

Projeto GENOLYPTUS

*Rede Brasileira de Pesquisa do Genoma de Eucalyptus
construindo recursos genômicos para o melhoramento genético do eucalipto em
uma parceria público-privada*



Dario Grattapaglia - Coordenador

✓ EMBRAPA Recursos Genéticos e Biotecnologia e Universidade Católica de Brasília

Projetos integrados floresta - industria



Floresta

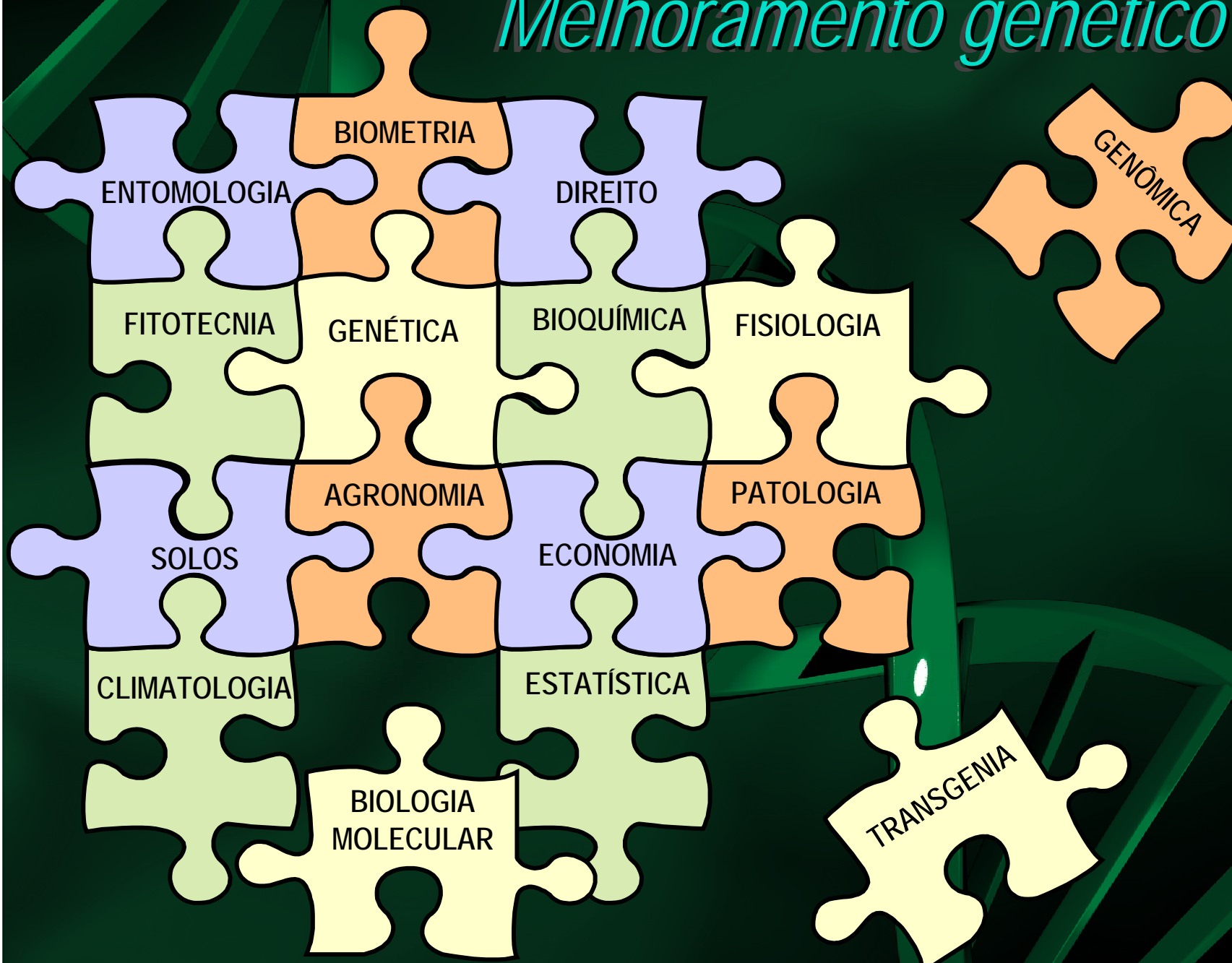


Fábrica

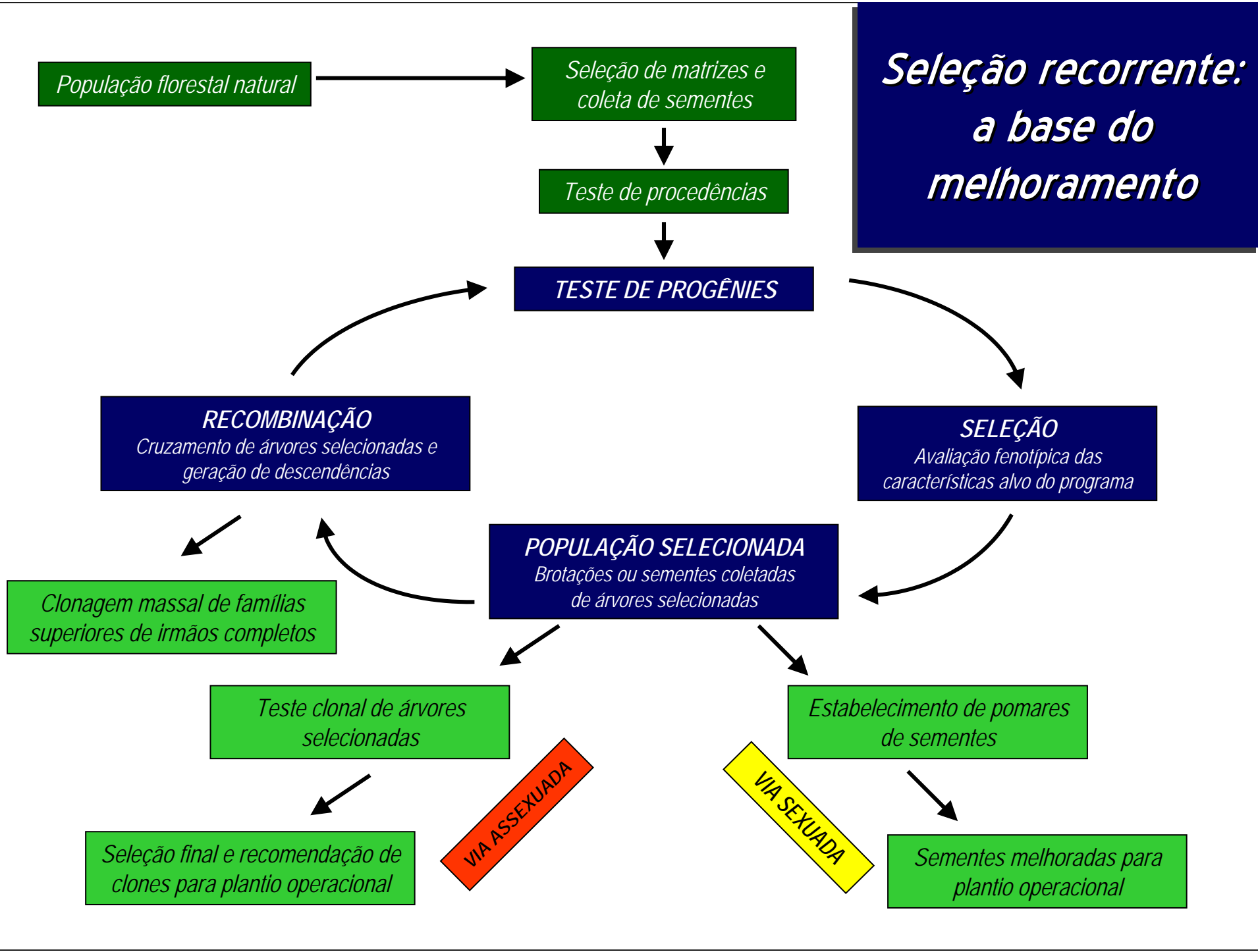


Transporte

Melhoramento genético



**Seleção recorrente:
a base do
melhoramento**



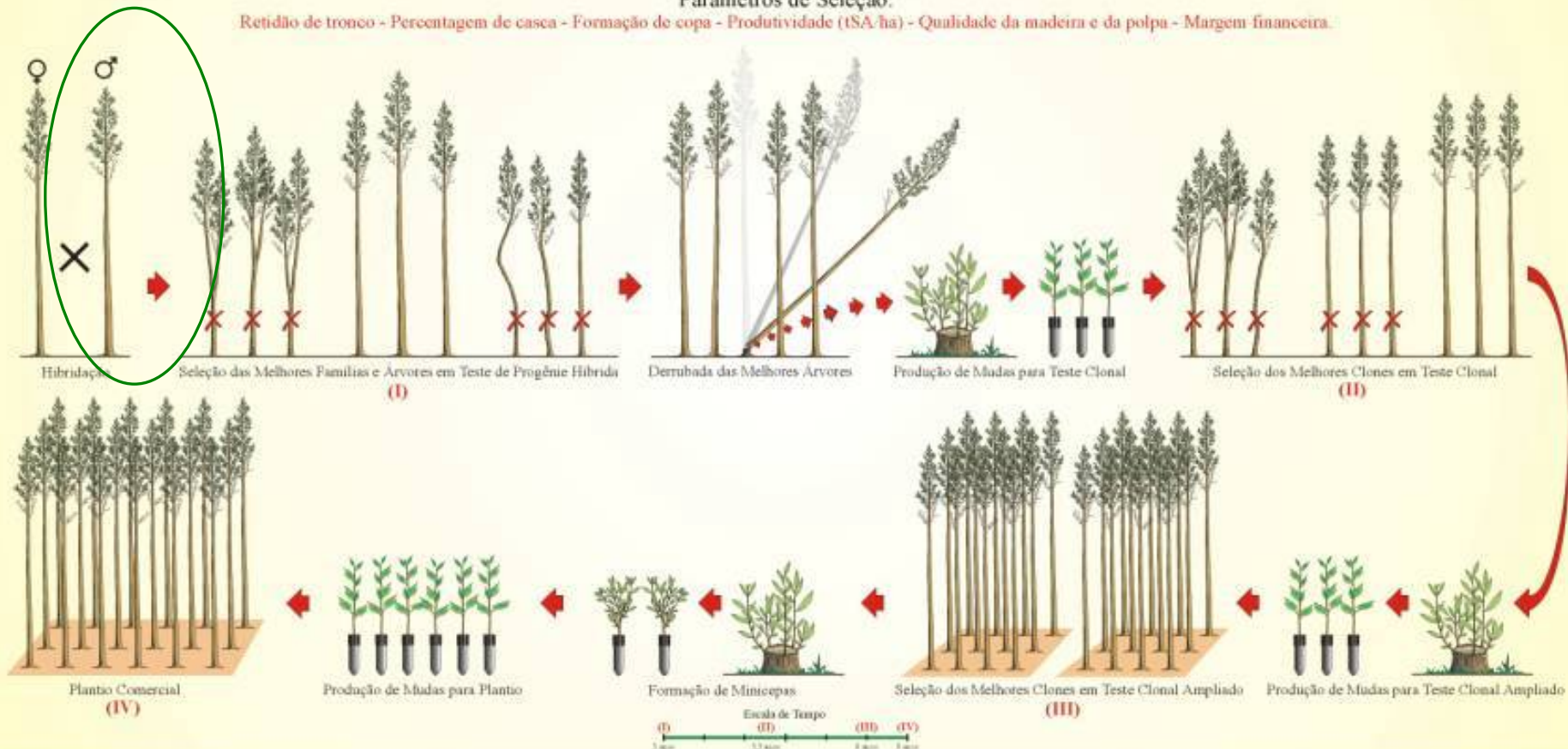
Seleção e clonagem de árvores elite

certificação de identidade e parentesco ao longo do processo

Seleção de Material Genético

Parâmetros de Seleção:

Retidão de tronco - Percentagem de casca - Formação de copa - Produtividade (tSA/ha) - Qualidade da madeira e da polpa - Margem financeira.



Brasil é líder na inovação e adoção de tecnologias para o melhoramento de *Eucalyptus*



Cruzamentos controlados



Testes de progênie



Seleção de árvores elite



Marcadores moleculares em apoio ao melhoramento



Plantios clonais de alta produtividade



Clonagem de árvores elite

Conceito de árvore como “fábrica” de celulose

- ✓ Redução do consumo específico de madeira para produção de celulose: m^3 de madeira por tonelada de celulose
- ✓ Aumento de produtividade florestal: ton de celulose/hectare
- ✓ Combinação de grande volume de madeira e baixo consumo específico de madeira garante o baixo custo desse insumo na composição dos custos industriais
- ✓ *Melhoramento genético é a chave para a competitividade*
- ✓ *Genômica é mais uma ferramenta para o melhorista*

Projeto genoma no melhoramento florestal

✓ Pré-requisitos para o sucesso de projetos genoma:

- ✓ *Grupo multidisciplinar de pesquisadores*
- ✓ *Material genético e populações experimentais adequadas ******
- ✓ *Recursos genômicos: sequências de genes, marcadores moleculares (SSR e SNPs), mapas genéticos, mapa físico, bioinformática*
- ✓ *Fenotipagem em larga escala: qualidade da madeira, resistência a doenças etc. ******
- ✓ *Abordagens experimentais integrativas de genética & genômica*
- ✓ *Interconexão com trabalhos intensivos de campo, genética quantitativa e estratégias de melhoramento*

Porque o projeto GENOLYPTUS?

Projeto Genolyptus



- ✓ Pesquisa em genética e melhoramento tem forte histórico de geração de ganhos e competitividade na indústria de base florestal
- ✓ Manutenção de vantagem competitiva do Brasil
- ✓ Existência de competência técnico-científica no setor privado
- ✓ Existência de competência científica nas Universidades reconhecida no setor privado
- ✓ Consórcio pré-competitivo: maneira mais eficiente
- ✓ Equipe multidisciplinar e parceria com empresas
- ✓ Projeto com visibilidade internacional – Iniciativa inédita no mundo

Pontos positivos que viabilizaram o projeto

- ✓ Pesquisa em genética e melhoramento tem forte histórico de geração de ganhos e competitividade na indústria de base florestal
- ✓ Existência de competência técnico-científica no setor privado
- ✓ Existência de competência científica nas Universidades reconhecida no setor privado
- ✓ Envolvimento dos cientistas das empresas desde o início da concepção do projeto
- ✓ Atribuição clara de responsabilidades no plano de trabalho
- ✓ Projeto na interface entre genômica e melhoramento: todos se relacionam com o que está sendo feito
- ✓ Perspectivas de crescimento e atualização profissional individual dos cientistas seja das Universidades como das empresas

Projeto Genolyptus



1. *Geração de um plataforma de informações e recursos genômicos*
2. *Estabelecimento de uma rede de populações experimentais para pesquisa genômica*
3. *Mapeamento, identificação e determinação de função de genes envolvidos na formação da madeira e resistência a doenças*
4. *Estabelecimento das bases para iniciativas empresariais visando a incorporação de tecnologias genômicas nos programas de melhoramento*

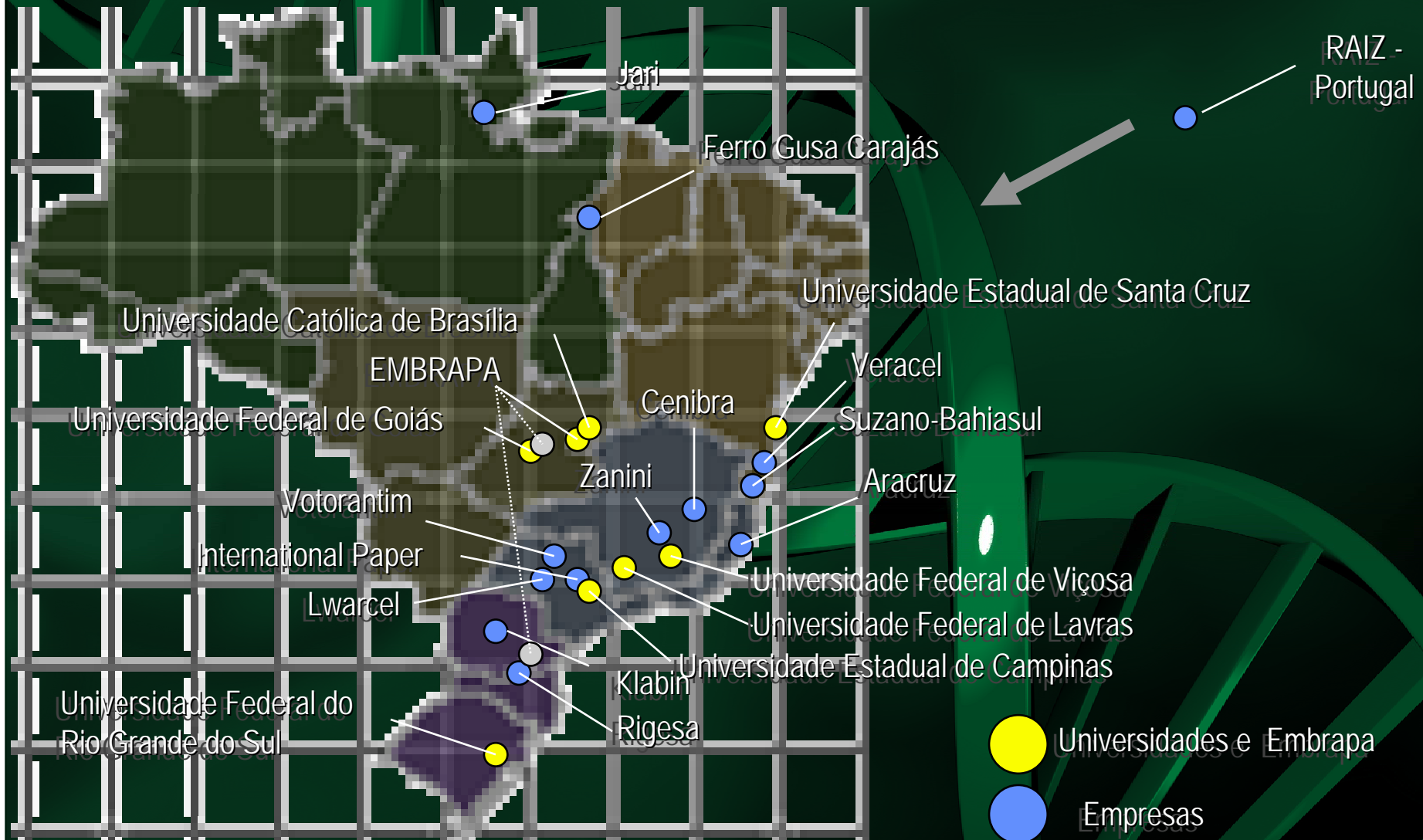


CT BRASIL
Ministério da Ciência e Tecnologia

35 ANOS FINEP
Financiadora de Estudos e Projetos

CNPq
CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO

Instituições participantes do projeto GENOLYPTUS





Rosana Brondani
Genetic mapping



Alexandre Coelho
Statistical genomics



Sergio Brommonschenkel
BAC library/physical mapping



Acelino Alfenas
Phytopathology



Julio Cascardo
EST sequencing



Giancarlo Pasquali
cDNA libraries/Microarrays



Georgios Pappas
Bioinformatics



Gonçalo Pereira
Bioinformatics



*José Livio Gomide &
Jorge Colodette*
Wood technology

Projeto Genolyptus



CT BRASIL
Ministério da Ciência e Tecnologia

35 ANOS FINEP
Financiadora de Estudos e Projetos

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico

Projeto lançado em
20/02/2002



Subprojetos componentes com forte interfaceamento entre eles

- ✓ Subprojeto 1: Instalação e avaliação continuada de uma rede experimental de campo
- ✓ Subprojeto 2: Internalização de tecnologias de alto desempenho para avaliação da qualidade da madeira
- ✓ Subprojeto 3: Base genética e identificação de genes que conferem resistência a doenças em *Eucalyptus*
- ✓ Subprojeto 4: Construção de mapas genéticos e mapeamento de QTLs
- ✓ Subprojeto 5: Construção de mapas físicos localizados no genoma de *Eucalyptus*
- ✓ Subprojeto 6: Sequenciamento do transcriptoma de *Eucalyptus*.
- ✓ Subprojeto 7: Análise de expressão gênica em microarranjos (microarrays)
- ✓ Subprojeto 8: Bioinformática para a análise, integração e disponibilização de dados genômicos
- ✓ Subprojeto 9: Genética estatística e desenvolvimento de ferramentas analíticas

Termo de compromisso firmado entre todas as instituições participantes

- ✓ Sugestões de todos os jurídicos - redação final pela SPRI da EMBRAPA
- ✓ Plano de trabalho detalhado com atribuições claras de metas e responsabilidades de cada instituição é anexo do Termo
- ✓ Conselho deliberativo: 20 membros votantes, cada um com 1 voto
- ✓ Propriedade intelectual: acesso restrito aos participantes, co-titularidade de patentes e material genético desenvolvido no projeto
- ✓ Possibilidade de desenvolvimentos competitivos ao longo do projeto
- ✓ Termo de confidencialidade e sistema ágil para aprovação de publicações
- ✓ Comitê externo de revisão e acompanhamento
- ✓ Mecanismos de avaliação de desempenho e desligamento do projeto

Estrutura do projeto

- ✓ **Estrutura técnico-científica**
 - ✓ Projeto organizado em nove subprojetos com coordenadores específicos
 - ✓ Intenso interfaceamento entre os subprojetos
- ✓ **Gerenciamento administrativo**
 - ✓ Contratação da Funarbe para controle contábil e de compras/importação
 - ✓ Termo de compromisso firmado entre todos os participantes
 - ✓ Conselho deliberativo com 20 membros representantes e comitê técnico
- ✓ **Acompanhamento, revisão e avaliação do projeto**
 - ✓ Comitê externo de revisão e acompanhamento
 - ✓ Reuniões semestrais internas e anuais externas
- ✓ **Transferência de tecnologia**
 - ✓ Cursos técnicos científicos semestrais por ocasião de reuniões
 - ✓ Workshops anuais e em eventos nacionais
- ✓ **Financiamento**
 - ✓ Orçamento de 10,1 milhões de reais
 - ✓ Empresas 37% , MCT 63%

Financiamento do projeto

- ✓ **Investimento direto: ~R\$ 10,1 milhões**
- ✓ MCT FINEP - Fundo Verde Amarelo: R\$ 5,0 milhões (equipamentos e consumo)
- ✓ CNPq (bolsas): R\$ 700k (bolsas DTI e bolsas ITI) por 24 meses
- ✓ Empresas:
 - ✓ ~ R\$ 3,0 milhões no total 14 empresas em 5 parcelas anuais
 - ✓ ~ R\$ 1,0 milhão em infra e pessoal para geração, instalação, mensuração e coleta de amostras nos experimentos de campo
- ✓ **Contratação da Fundação Arthur Bernardes, Funarbe (UFV) para gerenciamento contábil de recursos governamentais e das empresas, operações de compras e importação e contratação de pessoal temporário para o projeto**

Fenotipagem em população segregante

Mapa genético (SSR, AFLP, RAPD)

Mapa genético de alta resolução (AFLP, SNPs)

Mapa físico trilha mínima (Fingerprinting de BAC)

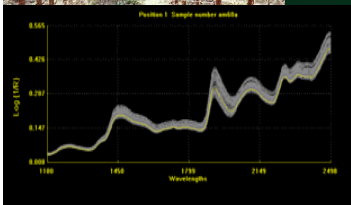
Identificação e validação do gene-alvo

Mapeamento de genes candidatos derivados de análises de bioinfo ou QTLs de expressão

Mapeamento de associação (*whole genome scan*, genes candidatos)

Sequenciamento *shotgun* de baixa cobertura e anotação

Testes funcionais via transformação genética (complementação, superexpressão, silenciamento)



QTL mapeado LOD >>3,0

LOD = 3,0

~ 10 cM

< 1,0 cM

~ 500 a 1000 kpb

Abordagem de genética direta (*forward genetics*) de fenótipos a genes

Variabilidade interespecífica em *Eucalyptus*

ESPÉCIE	AMI m³/ha/ ano	PULP YIELD (%)	DENS (g/cm³)	S5 (%)	SPEC. CONS. (m³/ton)	LIGNIN (%)
<i>E. grandis</i>	42	50	0,42	10	4,7	25,5
<i>E. tereticornis</i>	25	46	0,56	9	3,8	30
<i>E. saligna</i>	36	49	0,46	9	4,4	26
<i>E. urophylla</i>	30	49	0,48	11	4,2	28
<i>E. globulus</i>	25	54	0,55	15	3,3	22

Dados de Teotônio F. De Assis 2003

Eucalyptus globulus:

Paradigma de qualidade da madeira para celulose

- ✓ Melhor combinação de propriedades da madeira para papel e celulose
- ✓ Requer cerca de 25% menos madeira para produzir uma ton. de celulose
 - ✓ *E. grandis*: 3,89 m³/ton de celulose
 - ✓ *E. globulus*: 2,98 m³/ton de celulose
- ✓ Maior comprimento e espessura de fibra
- ✓ Maior conteúdo de holocelulose
- ✓ Maior flexibilidade de fibra
- ✓ Densidade na faixa de 550 kg/m³
- ✓ Menor necessidade de energia no processo industrial
- ✓ Melhor preservação das fibras no processo

GRANDIS X GLOBULUS

PORQUE SÃO TÃO DIFERENTES ??

