. 2. Eucalipto –
Marcadores moleculares. 3. Eucalipto – Genotipagem por
microsatélites. 4. Microsatélites – Uso. I. Universidade
Federal de Viçosa. II. Título.

CDO. adapt. CDD. 634.91653

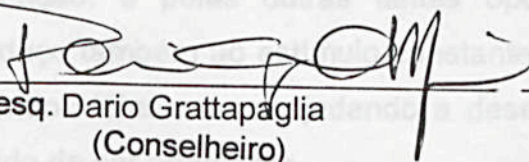
MATIAS KIRST

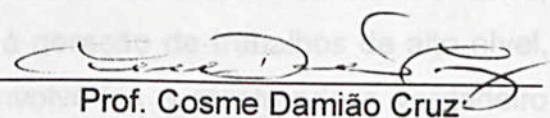
DESENVOLVIMENTO DE SISTEMAS DE GENOTIPAGEM MULTILOCOS
SEMI-AUTOMATIZADOS, BASEADOS EM MARCADORES
MICROSSATÉLITES, E SUA APLICAÇÃO EM
ESPÉCIES DO GÊNERO *Eucalyptus*

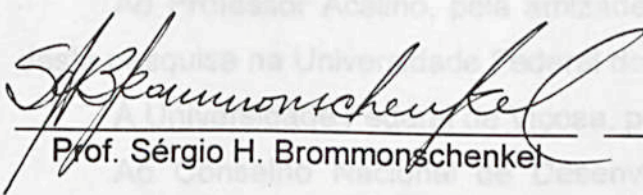
AGRADECIMENTO


Tese apresentada à Universidade
Federal de Viçosa, como parte das
exigências do Curso de Genética e
Melhoramento, para obtenção do
título de *Magister Scientiae*.

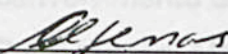
APROVADA: 1º de março de 1999.


Pesq. Dario Grattapaglia
(Conselheiro)


Prof. Cosme Damião Cruz
(Conselheiro)


Prof. Sérgio H. Brommonschenkel


Pesq. Luiz Antônio dos Santos Dias


Prof. Acelino Couto Alfenas
(Orientador)

AGRADECIMENTO

Aos meus pais, pelo estímulo que têm dado à continuação da minha carreira acadêmica e pelo apoio, principalmente durante o início do Curso.

À Mariana, companheira, amiga, digitadora, revisora e crítica, pela paciência e pelo apoio, que foram muito importantes durante estes dois anos e que certamente, o serão nos anos vindouros.

Ao Dario, por ter aberto as portas do Laboratório de Genética de Plantas da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA - Recursos Genéticos e Biotecnologia) para o meu aperfeiçoamento, no final da graduação, e pelas outras tantas oportunidades oferecidas nestes anos, agradeço também ao estímulo constante à geração de trabalhos de alto nível, trazendo idéias novas, ajudando a desenvolvê-las e mostrando o verdadeiro sentido de ser orientador.

Ao Professor Acelino, pela amizade e pela oportunidade de realização desta pesquisa na Universidade Federal de Viçosa, em Viçosa, MG.

À Universidade Federal de Viçosa, pela oportunidade que me foi dada.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela bolsa de estudos.

À empresa Aracruz Celulose, pelo fornecimento do material genético e pelo apoio financeiro no desenvolvimento deste trabalho.

Às empresas Bahia Sul Celulose, Mannesmann Florestal e Riocell S.A. e ao Instituto CRC-THF (Austrália), pelo fornecimento do material genético.

Aos Professores Cosme Damião Cruz, Luiz Antônio Dias e Sérgio Brommonschenkel, pelas sugestões.

Aos colegas do Laboratório de Genética de Plantas, pelo auxílio no desenvolvimento desta tese.

Aos colegas Márcio, Jean, Ricardo e Rony, pela amizade e pelo apoio durante a minha estada na Cidade de Viçosa.

BIOGRAFIA

Melias Kirst, filho de Nelson Kirst e Heide Kirst, nasceu a 2 de janeiro de 1973, em Porto Alegre, no Estado do Rio Grande do Sul.

Em janeiro de 1997, graduou-se em Engenharia Florestal pela Universidade Federal de Santa Maria, em Santa Maria, RS.

Em março desse mesmo ano, iniciou o Curso de Mestrado em Genética e Melhoramento na Universidade Federal de Viçosa, em Viçosa, MG.

BIOGRAFIA

	Página
LISTA Matias Kirst, filho de Nelson Kirst e Heide Kirst, nasceu a 2 de janeiro de 1973, em Porto Alegre, no Estado do Rio Grande do Sul.	7
ABST Em janeiro de 1997, graduou-se em Engenharia Florestal pela Universidade Federal de Santa Maria, em Santa Maria, RS.	1
2. REV Em março desse mesmo ano, iniciou o Curso de Mestrado em Genética e Melhoramento na Universidade Federal de Viçosa, em Viçosa, MG.	5
2.2. Marcadores moleculares	6
2.2.1. Marcadores moleculares baseados em PCR	7
2.2.2. Marcadores moleculares baseados em microssatélites	8
2.2.3. Conservação de locos microssatélites	9
2.2.4. Sistemas semi-automatizados de genotipagem baseados em microssatélites	10
2.2.5. Diversidade gênica dos marcadores moleculares	11
CAPÍTULO 1	15
DESENVOLVIMENTO DE SISTEMAS DE GENOTIPAGEM MULTILOCOS SEMI-AUTOMATIZADOS BASEADOS EM MARCADORES MICROSSATÉLITES	15
1. INTRODUÇÃO	15
2. MATERIAL E MÉTODOS	18
2.1. Seleção inicial de locos microssatélites	18

	Página
LISTA DE ABREVIATURAS	ix
EXTRATO	xi
ABSTRACT	xiii
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	5
2.1. Gênero <i>Eucalyptus</i>	5
2.2. Marcadores moleculares	6
2.2.1. Marcadores moleculares baseados em PCR	7
2.2.2. Marcadores moleculares baseados em microssatélites	8
2.2.3. Conservação de locos microssatélites	9
2.2.4. Sistemas semi-automatizados de genotipagem baseados em microssatélites	10
2.2.5. Diversidade gênica dos marcadores moleculares	11
CAPÍTULO 1	15
DESENVOLVIMENTO DE SISTEMAS DE GENOTIPAGEM MULTILOCOS SEMI-AUTOMATIZADOS, BASEADOS EM MARCADORES MICROSSATÉLITES	15
1. INTRODUÇÃO	15
2. MATERIAL E MÉTODOS	18
2.1. Seleção inicial de locos microssatélites	18

	Página
2.1.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificados	18
2.2. Avaliação dos locos microssatélites	19
2.2.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificados	19
2.3. Desenvolvimento de sistemas de genotipagem multiloco e otimização das reações de co-amplificação via PCR	21
2.3.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificados	22
2.3.2. Seleção de três sistemas Multiplex e ajuste do protocolo de amplificação via PCR	23
2.3.3. Variação no tamanho dos fragmentos detectados em diferentes condições de PCR	25
3. RESULTADOS E DISCUSSÃO	26
3.1. Seleção inicial de locos microssatélites	26
3.1.1. Qualidade de amplificação e padrão de segregação dos microssatélites em <i>E. grandis</i> e <i>E. urophylla</i>	26
3.1.2. Conservação dos locos microssatélites em outras espécies do gênero <i>Eucalyptus</i>	27
3.2. Avaliação dos locos microssatélites	28
3.2.1. Número de alelos e amplitude alélica	28
3.3. Desenvolvimento de sistemas Multiplex	30
3.3.1. Classificação dos locos microssatélites	30
3.3.2. Seleção de três sistemas de genotipagem multilocos e otimização das reações de PCR	31
3.3.3. Variação no tamanho dos fragmentos detectados em diferentes condições de PCR	37
4. RESUMO E CONCLUSÕES	40
CAPÍTULO 2	42
DETERMINAÇÃO DO TAMANHO MÍNIMO DE AMOSTRA PARA ESTIMAÇÃO DA HETEROZIGOSIDADE ESPERADA EM LOCOS MICROSSATÉLITES	42

	Página
1. INTRODUÇÃO	42
2. MATERIAL E MÉTODOS	44
2.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificados	44
2.2. Influência do tamanho da amostra sobre as estimativas de heterozigosidade esperada	46
2.2.1. Método de reamostragem	46
2.2.2. Fórmula da variância da heterozigosidade esperada	47
3. RESULTADOS E DISCUSSÃO	49
4. RESUMO E CONCLUSÕES	55
CAPÍTULO 3	57
BANCOS DE DADOS DE FREQUÊNCIAS ALÉLICAS E ESTIMATIVA DE PARÂMETROS DE INFORMAÇÃO GENÉTICA DOS SISTEMAS DE GENOTIPAGEM MULTILOCOS, EM SEIS ESPÉCIES DO GÊNERO <i>Eucalyptus</i>	57
1. INTRODUÇÃO	57
2. MATERIAL E MÉTODOS	60
2.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificados	60
2.2. Análises estatísticas	62
3. RESULTADOS E DISCUSSÃO	64
3.1. Parâmetros de informação genética dos locos microssatélites em seis espécies de <i>Eucalyptus</i>	64
3.2. Diferenciação gênica entre procedências de <i>E. grandis</i> e entre seis espécies de <i>Eucalyptus</i>	78
3.3. Distância genética e análise de variância molecular entre procedências de <i>E. grandis</i>	79
4. RESUMO E CONCLUSÕES	82
CAPÍTULO 4	84
COMPARAÇÃO ENTRE GENOTIPAGENS MANUAL E SEMI- AUTOMATIZADA DE MICROSSATÉLITES EM POPULAÇÕES DE MELHORAMENTO DE <i>Eucalyptus</i>	84

	Página
1. INTRODUÇÃO	84
2. MATERIAL E MÉTODOS	87
2.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificados	87
2.2. Similaridade genética entre os indivíduos das duas populações de melhoramento	90
3. RESULTADOS E DISCUSSÃO	91
3.1. Caracterização genotípica das populações de melhoramento de <i>E. urophylla</i> e <i>E. grandis</i>	91
3.2. Similaridade genética entre os clones das populações de melhoramento de <i>E. urophylla</i> e <i>E. grandis</i>	92
3.3. Caracterização genética dos locos EMBRA nas populações de <i>E. urophylla</i> e <i>E. grandis</i>	93
3.4. Comparação entre as eficiências da genotipagem manual e da genotipagem semi-automatizada	95
4. RESUMO E CONCLUSÕES	98
3. RESUMO E CONCLUSÕES	100
REFERÊNCIAS BIBLIOGRÁFICAS	104
APÊNDICES	113

* EDTA = ácido etilendiaminotetraacético dissódico

* EMBRA = microsátélites de *Eucalyptus demissus* no Brasil

* 5-FAM = 5-carboxifluoresceína

* H_e = heterozigosidade esperada

* HEX = hexafluoro-6-carboxifluoresceína

* H_o = heterozigosidade observada

* $-I$ = probabilidade de identidade

* M = índice Multiplex

* Kb = quilibase

* μ L = microlitro

* μ M = micromolar

* mA = milíamper

* mg = miligramas

LISTA DE ABREVIATURAS

- pu = pares de bases
- PCR = reação de polimerização em cadeia
- PIC = conteúdo de informação de polimorfismo
- Q = poder de exclusão
- QTL = locos que controlam características quantitativas
- RAPD = DNA polimórfico amplificado ao acaso
- SSR = sequência repetitiva curta
- RFLP = polimorfismo de tamanho do fragmento de restrição
- TAMRA = N,N,N',N'-tetrametil-5-carboxirodamina
- TET = tetracloro-5-carboxifluoresceína
- AFLP = polimorfismo de tamanho do fragmento amplificado
- AMOVA = análise de variância molecular
- dATP = desoxiadenosina trifosfato
- BSA = soro de albumina bovina
- DMSO = dimetil sulfóxido
- DNA = ácido desoxirribonucléico
- dNTP = dinucleotídeos trifosfatados
- EDTA = ácido etilenodiaminotetracético dissódico
- EMBRA = microssatélites de *Eucalyptus* desenvolvidos no Brasil
- 6-FAM = 6-carboxifluoresceína
- H_e = heterozigosidade esperada
- HEX = hexacloro-6-carboxifluoresceína
- H_o = heterozigosidade observada
- I = probabilidade de identidade
- IM = índice Multiplex
- Kb = kilobase
- μ L = microlitro
- μ M = micromolar
- mA = miliamper
- mg = miligrama

- mM = milimolar
- M = molar
- ng = nanograma
- pb = pares de bases
- PCR = reação de polimerização em cadeia
- PIC = conteúdo de informação de polimorfismo
- Q = poder de exclusão
- QTL = locos que controlam características quantitativas
- RAPD = DNA polimórfico amplificado ao acaso
- SSR = seqüência repetitiva curta
- RFLP = polimorfismo de tamanho do fragmento de restrição
- TAMRA = N,N,N',N'-tetrametil-6-carboxirodamina
- TET = tetracloro-6-carboxifluoresceína
- U = unidade
- V = volt
- W = watt

Foram desenvolvidos três sistemas de genotipagem multilocus (sistemas Multiplex) semi-automatizados, baseados na amplificação e análise simultânea de locos microssatélites via fluorescência, para análise genética de espécies do gênero *Eucalyptus*. Tabelas de frequências alélicas e estimativas do conteúdo de informação genética foram geradas para seis espécies comercialmente importantes do gênero *Eucalyptus* (subgênero *Symphycorymbus*). Incidentalmente, microssatélites desenvolvidos a partir de genótipos de *E. grandis* e *E. urophylla* foram avaliados quanto à especificidade de amplificação. Dentre os locos que apresentaram boa qualidade de amplificação, 11 foram selecionados para o desenvolvimento de sistemas Multiplex. Três sistemas Multiplex, comportando cada um a análise simultânea de três locos, foram desenvolvidos. Foi estudado o tamanho mínimo do amplicon necessário para se obterem estimativas acuradas de heterozigosidade de cada locos. Para isso foram genotipados 192 indivíduos de *E. grandis* provenientes de uma população de melhoramento. Com essa informação avaliou-se o comportamento de substituições para

EXTRATO

KIRST, Matias, M. S., Universidade Federal de Viçosa, julho de 1999.
Desenvolvimento de sistemas de genotipagem multilocos semi-automatizados, baseados em microssatélites, e sua aplicação em espécies do gênero *Eucalyptus*. Orientador: Acelino Couto Alfenas. Conselheiros: Dario Grattapaglia e Cosme Damião Cruz.

Foram desenvolvidos três sistemas de genotipagem multilocos (sistemas Multiplex) semi-automatizados, baseados na amplificação e análise simultânea de locos microssatélites via fluorescência, para análise genética de espécies do gênero *Eucalyptus*. Tabelas de frequências alélicas e estimativas do conteúdo de informação genética foram geradas para seis espécies comercialmente importantes do gênero *Eucalyptus* (subgênero *Symphyomyrtus*). Inicialmente, microssatélites desenvolvidos a partir de genótipos de *E. grandis* e *E. urophylla* foram avaliados quanto à especificidade de amplificação. Dentre os locos que apresentaram boa qualidade de amplificação, 11 foram selecionados para o desenvolvimento de sistemas Multiplex. Três sistemas Multiplex, comportando cada um a análise simultânea de três locos, foram selecionados. Foi estudado o tamanho mínimo de amostra necessário para se obterem estimativas acuradas da heterozigosidade de cada loco. Para isso, foram genotipados 192 indivíduos de *E. grandis* provenientes de uma população de melhoramento. Com essa informação, avaliou-se o comportamento de estimadores para

heterozigidade esperada com vários tamanhos de amostra, analisando-se a sua variância por meio da sua fórmula exata e por reamostragem numérica. Concluiu-se que uma amostra de 64 indivíduos dessa população seria satisfatória para estimar a heterozigidade esperada com boa precisão. Os três sistemas Multiplex foram utilizados para genotipar indivíduos provenientes de cinco procedências de *E. grandis*. Com os dados genotípicos, foram construídas tabelas de frequências alélicas para cada loco e procedência, além de serem registrados o número de alelos/loco e a amplitude alélica e estimados os parâmetros heterozigidade esperada, conteúdo de informação de polimorfismo, probabilidade de identidade e poder de exclusão, os quais indicaram que os locos utilizados possuía alto conteúdo de informação genética. Foram ainda estimadas as distâncias genéticas de Nei entre as procedências; uma AMOVA realizada entre as procedências indicou variação significativa entre elas, embora a maior porção da variabilidade fosse encontrada dentro de procedências. Em seguida, os sistemas Multiplex foram otimizados e caracterizados para outras cinco espécies do gênero *Eucalyptus*: *E. camaldulensis*, *E. dunnii*, *E. globulus*, *E. saligna* e *E. urophylla*. As estimativas dos parâmetros de informação genética da maioria dos locos indicaram que eles são apropriados para genotipagem de indivíduos dessas espécies. Finalmente, os sistemas Multiplex foram utilizados na avaliação de uma população de melhoramento de *E. urophylla* e de uma coleção de clones-elite, em que foi possível a detecção de erros de identificação do material genético, além de uma avaliação da diversidade genética e da similaridade entre clones.

... were generated... genetic information... to be highly informative... average values of H_e (0.85), PIC (0.84), combined Q (99.99%) and Q_{max} (10^{-14}). Nei's genetic distances estimated among provenances. An AMOVA showed that there is significant genetic variation among provenances, although the majority of the variation was found within provenances. The multiplex systems were then optimized and evaluated for other five species of *Eucalyptus*: *E. camaldulensis*, *E. dunnii*, *E. globulus*, *E. saligna* and *E. urophylla*. The estimates of genetic diversity parameters showed once more that these loci are highly appropriate for genotyping *Eucalyptus* species. As a final step to evaluate the usefulness of these genotyping systems, a breeding population of *E. urophylla* and a collection of elite clones were

ABSTRACT

KIRST, Matias, M. S., Universidade Federal de Viçosa, July 1999.
Development and application of semi-automated multilocus genotyping systems, based on microsatellites, for *Eucalyptus* species. Adviser: Acelino Couto Alfenas. Committee members: Dario Grattapaglia and Cosme Damião Cruz.

Three semi-automated multilocus genotyping systems, based on microsatellites, were developed for the analysis of *Eucalyptus* species. Allele frequency tables and estimates of the genetic information content of these loci were generated for six economically important species of *Eucalyptus* (subgenus *Symphyomyrtus*). A set of microsatellites, developed from *E. grandis* and *E. urophylla* genotypes, was evaluated based on their PCR amplification quality and transferability across species. Eleven microsatellites and fifteen combinations of three loci were tested. Among these, three triplex systems were chosen. The minimum number of individuals necessary to obtain accurate estimates of H_e was evaluated. A population of 192 individuals of *E. grandis* was typed with six microsatellites and the behavior of the estimate of H_e with increased sample sizes was analyzed using the gene diversity variance formula and numeric resampling. Congruent results using both approaches showed that a sample size of 64 individuals would be appropriate to estimate H_e with good accuracy. The multiplex systems were used to type individuals from five

Eucalyptus provenances. Allele frequency tables were generated for every locus, and estimates of parameters of genetic information content were obtained (He, PIC, I and Q). All loci showed to be highly informative with average values of He (0,85), PIC (0,84), combined Q (99,99%) and combined I (10^{-14}). Nei's genetic distances estimated among provenances. An AMOVA showed that there is significant genetic variation among provenances, although the majority of the variation was found within provenances. The multiplex systems were then optimized and evaluated for other five species of *Eucalyptus*: *E. camaldulensis*, *E. dunnii*, *E. globulus*, *E. saligna* and *E. urophylla*. The estimates of genetic information content parameters showed once more that these loci are highly appropriate for genotyping *Eucalyptus* species. As a final step to evaluate the usefulness of these genotyping systems, a breeding population of *E. urophylla* and a collection of elite clones were analyzed. Several clonal identification errors were detected within this population. Based on the multilocus genotypes, an evaluation of the genetic variability and similarity among clones was also performed.

1. INTRODUÇÃO

O gênero *Eucalyptus* (Myrtaceae) compreende mais de 500 espécies, que se distribuem sobre 40 milhões de hectares de florestas nativas na Austrália e ilhas adjacentes (ELDRIDGE et al., 1994). A ampla dispersão, em condições ecológicas variadas, faz com que este gênero apresente grande adaptabilidade, além de características favoráveis como elevada taxa de crescimento e boa qualidade da madeira, associadas à simplicidade de cultivo e manejo (FAO, 1979).

O eucalipto representa a espécie florestal de maior importância econômica no mundo. Estimativas do ano de 1985 indicaram que a área cultivada mundialmente era superior a seis milhões de hectares (ELDRIDGE et al., 1994), sendo certamente muito maior nos dias de hoje. Embora dados atuais sejam inexistentes, sabe-se que a área plantada com eucalipto no Brasil aumentou de aproximadamente 2,5 milhões de hectares em 1985 (ELDRIDGE et al., 1994) para mais de 3,5 milhões em 1990 (FERREIRA e SANTOS, 1997). O aumento de área cultivada demonstra claramente o interesse da indústria nacional pela espécie. De fato, o Brasil é hoje o país com maior área plantada de eucalipto no mundo. A indústria baseada na exploração de recursos florestais é igualmente expressiva, considerando-se que o país é o sétimo maior produtor mundial de celulose e o 11º colocado entre os países fabricantes de papel. As exportações de produtos florestais excederam a

quatro bilhões de dólares em 1995, correspondendo a 8,5% das exportações nacionais.

Apesar do interesse da indústria nacional pela espécie, a maioria dos programas de melhoramento genético de *Eucalyptus* ainda se encontra nos estádios de estabelecimento das populações de melhoramento. Nessa etapa, o conhecimento do nível de variabilidade genética entre os indivíduos que compõem, ou irão compor, essa população, assim como o grau de similaridade genética entre eles, é de grande importância. A composição de uma população de melhoramento com ampla base genética tende a possibilitar maiores ganhos e maior flexibilidade na seleção em longo prazo.

Embora a incorporação de tecnologias de análise genômica seja relativamente recente na área florestal, marcadores moleculares já têm sido utilizados por algumas empresas para avaliar a estrutura de suas populações de melhoramento atuais e planejar o estabelecimento ou o enriquecimento de novas populações. O conhecimento da divergência genética entre seus clones-elite também tem permitido o direcionamento dos cruzamentos. Isso é particularmente importante se forem considerados o custo e o tempo envolvidos nas polinizações controladas em *Eucalyptus*.

Os marcadores moleculares utilizados até o momento têm permitido avaliações importantes a respeito da estrutura e dinâmica das populações de melhoramento de *Eucalyptus*, entretanto inúmeras questões permanecem em aberto devido à falta de ferramenta mais potente de análise genética. Os marcadores RAPD, certamente os mais usados na atualidade para esse tipo de análise, são informativos, mas seguramente não são os mais apropriados para fins de estudos de diversidade genética. Questões simples e freqüentemente encontradas, como a necessidade de discriminação de clones, a determinação precisa do parentesco de indivíduos selecionados ou, ainda, a estimativa da taxa de fecundação cruzada, são de resolução difícil com RAPD em algumas situações. Marcadores co-dominantes como as isoenzimas ou RFLPs são interessantes, mas a sua capacidade de solucionar essas questões é limitada no caso das isoenzimas ou demorada e cara no caso dos RFLPs.

Marcadores moleculares baseados na amplificação de regiões microssatélites têm-se mostrado a ferramenta mais eficaz para análise genética. Além de apresentarem padrão de herança e segregação co-

dominante, são altamente multialélicos, fornecendo, com isso, elevado conteúdo de informação genética por loco (VOGEL et al., 1996; POWELL et al., 1996). A isso ainda pode ser adicionado o fato de que esses marcadores podem ser obtidos via PCR, simplificando e acelerando todo o processo de geração de dados. A análise de pequeno número de locos microssatélites tem permitido a discriminação inquestionável de variedades de culturas agrícolas (CREGAN et al., 1994; BOWERS et al., 1996; MICHELL et al., 1997), gerando subsídios para a proteção de cultivares (DIWAN e CREGAN, 1997). A reconstituição do "pedigree" de cultivares e de raças de animais domésticos também tem sido possível com o uso desses marcadores (GLOWATZKI-MULLIS et al., 1995; PÉPIN et al., 1995; RUSSELL et al., 1997).

Considerando os problemas relacionados com a identificação de material genético muitas vezes encontrados pelas empresas florestais, os microssatélites apresentam-se como a ferramenta mais eficaz para resolução dessas questões. A enorme diversidade genética encontrada nesses locos em *Eucalyptus* sp. (BYRNE et al., 1996; BRONDANI et al., 1998) demonstra imenso potencial de uso na determinação precisa do parentesco de indivíduos selecionados, permitindo ao melhorista saber exatamente quais os indivíduos que geraram determinada progênie. Além disso, a padronização dos locos analisados torna possível o desenvolvimento de sistemas universais de identificação de clones, facilitando significativamente todos os procedimentos de coleta, troca e comparação de material genético, assim como a proteção de clones-elite, da mesma forma como foi proposto para outras culturas (THOMAS et al., 1994; DIWAN e CREGAN, 1997).

A análise desses locos tem sido otimizada através do uso de sistemas em que vários microssatélites são amplificados e caracterizados via PCR simultaneamente, utilizando sistemas de detecção de fluorescência. Tais sistemas permitem alto desempenho e precisão na determinação de genótipos multilocos e tipicamente possuem excelente poder de discriminação genética de indivíduos. Embora esses sistemas de genotipagem já se encontrem disponíveis comercialmente para análise de humanos (HOCHMEISTER et al., 1995) e algumas espécies de animais domésticos, como bovinos e eqüinos, seu desenvolvimento é absolutamente inédito na área florestal.

No presente trabalho foram relatados o desenvolvimento e a aplicação dos primeiros sistemas de genotipagem semi-automatizados, baseados em microssatélites, para uma espécie florestal. Esses sistemas foram otimizados para análise de seis espécies economicamente importantes do gênero *Eucalyptus*, para as quais foi construído um banco de frequências alélicas. Além disso, foi demonstrada a aplicabilidade prática desses sistemas na caracterização de duas populações de melhoramento de uma empresa florestal e no estudo da estrutura de populações naturais de *E. grandis*. O desenvolvimento e a caracterização desses sistemas de alto desempenho constituem, portanto, uma proposta de padronização internacional de metodologia para genotipagem rápida e precisa de indivíduos das espécies de *Eucalyptus* de maior importância econômica mundial (ELDRIDGE et al., 1994).

2.1. Gênero *Eucalyptus*

O gênero *Eucalyptus* (Myrtaceae) compreende mais de 500 espécies originárias principalmente da Austrália, além da Indonésia e Papua Nova Guiné (ELDRIDGE et al., 1994).

Apesar de a sua origem estar confinada apenas ao continente australiano e ilhas adjacentes, este gênero apresenta grande amplitude de adaptações morfológicas, fisiológicas e ecológicas. Admite-se que o ancestral comum das espécies atuais ocorreu nas regiões de florestas tropicais úmidas ou nas suas margens, de onde se expandiu para regiões áridas e semi-áridas. No passo que outro grupo se adaptou às condições mais frias e úmidas do sudeste daquele continente (ELDRIDGE et al., 1994).

Atualmente, as espécies mantêm sua distribuição principalmente devido ao isolamento geográfico e à especiação floral, assim como à incompatibilidade pós-estigma. Entretanto, cruzamentos férteis podem ocorrer quando há sobreposição de distribuição das espécies e ausência de barreiras reprodutivas.

Embora diversas classificações do gênero tenham sido propostas desde o século passado, a mais usada na atualidade é a de PRYOR e JOHNSON (1971), baseada no compartilhamento de características morfológicas variadas. Essa classificação agrupa as espécies em subgêneros (espécies persistentes de

2. REVISÃO DE LITERATURA

2.1. Gênero *Eucalyptus*

O gênero *Eucalyptus* (Myrtaceae) comporta mais de 500 espécies originárias principalmente da Austrália, além da Indonésia e Papua Nova Guiné (ELDRIDGE et al., 1994).

Apesar de a sua origem estar confinada apenas ao continente australiano e ilhas adjacentes, este gênero apresenta grande amplitude de adaptações morfológicas, fisiológicas e ecológicas. Acredita-se que o ancestral comum das espécies atuais ocorria nas regiões de florestas tropicais úmidas ou nas suas margens, de onde se expandiu para regiões áridas e semi-áridas, ao passo que outro grupo se adaptou às condições mais frias e úmidas do sudeste daquele continente (ELDRIDGE et al., 1994).

Atualmente, as espécies mantêm sua distinção principalmente devido ao isolamento geográfico e à assincronia floral, assim como à incompatibilidade pólen-estigma. Entretanto, cruzamentos férteis podem ocorrer quando há sobreposição da distribuição das espécies e ausência de barreiras reprodutivas.

Embora diversas classificações do gênero tenham sido propostas desde o século passado, a mais aceita na atualidade é a de PRYOR e JOHNSON (1971), baseada no compartilhamento de características morfológicas variadas. Essa classificação agrupa as espécies em subgênero (espécies passíveis de

produzir híbridos férteis entre si), seção, série, subsérie, superespécie, espécie e subespécie.

Os dois primeiros subgêneros a serem descritos foram o *Monocalyptus* e o *Symphyomyrtus*, que se diferenciam pela presença de um e dois opérculos, respectivamente (CARR e CARR, 1963). Juntos, estes dois subgêneros compreendem aproximadamente 400 espécies, ou 4/5 do gênero *Eucalyptus*, sendo o último subgênero a ser incluído o *Telocalyptus*, por JOHNSON (1976). Os cinco subgêneros restantes são: *Blakella*, *Corymbia*, *Eudesmia*, *Gaubaea* e *Idiogenes*, que agrupam menos de 100 espécies.

Vários autores têm proposto alterações nessa classificação taxonômica, contestando principalmente a relação entre os subgêneros. Entretanto, algumas conclusões podem ser estabelecidas com base em estudos realizados em níveis morfológico e molecular. Os subgêneros *Blakella* e *Corymbia* formam um grupo monofilético com o gênero *Angophora* (LADIGES, 1983; LADIGES et al., 1995; SALE et al., 1996). Muitos taxonomistas têm, inclusive, sugerido a exclusão de *Corymbia* do grupo dos *Eucalyptus* (LADIGES et al., 1995). Os seis subgêneros restantes (*Eudesmia*, *Monocalyptus*, *Gaubaea*, *Idiogenes*, *Symphyomyrtus* e *Telocalyptus*) têm sido descritos como um grupo monofilético (SALE et al., 1996). Entretanto, outros trabalhos separaram os chamados *Eucalyptus sensu stricto* (que incluem todos os subgêneros, exceto *Blakella* e *Corymbia*) em três grupos: "symphyomyrts" (*Symphyomyrtus* e *Telocalyptus*), "monocalypts" (*Gaubaea*, *Monocalyptus* e *Idiogenes*) e "eudesimid" (*Eudesmia*) (LADIGES e HUMPHRIES, 1983; LADIGES et al., 1995).

Dentre os oito subgêneros de *Eucalyptus*, destaca-se o *Symphyomyrtus*, por sua grande importância comercial, em nível mundial. Esse subgênero inclui as espécies mundialmente mais plantadas (ELDRIDGE et al., 1994), dentre as quais estão: *E. camaldulensis*, *E. dunnii*, *E. globulus*, *E. grandis*, *E. saligna* e *E. urophylla*, utilizadas neste estudo.

2.2. Marcadores moleculares

O desenvolvimento da técnica de eletroforese de proteínas, durante a década de 50, permitiu que diferentes formas moleculares de uma mesma enzima, ou isoenzimas, passassem a ser exploradas como fonte de

polimorfismo para análise molecular. Dessa técnica resultou uma série de estudos pioneiros voltados para a quantificação da variação genética em populações (HUBBY e LEWONTIN, 1966; JOHNSON et al., 1966). Atualmente, as isoenzimas mantêm uma grande importância em relação às novas classes de marcadores moleculares, devido ao custo relativamente reduzido de implantação e manutenção, acessibilidade tecnológica e facilidade de aplicação a partir do momento em que as condições estejam otimizadas para a espécie em estudo (ALFENAS, 1998). Entretanto, o reduzido conteúdo de informação genética presente em locos isoenzimáticos de espécies florestais como as do gênero *Eucalyptus* (MARTINS-CORDER e LOPES, 1997), assim como o número restrito de locos disponíveis, contrasta com o elevado poder de informação das novas classes de marcadores como os microssatélites (BYRNE et al., 1996; BRONDANI et al., 1998). Outras críticas têm sido levantadas a respeito do uso desses marcadores (ROBINSON, 1998).

O desenvolvimento e a aplicação de RFLP (BOTSTEIN et al., 1980) trouxeram uma opção para o uso de marcadores isoenzimáticos. Embora o polimorfismo detectado através dessa técnica tenha melhorado o poder de análise em espécies do gênero *Eucalyptus* (WOLFF et al., 1993), várias limitações de ordem técnica têm restringido a utilização dessa tecnologia para análise em larga escala.

2.2.1. Marcadores moleculares baseados em PCR

O advento da técnica da reação de polimerização em cadeia (PCR) (MULLIS e FALOONA, 1987) acelerou o desenvolvimento de novas classes de marcadores moleculares.

A técnica denominada RAPD (WILLIAMS et al., 1990), baseada na amplificação via PCR de segmentos anônimos de DNA genômico, através de iniciadores arbitrários, tornou-se muito popular, principalmente por sua acessibilidade tecnológica e seu baixo custo. Entretanto, como o polimorfismo de RAPD está baseado em mutações nos sítios de anelamento dos iniciadores, ele se caracteriza pela presença ou ausência de produtos amplificados, ou seja, é um marcador com padrão de segregação dominante, com o qual não é possível a distinção entre homozigotos e heterozigotos. Apesar dessas

limitações, os marcadores RAPD foram amplamente utilizados em estudos de análise genética de *Eucalyptus* envolvendo o mapeamento e a estabilidade de QTLs (GRATTAPAGLIA et al., 1996; VERHAGEN et al., 1997) e a seleção assistida por marcadores (SQUILLASSI e GRATTAPAGLIA, 1997), além de numerosos trabalhos de avaliações da variabilidade genética em populações de melhoramento.

Os marcadores moleculares AFLP (VOS et al., 1995) exploram o polimorfismo decorrente da variação na posição de sítios de enzimas de restrição em diferentes organismos, associado à variabilidade na seqüência de nucleotídeos que flanqueiam essa seqüência. Apesar de o AFLP não apresentar as mesmas vantagens do RAPD em relação a acessibilidade tecnológica e custo, um número muito superior de locos polimórficos pode ser detectado por reação. Além disso, as variações nas combinações de enzimas de restrição e iniciadores disponíveis permitem a análise de um número de locos quase ilimitado. Marcadores AFLP foram utilizados na avaliação da variabilidade genética e da taxa de fecundação cruzada em populações de melhoramento de *Eucalyptus* (GAIOTTO et al., 1997) e na construção de mapas genéticos (MARQUES et al., 1996). Porém, essa classe de marcadores também apresenta limitações relacionadas ao padrão de segregação dominante dos marcadores RAPD, restringindo o seu uso para fins de discriminação de indivíduos.

2.2.2. Marcadores moleculares baseados em microssatélites

Os microssatélites, também conhecidos como SSR, são marcadores moleculares baseados na amplificação de seqüências repetitivas (1 a 6 pb), abundantemente encontradas no genoma de eucariotos (HAMADA et al., 1982).

Locos microssatélites têm sido descritos em grande variedade de culturas agrícolas, como cevada (WEISING et al., 1989), arroz (WU e TANKSLEY, 1993) e *Brassica* (KRESOVICH et al., 1995).

Recentemente, a identificação de locos microssatélites se estendeu a espécies florestais de interesse comercial como *Pinus* (SMITH e DEVEY, 1994; KOSTIA et al., 1995) e *Eucalyptus* (BYRNE et al., 1996; BRONDANI et al.,

1998), assim como a outras espécies florestais como o mogno (WHITE e POWELL, 1997).

Dentre as classes de DNA, a do DNA repetitivo é a que apresenta distribuição mais uniforme no genoma (STEPHAN e CHO, 1994). Em plantas, um microsatélite ocorre, em média, a cada 20 a 23 kb, sendo AA/TT e AT/TA as unidades de repetição mais freqüentes (LAGERCRANTZ et al., 1993; WANG et al., 1994).

A elevada taxa de mutações em regiões ricas em seqüências repetitivas é devida à alta incidência de escorregamento na replicação, como foi demonstrado "in vitro" por SCHLÖTTERER e TAUTZ (1992). O número variável de repetições gera locos multialélicos que podem ser analisados via PCR, utilizando um par de iniciadores específicos para cada loco que contenha uma região de DNA repetitivo (LITT e LUTY, 1989; WEBER e MAY, 1989).

Devido à abundância e à distribuição regular no genoma de eucariotos, associadas ao multialelismo e ao alto nível de polimorfismo com padrão de segregação co-dominante, os microsatélites têm tido ampla aplicação em casos em que há necessidade de discriminação de indivíduos e identificação de parentesco. Esses marcadores têm sido empregados em estudos de genética de populações (WENBURG et al., 1996), na genotipagem de indivíduos para discriminação de genótipos (CREGAN et al., 1994; RUSSELL et al., 1997), em estudos de parentesco (MARKLUND et al., 1994; GLOWATZKI-MULLIS et al., 1995) e em proteção de variedades de plantas (DIWAN e CREGAN, 1997).

2.2.3. Conservação de locos microsatélites

O uso de marcadores baseados em microsatélites em análise genômica tem sido restrito pelo alto custo e tempo necessário para o desenvolvimento dos iniciadores específicos, que flanqueiam as regiões de DNA repetitivo em cada espécie (RAFALSKI e TINGEY, 1993). Entretanto, tem sido observada a transferibilidade desses marcadores entre espécies pertencentes ao mesmo gênero e até mesmo espécies de mesma família, mas de gêneros diferentes. Essa possibilidade decorre do fato de que, embora as regiões microsatélites estejam sujeitas à alta taxa de mutação, as regiões

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

Estudo realizado por BREEN et al. (1994) demonstrou que um conjunto de 13 locos SSR, desenvolvidos a partir de DNA genômico de cavalo doméstico (*Equus caballus*), amplificou, com sucesso, regiões microssatélites de diversas espécies pertencentes ao mesmo gênero, incluindo asno (*Equus asinus*) e diversas espécies de zebras. KEMP et al. (1993), avaliando um conjunto de seis microssatélites desenvolvidos a partir de DNA genômico de *Bos taurus*, observaram que quatro deles amplificavam produtos polimórficos no gênero *Ovis* e dois no gênero *Capra*. Outros estudos demonstraram essa transferibilidade entre organismos de mesmo gênero e gêneros diferentes (LEVIN et al. 1995; PÉPIN et al., 1995; MENOTTI-RAYMOND e O'BRIEN, 1995).

Embora os estudos relatando a conservação de locos microssatélites em plantas sejam menos freqüentes, a transferibilidade entre espécies de mesmo gênero (BYRNE et al., 1996) e mesmo entre gêneros diferentes (KIJAS et al., 1995) já foi observada.

2.2.4. Sistemas semi-automatizados de genotipagem baseados em microssatélites

O uso de marcadores moleculares baseados em microssatélites para fins de genotipagem de indivíduos evoluiu para o desenvolvimento de sistemas de análise semi-automatizada. Estes sistemas apresentam inúmeras vantagens sobre as técnicas convencionais, incluindo rapidez e precisão na geração de dados, análise simultânea de vários locos, minimização da manipulação manual da informação e alto potencial de transferibilidade entre laboratórios.

Desenvolvidos inicialmente para estudos de genética humana (EDWARDS et al., 1991; LEVITT et al., 1994; MANSFIELD et al., 1994), esses sistemas baseiam-se na amplificação, via PCR, de um ou mais locos microssatélites, utilizando iniciadores marcados com fluorocromos. Os fragmentos amplificados são posteriormente separados e visualizados em equipamentos de alta resolução, que permitem a detecção de fluorescência.

Recentemente, sistemas Multiplex foram descritos para espécies de plantas. THOMAS et al. (1994) desenvolveram um sistema para *Vitis sp.*, no qual sete locos microssatélites foram amplificados separadamente via PCR, utilizando iniciadores marcados com diferentes fluorocromos e combinados anteriormente na análise em seqüenciador de DNA. O sistema foi utilizado para discriminar 80 cultivares de videira, muitos indistinguíveis em nível de fenótipo, e criar um sistema universal de diferenciação de genótipos do gênero *Vitis*. DIWAN e CREGAN (1997) propuseram um sistema semelhante, no qual 20 microssatélites eram amplificados via PCR, por meio de iniciadores marcados com fluorocromos e analisados em seqüenciador de DNA, permitindo a discriminação entre cultivares indistinguíveis em relação a características morfológicas e 17 locos RFLP.

Sistemas semelhantes têm sido propostos para genotipagem de raças de gado (*Bos taurus*) (GLOWATZKI-MULLIS et al., 1995) e cultivares de várias espécies do gênero *Brassica* (SZEWC-MCFADDEN et al., 1996).

2.2.5. Diversidade gênica dos marcadores moleculares

A diversidade gênica refere-se à probabilidade de ocorrência de polimorfismo entre indivíduos tomados ao acaso (RAFALSKI et al., 1996). Para estimar essa diversidade, vários índices que representam o conteúdo informativo de um loco para diferentes aplicações na identificação individual e na análise de parentesco têm sido utilizados.

- Heterozigosidade esperada

A heterozigosidade esperada, H_e , expressa a probabilidade de dois gametas, tomados ao acaso de uma população, terem alelos diferentes em determinado loco.

O índice H_e é expresso por

$$H_e = 1 - \sum_u p_u^2$$

em que " p_u " é a freqüência do u -ésimo alelo em uma população.

Recentemente, sistemas Multiplex foram descritos para espécies de plantas. THOMAS et al. (1994) desenvolveram um sistema para *Vitis sp.*, no qual sete locos microssatélites foram amplificados separadamente via PCR, utilizando iniciadores marcados com diferentes fluorocromos e combinados anteriormente na análise em seqüenciador de DNA. O sistema foi utilizado para discriminar 80 cultivares de videira, muitos indistinguíveis em nível de fenótipo, e criar um sistema universal de diferenciação de genótipos do gênero *Vitis*. DIWAN e CREGAN (1997) propuseram um sistema semelhante, no qual 20 microssatélites eram amplificados via PCR, por meio de iniciadores marcados com fluorocromos e analisados em seqüenciador de DNA, permitindo a discriminação entre cultivares indistinguíveis em relação a características morfológicas e 17 locos RFLP.

Sistemas semelhantes têm sido propostos para genotipagem de raças de gado (*Bos taurus*) (GLOWATZKI-MULLIS et al., 1995) e cultivares de várias espécies do gênero *Brassica* (SZEWC-MCFADDEN et al., 1996).

2.2.5. Diversidade gênica dos marcadores moleculares

A diversidade gênica refere-se à probabilidade de ocorrência de polimorfismo entre indivíduos tomados ao acaso (RAFALSKI et al., 1996). Para estimar essa diversidade, vários índices que representam o conteúdo informativo de um loco para diferentes aplicações na identificação individual e na análise de parentesco têm sido utilizados.

- Heterozigosidade esperada

A heterozigosidade esperada, H_e , expressa a probabilidade de dois gametas, tomados ao acaso de uma população, terem alelos diferentes em determinado loco.

O índice H_e é expresso por

$$H_e = 1 - \sum_u p_u^2$$

em que " p_u " é a freqüência do u-ésimo alelo em uma população.

Em outras palavras, quanto maior for o valor de heteroziguidade, maior será a probabilidade de ser detectado polimorfismo (RAFALSKI et al., 1996). Esse índice também é denominado diversidade gênica (WEIR, 1996).

Os valores do índice H_e encontrados em diferentes classes de marcadores têm demonstrado que microssatélites apresentam heteroziguidade superior à de outros marcadores de herança co-dominante (RFLP) e significativamente superior à de marcadores com padrão de herança dominante (RAPD e AFLP) (POWELL et al., 1996; VOGEL et al., 1996).

Em estudo realizado por DIWAN e CREGAN (1997), no qual um conjunto de genótipos silvestres de soja foi avaliado com 20 locos microssatélites, foram observadas heteroziguidades esperadas de 0,64 a 0,91, com média de 0,80. Quando um conjunto de genótipos de mesma origem foi avaliado com 17 locos RFLP, apenas dois apresentaram mais de dois alelos; em apenas um, a heteroziguidade esperada foi superior a 0,5.

- Conteúdo de informação de polimorfismo

O conteúdo de informação de polimorfismo, PIC, foi proposto por BOTSTEIN et al. (1980) e tem sido utilizado para indicar o nível de informação de determinado loco, para fins de estudo de mapeamento.

O conteúdo de informação de polimorfismo é calculado por

$$PIC = 1 - \left(\sum_u p_u^2 \right) - \sum_u \sum_{v>u} 2p_u^2 p_v^2$$

em que " p_u " e " p_v " são as freqüências do u-ésimo e v-ésimo alelo em uma população, sendo u diferente de v.

BOTSTEIN et al. (1980) consideraram altamente informativos locos que apresentassem valor de PIC superior a 0,5, no qual se enquadraram apenas 14% dos locos RFLP utilizados por esses pesquisadores para mapeamento do genoma humano.

Estudos mais recentes envolvendo marcadores microssatélites em humanos têm demonstrado que estes apresentam geralmente valores de PIC superiores a 0,5 (SPRECHER et al., 1996), indicando o maior conteúdo de informação presente em locos microssatélites em relação ao de locos RFLP.

Resultados semelhantes têm sido encontrados em estudos envolvendo animais domésticos (PÉPIN et al., 1995; GLOWATZKI-MULLIS et al., 1995) e culturas agrícolas (RUSSEL et al., 1997).

- Probabilidade de identidade

A probabilidade de identidade, I (PAETKAU et al., 1995), expressa a probabilidade de se encontrarem em uma população dois indivíduos com o mesmo genótipo para determinado loco.

A probabilidade de identidade é expressa por

$$I = 1 - \sum_u p_u^4 - \sum_u \sum_{v>u} (2p_u p_v)^2$$

em que " p_u " e " p_v " são as freqüências do u -ésimo e v -ésimo alelo em uma população, sendo u diferente de v .

O produto das probabilidades de identidade de vários locos que compõem um sistema de genotipagem gera a probabilidade de identidade combinada, que indica a probabilidade de serem encontrados dois indivíduos com o mesmo genótipo multiloco. Esse índice tem sido utilizado para indicar o poder de discriminação de sistemas de genotipagem multilocos, baseados em microsatélites, para identificação individual em investigações forenses.

PAETKAU et al. (1995), analisando um sistema de genotipagem de ursos polares (*Ursus maritimus*) baseado em marcadores microsatélites, determinou que, quando a probabilidade de identidade dos oito locos que compunham o sistema era considerada, o seu valor combinado era igual ou inferior a 10^{-6} . Isso indica que a probabilidade de dois indivíduos na população possuírem o mesmo genótipo multiloco para os locos amostrados era inferior a 1 em um milhão.

Valores ainda mais expressivos foram encontrados por WENBURG et al. (1996), avaliando três populações de salmão (*Oncorhynchus mykiss*) e truta (*Oncorhynchus clarki clarki*) com um sistema Multiplex composto de 10 locos microsatélites. Embora a probabilidade de identidade combinada tenha atingido valores muito variáveis entre as populações, devido às diferenças de

freqüências alélicas observadas, ela foi sempre inferior a 10^{-7} , chegando a 10^{-16} em algumas populações.

- Poder de exclusão

O poder de exclusão, Q, indica a capacidade de um loco excluir um indivíduo erroneamente acusado de paternidade (WEIR, 1996). Esse índice é calculado, considerando-se todas as possíveis configurações genótípicas entre a mãe, o filho e o suposto pai e quais deles são excludentes, conforme apresentado por WEIR (1996). O poder de exclusão combinado é calculado, levando-se em conta todos os locos que tenham sido analisados.

O índice Q de um loco microssatélite é calculado através da fórmula

$$Q = \sum_u p_u (1 - p_u)^2 - \sum_u \sum_{v>u} p_u^2 p_v^2 (4 - 3p_u - 3p_v)$$

em que " p_u " e " p_v " são as freqüências do u-ésimo e v-ésimo alelo em uma população, sendo u diferente de v.

Estudos com animais domésticos têm mostrado que os microssatélites apresentam alto poder de exclusão. MARKLUND et al. (1994) verificaram que o poder de exclusão combinado, considerando oito locos, variou de 0,96 a 0,99, em quatro raças de cavalo.

Outros estudos envolvendo a caracterização de microssatélites em raças de cavalo (BREEN et al., 1994) e gado (GLOWATSKI-MULLIS et al., 1995) mostraram poder de exclusão combinado semelhante.

Estudos com plantas também têm mostrado que os microssatélites apresentam alto poder de exclusão. LITT e LUTY (1999), WEBER e MAY (1991), YAN e WANG (1994) e WANG et al. (1994) verificaram que o poder de exclusão combinado, considerando oito locos, variou de 0,96 a 0,99, em quatro raças de cavalo. Outros estudos envolvendo a caracterização de microssatélites em raças de cavalo (BREEN et al., 1994) e gado (GLOWATSKI-MULLIS et al., 1995) mostraram poder de exclusão combinado semelhante.

Os microssatélites são abundantes (LAGERCRANTZ et al., 1993; WANG et al., 1994) e encontram-se regularmente distribuídos no genoma de plantas (STEPHAN e CHO, 1990). Além disso, apresentam multialelismo e alto nível de polimorfismo, resultado de alta incidência de mutação na região, como foi demonstrado "in vitro" por SCHLÖTTERER e TAUTZ (1994).

CAPÍTULO 1

DESENVOLVIMENTO DE SISTEMAS DE GENOTIPAGEM MULTILOCOS SEMI-AUTOMATIZADOS, BASEADOS EM MARCADORES MICROSSATÉLITES

1. INTRODUÇÃO

Os marcadores microsatélites são baseados na amplificação de regiões compostas por um número variável de nucleotídios, normalmente de 1 a 6 pb, repetidos várias vezes e encontrados no genoma de eucariotos (HAMADA et al., 1982).

Embora o polimorfismo presente nessas regiões tenha sido explorado inicialmente na área humana (LITT e LUTY, 1989; WEBER e MAY, 1989), locos baseados em microsatélites foram posteriormente descritos em uma grande variedade de culturas agrícolas (WEISING et al., 1989; WU e TANKSLEY, 1993; KRESOVICH et al., 1995) e algumas espécies florestais de importância econômica (SMITH e DEVEY, 1994; BRONDANI et al., 1998).

Os microsatélites são abundantes (LAGERCRANTZ et al., 1993; WANG et al., 1994) e encontram-se regularmente distribuídos no genoma de plantas (STEPHAN e CHO, 1994). Além disso, apresentam multialelismo e alto nível de polimorfismo, resultante da alta incidência de escorregamento na replicação, como foi demonstrado, "in vitro", por SCHLÖTTERER e TAUTZ

(1992). Todas essas características, aliadas ao padrão de segregação co-dominante, têm feito dos microssatélites o marcador mais apropriado para o desenvolvimento de "fingerprints" genéticos únicos e a discriminação de cultivares (CREGAN et al., 1994; THOMAS et al., 1994; RUSSEL et al., 1997; DIWAN e CREGAN, 1997).

O uso de marcadores moleculares baseados em microssatélites para fins de genotipagem individual evoluiu para o desenvolvimento de sistemas de análise multilocos semi-automatizados. Esses sistemas foram desenvolvidos inicialmente para estudos de genética humana (EDWARDS et al., 1991). Eles se baseiam na amplificação via PCR de um ou mais locos microssatélites, simultaneamente, utilizando iniciadores marcados com fluorocromos e separação e visualização dos fragmentos em máquinas que permitam a detecção de fluorescência.

Recentemente, sistemas semelhantes foram desenvolvidos para análise de espécies de importância agrícola. O objetivo tem sido estabelecer sistemas universais de diferenciação de cultivares e variedades, conforme foi proposto para *Vitis sp.* (THOMAS et al., 1994) e soja (DIWAN e CREGAN, 1997). Entre as vantagens desses sistemas estão: 1) rapidez e acurácia na geração de dados; 2) avaliação simultânea de vários locos SSR em uma linha do gel, reduzindo o número de análises necessárias em relação às técnicas convencionais; 3) aquisição e armazenamento dos dados (computadorizados) em tempo real, reduzindo a manipulação manual da informação; 4) possibilidade de exportação direta dos dados para programas de análise genética; e 5) transferibilidade do sistema para outros laboratórios, permitindo a padronização dos locos do genoma amostrados e a comparação e o compartilhamento de resultados.

O desenvolvimento e a aplicação mais ampla de microssatélites em sistemas Multiplex têm sido limitados pelo custo e tempo necessários para o desenvolvimento desses marcadores em cada espécie. Essa limitação tem sido contornada em algumas situações em que existe conservação suficiente das seqüências flanqueadoras dos microssatélites entre espécies pertencentes ao mesmo gênero e até mesmo gêneros pertencentes à mesma família, o que possibilita a transferibilidade de marcadores (KEMP et al., 1993; BREEN et al., 1994; KIJAS et al., 1995). A transferibilidade de microssatélites entre diferentes

espécies do gênero *Eucalyptus* foi relatada inicialmente em número reduzido de locos por BYRNE et al. (1996) e confirmada em maior número de locos entre *E. grandis* e *E. urophylla* por BRONDANI et al. (1998). Tais resultados indicaram o grande potencial de desenvolvimento de sistemas de genotipagem multilocos baseados em microssatélites para análise genética de várias espécies de importância econômica dentro do gênero *Eucalyptus*.

O objetivo geral desse projeto foi o desenvolvimento de sistemas de genotipagem semi-automatizados de alto desempenho, baseados em microssatélites, com detecção fluorescente. Inicialmente, foram testados 170 pares de iniciadores originalmente desenvolvidos e desenhados por BRONDANI et al. (1998), a partir de bibliotecas enriquecidas de *Eucalyptus grandis* e *E. urophylla*. Nessa triagem foi selecionado um conjunto de pares de iniciadores que amplificavam, de forma específica e robusta, locos polimórficos com herança mendeliana. No processo de triagem foi avaliada, ainda, a transferibilidade desses locos, principalmente, para espécies do subgênero *Symphyomyrtus*, bem como para outras espécies comercialmente importantes de outros subgêneros. A partir da bateria de microssatélites selecionados, foram desenvolvidos três sistemas de genotipagem multilocos. Cada um desses sistemas foi composto de três locos microssatélites, co-amplificados em mesma reação de PCR e visualizados simultaneamente através de sistemas de detecção fluorescente, com determinação automatizada dos genótipos a cada loco.

Brasil, para avaliação de transferibilidade foram utilizados indivíduos das espécies *E. nitens* (subgênero *Symphyomyrtus*), *E. pellita* (subgênero *Monocalyptus*), *E. obesa* (subgênero *Idiogenes*) e *E. celastroides* (subgênero *Corymbia*), originários de estívolos de introdução de estívolos da EMBRAPA/CPPAC. A extração de DNA segue o protocolo descrito por DOYLE e DOYLE (1987).

Foram avaliados 170 pares de iniciadores desenvolvidos a partir de genótipos de *E. grandis* e *E. urophylla*. A descoberta e isolamento dos microssatélites e o sequenciamento e desenho dos iniciadores foram descritos por BRONDANI et al. (1998). Os microssatélites foram amplificados via PCR em reações de 13 μ L contendo 7,5 ng de DNA genômico, 200 μ M de dNTP, 1,3 μ L de DMSO 20%, 1,3 μ L de tampão de PCR 10X contendo 16 mM de

2. MATERIAL E MÉTODOS

2.1. Seleção inicial de locos microsatélites

2.1.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificados

O DNA genômico foi extraído de folhas de parentais (*E. grandis* e *E. urophylla* - subgênero *Symphyomyrtus*) e de seis progênies F1, provenientes de populações de melhoramento da empresa Aracruz Celulose S.A. (Espírito Santo, Brasil). Para avaliação de transferibilidade, foram utilizados indivíduos das espécies *E. dunnii* (subgênero *Symphyomyrtus*), *E. pilularis* (subgênero *Monocalyptus*), *E. cloeziana* (subgênero *Idiogenes*) e *E. citriodora* (subgênero *Corymbia*), originários de ensaios de introdução de espécies da EMBRAPA/CPAC. A extração de DNA seguiu o protocolo descrito por DOYLE e DOYLE (1987).

Foram analisados 170 pares de iniciadores desenvolvidos a partir de genótipos de *E. grandis* e *E. urophylla*. A descoberta e isolamento dos microsatélites e o seqüenciamento e desenho dos iniciadores foram descritos por BRONDANI et al. (1998). Os microsatélites foram amplificados via PCR em reações de 13 μ L contendo 7,5 ng de DNA genômico, 200 μ M de dNTP, 1,3 μ L de DMSO 50%, 1,3 μ L de tampão de PCR 10X contendo 15 mM de

MgCl₂, 1,34 µL de água desionizada, 1 mM de MgCl₂, 0,29 µM de iniciadores anterior e posterior e 1 U de enzima Taq polimerase. O protocolo de PCR incluiu uma etapa inicial de desnaturação a 96°C por dois minutos, seguida de 34 ciclos constituídos de uma etapa de desnaturação a 96°C por um minuto, anelamento a 60-56°C por um minuto e extensão a 72 °C por 1 minuto. Durante os primeiros quatro ciclos, a temperatura de anelamento decresceu 2°C a cada dois ciclos até atingir 56°C. Os fragmentos amplificados foram separados em gel de agarose 2,5% (comum ou Metaphor-FMC Bioproducts), em TBE 1X a 130 V, por 2,0-2,5 horas, e corados com brometo de etídio.

A qualidade dos produtos de amplificação dos microssatélites, assim como o seu padrão de segregação, foi avaliada nos parentais (*E. grandis* e *E. urophylla*) e em seis progênies F1. Os microssatélites foram classificados de acordo com os seguintes critérios subjetivos, descritos por PÉPIN et al. (1995): 1) amplificação específica da região microssatélite, com presença de apenas uma (homozigoto) ou duas (heterozigoto) bandas; 2) amplificação específica da região microssatélite, com a presença de amplificação inespecífica; 3) amplificação inespecífica, sem distinção do microssatélite; e 4) ausência de amplificação.

Um conjunto de 100 locos microssatélites também foi avaliado quanto à sua transferibilidade para *E. dunnii*, *E. pilularis*, *E. cloeziana* e *E. citriodora*, observando-se se estes amplificavam locos SSR nessas espécies.

Com base nos resultados de análise de qualidade de amplificação das regiões microssatélites, 17 locos foram selecionados, para os quais foram sintetizados iniciadores marcados com fluorocromos de diferentes cores.

2.2. Avaliação dos locos microssatélites

2.2.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificados

O DNA genômico de três indivíduos de cada uma de duas procedências das espécies *E. camaldulensis*, *E. dunnii*, *E. grandis*, *E. saligna* e *E. urophylla*, todas pertencentes ao subgênero *Symphyomyrtus*, foi extraído de tecido foliar

de mudas originadas de sementes coletadas em populações naturais na Austrália e Indonésia (Quadro 1).

Quadro 1 – Procedências geográficas dos materiais de cada espécie, utilizados na triagem dos locos EMBRA (ND = não disponível)

Espécie	Procedência	Localização geográfica		País
		Latitude	Longitude	
<i>E. camaldulensis</i>	Irvinebank to Petford	17° 24'	145° 09'	Austrália
<i>E. camaldulensis</i>	Walsh River	16° 59'	144° 18'	Austrália
<i>E. dunnii</i>	ND	ND	ND	Austrália
<i>E. dunnii</i>	ND	ND	ND	Austrália
<i>E. grandis</i>	Atherton	17° 19'	145° 25'	Austrália
<i>E. grandis</i>	Wild River	17° 22'	145° 24'	Austrália
<i>E. saligna</i>	Chaelundi	30° 00'	152° 22'	Austrália
<i>E. saligna</i>	Sheas Nob S. F.	30° 02'	152° 32'	Austrália
<i>E. urophylla</i>	Timor	ND	ND	Indonésia
<i>E. urophylla</i>	Flores	ND	ND	Indonésia

Os indivíduos de *E. dunnii*, *E. grandis* e *E. saligna* foram escolhidos, ao acaso, de plantas originadas de uma mistura de sementes provenientes de 10 famílias de meios-irmãos, coletadas em cada procedência, cedidas pela empresa Riocell S.A. As plantas de *E. camaldulensis* e *E. urophylla* constituem indivíduos geneticamente não-relacionados, cujas sementes foram fornecidas pelas empresas Mannesmann Florestal LTDA., CAPEF e Bahia Sul Celulose S.A., respectivamente. Para extração de DNA genômico foi utilizada a metodologia descrita por DOYLE e DOYLE (1987).

Nessa etapa foram avaliados 17 locos EMBRA, escolhidos entre aqueles que apresentaram boa qualidade de amplificação e padrão de segregação na etapa anterior de triagem. A amplificação dos microssatélites ocorreu em reações de PCR, com volume final de 12,5 µL contendo 160 µM de dNTP e 1,25 µL de tampão de PCR 10X contendo 15 mM de MgCl₂, 5,63 µL de água desionizada, 1 U de enzima Taq polimerase, 0,2 µM de cada um dos iniciadores anterior e posterior e 12 ng de DNA genômico.

Um dos iniciadores de cada par teve seu terminal 5' marcado com um dos fluorocromos: 6-carboxifluoresceína (6-FAM), tetracloro-6-carboxifluoresceína (TET) ou hexacloro-6-carboxifluoresceína (HEX) (OPC, Perkin-Elmer/Applied Biosystems). A amplificação dos microssatélites passou por uma etapa inicial de desnaturação a 96°C por dois minutos, seguida de 30 ciclos com etapas de desnaturação (96°C), anelamento (56°C) e extensão (72°C), por um minuto cada. Ao final do último ciclo ocorreu uma nova etapa de extensão a 72°C, por sete minutos. Os fragmentos amplificados foram, então, diluídos a 1:10 em água desionizada estéril.

Um μL da diluição foi combinado com 2,75 μL de tampão de carregamento desnaturante (98% de formamida; 10 mM EDTA; 0,25 mg de xileno-cianol e 0,25 mg de azul-de-bromofenol) e 0,25 μL de marcador-padrão interno GeneScan Tamra 350 ("Perkin-Elmer/Applied Biosystems"), marcado com N,N,N',N'-tetrametil-6-carboxirodamina (TAMRA). As amostras foram, então, desnaturadas a 95°C por cinco minutos e colocadas em gelo. Os fragmentos amplificados foram separados (2 μL /poço) em gel de poliacrilamida (5%) contendo 6 M de uréia, a 3.000 V e 50-30 mA por duas horas, instalado em um seqüenciador de DNA ("Perkin-Elmer/Applied Biosystems", modelo 377). O tamanho dos fragmentos foi estimado através do programa GeneScan ("Perkin-Elmer/Applied Biosystems"), sendo os dados posteriormente aferidos através do programa Genotyper ("Perkin-Elmer/Applied Biosystems").

Com base nos genótipos gerados para os indivíduos avaliados nessa etapa, foi registrado o número de alelos observados em cada loco, assim como a amplitude de tamanho dos alelos em pares de bases.

2.3. Desenvolvimento de sistemas de genotipagem multiloco e otimização das reações de co-amplificação via PCR

Inicialmente, os 17 locos EMBRA, selecionados e avaliados nas etapas anteriores, foram classificados de acordo com os critérios: 1) amplitude de tamanho dos alelos em pares de bases; 2) grupo de ligação em que o loco está localizado, quando disponível (BRONDANI et al., 1998); e 3) marcação fluorescente (HEX, 6-FAM e TET) do iniciador.

A amplitude alélica indica qual a faixa de tamanho de alelos que pode ser encontrada para um loco microssatélite amplificado com determinado par de iniciadores. No desenvolvimento de cada sistema de genotipagem multiloco, procurou-se co-amplificar locos que apresentassem amplitude alélica semelhante, para possibilitar a posterior combinação de dois sistemas com faixa de amplificação diferente e, dessa forma, a análise de até seis locos por linha do gel. Buscou-se, também, combinar locos que estivessem localizados em diferentes cromossomos e marcados com diferentes fluorocromos.

Os 17 locos EMBRA foram classificados com base na faixa de amplificação de produtos acima ou abaixo de 150 pb. Locos que apresentaram faixa muito ampla de tamanhos de alelos, ou seja, com produtos acima e abaixo de 150 pb, foram descartados. Co-amplificações envolvendo microssatélites com amplitude alélica semelhante, mas marcados com o mesmo fluorocromo ou, ainda, localizados em um mesmo grupo de ligação, também foram descartadas.

2.3.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificados

Neste etapa, as combinações de locos microssatélites foram avaliadas com cinco plantas, sendo uma de cada uma das seguintes espécies: *E. camaldulensis* (procedência Kennedy River), *E. dunnii* (procedência desconhecida), *E. grandis* (procedência Wild River), *E. saligna* (procedência Chaelundi) e *E. urophylla* (procedência Flores).

Foi avaliado um total de 15 combinações de 11 locos microssatélites marcados com diferentes fluorocromos. Cada combinação teve seus locos microssatélites co-amplificados via PCR em reações de 25 μ L contendo 24 ng de DNA genômico e 160 μ M de dNTP; 2,5 μ L de tampão de PCR 10X contendo 15 mM de $MgCl_2$; e 9,83 μ L de água desionizada contendo 0,4 μ M de iniciadores anterior e posterior marcados com fluorocromos HEX e TET e 0,12 μ M de iniciadores anterior e posterior marcados com o fluorocromo 6-FAM, além de 1,875 U da enzima Taq polimerase. O protocolo de PCR incluiu uma etapa inicial de desnaturação a 96°C por dois minutos, seguida de 30 ciclos

constituídos de uma etapa de desnaturação a 96°C, anelamento a 56°C e extensão a 72°C, por um minuto cada. Após o último ciclo ocorreu nova etapa de extensão a 72°C por sete minutos. Após a amplificação, os produtos foram diluídos a 1:4, em água desionizada estéril.

Um μL da diluição foi combinado com 2,75 μL de tampão de carregamento desnaturante e 0,25 μL de marcador padrão interno marcado com TAMRA. Após a desnaturação, as amostras amplificadas foram separadas (2 μL /poço) em gel de poliacrilamida (5%) contendo 6 M de uréia, a 3.000 V e 50-30 mA por duas horas, instalado no seqüenciador de DNA. A análise dos dados foi feita por meio dos programas GeneScan e Genotyper.

Após a seleção dos três melhores sistemas Multiplex, as condições de amplificação via PCR dos locos que os compunham foram otimizadas em relação às concentrações de enzima Taq polimerase (1,25 U; 1,565 U; e 1,875 U) e dos iniciadores marcados com o fluorocromo 6-FAM (0,08 μM e 0,12 μM). Os indivíduos analisados e as condições de amplificação, eletroforese e detecção dos fragmentos amplificados seguiram o mesmo procedimento descrito anteriormente.

2.3.2. Seleção de três sistemas Multiplex e ajuste do protocolo de amplificação via PCR

Para escolher três sistemas dentro do conjunto de combinações testadas, alguns critérios foram avaliados: 1) intensidade do sinal dos fluorocromos (HEX, 6-FAM e TET) ligados aos fragmentos amplificados, 2) coeficiente de variação entre as intensidades de fluorescência produzidas pelos fluorocromos (HEX, 6-FAM e TET) nas cinco espécies avaliadas e 3) presença mínima de sinais provenientes de amplificações inespecíficas após análise e filtragem dos sinais por meio do programa Genotyper.

Uma intensidade adequada (superior a 100 unidades de fluorescência) do sinal dos fluorocromos ligados aos fragmentos amplificados é importante para que o sistema de detecção do seqüenciador de DNA discrimine, com clareza, os sinais provenientes dos alelos amplificados, sendo esse o primeiro critério avaliado nesta pesquisa. O segundo critério indicou quanto apropriado é

o sistema multiloco na análise de todas as espécies de *Eucalyptus* para as quais ele foi desenvolvido. Por último, a presença mínima de sinais provenientes de ampliações inespecíficas auxiliou na avaliação da presença de sinais provenientes de ampliações que não correspondem aos alelos dos locos EMBRA que compunham o sistema Multiplex.

Para reunir todos esses critérios em um mesmo parâmetro, foi criada uma equação que levasse em conta todos os aspectos descritos anteriormente. Esse parâmetro, denominado índice Multiplex (IM), foi calculado da seguinte maneira:

$$IM = \overline{IF} \cdot (1 - \overline{CV}) \cdot (1 - \overline{A})$$

sendo

$$\overline{IF} = \frac{\sum_{i=1}^5 \sum_{j=1}^3 IF_{ij}}{15}$$

$$\overline{CV} = \frac{\sum_{j=1}^3 CV_j}{3}$$

$$\overline{A} = \frac{\sum_{j=1}^3 H_j}{3}$$

$$H_j = \frac{|O_j - E_j|}{E_j}$$

em que

IF_{ij} = intensidade de fluorescência, detectada por meio do programa GeneScan, da i -ésima espécie no j -ésimo loco do sistema Multiplex;

CV_j = coeficiente de variação observado (entre os indivíduos representantes das cinco espécies estudadas) na intensidade de fluorescência no j -ésimo loco do sistema Multiplex;

O_j = número de fragmentos observados no j -ésimo loco do sistema Multiplex; e

E_j = número de fragmentos esperados no j -ésimo loco do sistema Multiplex.

Considerando que intensidade de fluorescência igual ou superior a 500 unidades é suficiente para que haja visualização adequada dos fragmentos amplificados, tomou-se esse valor, na equação de cálculo do índice Multiplex,

como o valor máximo a ser considerado na intensidade de fluorescência detectada em cada fluorocromo.

Após a seleção de três sistemas Multiplex, as diferentes condições de amplificação testadas foram igualmente avaliadas através do cálculo do seu índice Multiplex.

2.3.3. Variação no tamanho dos fragmentos detectados em diferentes condições de PCR

Com base no tamanho dos fragmentos que foram detectados na avaliação individual dos 17 locos microssatélites, foi feita uma comparação entre os tamanhos dos alelos que haviam sido observados na ocasião e o tamanho daqueles detectados nas diferentes condições de PCR testadas.

Registrou-se a diferença máxima entre a estimativa de tamanho dos alelos dos locos que compunham os sistemas Multiplex e a estimativa obtida quando esses locos foram analisados em reações de PCR em que foram amplificados isoladamente, assim como a diferença média e o desvio-padrão.

Vários dos 170 pares de iniciadores avaliados amplificaram especificamente os locos microssatélites. Outros 41,8% originaram produtos de tamanho inesperado, provavelmente devido a amplificação de regiões não especificamente alvo. Para esses pares de iniciadores foi possível observar o padrão de ligação esperado. Dos 170 pares de iniciadores testados, 10,6% não amplificaram qualquer fragmento, ao passo que outros 18,8% amplificaram regiões inespecíficas do genoma, não possibilitando a obtenção de microssatélites.

Resultados semelhantes foram relatados em outros estudos (OLSEN et al., 1995; WENNING et al., 1999), nos quais foi observada a ausência de amplificação dos microssatélites ou, ainda, a amplificação de regiões inesperadas sob determinadas condições. Em alguns casos, a amplificação de regiões inesperadas pode ser atribuída a reações de PCR em que grande quantidade de DNA não específico é amplificado, o que pode ocorrer devido a

3. RESULTADOS E DISCUSSÃO

3.1. Seleção inicial de locos microssatélites

3.1.1. Qualidade de amplificação e padrão de segregação dos microssatélites em *E. grandis* e *E. urophylla*

Verificou-se que 28,8% dos pares de iniciadores avaliados amplificaram especificamente seqüências microssatélites. Outros 41,8% originaram produtos claramente oriundos de amplificação dos locos microssatélites, além de amplificações inespecíficas. Para esses pares de iniciadores foi possível observar o padrão de segregação esperado. Dos 170 pares de iniciadores testados, 10,6% não amplificaram qualquer fragmento, ao passo que outros 18,8% amplificaram regiões inespecíficas do genoma, não possibilitando a distinção do microssatélite.

Resultados semelhantes foram relatados em outros estudos (OLSEN et al., 1996; WENBURG et al., 1996), nos quais foi observada a ausência de amplificação dos microssatélites ou, ainda, a amplificação de regiões inespecíficas, observação comum em, essencialmente, todos os trabalhos de desenvolvimento desses marcadores, independentemente da espécie considerada. O padrão de amplificação em que grande quantidade de fragmentos é amplificada, com ou sem a distinção clara da amplificação de regiões de DNA repetitivo, indica a não-especificidade, ou a especificidade

parcial, dos iniciadores em relação às regiões flangeadoras desses locos. Uma solução para esse problema seria testar temperaturas crescentes de anelamento dos iniciadores, com o intuito de aumentar a estringência da reação e, conseqüentemente, a especificidade do anelamento (WEISING et al., 1995). Entretanto, como o objetivo do presente trabalho foi o desenvolvimento de sistemas nos quais vários locos microssatélites fossem amplificados simultaneamente, a otimização de temperaturas de anelamento individualmente para cada loco não era desejável.

3.1.2. Conservação dos locos microssatélites em outras espécies do gênero *Eucalyptus*

Na avaliação da transferibilidade dos locos microssatélites para outras espécies do gênero *Eucalyptus*, observou-se que, de um conjunto de 100 pares de iniciadores, 81,7% daqueles que produziram amplificação clara e interpretável (classificações 1 e 2) em *E. grandis* e *E. urophylla* também amplificaram regiões microssatélites em *E. dunnii*, que pertence ao mesmo subgênero (*Symphyomyrtus*). A conservação foi menor nas espécies *E. cloeziana* (subgênero *Idiogenes*) e *E. pilularis* (subgênero *Monocalyptus*), 56,3 e 54,9%, respectivamente. Na espécie *E. citriodora* (subgênero *Corymbia*), com a qual foi avaliada a transferibilidade de 15 microssatélites, apenas seis amplificaram produtos interpretáveis.

A alta transferibilidade (81,7%) desses locos, desenvolvidos a partir de *E. grandis* e *E. urophylla*, para *E. dunnii* claramente indicou que existe grande potencial de transferibilidade desses marcadores entre diferentes espécies do gênero *Eucalyptus*, especialmente dentro do mesmo subgênero. Isso é particularmente importante ao se levar em conta que a maioria das espécies utilizadas em plantios comerciais no mundo pertence ao subgênero *Symphyomyrtus* (ELDRIGE et al., 1994) e que o desenvolvimento de marcadores microssatélites ainda apresenta custos financeiro e tecnológico proibitivos para a maioria dos laboratórios.

Os resultados discordaram parcialmente dos obtidos por BYRNE et al. (1996), que observaram que um conjunto de quatro locos microssatélites desenvolvidos a partir de genótipos de *E. nitens* amplificou, com sucesso,

regiões microssatélites em outras três espécies do mesmo subgênero (*Symphomyrtus*). No entanto, a transferibilidade para o subgênero *Monocalyptus* foi de 50% e, para o subgênero *Corymbia*, nula. Essa incoerência de resultados foi, provavelmente, devida à diferença no tamanho da amostra analisada.

A conservação dos locos microssatélites foi significativamente maior entre *E. grandis* e *E. urophylla* (*Symphomyrtus*), em relação às espécies dos subgêneros *Idiogenes* (*E. cloeziana*) e *Monocalyptus* (*E. pilularis*), confirmando estudos que indicaram maior similaridade genética entre esses subgêneros, em relação ao *Corymbia* (*E. citriodora*). O fato de a maioria dos locos microssatélites ter sido conservada na espécie do subgênero *Idiogenes* e também no subgênero *Monocalyptus* se explica em razão de esses dois subgêneros serem muito próximos filogeneticamente (LADIGES, 1983; Udivicic, McFadden e Ladiges, citados por LADIGES et al., 1995). Finalmente, existem evidências filogenéticas suficientes, a partir de estudos moleculares, que têm levado à separação do subgênero *Corymbia* em um novo gênero (LADIGES et al., 1995). Com base nessas evidências, observou-se que 17 locos EMBRA (Quadro 2) foram selecionados com base na robustez e alta transferibilidade entre espécies para desenvolvimento e otimização de sistemas de genotipagem multilocos.

3.2. Avaliação dos locos microssatélites

3.2.1. Número de alelos e amplitude alélica

Os 17 locos microssatélites permitiram a amplificação de um total de 284 fragmentos, com uma média de 16,71 alelos por loco (Quadro 2). O menor número de alelos foi encontrado no loco EMBRA22 e o maior, no loco EMBRA16, com 11 e 27 alelos, respectivamente. A maior amplitude alélica foi também observada no loco EMBRA16 (com 104 pb) e a menor, no EMBRA6 (com 40 pb).

O número de alelos detectados nesse grupo de indivíduos foi semelhante ao encontrado em outros trabalhos de caracterização de locos microssatélites, desenvolvidos para espécies do gênero *Eucalyptus*.

Quadro 2 – Avaliação dos 17 locos EMBRA selecionados em relação ao número de alelos e à amplitude alélica observada em 30 indivíduos provenientes de cinco espécies de *Eucalyptus*

EMBRA	Número de alelos	Amplitude alélica (pb)
27	15	179-229
28	19	109-173
21	14	110-176
22	11	143-197
5	22	91-151
6	13	113-153
4	18	83-135
23	13	107-149
24	18	183-227
10	14	102-154
11	13	99-153
25	21	104-194
26	16	112-172
16	27	108-212
15	19	71-125
12	14	114-160
18	17	96-136
Média	16,71	

BYRNE et al. (1996), avaliando quatro microssatélites em 20 indivíduos de *E. nitens*, encontraram 5 a 16 alelos por loco. BRONDANI et al. (1998), ao caracterizarem 15 locos, dos quais nove também foram utilizados neste estudo, observaram de 9 a 26 alelos por loco, com uma média de 16,3 alelos. O número médio de alelos detectados em *Eucalyptus* foi expressivamente maior do que o normalmente obtido na caracterização de microssatélites em culturas agrícolas, como brássicas (KRESOVICH et al., 1995), soja (MORGANTE et al., 1994), arroz (WU e TANKSLEY, 1993) e cevada (BECKER e HEUN, 1995). A razão mais provável dessa diferença está no fato de esses microssatélites terem sido caracterizados em variedades ou cultivares de espécies com longa história de domesticação e, portanto, de base genética estreita. Entretanto, cabe destacar que locos altamente polimórficos, com mais de 30 alelos, têm sido descritos em genótipos silvestres de cevada (SAGHAI-MAROOF et al., 1994), provavelmente devido à maior variabilidade genética presente em materiais selvagens em relação aos cultivados.

O número variável de alelos observados em cada loco pode ser explicado pela distribuição ampla e regular dos microssatélites no genoma de plantas (WANG et al., 1994), resultando sua localização em regiões mais ou menos conservadas e, conseqüentemente, sujeitas a mais ou menos variação. De fato, várias seqüências repetitivas de dinucleotídeos foram descritas posicionadas em genes ou próximas a estes (LITT e LUTY, 1989; VAIMAN et al., 1995).

3.3. Desenvolvimento de sistemas Multiplex

Três sistemas de genotipagem multilocos baseados em marcadores microssatélites foram criados para *E. camaldulensis*, *E. dunnii*, *E. grandis*, *E. saligna* e *E. urophylla*. Cada sistema foi desenvolvido, buscando-se a co-amplificação via PCR de três locos microssatélites localizados em diferentes cromossomos, com amplitude alélica semelhante, e marcados com diferentes fluorocromos. Ao todo, nove locos microssatélites foram otimizados em três reações de PCR.

3.3.1. Classificação dos locos microssatélites

Os microssatélites que apresentaram grande amplitude alélica, com alelos de tamanho acima e abaixo de 150 pb, foram descartados. Enquadraram-se nesse grupo sete dos 17 locos anteriormente avaliados: EMBRA28, EMBRA21, EMBRA25, EMBRA26, EMBRA16 e EMBRA12 (Quadro 2). Os locos selecionados, que amplificaram produtos na faixa abaixo de 150 pb, foram: EMBRA5, EMBRA6, EMBRA4, EMBRA23, EMBRA10, EMBRA11, EMBRA15 e EMBRA18. Foi permitida uma tolerância de até 5 pb, selecionado-se locos que amplificassem produtos com até 155 pb. Caso contrário, os locos EMBRA5, EMBRA6, EMBRA10 e EMBRA11 teriam de ser descartados do conjunto, e a gama de possíveis combinações testadas seria reduzida significativamente.

Foram encontrados apenas dois locos amplificando produtos somente na faixa superior a 150 pb: EMBRA24 e EMBRA27. De forma a possibilitar o desenvolvimento de um sistema de genotipagem envolvendo a co-amplificação

de três locos e a análise de produtos acima de 150 pb, o loco EMBRA22 foi incluído nesse grupo. Esse loco, marcado com fluorocromo TET e com alelos de tamanho entre 143 e 197 pb, poderia eventualmente ser combinado com os produtos da amplificação do loco EMBRA15, marcado com o mesmo fluorocromo, mas que amplifica produtos com, no máximo, 125 pb.

No Quadro 3, encontram-se descritos os microssatélites selecionados para o desenvolvimento dos sistemas de genotipagem multilocos, com a respectiva faixa de amplificação, marcação fluorescente do iniciador e o grupo de ligação no qual o loco está localizado, segundo BRONDANI et al. (1998).

Quadro 3 – Locos EMBRA selecionados para desenvolvimento de sistemas de genotipagem multilocos, classificados de acordo com a amplitude alélica (+ = acima de 150 pb; - = abaixo de 150 pb), a marcação fluorescente do iniciador (HEX, 6-FAM e TET) e o grupo de ligação em que o loco está posicionado, segundo BRONDANI et al. (1998) (ND = não disponível)

Microssatélite	Fluorocromo	Amplitude alélica	Grupo de ligação
EMBRA27	6-FAM	+	ND
EMBRA22	TET	+	ND
EMBRA5	6-FAM	-	5
EMBRA6	TET	-	1
EMBRA4	HEX	-	4
EMBRA23	HEX	-	ND
EMBRA24	HEX	+	ND
EMBRA10	HEX	-	10
EMBRA11	HEX	-	1
EMBRA15	TET	-	8
EMBRA18	6-FAM	-	9

3.3.2. Seleção de três sistemas de genotipagem multilocos e otimização das reações de PCR

Embora os locos EMBRA11 e EMBRA6 apresentassem alelos na mesma faixa de tamanho e seus iniciadores estivessem marcados com diferentes fluorocromos, a combinação de ambos foi descartada por eles estarem localizados em um mesmo grupo de ligação. Os resultados referentes

à avaliação das 15 combinações de locos EMBRA testadas encontram-se no Quadro 4. O sistema de genotipagem multiloco K, composto pelos locos EMBRA10, EMBRA18 e EMBRA6, foi considerado o melhor entre as combinações testadas, na faixa abaixo de 150 pb. O Multiplex O, único com locos EMBRA amplificando na faixa acima de 150 pb, também apresentou bons resultados na amplificação dos microssatélites nas cinco espécies consideradas. Essa combinação foi inferior ao Multiplex K, quanto à intensidade de fluorescência detectada e ao coeficiente de variação entre as intensidades de fluorescência detectadas nas cinco espécies. O terceiro sistema foi selecionado no conjunto de combinações que não continham os locos EMBRA10, EMBRA18 e EMBRA6, já incluídos no Multiplex K. Restaram, assim, as combinações B, F e M, dentre as quais o Multiplex B, composto pelos locos EMBRA4, EMBRA5 e EMBRA15, foi selecionado, por apresentar o maior índice Multiplex. A Figura 1 mostra dois sistemas Multiplex testados e o contraste da qualidade de amplificação em boa e má combinações de locos EMBRA.

Com base nos resultados obtidos na avaliação das diferentes condições de PCR para amplificação dos locos EMBRA que compõem os sistemas Multiplex, determinaram-se, quanto ao índice Multiplex detectado, as concentrações de enzima Taq polimerase e os iniciadores marcados com fluorocromo 6-FAM (Quadro 5). A Figura 2 mostra os resultados obtidos na amplificação de um mesmo sistema Multiplex com duas concentrações diferentes de enzima Taq polimerase. Os resultados referentes às combinações de reagentes nas reações de PCR de cada sistema Multiplex encontram-se no Quadro 6.

Além da otimização das concentrações da enzima Taq polimerase e do iniciador marcado com o fluorocromo 6-FAM, optou-se ainda, posteriormente, pela adição de 0,25 mg/ml de BSA às reações, com o intuito de melhorar a qualidade de amplificação dos locos microssatélites.

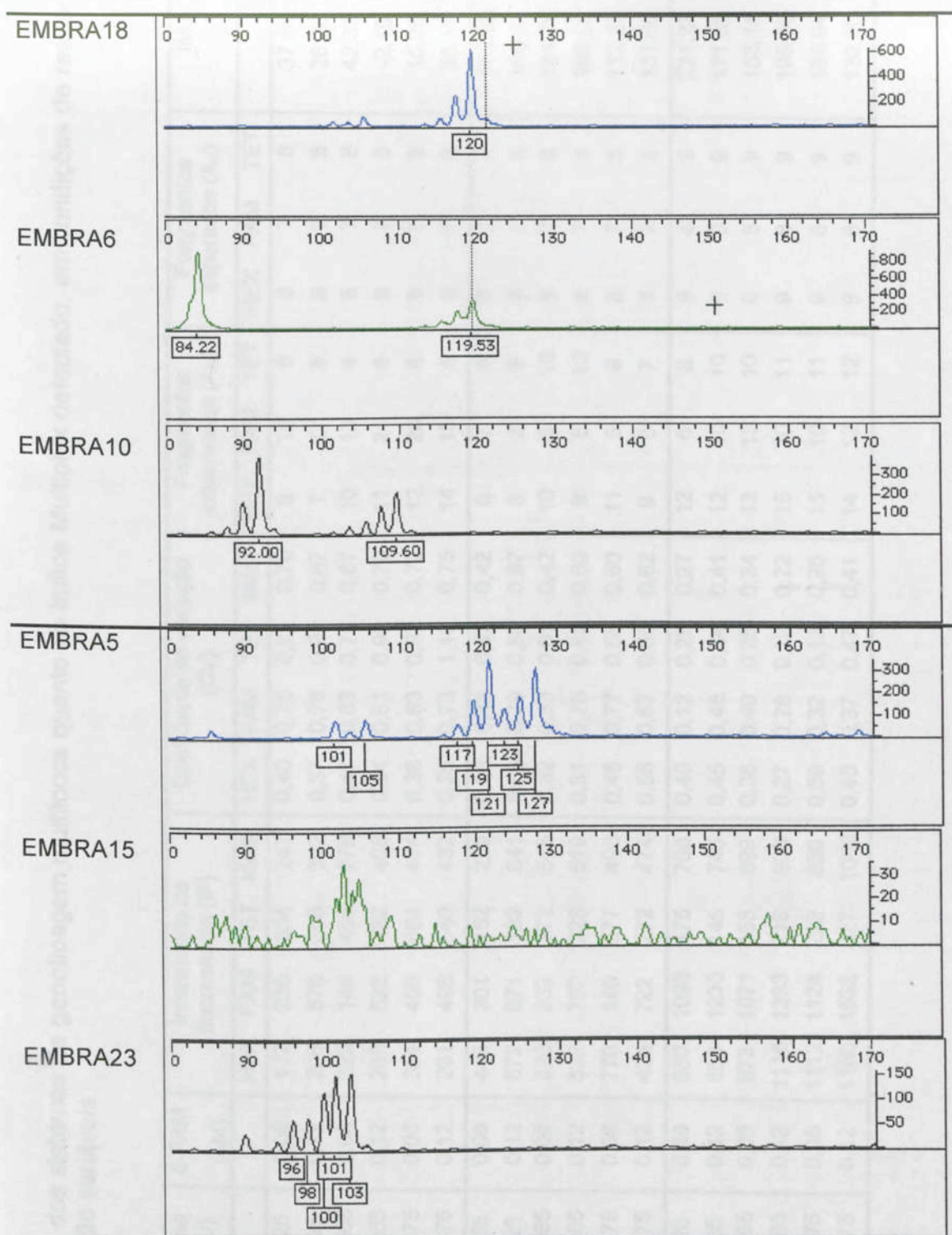


Figura 1 – Contraste entre a qualidade de amplificação de dois sistemas Multiplex. A combinação entre os locos EMBRA10, EMBRA18 e EMBRA6 gera produtos de amplificação específicos, o que não é observado no Multiplex composto pelos locos EMBRA23, EMBRA5 e EMBRA15.

Quadro 5 – Avaliação dos sistemas de genotipagem multilocos quanto ao Índice Multiplex detectado, em condições de reação de amplificação variáveis

Multiplex EMBRA	Taq (U)		6-FAM (μ M)	Intensidade de fluorescência (IF)			Coeficiente de variação (CV)			Fragmentos observados (A_0)			Fragmentos esperados (A_E)			IM		
	HEX	FAM		TET	HEX	FAM	TET	Médi	HEX	FAM	TET	HEX	FAM	TET				
4	5	15	1,25	172	236	334	247	0,40	0,75	0,93	0,70	9	14	5	8	7	8	37,68
			1,25	206	375	356	312	0,37	0,76	0,90	0,67	7	17	5	8	7	8	36,28
			1,565	229	144	455	276	0,47	0,83	0,71	0,67	10	13	4	8	7	8	42,38
			1,565	266	526	432	408	0,54	0,61	0,96	0,70	11	24	6	8	7	8	-2,10
			1,875	334	405	684	474	0,36	0,83	0,97	0,72	12	20	6	8	7	8	15,26
24	27	22	1,25	445	301	152	229	0,36	0,43	0,47	0,42	8	7	8	8	7	8	173,95
			1,25	572	871	182	541	0,56	0,79	0,36	0,57	8	8	8	8	7	8	160,04
			1,565	529	832	272	545	0,32	0,33	0,61	0,42	10	9	10	8	7	8	181,78
			1,565	532	797	228	519	0,31	0,78	0,42	0,50	9	8	10	8	7	8	169,00
			1,875	770	140	577	496	0,45	0,77	0,58	0,60	11	5	9	8	7	8	112,25
10	18	6	1,25	426	722	272	474	0,55	0,67	0,65	0,62	9	8	7	8	7	8	131,66
			1,25	850	1093	175	706	0,40	0,12	0,28	0,27	12	6	9	9	8	9	231,23
			1,25	860	1230	145	745	0,45	0,45	0,34	0,41	12	10	10	9	8	9	171,81
			1,565	873	1071	153	699	0,38	0,40	0,25	0,34	13	13	10	9	8	9	153,14
			1,565	1115	1393	218	908	0,27	0,28	0,12	0,22	15	10	11	9	8	9	195,66
			1,875	1112	1124	222	820	0,30	0,32	0,13	0,25	15	10	11	9	8	9	189,94
			1,875	1166	1532	327	1008	0,40	0,37	0,47	0,41	14	12	12	9	8	9	139,90

Figura 2 – Componentes da reação de PCR dos sistemas Multiplex nos quais as primers devem ser ajustadas para 25 µL com o volume adequado para a reação.

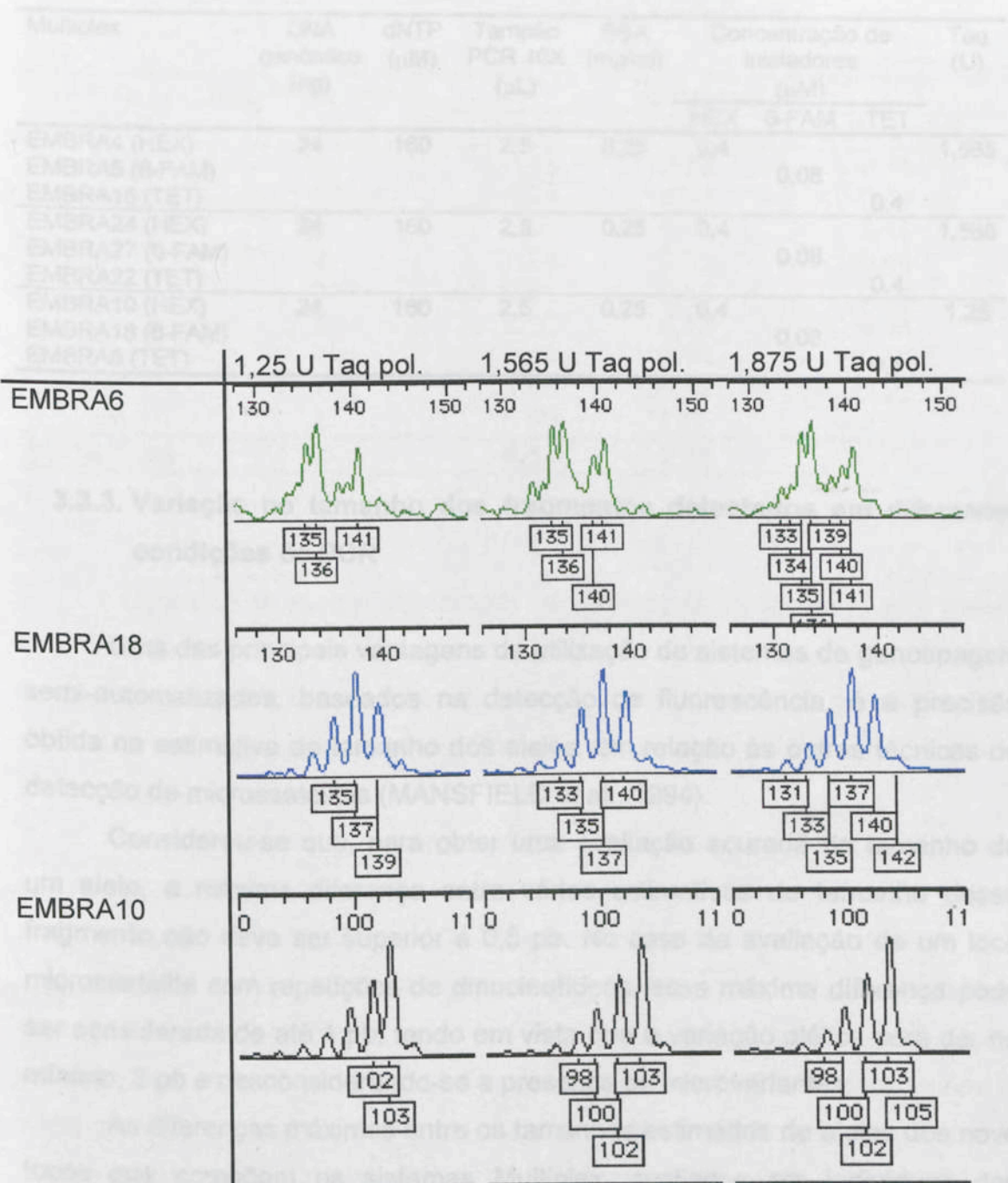


Figura 2 – Aumento do número de fragmentos inespecíficos detectados após a filtragem dos dados pelo programa Genotyper, com incremento da concentração da enzima Taq polimerase na reação de amplificação do Multiplex composto pelos locos EMBRA10, EMBRA18 e EMBRA6.

Quadro 6 – Componentes da reação de PCR dos sistemas Multiplex (os volumes finais devem ser ajustados para 25 μ L com água desionizada estéril)

Multiplex	DNA genômico (ng)	dNTP (μ M)	Tampão PCR 10X (μ L)	BSA (mg/ml)	Concentração de iniciadores (μ M)			Taq (U)
					HEX	6-FAM	TET	
EMBRA4 (HEX)	24	160	2,5	0,25	0,4	0,08	0,4	1,565
EMBRA5 (6-FAM)								
EMBRA15 (TET)								
EMBRA24 (HEX)	24	160	2,5	0,25	0,4	0,08	0,4	1,565
EMBRA27 (6-FAM)								
EMBRA22 (TET)								
EMBRA10 (HEX)	24	160	2,5	0,25	0,4	0,08	0,4	1,25
EMBRA18 (6-FAM)								
EMBRA6 (TET)								

3.3.3. Variação no tamanho dos fragmentos detectados em diferentes condições de PCR

Uma das principais vantagens da utilização de sistemas de genotipagem semi-automatizados, baseados na detecção de fluorescência, é a precisão obtida na estimativa do tamanho dos alelos, em relação às outras técnicas de detecção de microssatélites (MANSFIELD et al., 1994).

Considerou-se que, para obter uma avaliação acurada do tamanho de um alelo, a máxima diferença entre várias estimativas de tamanho desse fragmento não deve ser superior a 0,5 pb. No caso da avaliação de um loco microssatélite com repetições de dinucleotídeos, essa máxima diferença pode ser considerada de até 1 pb, tendo em vista que a variação alélica será de, no mínimo, 2 pb e desconsiderando-se a presença de microvariantes.

As diferenças máximas entre os tamanhos estimados de alelos dos nove locos que compõem os sistemas Multiplex, avaliados em indivíduos das espécies *E. dunnii*, *E. grandis*, *E. saligna* e *E. urophylla*, foram, salvo no loco EMBRA27, superiores a 0,5pb. Outros cinco locos, EMBRA5, EMBRA4, EMBRA24, EMBRA15 e EMBRA18, apresentaram diferenças entre 0,5 e 1 pb. Diferenças superiores a 1 pb foram observadas nos três locos restantes, chegando a até 3,13 pb no loco EMBRA10 (Quadro 7).

Quadro 7 – Diferenças observadas na estimativa do tamanho de alelos na análise de indivíduos de quatro espécies de *Eucalyptus*, com nove locos EMBRA, co-amplificados nos sistemas Multiplex, em condições variadas de PCR, em relação aos tamanhos de alelos detectados na amplificação isolada dos locos

Loco EMBRA	Número de alelos	Diferença (pb)		
		Máxima	Média	Desvio-padrão
27	6	0,48	0,09	0,11
22	7	1,38	0,68	0,47
5	5	0,66	0,33	0,10
6	8	1,28	0,69	0,48
4	6	0,97	0,52	0,26
24	7	0,50	0,13	0,14
10	7	3,13	0,34	0,61
15	4	0,83	0,22	0,22
18	6	0,80	0,65	0,08
Média/Total	56	1,11	0,41	0,27

Embora as variações detectadas na estimativa de tamanho dos alelos que compõem os sistemas sejam expressivas, deve-se observar que elas foram comparadas com os tamanhos de alelos estimados após a amplificação dos locos isolados e com a co-amplificação de vários locos em condições de PCR variadas.

Apesar de a condição de co-amplificação simultânea de vários locos não ser reconhecida como causadora de variações na estimativa do tamanho dos alelos, alterações na reação da PCR têm grande importância nesse fato.

OLSEN et al. (1996) detectaram variações superiores a 1 pb com a alteração da concentração do par de iniciadores responsáveis pela amplificação do loco em questão. A mesma variação foi encontrada com a alteração da temperatura de anelamento de um par de iniciadores, que levou à eliminação de bandas inespecíficas (*stutter*) e a um acréscimo superior a 1 pb na estimativa de tamanho de um dos alelos. KIJAS et al. (1995) também descreveram diferenças de até 1,37 pb na avaliação do tamanho de um alelo, detectado em diferentes condições de separação de fragmentos. Outro fator que pode explicar as variações observadas no tamanho estimado dos alelos é

a tendência da enzima Taq polimerase em adicionar um dATP ao terminal 3' dos fragmentos amplificados, durante a polimerização (CLARK, 1988).