- ✓ *Projeto se baseia na exploração da ampla variabilidade genética natural para propriedades da madeira existente no gênero*
- ✓ *Investigar as bases genéticas e moleculares das diferenças fenotípicas*
 - ✓ *Diferença em regiões codificantes: enzimas ?*
 - ✓ *Diferenças em regiões reguladoras : promotores, miRNA ?*
- ✓ *Entendendo as diferenças será possível explorar a variabilidade natural de forma mais direcionada no melhoramento*

Cruzamentos para geração de descendências no projeto GENOLYPTUS

	<i>E. grandis</i> Atherton Aracruz	<i>E. urophylla</i> Timor Cenibra	<i>E. globulus</i> K-Riocell	<i>E. dunni</i> K-Riocell	<i>E. camaldulensis</i> V-Mannesmann	<i>E. uro. x E. glob.</i> K-Riocell
<i>E. grandis</i> Coffs Harb. VCP	G1 x G2 (est. VCP x pó. AR)	G1 x U2 (est. VCP x pó. CE)	G1 x GL2 (est. VCP x pó. K-R)	G1 x D2 (est. VCP x pó. K-R)		G1 x (UxGL) (est. VCP x pó. K-R)
<i>E. urophylla</i> (Flores) IP	U1 x G2 (est. IP x pó. AR)	U1 x U2 (est. IP x pó. CE)	U1 x GL2 (est. IP x pó. K-R)	U1 x D2 (est. IP x pó. K-R)	U1 x C2 (est. IP x pó. V-M)	U1 x (UxGL) (est. IP x pó. K-R)
<i>E. globulus</i> K-Riocell	G2 x GL1 (est. AR x pó. K-R)	U2 x GL1 (est. CE x pó. K-R)			C2 x GL1 (est. V-M x pó. K-R)	
<i>E. dunni</i> Rigesa	D1 x G2 (est. RG x pó. AR)	D1 x U2 (est. RG x pó. CE)	D1 x GL2 (est. RG x pó. K-R)	D1 x D2 (est. RG x pó. K-R)		D1 x (UxGL) (est. RG x pó. K-R)
<i>E. camaldulensis</i> V-Mannesmann	G2 x C1 (est. AR x pó. V-M)	U2 x C1 (est. CE x pó. V-M)	C1 x GL2 (est. V-M x pó. K-R)	C1 x D2 (est. V-M x pó. K-R)	C1 x C2 (est. V-M x pó. V-M)	C1 x (UxGL) (est. V-M x pó. K-R)
<i>E. gran. x E. dunni</i> K-Riocell	(GxD) x G2 (est. K-R x pó. AR)	(GxD) x U2 (est. K-R x pó. CE)	(GxD) x GL2 (est. K-R x pó. K-R)	(GxD) x D2 (est. K-R x pó. K-R)	(GxD) x C2 (est. K-R x pó. V-M)	(GxD) x (UxGL) (est. K-R x pó. K-R)



Cruzamentos entre materiais genéticos elite de diferentes empresas

Rede de experimentos de campo do projeto GENOLYPTUS: *instalada em 2003*

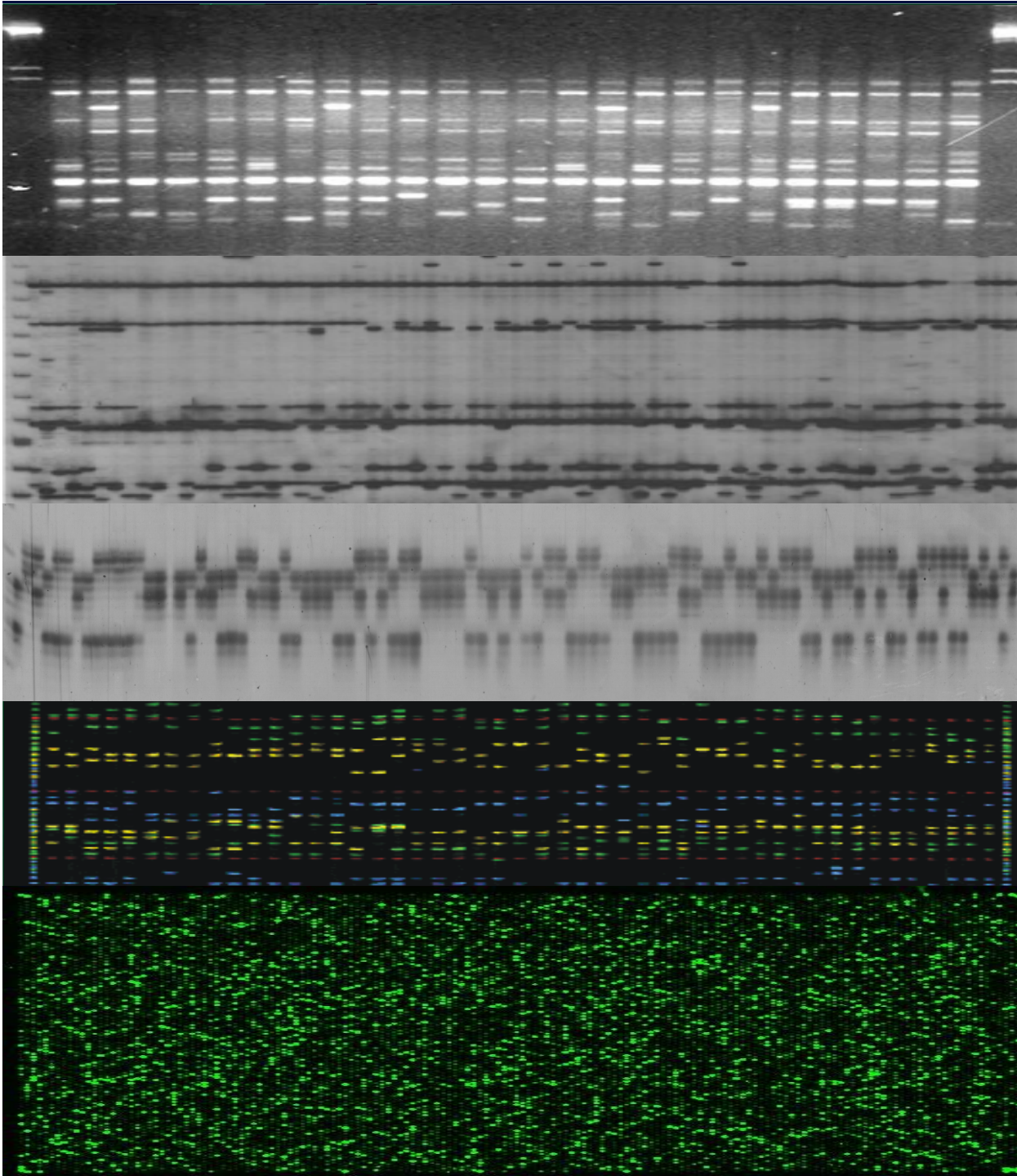


● Locais principais de instalação dos experimentos

● Locais complementares de instalação dos experimentos



Ampla variabilidade fenotípica para crescimento e qualidade da madeira



1990 - 1996 *RAPD*



1996 - 1997 *AFLP*



1997 - 2002 *microsats*



2002 - 2007

*Microsats com detecção
fluorescente*



*SNPs, SFPs, DArT,
Tecnologias genome-wide*

Clone 1

Clone 2

Clone 3

Clone 4

Clone 5

Clone 6

Clone 7

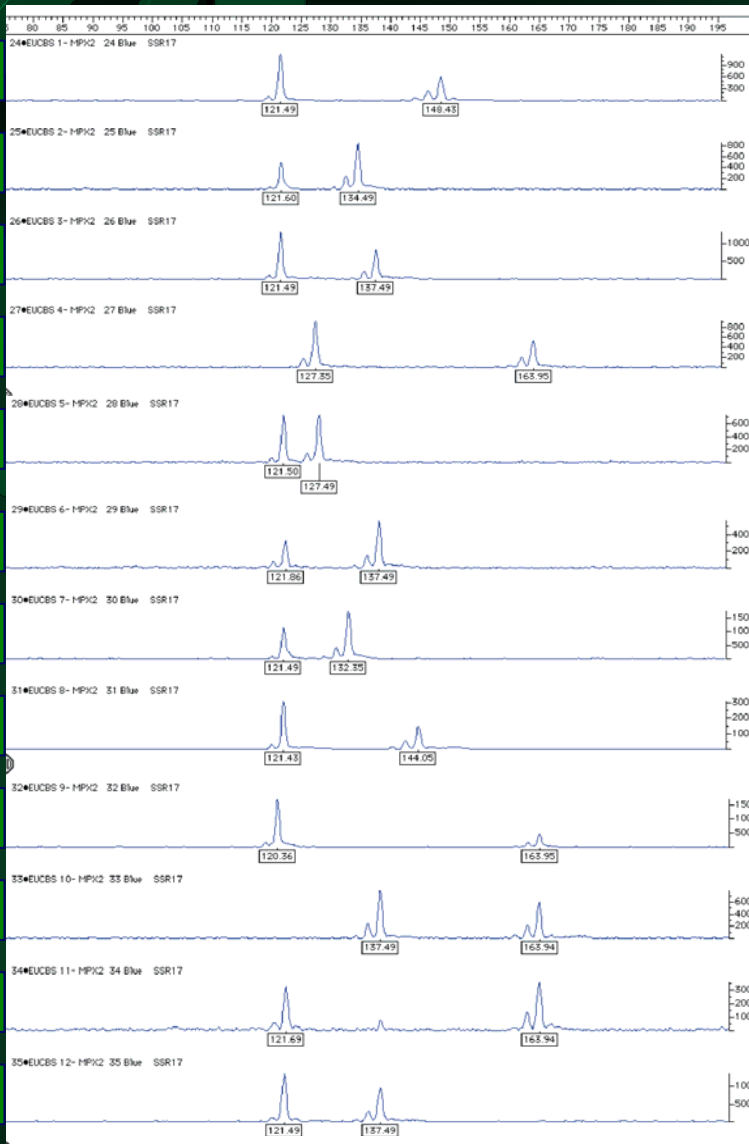
Clone 8

Clone 9

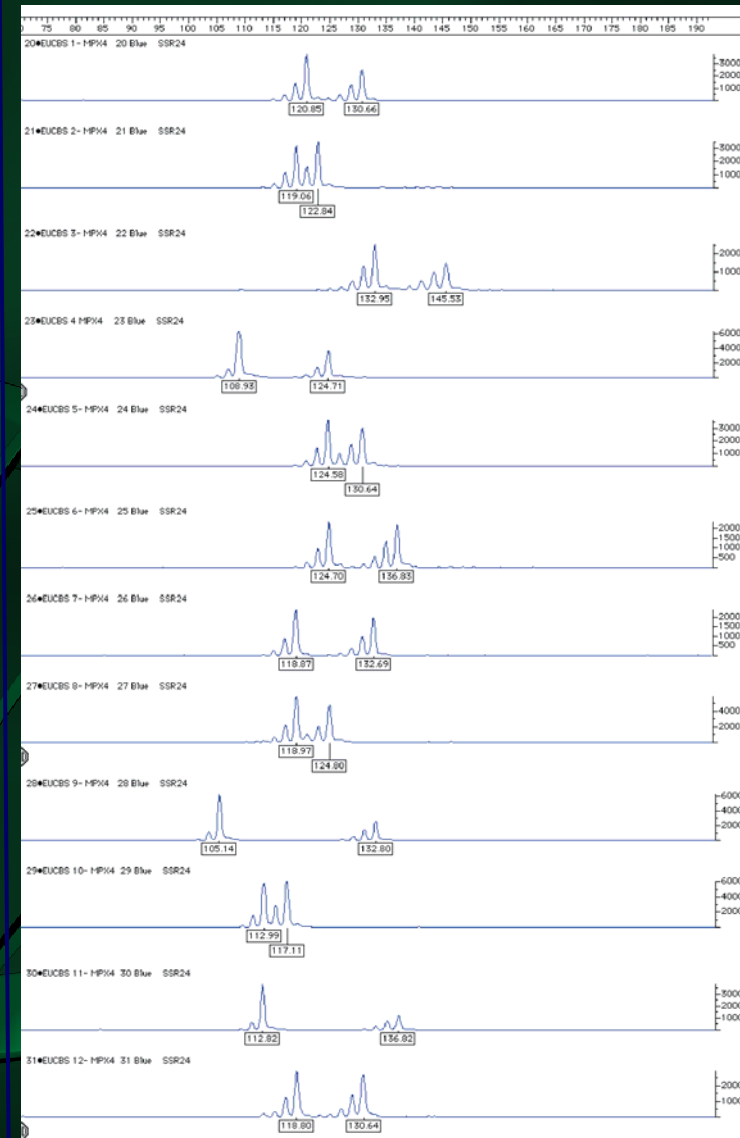
Clone 10

Clone 11

Clone 12



Microsatélite EMBRA21



Microsatélite EMBRA5

Identificação de clones de *Eucalyptus* com microssatélites

IMPrensa Nacional
A fonte oficial da informação
Governou Federal Casa Civil da Presidência da República

DIÁRIO OFICIAL - Nº24 - Seção

Imprimir matéria



Serviço Nacional de Proteção de Cultivares

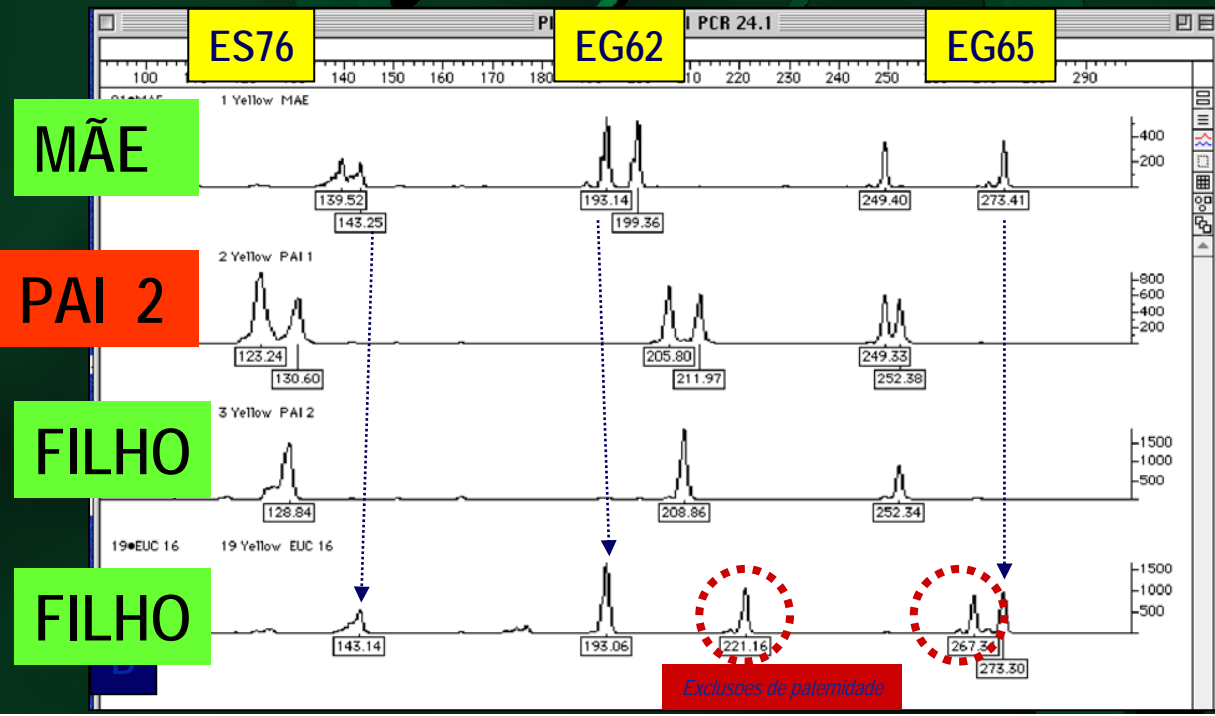
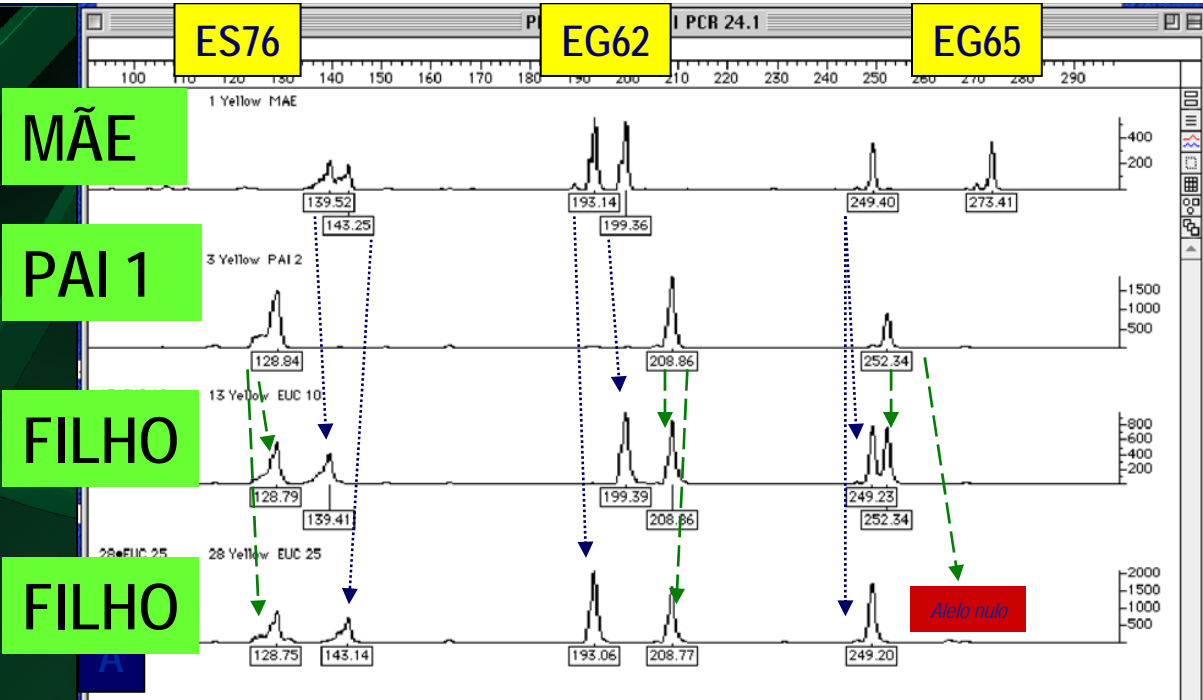
ATOS DE 11 DE JANEIRO DE 2002

Em cumprimento ao disposto no § 2º, do art. 4º, da Lei nº 9.456, de 25 de abril de 1997, e no inciso III, do art. 3º, do Decreto nº 2.366, de 5 de novembro de 1997, e no que consta do Processo nº 21806.000865/2001-60, o Serviço Nacional de Proteção de Cultivares divulga, para fins de abertura de pedidos de proteção de cultivares para a espécie EUCALIPTO, Gênero Eucalyptus, Sub-gênero Symphyomyrthus, Seções Transversaria, Exsertaria, Maidenaria, os descritores definidos na forma do Anexo I.

VIII. INFORMAÇÕES ADICIONAIS

1. MARCADORES MOLECULARES

Inclui-se marcadores moleculares como "Informações Adicionais", sendo que sua apresentação é facultativa. Sugere-se que o obtentor utilize os marcadores de acordo com metodologia a ser publicada pelo SNPC, de maneira a permitir a comparação de resultados.



Verificação de parentesco em cruzamentos controlados

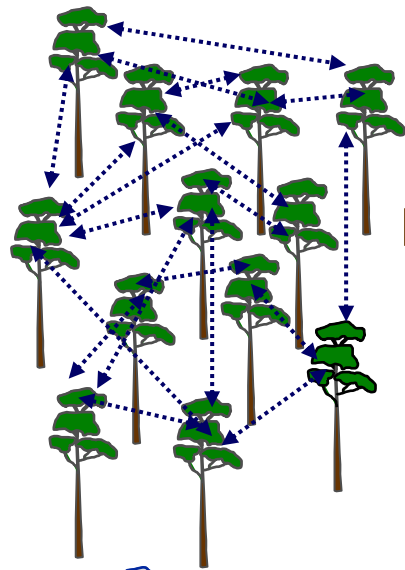
Responder com alterações... Finalizar revisão...