Quando foram analisadas, entretanto, as variações ocorridas na avaliação de locos amplificados em condições de amplificação, separação e análise de fragmentos constantes, observaram-se, em geral, variações inferiores a 0,5 pb (GLOWATZK-MULLIS et al., 1995; OLSEN et al., 1996; MICHELL et al., 1997). Isso demonstra a importância da manutenção das condições de análise de locos microssatélites por ocasião da sua transferência para outros laboratórios.

4. RESUMO E CONCLUSÕES

Sistemas de genótipagem multilocus baseados em marcadores microssatélites, associados a sistemas de detecção de fluorescência, foram desenvolvidos para uso em várias espécies economicamente importantes do gênero *Eucalyptus*.

Para o desenvolvimento desses sistemas, procedeu-se, inicialmente, à seleção de um conjunto de locos microssatélites, com base na qualidade de amplificação desses locos pelos iniciadores desenhados. Dos 170 locos testados, mais de 70% amplificaram, especificamente, regiões microssatélites.

Também, avaliou-se a transferibilidade desses locos para outras espécies do gênero *Eucalyptus*. Observou-se maior grau de conservação dos microssatélites em *E. dunnii*, espécie do subgênero *Symphomyrtus*, ao qual pertencem também as espécies a partir das quais os microssatélites foram desenvolvidos, ou seja, *E. grandis* e *E. teretifolia*. A conservação foi menor entre espécies de outros subgêneros.

Um conjunto de locos foi selecionado para desenvolvimento de sistemas que envolvessem a co-amplificação de vários microssatélites em mesma reação de PCR. Iniciadores marcados com diferentes fluorocromos foram sintetizados para a amplificação de 17 microssatélites. Após uma seleção inicial, 11 locos foram excluídos e várias combinações de locos co-amplificados em mesma reação de PCR foram testadas. Obtiveram-se, no final, três sistemas Multiplex, compostos por três locos cada. Tais sistemas tiveram

4. RESUMO E CONCLUSÕES

Sistemas de genotipagem multilocos baseados em marcadores microssatélites, associados a sistemas de detecção de fluorescência, foram desenvolvidos para uso em várias espécies economicamente importantes do gênero *Eucalyptus*.

Para o desenvolvimento desses sistemas, procedeu-se, inicialmente, à seleção de um conjunto de locos microssatélites, com base na qualidade de amplificação desses locos pelos iniciadores desenhados. Dos 170 locos testados, mais de 70% amplificaram, especificamente, regiões microssatélites.

Também, avaliou-se a transferibilidade desses locos para outras espécies do gênero *Eucalyptus*. Observou-se maior grau de conservação dos microssatélites em *E. dunnii*, espécie do subgênero *Symphyomyrtus*, ao qual pertencem também as espécies a partir das quais os microssatélites foram desenvolvidos, ou seja, *E. grandis* e *E. urophylla*. A conservação foi menor entre espécies de outros subgêneros.

Um conjunto de locos foi selecionado para desenvolvimento de sistemas que envolvessem a co-amplificação de vários microssatélites em mesma reação de PCR. Iniciadores marcados com diferentes fluorocromos foram sintetizados para a amplificação de 17 microssatélites. Após uma seleção inicial, 11 locos foram escolhidos e várias combinações de locos co-amplificados em mesma reação de PCR foram testadas. Obtiveram-se, no final, três sistemas Multiplex, compostos por três locos cada. Tais sistemas tiveram

suas reações de PCR otimizadas de forma a obter uma amplificação adequada em todos os locos, nas espécies para as quais eles foram desenvolvidos.

Os sistemas de genotipagem descritos neste trabalho assemelham-se aos desenvolvidos para diversas espécies de animais, como bovinos (GLOWATZKI-MULLIS et al., 1995), urso polar (PAETKAU et al., 1995) e truta (WENBURG et al., 1996), e culturas agrícolas como brássicas (MICHELL et al., 1997), soja (DIWAN e CREGAN, 1997) e videira (THOMAS et al., 1994).

CAPÍTULO 2

DETERMINAÇÃO DO TAMANHO MÍNIMO DE AMOSTRA PARA ESTIMAÇÃO DA HETEROZIGIDADE ESPERADA EM LOCOS MICROSSATÉLITES

1. INTRODUÇÃO

Um dos parâmetros mais amplamente utilizados para descrever o conteúdo de informação genética de um loco-marcador, ou de uma classe de marcadores moleculares, é a heterozigidade esperada, também conhecida como diversidade gênica (WEIR, 1996).

A heterozigidade esperada expressa a probabilidade de dois-garotas, tomados ao acaso em uma população, serem de diferentes alelos. Em outras palavras, quanto maior for o valor de heterozigidade esperada, maior será a probabilidade de ser detectado polimorfismo entre dois indivíduos (RAPALSKI et al., 1996).

A comparação entre os valores de heterozigidade esperada encontrados em diferentes classes de marcadores tem demonstrado que locos microsatelitais com diferentes tamanhos de repetição apresentam maior heterozigidade esperada em relação à heterozigidade esperada em outros marcadores co-dominantes, como

CAPÍTULO 2

DETERMINAÇÃO DO TAMANHO MÍNIMO DE AMOSTRA PARA ESTIMAÇÃO DA HETEROZIGOSIDADE ESPERADA EM LOCOS MICROSSATÉLITES

1. INTRODUÇÃO

Um dos parâmetros mais amplamente utilizados para descrever o conteúdo de informação genética de um loco marcador, ou de uma classe de marcadores moleculares, é a heterozigiosidade esperada, também conhecida como diversidade gênica (WEIR, 1996).

A heterozigiosidade esperada expressa a probabilidade de dois gametas, tomados ao acaso em uma população, terem alelos diferentes para determinado loco. Em outras palavras, quanto maior for o valor de heterozigiosidade, maior será a probabilidade de ser detectado polimorfismo entre dois indivíduos (RAFALSKI et al., 1996).

A comparação entre os valores de heterozigiosidade esperada encontrados em diferentes classes de marcadores tem demonstrado que locos baseados em seqüências microssatélites apresentam maior heterozigiosidade em relação à encontrada em outros marcadores co-dominantes, como

isoenzimas e RFLP. Além disso, esses locos apresentam heterozigosidade esperada normalmente até duas vezes superior à encontrada em marcadores com padrão de herança dominante (RAPD e AFLP) (POWELL et al., 1996; VOGEL et al., 1996). Entretanto, estimativas desse parâmetro em locos microssatélites desenvolvidos para espécies florestais tropicais têm sido feitas com base na avaliação de um número extremamente variável de indivíduos, desde 20 amostras para *Pithecellobium elegans* (CHASE et al., 1996) até 80 para *Swietenia humilis* (WHITE e POWELL, 1997). Existem sugestões vagas na literatura sobre o número mínimo de amostras que seria adequado para se obterem estimativas acuradas de freqüências alélicas e, conseqüentemente, do parâmetro de heterozigosidade esperada. WEIR (1996) mencionou que a amostragem deve-se basear nos genótipos e em freqüências alélicas esperadas, ou seja, nas próprias características do loco analisado. Nenhum estudo até o momento tem buscado determinar qual a amostra mínima necessária para estimar esse parâmetro, com acurácia, para determinado loco microssatélite em uma população ou espécie.

Nesta etapa do trabalho, procurou-se determinar, para um grupo de indivíduos que compunham uma população de melhoramento de *E. grandis*, qual seria o tamanho mínimo de amostra necessário para obter estimativas precisas do parâmetro heterozigosidade esperada e, com isso, fazer uma análise comparativa adequada do poder de resolução de locos candidatos a compor os sistemas Multiplex.

Seis microssatélites de *Eucalyptus* (EMBRA) foram escolhidos para o presente estudo: EMBRA4, EMBRA5, EMBRA10, EMBRA15, EMBRA16 e EMBRA11 (BRONDANI et al., 1998). Esses locos foram amplificados via PCR em reações de 15 μ L contendo 7,5 ng de DNA genômico, 200 μ M de dNTP e 1,3 μ L de DMSO 99% e 1,2 μ L de tampão de PCR 10X contendo 16 mM de MgCl₂, 1,34 μ L de água desionizada, 1 mM de MgCl₂, 0,29 μ M de iniciadores anterior e posterior e 1 U de Taq polimerase. A reação de amplificação via PCR passou por três etapas: 94°C de desnaturação e 30" por cinco minutos, seguida de 23 ciclos contínuos dos passos de desnaturação (94°C),

2. MATERIAL E MÉTODOS

2.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificados

Foram genotipados 192 indivíduos de *E. grandis*, provenientes de 18 procedências (Quadro 1) e selecionados a partir de ensaios de procedências/progênieis estabelecidos na empresa Aracruz Celulose S.A. (Espírito Santo, Brasil). Os ensaios foram estabelecidos com sementes coletadas de indivíduos em populações naturais da Austrália e de um indivíduo proveniente de plantios comerciais no Zimbábue. O DNA genômico dos indivíduos foi extraído de tecido foliar de plantas adultas, conforme a metodologia descrita por DOYLE e DOYLE (1987).

Seis microssatélites de *Eucalyptus* (EMBRA) foram escolhidos para o presente estudo: EMBRA4, EMBRA5, EMBRA10, EMBRA15, EMBRA16 e EMBRA11 (BRONDANI et al., 1998). Esses locos foram amplificados via PCR em reações de 13 μ L contendo 7,5 ng de DNA genômico, 200 μ M de dNTP e 1,3 μ L de DMSO 50% e 1,3 μ L de tampão de PCR 10X contendo 15 mM de $MgCl_2$, 1,34 μ L de água desionizada, 1 mM de $MgCl_2$, 0,29 μ M de iniciadores anterior e posterior e 1 U de Taq polimerase. A reação de amplificação via PCR passou por uma etapa inicial de desnaturação a 96°C por cinco minutos, seguida de 29 ciclos constituídos dos passos de desnaturação (94°C),

Quadro 1 – Procedências de *E. grandis* avaliadas no estudo de determinação do número mínimo de indivíduos para estimativa do parâmetro heterozigidade esperada nos locos microssatélites e do número de indivíduos disponíveis por procedência

Procedência	Número de indivíduos	Procedência	Número de indivíduos
Gadgarra	41	Davies	4
Windsor Tableland	32	Mount Lewis	2
Mount Spec	26	Mount Windsor	1
Cairns Dam	20	Boar	1
Herberton	17	Mount Pandanus	1
Yungaburra / Gordonvale	15	Mount Carbine	1
Dinbula	14	Douglas Creek	1
Lamb. Range – Cairns	9	Mount Fox	1
Atherton	5	Zimbabwe	1

anelamento (56°C) e extensão (72°C), todos por um minuto; e de uma etapa final de extensão a 72°C, por sete minutos.

Cinco μ L de tampão de carregamento desnaturante foram adicionados aos produtos amplificados. Estes foram separados em géis de poliacrilamida (4%) contendo 7 M de uréia, a 80 W, por 90 minutos, após desnaturação a 95°C por cinco minutos. Após a separação dos fragmentos por eletroforese, estes foram corados com nitrato de prata, de acordo com a metodologia descrita por BASSAN et al. (1991). Três produtos de PCR, resultantes da amplificação da mesma amostra de DNA, mas com iniciadores associados a três locos microssatélites diferentes, foram separados na mesma linha do gel, simultaneamente (Figura 1). O tamanho dos fragmentos amplificados foi estimado com base na comparação com o marcador-padrão 10 pb (Gibco BRL).

Os genótipos dos indivíduos amostrados, assim como o número de alelos detectados e as heterozigidades esperada e observada em cada loco, estão descritos no Capítulo 4.

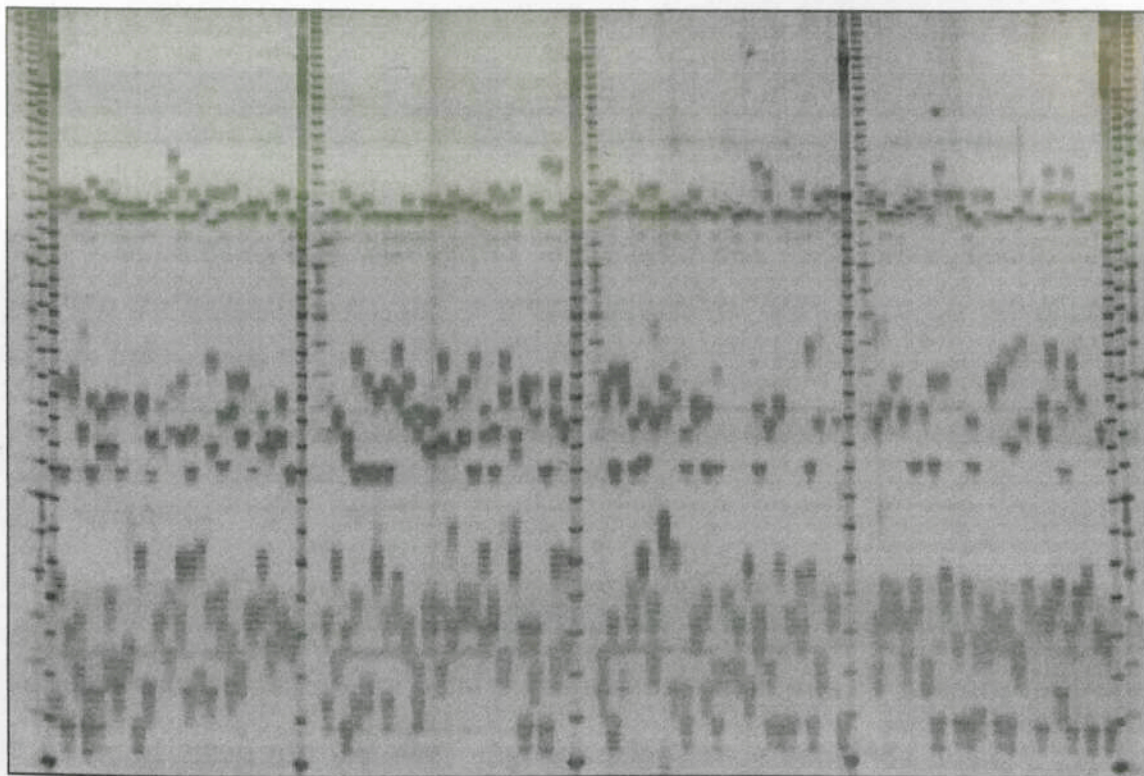


Figura 1 – Análise de 98 indivíduos de *E. grandis*, com os locos microssatélites EMBRA5, EMBRA10 e EMBRA11 analisados simultaneamente na mesma linha do gel.

2.2. Influência do tamanho da amostra sobre as estimativas de heterozigidade esperada

A heterozigidade esperada de um loco, He_l , é definida como

$$He_l = 1 - \sum_u p_{lu}^2$$

em que " p_{lu} " é a frequência do alelo "u" no loco "l".

2.2.1. Método de reamostragem

Para avaliação do comportamento desse estimador em função da amostra de indivíduos genotipados, assumiu-se o modelo de Wright para as frequências genotípicas em cada loco: g_{lj} , $j = 1$ a k , sendo k o número de classes de genótipos possíveis em cada loco. Os valores de frequências

alélicas e o valor do índice de fixação “ f ”, necessários para serem obtidas as frequências genóticas, foram tomados da amostra original. Tais valores foram adotados como valores paramétricos para avaliação do comportamento do tamanho da amostra. Foram geradas 1.000 amostras de tamanho $n = 16, 32, 64$ e 128 , com base na distribuição de probabilidade multinomial, sendo “ n ” o número de indivíduos e “ g_{ij} ” a probabilidade de ocorrência do genótipo da classe “ j ” no loco “ i ”.

Adotou-se o seguinte estimador do tipo “plug-in” (EFRON e TIBSHIRANI, 1993) para esses parâmetros:

$$\hat{H}_{e_i} = 1 - \sum_u \hat{p}_{iu}^2$$

Para cada uma das 24 distribuições geradas (seis locos e quatro tamanhos de amostras), foi obtida uma estimativa do desvio-padrão, $\hat{S}_{e_i}(\hat{H}_{e_i})$, e da tendenciosidade, $\hat{T}d_i(\hat{H}_{e_i})$, da heterozigiosidade esperada, ou seja:

$$\hat{S}_{e_i}(\hat{H}_{e_i}) = \sqrt{\frac{\sum (\hat{H}_{e_i} - H_{e_i})^2}{999}}$$

$$\hat{T}d_i(\hat{H}_{e_i}) = \tilde{H}_{e_i} - H_{e_i}$$

sendo \tilde{H}_{e_i} a média das 1.000 estimativas de H_{e_i} .

O coeficiente de variação do erro quadrático médio, CV_{eqm} (1), uma medida de acurácia, foi tomado para avaliação do tamanho de amostra. CV_{eqm} considera a variação total em torno do verdadeiro valor do parâmetro, ou seja, $\hat{T}d_i(\hat{H}_{e_i})$ e $\hat{S}_{e_i}(\hat{H}_{e_i})$, como se segue:

$$(1) \quad CV_{eqm}(l) = \frac{\sqrt{[\hat{T}d_i(\hat{H}_{e_i})]^2 + \hat{S}_{e_i}(\hat{H}_{e_i})^2}}{H_{e_i}}$$

2.2.2. Fórmula da variância da heterozigiosidade esperada

A variância da heterozigiosidade esperada em função do tamanho da amostra foi estimada também por meio da fórmula exata da variância da diversidade gênica, ou heterozigiosidade esperada (WEIR, 1996).

Nesta análise, a variância da heterozigosidade esperada, $Var(\hat{H}e_i)'$, foi calculada através da seguinte equação:

$$Var(\hat{H}e_i)' = \frac{2(1+f)}{n} \left[\sum_u p^3_{lu} - \left(\sum_u p^2_{lu} \right)^2 \right]$$

sendo "f" o índice de fixação de Wright e "n" o tamanho da amostra.

Para cada loco foram estimados o desvio-padrão, $\hat{S}e_i(\hat{H}e_i)'$ (1); a esperança da heterozigosidade esperada, $\varepsilon(\hat{H}e_i)'$ (2); e a tendenciosidade, $\hat{T}d_i(\hat{H}e_i)'$ (3), ou seja:

$$(1) \quad \hat{S}e_i(\hat{H}e_i)' = \sqrt{Var(\hat{H}e_i)}$$

$$(2) \quad \varepsilon(\hat{H}e_i)' = \left(1 - \frac{1+f}{2n} \right) H e_i$$

$$(3) \quad \hat{T}d_i(\hat{H}e_i)' = \varepsilon(\hat{H}e_i) - (H e_i)$$

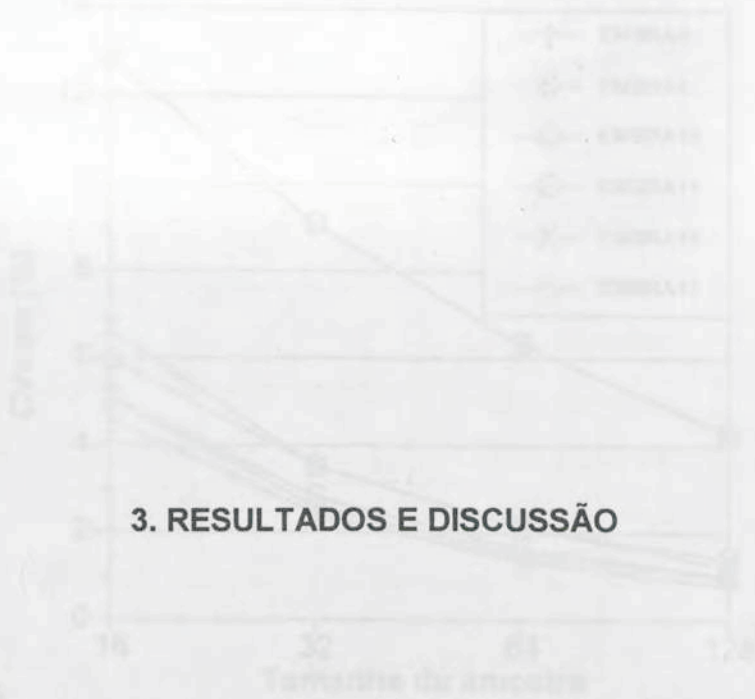
Uma medida de acurácia idêntica à adotada na estimação da heterozigosidade esperada em função do tamanho da amostra pelo método de reamostragem foi empregada. O CV_{eqm} foi obtido segundo (1), usando-se (2) e (3).

Como os desvios-padrão segundo WEIR (1986), encontram-se na Figura 4.

Os resultados observados concordaram com os de WEIR (1986), que afirmou que a variância da heterozigosidade esperada está relacionada com as frequências alélicas observadas em determinado loco. Locos microsatélites com grande número de alelos em baixa frequência apresentaram menor variância do que aqueles onde havia menor número de alelos, como no caso do loco EMBRA11.

Os resultados gerados por reamostragem e através da fórmula exata da variância da heterozigosidade esperada (WEIR, 1986) indicaram que, para locais com maior número de alelos e na presença de um índice de fixação significativo, o CV_{eqm} sobre as estimativas de heterozigosidade esperada tende a ser inferior a 2% a partir de amostras superiores ou iguais a 64 indivíduos.

Em locais com número reduzido de alelos, como é o caso do loco EMBRA11, para atingir valores de coeficientes de variação igualmente reduzidos, os tamanhos de amostra tendem a ser significativamente maiores, mesmo quando se trata de amostras de indivíduos, bem como de número avaliado neste trabalho.



Ambas as metodologias utilizadas para a estimativa da variância da heterozigosidade esperada com diferentes tamanhos de amostra produziram resultados congruentes, seja em termos de coeficiente de variação, seja em termos de tendenciosidade (Figuras 2 e 3). Os resultados referentes à esperança da heterozigosidade esperada encontrada em cada loco, assim como os desvios-padrão segundo WEIR (1996), encontram-se na Figura 4.

Os resultados observados concordaram com os de WEIR (1996), que afirmou que a variância da heterozigosidade esperada está relacionada com as frequências alélicas observadas em determinado loco. Locos microssatélites com grande número de alelos em baixa frequência apresentaram menor variância do que aqueles onde havia menor número de alelos, como no caso do loco EMBRA11.

Os resultados gerados por reamostragem e através da fórmula exata da variância da heterozigosidade esperada (WEIR, 1996) indicaram que, para locos com maior número de alelos e na presença de um índice de fixação significativo, o CV_{eqm} sobre as estimativas de heterozigosidade esperada tende a ser inferior a 2% a partir de amostras superiores ou iguais a 64 indivíduos.

Em locos com número reduzido de alelos, como é o caso do loco EMBRA11, para atingir valores de coeficientes de variação igualmente reduzidos, os tamanhos de amostra teriam que ser significativamente maiores, da ordem de centenas de indivíduos, bem acima do número avaliado neste trabalho.

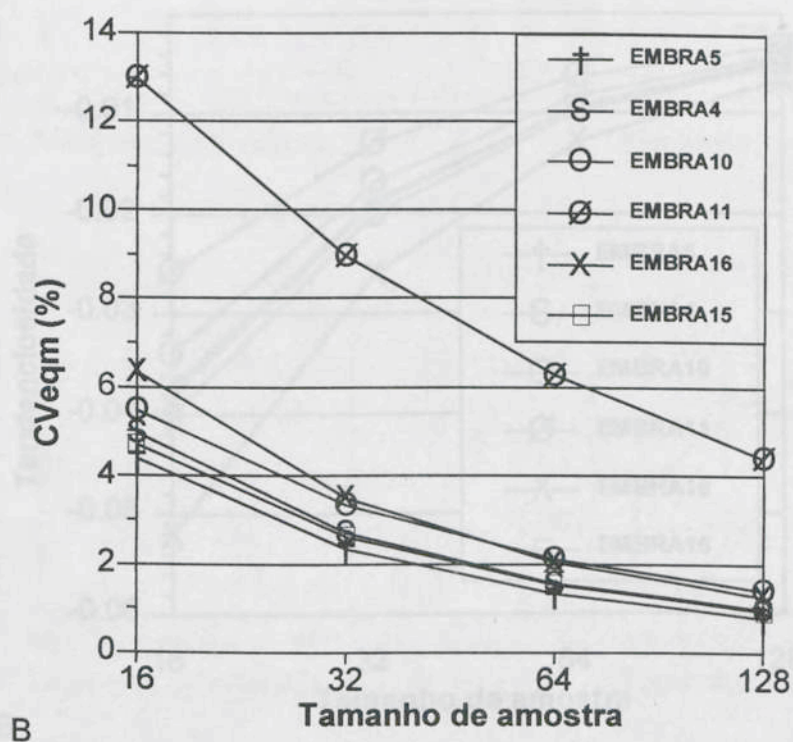
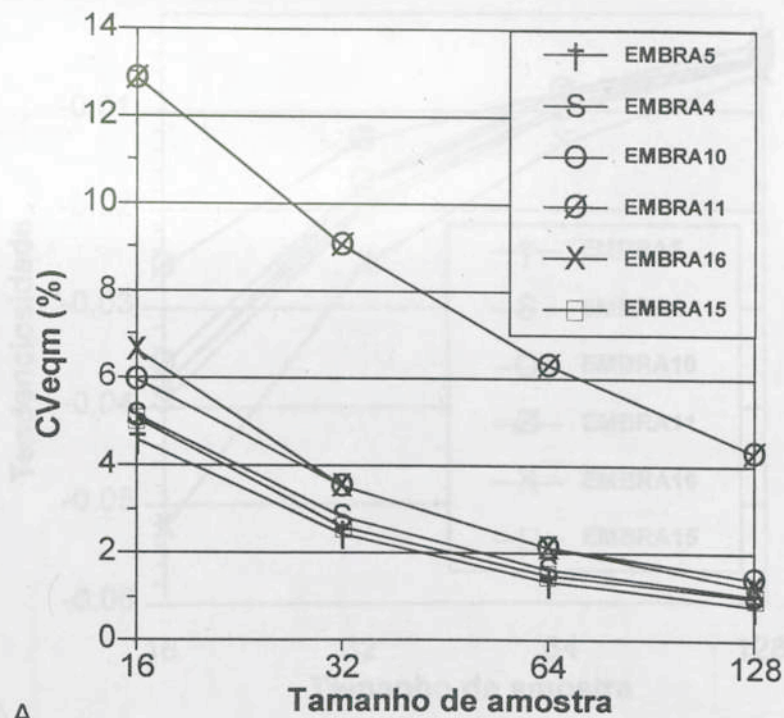
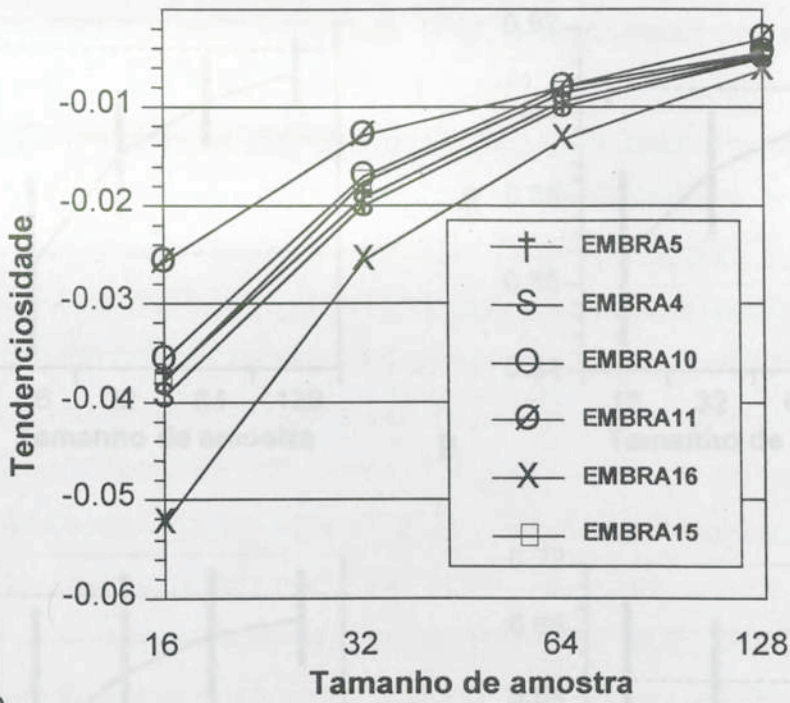
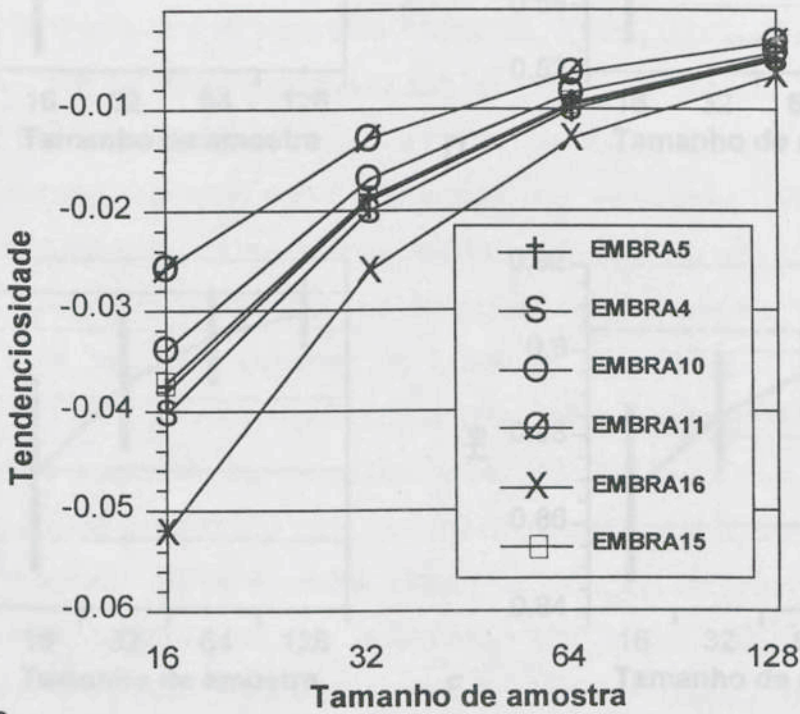


Figura 2 – Redução no coeficiente de variação do erro quadrático médio (CV_{eqm}) da heterozigidade esperada com o incremento do tamanho da amostra, para seis locos EMBRA, e estimativas do CV_{eqm} obtidas por meio de (A) reamostragem numérica e da (B) fórmula de variância da heterozigidade esperada (WEIR, 1996).



A



B

Figura 3 – Tendenciosidade estimada da heterozigosidade esperada com o incremento do tamanho da amostra, para seis locos EMBRA, e estimativas da tendenciosidade obtidas por meio de (A) reamostragem numérica e da (B) fórmula de variância da heterozigosidade esperada (WEIR, 1996).

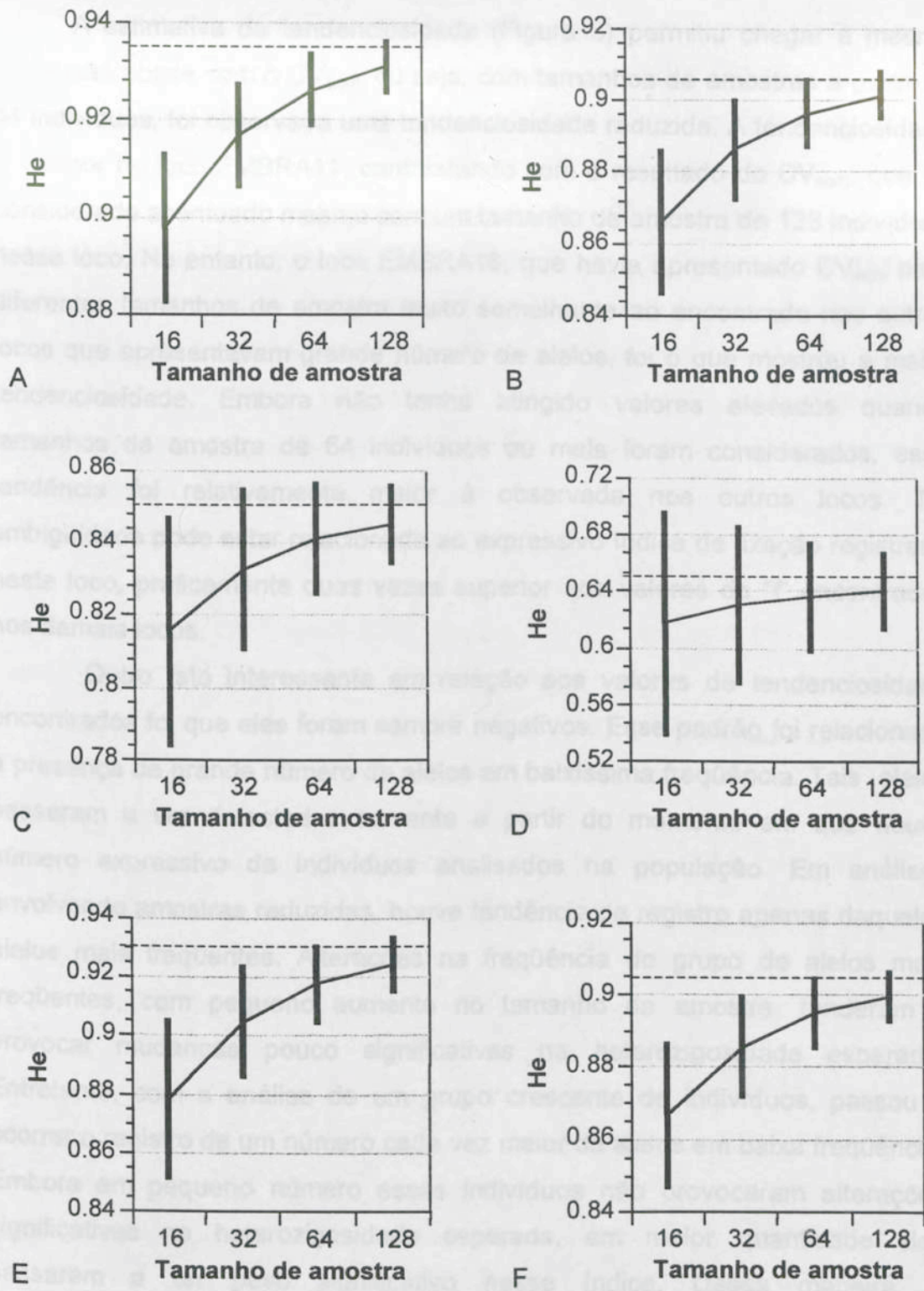


Figura 4 – Esperanças da heterozigosidade esperada (WEIR, 1996) e respectivos desvios-padrão, calculados para os locos (A) EMBRA5, (B) EMBRA4, (C) EMBRA10, (D) EMBRA11, (E) EMBRA16 e (F) EMBRA12 (linha pontilhada corresponde ao valor paramétrico de H_e).

A estimativa de tendenciosidade (Figura 3) permitiu chegar à mesma conclusão obtida com o CV_{eqm} , ou seja, com tamanhos de amostras a partir de 64 indivíduos, foi observada uma tendenciosidade reduzida. A tendenciosidade foi menor no loco EMBRA11, contrastando com o resultado do CV_{eqm} , que foi considerado acentuado mesmo com um tamanho de amostra de 128 indivíduos nesse loco. No entanto, o loco EMBRA16, que havia apresentado CV_{eqm} para diferentes tamanhos de amostra muito semelhante ao encontrado nos outros locos que apresentavam grande número de alelos, foi o que mostrou a maior tendenciosidade. Embora não tenha atingido valores elevados quando tamanhos de amostra de 64 indivíduos ou mais foram considerados, essa tendência foi relativamente maior à observada nos outros locos. Tal ambigüidade pode estar relacionada ao expressivo índice de fixação registrado neste loco, praticamente duas vezes superior aos valores de "f" encontrados nos demais locos.

Outro fato interessante em relação aos valores de tendenciosidade encontrados foi que eles foram sempre negativos. Esse padrão foi relacionado à presença de grande número de alelos em baixíssima freqüência. Tais alelos passaram a ser detectados somente a partir do momento em que houve número expressivo de indivíduos analisados na população. Em análises envolvendo amostras reduzidas, houve tendência de registro apenas daqueles alelos mais freqüentes. Alterações na freqüência do grupo de alelos mais freqüentes, com pequeno aumento no tamanho da amostra, tenderam a provocar mudanças pouco significativas na heterozigosidade esperada. Entretanto, com a análise de um grupo crescente de indivíduos, passou a ocorrer o registro de um número cada vez maior de alelos em baixa freqüência. Embora em pequeno número esses indivíduos não provocaram alterações significativas na heterozigosidade esperada, em maior quantidade eles passaram a ter peso significativo nesse índice. Dessa maneira, a heterozigosidade esperada apresentou a tendência de sempre mostrar valores inferiores aos do parâmetro da população. Assim, pôde-se concluir que, mesmo quando altos valores de heterozigosidade esperada forem registrados em amostras pequenas de indivíduos de *Eucalyptus*, o valor paramétrico será sempre superior a este.

Os resultados encontrados indicaram claramente que vários estudos de estimativas de heterozigosidade esperada em locos microssatélites de espécies florestais tropicais podem estar incorrendo em erros de grande magnitude. Frequentemente, estudos de caracterização de microssatélites têm-se baseado na análise de grupos menores do que 20 indivíduos (BYRNE et al., 1996; CHASE et al., 1996). Observou-se ainda que, embora estimativas desse parâmetro tenham sido frequentemente geradas para indicar o conteúdo de informação de locos microssatélites desenvolvidos para espécies florestais alógamas (CHASE et al., 1996; WHITE e POWELL, 1997), nenhuma atenção havia sido dada à questão do tamanho de amostra necessário para estimar esse parâmetro com acurácia.

Foi investigada a influência do tamanho da amostra na estimativa do parâmetro heterozigosidade esperada em uma população de melhoramento de E. grandis. As duas formas de análise, por meio de reamostragem e da fórmula exata de variância da diversidade genética de WEIR (1986), apresentaram resultados congruentes.

Em locos microssatélites com grande número de alelos detectados, tamanho da amostra superior a 64 indivíduos acarretou alterações pouco relevantes no nível de acurácia de estimativa da heterozigosidade esperada. Entretanto, em locos com menor número de alelos, a tendência indicou que o mesmo nível de acurácia seria atingido somente após a análise de um número superior de indivíduos.

O nível de acurácia também apresentou estabilização após a inclusão de 64 indivíduos na estimativa da heterozigosidade. A maior tendência foi observada no loco com maior índice de fixação de Wright. Finalmente, detectou-se ainda que a tendência foi sempre negativa, indicando que estimativas do parâmetro de heterozigosidade esperada tendem a ser sempre subestimadas em relação ao valor paramétrico da população. Esta foi a primeira indicação do número mínimo de indivíduos necessários para obter estimativas adequadas de heterozigosidade esperada em locos microssatélites. A análise de resultados para outros locos e espécies também é apresentada.

4. RESUMO E CONCLUSÕES

Foi investigada a influência do tamanho de amostra na estimativa do parâmetro heterozigidade esperada em uma população de melhoramento de *E. grandis*. As duas formas de análise, por meio de reamostragem e da fórmula exata de variância da diversidade gênica de WEIR (1996), apresentaram resultados congruentes.

Em locos microsatélites com grande número de alelos detectados, tamanho de amostra superior a 64 indivíduos acarretou alterações pouco relevantes no nível de acurácia da estimativa da heterozigidade esperada. Entretanto, em locos com menor número de alelos, a tendência indicou que o mesmo nível de acurácia seria atingido somente após a análise de um número superior de indivíduos.

O nível de tendenciosidade também apresentou estabilização após a inclusão de 64 indivíduos na estimativa da heterozigidade. A maior tendenciosidade foi observada no loco com maior índice de fixação de Wright. Finalmente, detectou-se ainda que a tendenciosidade foi sempre negativa, indicando que estimativas do parâmetro de heterozigidade esperado tendem a ser sempre subestimados em relação ao valor paramétrico da população. Essa foi a primeira indicação do número mínimo de indivíduos necessários para obter estimativas acuradas da heterozigidade esperada para locos microsatélites baseados em dinucleotídeos. *A priori* não existem dados que validam tais resultados para outras espécies. Considerando a amostra

estudada com seis locos de peculiaridades diferentes e o fato de que a natureza molecular dos microssatélites faz com que estes tenham comportamento extremamente semelhante em termos de distribuição de frequências alélicas entre organismos, acredita-se que os resultados do presente estudo tenham ampla relevância para locos microssatélites baseados em dinucleotídeos em eucariotos. Além disso, o fato claro de que a tendenciosidade é exclusivamente negativa indica que, apesar de as estimativas com pequeno número de indivíduos não estarem dentro de um nível de acurácia adequado, a expectativa é de que o valor paramétrico da população assuma sempre valores superiores aos estimados.

BANCOS DE DADOS DE FREQUÊNCIAS ALÉLICAS E ESTIMATIVA DE PARÂMETROS DE INFORMAÇÃO GENÉTICA DOS SISTEMAS DE GENOTIPAGEM MULTILOCCOS, EM SEIS ESPÉCIES DO GÊNERO

Eucalyptus

1. INTRODUÇÃO

O gênero *Eucalyptus* (Myrtaceae) inclui mais de 500 espécies originárias principalmente da Austrália, além da Indonésia e Papua Nova Guiné (ELDRIDGE et al., 1994). Este gênero foi dividido em oito subgêneros, dentre os quais o mais importante é o *Symphyomyrtus*. Destaca-se pela expressiva importância comercial de algumas das 300 espécies que ele compreende. Dentre estas estão *E. camaldulensis*, *E. dunnii*, *E. globulus*, *E. grandis*, *E. saligna* e *E. urophylla*, que são algumas das espécies de *Eucalyptus* mais plantadas mundialmente (ELDRIDGE et al., 1994).

A estratégia dos programas mais avançados de melhoramento de *Eucalyptus* tem buscado associar a seleção recorrente recíproca com o plantio clonal de indivíduos selecionados. Entretanto, essa estratégia depara-se frequentemente com obstáculos relacionados à falta de conhecimento do

CAPÍTULO 3

BANCOS DE DADOS DE FREQUÊNCIAS ALÉLICAS E ESTIMATIVA DE PARÂMETROS DE INFORMAÇÃO GENÉTICA DOS SISTEMAS DE GENOTIPAGEM MULTILOCOS, EM SEIS ESPÉCIES DO GÊNERO

Eucalyptus

1. INTRODUÇÃO

O gênero *Eucalyptus* (Myrtaceae) inclui mais de 500 espécies originárias principalmente da Austrália, além da Indonésia e Papua Nova Guiné (ELDRIDGE et al., 1994). Este gênero foi dividido em oito subgêneros, dentre os quais o mais importante é o *Symphyomyrtus*. Destaca-se pela expressiva importância comercial de algumas das 300 espécies que ele compreende. Dentre estas estão *E. camaldulensis*, *E. dunnii*, *E. globulus*, *E. grandis*, *E. saligna* e *E. urophylla*, que são algumas das espécies de *Eucalyptus* mais plantadas mundialmente (ELDRIDGE et al., 1994).

A estratégia dos programas mais avançados de melhoramento de *Eucalyptus* tem buscado associar a seleção recorrente recíproca com o plantio clonal de indivíduos selecionados. Entretanto, essa estratégia depara, freqüentemente, com obstáculos relacionados à falta de conhecimento do

"pedigree" dos indivíduos que compõem as populações de melhoramento. Erros de identificação de material genético, ocorridos no estabelecimento dessas populações, são igualmente comuns. Tais dificuldades geram a necessidade de uma ferramenta que permita a determinação precisa das relações de parentesco entre os clones-elite, assim como o estabelecimento de um procedimento de identificação individual.

Metodologias voltadas para atender a necessidades semelhantes foram criadas recentemente para videira (THOMAS et al., 1994), brássicas (SZEWC-MCFADDEN et al., 1996) e soja (DIWAN e CREGAN, 1997), mediante o desenvolvimento de sistemas de genotipagem multilocos, baseados em marcadores microssatélites, com visualização e análise em sistemas de detecção de fluorescência. Da mesma maneira como foi feito nos primeiros estudos em humanos, em que os locos amostrados nesse sistemas foram caracterizados quanto às freqüências alélicas observadas em diferentes populações, os trabalhos de THOMAS et al. (1994), SZEWC-MCFADDEN et al. (1997) e DIWAN e CREGAN (1997) também procuraram estabelecer a utilidade dos locos que compunham os sistemas de genotipagem para diferentes espécies e variedades das culturas de interesse.

Embora esses sistemas apresentem freqüentemente alta transferibilidade para outras espécies relacionadas àquela para qual eles foram desenvolvidos, observou-se que alguns locos com alto conteúdo de informação em uma espécie apresentavam poucos alelos ou até mesmo alelos fixados e, conseqüentemente, poder informativo nulo em outras (WENBURG et al., 1996).

O conteúdo de informação ou diversidade gênica dos microssatélites está relacionado com a probabilidade de se encontrar polimorfismo entre indivíduos (RAFALSKI et al., 1996). Para estimar essa diversidade, vários índices que expressam a utilidade dos marcadores moleculares têm sido propostos. Uma maneira de avaliar o conteúdo de informação de um loco microssatélite é por meio da sua heterozigosidade esperada, ou diversidade gênica (WEIR, 1996), que indica a probabilidade de se encontrarem dois alelos diferentes em um loco analisado, em dois gametas escolhidos ao acaso na população. Outros índices, como o conteúdo de informação de polimorfismo (BOTSTEIN et al., 1980), a probabilidade de identidade (PAETKAU et al., 1995) e o poder de exclusão (WEIR, 1996), têm sido propostos, os quais expressam

essencialmente a capacidade de cada loco de discriminar indivíduos, determinar relações de parentesco ou fornecer informação genética de segregação em experimentos de detecção de ligação gênica.

Todos esses parâmetros de conteúdo informativo se baseiam, fundamentalmente, no número de alelos no loco e em suas freqüências relativas. De maneira geral, quanto maior o número de alelos, mais informativo é o loco. Entretanto, a freqüência desses alelos é o fator que, em última análise, determina o poder de informação genética do loco. A caracterização de uma bateria de locos quanto à distribuição de freqüência de seus alelos torna-se, portanto, um passo essencial no desenvolvimento e na recomendação de um sistema de genotipagem.

Neste trabalho foi gerado um banco de dados de freqüências alélicas dos locos incluídos nos sistemas de genotipagem desenvolvidos. Esse banco de dados foi construído para as seis espécies do gênero *Eucalyptus* de maior importância econômica mundial. Com base nessas freqüências alélicas, a diversidade desses locos e o seu poder de resolução para fins de identificação individual e estudos de parentesco foram estabelecidos para cada espécie. Foram realizadas análises comparativas das distribuições de freqüências alélicas nas seis espécies, e, com base em uma análise ampliada para *E. grandis*, foi avaliada, ainda, a estrutura da diversidade genética entre procedências e dentro delas.

As originais de uma mistura de sementes provenientes de 10 famílias de melaleucae selecionadas, nas procedências citadas, pela empresa Ricoll S.A. Para análise de *E. camaldulensis*, *E. durnii*, *E. globulus*, *E. saligna* e *E. urophylla*, foram utilizadas mudas originadas de sementes coletadas em populações naturais na Austrália e Indonésia (Quadro 1). Para *E. durnii* e *E. saligna*, foram utilizadas plantas originadas de uma mistura de sementes cedidas pela empresa Ricoll S.A., provenientes de 10 famílias de melaleucae e coletadas em quatro procedências. As plantas de *E. camaldulensis* e *E. urophylla* constituem indivíduos sem origem conhecida, cujas sementes foram fornecidas pelas empresas Madeirense Florestal S.A. e Bahia Sul S.A., respectivamente. Para todas estas espécies, foi utilizada uma amostra de 68 indivíduos, sendo 17 de cada uma das quatro procedências, à exceção de *E. urophylla* para a qual foram utilizados 34 indivíduos de cada uma das duas procedências. Para *E. globulus*, somente

Quadro 1 - Proveniências e localização geográfica dos locais de coleta de sementes, com origem em 68 indivíduos de seis espécies de eucaliptos originados com sistemas de propagação multiplasmígonos (ND = não disponível).

Espécie	Procedência	Localização geográfica		País
		Latitude	Longitude	
<i>E. camaldulensis</i>	Innesbank to Pefferd	17° 24'	145° 08'	Austrália
<i>E. camaldulensis</i>	Katherine	14° 23'	132° 15'	Austrália
<i>E. camaldulensis</i>	Kennedy River	15° 05'	144° 10'	Austrália
<i>E. camaldulensis</i>	Walsh River	16° 59'	144° 18'	Austrália
<i>E. dunnii</i>	Molton	ND	ND	Austrália
<i>E. dunnii</i>	ND	ND	ND	Austrália
<i>E. dunnii</i>	ND	ND	ND	Austrália
<i>E. dunnii</i>	ND	ND	ND	Austrália
<i>E. globulus</i>	Central Planters Island	10° 00'	147° 55'	Austrália
<i>E. globulus</i>	Central King Island	37° 53'	144° 00'	Austrália
<i>E. globulus</i>	Forest	35° 22'	148° 34'	Austrália
<i>E. globulus</i>	Atherton	17° 12'	143° 25'	Austrália

2. MATERIAL E MÉTODOS

2.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificados

Para *E. grandis* foi realizado um estudo detalhado, ampliando-se a amostragem das procedências. Foram utilizados 48 indivíduos de cada uma das procedências Atherton, Davies Creek, Mount Spec, Tinaroo e Wild River (Quadro 1), totalizando-se 240 indivíduos analisados. Estes indivíduos foram escolhidos ao acaso de mudas originadas de uma mistura de sementes provenientes de 10 famílias de meios-irmãos coletadas, nas procedências citadas, pela empresa Riocell S.A. Para análise de *E. camaldulensis*, *E. dunnii*, *E. globulus*, *E. saligna* e *E. urophylla*, foram utilizadas mudas originadas de sementes coletadas em populações naturais na Austrália e Indonésia (Quadro 1). Para *E. dunnii* e *E. saligna*, foram utilizadas plantas originadas de uma mistura de sementes cedidas pela empresa Riocell S.A., provenientes de 10 famílias de meios-irmãos e coletadas em quatro procedências. As plantas de *E. camaldulensis* e *E. urophylla* constituem indivíduos sem origem conhecida, cujas sementes foram fornecidas pelas empresas Mannesmann Florestal S.A. e Bahia Sul S.A., respectivamente. Para todas essas espécies, foi utilizada uma amostra de 68 indivíduos, sendo 17 de cada uma das quatro procedências, à exceção de *E. urophylla*, para a qual foram utilizados 34 indivíduos de cada uma das duas procedências. Para *E. globulus*, somente

Quadro 1 - Procedências e localização geográfica dos locais de coleta de sementes que originaram os indivíduos de seis espécies de *Eucalyptus*, avaliados com sistemas de genotipagem multilocus (ND = não disponível)

Espécie	Procedência	Localização geográfica		País
		Latitude	Longitude	
<i>E. camaldulensis</i>	Irvinebank to Petford	17° 24'	145° 09'	Austrália
<i>E. camaldulensis</i>	Katherine	14° 29'	132° 15'	Austrália
<i>E. camaldulensis</i>	Kennedy River	15° 05'	144° 19'	Austrália
<i>E. camaldulensis</i>	Walsh River	16° 59'	144° 18'	Austrália
<i>E. dunnii</i>	Moleton	ND	ND	Austrália
<i>E. dunnii</i>	ND	ND	ND	Austrália
<i>E. dunnii</i>	ND	ND	ND	Austrália
<i>E. dunnii</i>	ND	ND	ND	Austrália
<i>E. globulus</i>	Central Flinders Island	40° 00'	147° 55'	Austrália
<i>E. globulus</i>	Central King Island	39° 55'	144° 00'	Austrália
<i>E. globulus</i>	Jeeralang North	38° 22'	146° 34'	Austrália
<i>E. globulus</i>	South Geeveston	43° 12'	146° 58'	Austrália
<i>E. globulus</i>	Otway State Forest	38° 22'	146° 34'	Austrália
<i>E. grandis</i>	Atherton	17° 19'	145° 25'	Austrália
<i>E. grandis</i>	Davies Creek	17° 03'	145° 37'	Austrália
<i>E. grandis</i>	Mount Spec	19° 01'	146° 08'	Austrália
<i>E. grandis</i>	Tinaroo Creek	17° 05'	145° 36'	Austrália
<i>E. grandis</i>	Wild River	17° 22'	145° 24'	Austrália
<i>E. saligna</i>	Chaelundi	30° 00'	152° 22'	Austrália
<i>E. saligna</i>	Clouds Creek	30° 04'	152° 37'	Austrália
<i>E. saligna</i>	Sheas Nob S. F.	30° 02'	152° 32'	Austrália
<i>E. saligna</i>	Wild C. Creek	30° 06'	152° 45'	Austrália
<i>E. urophylla</i>	Timor	ND	ND	Indonésia
<i>E. urophylla</i>	Flores	ND	ND	Indonésia

pequena amostra de semente foi obtida no Instituto CRC-THF, da Austrália. Em razão da limitação de sementes, foram utilizadas de cinco a oito plantas de cada procedência, totalizando-se somente 32 plantas naquela espécie. Para extração de DNA genômico, foi utilizada a metodologia descrita por DOYLE e DOYLE (1987).

Os procedimentos de amplificação dos sistemas de genotipagem multilocos, eletroforese e detecção dos fragmentos seguiram o protocolo estabelecido na etapa de desenvolvimento desses sistemas (Capítulo 1).

2.2. Análises estatísticas

Com base nos genótipos dos indivíduos, foram obtidas as freqüências alélicas para os locos microssatélites. A partir das estimativas de freqüências alélicas, foram estimados os parâmetros de informação genética heterozigidade esperada, H_e ; conteúdo de informação de polimorfismo, PIC; probabilidade de identidade, I ; e poder de exclusão, Q . Registraram-se, ainda, o número de alelos observados, a amplitude alélica e a heterozigidade observada, H_o , nas espécies. Tais estimativas foram obtidas para as seis espécies de *Eucalyptus*, sendo em *E. grandis* a análise também realizada para cada procedência individualmente. Os parâmetros de informação genética foram estimados de acordo com as seguintes fórmulas:

$$H_e = 1 - \sum_u p_u^2$$

$$PIC = 1 - \left(\sum_{u=1}^n p_u^2 \right) - \sum_u \sum_{v>u} 2p_u^2 p_v^2$$

$$I = 1 - \sum_u p_u^4 - \sum_u \sum_{v>u} (2p_u p_v)^2$$

$$Q = \sum_u p_u (1 - p_u)^2 - \sum_u \sum_{v>u} p_u^2 p_v^2 (4 - 3p_u - 3p_v)$$

em que " p_u " e " p_v " são as freqüências do u -ésimo e v -ésimo alelo em uma população, sendo u diferente de v .

Foi testada, ainda, a existência de diferença significativa nas distribuições de freqüências alélicas nos locos microssatélites, nas diferentes

procedências de *E. grandis* e nas seis espécies de *Eucalyptus*. Para cada loco, uma estimativa não-viesada do valor P do teste de probabilidade (teste exato de Fisher) foi gerada (RAYMOND e ROUSSET, 1995a). O teste foi feito entre todas as procedências de *E. grandis* e espécies, duas a duas, por meio do programa GenePop, versão 1.2 (RAYMOND e ROUSSET, 1995b).

Para a análise das procedências de *E. grandis*, uma matriz de distâncias genéticas baseada no índice de Nei (1972) foi gerada a partir das freqüências alélicas observadas nos nove locos. As análises foram feitas com o programa NT-SYS ("Applied Biostatistics Inc"., versão 2.0) e manualmente (DIAS, 1998). Foi realizada ainda análise de variância molecular, AMOVA (EXCOFFIER et al., 1992), entre as cinco procedências de *E. grandis*, para avaliar a distribuição da variabilidade genética entre procedências e dentro delas. As análises foram realizadas por meio do programa Arlequin, versão 1.1 (SCHNEIDER et al., 1997).

especies da Eucalyptus

Na caracterização dos locos microsatélites em *E. grandis* (Quadro 2), valores de H_e superiores a 0,9 foram encontrados em três locos: EMBRA22, EMBRA5 e EMBRA24, refletindo o grande número de alelos detectados nesses microsatélites. O valor de H_e foi inferior a 0,5 nos locos EMBRA8 e EMBRA18. No loco EMBRA18, a baixa H_e foi relacionada ao reduzido número de alelos detectados e à distribuição irregular da freqüência alélica em relação aos outros locos. O loco EMBRA9, com 19 alelos observados, apresentou baixa H_e devido à distribuição irregular das freqüências alélicas, com o presença do alelo 120 pb em alta freqüência em todas as procedências. O sistema Multiplex formado pelos dois locos anteriormente citados, além do EMBRA10, teve o menor H_e média (0,75), seguido do sistema composto pelos locos EMBRA4, EMBRA5 e EMBRA16 (0,89) e EMBRA24, EMBRA27 e EMBRA22 (0,91), cujos resultados foram muito semelhantes aos encontrados na caracterização de quatro locos microsatélites, em um grupo de 20 indivíduos não-relacionados de *E. nitens* (BYRNE et al., 1998).

No presente trabalho, os H_e variaram de 0,72 a 0,91, com H_e média de 0,82, mostrando que microsatélites são, de maneira geral, altamente informativos em espécies do gênero *Eucalyptus* BRONDANI et al. (1998).

3. RESULTADOS E DISCUSSÃO

3.1. Parâmetros de informação genética dos locos microssatélites em seis espécies de *Eucalyptus*

Na caracterização dos locos microssatélites em *E. grandis* (Quadro 2), valores de H_e superiores a 0,9 foram encontrados em três locos: EMBRA22, EMBRA5 e EMRA24, refletindo o grande número de alelos detectados nesses microssatélites. O valor de H_e foi inferior a 0,8 nos locos EMBRA6 e EMBRA18. No loco EMBRA18, a baixa H_e foi relacionada ao reduzido número de alelos detectados e à distribuição irregular da frequência alélica em relação aos outros locos. O loco EMBRA6, com 19 alelos observados, apresentou baixa H_e devido à distribuição irregular das frequências alélicas, com a presença do alelo 120 pb em alta frequência em todas as procedências. O sistema Multiplex formado pelos dois locos anteriormente citados, além do EMBRA10, teve a menor H_e média (0,76), seguido do sistema composto pelos locos EMBRA4, EMBRA5 e EMBRA15 (0,89) e EMBRA24, EMBRA27 e EMBRA22 (0,91), cujos resultados foram muito semelhantes aos encontrados na caracterização de quatro locos microssatélites, em um grupo de 20 indivíduos não-relacionados de *E. nitens* (BYRNE et al., 1996).

No presente trabalho, as H_e variaram de 0,72 a 0,91, com H_e média de 0,83, mostrando que microssatélites são, de maneira geral, altamente informativos em espécies do gênero *Eucalyptus*. BRONDANI et al. (1998),

Quadro 2 – Parâmetros de informação genética dos locos EMBRA que compõem os sistemas Multiplex, caracterizados em 240 indivíduos provenientes de cinco procedências de *E. grandis*. Os parâmetros avaliados foram número de alelos detectados, amplitude alélica (pb), H_e , H_o , I, Q e PIC

Loco	Número de alelos	Amplitude alélica (pb)	H_e	H_o	I	Q	PIC
EMBRA27	21	173-235	0,8869	0,6201	0,0230	0,7820	0,8767
EMBRA22	31	140-230	0,9068	0,5392	0,0152	0,8215	0,9002
EMBRA24	28	167-247	0,9249	0,6881	0,0104	0,8535	0,9202
	26,67		0,9062	0,6158	10^{-6}	0,9943	0,8990
EMBRA5	30	97-161	0,9334	0,7845	0,0083	0,8692	0,9296
EMBRA4	19	56-148	0,8610	0,3863	0,0318	0,7411	0,8485
EMBRA15	24	76-134	0,8634	0,6422	0,0316	0,7437	0,8505
	24,33		0,8859	0,6043	10^{-6}	0,9913	0,8762
EMBRA6	19	114-162	0,7557	0,5760	0,0737	0,6037	0,7418
EMBRA10	12	115-151	0,8043	0,6376	0,0662	0,6432	0,7804
EMBRA18	13	112-144	0,7165	0,5297	0,1105	0,5293	0,6863
	14,67		0,7588	0,5811	10^{-4}	0,9334	0,7362
	21,89 ^a		0,8503 ^b	0,6004 ^c	10^{-14} ^d	0,999997 ^e	0,8371 ^f

^aNúmero médio de alelos observados por loco, ^b heterozigidade média esperada, ^c heterozigidade média observada, ^d probabilidade de identidade combinada (PAETKAU et al., 1995), ^e poder de exclusão combinado (WEIR, 1996) e ^f conteúdo médio de informação de polimorfismo.

caracterizando 15 locos EMBRA, dentre os quais aqueles que compõem dois sistemas Multiplex (EMBRA4, EMBRA5, EMBRA15 e EMBRA10, EMBRA18 e EMBRA6), com 16 indivíduos de *E. grandis*, encontraram uma H_e média praticamente idêntica (0,83) à observada neste trabalho. Cabe salientar ainda que, na caracterização dos locos que foram avaliados em comum nos dois estudos, observou-se que aqueles com maior H_e , na avaliação de BRONDANI et al. (1998), também foram os mais informativos neste trabalho (EMBRA5, EMBRA4, EMBRA15, EMBRA10, EMBRA6 e EMBRA18, em ordem decrescente de H_e). Entretanto, salvo o loco EMBRA5, que apresentou a mesma H_e nos dois estudos (0,93), os outros locos sempre mostraram valores inferiores. Vários fatores podem explicar essa diferença, como a forma de amostragem do material genético, o procedimento de visualização dos fragmentos amplificados e a origem do material genético caracterizado.

Os valores de H_o em *E. grandis* foram expressivamente inferiores aos de H_e para todos os locos, principalmente no EMBRA4 (Quadro 2). Alguns fatores estão relacionados com a redução no número de heterozigotos na população (HARTL e CLARK, 1997). O primeiro é o efeito da estruturação das populações, e o segundo, que pode ser consequência da estruturação, é a ocorrência de endogamia. Embora espécies de *Eucalyptus* em populações naturais sejam preferencialmente alógamas (MORAN e BELL, 1983), a autofecundação entre as plantas é possível, embora resulte freqüentemente em progênies com reduzida capacidade de sobrevivência (GRIFFIN e COTTERIL, 1988). Entretanto, pelo fato de terem sido utilizadas mudas como material genético, estas podem não ter sido submetidas à pressão de seleção, que resultaria na eliminação, ao menos parcial, das plantas originadas de autofecundação. Isso explica também por que BRONDANI et al. (1998), analisando um grupo de indivíduos adultos da mesma espécie e com os mesmos locos, encontraram diferenças menores entre os valores de H_e e H_o , em relação às deste estudo.

O PIC detectado em *E. grandis* variou de 0,69 no loco EMBRA18 a 0,93 no loco EMBRA5, com um valor médio de 0,84 (Quadro 2). Embora alguns locos tenham apresentado PIC relativamente baixo em relação aos microsatélites mais informativos, o valor foi expressivamente superior ao considerado por BOTSTEIN et al. (1980) como indicativo de um loco altamente

informativo, ou seja, com PIC superior a 0,5. De fato, esses pesquisadores, avaliando locos RFLP para fins de mapeamento, puderam considerar apenas 14% deles nessa categoria. Estudos mais recentes, envolvendo a caracterização de marcadores microssatélites em humanos, têm demonstrado que esses locos apresentam, em sua maioria, PIC superior a 0,5 (SPRECHER et al., 1996), sendo, assim, muito mais informativos em relação a locos RFLP. Resultados semelhantes têm sido encontrados em estudos envolvendo animais domésticos (PÉPIN et al., 1995; GLOWATZKI-MULLIS et al., 1995) e culturas agrícolas (RUSSELL et al., 1997), demonstrando a maior diversidade genética encontrada em locos microssatélites, em relação a outros marcadores moleculares.

A menor probabilidade de identidade (0,0083) observada em *E. grandis* foi encontrada no loco EMBRA5 (Quadro 2). Isso indica que a probabilidade de se encontrarem dois indivíduos com genótipo idêntico, para este loco, é inferior a 1, em 100; a maior probabilidade (0,1105) foi observada no loco EMBRA18. Apesar do menor potencial para discriminação de indivíduos desse loco, ele ainda pode ser considerado altamente informativo para uso em sistemas de genotipagem. A I combinada, que considera a probabilidade de serem encontrados indivíduos idênticos, em relação a mais de um loco, foi inferior a 10^{-5} em dois sistemas Multiplex (EMBRA4, EMBRA5, EMBRA15, EMBRA24, EMBRA27 e EMBRA22) e a 10^{-3} no sistema formado pelos locos EMBRA10, EMBRA18 e EMBRA6. Quando todos os locos que compõem os sistemas Multiplex foram considerados, esse valor foi de 10^{-14} , confirmando ser praticamente nula a possibilidade de serem encontrados dois indivíduos com genótipo idêntico em relação a esses locos. O grande número de alelos observados e o relativo equilíbrio na distribuição de freqüências alélicas resultantes da variabilidade genética no grupo de indivíduos analisados se refletem na pequena I desses locos. Embora outros estudos tenham mostrado freqüentemente valores igualmente reduzidos de I (PAETKAU et al., 1995; WENBURG et al., 1996; OLSEN et al., 1996; RUSSELL et al., 1997), poucos têm apresentado locos tão informativos quanto os observados no presente trabalho.

A probabilidade de se excluir um indivíduo escolhido ao acaso nessas populações de *E. grandis* de uma paternidade, ou seja, o Q dos locos EMBRA,

variou de 0,53 (EMBRA18) a 0,87 (EMBRA5) (Quadro 2). O Q combinado, que indica a probabilidade de exclusão considerando mais de um loco, foi superior a 99% em dois sistemas Multiplex e a 93% no terceiro. Quando os nove locos que compõem os sistemas Multiplex foram combinados, esse valor chegou a 99,9997%.

As distribuições das freqüências alélicas observadas nos nove locos EMBRA, em relação às seis espécies de *Eucalyptus* analisadas, mostraram um padrão irregular e multimodal. Tal comportamento é tipicamente observado em marcadores microssatélites em vários organismos, nos quais as freqüências alélicas de locos SSR já foram caracterizadas (VALDES et al., 1993; BREEN et al., 1994; BOWERS et al., 1996). As freqüências alélicas dos locos que compõem os sistemas Multiplex, caracterizadas nas seis espécies de *Eucalyptus*, incluindo as cinco procedências de *E. grandis*, estão descritas no Apêndice A (Figuras 1A a 18A).

Considerando apenas as diferentes procedências de *E. grandis* que foram analisadas, não se distinguiu um padrão de freqüência alélica comum a todos os locos. A maioria deles apresentou o mesmo alelo, ou um pequeno número de alelos, como os mais freqüentes em todas as procedências (alelo 120 pb no loco EMBRA6, por exemplo). Esse padrão se repetiu nos locos EMBRA4 e EMBRA10, com os alelos predominantes 102 pb e 129 pb, respectivamente. O mesmo não foi observado em outros locos, como EMBRA24, em que, salvo algumas exceções, os alelos mais freqüentes foram distintos em cada procedência. O mesmo comportamento foi observado em locos microssatélites caracterizados em populações de espécies arbóreas tropicais (CHASE et al., 1996), populações de ursos polares (PAETKAU et al., 1995), raças de gado (KEMP et al., 1993) e humanos (EDWARDS et al., 1991). Essa característica de cada loco pode estar associada à sua localização em partes mais ou menos conservadas do genoma. A presença de um microssatélite ligado a um gene essencial para a sobrevivência do organismo, o qual ficaria parcial ou totalmente comprometido com qualquer alteração no número de repetições do SSR, levaria, necessariamente, à conservação daquele alelo em alta freqüência em todas as populações. Também, foi detectado que procedências geograficamente mais próximas compartilhavam, muitas vezes, o mesmo ou os mesmos alelos mais freqüentes. Isso foi

observado principalmente entre as procedências Atherton e Wild River, localizadas a poucos quilômetros de distância entre si (Figura 1), que compartilhavam, geralmente, dois a três alelos, dentre os mais freqüentes. A razão para isso está, provavelmente, no fato de haver maior fluxo gênico entre essas procedências, levando a uma troca mais intensa de alelos. O mesmo foi encontrado em um estudo envolvendo a caracterização de 10 microssatélites em três populações de *Oncorhynchus mykiss* (WENBURG et al., 1996). Observou-se que duas, das populações geograficamente próximas, compartilhavam o mesmo alelo mais freqüente em oito locos, enquanto o mesmo ocorria em apenas um microssatélite com a terceira população, localizada em outro continente.

Embora alelos exclusivos a determinada procedência de *E. grandis* tenham sido observados, estes estavam geralmente em freqüências inferiores a 0,1, salvo raras exceções (alelo 78 pb, no loco EMBRA4, exclusivo da procedência Davies Creek, por exemplo) (Apêndice A - Figura 5A). Entretanto, embora não fossem exclusivos, alguns alelos foram encontrados em alta freqüência (superiores a 0,5) em algumas procedências, ao passo que estavam quase ausentes em outras. Dois exemplos são os alelos 181 pb no loco EMBRA27 e 96 pb no loco EMBRA4, praticamente exclusivos da procedência Mount Spec. Entretanto, a ocorrência quase única desses alelos com alta freqüência nesta procedência pode estar associada à maior distância genética entre elas. Possivelmente, ao analisar material genético proveniente de locais menos distantes de Mount Spec, verificou-se que estes alelos poderiam ser detectados com maior freqüência.

A distribuição das freqüências alélicas nas outras cinco espécies de *Eucalyptus* assumiu o mesmo padrão multimodal que havia sido detectado em *E. grandis*. Foi observada pouca similaridade entre os alelos mais freqüentemente registrados nas várias espécies.

De maneira geral, o número médio de alelos detectados nos locos EMBRA foi inferior nas outras espécies em relação ao observado em *E. grandis*. Esse resultado está, certamente, relacionado com o menor número de indivíduos amostrados em cada espécie. Embora uma amostra de aproximadamente 60 indivíduos tenha sido suficiente para gerar estimativas acuradas de H_e , conforme foi visto no Capítulo 2, uma amostra de tal

região não permite a detecção de todo o conjunto de alelos presentes na população, principalmente daqueles em frequência muito baixa. Entretanto, devido ao tamanho das amostras em que foi detectado o menor multiallelismo ($H_e = 0,47$ em *E. caliginosa* e $H_e = 0,48$ em *E. grandis*), o número médio de alelos detectados é próximo ao observado em outros estudos realizados em espécies florestais tropicais (CHASE et al., 1998; WHITE e POWELL, 1997).

Devido todas as espécies estudadas, a maior H_e média foi encontrada em *E. camaldulensis* (0,87), na qual foram detectados 19,22 alelos por locus (Quadro 3). Apesar de o número médio de alelos ter sido inferior ao observado

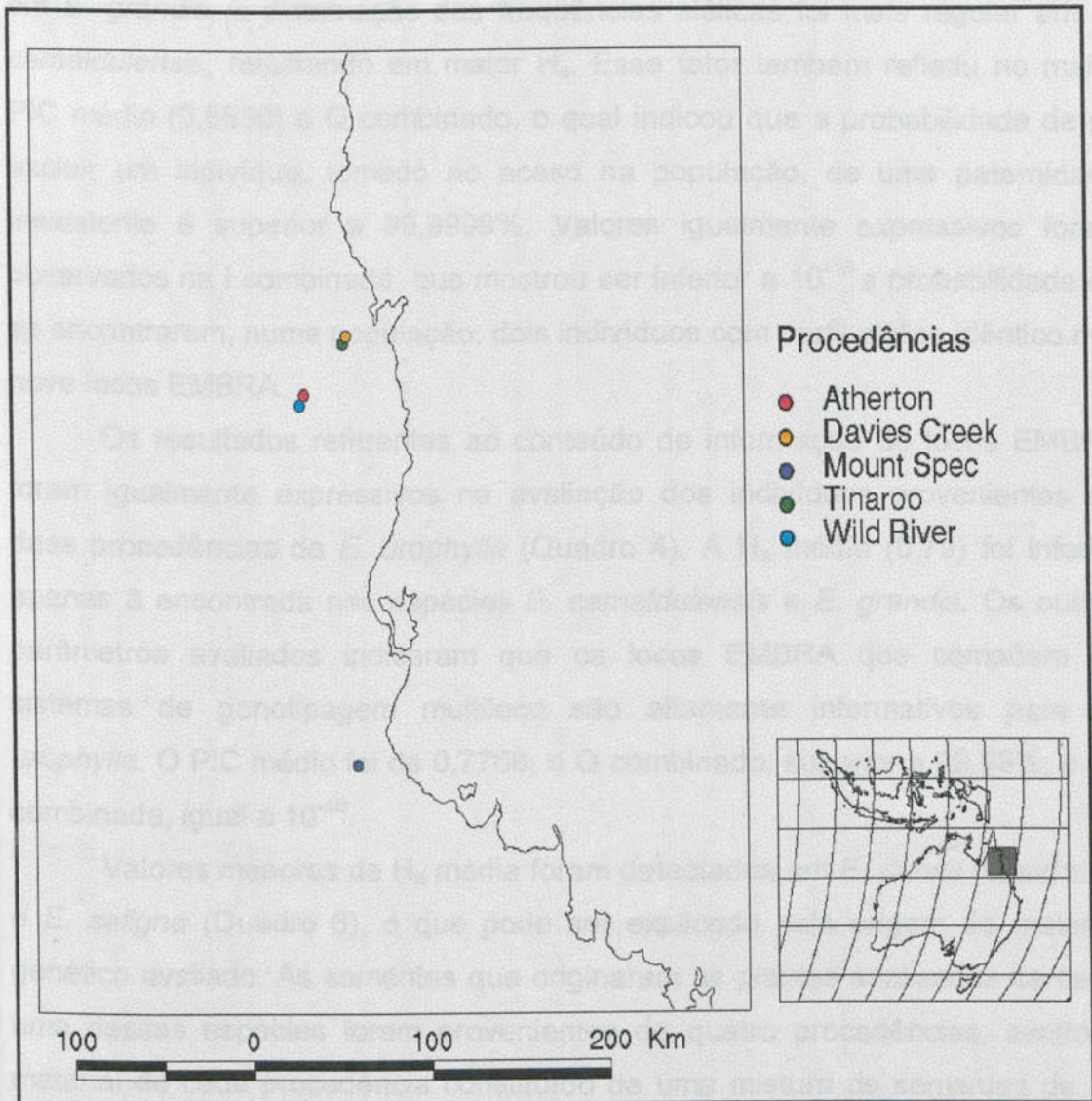


Figura 1 - Localização geográfica das procedências de *E. grandis* analisadas.

magnitude não permite a detecção de todo o conjunto de alelos presentes na população, principalmente daqueles em frequência muito baixa. Entretanto, mesmo nas espécies em que foi detectado o menor multialelismo (*E. dunnii*, *E. globulus* e *E. saligna*), o número médio de alelos detectados e os outros parâmetros avaliados foram superiores aos encontrados em estudos similares, realizados com espécies florestais tropicais (CHASE et al., 1996; WHITE e POWELL, 1997).

Dentre todas as espécies estudadas, a maior H_e média foi encontrada em *E. camaldulensis* (0,87), na qual foram detectados 19,22 alelos por loco (Quadro 3). Apesar de o número médio de alelos ter sido inferior ao observado em *E. grandis*, a distribuição das frequências alélicas foi mais regular em *E. camaldulensis*, resultando em maior H_e . Esse fator também refletiu no maior PIC médio (0,8650) e Q combinado, o qual indicou que a probabilidade de se excluir um indivíduo, tomado ao acaso na população, de uma paternidade inexistente é superior a 99,9999%. Valores igualmente expressivos foram observados na I combinada, que mostrou ser inferior a 10^{-15} a probabilidade de se encontrarem, numa população, dois indivíduos com perfil alélico idêntico nos nove locos EMBRA.

Os resultados referentes ao conteúdo de informação do locos EMBRA foram igualmente expressivos na avaliação dos indivíduos provenientes de duas procedências de *E. urophylla* (Quadro 4). A H_e média (0,79) foi inferior apenas à encontrada nas espécies *E. camaldulensis* e *E. grandis*. Os outros parâmetros avaliados indicaram que os locos EMBRA que compõem os sistemas de genotipagem multiloco são altamente informativos para *E. urophylla*. O PIC médio foi de 0,7766; o Q combinado, superior a 99,99%; e a I combinada, igual a 10^{-10} .

Valores menores de H_e média foram detectados em *E. dunnii* (Quadro 5) e *E. saligna* (Quadro 6), o que pode ser explicado pela origem do material genético avaliado. As sementes que originaram as plantas analisadas de cada uma dessas espécies foram provenientes de quatro procedências, sendo o material de cada procedência constituído de uma mistura de sementes de 10 famílias de meios-irmãos. Como foram amostradas 17 plantas por procedência, deve-se considerar que alguns indivíduos foram originados de, pelo menos, um mesmo parental. Esse fato levou, obrigatoriamente, a uma superestimativa da

Quadro 3 – Parâmetros de informação genética dos locos EMBRA que compõem os sistemas Multiplex, caracterizados em 68 indivíduos provenientes de quatro procedências de *E. camaldulensis*. Os parâmetros avaliados foram número de alelos detectados, amplitude alélica (pb), H_e , H_o , I, Q e PIC

Loco	Número de alelos	Amplitude alélica (pb)	H_e	H_o	I	Q	PIC
EMBRA27	13	181-231	0,8436	0,3091	0,0412	0,7076	0,8269
EMBRA22	15	140-182	0,7504	0,6047	0,0776	0,5948	0,7351
EMBRA24	21	177-231	0,9092	0,7344	0,0145	0,8254	0,9029
	16,33		0,8344	0,5494	10^{-5}	0,9793	0,8216
EMBRA5	23	85-157	0,9165	0,7143	0,0129	0,8368	0,9105
EMBRA4	21	74-134	0,9181	0,3191	0,0122	0,8408	0,9126
EMBRA15	25	76-126	0,9254	0,3770	0,0100	0,8556	0,9210
	23,00		0,9200	0,4701	10^{-6}	0,9962	0,9147
EMBRA6	22	114-170	0,8994	0,7667	0,0179	0,8071	0,8917
EMBRA10	17	123-159	0,9032	0,7761	0,0170	0,8119	0,8955
EMBRA18	16	96-134	0,8057	0,6667	0,0543	0,6603	0,7891
	18,33		0,8694	0,7365	10^{-5}	0,9877	0,8588
	19,22 ^a		0,8746 ^b	0,5853 ^c	10^{-15d}	0,999999 ^e	0,8650 ^f

^a Número médio de alelos observados por loco, ^b heterozigiosidade média esperada, ^c heterozigiosidade média observada, ^d probabilidade de identidade combinada (PAETKAU et al., 1995), ^e poder de exclusão combinado (WEIR, 1996) e ^f conteúdo médio de informação de polimorfismo.

Quadro 4 – Parâmetros de informação genética dos locos EMBRA que compõem os sistemas Multiplex, caracterizados em 68 indivíduos provenientes de duas procedências de *E. urophylla*. Os parâmetros avaliados foram número de alelos detectados, amplitude alélica (pb), H_e , H_o , I, Q e PIC

Loco	Número de alelos	Amplitude alélica (pb)	H_e	H_o	I	Q	PIC
EMBRA27	16	179-223	0,7860	0,5500	0,0649	0,6311	0,7669
EMBRA22	12	144-186	0,8536	0,4444	0,0355	0,7268	0,8395
EMBRA24	21	177-245	0,8378	0,6774	0,0406	0,7068	0,8235
	16,33		0,8258	0,5573	10^{-5}	0,9705	0,8100
EMBRA5	13	107-155	0,8224	0,7551	0,0534	0,6696	0,8005
EMBRA4	17	70-112	0,9078	0,4734	0,0158	0,8196	0,9005
EMBRA15	15	76-128	0,8443	0,3913	0,0388	0,7141	0,8297
	15,00		0,8582	0,5399	10^{-5}	0,9830	0,8436
EMBRA6	17	114-164	0,7384	0,3636	0,0847	0,5791	0,7222
EMBRA10	13	119-157	0,8084	0,6508	0,0599	0,6509	0,7852
EMBRA18	11	104-136	0,5377	0,5238	0,2299	0,3654	0,5215
	13,67		0,6948	0,5127	10^{-3}	0,9068	0,6763
	15,00 ^a		0,7929 ^b	0,5366 ^c	10^{-12d}	0,999953 ^e	0,7766 ^f

^a Número médio de alelos observados por loco, ^b heterozigosidade média esperada, ^c heterozigosidade média observada, ^d probabilidade de identidade combinada (PAETKAU et al., 1995), ^e poder de exclusão combinado (WEIR, 1996) e ^f conteúdo médio de informação de polimorfismo.

Quadro 5 – Parâmetros de informação genética dos locos EMBRA que compõem os sistemas Multiplex, caracterizados em 68 indivíduos provenientes de quatro procedências de *E. dunnii*. Os parâmetros avaliados foram número de alelos detectados, amplitude alélica (pb), H_e , H_o , I, Q e PIC

Loco	Número de alelos	Amplitude alélica (pb)	H_e	H_o	I	Q	PIC
EMBRA27	2	181-207	0,5000	1,0000	0,3750	0,3750	0,3750
EMBRA22	7	144-166	0,6522	0,5600	0,1498	0,4611	0,6234
EMBRA24	14	183-227	0,8337	0,6727	0,0461	0,6913	0,8152
	7,67		0,6620	0,7442	10^{-3}	0,8960	0,6045
EMBRA5	16	107-155	0,8755	0,7368	0,0267	0,7642	0,8643
EMBRA4	13	74-108	0,8781	0,7636	0,0264	0,7658	0,8666
EMBRA15	15	70-134	0,8733	0,4167	0,0275	0,7603	0,8618
	14,67		0,8756	0,6390	10^{-5}	0,9868	0,8642
EMBRA6	13	110-156	0,8352	0,8095	0,0461	0,6917	0,8162
EMBRA10	15	107-155	0,8949	0,6912	0,0200	0,7964	0,8859
EMBRA18	15	102-134	0,7200	0,6269	0,1031	0,5445	0,6953
	14,33		0,8167	0,7092	10^{-5}	0,9714	0,7991
	12,22 ^a		0,7848 ^b	0,6975 ^c	10^{-12d}	0,999961 ^e	0,7560 ^f

^a Número médio de alelos observados por loco, ^b heterozigidade média esperada, ^c heterozigidade média observada, ^d probabilidade de identidade combinada (PAETKAU et al., 1995), ^e poder de exclusão combinado (WEIR, 1996) e ^f conteúdo médio de informação de polimorfismo.

Quadro 6 – Parâmetros de informação genética dos locos EMBRA que compõem os sistemas Multiplex, caracterizados em 68 indivíduos provenientes de quatro procedências de *E. saligna*. Os parâmetros avaliados foram número de alelos detectados, amplitude alélica (pb), H_e , H_o , I, Q e PIC

Loco	Número de alelos	Amplitude alélica (pb)	H_e	H_o	I	Q	PIC
EMBRA27	8	181-209	0,4009	0,3594	0,3742	0,2458	0,3856
EMBRA22	11	144-176	0,7706	0,6071	0,0711	0,6128	0,7522
EMBRA24	10	185-223	0,6766	0,8077	0,1547	0,4592	0,6265
	9,67		0,6160	0,5914	10^{-3}	0,8421	0,5881
EMBRA5	12	101-157	0,7553	0,4595	0,0840	0,5835	0,7312
EMBRA4	7	72-98	0,5774	0,8491	0,2667	0,3213	0,4893
EMBRA15	17	68-130	0,8186	0,2500	0,0516	0,6730	0,7999
	12,00		0,7171	0,5195	10^{-3}	0,9076	0,6735
EMBRA6	16	120-158	0,9099	0,7308	0,0152	0,8232	0,9029
EMBRA10	12	113-147	0,8664	0,7069	0,0325	0,7416	0,8517
EMBRA18	16	96-136	0,8821	0,8545	0,0247	0,7735	0,8713
	14,67		0,8861	0,7641	10^{-5}	0,9897	0,8753
	12,11 ^a		0,7398 ^b	0,6250 ^c	10^{-11d}	0,999849 ^e	0,7123 ^f

^a Número médio de alelos observados por loco, ^b heterozigidade média esperada, ^c heterozigidade média observada, ^d probabilidade de identidade combinada (PAETKAU et al., 1995), ^e poder de exclusão combinado (WEIR, 1996) e ^f conteúdo médio de informação de polimorfismo.

freqüência dos alelos provenientes dos parentais em comum e, conseqüentemente, a uma subestimativa da H_e nessas espécies. A distribuição mais irregular das freqüências alélicas nos locos EMBRA, detectadas em *E. dunnii* e *E. saligna* e associadas ao menor número de alelos detectados, resultou igualmente em estimativas inferiores de PIC médio. O Q e a I combinada demonstraram que tais locos são altamente informativos para fins de genotipagem de indivíduos nessas espécies, apesar de terem poder de discriminação inferior ao observado nas espécies anteriormente citadas. Outro fator que contribuiu acentuadamente para as menores estimativas encontradas para os parâmetros médios de informação genética dos locos EMBRA, em *E. dunnii* e *E. saligna*, foi a presença de alguns locos particularmente pouco informativos; em *E. dunnii*, dois locos apresentaram esse comportamento. No loco EMBRA22 foram detectados apenas sete alelos, e o alelo 154 pb estava em freqüência superior a 0,5. Apesar de apresentar menor conteúdo de informação em relação aos outros microsatélites, o EMBRA22 ainda foi considerado um loco adequado para fins de genotipagem e discriminação de indivíduos. O mesmo não pôde ser concluído para o loco EMBRA27, que apresentou apenas dois alelos na população de *E. dunnii*. Em *E. saligna*, também foram detectados alguns locos com menor conteúdo de informação, dentre eles o EMBRA27, no qual foram detectados apenas oito alelos, dos quais um (207 pb) apresentou freqüência superior a 0,7. No loco EMBRA4, no qual foram detectados sete alelos, os alelos 74 pb e 98 pb contribuíram com quase 90% das freqüências alélicas observadas nessa população. Apesar de esse comportamento ter sido observado em pequeno número de locos em *E. dunnii* e *E. saligna*, ele contribuiu muito para que os parâmetros heterozigosidade média esperada, conteúdo médio de polimorfismo, poder de exclusão combinado e probabilidade de identidade combinada fossem expressivamente inferiores nessas espécies.

Estimativas dos parâmetros de conteúdo de informação genética dos locos EMBRA também foram geradas para a espécie *E. globulus* (Quadro 7), na qual apenas 32 indivíduos foram avaliados; a heterozigosidade média esperada foi de 0,76. Considerando-se que com esse tamanho de amostra há tendenciosidade relativamente acentuada na estimativa da heterozigosidade esperada (Capítulo 2), então o valor encontrado deve ser considerado com

Quadro 7 – Parâmetros de informação genética dos locos EMBRA que compõem os sistemas Multiplex, caracterizados em 32 indivíduos provenientes de cinco procedências de *E. globulus*. Os parâmetros avaliados foram número de alelos detectados, amplitude alélica (pb), H_e , H_o , I, Q e PIC

Loco	Número de alelos	Amplitude alélica (pb)	H_e	H_o	I	Q	PIC
EMBRA27	2	205-207	0,0328	0,0333	0,9361	0,0163	0,0322
EMBRA22	10	142-198	0,6882	0,4286	0,1187	0,5124	0,6668
EMBRA24	15	183-223	0,9170	0,4063	0,0130	0,8366	0,9109
	9,00		0,5460	0,2894	10^{-3}	0,9216	0,5366
EMBRA5	10	81-119	0,8050	0,4667	0,0611	0,6458	0,7819
EMBRA4	15	76-120	0,8944	0,6774	0,0206	0,7940	0,8849
EMBRA15	12	78-112	0,8595	0,2903	0,0339	0,7345	0,8454
	12,33		0,8530	0,4781	10^{-5}	0,9806	0,8374
EMBRA6	13	120-154	0,8835	0,7586	0,0239	0,7767	0,8732
EMBRA10	13	119-151	0,8999	0,7188	0,0186	0,8043	0,8913
EMBRA18	15	92-140	0,8926	0,6875	0,0207	0,7929	0,8834
	13,67		0,8920	0,7216	10^{-6}	0,9909	0,8826
	11,67 ^a		0,7637 ^b	0,4964 ^c	10^{-13a}	0,999986 ^e	0,7522 ^f

^a Número médio de alelos observados por loco, ^b heterozigidade média esperada, ^c heterozigidade média observada, ^d probabilidade de identidade combinada (PAETKAU et al., 1995), ^e poder de exclusão combinado (WEIR, 1996) e ^f conteúdo médio de informação de polimorfismo.

precaução. Entretanto, como essa tendenciosidade é sempre negativa, foi possível considerar que esse valor será crescente ao se analisar maior número de indivíduos daquela espécie. Os outros parâmetros avaliados mostraram que a maioria dos locos eram potencialmente úteis para genotipagem de indivíduos em *E. globulus*. O PIC médio (0,75), o Q combinado (99,99%) e a I combinada de 10^{-10} indicaram que praticamente todos os locos avaliados eram informativos. A única exceção foi o loco EMBRA27, que possuía um alelo (207 pb) praticamente fixado nos indivíduos avaliados. Conseqüentemente, todos os parâmetros estimados indicaram que esse loco não era apropriado para genotipagem de indivíduos de *E. globulus*, da mesma forma como foi observado em *E. dunnii*.

3.2. Diferenciação gênica entre procedências de *E. grandis* e entre seis espécies de *Eucalyptus*

Os valores estimados de P nos testes de probabilidade realizados entre todas as procedências de *E. grandis*, para todos os locos que compõem os sistemas de genotipagem, indicaram que a distribuição alélica não foi idêntica entre todas as procedências, para todos os locos ($P < 0,05$) (Apêndice B: Quadros 1B a 9B). Apenas um loco, EMBRA15, apresentou valor de P superior a 0,01 ($P = 0,01206$). Conforme o esperado, esse valor foi encontrado entre Atherton e Wild River, demonstrando novamente o maior fluxo gênico entre essas procedências, localizadas a poucos quilômetros de distância uma da outra (Figura 1). Nos locos restantes, o valor de P foi invariavelmente inferior a 0,01, indicando claramente que há distribuição alélica independente entre as procedências.

Resultados semelhantes foram encontrados na análise de diferenciação gênica entre as espécies de *Eucalyptus* analisadas. Salvo para o loco EMBRA22, em que foi encontrado um valor de P de 0,0234 entre *E. grandis* e *E. globulus*, no restante dos locos e em combinações de espécies o valor de P foi sempre inferior a 10^{-5} . Conforme era esperado e considerando que a distribuição alélica não foi idêntica entre diferentes procedências de *E. grandis*, verificou-se que esta também não o foi entre as diferentes espécies de *Eucalyptus*.

3.3. Distância genética e análise de variância molecular entre procedências de *E. grandis*

As distâncias genéticas de Nei (1972), calculadas entre as cinco procedências de *E. grandis* (Quadro 8), indicaram, de maneira geral, um padrão de crescente distância genética associada à crescente distância geográfica entre as procedências (Figura 1), analogamente ao que foi encontrado em estudos similares (SHEELY e MEAGHER, 1996). Considerando inicialmente a procedência Mount Spec em relação a Tinaroo e Davies Creek, observou-se que entre estas foram encontradas as maiores distâncias genéticas (0,6897 e 0,6267, respectivamente). Esse resultado foi concordante com a distância geográfica, considerando-se que a distância entre Mount Spec e Tinaroo/Davies Creek foi a maior entre todas as procedências analisadas. Observou-se também que as distâncias genéticas entre Mount Spec e Atherton e Wild River (0,3954 e 0,4224) e entre Davies Creek e Tinaroo (0,6267 e 0,6897) foram praticamente idênticas. De fato, ao observar a localização geográfica das procedências que compõem esses dois grupos (Atherton/Wild River e Davies Creek/Tinaroo), percebeu-se que elas estavam localizadas a poucos quilômetros uma da outra. O grupo Atherton/Wild River estava localizado mais ao sul e, conseqüentemente, mais próximo de Mount Spec, de maneira que o maior fluxo gênico entre essas procedências pode justificar a menor distância genética encontrada. Ao ser analisada a relação entre as procedências Atherton e Wild River, localizadas a poucos quilômetros de distância entre si, observou-se que, conforme era esperado, a distância genética detectada foi a menor de todas (0,2053). Entretanto, o mesmo padrão não ocorreu nas procedências Davies Creek e Tinaroo, as quais, embora localizadas igualmente próximas geograficamente, foram distantes geneticamente (0,4881).

Os resultados da AMOVA, com a distribuição da variação entre populações e dentro delas, considerando as cinco procedências de *E. grandis*, encontram-se no Quadro 9. Constatou-se que existe diferença significativa ($P < 0,05$) entre procedências, embora ela seja responsável por apenas 8,02% da variação. O restante (91,98%) está dentro das procedências. Esse padrão de distribuição da variação está em concordância com os resultados

Quadro 8 – Matriz de distância genética de Nei (1972) entre cinco procedências de *E. grandis*, determinada por meio da análise de 48 indivíduos de cada procedência, com nove locos EMBRA

	Atherton	Davies Creek	Mount Spec	Tinaroo	Wild River
Atherton	0				
Davies Creek	0,4825	0			
Mount Spec	0,3954	0,6267	0		
Tinaroo	0,3966	0,4881	0,6897	0	
Wild River	0,2053	0,5523	0,4224	0,6897	0

Quadro 9 – Análise de variância molecular (AMOVA) entre cinco procedências de *E. grandis*, avaliadas com nove locos EMBRA

F.V.	g.l.	s.q.	Componente de variância	% variação	P
Entre procedências	4	109,66	0,26	8,02	< 0,01
Dentro de procedências	463	1.386,34	2,99	91,98	
Total	467	1.496,00	3,25		

encontrados por NESBITT et al. (1995), que analisaram a estrutura genética de populações de *E. globulus*. Esses autores verificaram que 80,2% da variação total ocorria dentro das localidades e o restante, entre localidades e regiões. Outros estudos envolvendo espécies alógamas têm mostrado resultados semelhantes (ROSSETTO et al., 1995; MAGUIRE e SEDGLEY, 1997). Cabe observar, ainda, que, apesar de o padrão de distribuição da variância genética ser semelhante ao observado nos estudos citados anteriormente, a proporção da variação entre procedências foi, apesar de significativa, relativamente pequena em relação aos trabalhos de NESBITT et al. (1995). Dificilmente a classe de marcador utilizada seria responsável por essa diferença, tendo em vista que marcadores baseados em microssatélites apresentam, tradicionalmente, mais conteúdo de informação do que aqueles baseados na amplificação de seqüências aleatórias (RAPD). Entretanto, as características

das procedências avaliadas poderiam ter papel importante nessa questão. NESBITT et al. (1995) avaliaram várias procedências de *E. globulus* localizadas no continente australiano ou em ilhas do mar da Tasmânia, sujeitas a um isolamento que contribuiu com 3,5% da variância total (variância entre regiões). Assim, neste estudo foram avaliadas procedências relativamente próximas entre si, como no caso de Atherton e Wild River e Davies Creek e Tinaroo. A ausência de isolamento geográfico entre essas procedências permitiu maior fluxo gênico entre elas e, conseqüentemente, menor variação.

Os resultados indicaram que a coleta de material genético de *E. grandis*, para fins de conservação ou estabelecimento de programas de melhoramento genético, pode ser feita em menor número de procedências, tendo em vista que a maior parte da variação está dentro delas. Entretanto, a coleta em várias procedências permanece importante, principalmente entre grupos distantes geograficamente, pelo fato de se terem detectado variações significativas entre elas.

4. RESUMO E CONCLUSÕES

O conteúdo de informação dos locos EMBRA que compõem os sistemas de genotipagem multilocos foram avaliados em seis espécies de importância comercial pertencentes ao gênero *Eucalyptus*, subgênero *Symphomyrtus*.

Inicialmente, indivíduos provenientes de cinco procedências de *E. grandis* foram caracterizados. Com base nos genótipos observados, os índices H_e , PIC, I e Q, referentes aos locos EMBRA, foram calculados. Todos os parâmetros estimados indicaram que os locos avaliados são muito apropriados para genotipagem de indivíduos de *E. grandis*.

A distribuição das freqüências alélicas seguiu o padrão irregular, multimodal, característico dos marcadores microssatélites.

Considerando as informações de freqüências alélicas geradas para as cinco procedências de *E. grandis*, procedeu-se a uma avaliação da distância genética entre elas, assim como a uma análise de variância molecular. A distribuição da variância genética detectada na AMOVA seguiu o padrão normalmente observado em espécies alógamas, com mais de 90% da variância sendo encontrada dentro das procedências e menos de 10% entre elas.

Os locos EMBRA caracterizados nas outras cinco espécies, *E. camaldulensis*, *E. dunnii*, *E. globulus*, *E. saligna* e *E. urophylla*, apresentaram o mesmo padrão de distribuição multimodal de freqüências alélicas que havia sido observado em *E. grandis*.

A maioria dos locos apresentou-se com elevado conteúdo de informação, sendo considerados apropriados à genotipagem para fins de discriminação de indivíduos e estudos de parentesco. Entretanto, em alguns casos, o número de alelos detectados foi expressivamente reduzido, ocorrendo até mesmo um alelo praticamente fixado no loco EMBRA27, em *E. globulus*. Esse microssatélite também apresentou pouca diversidade em *E. dunnii* e relativamente pouca diversidade em *E. saligna*. Esse loco poderá ser substituído por outro microssatélite nos sistemas de genotipagem que forem desenvolvidos para serem utilizados na genotipagem de indivíduos dessas espécies.

Os resultados indicaram claramente a importância da caracterização dos locos EMBRA que compõem os sistemas de genotipagem, anteriormente à sua aplicação na análise de indivíduos de espécies ainda não avaliadas. Em algumas espécies, esses locos podem estar intimamente ligados a genes de importância fundamental para a sobrevivência da planta, de maneira que eles não apresentarão a mesma variabilidade observada em outras espécies.

Apesar da presença de locos com menor conteúdo de informação em alguns casos, os sistemas de genotipagem foram muito informativos nas espécies estudadas. De fato, houve alta transferibilidade desses sistemas para as espécies avaliadas e, potencialmente, também para outras espécies do subgênero *Symphyomyrtus*.

No presente trabalho, criou-se o primeiro banco de dados de freqüências alélicas de locos microssatélites em espécies do gênero *Eucalyptus*. Espera-se que as informações que foram geradas sejam incrementadas através da análise de maior número de indivíduos das espécies descritas, além da adição de novas espécies ao conjunto já estudado. A padronização dos locos analisados nessas espécies possibilitará o uso universal do sistema de genotipagem, facilitando todos os procedimentos de troca e identificação de material genético nas empresas e institutos de pesquisa, da mesma forma como foi proposto para outras espécies (THOMAS et al., 1994; DIWAN CREGAN, 1997).