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

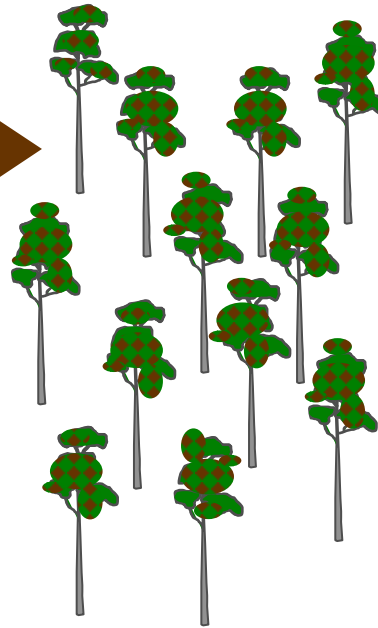
Arial Narrow 14

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	AMOSTRA	IDENTIFICAÇÃO			EMBRA2	EMBRA 28	EMBRA 3		EMBRA 11	EMBRA 128	EMBRA 38					
2	1	M1-GRA 277 MÃE			113,48	119,03	197,73	203,91	138,58	138,58	119,10	130,03	109,68	138,34	97,04	97,04
3	2	P1-URO 225 PAI			113,40	132,10	189,75	208,70	122,40	179,74	116,54	125,25	119,61	125,12	96,53	114,28
4	3	1.1 FILHO 01			113,55	119,13	189,20	204,04	138,87	179,19	119,23	125,89	124,95	137,51	95,63	97,18
5	4	1.2 FILHO 02			113,06	131,09	203,62	208,90	137,98	178,92	118,88	124,65	126,05	137,00	96,91	114,63
6	5	1.3 FILHO 03			113,67	133,04	197,38	209,05	136,95	178,65	119,43	125,90	109,57	126,71	98,78	114,10
7	6	1.4 FILHO 04			113,69	118,76	197,20	208,34	138,27	179,10	115,73	120,53	126,42	137,05	96,56	96,56
8	7	1.5 FILHO 05			112,69	118,09	190,01	204,00	122,65	138,58	118,49	126,21	120,49	137,45	96,53	96,53
9	8	1.6 FILHO 06			113,51	119,13	203,95	208,56	122,29	138,87	120,05	125,54	109,79	126,79	96,64	96,64
10	9	1.7 FILHO 07			119,14	132,26	189,24	203,71	122,61	138,58	116,58	125,37	109,60	119,78	97,35	113,94
11	10	1.8 FILHO 08			113,08	119,48	189,18	203,05	138,72	179,44	117,13	119,13	125,73	137,40	98,19	114,11
12	11	1.9 FILHO 09			113,55	132,13	203,28	208,53	122,65	138,97	124,00	129,98	109,49	126,62	96,74	114,36
13	12	1.10 FILHO 10			114,54	114,54	189,70	203,70	122,65	138,80	119,14	125,05	126,05	137,73	96,53	96,53
14	13	M2 URO 134 MÃE			109,44	126,76	195,89	211,71	122,65	179,01	132,61	148,61	124,49	124,49	110,08	110,08
15	14	P2-GRA 719 PAI			140,59	140,59	189,98	203,17	138,16	138,16	127,45	146,22	107,15	114,83	95,01	104,52
16	15	2.1 FILHO 1			108,65	141,22	195,43	203,32	138,32	178,87	132,61	146,80	108,09	124,69	95,99	109,01
17	16	2.2 FILHO 2			125,47	141,69	189,62	195,40	122,75	139,08	128,25	148,61	113,62	121,71	103,03	109,01
18	17	2.3 FILHO 3			108,87	141,87	197,73	204,21	139,11	179,40	131,20	145,05	113,82	125,70	94,98	110,25
19	18	2.4 FILHO 4			108,73	141,15	197,73	203,32	122,50	138,80	128,07	148,40	107,55	124,70	104,31	110,08
20	19	2.5 FILHO 5			108,74	141,17	189,97	195,97	122,52	138,68	131,23	145,90	107,57	124,92	95,14	110,58
21	20	2.6 FILHO 6			125,33	141,75	189,51	195,28	122,73	139,11	126,25	148,50	107,11	124,46	105,90	110,58
22	21	2.7 FILHO 7			126,10	140,36	196,58	202,89	123,56	139,52	126,74	148,54	107,17	124,40	104,44	110,75
23	22	2.8 FILHO 8			125,14	141,90	196,72	204,07	122,84	138,96	128,63	148,65	108,09	125,47	105,55	110,76
24	23	2.9 FILHO 9			108,69	141,10	196,45	203,92	138,89	178,86	132,36	145,68	107,56	123,99	105,11	110,81
25	24	2.10 FILHO 10			126,29	141,14	196,45	202,73	122,80	138,50	131,26	146,50	114,89	125,05	94,33	110,08
26	25	M3- URO 141 MÃE			108,71	108,71	192,05	192,05	130,78	130,78	131,50	134,49	126,35	126,35	96,36	96,36
27	26	P3- GRA 245 PAI			136,90	136,90	197,45	202,65	122,44	137,35	120,35	125,83	107,52	107,52	106,49	108,42
28	27	3.1 FILHO 1			108,56	138,24	192,31	202,73	131,05	137,02	125,71	131,76	107,58	126,90	96,50	108,48
29	28	3.2 FILHO 2			109,12	138,93	191,73	197,86	123,84	136,38	120,17	134,49	107,24	126,58	96,50	108,62
30	29	3.3 FILHO 3			108,76	137,24	191,16	196,76	131,87	136,19	126,28	134,29	107,67	126,58	96,51	108,44
31	30	3.4 FILHO 4			108,58	137,06	192,59	203,08	122,47	130,10	126,02	133,98	107,71	127,14	96,75	108,62
32	31	3.5 FILHO 5			108,62	137,24	191,14	197,17	130,88	137,13	125,79	133,74	107,55	126,93	96,50	108,45
33	32	3.6 FILHO 6			108,48	137,80	191,57	197,45	122,60	130,70	126,13	133,96	107,07	126,78	96,64	108,62

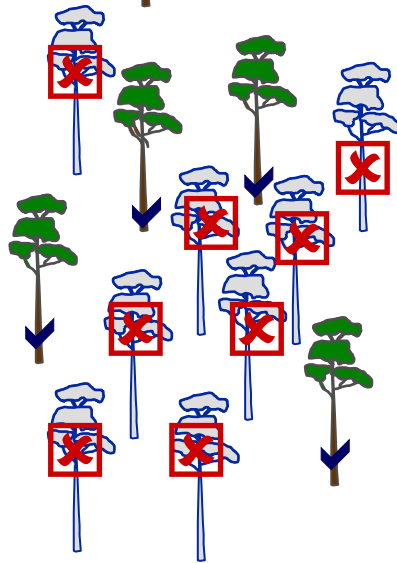
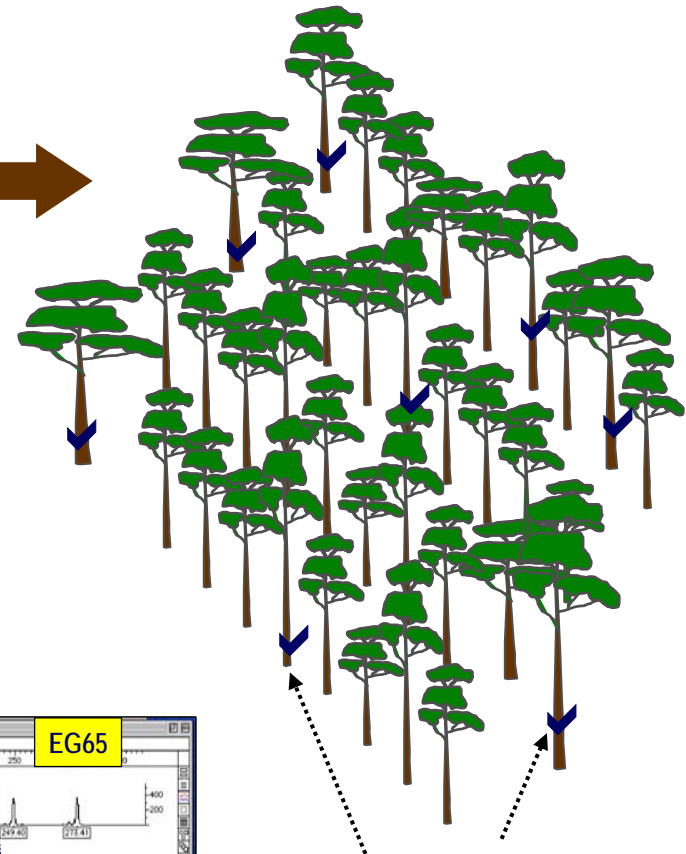
Pomar de sementes



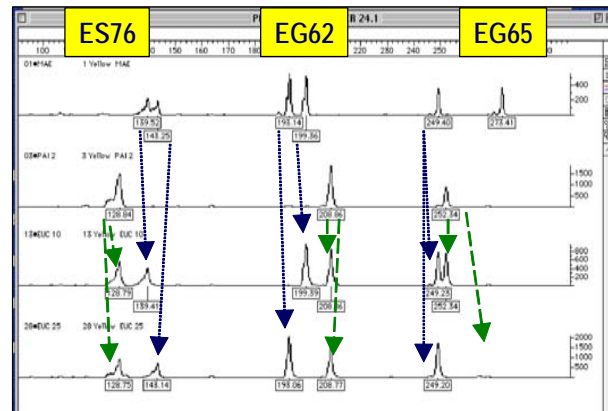
Colheita de sementes no pomar, separadamente por genitor materno (controle materno)



Teste de progênies ou plantio comercial em talhões de famílias de meios-irmãos



Seleção retrospectiva de genitores do pomar de sementes que apresentam elevada capacidade específica de combinação para cruzamentos controlados ou eliminação de árvores de baixa capacidade de combinação do pomar



Teste de paternidade de árvores superiores para a(s) característica(s) de interesse e identificação dos seus pais

Identificação de árvores superiores para a(s) característica(s) de interesse

Thore Appl Genet (2004) 109:192–199
DOI 10.1007/s11292-004-9117-9

ORIGINAL PAPER

D. Grattapaglia · V. J. Ribeiro · G. D. S. P. Rezende

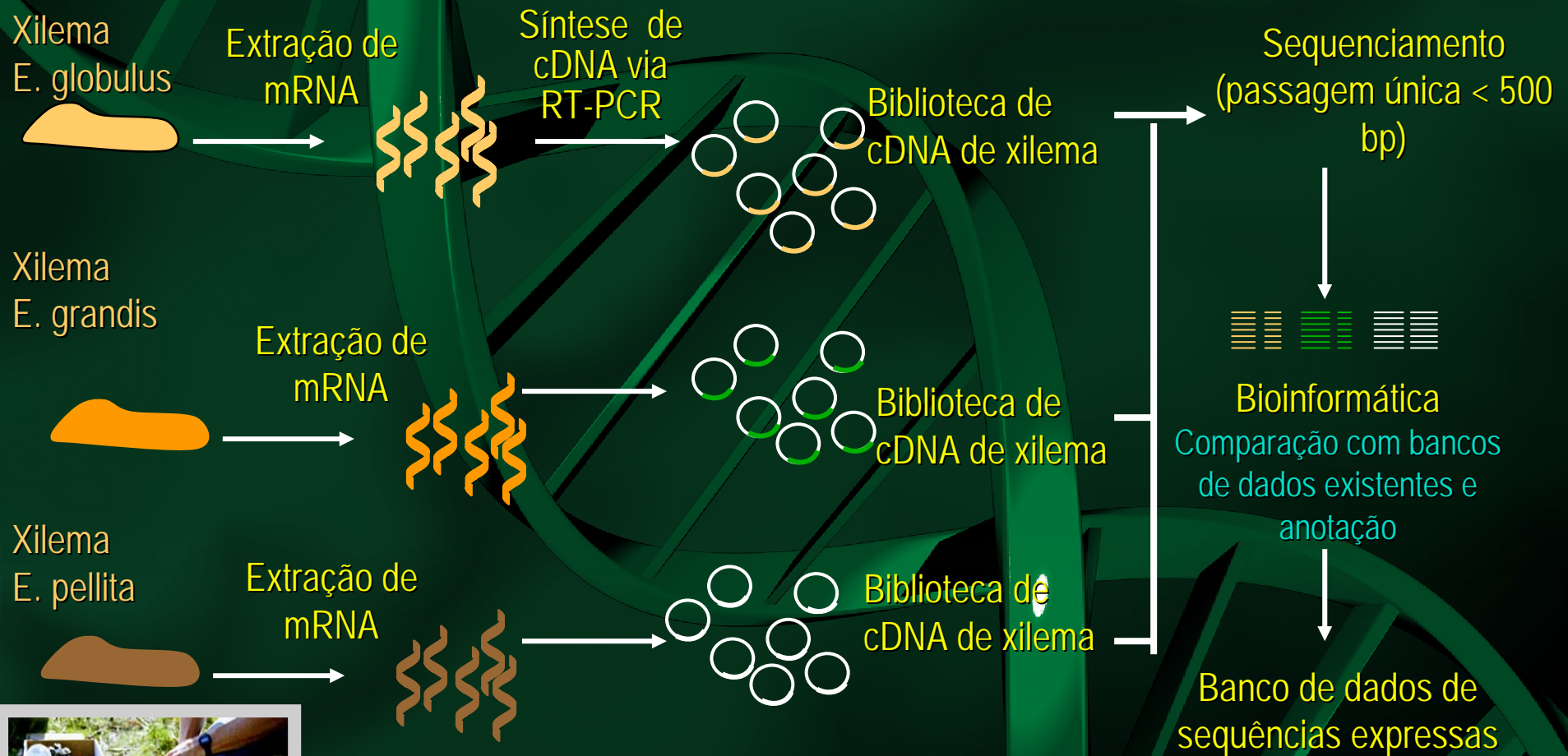
Retrospective selection of elite parent trees using paternity testing with microsatellite markers: an alternative short term breeding tactic for *Eucalyptus*

Programa operacional de diversas empresas brasileiras

Gerenciamento de diversidade e certificação de identidade genética e parentesco

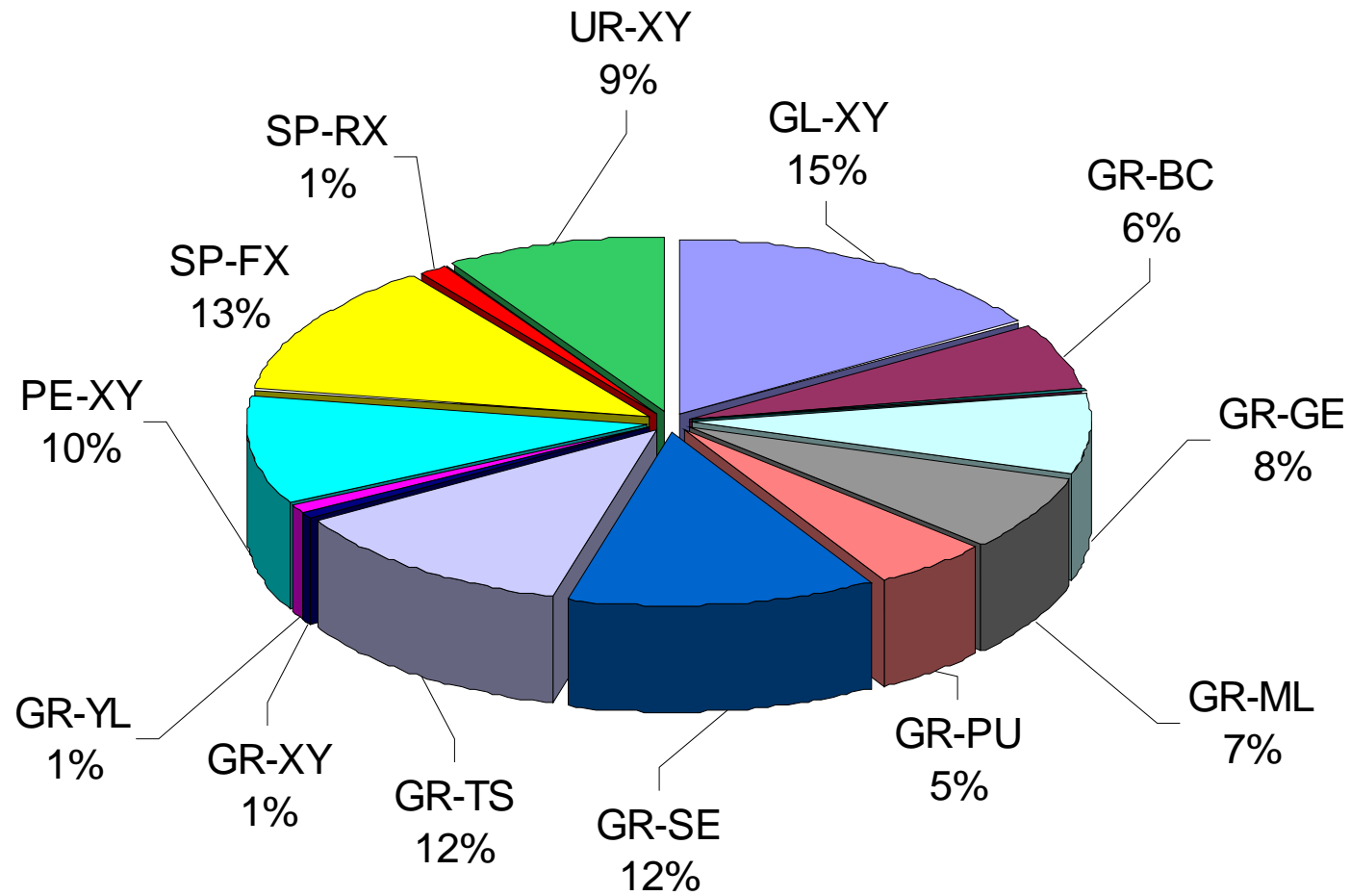
- ✓ Caracterização de diversidade, estrutura e relacionamento genético entre indivíduos das populações de melhoramento
- ✓ Certificação de identidade de coleção de matrizes nos pomares de hibridação
- ✓ Monitoramento de tamanho efetivo (N_e) com a seleção nos testes de progênies puras (famílias de MI) para o avanço das gerações
- ✓ Certificação de parentesco dos testes de progênies híbridas
- ✓ Monitoramento do tamanho efetivo populacional (N_e) no programa de conservação genética das espécies puras (*E. grandis* e *E. urophylla*)
- ✓ Determinação de parentesco de árvores excepcionais selecionadas em testes de progênies de famílias de meios-irmãos, famílias derivadas de polymix multiespécies ou em plantios comerciais cujas sementes são oriundas de pomares de sementes da empresa
- ✓ Monitoramento e certificação da identidade genética de propágulos durante testes clonais e propagação comercial
- ✓ Verificação de identidade em processos de registro e proteção de clones e contratos de propagação, compra e intercâmbio de clones

Formação da madeira : **variabilidade** "transversal" de genes expressos

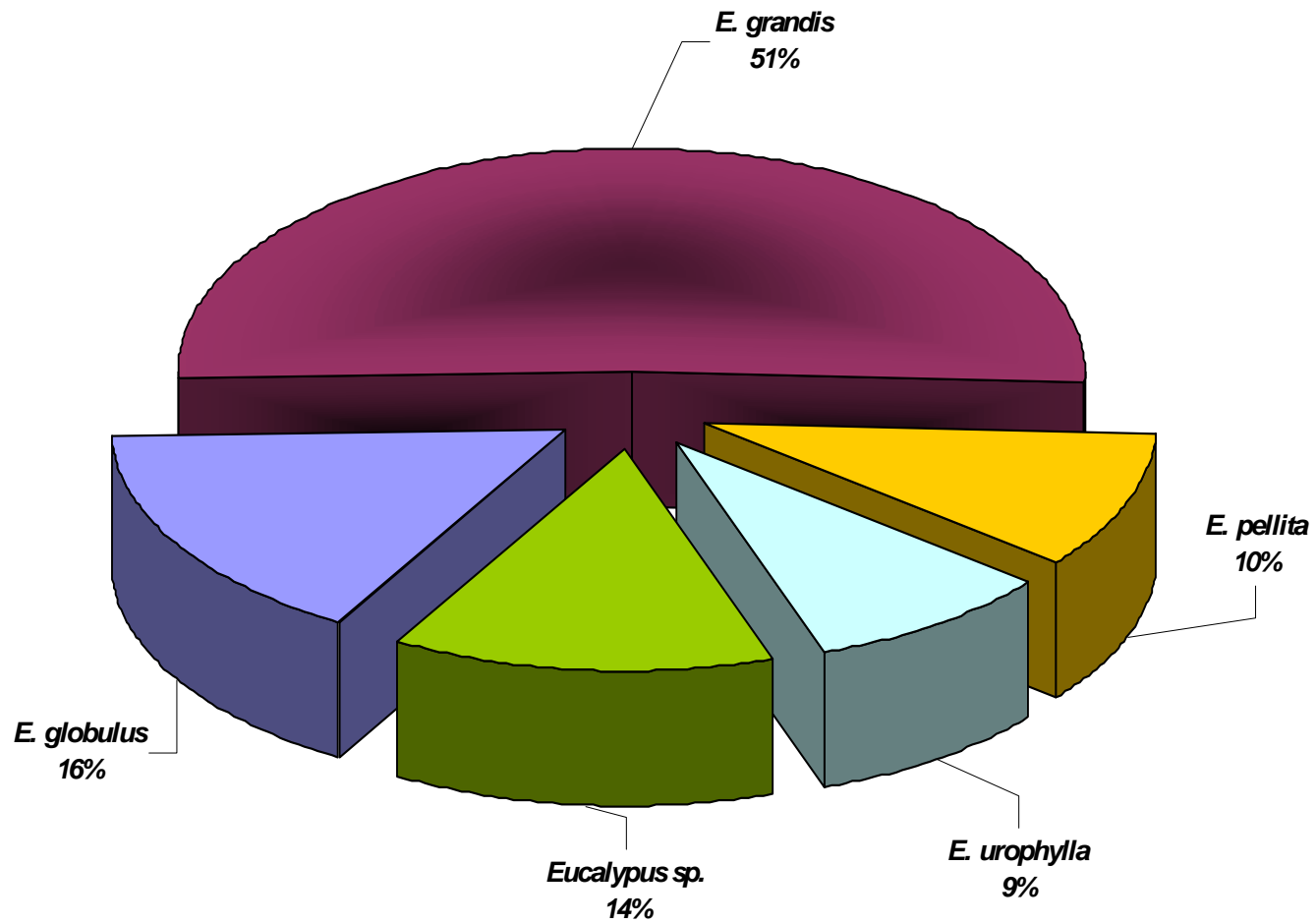


.... e outras espécies com propriedades físico-químicas contrastantes

Genolyptus libraries



Sequences by species



Seqüenciamento *GENOLYPTUS*

ESTs

Total	124.851
Aceitas	96.493 (77,3%)
<i>Full-length</i>	18.232

Seq. Genômicas

Total	10.298
Aceitas	6.432 (62,5%)