CAPÍTULO 4

COMPARAÇÃO ENTRE GENOTIPAGENS MANUAL E SEMI-AUTOMATIZADA DE MICROSSATÉLITES EM POPULAÇÕES DE MELHORAMENTO DE *Eucalyptus*

1. INTRODUÇÃO

A utilização de marcadores moleculares tem sido incorporada, de forma crescente, em alguns programas de melhoramento genético e produção florestal de *Eucalyptus* no Brasil. Esta tecnologia tem encontrado aplicação operacional principalmente na avaliação da diversidade genética de populações-base, na identificação e discriminação de clones (COSTA E SILVA et al., 1997) e no monitoramento da taxa de fecundação cruzada em pomares e populações de melhoramento (GAIOTTO et al., 1997). Outras potenciais aplicações da tecnologia de análise genômica no melhoramento têm sido demonstradas envolvendo mapeamento de QTLs para características diretamente relacionadas à produção (GRATTAPAGLIA e SEDEROFF, 1996), seleção assistida por marcadores (SQUILLASSI e GRATTAPAGLIA, 1997) e identificação de genes controlando a resistência a patógenos (JUNGHANS et al., 1998).

Tendo em vista a utilização crescente de florestas clonais de eucalipto, observou-se que tem aumentado significativamente a necessidade de

tecnologias extremamente precisas para discriminação individual de clones, mesmo em situações em que existe possível ancestralidade comum entre indivíduos. Embora marcadores moleculares RAPD tenham sido utilizados, com sucesso, para tal finalidade em alguns casos (KEIL e GRIFFIN, 1994), a sua natureza dominante e o baixo nível de polimorfismo tipicamente existente entre indivíduos geneticamente relacionados e a repetibilidade parcial dos perfis de marcadores gerados não permitem afirmar, de forma categórica, que dois indivíduos sejam de fato o mesmo. Frequentemente, o mesmo indivíduo pode apresentar bandas RAPD a mais ou a menos, dependendo da qualidade do DNA e das condições específicas de amplificação utilizadas, indicando que se trata de dois indivíduos distintos, quando, na verdade, eles são idênticos.

Recentemente, marcadores geneticamente mais informativos, baseados em polimorfismos de microssatélites, foram desenvolvidos para *Eucalyptus* (BYRNE et al., 1996; BRONDANI et al., 1998). Devido à abundância e à distribuição regular dessas seqüências no genoma de eucariotos, associadas ao multialelismo e ao padrão de segregação co-dominante, os microssatélites têm tido ampla aplicação em estudos de genética de populações (CHASE et al., 1996; WHITE e POWELL, 1997) e na avaliação de recursos genéticos (MICHELL et al., 1997). Essa tecnologia permitiu avanço significativo na determinação de parentesco (MARKLUND et al., 1994; GLOWATZKI-MULLIS et al., 1995) e na discriminação individual e proteção de cultivares (THOMAS et al., 1994; RUSSELL et al., 1997; DIWAN e CREGAN, 1997).

O uso de marcadores moleculares baseados em microssatélites tem, portanto, aplicação potencial imediata, principalmente nos programas de produção florestal de *Eucalyptus*. Devido a sua maior robustez e seu maior conteúdo de informação e facilidade de comparação entre laboratórios em relação a marcadores dominantes, os microssatélites permitem a determinação de genótipos multilocos únicos, possibilitando a identificação e proteção precisas e eficientes de genótipos superiores. Esses dados podem ser empregados para entender a origem, a divergência e a relação entre clones de procedência desconhecida para resolver problemas de identidade clonal em pomares de sementes e para quantificar precisamente os níveis de variabilidade genética entre procedências e progênes e dentro delas,

otimizando-se, assim, os trabalhos de coleta de germoplasma e enriquecimento de populações-base.

No presente trabalho, os sistemas semi-automatizados de genotipagem multilocos desenvolvidos nas etapas anteriores foram aplicados na genotipagem dos indivíduos que compõem duas populações de melhoramento da empresa Aracruz Celulose S.A. Foram determinados genótipos multilocos únicos para todos os indivíduos, gerando-se, assim, um banco de dados extremamente úteis para o entendimento da estrutura dessas populações, para a proteção de clones e para o controle de qualidade de cruzamentos controlados. Além disso, foi realizada uma análise comparativa da eficiência, em termos de rapidez, do procedimento de genotipagem manual, utilizando-se coloração com nitrato de prata; e da genotipagem semi-automatizada, com fluorescência em seqüenciador automático.

2.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificadas

Dois populações de melhoramento da empresa Aracruz Celulose S.A. foram analisadas.

A primeira população era composta de 192 indivíduos de *E. grandis*, provenientes de 18 procedências, conforme foi descrito no Capítulo 2. A extração do DNA genômico foi feita de tecido foliar de plantas adultas, relacionadas a partir de etiquetas de procedências próprias estabelecidas na empresa. Esta população foi analisada com seis microsatélites, EMBR14, EMBR15, EMBR16, EMBR17, EMBR18 e EMBR19 (BROWN *et al.*, 1993). Os procedimentos de amplificação via PCR descritos e a separação dos fragmentos amplificadas foram descritos anteriormente (Capítulo 2). Após serem corados com nitrato de prata, segundo o método descrito por BASSAN *et al.* (1991), a análise dos fragmentos amplificados foi realizada manualmente com base na comparação com o marcador padrão 100 bp (GeneAid).

A segunda população analisada foi composta por 165 indivíduos de *E. grandis*, relacionados a partir de etiquetas de procedências próprias estabelecidas na Aracruz Celulose S.A. (Estação Santa Inês), além de 40 clones de mesma espécie, plantados comercialmente por esta empresa. O DNA

2. MATERIAL E MÉTODOS

2.1. Material vegetal, reação de PCR, eletroforese e detecção dos fragmentos amplificados

Duas populações de melhoramento da empresa Aracruz Celulose S.A. foram analisadas.

A primeira população era composta de 192 indivíduos de *E. grandis*, provenientes de 18 procedências, conforme foi descrito no Capítulo 2. A extração de DNA genômico foi feita de tecido foliar de plantas adultas, selecionadas a partir de ensaios de procedências/progênes estabelecidos na empresa. Esta população foi analisada com seis microssatélites: EMBRA4, EMBRA5, EMBRA10, EMBRA15, EMBRA16 e EMBRA11 (BRONDANI et al., 1998). Os procedimentos de amplificação via PCR desses locos e a separação dos fragmentos amplificados foram descritos anteriormente (Capítulo 2). Após serem corados com nitrato de prata, seguindo-se a metodologia descrita por BASSAN et al. (1991), o tamanho dos fragmentos amplificados foi estimado manualmente, com base na comparação com o marcador padrão 10 pb (Gibco BRL).

A segunda população analisada foi composta por 185 indivíduos de *E. urophylla*, selecionados a partir de testes de procedências/progênes, instalados na Aracruz Celulose S. A. (Espírito Santo, Brasil), além de 40 clones da mesma espécie, plantados comercialmente por essa empresa. O DNA

genômico desses indivíduos foi extraído conforme a metodologia descrita por DOYLE e DOYLE (1987). Nesta população foram analisados seis locos microssatélites que compõem dois sistemas de genotipagem multilocos desenvolvidos nas etapas anteriores: EMBRA27, EMBRA22, EMBRA5, EMBRA4, EMBRA24 e EMBRA15. Os procedimentos de amplificação desses locos e a separação dos fragmentos seguiram o protocolo desenvolvido no Capítulo 1. Diferentemente do processo de genotipagem manual dos indivíduos, através da determinação visual do tamanho dos fragmentos amplificados, utilizado na avaliação da população de *E. grandis*, os genótipos dessa população foram gerados por meio da estimativa automatizada do tamanho dos fragmentos. Nesse processo, os produtos da amplificação via PCR dos locos microssatélites foram registrados através de um sistema de detecção de fluorescência (seqüenciador de DNA), e o tamanho desses fragmentos foi estimado através de programas específicos (GeneScan e Genotyper) (Figuras 1 e 2).

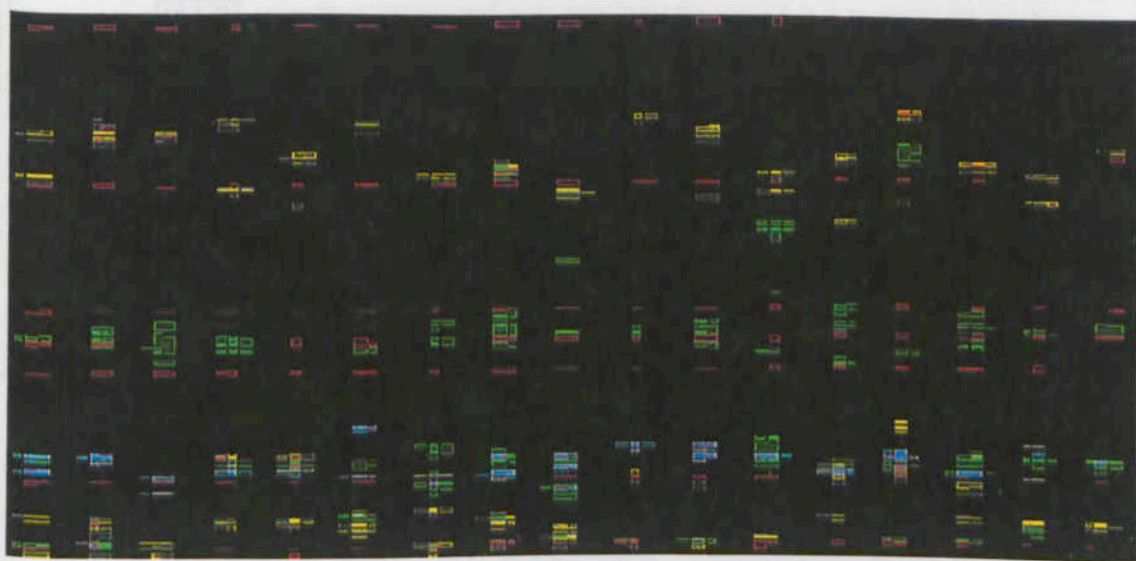


Figura 1 – Imagem do gel mostrando 17 genótipos de *E. grandis* analisados com cinco locos EMBRA, tomada do seqüenciador de DNA ABI Prism, modelo 377, por meio do programa de análise GeneScan.

2.7. Análise de similaridade entre os indivíduos das duas populações de melhoramento

Com as duas matrizes de genótipos foram geradas matrizes de similaridade genéticas, baseadas no coeficiente de Dice, para os indivíduos que compõem as duas populações de melhoramento. Com base nas similaridades geradas, foi realizada uma análise de agrupamento, usando-se o método SHAN, baseado em UPGMA (*unweighted pair group analysis*). A análise foi feita com o programa NT-SYS (Applied Biostatistics Inc., versão 2.0).

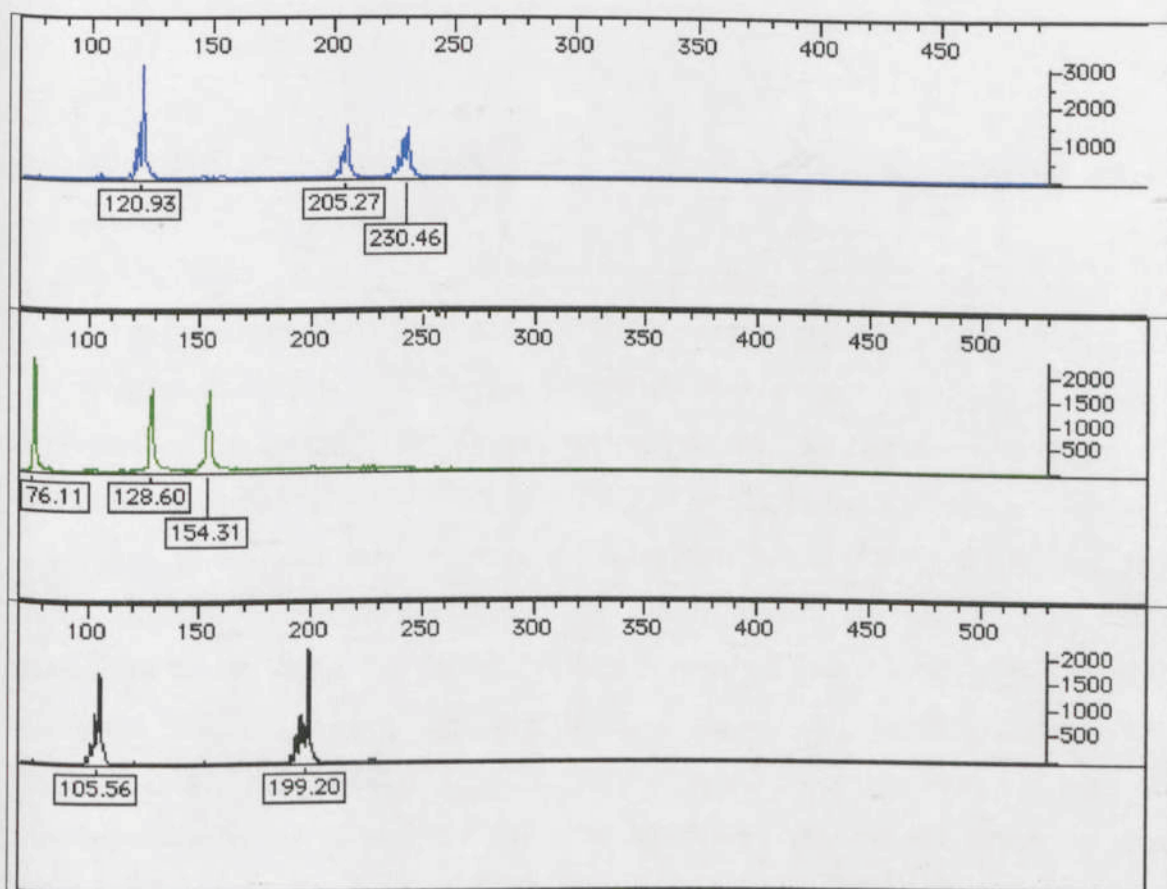


Figura 2 – Resolução de fragmentos amplificados de seis locos EMBRA que compõem dois sistemas Multiplex, separados em uma linha do gel e analisados através dos programas GeneScan e Genotyper. Na faixa superior, podem ser observados os locos EMBRA5 (homozigoto, 121 pb) e EMBRA27 (heterozigoto, 205 pb e 230 pb); na faixa do meio, os locos EMBRA15 (homozigoto, 76 pb) e EMBRA22 (heterozigoto, 129 pb e 154 pb); e na faixa inferior, os locos EMBRA4 (homozigoto, 106 pb) e EMBRA24 (homozigoto, 199 pb).

2.2. Similaridade genética entre os indivíduos das duas populações de melhoramento

Com o uso das informações de genotipagem foram geradas matrizes de similaridade genéticas, baseadas no coeficiente de Dice, para os indivíduos que compõem as duas populações de melhoramento. Com base nas similaridades geradas, foi realizada uma análise de agrupamento, usando-se o método SHAN, baseado em UPGMA (*unweighted pair group analysis*). As análises foram feitas com o programa NT-SYS ("Applied Biostatistics Inc"., versão 2.0).

3.1. Caracterização genotípica das populações de melhoramento de *E. urophylla* e *E. grandis*

Os genótipos multialélicos das 225 indivíduos da população de *E. urophylla* foram determinados com base nos alelos observados em cinco locos EMRA (Apêndice C - Quadro 1C). O número mínimo de alelos detectados na caracterização de locos EMRA22 levou à sua exclusão da análise.

Foi observado que, apesar da elevada multialélica detectada nos microsatélites, na população de *E. urophylla* não foram observados alelos com frequência maior de 54 de 55 e 36 de 51. A ocorrência de indivíduos com genótipos idênticos indica que os indivíduos não são considerados como indivíduos únicos na realidade, pois o perfil genotípico observado em alguns indivíduos pode ser idêntico ao observado em outros indivíduos. Portanto, neste trabalho, a identificação dos indivíduos foi baseada no método de agrupamento de indivíduos com base nos dados de genotipagem em locos de microsatélites.

É importante observar, também, que alguns genótipos, como 145/147/148/150, CP3/CP23, CP6/CP21 e GP34/CP36, apresentaram diferenças genotípicas em relação a um número reduzido de alelos (um ou dois). Tais diferenças foram geradas de apenas um par de bases, quando, portanto, há diferença entre os indivíduos observados em outros locos.

3. RESULTADOS E DISCUSSÃO

3.1. Caracterização genotípica das populações de melhoramento de *E. urophylla* e *E. grandis*

Os genótipos multilocos dos 225 indivíduos da população de *E. urophylla* foram determinados com base nos alelos observados em cinco locos EMBRA (Apêndice C - Quadro 1C). O excessivo número de falhas detectadas na caracterização do loco EMBRA22 levou à sua exclusão da análise.

Foi observado que, apesar do elevado multialelismo detectado nos microssatélites, naquela população os cinco locos não discriminaram os genótipos 64 de 65 e 36 de 51. A ocorrência de indivíduos com genótipo multiloco idêntico indicou claramente que os clones considerados como distintos eram, na realidade, iguais. O problema provavelmente ocorreu no estabelecimento da população de melhoramento ou na identificação dos indivíduos que a compõem. Entretanto, existe também a possibilidade de ter havido erro na coleta do material na empresa ou, ainda, durante os procedimentos de extração de DNA no laboratório.

É importante observar, também, que alguns genótipos, como 145/147/148/150, CP3/CP23, CP8/CP21 e CP34/CP35, apresentaram diferenças genotípicas em relação a um número muito reduzido de alelos (um ou dois). Tais diferenças foram, geralmente, de apenas um par de bases, estando, portanto, dentro da margem de erro observada na genotipagem semi-

automatizada, com a utilização de um seqüenciador de DNA. Calculou-se a probabilidade de se terem dois genótipos idênticos em relação aos locos nos quais os clones CP foram iguais (ou seja, as probabilidades de identidade combinadas). Observou-se que essa probabilidade era inferior a 10^{-8} para os genótipos CP3/CP23 e CP8/CP21 e 10^{-7} para CP34/CP35. Considerou-se que, se estes genótipos foram idênticos em relação a esses locos, então há grande probabilidade de que eles sejam, na realidade, os mesmos indivíduos ou indivíduos com alto grau de parentesco.

O restante dos indivíduos da população de *E. urophylla* foi discriminado adequadamente por meio dos locos analisados, apresentando genótipos multilocos únicos.

Os genótipos multilocos dos 192 indivíduos da população de *E. grandis* foram determinados com base nos alelos observados em seis locos EMBRA e encontram-se descritos no Apêndice C (Quadro 2C). Em razão do elevado conteúdo informativo encontrado nesses locos, os 192 indivíduos foram facilmente discriminados uns dos outros. Verificou-se, inclusive, que com apenas os locos EMBRA4, EMBRA5 e EMBRA16 era possível estabelecer um perfil alélico exclusivo para cada indivíduo da população de melhoramento analisada.

Ao contrário do que foi detectado na população de melhoramento de *E. urophylla*, não foram observados dois clones com perfil alélico idêntico na população de *E. grandis*.

3.2. Similaridade genética entre os clones das populações de melhoramento de *E. urophylla* e *E. grandis*

Com base nos genótipos detectados nos 225 clones de *E. urophylla*, o coeficiente de similaridade de Dice foi calculado entre todos os indivíduos desta população de melhoramento. O coeficiente de similaridade médio foi de 0,0764, muito inferior ao observado em outro estudo, que avaliou genótipos selecionados de *E. urophylla* da Aracruz Celulose S.A. com 62 marcadores RAPD, em que se obteve um coeficiente de similaridade médio de Dice de 0,529 (RIBEIRO et al., 1997). A diferença entre os dois resultados está provavelmente relacionada à classe de marcador molecular empregada. De

fato, os microsatélites apresentaram, reconhecidamente, maior diversidade do que os marcadores baseados em marcadores RAPD. O fenograma gerado encontra-se no Apêndice D (Figura 1D).

Foi observado que os clones plantados comercialmente (CP) formaram um agrupamento no qual foram encontrados, praticamente, $\frac{3}{4}$ dos clones incluídos nessa categoria, mostrando haver maior similaridade entre eles, em relação ao resto da população de melhoramento. Os 10 clones-elite restantes foram observados isolados ou formando agrupamentos de dois a três indivíduos.

A população de melhoramento de *E. grandis* também apresentou coeficiente médio de similaridade de Dice muito baixo, atingindo o valor de 0,1357. Apesar de o valor médio ser aproximadamente duas vezes superior ao que foi encontrado na população de *E. urophylla*, ele permaneceu muito reduzido. O fenograma representativo dos coeficientes de similaridade gerados entre os indivíduos encontra-se no Apêndice D (Figura 2D).

Os resultados encontrados têm aplicação imediata no programa de melhoramento da empresa, através do direcionamento dos cruzamentos. Cruzamentos entre indivíduos mais similares ou menos similares podem ser indicados, dependendo dos objetivos do melhorista. Cruzamentos entre indivíduos mais divergentes apresentam grande potencial de gerar progênies altamente segregantes, incrementando as possibilidades de ganho na seleção pela escolha dos indivíduos localizados no extremo da variação fenotípica. Inversamente, o cruzamento entre indivíduos menos divergentes geraria maior uniformidade das progênies, tendo em vista que envolveria clones mais semelhantes geneticamente. As sementes resultantes desses cruzamentos tenderiam a apresentar bom desempenho, embora associado ao risco de depressão por endogamia, e poderiam ser úteis em plantios comerciais através de sementes.

3.3. Caracterização genética dos locos EMBRA nas populações de *E. urophylla* e *E. grandis*

Foram detectados 124 alelos nos cinco locos analisados na população de *E. urophylla*, com uma média de 24,8 alelos por loco (Quadro 1). O maior

Quadro 1 – Número de alelos detectados, amplitude alélica, H_e , H_o , PIC, I e Q observados nos locos EMBRA27, EMBRA5, EMBRA4, EMBRA24 e EMBRA15, na avaliação de uma população de melhoramento de *E. urophylla*

Loco	Número de alelos	Amplitude alélica	H_e	H_o	PIC	I	Q
EMBRA27	26	179-230	0,8882	0,6758	0,8787	0,0220	0,7863
EMBRA5	28	80-160	0,9166	0,8578	0,9111	0,0125	0,8386
EMBRA4	21	70-113	0,9127	0,6484	0,9064	0,0139	0,8303
EMBRA24	25	175-245	0,9028	0,7788	0,9022	0,0149	0,8193
EMBRA15	24	70-136	0,8762	0,3777	0,8682	0,0234	0,7734
	24,8*		0,9005*	0,6677*	0,8933*	10^{-9} **	0,9978**

* Média dos valores observados em cada loco e ** probabilidade de identidade combinada e poder de exclusão combinado, considerando-se os cinco locos avaliados.

número de alelos (28) foi observado no loco EMBRA5 e o menor, no EMBRA4 (21). A heterozigosidade esperada e o conteúdo de informação de polimorfismo foram, em média, 0,90 e 0,89, respectivamente. O poder de exclusão exibiu valor combinado superior a 99%. Finalmente, a probabilidade de identidade variou de 0,0125 no loco EMBRA5 até 0,0234 no loco EMBRA15. Deve-se salientar que a probabilidade de identidade combinada, ou seja, considerando os cinco locos microssatélites estudados, foi inferior a 10^{-9} , evidenciando que a probabilidade de haver dois genótipos idênticos nessa população, para esses locos, é inferior a 1 em um bilhão.

Um total de 119 alelos foi detectado nos seis locos EMBRA caracterizados nos 192 indivíduos de *E. grandis* (Quadro 2). O maior número de alelos foi verificado no loco EMBRA16, que apresentou 33 alelos, ao passo que apenas seis alelos foram observados em EMBRA11. A heterozigosidade esperada variou de 0,65 a 0,94, nos locos EMBRA11 e EMBRA5, respectivamente. A menor probabilidade de identidade foi encontrada no loco EMBRA5, coincidindo com o loco de maior heterozigosidade, e a probabilidade de identidade mais alta ocorreu no EMBRA11. Levando-se em conta todos os seis locos avaliados, a probabilidade de identidade combinada foi de 10^{-9} . O poder de exclusão combinado foi superior a 0,9999, com o maior índice sendo encontrado no loco EMBRA5 (0,87) e, o menor, no EMBRA11 (0,43).

Quadro 2 – Número de alelos detectados, amplitude alélica, H_e , H_o , PIC, I e Q observados nos locos EMBRA4, EMBRA5, EMBRA10, EMBRA15, EMBRA16 e EMBRA11, na avaliação de uma população de melhoramento de *E. grandis*

Loco	Número de alelos	Amplitude alélica	H_e	H_o	PIC	I	Q
EMBRA4	19	74-120	0,9064	0,5355	0,8981	0,0749	0,8177
EMBRA5	26	104-160	0,9360	0,6630	0,9327	0,0073	0,8745
EMBRA10	15	116-150	0,8512	0,6290	0,8350	0,0383	0,7190
EMBRA15	20	110-154	0,9027	0,6236	0,8949	0,0173	0,8111
EMBRA16	33	106-194	0,9297	0,2011	0,9105	0,0383	0,8643
EMBRA11	6	124-146	0,6451	0,4658	0,5236	0,1684	0,4339
	19,83*		0,8619*	0,5197*	0,8325*	10^{-9} **	0,9999**

* Média dos valores observados em cada loco e ** probabilidade de identidade combinada e poder de exclusão combinado, considerando-se os cinco locos avaliados.

3.4. Comparação entre as eficiências da genotipagem manual e da genotipagem semi-automatizada

A eficiência do sistema de genotipagem semi-automatizado oferece uma série de vantagens sobre os métodos manuais de avaliação dos microssatélites amplificados. Em primeiro lugar, a acurácia dos dados é melhorada, pelo fato de que o tamanho dos alelos é determinado através de algoritmos que comparam o tamanho dos fragmentos do marcador-padrão interno, que são conhecidos, com o tamanho dos fragmentos detectados. Além disso, a passagem dos dados para programas de análise genética fica isenta de erros decorrentes da digitação dos genótipos. Entretanto, a principal vantagem que o sistema de genotipagem semi-automatizado apresenta sobre o processo manual está na rapidez da geração dos dados.

A análise de 192 indivíduos provenientes da população de melhoramento de *E. grandis* consumiu um período de aproximadamente três semanas. Quatro a cinco dias foram necessários para os processos de extração, quantificação e diluição do DNA. Na etapa seguinte foram preparadas 1.152 reações de PCR (192 amostras analisadas com seis locos

microssatélites), o que tomou mais dois a três dias de trabalho. O processo de separação dos fragmentos amplificados foi bastante eficiente, devido à disponibilidade de cubas de eletroforese com a capacidade de análise simultânea de 100 amostras. Considerando que até três eletroforeses foram feitas diariamente, com a análise simultânea de dois locos em cada gel, chegou-se a um total de 100 indivíduos analisados com seis locos por dia. Assim, o tempo necessário para analisar a população de 192 indivíduos foi de dois dias. Finalmente, a etapa mais demorada e exaustiva, a obtenção da estimativa do tamanho dos fragmentos amplificados, tomou aproximadamente um dia para avaliação de cada loco, totalizando-se, assim, mais seis dias de trabalho. No total foram contabilizados 16 dias trabalhados, aproximadamente.

O processo de extração de DNA, quantificação e diluição das amostras da população de melhoramento de *E. urophylla* consumiu um tempo semelhante ao da análise da população de *E. grandis*, ou seja, aproximadamente cinco dias. O número de reações de PCR que foram preparadas foi reduzido a 1/3, pelo fato de três locos microssatélites terem sido co-amplificados em mesma reação e, conseqüentemente, apenas um dia foi consumido nessa etapa. A capacidade de análise do seqüenciador de DNA ("Perkin-Elmer/Applied Biosystems", modelo 377) disponível no laboratório onde foi desenvolvido o presente trabalho era de 64 amostras por gel de eletroforese. Considerando-se que o tempo de preparo das amostras, aplicação e separação dos fragmentos foi de aproximadamente duas horas e meia por gel, três a quatro géis podem ser analisados por dia. Como foram utilizados os sistemas de genotipagem multilocos que amplificaram produtos de tamanhos distintos, seis locos EMBRA de cada amostra foram analisados em cada linha do gel. A definição dos genótipos foi praticamente imediata, através do uso de programas específicos que recolhiam os dados e filtravam a informação. Dessa forma, o processo de separação e análise dos fragmentos amplificados da população de 225 indivíduos de *E. urophylla* consumiu um a dois dias. No total foram 9 a 10 dias consumidos na análise dessa população.

Como foi observado, o sistema de genotipagem semi-automatizado utilizando os sistemas desenvolvidos anteriormente foi muito mais eficiente do que o sistema de genotipagem manual. Embora o custo de análise semi-automatizada seja superior ao da análise da forma tradicional, houve grande

economia no número de reações de PCR preparadas. Paralelamente, ocorreu redução de custos no preparo de géis de poliacrilamida, tendo em vista que o número de locos que poderiam ser analisados em cada linha do gel era triplicado e, conseqüentemente, o número de géis a serem preparados era reduzido a 1/3. Seguindo essa mesma linha, a economia de tempo também foi significativa, considerando-se que a análise de 225 clones de *E. urophylla* tomou menos de 2/3 do tempo necessário para a análise de 192 clones de *E. grandis*.

4. RESUMO E CONCLUSÕES

Duas populações de melhoramento da empresa Aracruz Celulose S. A. foram avaliadas com marcadores microsatélites, com os objetivos de estabelecer genótipos únicos para cada clone e determinar a relação de similaridade genética entre eles.

A população de *E. urophylla* apresentou vários indivíduos com perfil alélico igual ou idêntico na quase totalidade dos locos, demonstrando a existência de clones duplicados, inclusive entre aqueles plantados comercialmente pela empresa. A análise de similaridade indicou um agrupamento dos clones que eram empregados em plantas comerciais na empresa, indicando uma possível relação mais estreita entre esses indivíduos com relação ao restante da população.

Todos os locos caracterizados nas duas populações apresentaram marcadores adequados para o uso na genotipagem de indivíduos da população de *E. urophylla*.

Finalmente, foi observado o extraordinário aumento na eficiência do processo de genotipagem utilizando o sistema semi-automatizado, empregado na população de *E. urophylla*, em relação ao processo de registro manual dos perfis alélicos, feito na população de *E. grandis*. Além do aumento na rapidez da geração de dados para a análise de seis locos EMBA foi necessário apenas 1/3 das reações que seriam feitas no sistema tradicional de gel, com a amplificação de um loco por reação. Além disso, o número de linhas do gel

4. RESUMO E CONCLUSÕES

Duas populações de melhoramento da empresa Aracruz Celulose S. A. foram avaliadas com marcadores microssatélites, com os objetivos de estabelecer genótipos multilocos únicos para cada clone e determinar a relação de similaridade genética entre eles.

A população de *E. urophylla* apresentou vários indivíduos com perfil alélico igual ou idêntico na quase totalidade dos locos, demonstrando a existência de clones duplicados, inclusive entre aqueles plantados comercialmente pela empresa. A análise de similaridade indicou um agrupamento dos clones que eram empregados em plantios comerciais pela empresa, indicando uma possível relação mais estreita entre esses indivíduos com relação ao restante da população.

Todos os locos caracterizados nas duas populações apresentaram amplo conteúdo de informação, mostrando-se, dessa forma, adequados para o uso na genotipagem de indivíduos das populações de *E. urophylla*.

Finalmente, foi observado o extraordinário aumento na eficiência do processo de genotipagem utilizando o sistema semi-automatizado, empregado na população de *E. urophylla*, em relação ao processo de registro manual dos perfis alélicos, feito na população de *E. grandis*. Além do aumento na rapidez da geração de dados, para a análise de seis locos EMBRA foi necessário apenas 1/3 das reações que seriam feitas no sistema tradicional, ou seja, com a amplificação de um loco por reação. Além disso, o número de linhas de gel

necessárias para a separação dos fragmentos ficou muito reduzido, tendo em vista que até seis locos poderiam ser separados simultaneamente, devido ao uso de fluorocromos com diferentes cores e locos que amplificam fragmentos de tamanhos distintos.

Os resultados encontrados justificaram claramente o uso de sistemas de genotipagem baseados em marcadores microssatélites na avaliação de populações de melhoramento de *Eucalyptus*. A identificação de clones duplicados será de suma importância para a empresa, redirecionando-se cruzamentos anteriormente considerados, assim como o estabelecimento de novos plantios com o emprego desses clones idênticos. As informações precisas do genótipo de cada indivíduo permitirão ainda, aos melhoristas, determinar a origem dos clones-elite, estabelecendo-se os seus "pedigrees".

No presente trabalho foram desenvolvidos três sistemas de genotipagem multilocus, baseados na amplificação de locos microssatélites, através de iniciadores marcados com fluorocromos, e na detecção por eletroforese, para análise de espécies do gênero *Eucalyptus*. Para isso, primeiro, inicialmente, a seleção de um conjunto de microssatélites a partir de um conjunto de 170 locos, considerando não apenas a qualidade da amplificação, mas também a sua transferibilidade para outras espécies do gênero *Eucalyptus*, observaram-se mais de 70% dos locos amplificados produzindo resultados interpretáveis. Grande parte desses locos (81%) estava conservada em *E. dunnii*, espécie do subgênero *Symphycarphus*, ao qual pertencem também as espécies a partir das quais os microssatélites foram desenvolvidos, ou seja *E. grandis* e *E. amplifolia*. A conservação foi menor entre espécies de outros subgêneros. Um conjunto de 17 locos foi utilizado no desenvolvimento de sistemas Multiplex envolvendo a co-amplificação de vários locos na mesma reação de PCR, com iniciadores marcados com diferentes fluorocromos. Foram obtidos, assim, três sistemas Multiplex que tiveram suas reações de PCR otimizadas de forma a obter amplificação adequada em todos os locos, nas espécies para as quais eles foram produzidos. Contribuição para a LITERATURA: THOMAS e GREGAN, 1997; e vide THOMAS et al., 1994.

3. RESUMO E CONCLUSÕES

No presente trabalho foram desenvolvidos três sistemas de genotipagem multilocos, baseados na amplificação de locos microssatélites, através de iniciadores marcados com fluorocromos, e na detecção semi-automatizada, para análise de espécies do gênero *Eucalyptus*. Para isso, procedeu-se, inicialmente, à seleção de um conjunto de microssatélites a partir de um conjunto de 170 locos, considerando não apenas a qualidade de amplificação, mas também a sua transferibilidade para outras espécies do gênero *Eucalyptus*; observaram-se mais de 70% dos locos amplificando produtos interpretáveis. Grande parte desses locos (81%) estava conservada em *E. dunnii*, espécie do subgênero *Symphyomyrtus*, ao qual pertencem também as espécies a partir das quais os microssatélites foram desenvolvidos, ou seja, *E. grandis* e *E. urophylla*. A conservação foi menor entre espécies de outros subgêneros. Um conjunto de 17 locos foi utilizado no desenvolvimento de sistemas Multiplex envolvendo a co-amplificação de vários locos na mesma reação de PCR, com iniciadores marcados com diferentes fluorocromos. Foram obtidos, assim, três sistemas Triplex, que tiveram suas reações de PCR otimizadas de forma a obter amplificação adequada em todos os locos, nas espécies para as quais eles foram produzidos, assemelhados aos desenvolvidos para diversas culturas agrícolas, como brássicas (MITCHELL et al., 1997), soja (DIWAN e CREGAN, 1997) e videira (THOMAS et al., 1994).

Foi investigada a influência do tamanho de amostra na estimativa do parâmetro heterozigidade esperada em uma população de melhoramento de *E. grandis*. Observou-se que, em locos com grande número de alelos, amostra de 64 indivíduos é adequada para estimativas de heterozigidade esperada com alto nível de acurácia. Isso ocorre somente após a análise de um número superior de indivíduos em locos com menor número de alelos. A tendência da estimativa desse parâmetro foi sempre negativa, indicando que estimativas de heterozigidade esperada tendem a ser sempre subestimadas em relação ao valor paramétrico da população. Considerando-se que a natureza molecular dos microssatélites faz com que estes tenham comportamento semelhante em termos de distribuição de freqüências alélicas entre organismos, o presente estudo tem ampla relevância para locos microssatélites baseados em dinucleotídeos em eucariotos. A tendência exclusivamente negativa das estimativas indica que o valor paramétrico da população assume sempre valores superiores aos estimados.

Os locos EMBRA que compõem os sistemas Multiplex foram caracterizados em seis espécies do gênero *Eucalyptus*. Inicialmente, indivíduos provenientes de cinco procedências de *E. grandis* foram genotipados e os índices H_e , PIC, I e Q, referentes aos locos EMBRA, calculados. Todos os parâmetros estimados indicaram que os locos avaliados não possuíam alto conteúdo de informação. Análise de variância molecular realizada com base nas freqüências alélicas geradas para as cinco procedências de *E. grandis* indicou que a variância seguiu o padrão normalmente observado em espécies alógamas, com mais de 90% de variância sendo encontrada dentro das procedências e, menos de 10%, entre elas.

Os locos EMBRA caracterizados nas outras cinco espécies – *E. camaldulensis*, *E. dunnii*, *E. globulus*, *E. saligna* e *E.* – apresentaram o mesmo padrão de distribuição multimodal de freqüências alélicas que haviam sido observados em *E. grandis*. Os locos normalmente apresentaram elevado conteúdo de informação. Entretanto, em alguns casos, o número de alelos detectados foi expressivamente reduzido, ocorrendo até mesmo um alelo praticamente fixado no loco EMBRA27, em *E. globulus*.

Assim, foi gerado o primeiro banco de dados de freqüências alélicas de locos microssatélites para espécies do gênero *Eucalyptus*, esperando-se, no

futuro, incremento desses dados através da análise de maior número de indivíduos das espécies descritas, além da adição de novas espécies ao conjunto estudado. A padronização dos locos analisados nessas espécies possibilitará o uso universal deste sistema de genotipagem, facilitando todos os procedimentos de troca e identificação de material genético nas empresas e nos institutos de pesquisa, da mesma forma como foi proposto para outras espécies (DIWAN e CREGAN, 1997).

Para avaliar a aplicabilidade dos sistemas de genotipagem desenvolvidos, duas populações de melhoramento da empresa Aracruz Celulose S. A. foram avaliadas. Com os dados gerados, foram possíveis estabelecer um genótipo multiloco único para cada clone e determinar a relação de similaridade genética entre eles. Observou-se que uma população de *E. urophylla* apresentou vários indivíduos com perfil alélico igual ou idêntico na quase totalidade dos locos, demonstrando a existência de clones duplicados. A análise de similaridade indicou agrupamento dos clones que são empregados em plantios comerciais pela empresa, evidenciando uma possível ligação mais estreita entre esses indivíduos em relação ao restante da população.

Finalmente, foi observado extraordinário aumento na eficiência do processo de genotipagem com a utilização do sistema semi-automatizado, empregado na população de *E. urophylla*, em relação ao processo de registro manual dos perfis alélicos, feito na população de *E. grandis*. Além do aumento na rapidez da geração de dados, para a análise de seis locos EMBRA foi necessário apenas 1/3 das reações que seriam feitas no sistema tradicional, ou seja, com a amplificação de um loco por reação. Além disso, o número de linhas de gel necessárias para a separação dos fragmentos ficou muito reduzida, tendo em vista que até seis locos podem ser separados simultaneamente, devido ao uso de fluorocromos com diferentes cores e locos que amplificam fragmentos de tamanhos distintos.

Os resultados encontrados justificaram claramente o uso de sistemas de genotipagem baseados em marcadores microssatélites na avaliação de populações de melhoramento de *Eucalyptus*. A identificação de clones duplicados será de suma importância para a empresa, que poderá redirecionar cruzamentos anteriormente considerados, assim como estabelecer novos

plantios, mediante o emprego de clones idênticos. Além disso, as informações precisas do genótipo de cada indivíduo ainda permitirão aos melhoristas determinar a origem dos clones-elite, estabelecendo-se os seus "pedigrees".

REFERÊNCIAS BIBLIOGRÁFICAS

- ALFENAS, A.C. (Ed.) *Eletroforeses de fosforimas e proteínas vivas*. Viçosa: UFV, 1995. 574p.
- BASSAM, B.J., CAETANO-ANOLLES, G., GRESSHOFF, P.M. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal. Biochem.*, v. 196, p.80-82, 1991.
- BECKER, J., HEUM, M. Barley microsatellites: allele variation and mapping. *Plant Mol. Biol.*, v.27, p.835-845, 1995.
- BUTSTEIN, D., WHITE, R.L., SKOLNICK, M.H., DAVIES, R.W. Construction of a genetic map in human using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, v.32, p.314-331, 1980.
- BOWERS, J.E., DANOL, G.S., VIGNANI, P., MEREDITH, C.P. Isolation and characterization of new polymorphic simple sequence-repeat loci in grape (*Vitis vinifera* L.). *Genome*, v.39, p.553-555, 1996.
- BREEN, M., DOVNS, P., IRWIN, J., DELL, K. Intraspecific amplification of horse microsatellite markers with emphasis on the Przewalski's horse (*E. przewalskii*). *Anim. Genet.*, v.23, p.401-405, 1994.
- BRONDANI, R.P.V., BRONDANI, C., TARCHINI, R., GRATTAPAGLIA, J. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. woodylia*. *Theor. Appl. Genet.*, v.97, p.516-527, 1998.
- BYRNE, M., MARQUEZ-DARCA, M.I., UREN, T., SMITH, D.S., MORAN, G.F. Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus*. *Aust. J. Bot.*, v.44, p.331-341, 1996.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALFENAS, A.C. (Ed.) **Eletoforese de isoenzimas e proteínas afins**. Viçosa: UFV, 1998. 574p.
- BASSAM, B.J., CAETANO-ANOLLÉS, G., GRESSHOFF, P.M. Fast and sensitive silver staining of DNA in polyacrilamide gels. **Anal. Biochem.**, v.196, p.80-83, 1991.
- BECKER, J., HEUN, M. Barley microsatellites: allele variation and mapping. **Plant Mol. Biol.**, v.27, p.835-845, 1995.
- BOTSTEIN, D., WHITE, R.L., SKOLNICK, M.H., DAVIES, R.W. Construction of a genetic map in human using restriction fragment length polymorphisms. **Am. J. Hum. Genet.**, v.32, p.314-331, 1980.
- BOWERS, J.E., DANGL, G.S., VIGNANI, R., MEREDITH, C.P. Isolation and characterization of new polymorphic simple sequence repeat loci in grape (*Vitis vinifera* L.). **Genome**, v.39, p.658-633, 1996.
- BREEN, M., DOWNS, P., IRVIN, Z., BELL, K. Intrageneric amplification of horse microsatellite markers with emphasis on the Przewalski's horse (*E. przewalskii*). **Anim. Genet.**, v.25, p.401-405, 1994.
- BRONDANI, R.P.V., BRONDANI, C., TARCHINI, R., GRATTAPAGLIA, D. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. **Theor. Appl. Genet.**, v.97, p.816-827, 1998.
- BYRNE, M., MARQUEZ-GARCIA, M.I., UREN, T., SMITH, D.S., MORAN, G.F. Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus*. **Aust. J. Bot.**, v.44, p.331-341, 1996.

- CARR, S.G.M., CARR, D.J. The taxonomic position of certain eucalypts. **Proc. Linn. Soc. Victoria**, v.77, p.207-216, 1963.
- CHASE, M., KESSELI, R., BAWA, K. Microsatellite markers for population and conservation genetics of tropical trees. **Am. J. Bot.**, v.83, p.51-57, 1996.
- CLARK, J.M. Novel non-template nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. **Nucleic Acids Res.**, v.16, p.9677-9686, 1988.
- COSTA E SILVA, C., GRATTAPAGLIA, D. RAPD relatedness of elite *Eucalyptus* clones: application in breeding, mapping and forest plantation. In: IUFRO CONGRESS ON SILVICULTURE AND IMPROVEMENT OF EUCALYPTS, 1997, Salvador, BA. **Proceedings...**, Colombo:EMBRAPA-CNPQ, 1997. v.2, p.161-166.
- CREGAN, P.B., BHAGWAT, A.A., AKKAYA, M.S., RONGWEN, J. Microsatellite fingerprinting and mapping of soybean. **Meth. Mol. Cell. Biol.**, v.5, p.49-61, 1994.
- DIAS, L.A.S. Análises multidimensionais. In: ALFENAS, A.C. (Ed.) **Eletroforese de isoenzimas e proteínas afins**. Viçosa: UFV, 1998. p.405-475.
- DIWAN, N., CREGAN, P.B. Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. **Theor. Appl. Genet.**, v.95, p.723-733, 1997.
- DOYLE, J.J., DOYLE, J.L. Isolation of plant DNA from fresh tissue. **Focus**, v.12, p.13-15, 1987.
- EDWARDS, A., CIVITELLO, A., HAMMOND, H.A., CASKEY, C.T. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. **Am. J. Hum. Genet.**, v.49, p.746-756, 1991.
- EFRON, B., TIBSHIRANI, R.J. **An introduction to the bootstrap**. New York: Chapman e Hall, 1993. 436p.
- ELDRIDGE, K., DAVIDSON, J., HARWOOD, C., VAN WYK, G. **Eucalypt domestication and breeding**. Oxford: Oxford University Press, 1994. 288p.
- EXCOFFIER, L., SMOUSE, P.E., QUATTRO, J.M. Analysis of molecular variance inferred from metric distance among DNA haplotypes: application to human mitochondrial DNA restriction sites. **Genetics**, v.131, p.479-491, 1992.
- FAO. **Eucalypts for planting**. Rome: FAO, 1979.

- FERREIRA, M., SANTOS, P.E.T. Melhoramento genético florestal dos *Eucalyptus* no Brasil - Breve histórico e perspectivas. In: IUFRO CONGRESS ON SILVICULTURE AND IMPROVEMENT OF EUCALYPTS, 1997, Salvador, BA. **Proceedings...**, Colombo: EMBRAPA-CNPQ, 1997. v.1, p.14-34.
- GAIOTTO, F.A., BRAMUCCI, M., GRATTAPAGLIA, D. Estimation of outcrossing rate in a breeding population of *Eucalyptus urophylla* with dominant RAPD and AFLP markers. **Theor. Appl. Genet.**, v.95, p.842-849, 1997.
- GLOWATZKI-MULLIS, M.L., GAILLARD, C., WIGGER, G., FRIES, R. Microsatellite-based parentage control in cattle. **Anim. Genet.**, v.26, p.7-12, 1995.
- GRATTAPAGLIA, D., SEDEROFF, R.R. Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers. **Genetics**, v.144, p.1205-1214, 1996.
- GRIFFIN, A.R., COTTERIL, P.P. Genetic variation in growth of outcrossed, selfed and open-pollinated progenies of *Eucalyptus regnans* and some implications for breeding strategy. **Silvae Genet.**, v.37, p.124-131, 1988.
- HAMADA, H., PETRINO, M.G., KAKUNAGA, T. A novel repeat element with Z-DNA-forming potential is widely found in evolutionary diverse eucariotic genomes. **Proc. Natl. Acad. Sci. USA**, v.79, p.6465-6469, 1982.
- HARTL, D.L., CLARK, A.G. **Principles of population genetics**. Sunderland: Sinauer Associates, 1997. 542p.
- HOCHMEISTER, M.N., BUDOWLE, B., BORER, U.V., RUDIN, O., BOHNERT, M., DIRNHOFER, R. Confirmation of the identity of human skeletal remains using multiplex PCR amplification and typing kits. **J. Forensic Sci.**, v.40, n.4, p.701-705, 1995.
- HUBBY, J.L., LEWONTIN, R.C. A molecular approach to the study of genic heterozygosity in natural populations, I. The number of alleles at different loci in *Drosophila pseudoobscura*. **Genetics**, v.54, p.577-594, 1966.
- JOHNSON, F.M., KANAPI, C.G., RICHARDSON, R.H., WHEELER, M.R., STONE, W.S. An analysis of polymorphisms among isozyme loci in dark and light *Drosophyla ananassae* strains from America and Western Samoa. **Proc. Natl. Acad. Sci. USA**, v.56, p.119-125, 1966.
- JOHNSON, L.A.S. Problems of species and genera in *Eucalyptus* (Myrtaceae). **Plant Syst. Evol.**, v.125, p.155-167, 1976.

- JUNGHANS, D.T., ALFENAS, A.C., BROMMONSCHENKEL, S.H., VALLE, L.A.C., CHEAVEAGATTI, A., ODA, S., MELLO, E.J. Inheritance of the rust (*Puccinia psidii*) resistance in *Eucalyptus grandis*. **Genet. Mol. Biol.**, v.21, n.3, p.185, 1998. Supplement.
- KEIL, M., GRIFFIN, A.R. Use of random amplified polymorphic DNA (RAPD) markers in the discrimination and verification of genotypes in *Eucalyptus*. **Theor. Appl. Genet.**, v.89, p.442-450, 1994.
- KEMP, S.J., BREZINSKY, L., TEALE, A.J. A panel of bovine, ovine and caprine polymorphic microsatellites. **Anim. Genet.**, v.24, p.363-365, 1993.
- KIJAS, J.M.H., FOWLER, J.C.S., THOMAS, M.R. An evaluation of sequence tagged microsatellite site markers for genetic analysis within *Citrus* and related species. **Genome**, v.38, p.349-355, 1995.
- KOSTIA, S., VARVIO, S., VAKKARI, P., PULKKINEN, P. Microsatellite sequences in a conifer, *Pinus sylvestris*. **Genome**, v.38, p.1244-1248, 1995.
- KRESOVICH, S., SZEWC-MCFADDEN, A.K., BLIEK, S.M., MCFERSON, J.R. Abundance and characterization of simple-sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. (rapeseed). **Theor. Appl. Genet.**, v.91, p.206-211, 1995.
- LADIGES, P.Y. A cladistic study of *Arillastrum*, *Angophora* and *Eucalyptus* (Myrtaceae). **Bot. J. Linn. Soc.**, v.87, p.105-134, 1983.
- LADIGES, P.Y., UDOVICIC, F., DRINNAN, A.N. Eucalypt phylogeny-molecules and morphology. **Aust. Syst. Bot.**, v.8, p.483-497, 1995.
- LAGERCRANTZ, U., ELLEGREN, H., ANDERSSON, L. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. **Nucleic Acids Res.**, v.21, p.1111-1115, 1993.
- LEVIN, I., CHENG, H.H., BAXTER-JONES, C., HILLEL, J. Turkey microsatellite DNA loci amplified by chicken-specific primers. **Anim. Genet.**, v.26, p.107-110, 1995.
- LEVITT, R.C., KISER, M.B., DRAGWA, C., JEDLICKA, A.E., XU, J., MEYERS, D.A., HUDSON, J.R. Fluorescence-based resource for semiautomated genomic analysis using microsatellite markers. **Genomics**, v.24, p.361-365, 1994.
- LITT, M., LUTY, J.A. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. **Am. J. Hum. Genet.**, v.44, p.397-401, 1989.

- MAGUIRE, T.L., SEDGLEY, M. Genetic diversity in *Banksia* and *Dryandra* (Proteaceae) with emphasis on *Banksia cuneata*, a rare and endangered specie. **Heredity**, v.79, p.394-401, 1997.
- MANSFIELD, D.C., BROWN, A.F., GREEN, D.K., CAROTHERS, A.D., MORRIS, S.W., EVANS, H.J., WRIGHT, A.F. Automation of genetic linkage analysis using fluorescent microsatellite markers. **Genomics**, v.24, p.225-233, 1994.
- MARKLUND, S., ELLEGREN, H., ERIKSSON, S., SANDBERG, K., ANDERSSON, L. Parentage testing and linkage analysis in the horse using a set of highly polymorphic microsatellites. **Anim. Genet.**, v.25, p.19-23, 1994.
- 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*. **Theor. Appl. Genet.**, v.96, p.727-737, 1996.
- MARTINS-CORDER, M.P., LOPES, C.R. Isozyme characterization of *Eucalyptus urophylla* (S.T. Blake) and *E. grandis* (Hill ex Maiden) populations in Brazil. In: IUFRO CONGRESS ON SILVICULTURE AND IMPROVEMENT OF EUCALYPTS, 1997, Salvador, BA. **Proceedings...**, Colombo: EMBRAPA-CNPQ, 1997. v.2, p.96-101.
- MENOTTI-RAYMOND, M.A., O'BRIEN, S.J. Evolutionary conservation of ten microsatellite loci in four species of Felidae. **J. Heredity**, v. 86, n. 4, p.319-322. 1995.
- MICHELL, S.E., KRESOVICH, S., JESTER, C.A., HERNANDEZ, C.J., SZEWC-MCFADDEN, A.K. Application of multiplex PCR and fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. **Crop Sci.**, v.37, p.617-624, 1997.
- MORAN, G.F., BELL, J.C. *Eucalyptus*. In: TANKSLEY, S.D., ORTON, T.J. (Eds.). **Isozymes in plant genetics and breeding. Part B**. Amsterdam: Elsevier Science Publishers, 1983. p.423-441.
- MORGANTE, M., RAFALSKI, A., BIDDLE, P., TINGEY, S., OLIVIERI, A.M. Genetic mapping and variability of seven soybean simple sequence repeat loci. **Genome**, v.37, p.763-769, 1994.
- MULLIS, K.B., FALOONA, F.A. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. **Methods Enzimol.**, v.155, p.335-350. 1987.
- NESBITT, K.A., POTTS, B.M., VAILLANCOURT, R.E., WEST, A.K., REID, J.B. Partitioning and distribution of RAPD variation in a forest tree species, *Eucalyptus globulus* (Myrtaceae). **Heredity**, v.74, p.628-637, 1995.

- OLSEN, J.B., WENBURG, J.K., BENTZEN, P. Semiautomated multilocus genotyping of Pacific salmon (*Oncorhynchus* spp.) using microsatellites. **Mol. Marine Biol. Biotech.**, v.5, p.259-272, 1996.
- PAETKAU, D., CALVERT, W., STIRLING, I., STROBECK, C. Microsatellite analysis of population structure in Canadian polar bears. **Mol. Ecol.**, v.4, p.347-354, 1995.
- PÉPIN, L., AMIGUES, Y., LÉPINGLE, A., BERTHIER, J., BENSALD, A., VAIMAN, D. Sequence conservation of microsatellites between *Bos taurus* (cattle), *Capra hircus* (goat) and related species. Examples of use in parentage testing and phylogeny analysis. **Heredity**, v.74, p.53-61, 1995.
- POWELL, W., MORGANTE, M., ANDRE, C., HANAFEY, M., VOGEL, J., TINGEY, S., RAFALSKI, A. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. **Mol. Breeding**, v.2, p.225-238, 1996.
- PRYOR, L.D., JOHNSON, L.A.S. **A classification of the Eucalypts**. Canberra: Australian National University, 1971. 102p.
- RAFALSKI, J.A., TINGEY, S.V. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. **Trends Genet.**, v.9, p.275-280, 1993.
- RAFALSKI, J.A., VOGEL, J.M., MORGANTE, M., POWELL, W., ANDRE, C., TINGEY, S.V. Generating and using DNA markers in plants. In: BIRREN, B., LAI, E. (Eds.) **Nonmammalian genomic analysis**. San Diego: Academic Press, 1996. p.75-134.
- RAYMOND, M., ROUSSET, F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. **J. Heredity**, v.86, p.248-249, 1995a.
- RAYMOND, M., ROUSSET, F. An exact test for population differentiation. **Evolution**, v.49, p.1280-1283, 1995b.
- RIBEIRO, V.J., BERTOLUCCI, F.L., GRATTAPAGLIA, D. RAPD marker - guide matings in a reciprocal recurrent selection program of Eucalyptus. In: IUFRO CONGRESS ON SILVICULTURE AND IMPROVEMENT OF EUCALYPTS, 1997, Salvador, BA. **Proceedings...**, Colombo: EMBRAPA-CNPQ, 1997. v.2, p.156-160.
- ROBINSON, I.P. Aloenzimas na genética de populações de plantas. In: ALFENAS, A.C. (Ed.). **Eletroforese de isoenzimas e proteínas afins**. Viçosa: UFV, 1998. p.329-480.
- ROSSETO, M., WEAVER, P.K., DIXON, K.W. Use of RAPD analysis in devising conservation strategies for the rare and endangered *Grevillea scapigera*. **Mol. Ecol.**, v.4, p.321-329, 1995.

- RUSSELL, J., FULLER, J., YOUNG, G., THOMAS, B., TARMINO, G., MACAULAY, M., WAUGH, R., POWELL, W. Discriminating between barley genotypes using microsatellite markers. **Genome**, v.40, p.442-450, 1997.
- SAGHAI-MAROOF, M.A., BIYASHEV, R.M., YANG, G.P., ZHANG, Q., ALLARD, R.W. Extraordinary polymorphic microsatellite DNA in barley: species diversity, chromosomal localtions and population dynamics. **Proc. Natl. Acad. Sci. USA**, v.91, p.5466-5470, 1994.
- SALE, M.M., POTTS, B.M., WEST, A.K., REID, J.B. Relationships within *Eucalyptus* (Myrtaceae) using PCR-amplification and Southern hybridisation of cloroplast DNA. **Aust. Syst. Bot.**, v.9, p.273-282, 1996.
- SCHLÖTTERER, C., TAUTZ, D. Slippage synthesis of simple sequence DNA. **Nucleic Acids Res.**, v.20, p.211-215, 1992.
- SCHNEIDER, S., KUEFFER, J.M., ROESSLI, D., EXCOFFIER, L. **Arlequin. ver.1.1: A software for population genetic data analysis**. Genève: Genetics and Biometrics Laboratory, Univerity of Geneva, Switzerland, 1997. 76p.
- SHEELY, D.L., MEAGHER, T.R. Genetic diversity in micronesian island populations of the tropical tree *Camposperma brevipetiolata* (Anacardiaceae). **Am. J. Bot.**, v.83, p.1571-1579, 1996.
- SMITH, D.N., DEVEY, M.E. Occurance and inheritance of microsatellites in *Pinus radiata*. **Genome**, v.37, p.977-983, 1994.
- SPRECHER, C.J., PUERS, C., LINS, A.M., SCHUMM, J.W. General approach to analysis of polymorphic short tandem repeat loci. **BioTechniques**, v.20, p.266-276, 1996.
- SQUILLASSI, M.G., GRATTAPAGLIA, D. A preliminary assessment of the efficiency of marker assisted selection and phenotypic selection in *Eucalyptus*. In: IUFRO CONGRESS ON SILVICULTURE AND IMPROVEMENT OF EUCALYPTS, 1997, Salvador, BA. **Proceedings....**, Colombo: EMBRAPA-CNPf, 1997. v.2, p.14-18.
- STEPHAN, W., CHO, S. Possible role of natural selection in the formation of tandem-repetitive noncoding DNA. **Genetics**, v.136, p.333-341, 1994.
- SZEWC-MCFADDEN, A.K., KRESOVICH, S., BILEK, S.M., MITCHELL, S.E., MCFERSON, J.R. Identification of polymorphic, conserved simple sequence repeats (SSRs) in cultivated *Brassica* species. **Theor. Appl. Genet.**, v.93, p.534-538, 1996.
- THOMAS, M.R., CAIN, P., SCOTT, N.S. DNA typing of grapevines: A universal methodology and database for describing cultivars and evaluating genetic relatedness. **Plant Mol. Biol.**, v.25, p.939-949, 1994.

- VAIMAN, D., PAILHOUX, E., PAAYEN, E., SAÏDI-MEHTAR, N., COTINOT, C. Evolutionary conservation of a microsatellite in the Wilms Tumor (WT) gene: mapping in sheep and cattle. **Cytogenet. Cell Genet.**, v.70, p.112-115, 1995.
- VALDES, A.M., SLATKIN, M., FREIMER, N. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. **Genetics**, v.133, p.737-749, 1993.
- VERHAGEN, D., GION, J.M., PLOMION, C., BOUVET, J.M., POINTEL, M. Quantitative trait dissection analysis in *Eucalyptus grandis* using RAPD markers: QTL detection of wood density and stability of QTL expression across different ages. In: IUFRO CONGRESS ON SILVICULTURE AND IMPROVEMENT OF EUCALYPTS, 1997, Salvador, BA. **Proceedings...**, Colombo: EMBRAPA-CNPQ, 1997. v.2, p.149-155.
- VOGEL, J.M., RAFALSKI, A., POWELL, W., MORGANTE, M., ANDRE, C., HANAFEY, M., TINGEY, S.V. Application of genetic diagnostics to plant genome analysis and plant breeding. **HortScience**, v.31, p.1107-1108, 1996.
- VOS, P., HOGERS, R., BLEEKER, M., REIJANS, M., LEE, T.V., HORNES, M., FRIJTERS, A., POT, J., PELEMAN, J., KUIPER, M., ZABEAU, M. AFLP: a new technique for DAN fingerprinting. **Nucleic Acids Res.**, v.23, p.4407-4414, 1995.
- WANG, Z., WEBER, J.L., ZHONG, G., TANKSLEY, S.D. Survey of plant short tandem DNA repeats. **Theor. Appl. Genet.**, v.88, p.1-6, 1994.
- WEBER, J.L., MAY, P.E. Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction. **Am. J. Hum. Genet.**, v.44, p.388-396, 1989.
- WEIR, B.S. **Genetic data analysis II**. Sunderland: Sinauer Associates, 1996. p.445.
- WEISING, K., WEIGAND, F., DRIESEL, A.J., KAHL, G., ZISCHLER, H., EPPLER, J.T. Polymorphic simple GATA/GACA repeats in plant genomes. **Nucleic Acids Res.**, v.17, p.10128, 1989.
- WEISING, K., ATKINSON, R.G., GARDNER, R.C. Genomic fingerprinting by microsatellite primed PCR: a critical evaluation. **PCR Methods and Applications**, v.4, p.249-255, 1995.
- WENBURG, J.K., OLSEN, J.B., BENTZEN, P. Multiplexed systems of microsatellites for genetic analysis in coastal cutthroat trout (*Oncorhynchus clarki clarki*) and steelhead (*Oncorhynchus mykiss*). **Mol. Marine Biol. Biotech.**, v.5, p.273-283, 1996.

- WHITE, G., POWELL, W. Isolation and characterization of microsatellite loci in *Swietenia humilis* (Meliaceae): an endangered tropical hardwood species. **Mol. Ecol.**, v.6, p.851-860, 1997.
- WILLIAMS, J.G.K., KUBELIK, A.R., LIVAK, K.J., RAFALSKI, J.A., TINGEY, S.V. DNA polymorphisms amplified by arbitrary primers are usefull as genetic markers. **Nucleic Acids Res.**, v.18, p.6531-6535, 1990.
- WOLFF, J.L.C., MCDOWELL, S.W., MACHADO, M.A. Analysis of the restriction fragment length polymorfism of *Eucalyptus urophylla*, *E. grandis* and *E. citriodora*. **R. Bras. Genet.**, v.16, p.431-439, 1993.
- WU, K., TANKSLEY, S.D. Abundance, polymorphism and genetic mapping of microsatellites in rice. **Mol. Gen. Genet.**, v.241, p.225-235, 1993.

APÊNDICES

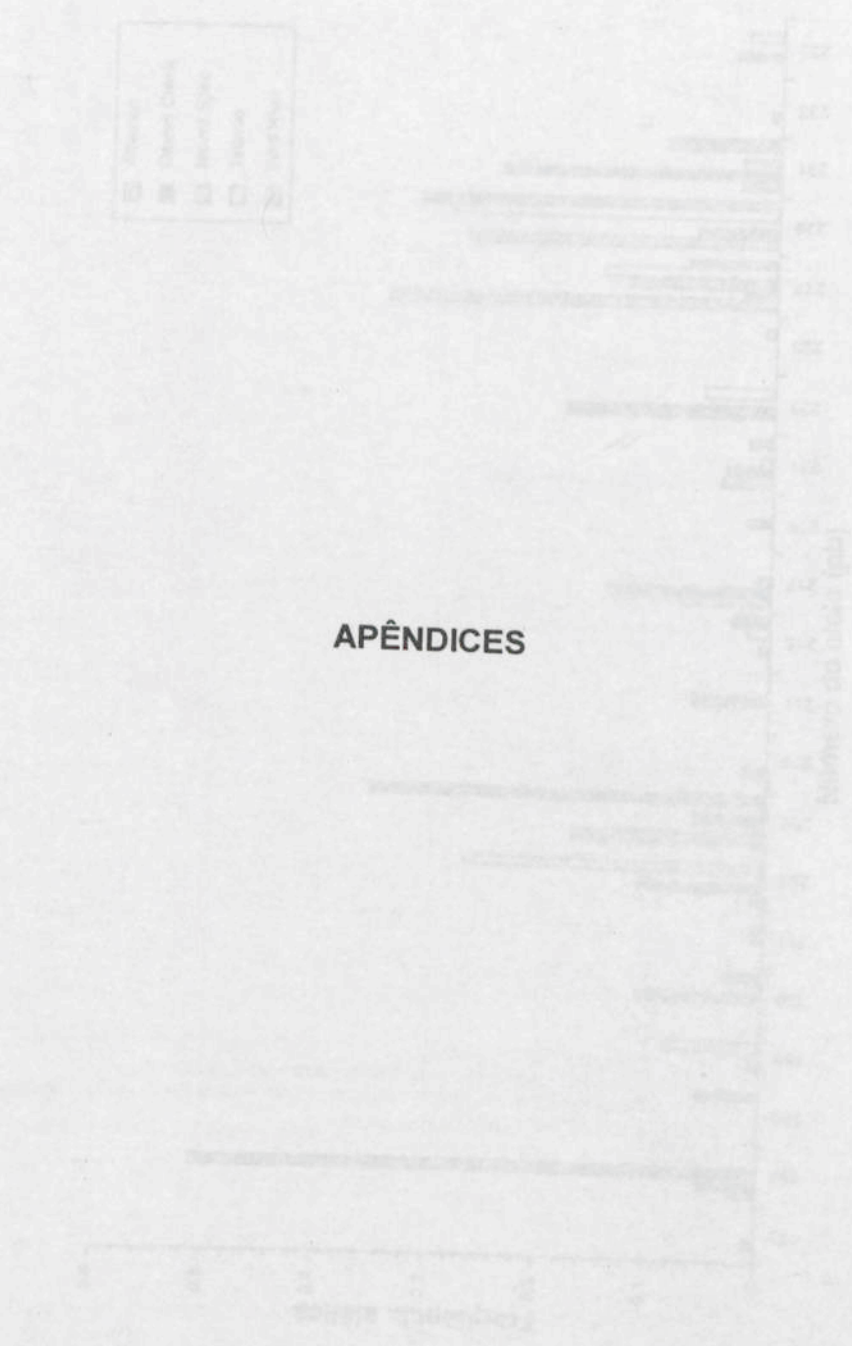


FIGURA 1A - Frequências absolutas de indivíduos no loco EYBERAZ7, gerados pela caracterização de 240 indivíduos provenientes de

APÊNDICE A

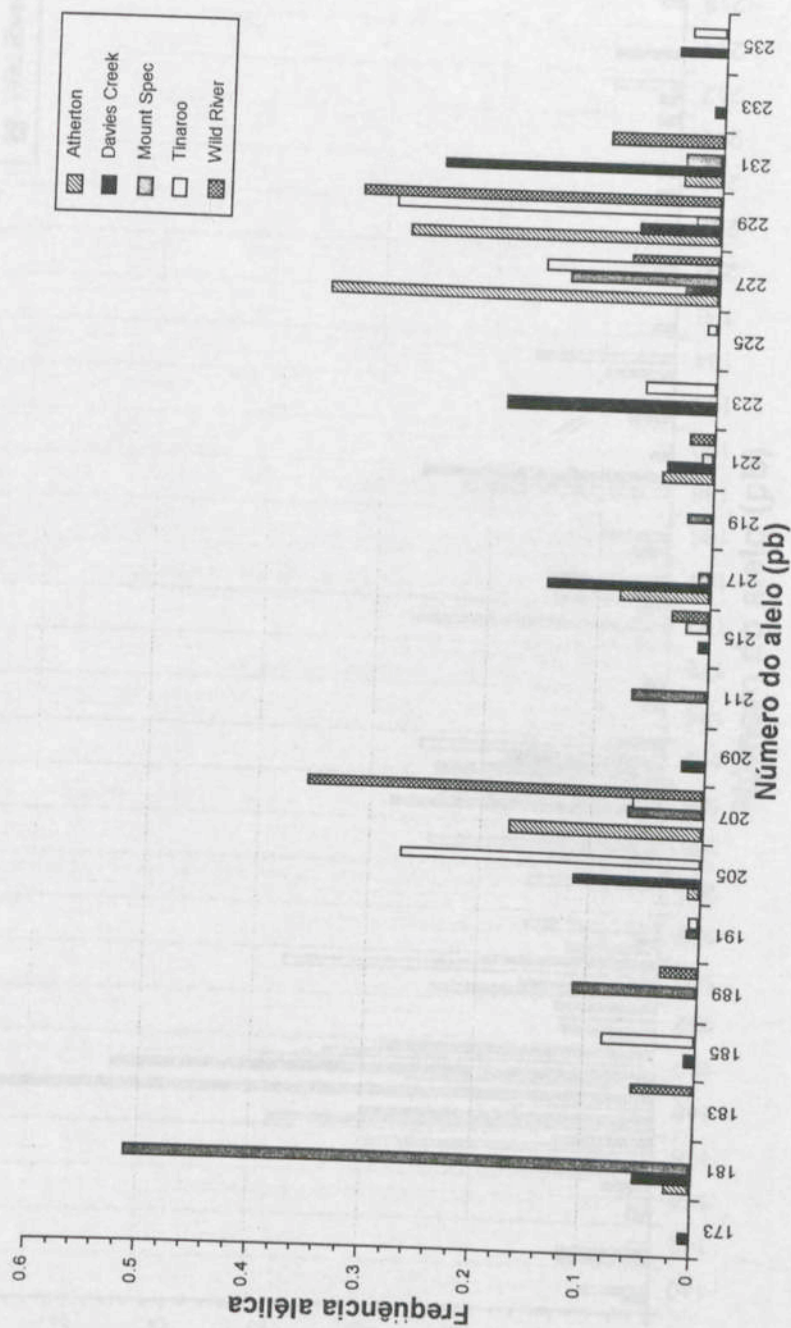


Figura 1A – Frequências alélicas observadas no loco EMBRA27, geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *Eucalyptus grandis*.

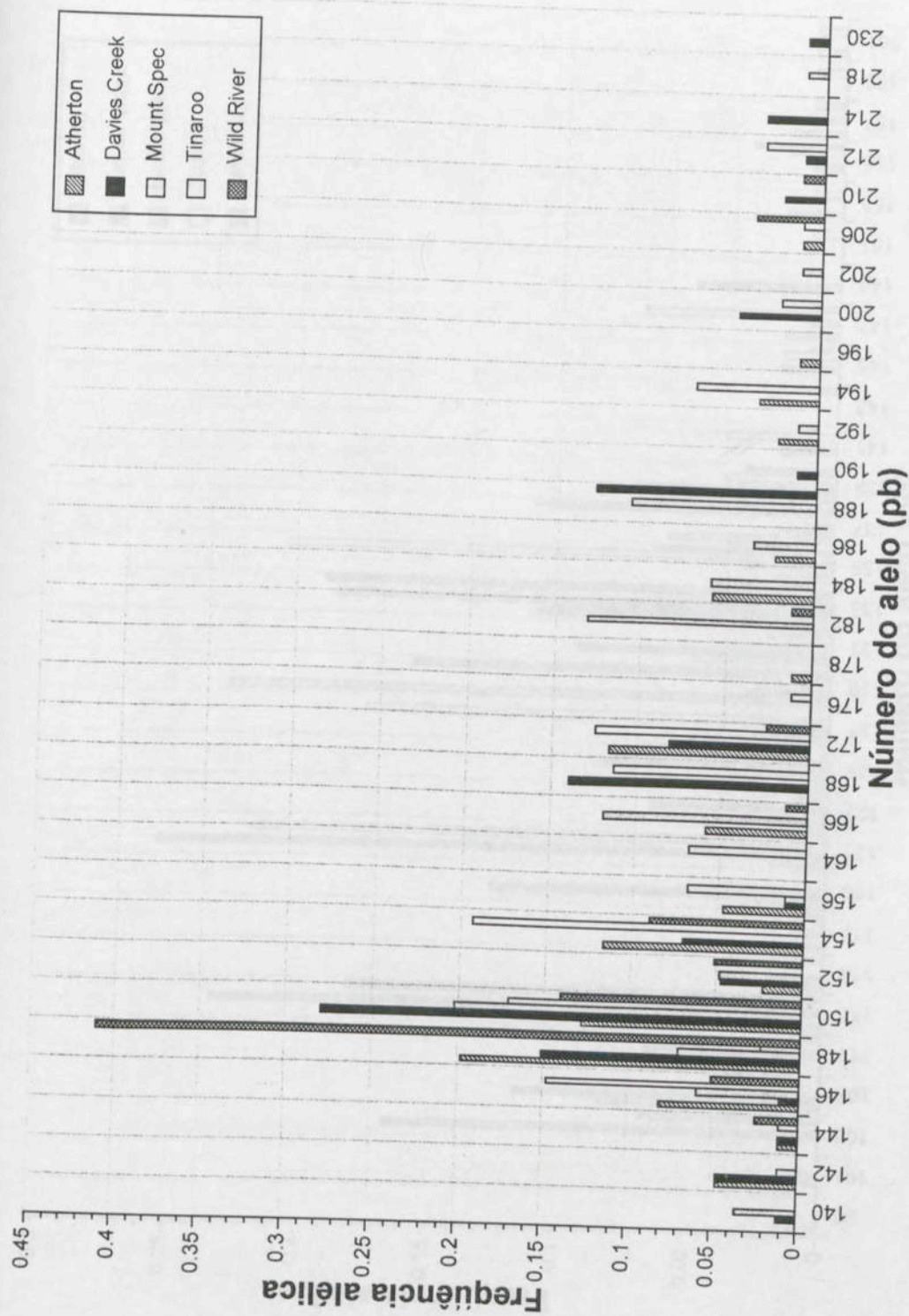


Figura 2A – Frequências alélicas observadas no loco EMBRA22, geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *Eucalyptus grandis*.

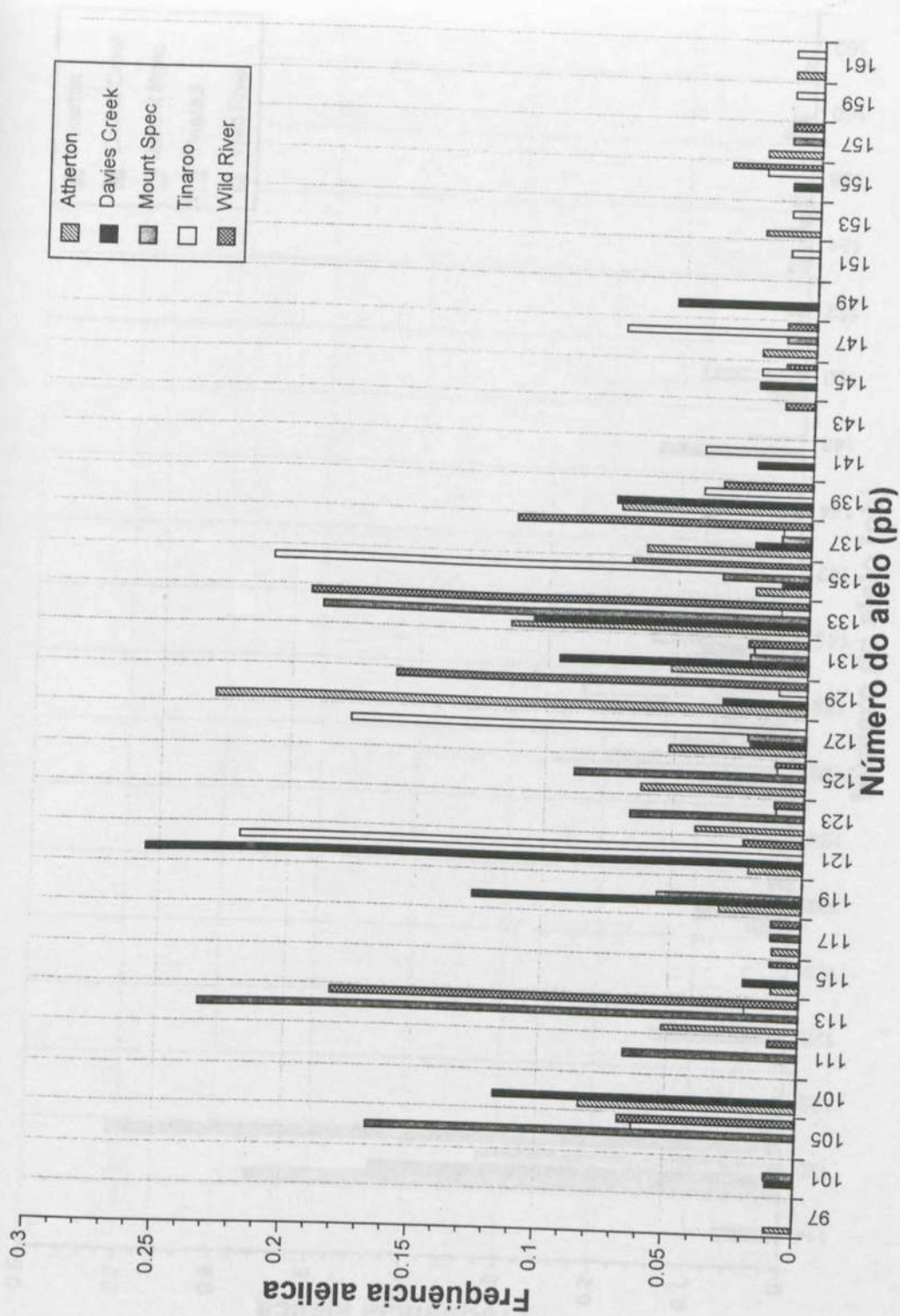


Figura 3A – Frequências alélicas observadas no loco EMBRA5, geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *Eucalyptus grandis*.

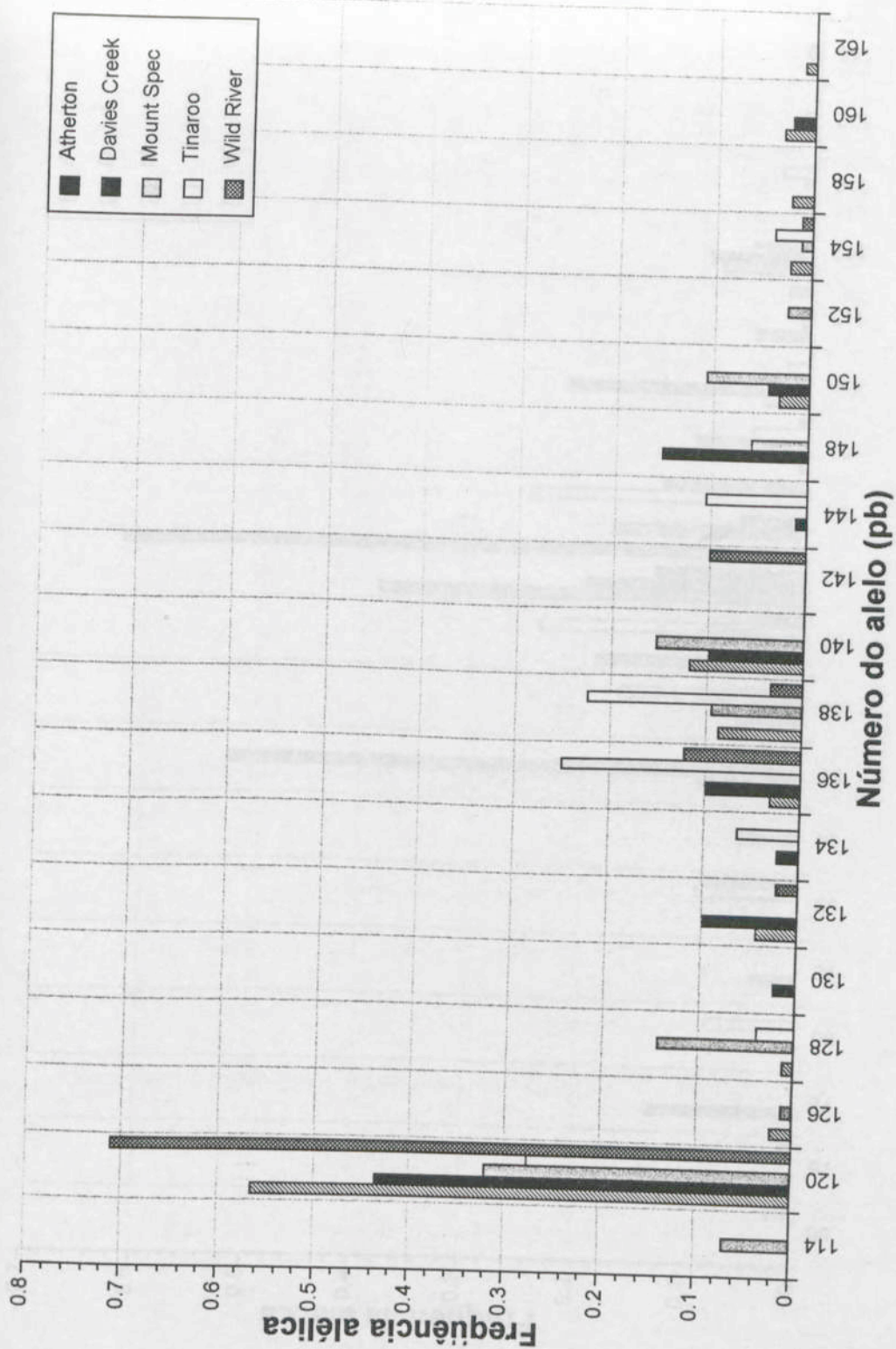


Figura 4A – Frequências alélicas observadas no loco EMBRA6, geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *Eucalyptus grandis*.

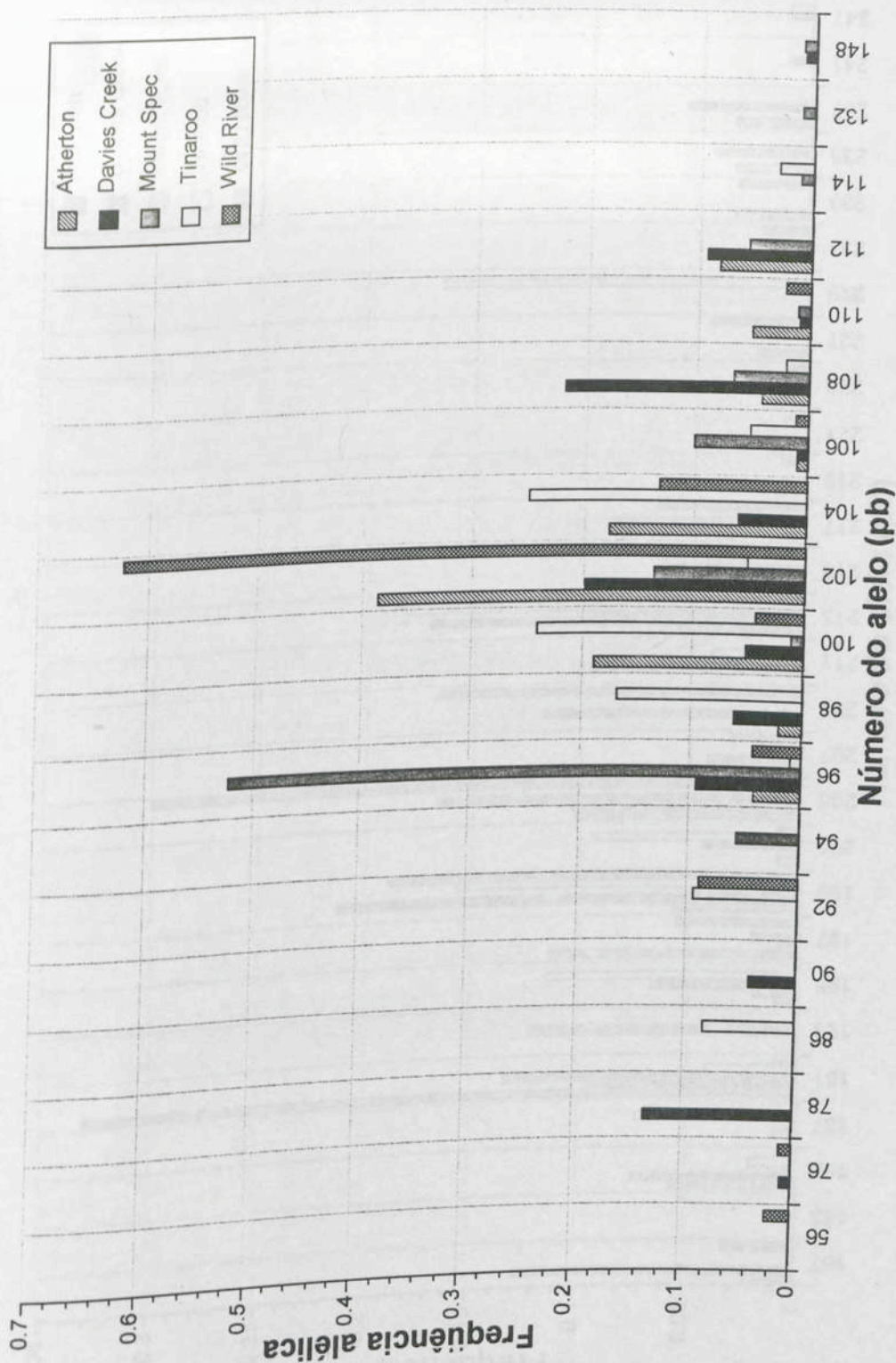


Figura 5A – Frequências alélicas observadas no loco EMBRA4, geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *Eucalyptus grandis*.

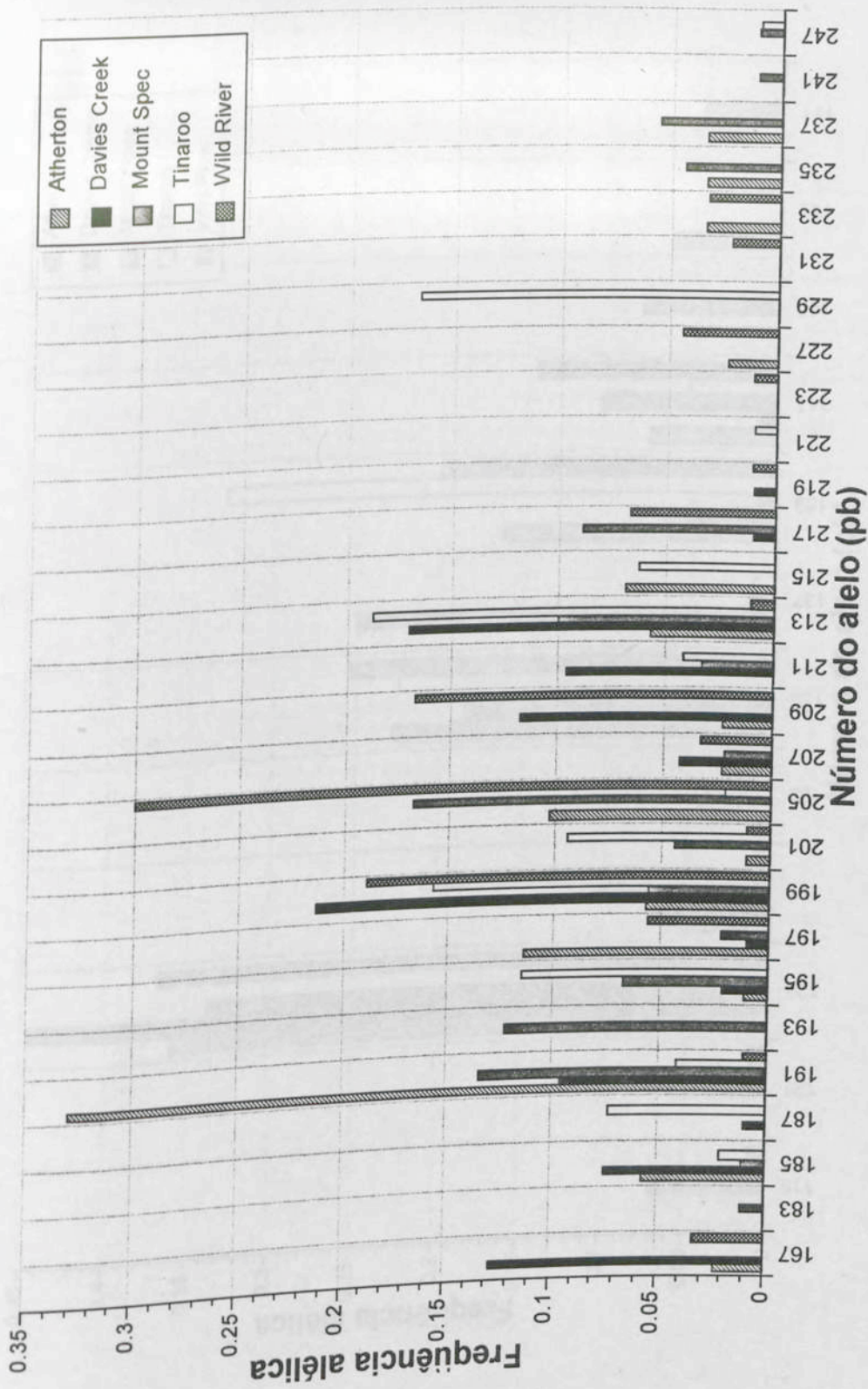


Figura 6A – Frequências alélicas observadas no loco EMBRA24, geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *Eucalyptus grandis*.

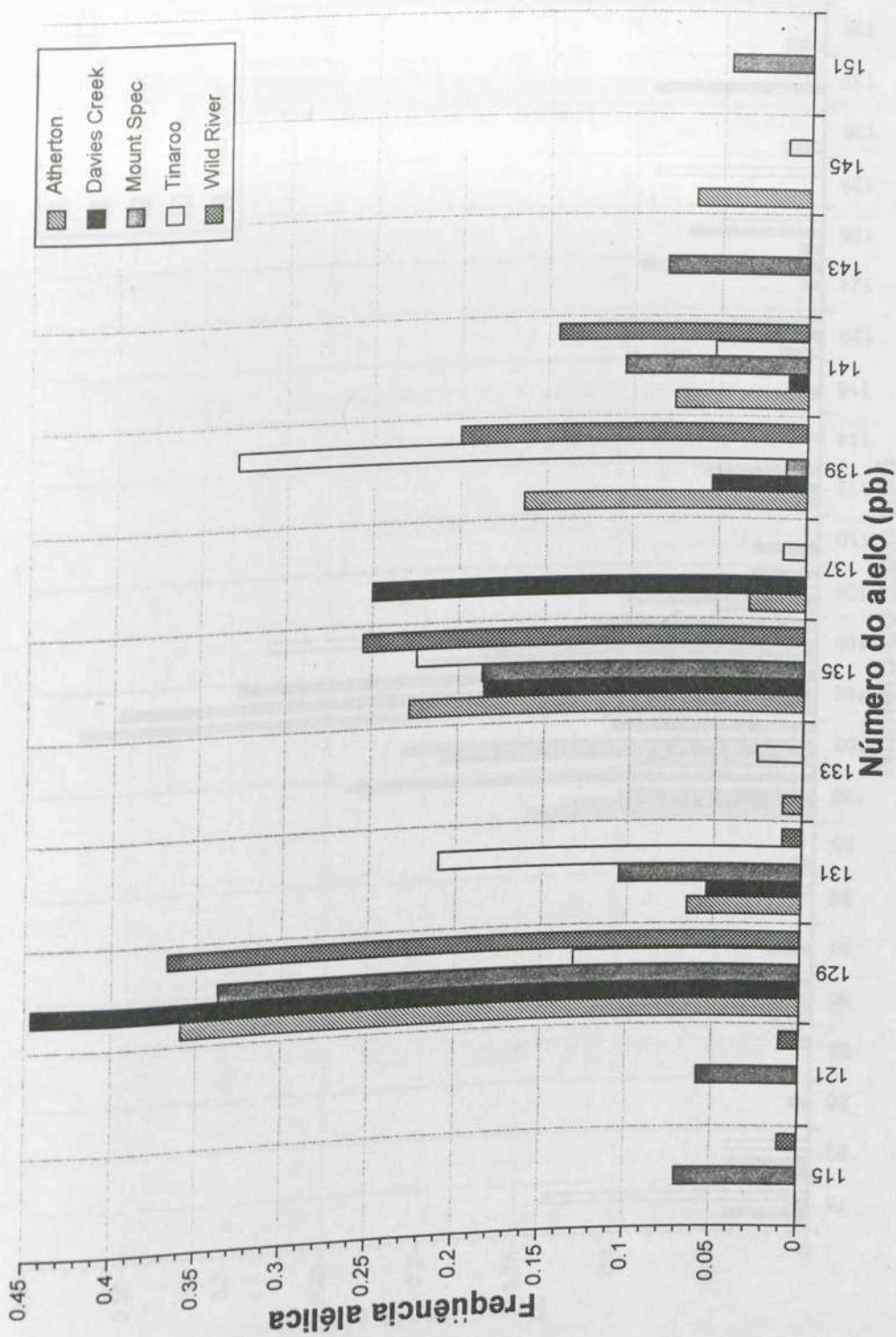


Figura 7A – Frequências alélicas observadas no loco EMBRA10, geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *Eucalyptus grandis*.

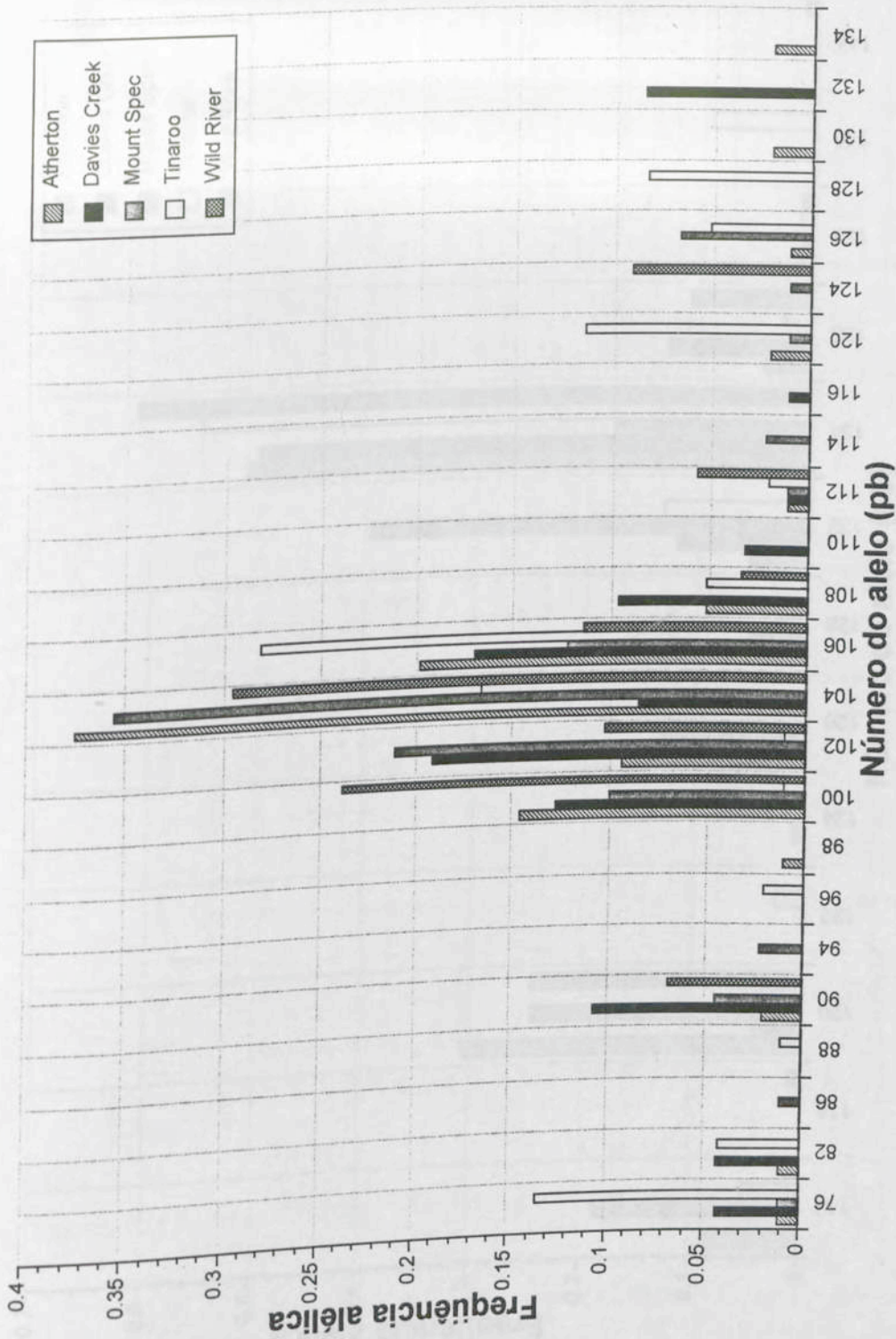


Figura 8A – Frequências alélicas observadas no loco EMBRA15, geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *Eucalyptus grandis*.

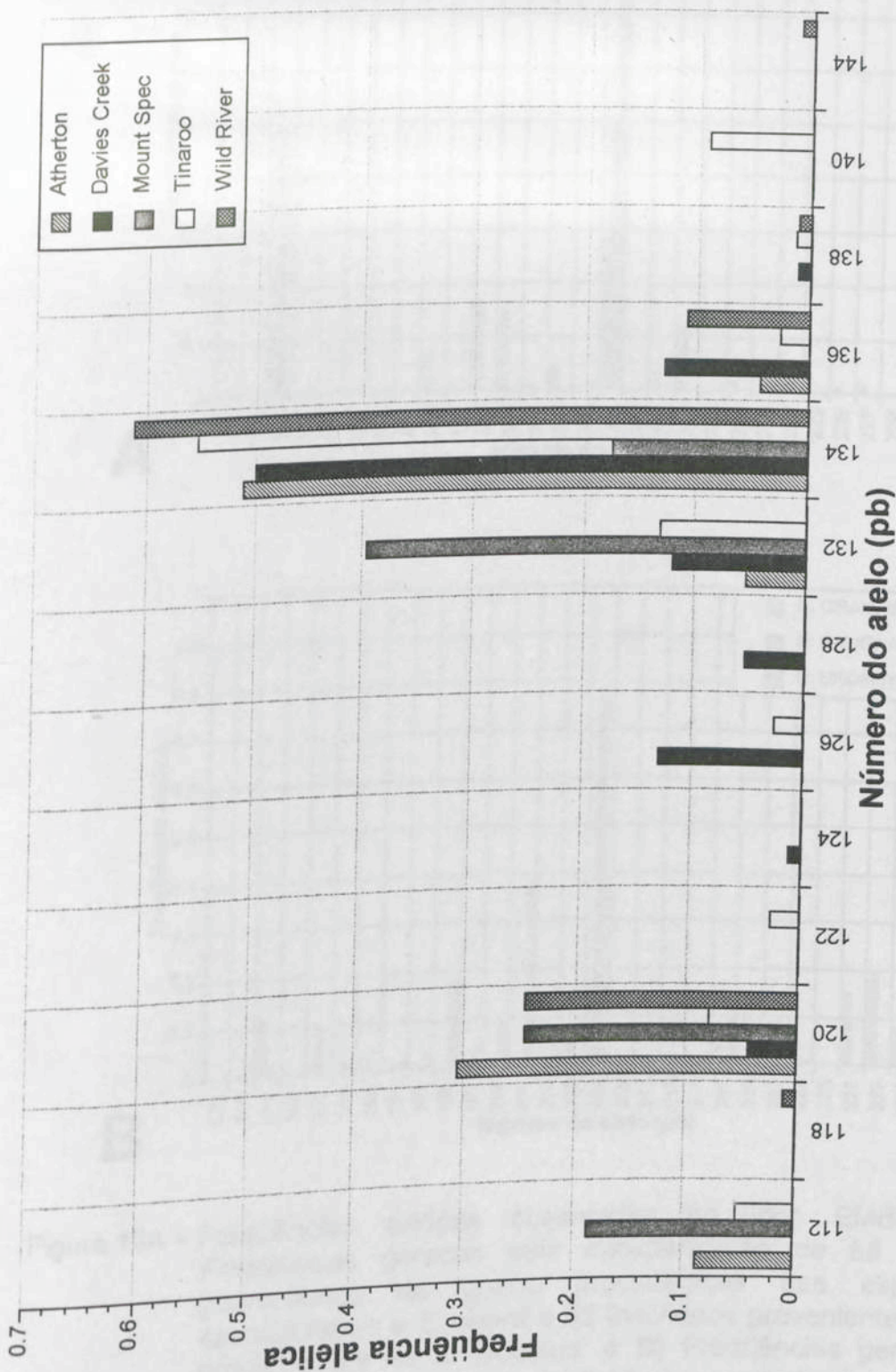


Figura 9A – Frequências alélicas observadas no loco EMBRA18, geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *Eucalyptus grandis*.