Extrem. BACs

Total	25.751
Aceitas	18.739 (72,8%)

121.664

<i>Clusters</i>	11.660
<i>Singlets</i>	29.089
Seq. únicas	40.749
ESTs únicas	21.442 (21.905)

Recursos genômicos para *Eucalyptus*

Expressed sequence tags

O impacto das novas tecnologias na geração de recursos genômicos

GENOLYPTUS EST database (Brazil)

- ✓ produzido entre 2002-2004
- ✓ 4 espécies ; 11 bibliotecas
- ✓ 124,851 reads ~ 50 Mbp
- ✓ 21,442 consensi
- ✓ A ser publicado em 2009
- ✓ 1.5 ano – 1.5 milhões de US\$

UNIVERSITY OF FLORIDA (Kirst lab)

- ✓ Produzido em 2007
- ✓ Pool de 21 árvores de *E. grandis*
- ✓ 1,024,251 reads 148.4 Mbp
- ✓ 29,000 consensi
- ✓ 33.742 SNPs descobertos
- ✓ 1.5 meses – 30 k US\$

Estudo comparativo
Sanger x 454 sequencing

BMC Genomics



Research article

Open Access

High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome

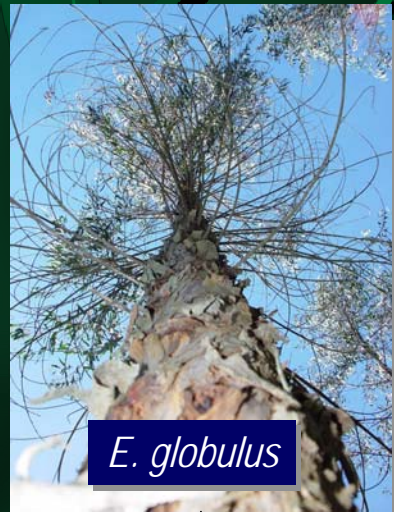
Evandro Novaes^{†1}, Derek R Drost^{†1,2}, William G Farmerie^{3,4},
Georgios J Pappas Jr^{5,6}, Dario Grattapaglia^{5,6}, Ronald R Sederoff⁷ and
Matias Kirst*^{1,2,4}

Address: ¹School of Forest Resources and Conservation, University of Florida, PO Box 110410, Gainesville, USA, ²Plant Molecular and Cellular Biology, University of Florida, Gainesville, USA, ³Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, USA, ⁴University of Florida Genetics Institute, University of Florida, Gainesville, USA, ⁵Graduate Program in Genomic Sciences and Biotechnology, Universidade Católica de Brasília, Brasília, Brazil, ⁶EMBRAPA Recursos Genéticos e Biotecnologia, Empresa Brasileira de Pesquisa Agropecuária, Brasília, Brazil and ⁷Department of Genetics, North Carolina State University, Raleigh, USA



Diferenças entre tecidos

Diferenças entre genótipos



Diferenças entre espécies

GENOLYPTUS *Chip*
Experimento piloto



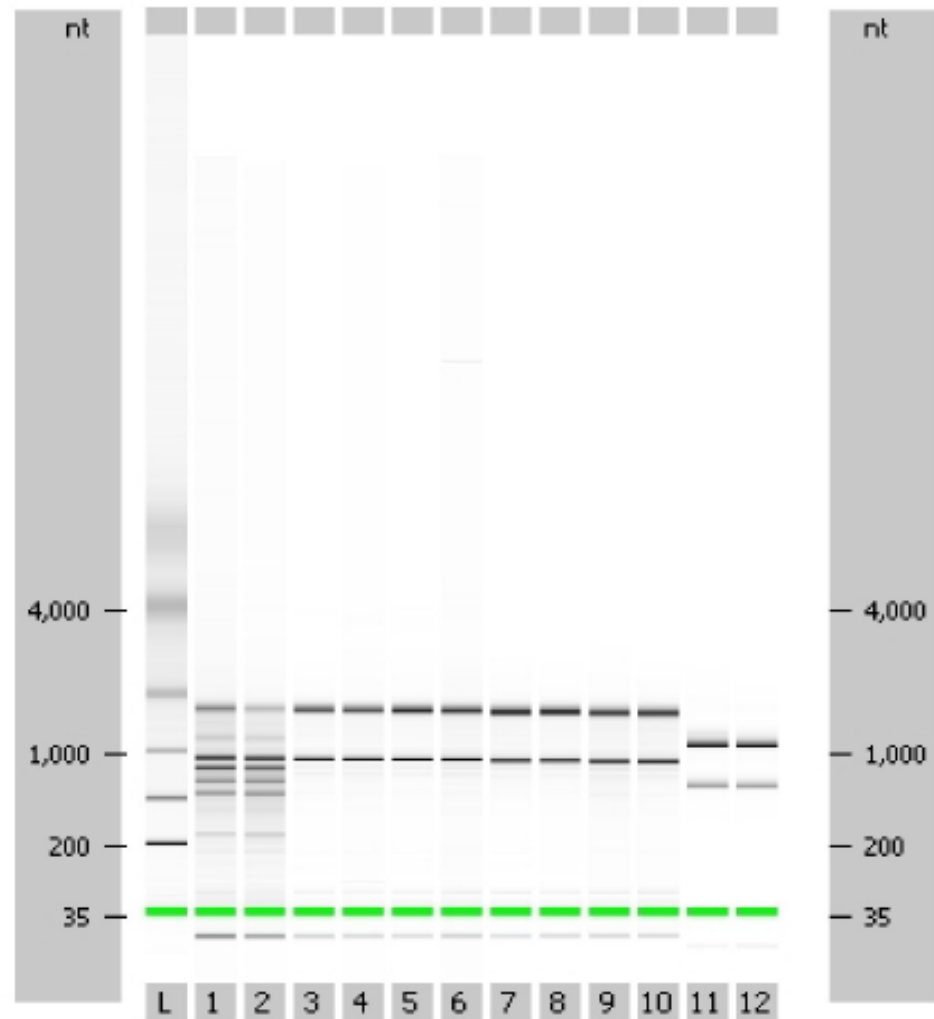
GENOLYPTUS *Chip Experimento piloto*

NimbleGen Systems, Inc.
Reykjavik, Iceland

- *"On-chip" síntes de sondas – 50mer*
- *Até 9 sondas/unigene contig*
- *Sondas replicadas (2X)*
- *21,442 sequências ("unigenes")*
- *385,856 features por chip*
- *10 chips idênticos*
- *Cy3-labelled cDNA*
- *Hibridização*
- *Lavagem*
- *Escaneamento*
- *Coleta de dados e normalização*

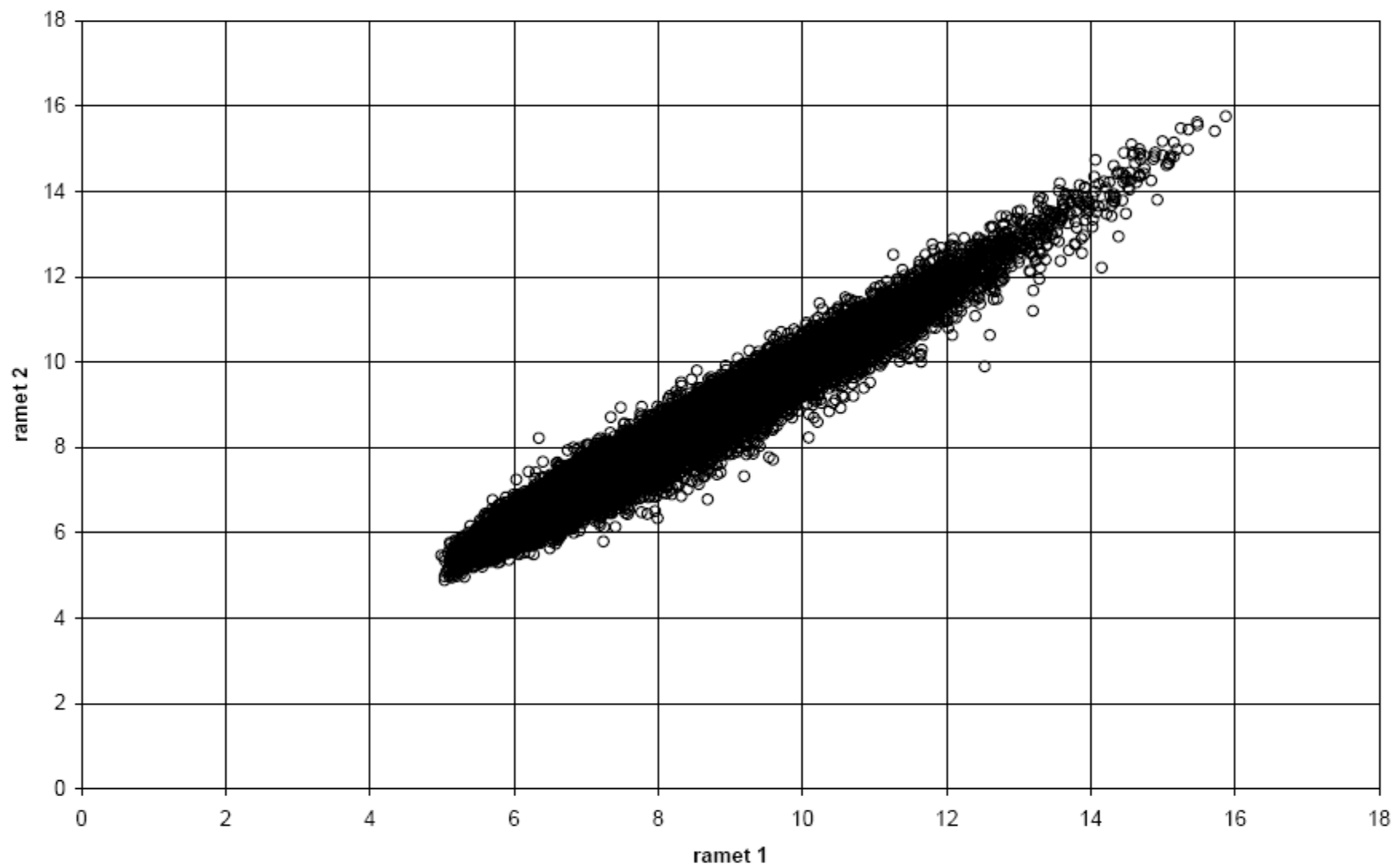
Electrophoresis File Run Summary

Bioanalyzer

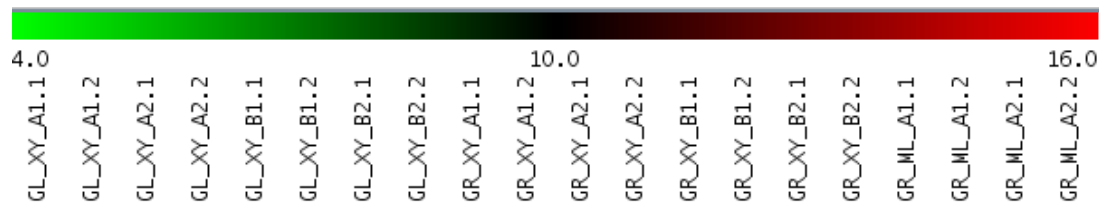
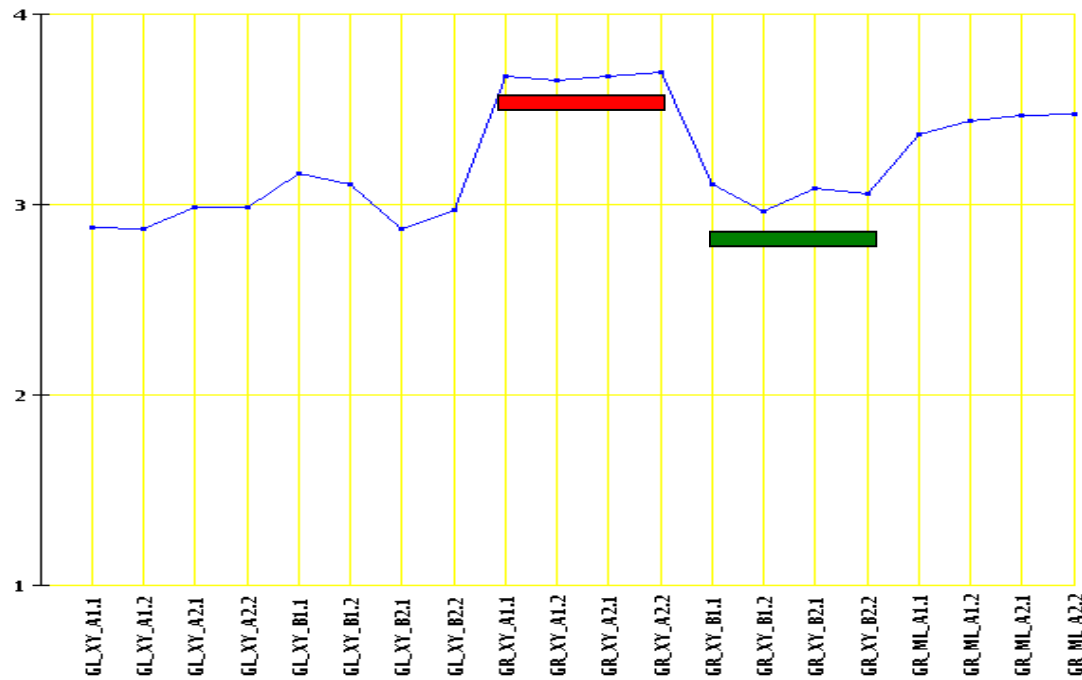


E. grandis, clone A, xilema

$r = 0,981$

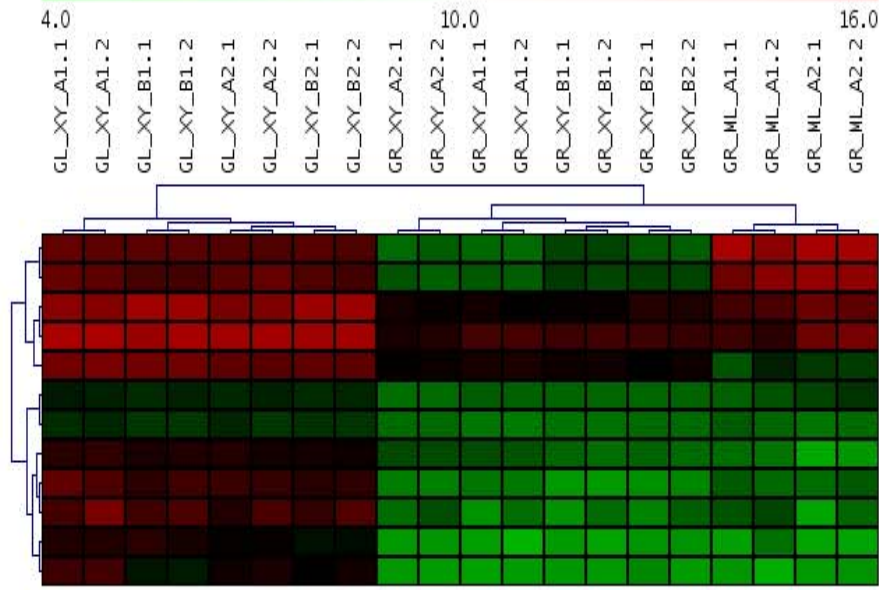


Variação entre diferentes genótipos de *E. grandis* – xilema

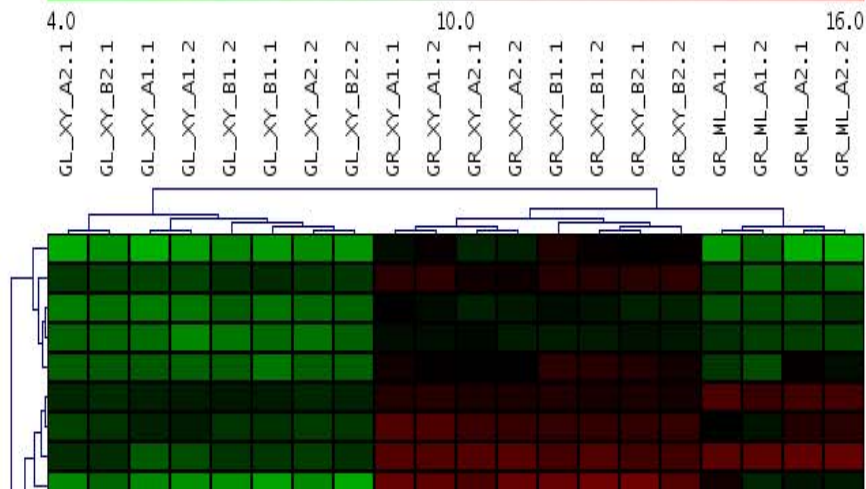


unknown protein [Arab
 dna-directed rna poly
 putative ras-GTPase-a
 Heavy metal transport
 nucleotide binding [A
 no hit
 no hit
 unknown protein [Arab
 no hit
 catalytic [Arabidopsi
 pathogenesis-related
 no hit
 no hit

Varição interespecífica: *E. globulus* x *E. grandis*



no hit
TIR-NBS disease resistance-like protein [(Populus tomentosa x P. bolleana) x P. tomentosa]
unknown protein [Arabidopsis thaliana] gb|AAF79910.1| Contains similarity to SCUTL1
putative aquaporin PIP1-2 [Vitis berlandieri x Vitis rupestris]
no hit
protein binding / ubiquitin-protein ligase/ zinc ion binding [Arabidopsis thaliana]
unknown protein [Arabidopsis thaliana] gb|AAM62838.1| Yippee-like protein [Arabidopsis thaliana]
unknown protein [Arabidopsis thaliana] gb|AAM97053.1| unknown protein [Arabidopsis thaliana]
Protein of unknown function DUF567 [Medicago truncatula]
no hit
no hit
SKP1 component [Medicago truncatula]



no hit
no hit
unknown protein [Arabidopsis thaliana] gb|AAD10685.1| Hypothetical protein [Arabidopsis thaliana]
alpha-mannosidase [Arabidopsis thaliana] gb|AAM47314.1| AT3g26720/MLJ15_12 [Arabidopsis thaliana]
Mal d 1-like [Malus x domestica]
nucleotide binding [Arabidopsis thaliana] gb|AAL47352.1| WD-repeat protein-like [Arabidopsis thaliana]
Pyridoxal phosphate biosynthetic protein pdxA [Rhizobium sp. NGR234]
Similar to gb|X84260 POS5 gene product from Saccharomyces cerevisiae. EST gb|W4387
zinc transporter [Eucalyptus grandis]

Microsoft Excel - novos microsats into genes selected

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda Digite uma pergunta

155 0

	A	B	C	D	E	F	G	H	I	J	
21	EMBRA1031R	AAGCTCGTT	CGGCGGCGGCGGCGGCGG	59.978	330	CGG	3	23	1	ORF	bZIP family transcription factor [Arabidopsis thaliana]
22	EMBRA1032F	TCGTCGGAG	TGATGATGATGAT	59.978	115	TGA	3	13	1	ORF	bZIP family transcription factor [Arabidopsis thaliana]
23	EMBRA1032R	CTCTTCCTCT	TGATGATGATGAT	59.945	115	TGA	3	13	1	ORF	bZIP family transcription factor [Arabidopsis thaliana]
24	EMBRA1040F	TCCGCACAA	CTCCTCTCCTCTCCTCTC	60.596	323	CTCCT	5	18	0	0	extensin like protein - black poplar d
25	EMBRA1040R	GCACACACC	CTCCTCTCCTCTCCTCTC	60.227	323	CTCCT	5	18	0	0	extensin like protein - black poplar d
26	EMBRA1041F	TCTCCTCTCC	GAGAGAGAGAGAGAGAGAG	59.817	332	GA	2	38	0	0	extensin like protein - black poplar d
27	EMBRA1041R	CACAGTCAC	GAGAGAGAGAGAGAGAGAG	59.940	332	GA	2	38	0	0	extensin like protein - black poplar d
28	EMBRA1056F	CCGCAAGCA	GAGAGAGAGAGAGAGAGAG	59.848	267	GA	2	22	1	5' UTR	ACC oxidase 1 [Fagus sylvatica]
29	EMBRA1056R	CTGTAGTGC	GAGAGAGAGAGAGAGAGAG	59.989	267	GA	2	22	1	5' UTR	ACC oxidase 1 [Fagus sylvatica]
30	EMBRA1062F	CGTCCGAAG	GAAGAAGAAGAAG	59.516	99	GAA	3	13	1	5' UTR	GTP-binding protein, ras-related - co
31	EMBRA1062R	GAAATTGGG	GAAGAAGAAGAAG	60.008	99	GAA	3	13	1	5' UTR	GTP-binding protein, ras-related - co
32	EMBRA1063F	CGAAGGAA	CCTCCTCCTCCTCCTCCTC	59.564	212	CCT	3	23	1	5' UTR	GTP-binding protein, ras-related - co
33	EMBRA1063R	CATTGATGT	CCTCCTCCTCCTCCTCCTC	60.111	212	CCT	3	23	1	5' UTR	GTP-binding protein, ras-related - co
34	EMBRA1068F	GGGCAGCA	CTCTCTCTCTCTCTCTCT	60.038	262	CT	2	18	1	5' UTR	transcription factor BTF3 (RNA poly
35	EMBRA1068R	TTCACCCCA	CTCTCTCTCTCTCTCTCT	60.042	262	CT	2	18	1	5' UTR	transcription factor BTF3 (RNA poly
36	EMBRA1069F	CTGAAGAGC	GACGACGACGACGACGAC	60.042	326	GAC	3	14	1	ORF	transcription factor BTF3 (RNA poly
37	EMBRA1069R	GTCTCGCCA	GACGACGACGACGACGAC	60.020	326	GAC	3	14	1	ORF	transcription factor BTF3 (RNA poly
38	EMBRA1070F	GGAACCGA	AGAGAGAGAGAGAGAGAGAG	60.074	259	AG	2	20	0	0	putative Pi starvation-induced protei
39	EMBRA1070R	CCAGGTGTT	AGAGAGAGAGAGAGAGAGAG	60.002	259	AG	2	20	0	0	putative Pi starvation-induced protei
40	EMBRA1074F	TTGAGGGTT	GATGACGATGACGATGACG	60.051	223	GATGAC	6	19	1	ORF	elongation factor 1-beta -related [Ara
41	EMB										
42	EMB										
43	EMB										
44	EMB										
45	EMB										
46	EMB										
47	EMB										
48	EMB										
49	EMB										
50	EMB										
51	EMB										
52	EMB										
53	EMB										
54	EMB										
55	EMB										

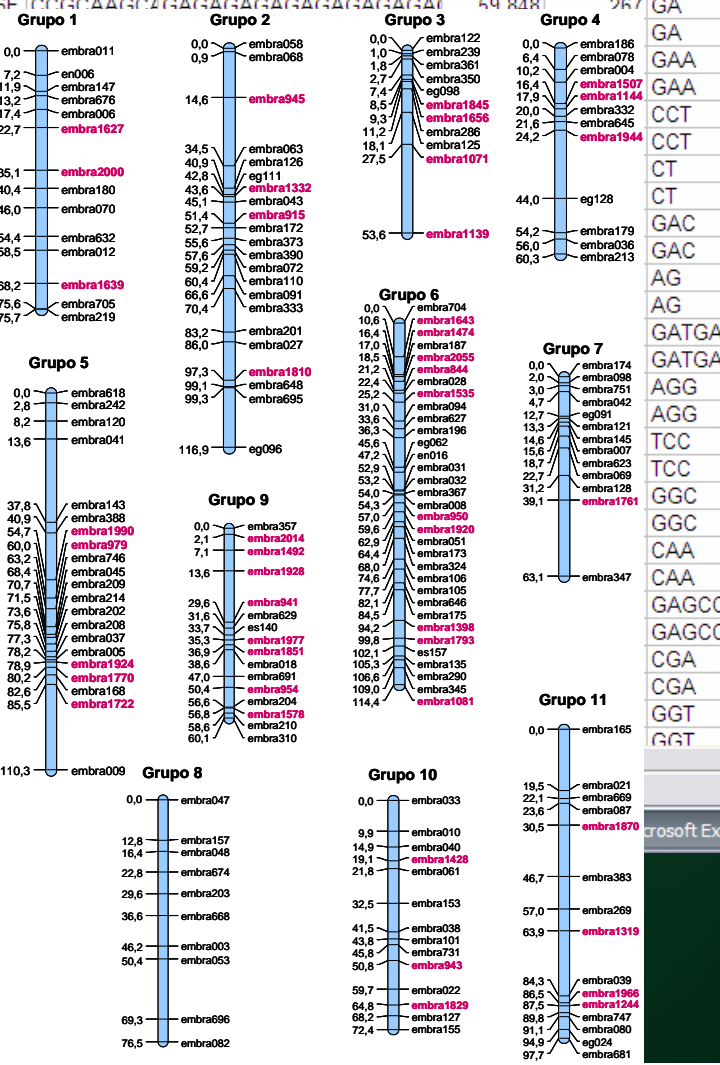
Triagem de microssatélites derivados de EST em um painel de 6 espécies de *Eucalyptus*

Microsoft Excel - novos microsats into genes selected

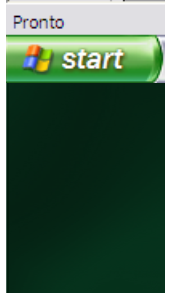
Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

155 fx 0

	A	B	C	D	E	F	G	H	I	J	
21	EMBRA1031R	AAGCTCGTT,CGGCGGGCGGGCGGGCGG		59.978	330	CGG	3	23	1	ORF	bZIP family transcription factor [Arabidopsis thaliana]
22	EMBRA1032F	TCGTCGGAGTGATGATGATGAT		59.978	115	TGA	3	13	1	ORF	bZIP family transcription factor [Arabidopsis thaliana]
23	EMBRA1032R	CTCTCTCTCTTGATGATGATGAT		59.945	115	TGA	3	13	1	ORF	bZIP family transcription factor [Arabidopsis thaliana]
24	EMBRA1040F	TCCGCACAACTCCTCTCCTCTCCTCTC		60.596	323	CTCCT	5	18	0	0	extensin like protein - black poplar d
25	EMBRA1040R	GCACACACCCTCCTCTCCTCTCCTCTC		60.227	323	CTCCT	5	18	0	0	extensin like protein - black poplar d
26	EMBRA1041F	TCTCCTCTCCGAGAGAGAGAGAGAGAGAG		59.817	332	GA	2	38	0	0	extensin like protein - black poplar d
27	EMBRA1041R	CACAGTCACTGAGAGAGAGAGAGAGAGAG		59.940	332	GA	2	38	0	0	extensin like protein - black poplar d
28	EMBRA1056F	CCGCAAGCTGAGAGAGAGAGAGAGAGAG		59.818	267	GA	2	22	1	5' UTR	ACC oxidase 1 [Fagus sylvatica]
29	EMBRA1056R					GA	2	22	1	5' UTR	ACC oxidase 1 [Fagus sylvatica]
30	EMBRA106					GAA	3	13	1	5' UTR	GTP-binding protein, ras-related - co
31	EMBRA106					GAA	3	13	1	5' UTR	GTP-binding protein, ras-related - co
32	EMBRA106					CCT	3	13			
33	EMBRA106					CCT					
34	EMBRA106					CT					
35	EMBRA106					CT					
36	EMBRA106					GAC					
37	EMBRA106					GAC					
38	EMBRA107					AG					
39	EMBRA107					AG					
40	EMBRA107					GATGAC					
41	EMBRA107					GATGAC					
42	EMBRA107					AGG					
43	EMBRA107					AGG					
44	EMBRA109					TCC					
45	EMBRA109					TCC					
46	EMBRA110					GGC					
47	EMBRA110					GGC					
48	EMBRA111					CAA					
49	EMBRA111					CAA					
50	EMBRA111					GAGCCG					
51	EMBRA111					GAGCCG					
52	EMBRA112					CGA					
53	EMBRA112					CGA					
54	EMBRA112					GGT					
55	EMBRA112					GGT					

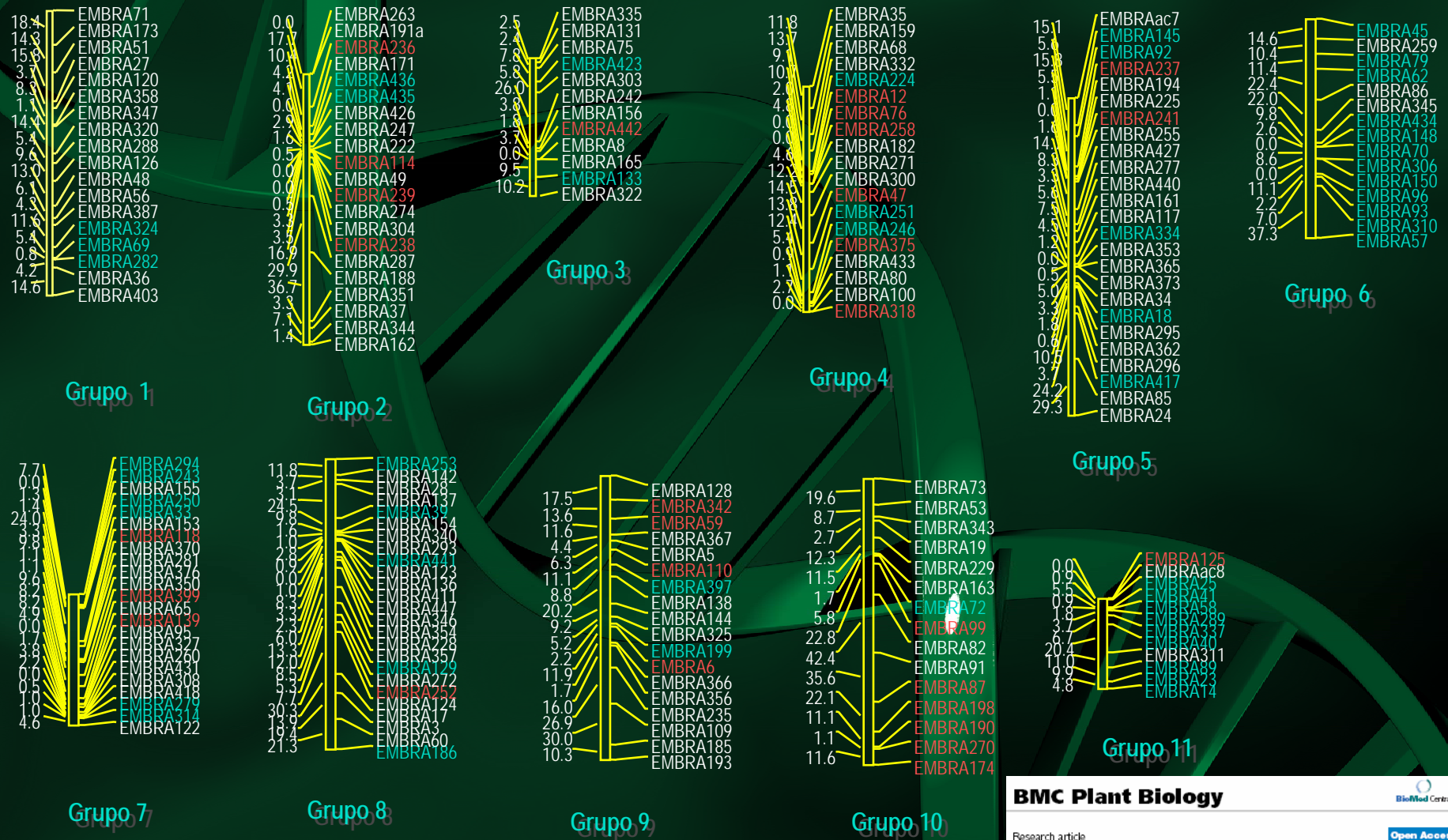


LOCO SSR	GENE COM MAIOR SIMILARIDADE NA ANÁLISE DE BLAST
EMBRA915	DNA binding protein [Elaeis oleifera]
EMBRA941	Skp1 [Medicago sativa]
EMBRA954	HB2 homeodomain protein [Populus tremula x Populus tremuloides]
EMBRA1081	zinc finger (C3HC4-type RING finger) protein family [Arabidopsis thaliana]
EMBRA1244	LIM domain protein PLIM1 [Nicotiana tabacum]
EMBRA1319	KH domain/zinc finger protein [Arabidopsis thaliana]
EMBRA1398	NOI protein, nitrate-induced [Arabidopsis thaliana]
EMBRA1428	wound induced protein kinase [Nicotiana tabacum]
EMBRA1474	senescence-associated protein -related [Arabidopsis thaliana]
EMBRA1627	glycosyltransferase family 43 [Arabidopsis thaliana]
EMBRA1639	PIN1-like auxin transport protein [Populus tremula x Populus tremuloides]
EMBRA1722	transfactor-related protein [Arabidopsis thaliana]
EMBRA1770	RHO GDP-dissociation inhibitor 1 -related [Arabidopsis thaliana]
EMBRA1829	histone H2A, putative [Arabidopsis thaliana]
EMBRA1870	auxin-induced (indole-3-acetic acid induced) protein, putative (SAUR_e) [Arabidopsis thaliana]
EMBRA1928	Heat Shock factor protein HSF24 [Lycopersicon peruvianum]
EMBRA1944	rac family GTP-binding protein (ARAC9) [Arabidopsis thaliana]
EMBRA1990	calcium-binding EF-hand family protein [Arabidopsis thaliana]



Desenvolvimento e mapeamento de microssatélites derivados de ESTs
Mamani et al. 2007

Um mapa genético referência para espécies de *Eucalyptus* baseado exclusivamente em microssatélites



Coleta de amostras de madeira para análise fenotípica via NIRS

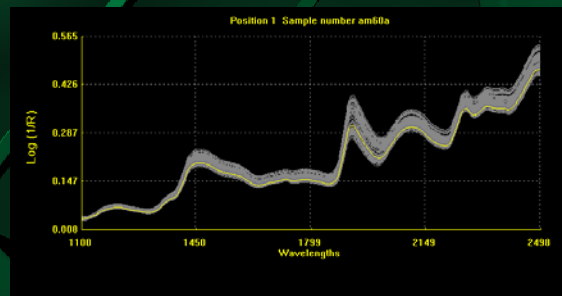


CALIBRAÇÃO do NIRS

Amostras de
Calibração



Resultados de análises de lab



espectro

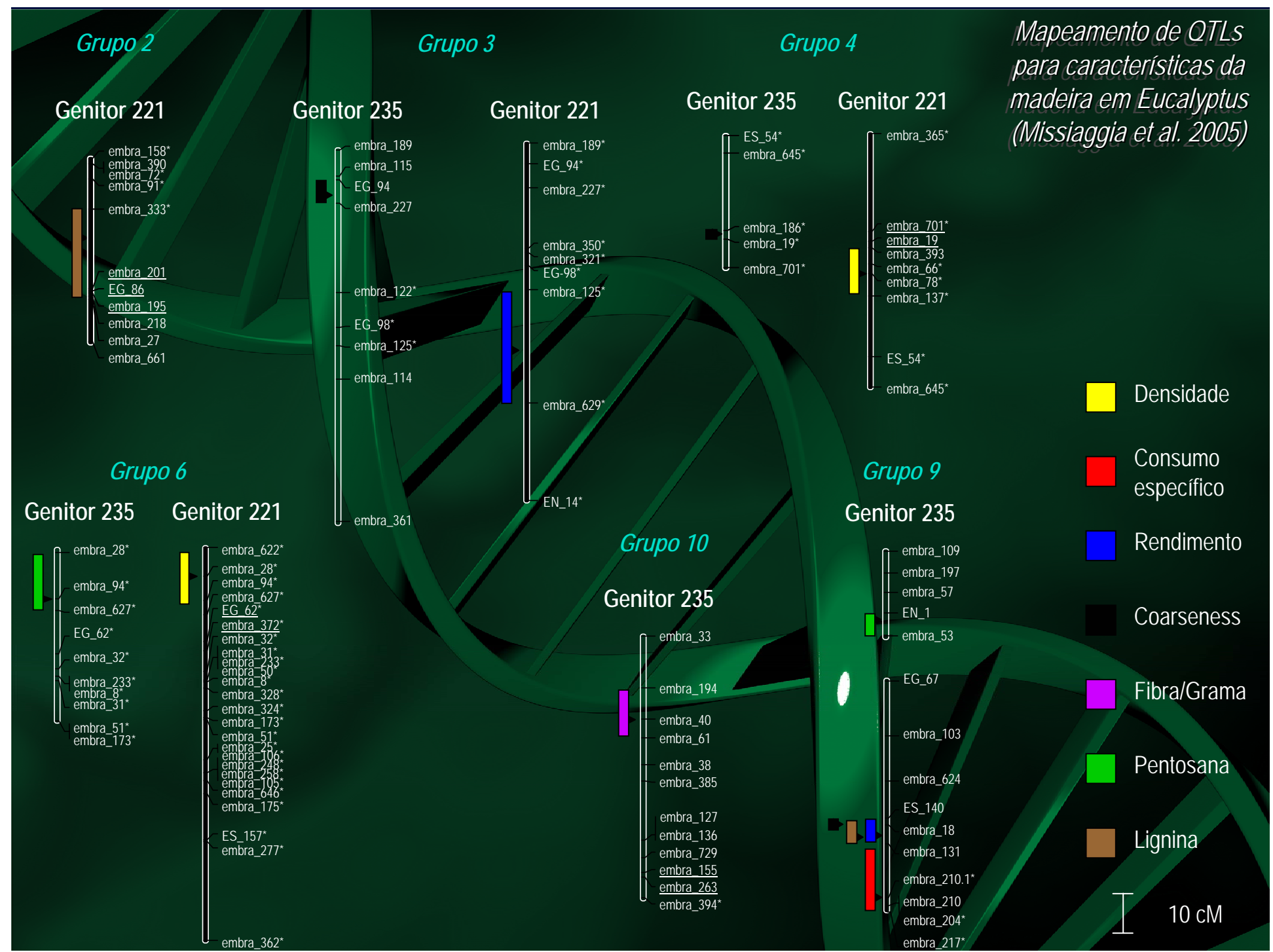
Equações
Preliminares

Análises de Rotina

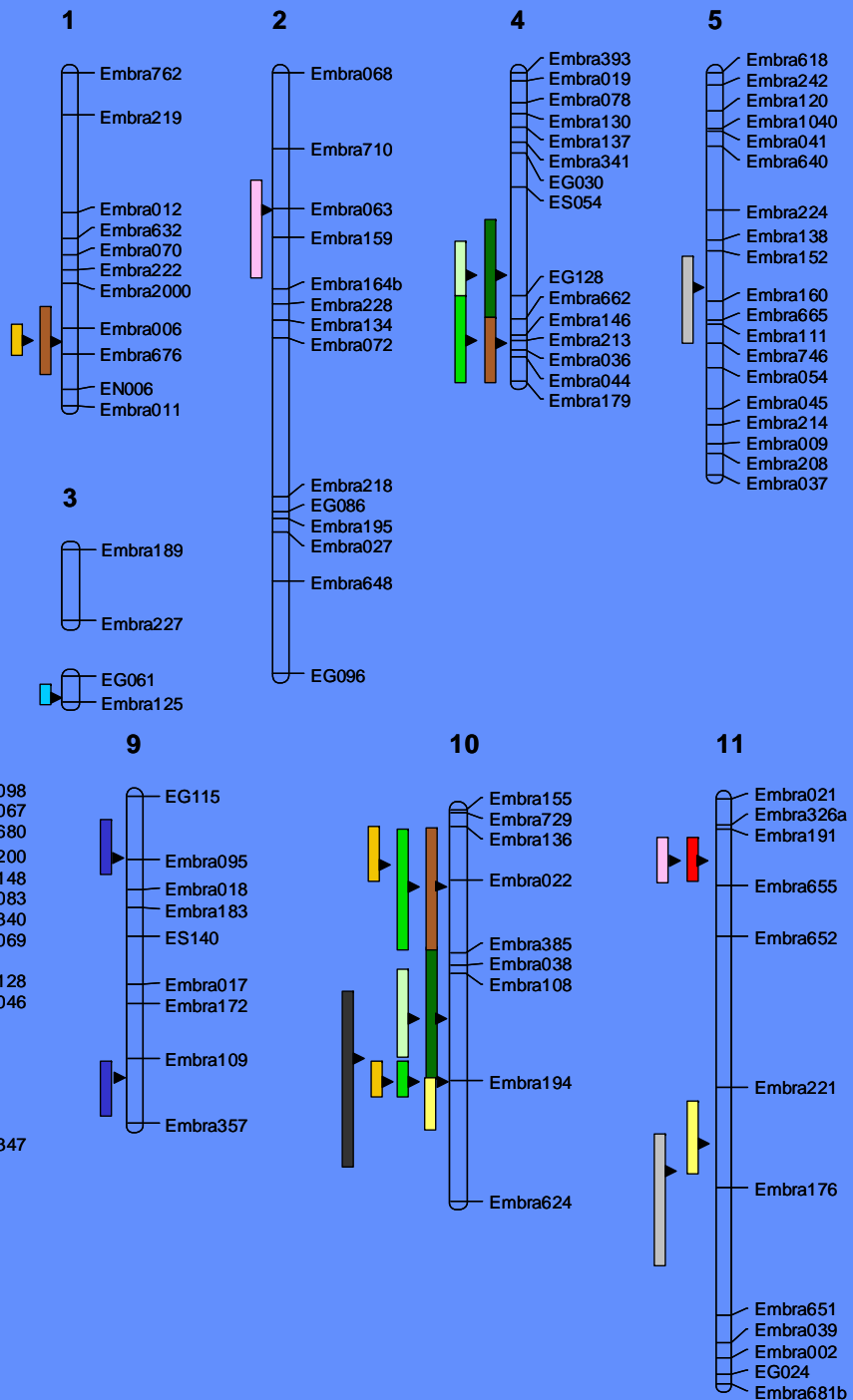
Validação



Mapeamento de QTLs para características da madeira em *Eucalyptus* (Missiaggia et al. 2005)

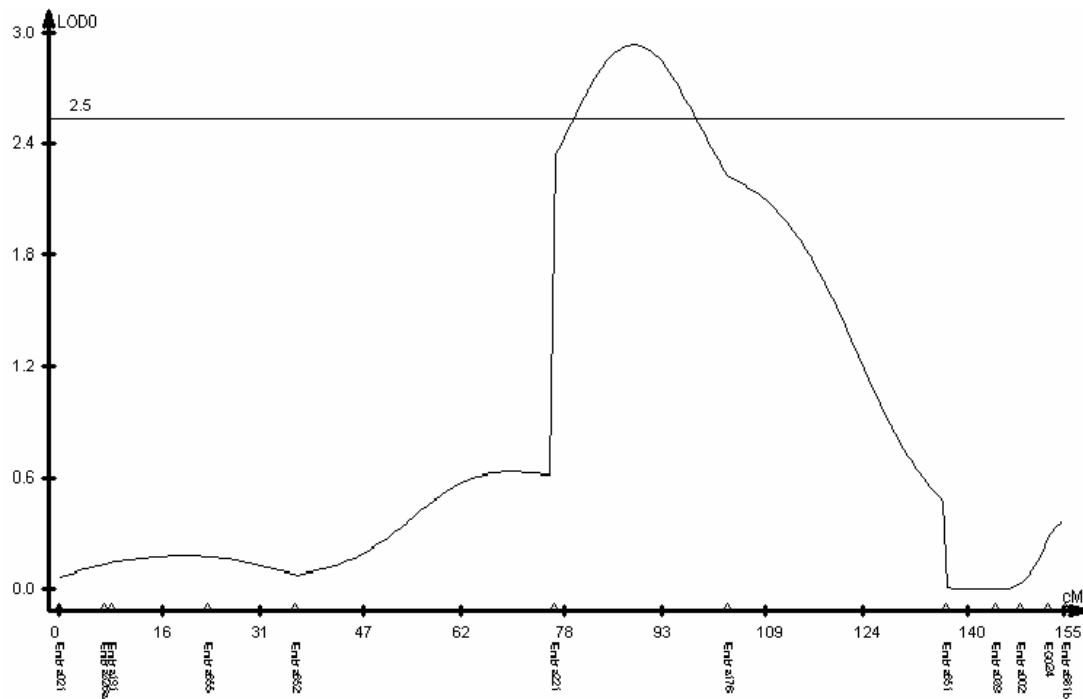


- DAP
- Altura
- Pilodyn
- Dens. Básica
- Glicana
- Xilana
- Ác. metilglucurônico
- Ác. galactourônico
- Lignina
- Siringil/Guaicil
- Álcali efetivo
- Rendimento depurado
- Viscosidade
- Extrativos