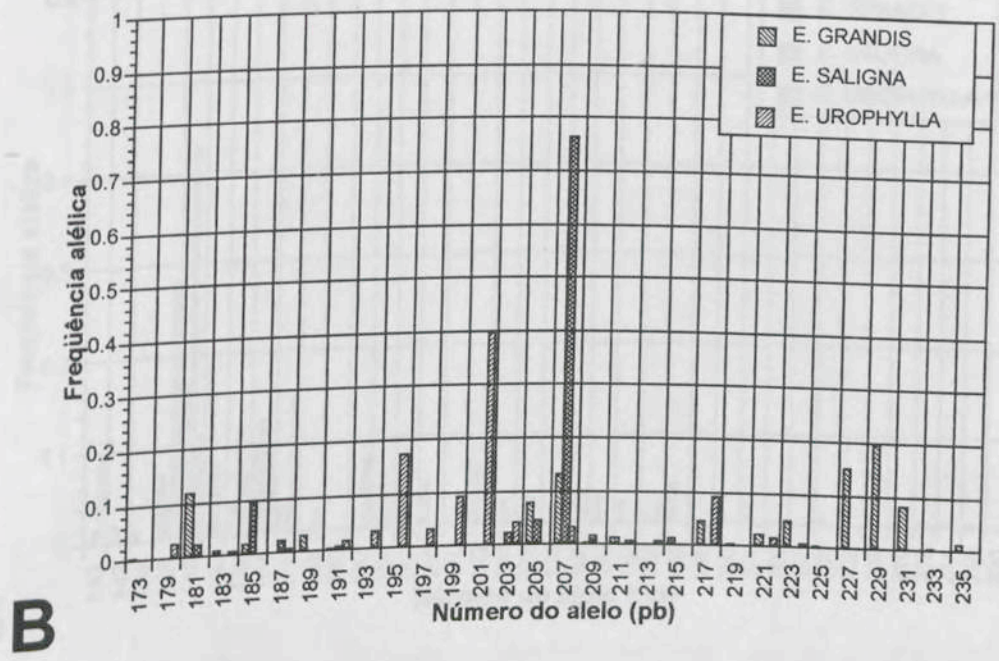
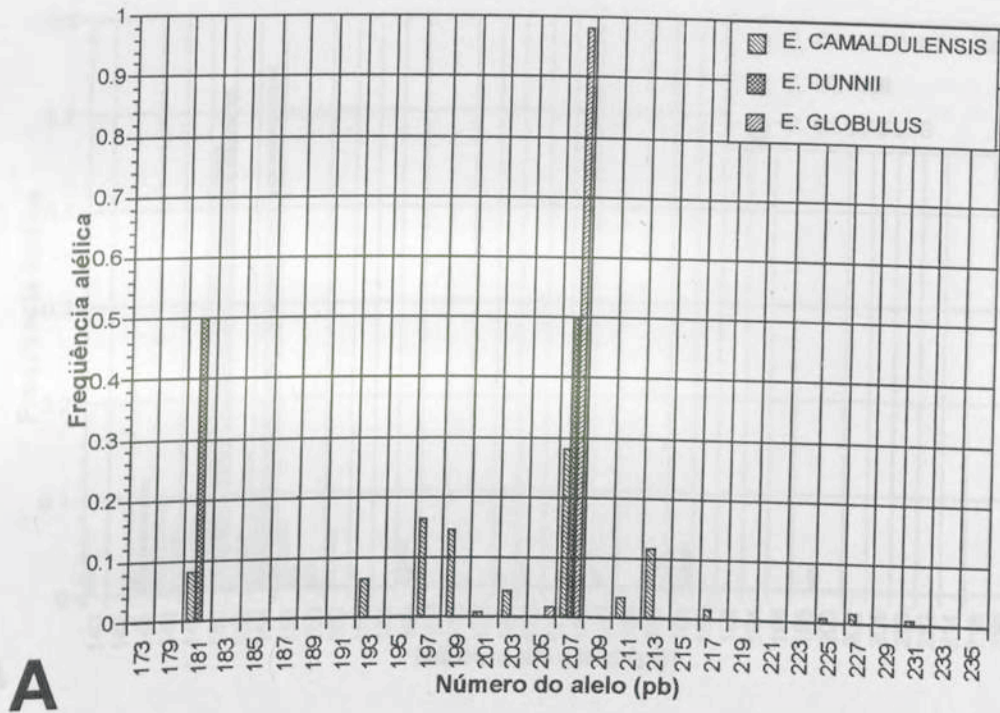


Figura 10A – Frequências alélicas observadas no loco EMBRA27: **A)** Frequências geradas pela caracterização de 68 indivíduos provenientes de quatro procedências das espécies *E. camaldulensis* e *E. dunnii* e 32 indivíduos provenientes de cinco procedências de *E. globulus*; e **B)** Frequências geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *E. grandis*, 68 indivíduos de quatro procedências de *E. saligna* e 68 indivíduos provenientes de duas procedências de *E. urophylla*.

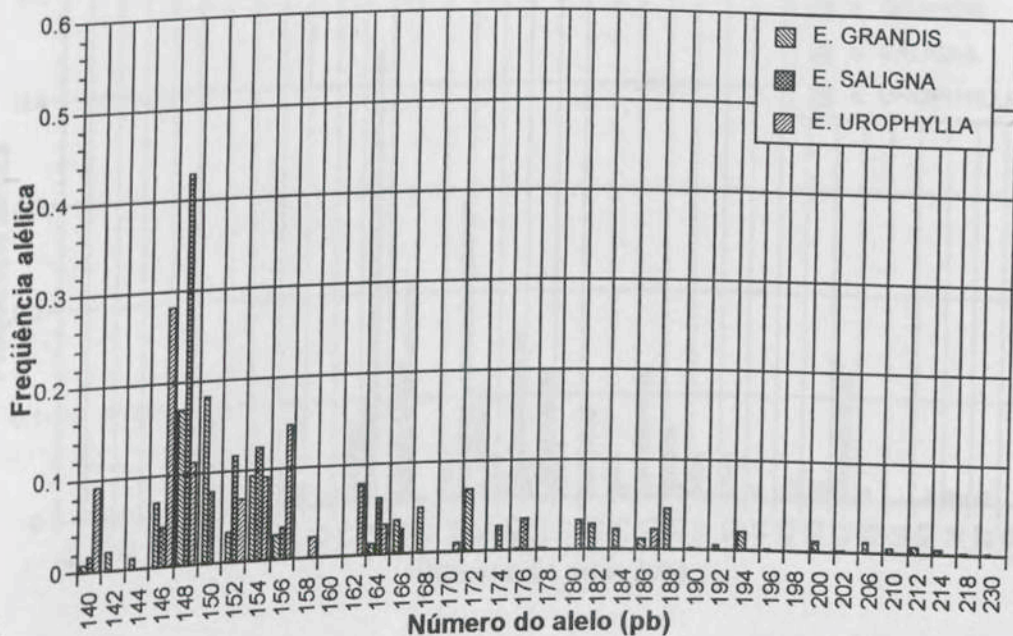
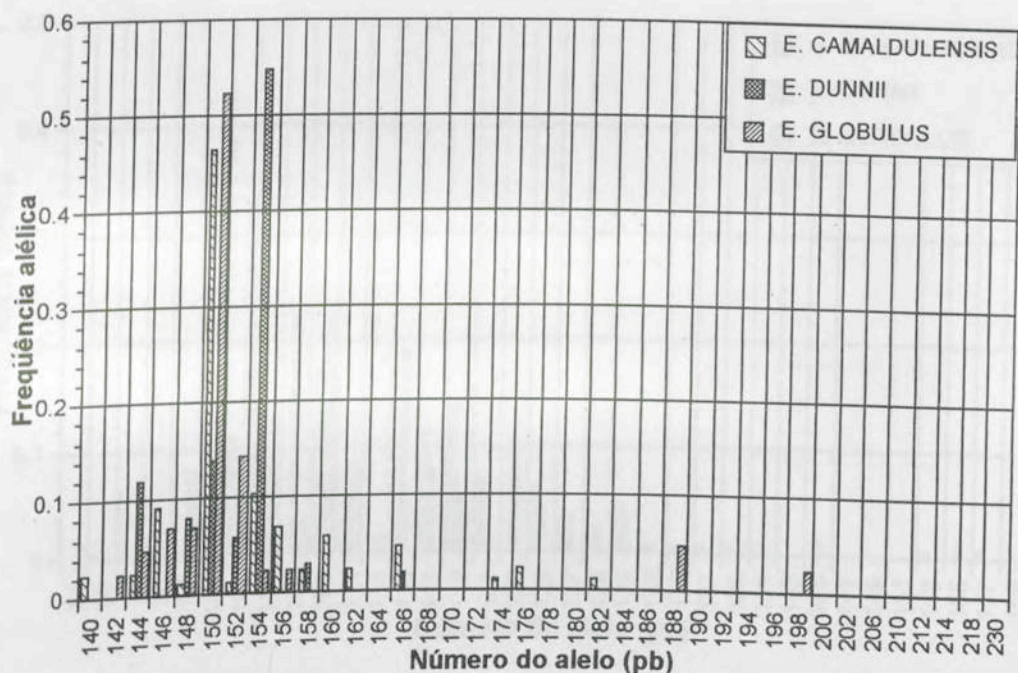


Figura 11A – Frequências alélicas observadas no loco EMBRA22: **A)** Frequências geradas pela caracterização de 68 indivíduos provenientes de quatro procedências das espécies *E. camaldulensis* e *E. dunnii* e 32 indivíduos provenientes de cinco procedências de *E. globulus*; e **B)** Frequências geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *E. grandis*, 68 indivíduos de quatro procedências de *E. saligna* e 68 indivíduos provenientes de duas procedências de *E. urophylla*.

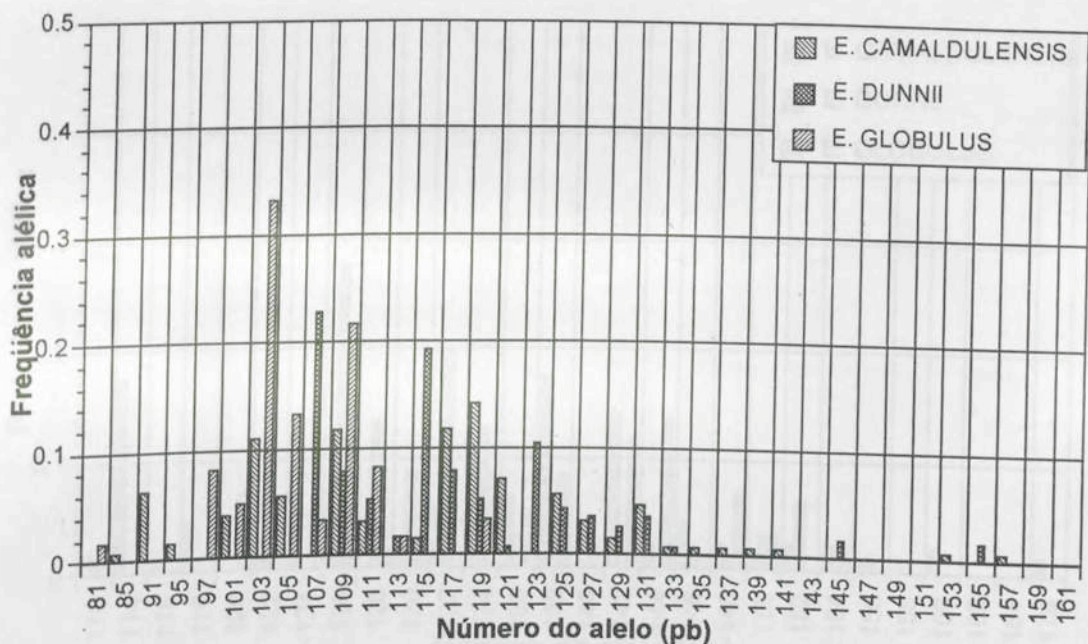
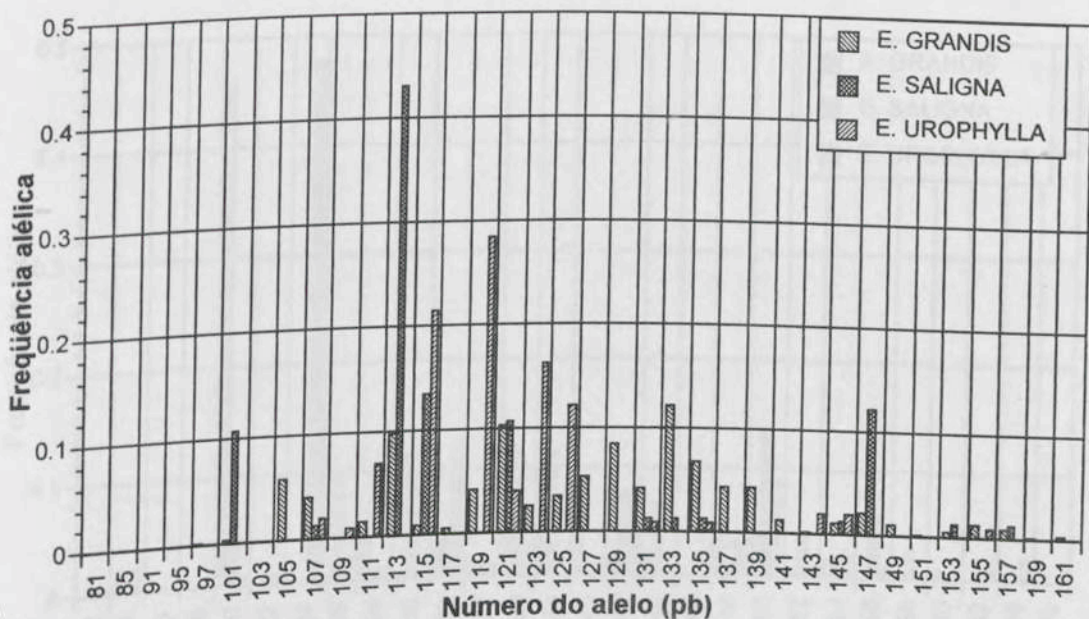
A**B**

Figura 12A – Frequências alélicas observadas no loco EMBRA5: **A)** Frequências geradas pela caracterização de 68 indivíduos provenientes de quatro procedências das espécies *E. camaldulensis* e *E. dunnii* e 32 indivíduos provenientes de cinco procedências de *E. globulus*; e **B)** Frequências geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *E. grandis*, 68 indivíduos de quatro procedências de *E. saligna* e 68 indivíduos provenientes de duas procedências de *E. urophylla*.

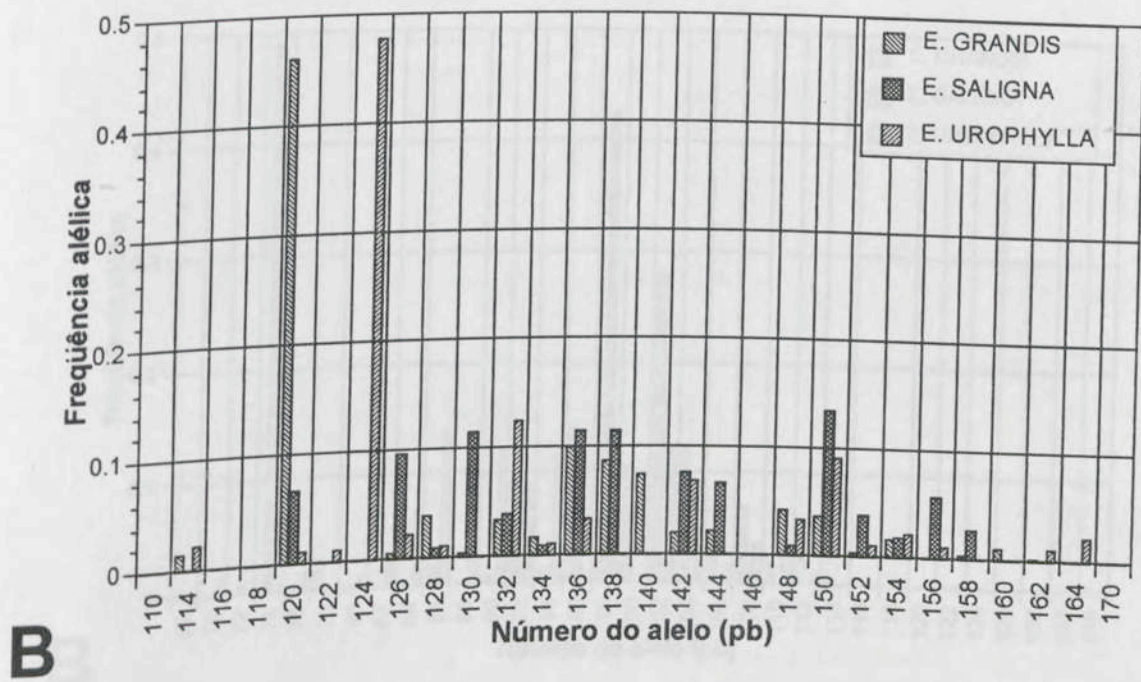
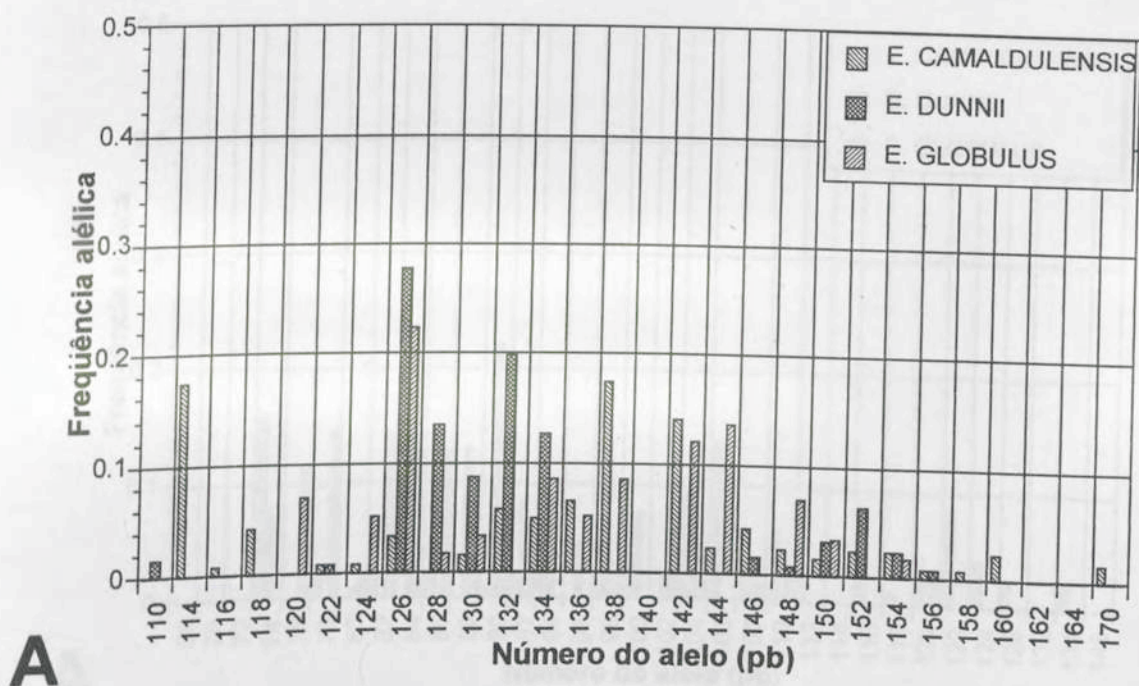


Figura 13A – Frequências alélicas observadas no loco EMBRA6: **A)** Frequências geradas pela caracterização de 68 indivíduos provenientes de quatro procedências das espécies *E. camaldulensis* e *E. dunnii* e 32 indivíduos provenientes de cinco procedências de *E. globulus*; e **B)** Frequências geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *E. grandis*, 68 indivíduos de quatro procedências de *E. saligna* e 68 indivíduos provenientes de duas procedências de *E. urophylla*.

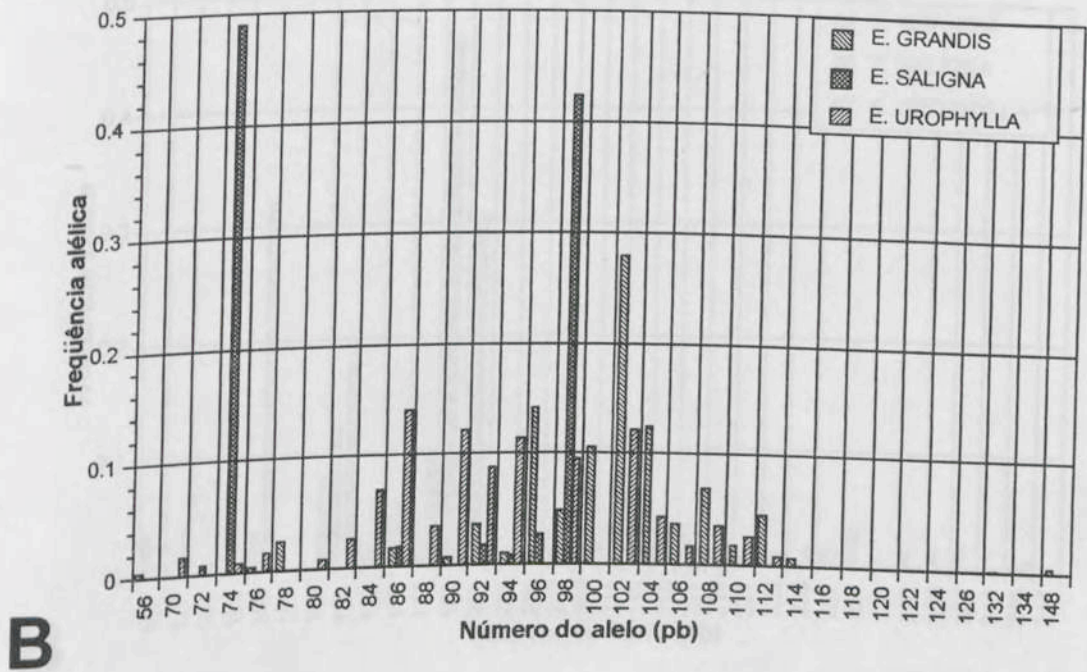
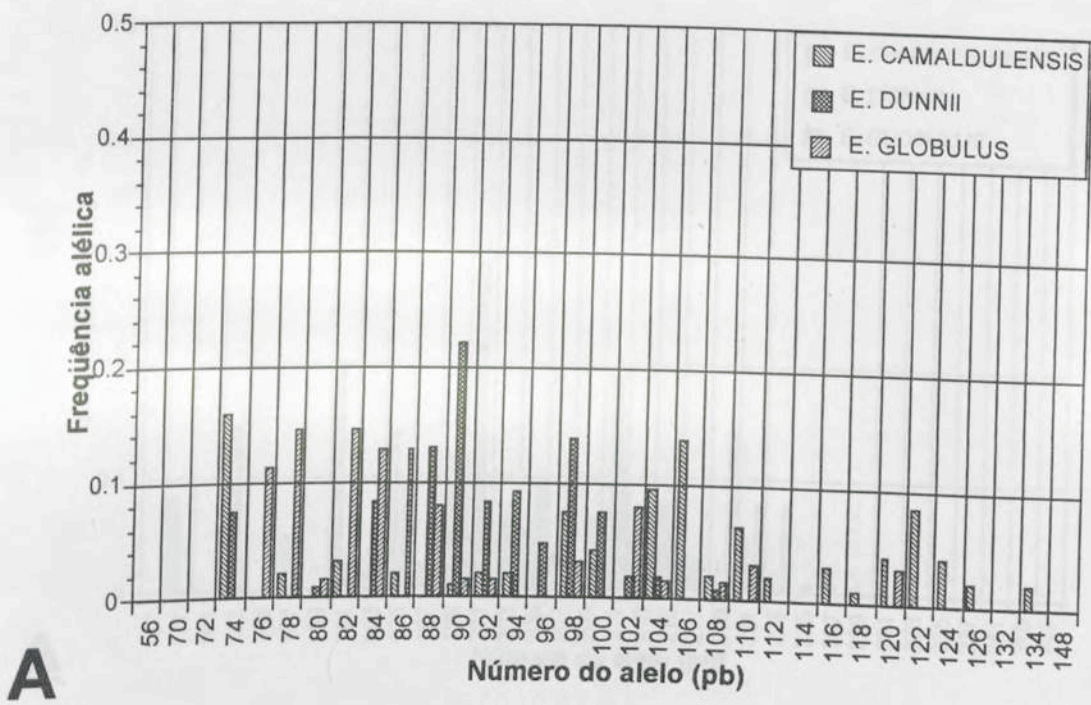
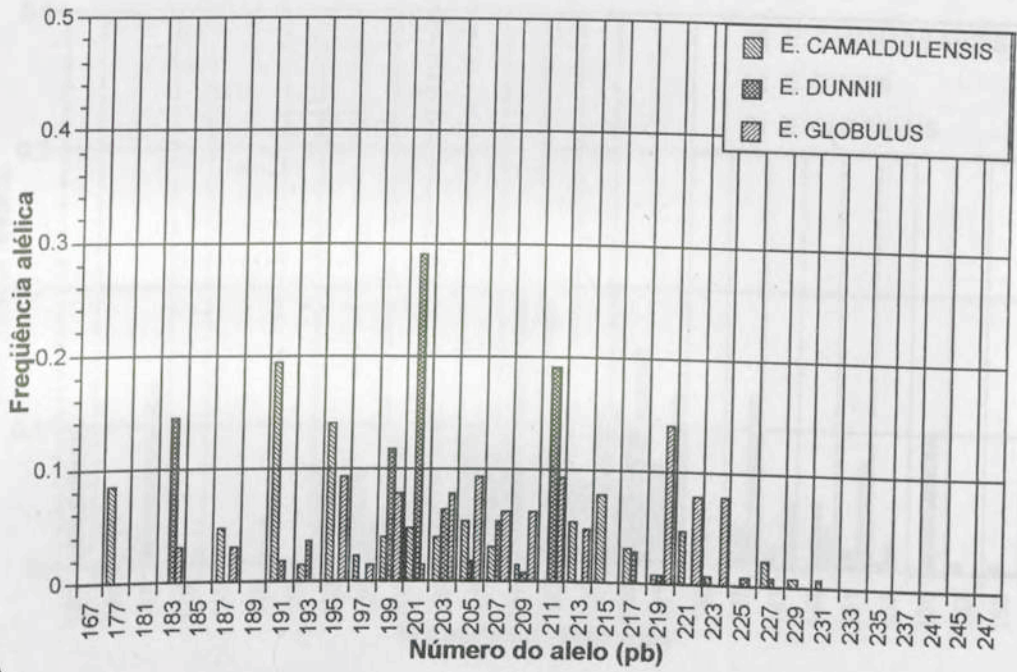
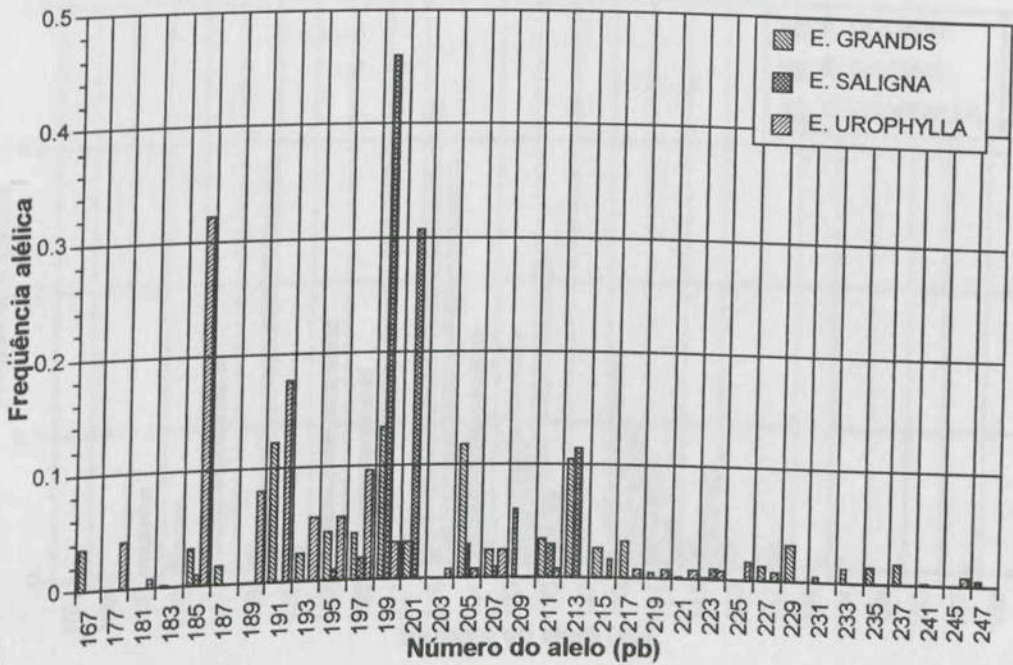


Figura 14A – Frequências alélicas observadas no loco EMBRA4: **A)** Frequências geradas pela caracterização de 68 indivíduos provenientes de quatro procedências das espécies *E. camaldulensis* e *E. dunnii* e 32 indivíduos provenientes de cinco procedências de *E. globulus*; e **B)** Frequências geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *E. grandis*, 68 indivíduos de quatro procedências de *E. saligna* e 68 indivíduos provenientes de duas procedências de *E. urophylla*.



A



B

Figura 15A – Frequências alélicas observadas no loco EMBRA24: **A)** Frequências geradas pela caracterização de 68 indivíduos provenientes de quatro procedências das espécies *E. camaldulensis* e *E. dunnii* e 32 indivíduos provenientes de cinco procedências de *E. globulus*; e **B)** Frequências geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *E. grandis*, 68 indivíduos de quatro procedências de *E. saligna* e 68 indivíduos provenientes de duas procedências de *E. urophylla*.

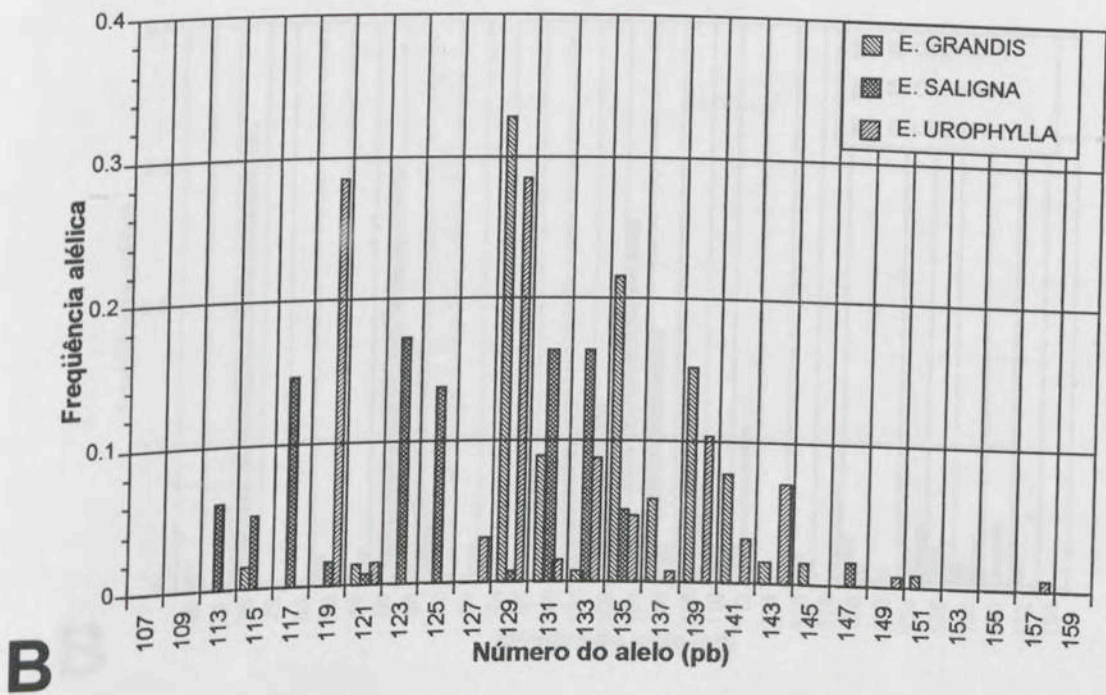
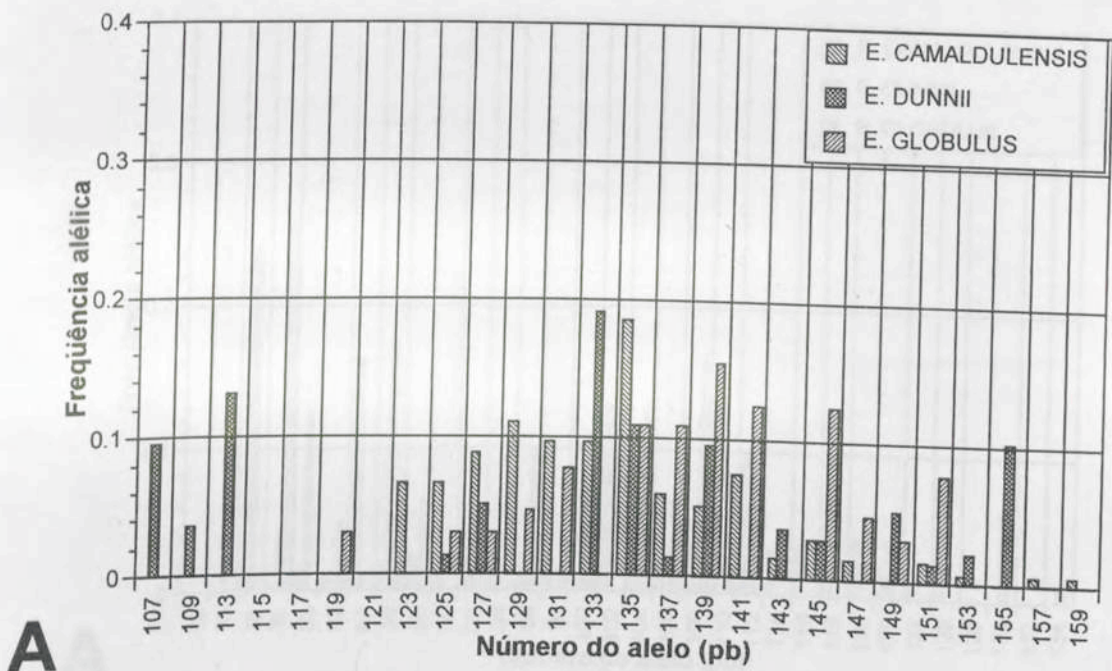


Figura 16A – Frequências alélicas observadas no loco EMBRA10: A) Frequências geradas pela caracterização de 68 indivíduos provenientes de quatro procedências das espécies *E. camaldulensis* e *E. dunnii* e 32 indivíduos provenientes de cinco procedências de *E. globulus*; e B) Frequências geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *E. grandis*, 68 indivíduos de quatro procedências de *E. saligna* e 68 indivíduos provenientes de duas procedências de *E. urophylla*.

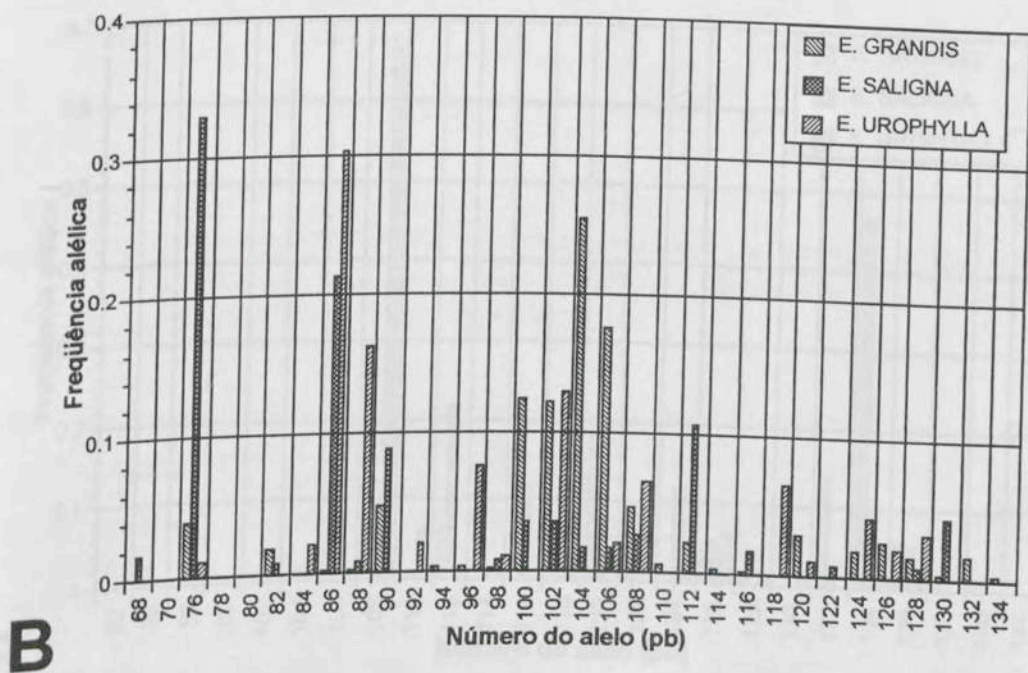
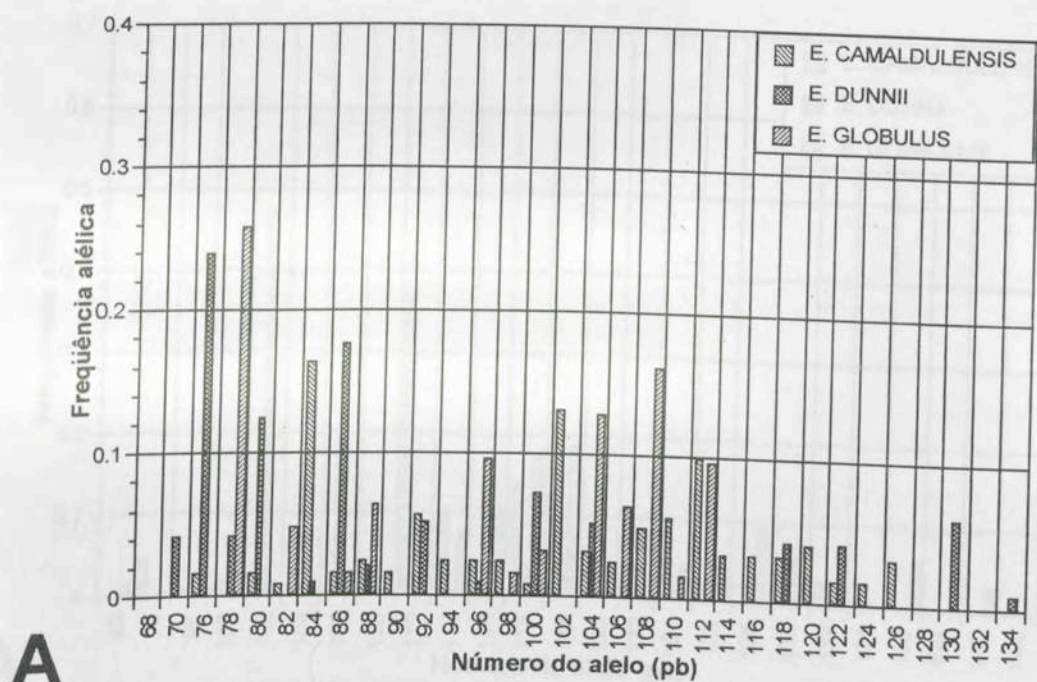


Figura 17A – Frequências alélicas observadas no loco EMBRA15: **A)** Frequências geradas pela caracterização de 68 indivíduos provenientes de quatro procedências das espécies *E. camaldulensis* e *E. dunnii* e 32 indivíduos provenientes de cinco procedências de *E. globulus*; e **B)** Frequências geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *E. grandis*, 68 indivíduos de quatro procedências de *E. saligna* e 68 indivíduos provenientes de duas procedências de *E. urophylla*.

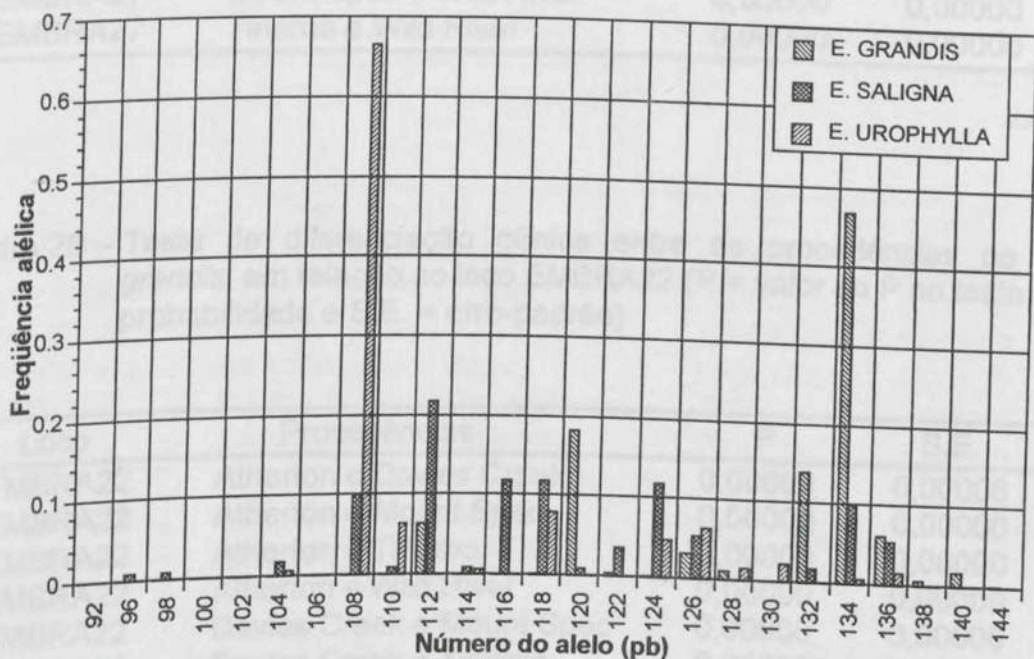
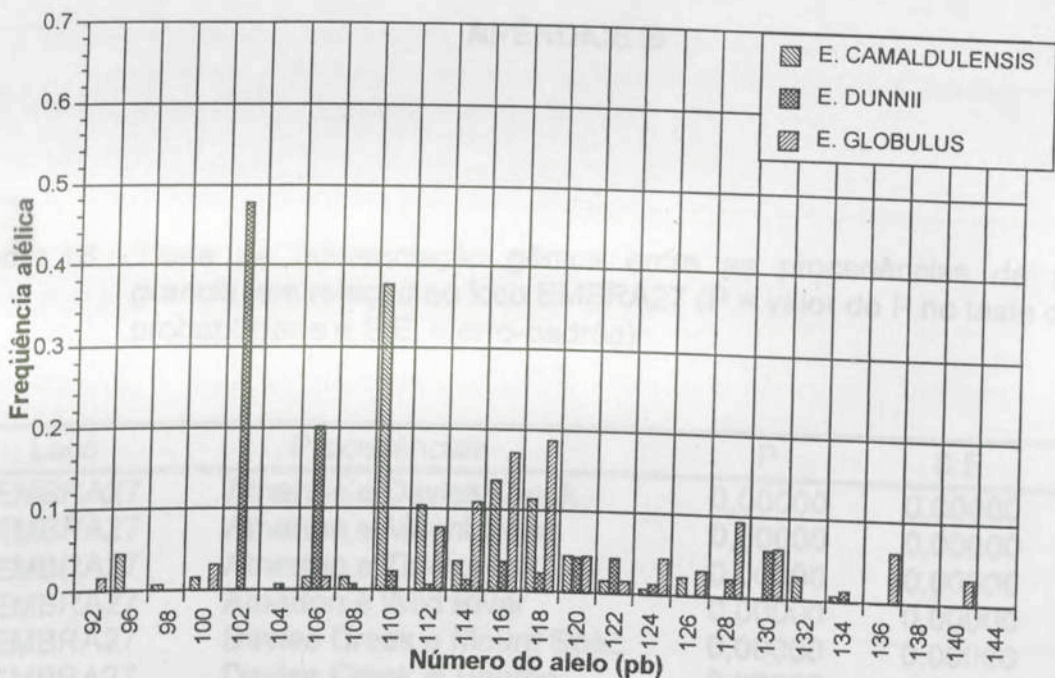


Figura 18A – Frequências alélicas observadas no loco EMBRA18: **A)** Frequências geradas pela caracterização de 68 indivíduos provenientes de quatro procedências das espécies *E. camaldulensis* e *E. dunnii* e 32 indivíduos provenientes de cinco procedências de *E. globulus*; e **B)** Frequências geradas pela caracterização de 240 indivíduos provenientes de cinco procedências de *E. grandis*, 68 indivíduos de quatro procedências de *E. saligna* e 68 indivíduos provenientes de duas procedências de *E. urophylla*.