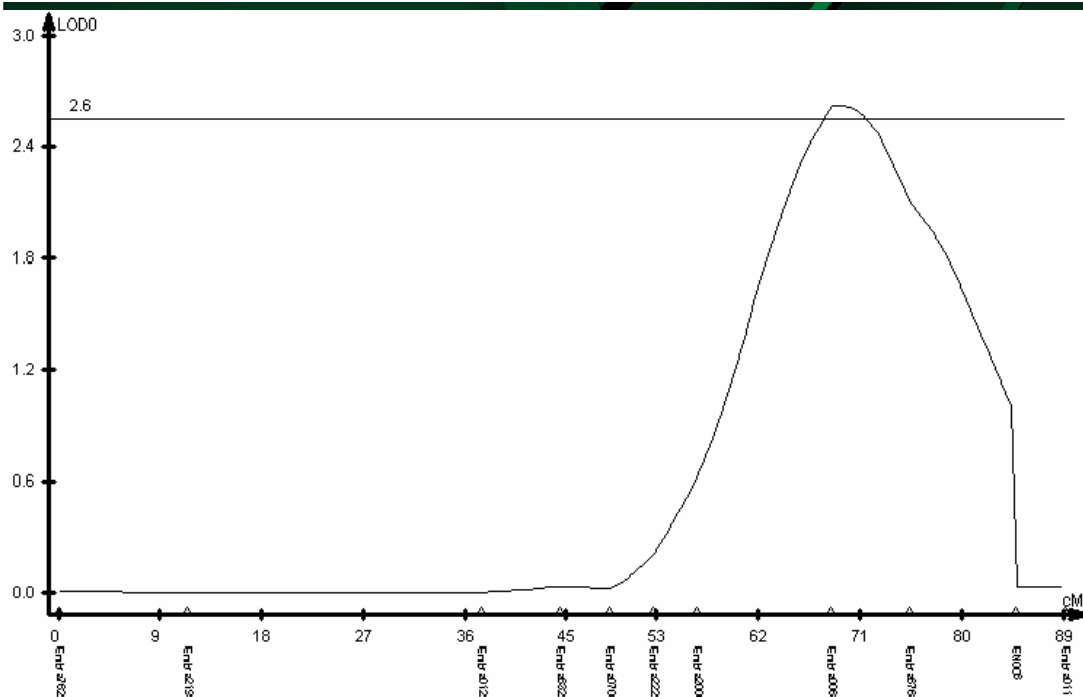
Exemplos de QTLs
mapeados para
características de
importância industrial

Novaes et al. 2006



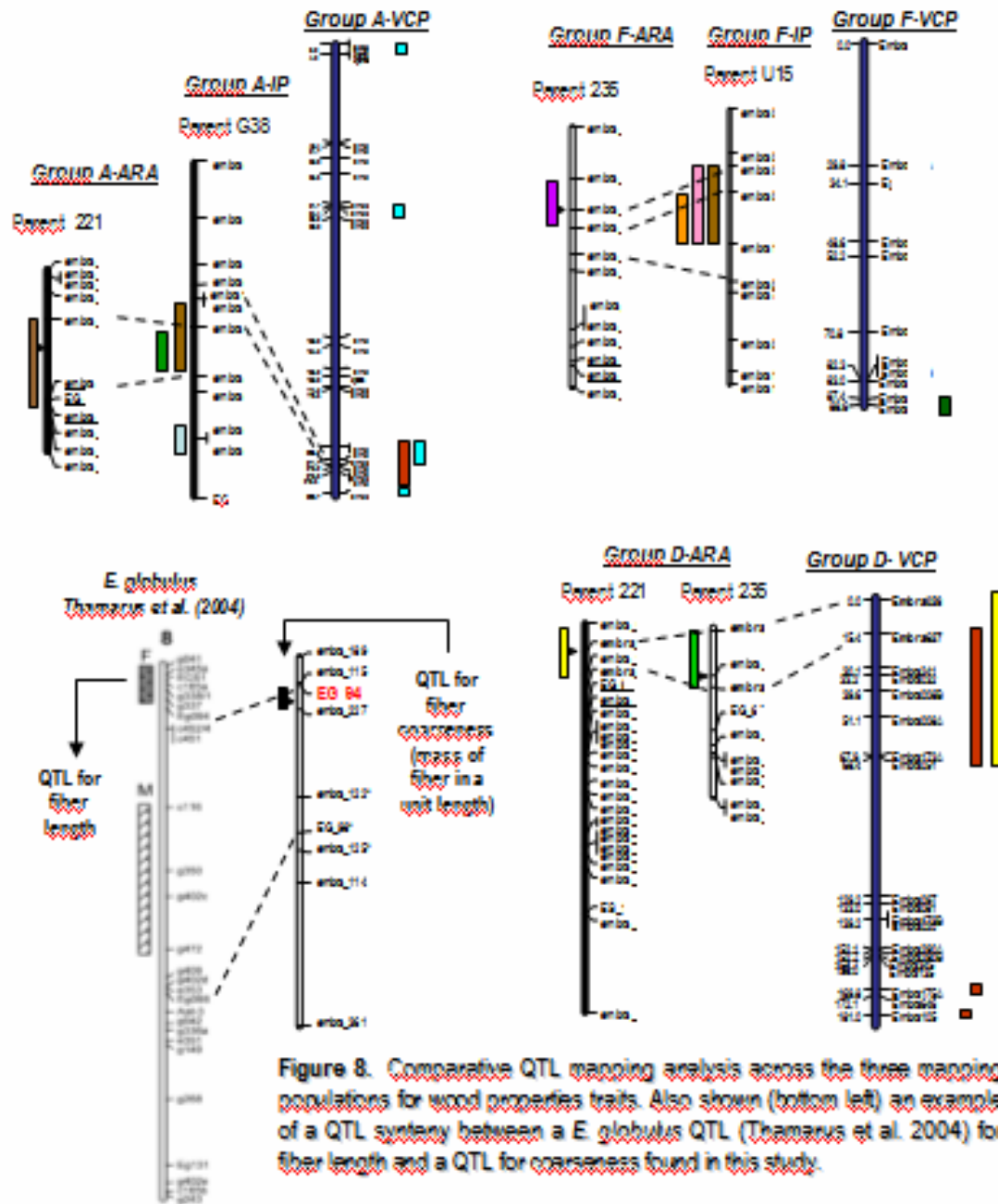
Exemplos de QTLs mapeados para características de importância industrial

QTL para teor de álcali efetivo



QTL para a relação Siringil/Guaicil lignina

Novaes et al. 2006



Mapeamento comparativo da localização de QTLs para propriedades físicas e químicas da madeira

3 famílias mapeadas:
 Aracruz
 Int. Paper
 Votorantim Cel. Pap.

Estas informações independentes de posicionamento de QTLs no mesmo segmento genômico são alvos de investigações correntes em outras famílias segregantes

DESENVOLVIMENTO DE MARCADORES E CONSTRUÇÃO DE MAPAS GENÉTICOS

- ✓ Desenvolvimento de 127 microssatélites de EST
- ✓ Desenvolvimento de 140 microssatélites de tetra e pentanucleotídeos
- ✓ Construção de mapas genéticos e posicionamento de QTL para 7 famílias
 - ✓ Família testemunha IP – 188 ind.; ~200 marcadores (Eva)
 - ✓ Família testemunha VCP – 188 ind.; ~180 marcadores (Evandro)
 - ✓ Família DG x UGL – 188 ind.; ~120 marcadores (Danielle)
 - ✓ Família C1 x UGL – 188 ind.; ~110 marcadores (Leo & Danielle)
 - ✓ Família DG x U2 – 188 ind.; ~100 marcadores (Cesar & Carolina)
 - ✓ Família DG x GL2 – 188 ind.; 110 marcadores (Medina)
 - ✓ Família G1 x UGL – 188 ind.; 123 marcadores (Gisele)

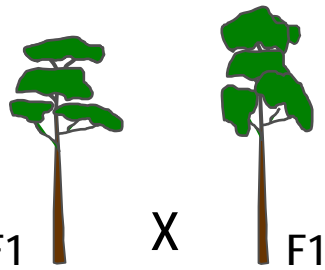


*Vigor híbrido em nível
de média de família*

*Vigor híbrido em nível
de árvore individual
dentro de família*

Fotos: Teotônio de Assis





ETAPA DE UTILIZAÇÃO DE QTLs PARA SELEÇÃO ASSISTIDA

Cruzamento entre híbridos para maximização da segregação na F2

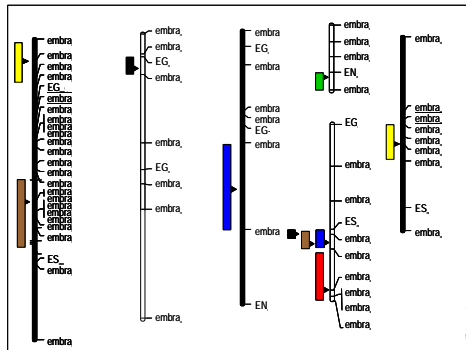
ETAPA DE DESCOBRIMENTO DE QTLs



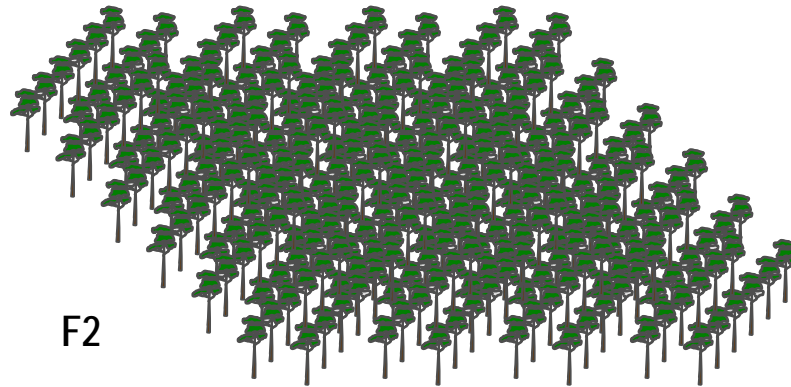
F2



Mapeamento de QTLs para características de qualidade da madeira com expressão tardia

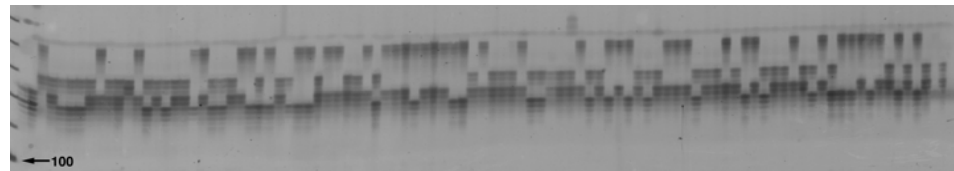


Informação de QTLs descobertos para características de expressão tardia (ex. qualidade da madeira)



F2

Geração de um grande número de descendentes F2 (n > 1000) da mesma família



Genotipagem de descendentes em idade ultra precoce (semanas) para marcadores ligados a QTLs para características de expressão tardia (ex. qualidade de fibra, teor de lignina, densidade básica) detectados na fase de mapeamento



Seleção assistida por marcadores para características de expressão tardia de descendentes. Estes são clonados para implantação de teste clonal. Teste clonal é otimizado pois envolve somente indivíduos pré-selecionados para características de expressão tardia.

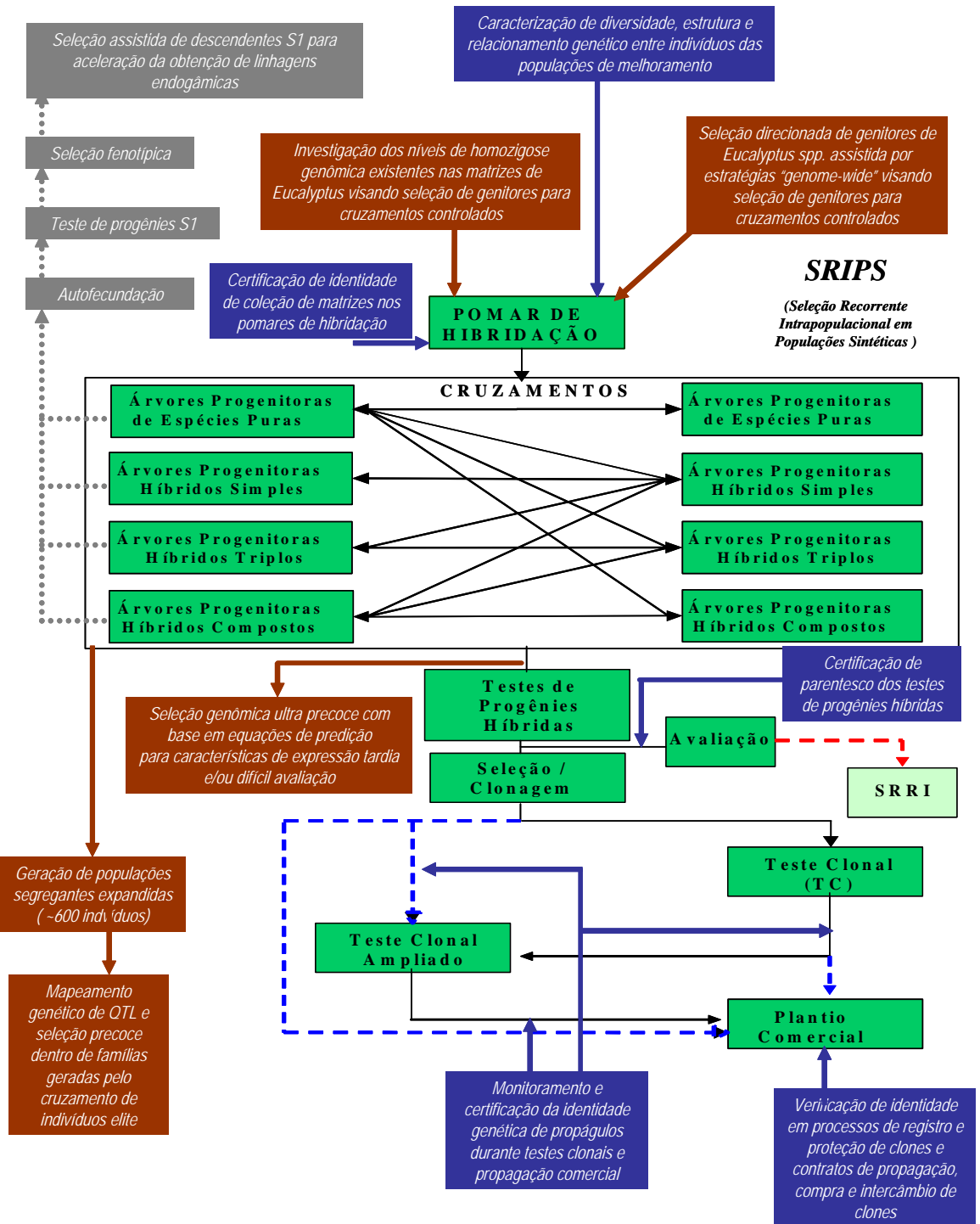


Integração de tecnologias genômicas em um programa de seleção recorrente interpopulacional de *Eucalyptus*

Ações de gerenciamento de diversidade e certificação de identidade genética e parentesco

Ações de seleção assistida por marcadores

Novas ações especificamente desenhadas para a integração de ferramentas genômicas



Conclusões e perspectivas

IMPACTO NO MELHORAMENTO

- ✓ Dependente da arquitetura genética da característica, dificuldade/custos de fenotipagem, possibilidade de seleção precoce, ganho por unidade de tempo
- ✓ *Benchmark* sempre será o melhoramento convencional

SELEÇÃO ASSISTIDA POR MARCADORES

- ✓ Desafio pressupõe a manipulação de características multifatoriais influenciadas por idades e ambientes variáveis
- ✓ A opção por SAM: caso a caso para características e populações específicas
- ✓ Novas tecnologias de genotipagem de alto desempenho estão rapidamente mudando a escala experimental e principalmente a relação de custos entre fenotipagem e genotipagem
- ✓ Whole-genome selection e seleção assistida por fenótipos ?
- ✓ Recursos genômicos não serão mais limitação

GENÔMICA E MELHORAMENTO

- ✓ *A genômica somente terá papel central no melhoramento genético quando puder contribuir não apenas para o entendimento da relação entre um gene e um fenótipo mutante mas sim para um entendimento mais global das relações complexas entre variação genética e variação fenotípica*

Produção científica e formação de talentos no âmbito do projeto até o momento

- ✓ Publicação de 25 artigos em revistas indexadas sendo 15 em revistas internacionais e mais 8 artigos submetidos ou em preparação
- ✓ Publicação de 10 capítulos de livros sendo 4 em livros internacionais
- ✓ Publicação de 22 artigos completos em congressos internacionais
- ✓ Defesa e publicação de 32 teses de mestrado ou doutorado
- ✓ Mais de 110 resumos publicados de apresentações em congressos nacionais e internacionais
- ✓ Publicação de um livro sobre Biotecnologia Florestal
- ✓ Treinamento de 40 bolsistas em diferentes níveis, principalmente DTI

Research article

Open Access

A microsatellite-based consensus linkage map for species of *Eucalyptus* and a novel set of 230 microsatellite markers for the genus
Rosana PV Brondani^{1,2,3}, Emlyn R Williams⁴, Claudio Brondani^{1,3} and Dario Grattapaglia^{1,5}

Mol Genet Genomics (2007) 267: 338–347
DOI 10.1007/s00438-002-0665-6

ORIGINAL PAPER

R.P.V. Brondani · C. Brondani · D. Grattapaglia

Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers

Theor Appl Genet (2004) 109: 192–199
DOI 10.1007/s00122-004-1617-9

ORIGINAL PAPER

D. Grattapaglia · V. J. Ribeiro · G. D. S. P. Rezende

Retrospective selection of elite parent trees using paternity testing with microsatellite markers: an alternative short term breeding tactic for *Eucalyptus*

Theor Appl Genet (2003) 108:175–180
DOI 10.1007/s00122-003-1415-9

ORIGINAL PAPER

D. T. Junghans · A. C. Alfenas · S. H. Brommonschenkel · S. Oda · E. J. Mello · D. Grattapaglia

Resistance to rust (*Puccinia psidii* Winter) in *Eucalyptus*: mode of inheritance and mapping of a major gene with RAPD markers

Theor Appl Genet (2002) 105:474–478
DOI 10.1007/s00122-002-0899-z

C.M. Marques · R.P.V. Brondani · D. Grattapaglia
R. Sederoff

Conservation and synteny of SSR loci and QTLs for vegetative propagation in four *Eucalyptus* species

Tree Genet Genomes (2005) 1: 79–84
DOI 10.1007/s12245-005-0011-3

ORIGINAL PAPER

A. A. Missiaglia · A. L. Piacuzzi · D. Grattapaglia

Genetic mapping of *Eef1*, a major effect QTL for early flowering in *Eucalyptus grandis*

Plant microsatellite genotyping with 4-color fluorescent detection using multiple-tailed primers

Alexandre Missiaglia^{1,2} and Dario Grattapaglia^{1,3}

¹Laboratório de Genética Vegetal, Embrapa - Recursos Genéticos e Biotecnologia, Parque Estação Biológica, Brasília, DF, Brasil

Journal of Heredity 2005:96(2):1–6
doi:10.1093/jhered/es023

© 2005 The American Genetic Association

Power of Microsatellite Markers for Fingerprinting and Parentage Analysis in *Eucalyptus grandis* Breeding Populations

M. KIRST, C. M. CORDEIRO, G. D. S. P. REZENDE, AND D. GRATTAPAGLIA

From the Plant Genetics Laboratory – EMBRAPA Recursos Genéticos e Biotecnologia, Parque Estação Biológica, Brasília 70770-970 DF, Brazil (Kirst, Cordeiro, and Grattapaglia); Research and Technology Center – Aracruz Celulose S.A., C.P. 331011 Aracruz 29197-000 ES, Brazil (Rezende); and Genomic Sciences Laboratory, Universidade Católica de Brasília – SGAN 916 módulo B, Brasília 70790-160 DF (Grattapaglia).

Integrating genomics into *Eucalyptus* breeding

Dario Grattapaglia

Plant Genetics Laboratory, Embrapa Recursos Genéticos e Biotecnologia, Caixa Postal 02372, 70770-900 Brasília, DF, Brasil, and Graduate Program in Genomic Sciences, Universidade Católica de Brasília, UCB, SGAN 916 Módulo B, 70790-160 Brasília, DF, Brasil
Corresponding author: Dario Grattapaglia
E-mail: dario@cenargen.embrapa.br

Aplicações Operacionais de Marcadores Moleculares

Aplicações Operacionais de Marcadores

Moleculares

Pesq. agropec. bras., Brasília, v. 42, n. 7, p.1007-1012, jul. 2007

Potencial de híbridos entre clones-elite de eucalipto por meio de marcadores microsatélites

Marcelo Steir de Aguiar¹*, Daniel Furtado Ferreira², Aurélio Mendes Aguiar³, Odair Bisoni⁴, Gabriel Dehon Sampaio Peçanha Rezende⁵ e Dario Grattapaglia⁶

¹Universidade Federal de Lavras (UFLA), Dep. de Biologia, Caixa Postal 37, CEP 37200-000 Lavras, MG. E-mail: aguiarml@hotmail.com
²UFLA, Dep. de Exatas. E-mail: daniel1f85@ufla.br
³Aracruz Celulose, Rua São Geraldo, 1600, CEP 92500-000 Guaiiba, RS. E-mail: amendes@aracruz.com.br
⁴Seminhas Agrárias, Rod. BR 264, km 04, CEP 78770-000 Alto Garças, MT. E-mail: odairbisoni@terra.com.br
⁵Aracruz Celulose, Rod. Aracruz, Barra do Riocho, s/n, CEP 29197-000 Aracruz, ES. E-mail: gda@aracruz.com.br
⁶Embrapa Recursos Genéticos e Biotecnologia, Pq. Estação Biológica, Av. W3 Norte, Caixa Postal 02372, CEP 70770-900 Brasília, DF. E-mail: dario@cenargen.embrapa.br

Mapas Genéticos e Seleção Assistida por Marcadores Moleculares

Mapas Genéticos e Seleção Assistida por

Marcadores Moleculares

11 **Genomics of *Eucalyptus*, a Global Tree**
12 **for Energy, Paper, and Wood**

13
14
15 **Dario Grattapaglia**
16
17

CHAPTER 4

4 Eucalypts

Alexander A. Myburg¹, Brad M. Potts², Christina M. Marques³, Matias Kirst⁴, Jean-Marc Gion⁵, Dario Grattapaglia⁶, and Jacqueline Grima-Pettenatt⁷

¹ Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0002, South Africa, e-mail: zander.myburg@fabi.up.ac.za

² School of Plant Science, and CRC for Forestry, University of Tasmania, Private Bag 55, Hobart, Tasmania 7001, Australia

³ RAIZ-Direção de Investigação Florestal ITQB II Av. República, Apartado 127, 2781-901 Oeiras, Portugal

⁴ School of Forest Resources and Conservation, 368 Newins-Ziegler Hall, University of Florida, Gainesville, FL 32611-0410, USA

⁵ Programme Arbres et Plantations, CIRAD Forêt TA 10°C, Campus de Baillarguet 34398, Montpellier Cedex 5, France

⁶ Graduate Programme in Genomic Sciences and Biotechnology, Universidade Católica de Brasília-SGAN 916 módulo B, 70790-160 DF Brasília, Brazil

⁷ UMR UPS/CNRS 5546, Pôle de Biotechnologies Végétales, 24 chemin de Borde Rouge, BP42617, Auzerville, Tolosane 31325 Castanet Tolosan, France

Crop Breeding and Applied Biotechnology 5:91-98, 2005
Brazilian Society of Plant Breeding. Printed in Brazil



Localized mapping of RAPD markers in *Eucalyptus grandis*

Eduardo Nogueira Campinho¹*, Marcelo Luiz de Lais², Dario Grattapaglia³, Fernando L. Bertolucci⁴, Acélio Couro Alfenas⁵, and Edgard Augusto de Toledo Picoli⁶

151

CHAPTER 14

Marker-assisted selection in *Eucalyptus*

MARKER-ASSISTED SELECTION

Current status and future perspectives
in crops, livestock, forestry and fish

Dario Grattapaglia

*Algumas publicações 2002-2007 do grupo sobre desenvolvimento e aplicações de marcadores moleculares em *Eucalyptus**

Guia do

EUCALIPTO

Oportunidades para um desenvolvimento sustentável



Conselho de Informações em biotecnologia
www.cib.org.br

EUCALIPTO

Oportunidades para um desenvolvimento sustentável



O mundo se curva diante da competitividade do setor florestal brasileiro, fruto de nossas condições climáticas e da tecnologia desenvolvida pelas empresas e instituições de pesquisa do País. Não se trata de força de expressão. Os números apontam que o segmento é responsável por 3,5% do PIB, US\$ 4,8 bilhões em Impostos e US\$ 6,1 bilhões em exportações. Para movimentar essa máquina, o setor emprega 4,6 milhões de pessoas, quase 5% da população economicamente ativa do Brasil.

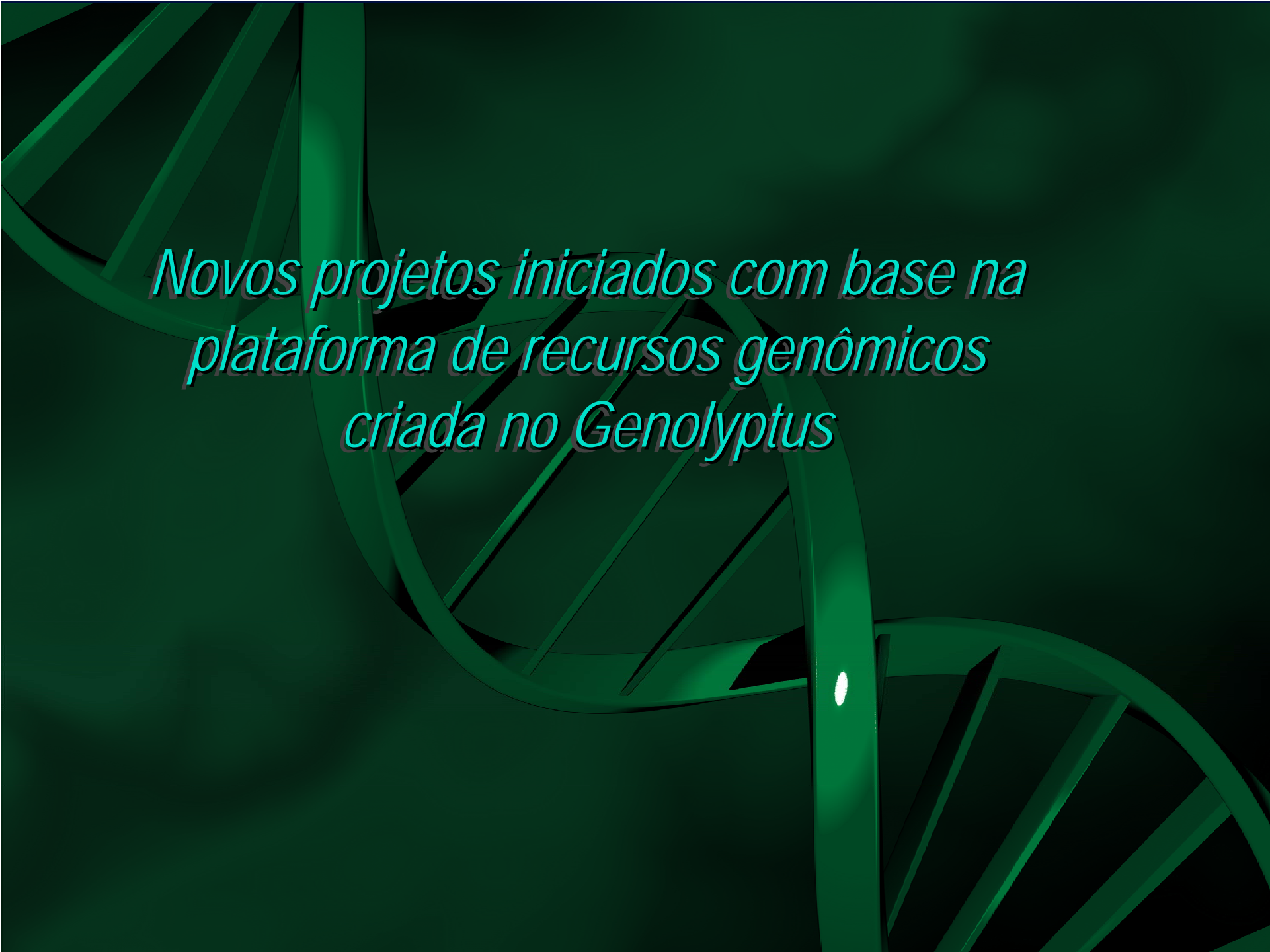
Nesse cenário, as plantações de eucalipto no Brasil ocupam mais de 3,7 milhões de hectares. E a tecnologia – neste caso, a Biotecnologia – pode contribuir ainda mais para que o Brasil se mantenha na posição de liderança no mercado mundial. Esse é o principal razão pela qual o Conselho de Informações sobre Biotecnologia decidiu editar o presente guia.

Aqui você encontrará informações desde a origem da cultura do eucalipto, passando por seu melhoramento genético convencional, até sua evolução pela Biotecnologia, que, ao contribuir para desenvolver novas características – a exemplo do maior rendimento industrial –, está estabelecendo uma realidade diferente para a produção florestal brasileira, particularmente para o eucaliptocultura.

Esperamos que este material sirva de fonte de informação e pesquisa para educadores, estudantes, agricultores, jornalistas e representantes da sociedade interessados no tema.

Bom leitura!





*Novos projetos iniciados com base na
plataforma de recursos genômicos
criada no Genolyptus*



De novo assembly results

454NewblerMetrics.txt

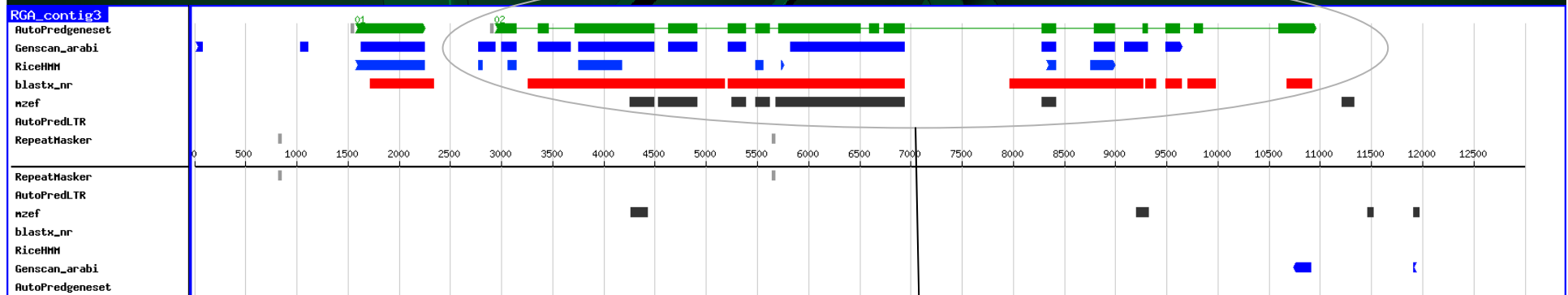
	RCI	CAD	FSH	RGA	EGPOOL
Number of Reads	14445	23789	30594	27777	89714
Number of Bases	3852819	6062028	7853206	7126351	22983486
Number of Reads Assembled	8490	15643	26254	22927	77618
Number of Large Contigs	10	6	34	11	126
Large Contig Bases	9188	21120	158652	76589	912842
N50 Size	986	6343	84011	10304	18359
Q40+ bases	9150	21061	154805	76507	906355
Avg Depth for Large Contigs	264X	199X	43.73X	78.94X	23.10X

Large Contig: a "Large" contigs is a contig with at least 500 bases in length

Q40+ bases: This is all the bases in the assembled contigs that carry a phred-equivalent quality score of 40 and above. A score of 40 is equivalent to an error rate of 10^{-4} , or an accuracy of 99.99%.

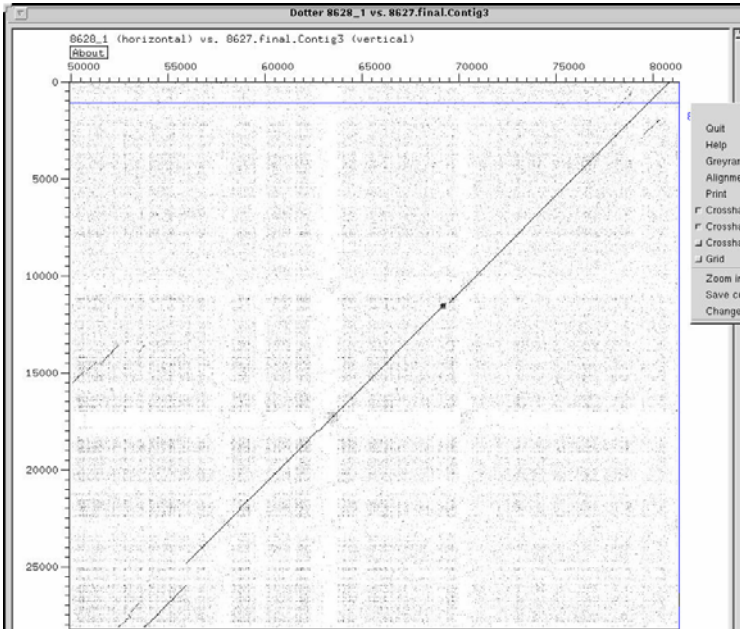
Clone	Locus	Accession	Chromosome	number of gene
F5H	---	---	---	13
No	Predicted Function by GFSelector			
01	hypothetical protein			
02	putative Lipase, active site			
03	putative Histone-lysine N-methyltransferase NSD3 (Nuclear SET domain-containing protein 3) (WHSC1-like protein 1) (Wolf-Hirschhorn syndrome candidate 1-like protein 1 homolog)			
04	hypothetical protein			
05	putative ZWI (ZWICHEL); calmodulin binding / microtubule motor			
06	putative Pectinesterase precursor (Pectin methylesterase) (PE)			
07	hypothetical protein			
08	hypothetical protein			
09	putative retrotransposon dell-46			
10	hypothetical protein			
11	putative serine/threonine protein kinase			
12	putative RNA-directed DNA polymerase homolog T13L16.7			
13	putative Integrase, catalytic region; Zinc finger, CCHC-type			

Análogo de gene de resistência

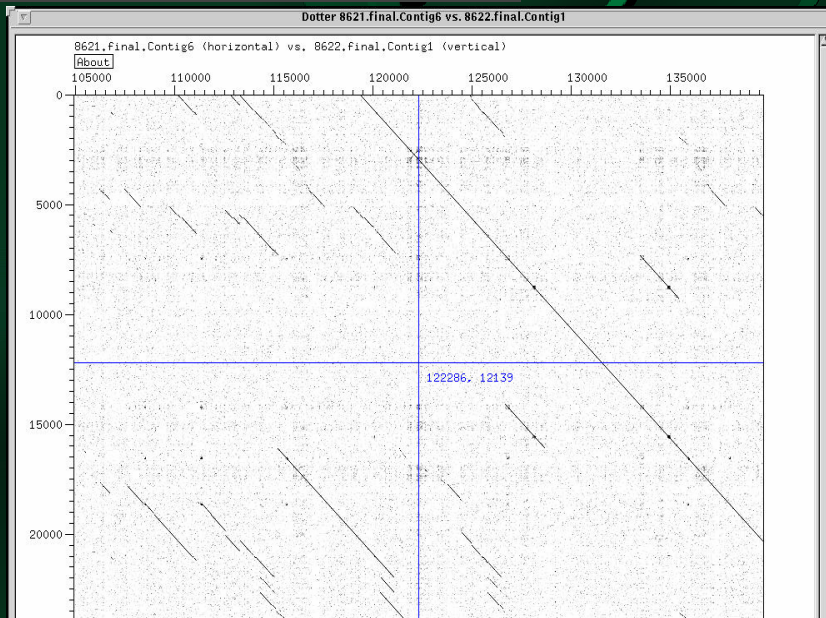


Disease Resistance gene analog (RGA)

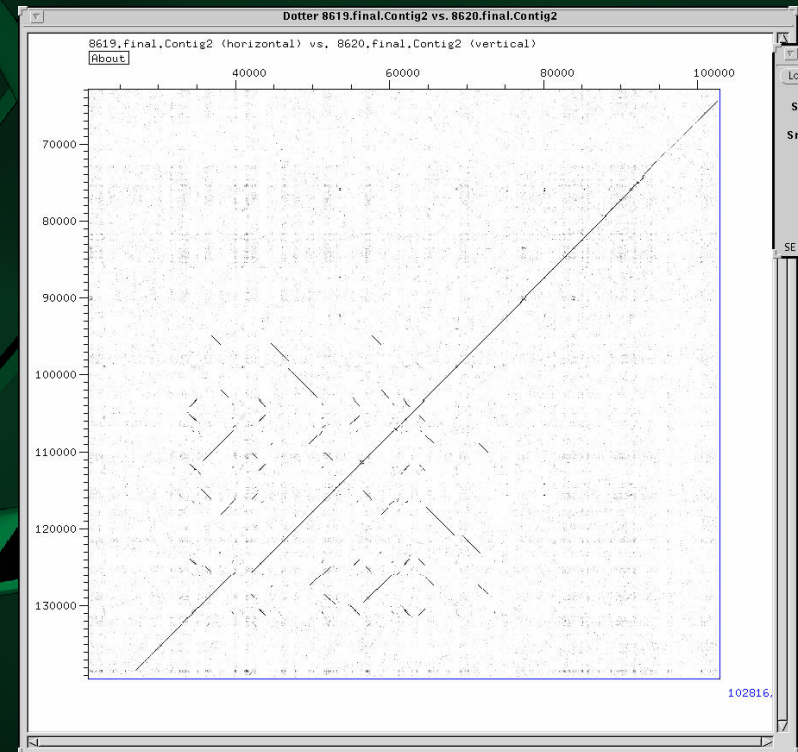
Sequenciamento e alinhamento dos dois haplótipos de uma árvore heterozigota de *Eucalyptus grandis*
Avaliação das perspectivas de montagem do genoma heterozigoto
(dados colaboração Genolyptus – Stanford Genome Technology Center)



Elevada microcolinearidade (27 kb)

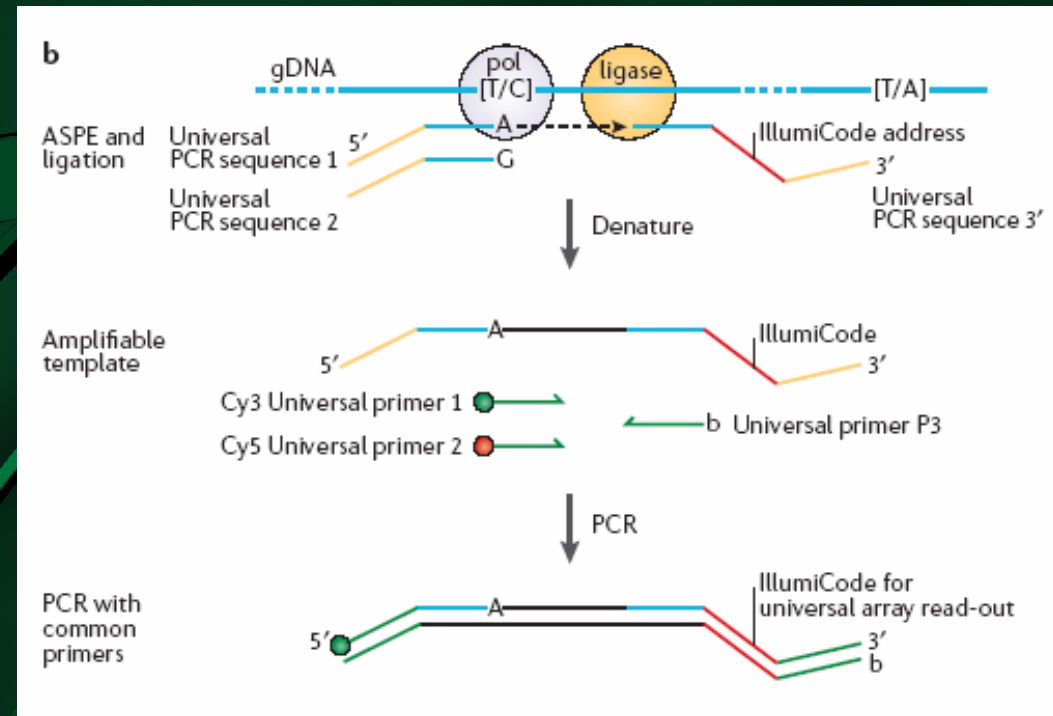
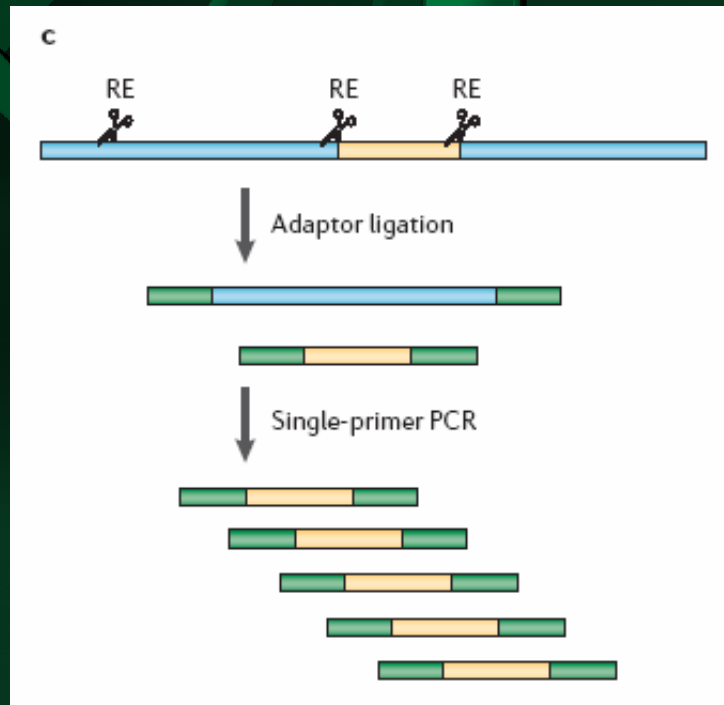


Presença de duplicações em tandem (25 kb)



Presença de duplicações invertidas (14 kb)

Tecnologias de genotipagem de alto desempenho



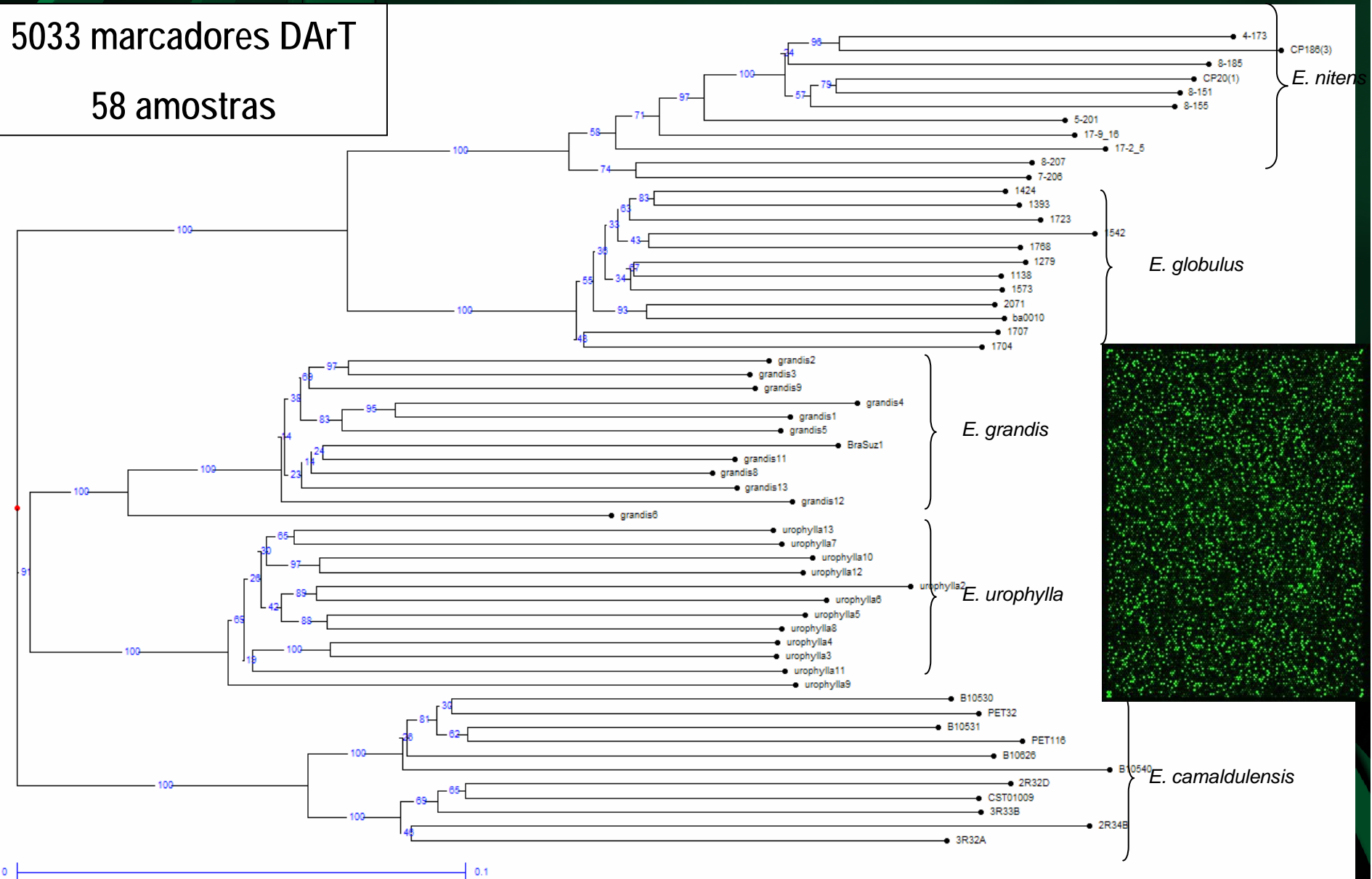
Microarranjos de representações genômicas reduzidas
(Ex. Diversity Arrays Technology DArT)
Milhares de marcadores por amostra a custo de ~ 0,005 US\$/datapoint