APÊNDICE B

Quadro 1B – Teste de diferenciação gênica entre as procedências de *E. grandis*, em relação ao loco EMBRA27 (P = valor do P no teste de probabilidade e S.E. = erro-padrão)

Loco	Procedências	P	S.E.
EMBRA27	Atherton e Davies Creek	0,00000	0,00000
EMBRA27	Atherton e Mount Spec	0,00000	0,00000
EMBRA27	Atherton e Tinaroo	0,00000	0,00000
EMBRA27	Atherton e Wild River	0,00000	0,00000
EMBRA27	Davies Creek e Mount Spec	0,00000	0,00000
EMBRA27	Davies Creek e Tinaroo	0,00000	0,00000
EMBRA27	Davies Creek e Wild River	0,00000	0,00000
EMBRA27	Mount Spec e Tinaroo	0,00000	0,00000
EMBRA27	Mount Spec e Wild River	0,00000	0,00000
EMBRA27	Tinaroo e Wild River	0,00000	0,00000

Quadro 2B – Teste de diferenciação gênica entre as procedências de *E. grandis*, em relação ao loco EMBRA22 (P = valor do P no teste de probabilidade e S.E. = erro-padrão)

Loco	Procedências	P	S.E.
EMBRA22	Atherton e Davies Creek	0,00009	0,00006
EMBRA22	Atherton e Mount Spec	0,00000	0,00000
EMBRA22	Atherton e Tinaroo	0,00000	0,00000
EMBRA22	Atherton e Wild River	0,00000	0,00000
EMBRA22	Davies Creek e Mount Spec	0,00000	0,00000
EMBRA22	Davies Creek e Tinaroo	0,00000	0,00000
EMBRA22	Davies Creek e Wild River	0,00000	0,00000
EMBRA22	Mount Spec e Tinaroo	0,00000	0,00000
EMBRA22	Mount Spec e Wild River	0,00000	0,00000
EMBRA22	Tinaroo e Wild River	0,00000	0,00000

Quadro 3B – Teste de diferenciação gênica entre as procedências de *E. grandis*, em relação ao loco EMBRA5 (P = valor do P no teste de probabilidade e S.E. = erro-padrão)

Loco	Procedências	P	S.E.
EMBRA5	Atherton e Davies Creek	0,00000	0,00000
EMBRA5	Atherton e Mount Spec	0,00000	0,00000
EMBRA5	Atherton e Tinaroo	0,00000	0,00000
EMBRA5	Atherton e Wild River	0,00000	0,00000
EMBRA5	Davies Creek e Mount Spec	0,00000	0,00000
EMBRA5	Davies Creek e Tinaroo	0,00000	0,00000
EMBRA5	Davies Creek e Wild River	0,00000	0,00000
EMBRA5	Mount Spec e Tinaroo	0,00000	0,00000
EMBRA5	Mount Spec e Wild River	0,00000	0,00000
EMBRA5	Tinaroo e Wild River	0,00000	0,00000

Quadro 4B – Teste de diferenciação gênica entre as procedências de *E. grandis*, em relação ao loco EMBRA6 (P = valor do P no teste de probabilidade e S.E. = erro-padrão)

Loco	Procedências	P	S.E.
EMBRA6	Atherton e Davies Creek	0,00000	0,00000
EMBRA6	Atherton e Mount Spec	0,00000	0,00000
EMBRA6	Atherton e Tinaroo	0,00000	0,00000
EMBRA6	Atherton e Wild River	0,00132	0,00047
EMBRA6	Davies Creek e Mount Spec	0,00000	0,00000
EMBRA6	Davies Creek e Tinaroo	0,00000	0,00000
EMBRA6	Davies Creek e Wild River	0,00000	0,00000
EMBRA6	Mount Spec e Tinaroo	0,00000	0,00000
EMBRA6	Mount Spec e Wild River	0,00000	0,00000
EMBRA6	Tinaroo e Wild River	0,00000	0,00000

Quadro 5B – Teste de diferenciação gênica entre as procedências de *E. grandis*, em relação ao loco EMBRA4 (P = valor do P no teste de probabilidade e S.E. = erro-padrão)

Loco	Procedências	P	S.E.
EMBRA4	Atherton e Davies Creek	0,00000	0,00000
EMBRA4	Atherton e Mount Spec	0,00000	0,00000
EMBRA4	Atherton e Tinaroo	0,00000	0,00000
EMBRA4	Atherton e Wild River	0,00000	0,00000
EMBRA4	Davies Creek e Mount Spec	0,00000	0,00000
EMBRA4	Davies Creek e Tinaroo	0,00000	0,00000
EMBRA4	Davies Creek e Wild River	0,00000	0,00000
EMBRA4	Mount Spec e Tinaroo	0,00000	0,00000
EMBRA4	Mount Spec e Wild River	0,00000	0,00000
EMBRA4	Tinaroo e Wild River	0,00000	0,00000

Quadro 6B – Teste de diferenciação gênica entre as procedências de *E. grandis*, em relação ao loco EMBRA24 (P = valor do P no teste de probabilidade e S.E. = erro-padrão)

Loco	Procedências	P	S.E.
EMBRA24	Atherton e Davies Creek	0,00000	0,00000
EMBRA24	Atherton e Mount Spec	0,00000	0,00000
EMBRA24	Atherton e Tinaroo	0,00000	0,00000
EMBRA24	Atherton e Wild River	0,00000	0,00000
EMBRA24	Davies Creek e Mount Spec	0,00000	0,00000
EMBRA24	Davies Creek e Tinaroo	0,00000	0,00000
EMBRA24	Davies Creek e Wild River	0,00000	0,00000
EMBRA24	Mount Spec e Tinaroo	0,00000	0,00000
EMBRA24	Mount Spec e Wild River	0,00000	0,00000
EMBRA24	Tinaroo e Wild River	0,00000	0,00000

Quadro 7B – Teste de diferenciação gênica entre as procedências de *E. grandis*, em relação ao loco EMBRA10 (P = valor do P no teste de probabilidade e S.E. = erro-padrão)

Loco	Procedências	P	S.E.
EMBRA10	Atherton e Davies Creek	0,00000	0,00000
EMBRA10	Atherton e Mount Spec	0,00000	0,00000
EMBRA10	Atherton e Tinaroo	0,00543	0,00148
EMBRA10	Atherton e Wild River	0,00000	0,00000
EMBRA10	Davies Creek e Mount Spec	0,00000	0,00000
EMBRA10	Davies Creek e Tinaroo	0,00000	0,00000
EMBRA10	Davies Creek e Wild River	0,00000	0,00000
EMBRA10	Mount Spec e Tinaroo	0,00000	0,00000
EMBRA10	Mount Spec e Wild River	0,00000	0,00000
EMBRA10	Tinaroo e Wild River	0,00000	0,00000

Quadro 8B – Teste de diferenciação gênica entre as procedências de *E. grandis*, em relação ao loco EMBRA15 (P = valor do P no teste de probabilidade e S.E. = erro-padrão)

Loco	Procedências	P	S.E.
EMBRA15	Atherton e Davies Creek	0,00000	0,00000
EMBRA15	Atherton e Mount Spec	0,00332	0,00138
EMBRA15	Atherton e Tinaroo	0,00000	0,00000
EMBRA15	Atherton e Wild River	0,01206	0,00274
EMBRA15	Davies Creek e Mount Spec	0,00000	0,00000
EMBRA15	Davies Creek e Tinaroo	0,00000	0,00000
EMBRA15	Davies Creek e Wild River	0,00000	0,00000
EMBRA15	Mount Spec e Tinaroo	0,00000	0,00000
EMBRA15	Mount Spec e Wild River	0,00023	0,00019
EMBRA15	Tinaroo e Wild River	0,00000	0,00000

Quadro 9B – Teste de diferenciação gênica entre as procedências de *E. grandis*, em relação ao loco EMBRA18 (P = valor do P no teste de probabilidade e S.E. = erro-padrão)

Loco	Procedências	P	S.E.
EMBRA18	Atherton e Davies Creek	0,00000	0,00000
EMBRA18	Atherton e Mount Spec	0,00000	0,00000
EMBRA18	Atherton e Tinaroo	0,00003	0,00003
EMBRA18	Atherton e Wild River	0,00000	0,00000
EMBRA18	Davies Creek e Mount Spec	0,00000	0,00000
EMBRA18	Davies Creek e Tinaroo	0,00003	0,00003
EMBRA18	Davies Creek e Wild River	0,00000	0,00000
EMBRA18	Mount Spec e Tinaroo	0,00000	0,00000
EMBRA18	Mount Spec e Wild River	0,00000	0,00000
EMBRA18	Tinaroo e Wild River	0,00000	0,00000

2	201	201	115	115	80	85	204	218	75	78
3	194	204	157	157	50	105	199	199	ND	ND
4	200	200	121	123	81	83	194	199	85	101
5	192	204	ND	ND	101	107	177	185	84	90
6	199	199	117	138	78	103	194	197	120	129
7	195	200	119	123	99	99	185	187	86	86
8	194	201	ND	ND	80	88	185	185	ND	ND
10	185	201	105	125	74	74	191	191	85	86
11	193	199	125	129	83	103	177	191	86	86
12	207	207	108	105	74	85	199	212	ND	ND
13	188	218	111	135	85	91	195	218	135	135
14	192	217	103	125	92	101	177	187	82	113
15	170	204	119	129	93	97	192	192	85	86
16	197	203	107	136	90	113	201	218	ND	ND
17	191	196	117	136	78	85	194	206	78	78
18	195	201	120	139	88	82	185	191	84	118
19	201	201	119	133	81	85	185	185	108	112
20	200	201	113	138	76	79	177	188	109	108
21	199	214	121	131	84	89	191	191	84	106
22	182	207	109	106	74	87	212	214	ND	ND
23	201	201	123	138	81	85	195	185	108	108
24	202	202	111	125	80	86	177	215	86	92
25	195	212	102	124	87	106	177	199	120	120
27	193	197	107	129	79	101	199	199	86	86
28	197	199	115	119	87	105	140	180	82	82
29	199	199	115	146	75	107	185	185	ND	ND
30	201	225	127	129	105	107	185	185	ND	ND
31	197	225	119	181	98	89	195	185	85	86
32	203	203	113	117	78	79	199	199	ND	ND
33	182	207	ND	ND	74	80	ND	ND	129	129
34	193	199	117	119	86	84	195	199	85	86

APÊNDICE C

Quadro 1C – Alelos detectados (em pares de bases) nos 225 genótipos de *Eucalyptus urophylla* avaliados com os microssatélites EMBRA27, EMBRA5, EMBRA4, EMBRA24 e EMBRA15 (clones 1 a 200 correspondem aos indivíduos que compõem a população de melhoramento, e clones CP 1 a 40 correspondem aos clones plantados comercialmente; ND=dados não disponíveis)

Clone	EMBRA27		EMBRA5		EMBRA4		EMBRA24		EMBRA15	
	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2
1	201	201	121	123	82	83	194	199	86	101
2	201	201	113	115	90	95	208	216	76	76
3	194	224	157	157	90	105	199	199	ND	ND
4	200	200	121	123	81	83	194	199	86	101
5	195	204	ND	ND	101	107	177	185	84	90
6	199	199	117	133	78	103	194	197	126	129
7	195	200	119	123	99	99	185	197	86	86
9	184	201	ND	ND	80	88	185	185	ND	ND
10	195	201	103	125	74	74	191	191	86	86
11	195	199	125	129	93	103	177	191	86	86
12	207	207	105	105	74	93	199	212	ND	ND
13	188	218	111	135	85	91	185	218	135	135
14	192	217	123	125	92	101	177	197	82	118
15	179	229	119	129	93	97	192	192	86	86
16	197	203	107	135	90	113	201	218	ND	ND
17	191	195	117	135	95	95	194	208	76	76
18	195	201	123	133	88	92	185	191	84	118
19	201	201	119	123	81	93	185	185	108	118
20	200	201	113	126	79	79	177	198	106	106
21	199	211	121	123	85	90	191	191	84	108
22	182	207	109	109	74	97	212	214	ND	ND
23	201	201	123	133	81	81	185	185	108	108
24	202	202	111	121	90	90	177	216	86	92
26	195	212	103	121	97	109	177	199	120	120
27	193	197	107	121	95	101	189	199	86	86
28	197	199	115	119	97	107	ND	ND	82	82
29	199	199	115	148	101	108	185	185	ND	ND
30	201	225	127	129	101	103	185	192	ND	ND
31	197	225	119	121	86	86	196	196	88	88
32	203	203	113	117	79	79	199	199	ND	ND
33	182	207	ND	ND	74	95	ND	ND	129	129
34	193	199	117	119	84	84	189	199	86	86

Continua...

Quadro 1C, Cont.

Clone	EMBRA27		EMBRA5		EMBRA4		EMBRA24		EMBRA15	
	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2
35	193	198	107	122	95	95	189	200	86	86
36	193	202	120	131	91	91	189	189	86	86
37	196	196	120	120	95	95	177	186	ND	ND
38	204	204	104	116	87	99	192	194	ND	ND
39	198	200	107	133	85	89	186	200	86	106
40	195	204	107	131	83	95	191	202	108	108
41	195	226	131	131	85	106	186	202	ND	ND
42	195	204	115	120	87	106	196	218	86	86
43	202	202	124	129	74	91	186	216	86	90
44	195	202	122	124	91	91	192	198	86	86
45	198	202	124	142	91	91	192	192	86	86
48	204	204	122	131	79	108	177	192	ND	ND
49	179	204	118	129	91	91	186	186	86	94
50	196	200	116	120	76	76	194	210	86	86
51	193	202	120	131	91	91	189	189	86	86
53	200	200	125	125	91	91	177	186	86	86
54	195	206	106	125	85	104	196	226	87	87
55	196	196	107	122	101	110	186	187	88	98
56	202	213	111	127	74	101	196	226	ND	ND
57	194	202	124	124	85	85	186	196	88	88
58	195	214	104	111	87	93	196	200	86	86
59	179	188	112	118	83	87	186	190	ND	ND
60	196	200	111	120	104	106	186	198	108	108
61	179	188	111	118	83	87	186	189	89	89
62	192	227	116	127	87	97	186	198	84	84
63	190	196	111	120	81	87	192	198	86	86
64	200	202	107	107	88	91	186	192	84	84
65	200	202	107	107	88	91	186	192	84	84
66	190	195	111	125	97	97	186	192	ND	ND
67	196	203	146	146	74	97	198	212	ND	ND
68	196	227	111	122	85	97	198	198	84	84
69	196	204	125	125	92	94	186	198	ND	ND
70	201	205	118	123	94	109	177	190	109	115
72	188	193	100	120	84	89	177	188	90	90
73	193	207	120	120	98	98	188	188	109	109
74	196	198	110	120	89	94	202	213	ND	ND
75	201	211	115	120	102	102	ND	ND	95	95
76	202	202	129	135	102	102	186	209	76	88
77	196	206	120	141	70	83	192	192	ND	ND
78	204	229	132	137	83	83	198	205	ND	ND
79	196	204	116	133	100	104	198	213	97	97
80	196	196	123	137	74	100	192	221	127	127
81	196	205	114	121	ND	ND	186	192	86	127

Continua...

Quadro 1C, Cont.

Clone	EMBRA27		EMBRA5		EMBRA4		EMBRA24		EMBRA15	
	Alelo1	alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2
82	196	196	114	137	74	104	198	198	76	115
83	198	202	114	125	92	92	196	200	ND	ND
84	201	212	115	125	88	104	198	211	105	105
85	196	196	121	133	83	83	199	221	ND	ND
86	194	205	118	118	83	90	177	204	101	101
87	194	225	137	157	90	106	200	200	ND	ND
88	202	202	115	125	94	94	177	194	86	86
89	198	198	108	115	83	106	199	213	ND	ND
90	202	205	115	125	88	92	188	199	76	104
91	ND	ND	ND	ND	83	87	186	186	107	107
92	196	200	127	133	70	88	188	221	94	97
93	184	184	120	139	88	88	186	186	94	107
94	228	228	115	133	98	104	ND	ND	97	97
95	205	205	125	137	ND	ND	192	198	104	129
96	202	202	115	133	74	102	192	192	109	120
97	179	228	100	133	92	92	201	221	76	90
98	194	228	113	137	83	83	175	175	ND	ND
99	ND	ND	110	110	ND	ND	192	196	86	86
100	194	201	115	121	84	107	190	190	88	88
101	203	211	115	121	88	106	177	196	102	114
102	182	208	121	125	74	88	ND	ND	102	104
103	201	230	125	125	ND	ND	193	210	87	87
104	196	204	119	119	82	82	210	210	87	87
105	202	204	115	138	93	102	198	212	88	88
107	215	215	123	125	87	102	177	177	ND	ND
108	200	200	110	125	86	97	191	213	ND	ND
109	208	215	125	125	75	82	191	212	77	77
110	201	201	111	115	84	89	177	191	ND	ND
111	201	201	117	121	75	98	189	189	126	126
112	196	204	119	125	70	70	191	219	87	103
113	213	213	115	129	ND	ND	191	191	108	108
114	200	200	ND	ND	86	86	189	189	108	113
115	202	204	113	125	84	93	191	191	ND	ND
116	201	201	111	149	87	87	198	219	ND	ND
117	201	201	113	149	97	97	187	198	ND	ND
118	ND	ND	80	111	94	103	177	177	81	81
119	201	204	ND	ND	84	87	187	191	80	80
120	201	216	115	115	87	91	177	198	86	86
122	193	193	103	109	87	87	194	198	86	86
123	195	211	113	121	87	95	177	189	128	128
124	ND	ND	103	113	89	100	198	198	81	81
125	193	193	ND	ND	87	91	193	198	86	86
126	202	209	ND	ND	84	87	184	184	94	94

Continua...

Quadro 1C, Cont.

Clone	EMBRA27		EMBRA5		EMBRA4		EMBRA24		EMBRA15	
	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2
127	204	228	145	159	70	107	191	191	ND	ND
129	201	201	103	127	ND	ND	177	191	115	115
130	179	179	115	117	98	101	177	198	118	118
132	ND	ND	ND	ND	75	101	ND	ND	86	106
135	179	198	103	121	80	84	177	193	86	110
136	ND	ND	115	119	75	96	194	196	90	101
137	204	204	114	122	74	85	208	208	86	86
138	207	207	120	120	74	74	194	214	86	86
139	196	196	116	120	85	102	185	200	108	110
140	210	210	114	114	97	108	194	198	86	86
141	194	200	123	129	91	112	216	218	108	108
142	203	205	114	116	79	95	185	194	88	88
143	198	204	116	123	103	106	188	218	90	108
144	196	196	120	120	89	89	177	198	88	90
145	196	204	120	125	83	83	192	204	83	103
146	196	204	120	125	70	70	192	219	86	103
147	196	204	120	125	70	83	192	204	103	103
148	196	196	120	125	83	83	192	204	103	103
149	202	218	124	153	91	97	196	198	96	96
150	196	204	120	120	70	83	192	204	103	103
151	183	202	120	126	83	93	187	187	110	110
152	206	208	129	142	83	87	185	185	106	106
153	194	198	119	128	83	104	222	222	88	126
154	202	214	111	116	87	91	177	229	88	100
155	198	203	112	116	88	93	196	227	86	106
156	190	194	103	119	97	97	189	192	76	76
157	201	211	119	134	99	104	185	192	86	86
158	200	200	118	133	77	104	194	198	77	126
159	182	207	106	122	74	97	ND	ND	ND	ND
160	195	195	119	134	91	91	185	200	76	76
161	195	201	120	132	89	89	198	202	76	98
162	196	204	122	131	91	97	189	216	76	86
163	196	199	125	130	93	104	177	192	86	86
165	195	204	112	120	95	95	208	238	76	76
166	179	230	116	129	91	91	185	214	88	106
167	195	204	113	121	83	95	185	194	76	82
168	193	201	108	120	83	91	196	196	86	86
169	193	201	118	118	91	106	196	210	98	98
170	193	201	108	108	83	96	190	196	86	86
171	182	196	111	121	87	102	186	190	86	98
172	196	196	107	121	102	102	186	188	88	98
173	186	194	119	123	109	109	190	190	86	88
174	196	205	121	125	94	102	186	190	98	98

Continua...

Quadro 1C, Cont.

Clone	EMBRA27		EMBRA5		EMBRA4		EMBRA24		EMBRA5	
	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2
175	186	215	107	123	102	102	186	190	88	88
176	188	196	117	119	90	92	178	208	88	88
177	200	200	111	127	89	104	175	190	84	133
178	200	202	110	127	89	98	188	190	82	82
179	194	194	121	121	81	87	190	200	86	107
180	196	200	103	127	91	91	192	192	129	129
181	202	206	123	137	91	95	177	186	100	100
182	230	230	121	125	74	91	ND	ND	94	100
183	201	201	115	115	85	85	212	214	86	113
184	202	208	121	121	87	87	198	198	ND	ND
185	192	195	109	111	87	96	196	198	86	86
186	201	201	121	121	87	87	198	226	84	84
187	201	201	112	127	89	95	198	198	103	103
188	194	201	111	115	74	90	175	192	94	94
189	201	201	108	125	81	87	194	200	76	76
190	182	207	119	129	89	92	177	194	127	127
192	194	201	109	139	87	87	177	182	94	94
193	201	203	ND	ND	89	92	192	192	ND	ND
195	201	201	145	145	83	87	178	212	94	94
196	193	210	ND	ND	98	98	186	217	98	129
197	214	222	ND	ND	96	102	218	245	94	118
198	201	201	ND	ND	98	111	192	214	129	129
199	204	227	ND	ND	91	103	194	210	96	129
200	204	204	113	123	90	90	192	202	107	127
TC1	194	194	113	125	83	83	197	197	ND	ND
TC2	203	229	146	146	83	83	191	209	86	86
TC3	204	204	125	138	83	87	191	197	103	128
TC4	203	203	118	120	83	91	191	203	100	100
TC5	204	204	109	125	83	91	197	197	86	100
TC6	204	227	120	130	97	97	191	220	92	104
TC7	187	187	120	134	83	101	199	205	100	100
TC8	195	225	134	138	99	103	197	220	96	96
TC9	204	204	109	125	83	91	197	203	99	99
TC10	203	227	120	130	91	104	193	210	96	136
TC11	215	227	134	138	74	97	197	209	104	128
TC12	203	227	111	134	91	91	189	197	76	76
TC13	204	207	120	134	83	110	195	197	103	103
TC14	195	211	108	130	97	97	195	210	103	103
TC15	195	227	103	130	70	70	190	210	70	104
TC16	195	195	118	119	83	83	191	210	108	108
TC17	195	229	109	125	70	91	203	205	99	99
TC18	195	195	109	125	91	91	205	218	100	100
TC19	204	229	125	144	83	91	193	193	86	100

Continua...

Quadro 1C - Cont.

Clone	EMBRA27		EMBRA5		EMBRA4		EMBRA24		EMBRA15	
	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2
TC20	195	195	118	120	83	83	191	209	107	107
TC21	195	227	134	138	99	103	197	220	96	96
TC22	195	227	125	130	83	83	203	210	104	104
TC23	203	203	125	138	83	87	191	197	103	128
TC24	195	211	125	130	70	70	191	209	96	96
TC25	201	223	116	134	97	101	210	226	86	96
TC26	195	211	125	134	103	103	220	244	96	96
TC27	204	225	105	134	103	103	210	214	104	108
TC28	228	228	121	138	99	99	191	199	105	105
TC29	203	211	119	134	70	70	193	220	104	104
TC30	197	204	106	120	83	83	191	207	99	103
TC31	208	229	109	125	70	91	193	205	100	108
TC32	203	227	146	160	70	108	191	191	ND	ND
TC33	229	229	109	125	91	91	197	203	100	103
TC34	204	204	115	125	93	105	189	222	86	103
TC35	204	204	115	126	93	105	189	222	86	103
TC36	204	217	120	126	112	112	212	218	86	97
TC37	202	217	107	130	95	95	210	212	97	97
TC38	196	229	109	126	83	83	204	210	100	100
TC39	188	196	132	137	101	101	191	200	108	108
TC40	206	218	107	154	95	105	210	212	94	97

Quadro 2C – Alelos detectados (em pares de bases) nos 192 genótipos de *Eucalyptus grandis* avaliados com os microssatélites EMBRA4, EMBRA5, EMBRA10, EMBRA15, EMBRA16 e EMBRA11 (ND = dados não disponíveis)

Clone	EMBRA5		EMBRA4		EMBRA10		EMBRA11		EMBRA16		EMBRA15	
	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2
1	108	108	92	92	132	136	124	136	124	124	126	136
2	130	130	76	92	132	138	ND	ND	134	154	144	144
3	106	128	92	104	130	136	136	136	ND	ND	ND	ND
4	114	144	92	92	138	146	128	130	128	128	134	148
5	108	108	76	98	ND	ND	ND	ND	124	140	126	148
6	108	108	108	114	132	138	ND	ND	120	132	130	136
7	114	150	110	110	132	146	130	130	156	156	136	136
8	108	132	106	106	132	136	128	128	128	128	130	152
9	114	114	102	102	130	146	124	124	166	166	134	134
10	124	160	102	102	124	138	124	124	120	120	126	130
11	122	136	110	110	132	146	124	128	128	128	134	142
12	108	116	100	104	130	138	124	136	120	128	134	144
13	108	116	76	110	132	148	124	128	120	140	122	122
14	108	140	106	106	136	138	124	130	120	140	124	124
15	124	140	106	106	132	144	130	130	138	138	136	136
16	114	132	92	92	136	136	130	136	118	118	126	134
17	108	124	110	110	130	136	124	146	148	148	134	134
18	118	124	76	76	130	142	124	130	128	128	122	148
19	124	124	108	108	132	136	124	128	170	170	136	144
20	128	160	106	114	130	136	124	130	168	168	122	122
21	132	132	108	108	132	138	136	136	166	166	122	124
22	108	152	76	104	136	142	124	130	124	124	126	144
23	116	120	104	104	130	144	ND	ND	128	128	134	142
24	116	132	76	102	130	134	ND	ND	128	128	126	126
25	118	118	106	118	130	136	124	124	128	146	136	148
26	136	136	108	118	136	136	124	130	128	128	130	148
27	126	142	92	120	ND	ND	136	136	136	136	130	134
28	ND	ND	76	108	136	136	130	130	126	126	128	134
29	136	150	104	104	130	138	124	146	ND	ND	136	136
30	114	120	92	106	136	142	ND	ND	130	130	138	146
31	118	140	76	112	136	140	130	130	128	138	136	152
32	114	114	100	100	130	130	136	136	126	126	124	148
33	116	150	92	102	130	150	124	136	128	128	130	148
34	114	128	96	104	142	142	ND	ND	168	168	128	128
35	120	154	76	104	136	142	130	130	136	136	130	138
36	104	118	108	108	136	136	124	130	130	130	ND	ND
37	ND	ND	102	108	132	132	ND	ND	120	120	136	152
38	116	128	114	114	130	130	124	124	166	166	116	148
39	108	114	104	110	132	132	124	124	122	122	136	148
40	116	140	76	120	136	146	120	136	128	128	130	132
41	108	124	92	92	134	146	ND	ND	124	124	138	138
42	ND	ND	92	92	136	150	136	136	134	134	124	150
43	120	134	94	104	136	136	146	146	122	122	126	148
44	116	134	104	104	132	136	124	128	136	136	132	132
46	134	142	92	110	136	136	124	136	136	136	110	144
47	118	130	92	102	130	142	124	136	156	156	126	130

Continua...

Quadro 2C, Cont.

Clone	EMBRA5		EMBRA4		EMBRA10		EMBRA11		EMBRA16		EMBRA15	
	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2
48	118	118	92	106	132	132	ND	ND	110	110	124	142
49	122	122	76	92	136	136	124	130	132	132	138	148
50	136	136	104	104	122	130	124	130	154	154	124	136
51	136	150	108	108	122	130	130	130	128	194	136	148
52	128	140	94	108	116	140	124	124	ND	ND	124	150
53	136	136	106	106	136	140	136	136	ND	ND	130	144
54	132	132	92	92	138	138	124	124	174	174	130	148
55	132	132	104	104	138	146	136	128	128	136	122	122
56	ND	ND	76	92	136	136	124	130	116	116	130	130
57	114	136	76	102	132	146	124	124	156	156	136	136
58	116	128	100	110	130	140	124	124	136	136	136	136
59	124	124	98	104	136	142	124	128	168	168	134	134
60	120	130	96	108	124	142	128	128	128	128	130	134
61	114	120	104	114	136	142	ND	ND	138	138	124	124
62	124	150	104	104	142	142	124	124	120	120	134	134
63	ND	ND	76	92	136	142	124	128	112	120	144	144
64	114	140	98	108	132	142	130	130	128	128	144	144
65	118	132	100	100	134	144	ND	ND	128	128	144	144
66	118	118	96	116	146	146	124	128	132	132	130	148
67	128	140	96	108	132	138	124	130	112	128	134	134
68	120	142	104	104	132	132	124	130	170	170	130	130
69	108	136	96	108	136	136	124	124	168	168	130	134
70	120	160	92	100	130	136	124	124	134	134	144	144
71	150	160	96	96	130	130	124	124	120	154	ND	ND
73	108	108	96	104	136	136	124	136	ND	ND	ND	ND
74	108	108	74	92	132	142	ND	ND	128	128	130	144
75	118	150	78	106	136	136	124	124	144	144	124	136
76	130	142	110	110	142	142	124	124	140	140	136	144
77	114	136	100	112	132	142	124	124	ND	ND	122	122
78	ND	ND	104	116	ND	ND	ND	ND	126	126	128	136
79	146	146	110	110	136	142	124	124	120	154	130	148
80	134	134	92	98	136	146	124	128	120	130	134	134
81	114	122	92	102	136	148	124	124	130	130	130	134
82	140	140	96	114	136	142	124	128	140	140	122	122
83	122	138	76	76	138	138	124	130	168	168	ND	ND
84	136	152	110	110	132	142	124	124	168	168	134	144
85	128	128	102	102	132	136	124	128	128	128	130	138
86	134	160	108	108	136	136	130	136	126	126	122	130
87	140	140	110	110	142	142	124	124	130	130	130	130
88	142	142	94	100	136	136	128	128	168	168	134	134
90	142	142	100	106	122	130	130	130	128	128	130	130
91	132	152	108	108	116	142	124	130	120	134	130	134
92	140	140	92	92	132	132	124	136	ND	ND	130	134
93	124	124	94	104	132	142	124	128	134	134	134	144
94	108	142	102	110	132	94	124	128	158	158	116	136
95	152	160	116	116	136	146	124	136	122	122	130	130
96	134	134	92	92	130	130	124	124	134	146	132	144
97	120	132	106	114	136	136	130	130	142	142	124	138
98	130	160	74	110	136	136	130	130	114	146	126	144

Continua...

Quadro 2C, Cont.

Clone	EMBRA5		EMBRA4		EMBRA10		EMBRA11		EMBRA16		EMBRA15	
	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2
99	122	122	106	118	130	136	124	124	126	126	134	148
100	136	136	ND	ND	132	136	124	124	144	144	136	136
101	128	128	100	108	136	136	ND	ND	130	130	150	150
102	108	108	102	102	130	142	124	146	128	128	124	130
104	124	150	106	116	132	132	124	146	118	118	136	144
105	128	150	78	96	132	136	124	130	164	164	136	136
106	116	124	98	108	130	148	124	136	142	142	136	148
107	108	134	76	110	136	146	124	124	138	138	154	154
108	128	142	106	106	132	136	124	124	142	142	124	134
109	120	120	76	96	142	142	124	124	150	150	126	130
110	134	142	100	110	136	136	124	130	132	164	130	130
111	150	150	102	112	130	142	124	136	156	156	138	144
112	124	138	76	104	136	142	124	136	128	128	134	142
113	160	160	92	96	132	138	124	130	132	132	134	148
114	116	152	ND	ND	130	142	124	124	138	138	124	136
116	116	142	86	98	132	138	ND	ND	120	120	136	136
117	108	116	104	104	136	136	130	120	128	128	134	144
118	134	134	96	110	138	146	124	130	114	114	136	150
119	134	134	102	102	124	132	124	136	152	152	126	152
120	108	150	96	102	132	132	130	130	ND	ND	116	144
121	116	140	92	92	136	136	124	124	138	138	124	136
122	108	122	92	92	132	132	124	124	ND	ND	ND	ND
123	122	140	86	98	120	132	130	130	146	146	116	144
124	122	134	ND	ND	132	144	124	124	ND	ND	ND	ND
125	116	150	92	92	136	146	124	124	134	134	116	136
126	108	150	ND	ND	138	138	124	124	142	142	134	134
127	128	140	92	106	140	140	ND	ND	128	128	130	130
128	130	136	92	100	132	142	124	124	128	158	148	148
129	108	122	ND	ND	132	136	136	146	168	168	140	148
130	108	136	92	92	130	138	124	124	128	128	132	132
131	108	130	76	98	140	140	124	124	ND	ND	126	126
132	132	146	100	100	138	144	124	124	120	168	132	132
133	108	148	92	92	132	132	124	124	138	138	130	148
134	142	142	92	92	132	144	124	124	140	140	126	146
135	122	136	108	108	132	142	ND	ND	128	138	126	136
136	124	124	108	108	132	132	124	136	128	158	130	130
137	140	140	104	104	132	132	130	130	138	138	126	148
138	108	130	76	110	138	144	ND	ND	128	128	126	130
139	108	108	108	108	132	132	124	136	ND	ND	130	130
141	140	140	76	104	130	136	124	124	ND	ND	130	130
142	ND	ND	102	102	ND	ND	ND	ND	150	150	150	150
143	122	134	98	98	130	144	130	136	170	170	ND	ND
144	132	132	110	110	132	132	ND	ND	138	138	134	148
145	124	124	92	110	142	146	136	136	154	154	124	136
146	136	150	104	104	136	136	124	124	ND	ND	136	144
147	124	124	76	104	130	134	ND	ND	132	132	126	150
148	142	142	110	110	144	144	124	128	138	138	132	132
149	140	140	104	104	132	132	124	124	ND	ND	136	136
150	108	140	116	116	132	132	130	130	134	134	124	130

Continua...

Quadro 2C, Cont.

Clone	EMBRA5		EMBRA4		EMBRA10		EMBRA11		EMBRA16		EMBRA15	
	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2	Alelo1	Alelo2
151	124	134	104	104	146	146	124	130	134	134	134	134
152	122	142	100	110	136	142	124	128	140	140	126	146
153	110	128	104	110	130	144	ND	ND	122	122	136	144
154	122	122	ND	ND	134	142	130	136	124	124	134	148
155	108	122	104	104	130	140	130	130	142	142	124	130
156	108	108	98	114	136	136	124	130	138	138	ND	ND
157	148	148	92	104	132	132	130	136	120	120	132	142
158	110	130	76	76	132	142	124	124	128	128	142	144
159	146	146	104	108	136	148	136	146	128	128	126	134
160	ND	ND	106	106	ND	ND	124	124	128	138	128	150
161	108	128	92	92	138	144	136	136	136	166	130	134
163	108	120	92	106	142	142	124	124	140	170	122	134
164	108	142	ND	ND	136	136	130	136	ND	ND	136	136
165	134	134	96	96	132	132	124	124	158	158	130	144
166	110	128	ND	ND	130	144	124	124	170	170	ND	ND
167	108	150	110	110	136	142	124	124	134	134	136	136
168	128	138	ND	ND	124	124	120	136	ND	ND	138	138
169	108	134	108	108	ND	ND	130	136	132	132	ND	ND
170	108	136	92	110	130	138	124	124	126	138	130	136
171	126	160	112	171	130	130	124	124	136	136	142	142
172	128	128	112	112	140	146	124	124	ND	ND	132	132
173	142	142	108	108	136	144	130	130	140	140	136	136
174	128	140	104	104	132	136	124	124	136	136	118	142
175	108	136	106	106	130	136	124	136	128	128	138	150
176	138	138	94	94	130	144	124	124	106	162	ND	ND
177	106	132	102	106	132	144	124	130	120	120	136	146
178	116	142	92	92	130	140	136	146	138	138	130	148
180	136	140	76	92	132	132	128	128	134	134	124	130
181	108	152	100	104	134	140	ND	ND	126	144	148	148
182	130	130	96	104	136	138	ND	ND	136	136	124	136
183	108	130	76	96	130	130	124	136	138	138	136	136
184	124	124	92	92	132	142	ND	ND	128	128	126	126
185	120	140	76	102	132	142	ND	ND	132	132	132	148
186	122	130	96	100	132	132	ND	ND	128	128	130	130
187	140	140	92	108	132	140	ND	ND	120	152	124	124
188	120	130	76	100	136	136	124	124	120	120	ND	ND
189	106	122	102	112	136	136	124	124	136	136	110	142
190	136	146	102	102	130	130	124	124	136	136	128	132
191	140	140	98	98	148	148	136	146	120	156	144	144
192	108	148	92	96	130	130	124	130	132	132	136	136
193	108	132	78	100	138	142	ND	ND	156	156	148	148
194	106	132	108	108	134	134	124	124	128	158	134	142
195	108	132	108	108	132	138	124	136	156	156	ND	ND
196	108	138	106	112	132	132	124	130	ND	ND	136	136
197	138	142	92	92	130	138	124	130	144	144	144	144
198	124	124	78	78	150	150	ND	ND	136	146	136	146
199	108	130	104	104	130	130	130	136	136	136	136	136
200	108	108	100	110	142	142	124	124	122	132	122	132

Continua...

APÊNDICE D

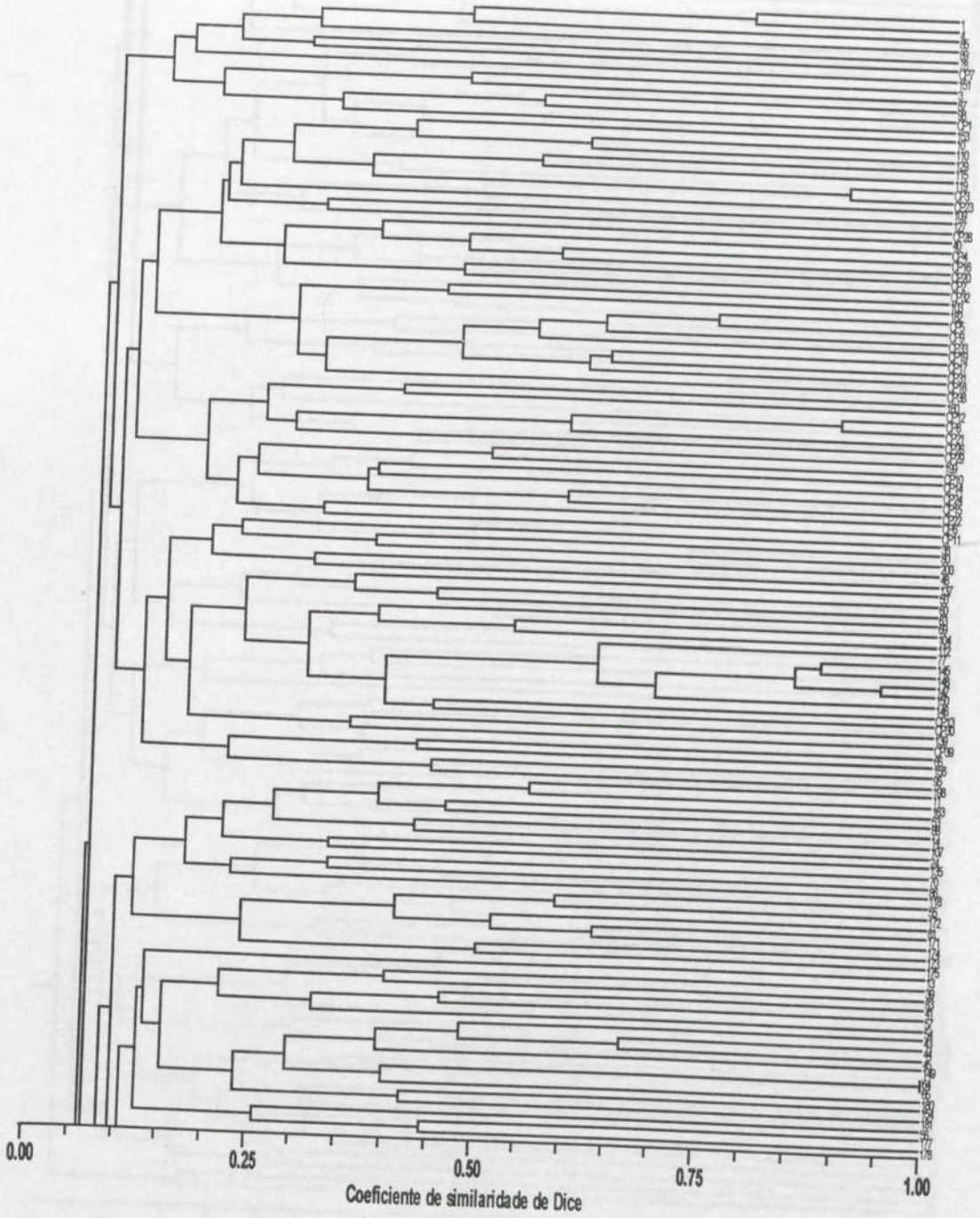


Figura 1D, Cont.

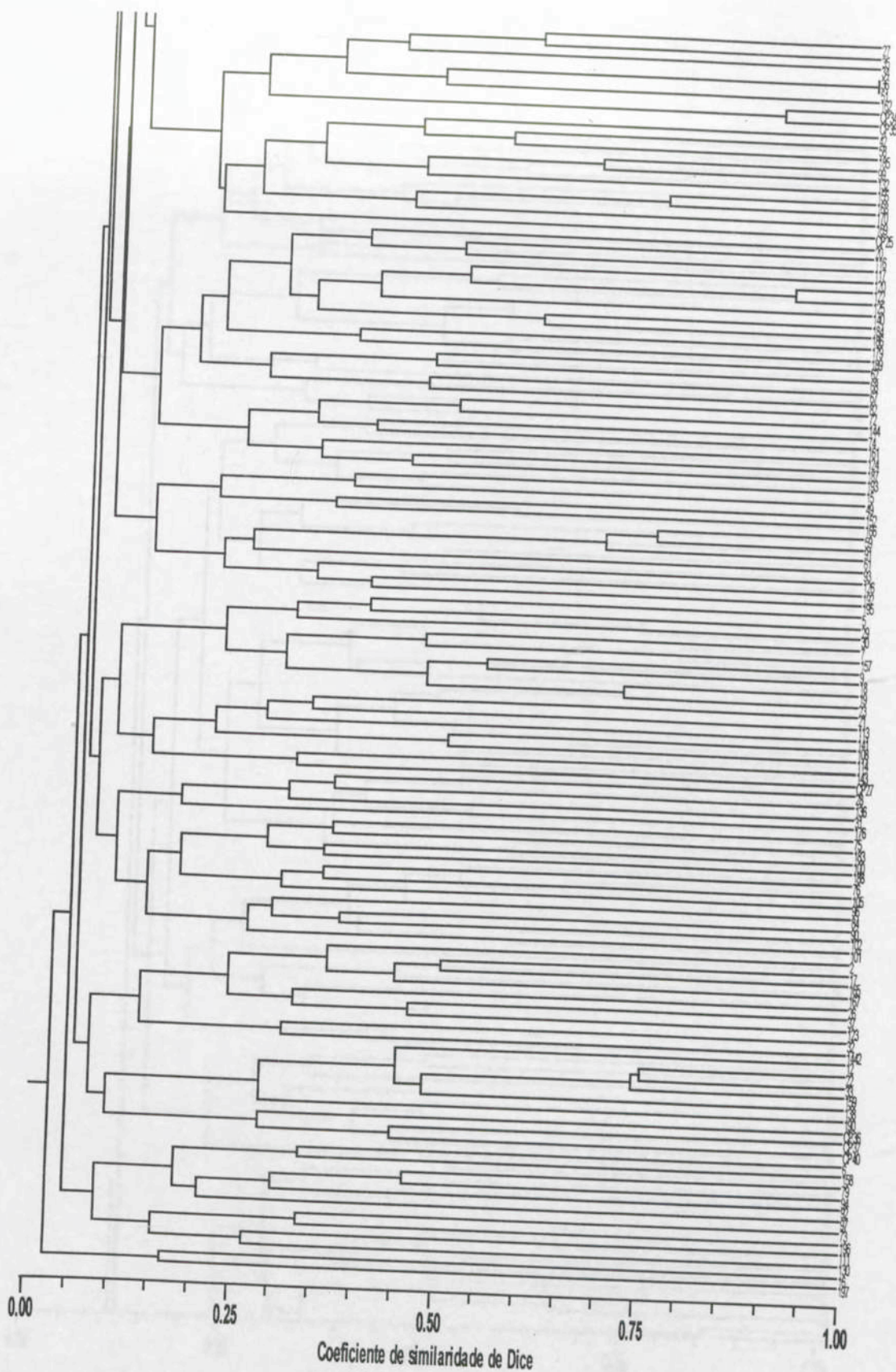


Figura 1D – Fenograma UPGMA gerado por meio dos coeficientes de similaridade de Dice entre 225 indivíduos de *E. urophylla*, analisados com cinco locos EMBRA.

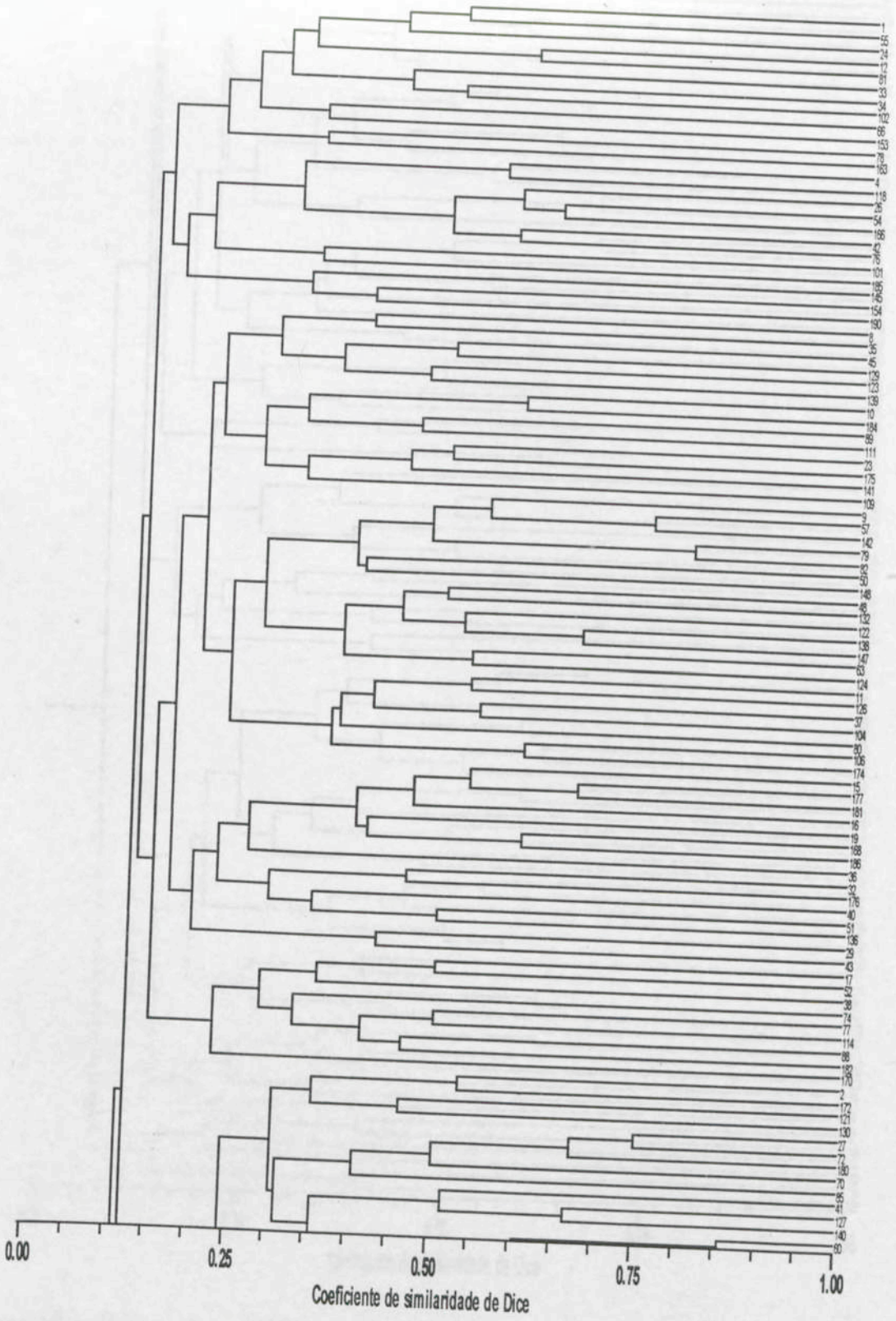


Figura 2D, Cont.

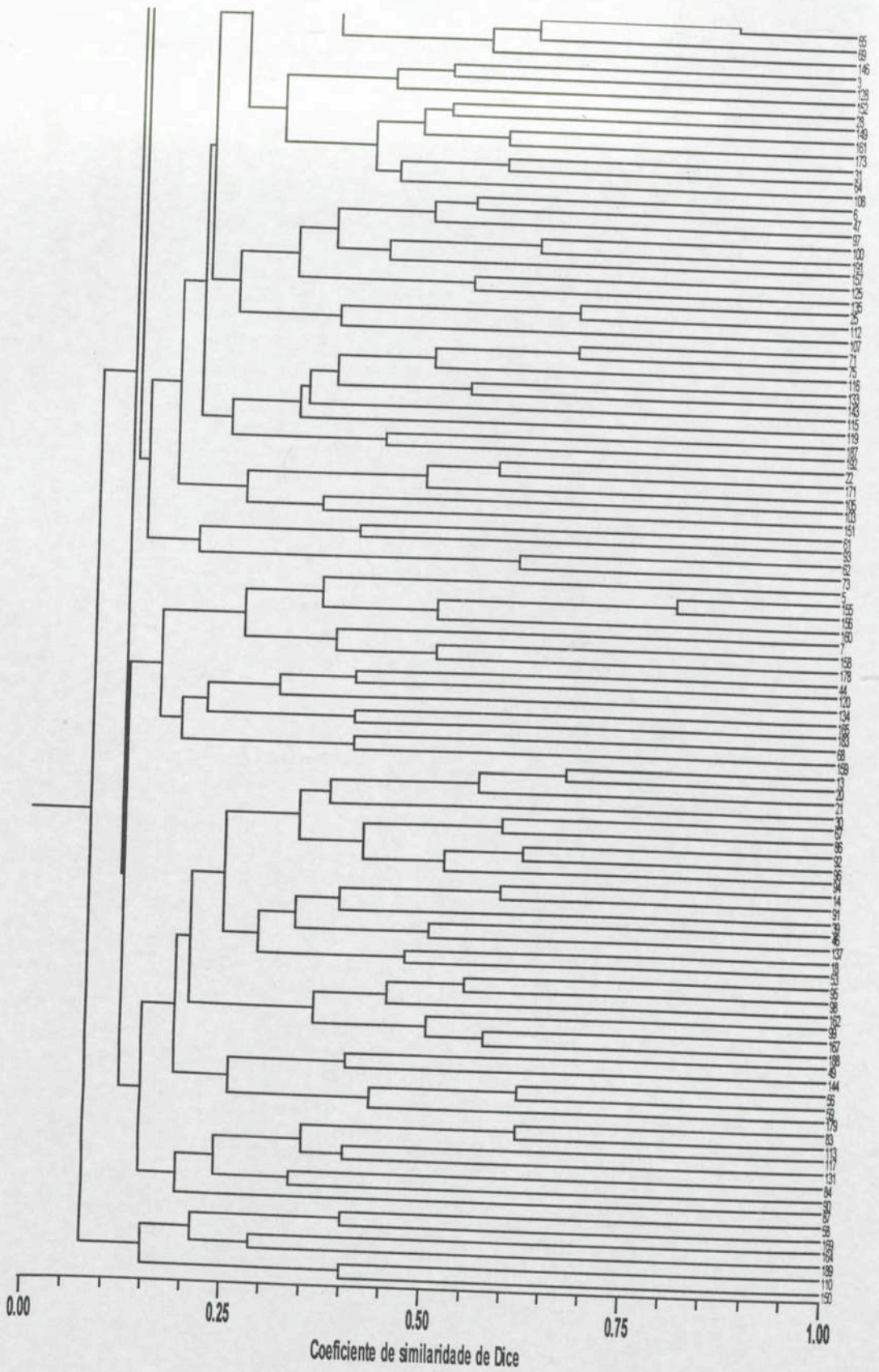


Figura 2D – Fenograma UPGMA mostrando os coeficientes de similaridade de DICE entre 192 indivíduos de *E. grandis*, analisados com seis locos EMBRA.