Extensão de ligação de oligonucleotídeos loco específicos sobre arranjos de contas
(Ex. GoldenGate)
Milhares de marcadores por amostra a custo de ~ 0,10 a 0,05 US\$/datapoint

Polimorfismo em *E. nitens*, *E. globulus*, *E. urophylla*, *E. grandis*, *E. camaldulensis* - array teste de 14.600 fragmentos

5033 marcadores DArT

58 amostras



Mineração e seleção de SNPs para desenvolvimento de plataforma de genotipagem de alto desempenho para *Eucalyptus*

(a)

Reference nucleotide (position)	A (158)	G (200)	T (258)	C (287)	C (431)	A (605)
# reads (forward)	27	29	18	14	13	35
# reads (reverse)	29	41	48	47	26	42
Alternative nucleotide	G	A	C	A	T	C
# reads (forward)	6	6	10	10	4	8
# reads (reverse)	4	9	12	7	12	7

(b)

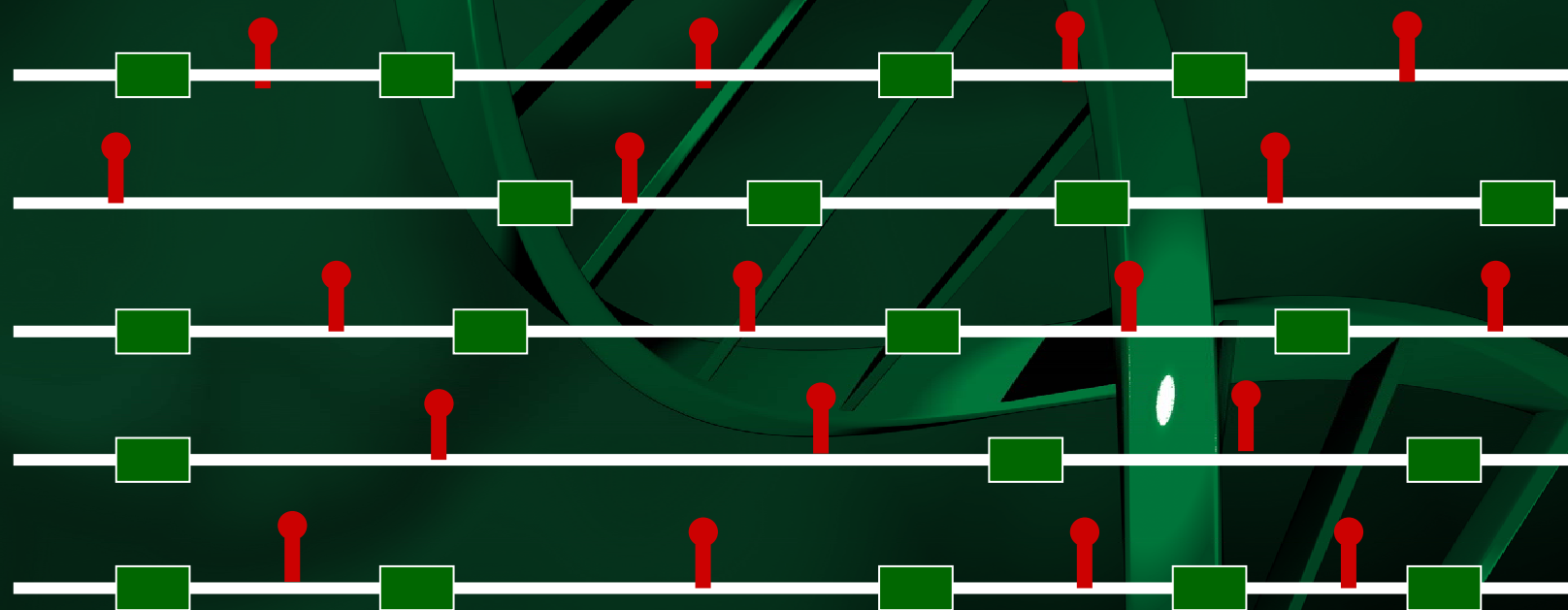
	Start	Strand	Sequence	End
Reference sequence	584	+	CC-ATGG-AACG-TACGACTTCATATT-CGT-GGAT-GCCGACAAGGA	625
EPE3D2102IPD9P	70	-	CCTATGG-AACG-TACGACTTCATCCTT-CGT-GGAT-GCCGACAAGGA	28
EPE3D2101A5YSB	137	+	CC-ATGG-AACG-TACGACTTCATCCTT-CGT-GGAT-GCCGACAAGGA	178
EPE3D2102JB6TU	137	+	CC-ATGG-AAC--TACGACTTCATCCTT-CGT-GGAT-GCCGACAAGGA	177
EPE3D2102G9XCJ	137	+	CC-ATGG-AACG-TACGACTTCATCCTT-CGT-GGAT-GCCGA-AAGGA	178
EPE3D2101B7Z57	137	+	CC-ATGG-AACG-TACGACTTCATCCTT-CGT-GGAT-GCCGACAAGGA	178
EPE3D2102GRHSL	136	-	CC-ATGG-AACG-TACGACTTCATCCTT-CGT-GGAT-GCCGACAAGGA	95
EPE3D2101B9176	110	+	CC-ATGG-AACG-TACGACTTCATCCTT-CGT-GGAT-GCCGACAAGGA	151
EPE3D2101EYG9P	136	-	CC-ATGG-AACG-TACGACTTCATCCTT-CGT-GGAT-GCCGACAAGGA	95
EPE3D2101AWNSN	102	-	CC-ATGG-AACG-TACGACTTCATCCTT-CGT-GGAT-GCCGACAAGGA	61
EPE3D2101BK40P	195	-	CC-ATGG-AACG-TACGACTTCATCCTT-CGT-GGAT-GCCGACAAGGA	154
EHZ8CPV02HE68U	48	-	CC-ATGGG--CG-TACGACTTCATCCTT-CGT-GGAT-GCCGACAAGGA	8
EE6U72G01B7TNW	39	+	CC-ATGG-AACG-TACGACTTCATCCTTACGT-GGAT-GCCGACAAGGA	81
EPE3D2102HFW0Y	11	+	CC-ATGG-AACG-TACGACTTCATCCTT-CGT-GGATTGCCGACAAGGA	53
EHZ8CPV01BU497	96	-	CC-ATGG-AACG-TACGACTTCATCCTT-CGT-GGAT-GCCGACAAGGA	55
EPE3D2102GBLRU	1	+	ATGGGAACGGTACGACTTCATCCTT-CGTAGGAT-GCCGACAAGGA	43

Fig. 6 Single nucleotide polymorphism (SNP) discovery by massively parallel sequencing of normalized cDNA from unrelated *Eucalyptus* trees. (a) SNPs identified in the *Eucalyptus grandis* homologs of the lignin gene caffeoyl-CoA 3-O-methyltransferase. The reference nucleotide at an SNP is defined by the most frequently observed allele. In this analysis, at least two reads containing the reference and alternative SNP nucleotide, in both forward and reverse directions, were required to declare the presence of an SNP. (b) Alignment of 15 sequences (eight forward and seven reverse sequences) that contain the alternative alleles for the SNP located in position 605 (box). The gray shading indicates other variable positions that have not been declared polymorphic because of the limited number of sequences to support them.

Seleção genômica ampla (SGA)

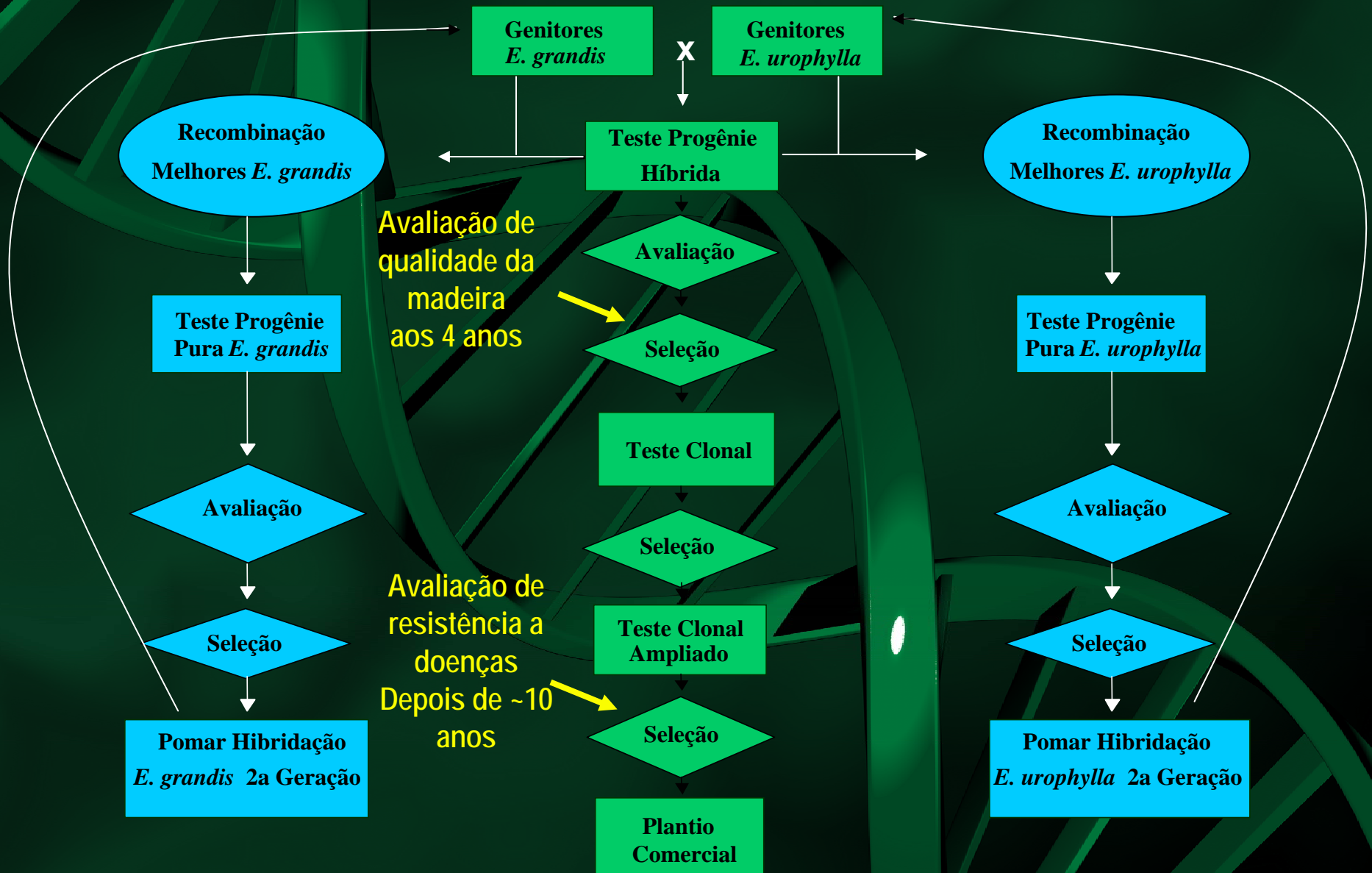
(Genomic selection ou Genome-wide selection)

SGA pode ser definida como sendo a seleção simultânea para centenas ou milhares de marcadores (SNPs), a depender do organismo e extensão do DL (desequilíbrio de ligação), cobrindo todo o genoma, de forma que todos os alelos de interesse estarão em desequilíbrio de ligação com pelo menos um ou mais marcadores genotipados e portanto devidamente capturados nos modelos preditivos.

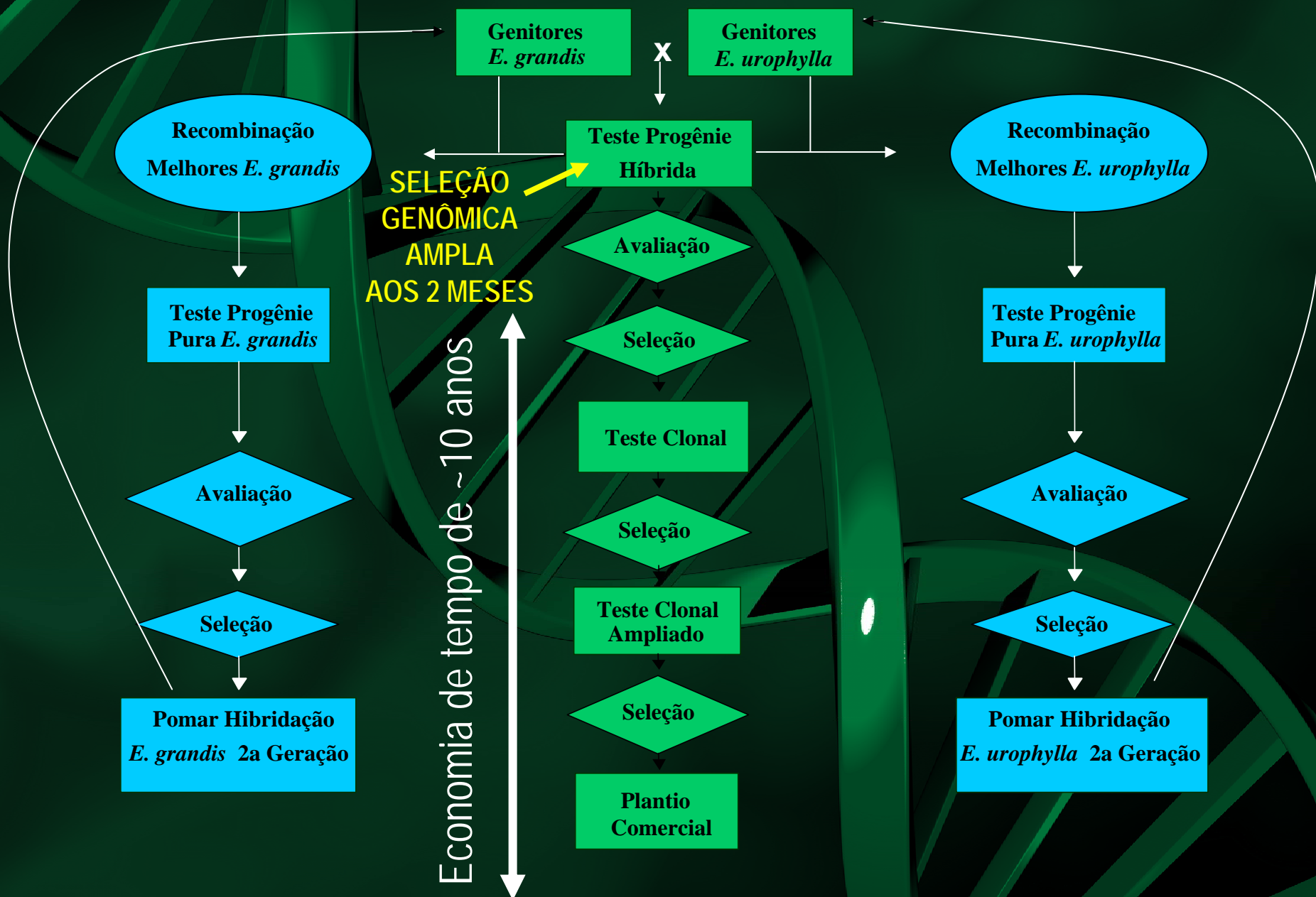


■ Gene/QTL que controla a característica alvo
● SNP genotipado

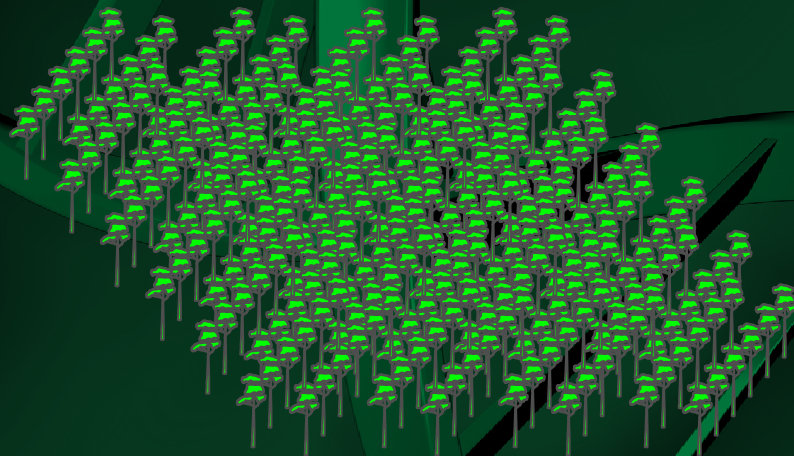
Programa de melhoramento de eucalipto para seleção de novos clones



Programa de melhoramento de eucalipto - Seleção recorrente



Seleção genômica ampla (*Genomic selection ou Genome-wide selection*)



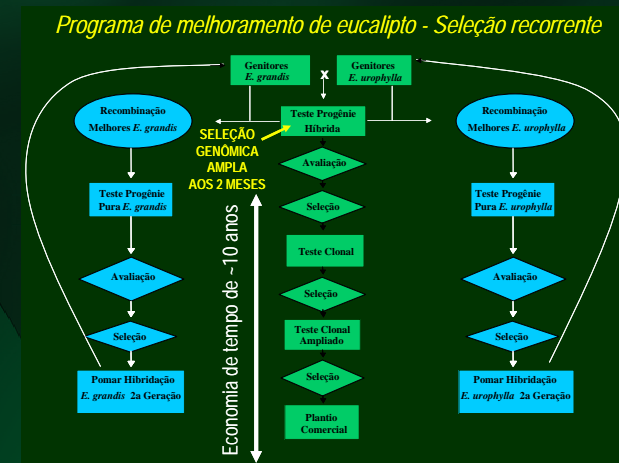
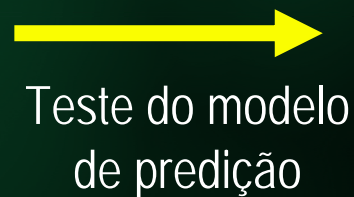
População de descoberta N = 800 clones



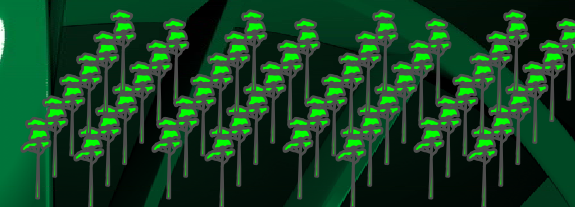
$$Y = Xb + Zh + e$$

Geração das equações de predição

Genotipagem de SNPs
à razão de $8Ne/100$ cMorgan



Aplicação na seleção assistida precoce



População de validação
N = 200

EUCAGEN - Eucalyptus Genome Network

Azores June 2007



~130 scientists in 18 countries

SEQUENCING OF THE *EUCALYPTUS* GENOME: A PROPOSAL TO DOE-JGI



Prepared by

Zander Myburg, Dario Grattapaglia, Brad Potts, Carlos Labate, Gerd Bossinger,
Margaret Byrne, Rena Vallancourt, Ron Sederoff, Simon Southerton

and other members of

the *Eucalyptus* Genome Network
(EUCAGEN)

18 June 2006

news

JGI in the News

Press Releases

Scientific Publications

Press Release: June 8, 2007

DOE Joint Genome Institute Announces 2008 Genome Sequencing Targets

Eucalyptus, Foxtail Millet, Red Algae, and Novel Microbial Communities Added to Growing Bioenergy and Carbon Cycling Portfolio

WALNUT CREEK, CA—Toward the goal of harnessing the power of nature through DNA sequencing, the DOE Joint Genome Institute (DOE JGI) has announced the latest Community Sequencing Program (CSP) portfolio. These plant and microbial targets—most with implications for helping wean the nation's dependence on fossil fuel—total some 21 billion nucleotides of DNA sequence capacity allocated to public projects submitted through the CSP for fiscal year 2008.

"This year's selections are completely aligned with the CSP mission, that is, selecting DOE-relevant organisms with the large and diverse communities of investigators," said Jim Bristow, DOE JGI Deputy Director and manager of the CSP. "The response to this year's program, with over 120 submissions, demonstrates an increasing desire to fuel discovery with DNA sequence information—which DOE JGI makes freely available through its web portals and the public databases."

Among the highest profile of these projects, and largest, with a 600-million-nucleotide genome, is the eucalyptus tree genome—geared to the generation of resources for renewable energy—led by Alexander Myburg of the University of Pretoria, South Africa, with Gerald Tuskan of Oak Ridge National Laboratory (and DOE JGI), and Dario Grattapaglia, of EMBRAPA Genetic Resources and Biotechnology (Brazil).



Eucalyptus: sequencing a global tree genome for energy, fiber, and wood



Alexander A. Myburg, Dario Grattapaglia, Gerald A. Tuskan,
Jeremy Schmutz, Daniel S. Rokhsar, Kerrie Barry, Jim Bristow,
The Eucalyptus Genome Network (EUCAGEN)

GENOMICS OF ENERGY & ENVIRONMENT

U.S. DEPARTMENT OF ENERGY
JOINT GENOME INSTITUTE (JGI)

THIRD ANNUAL
USER MEETING

MARRIOTT HOTEL
WALNUT CREEK, CALIFORNIA
MARCH 26-28, 2008



Eucalypts in tropical sites accumulate biomass 4X faster than conifers in temperate and 2X in tropical sites

Pinus taeda (Loblolly pine)

15 m³/ha/yr (US) 30 m³/ha/yr (Brazil)

Eucalyptus grandis

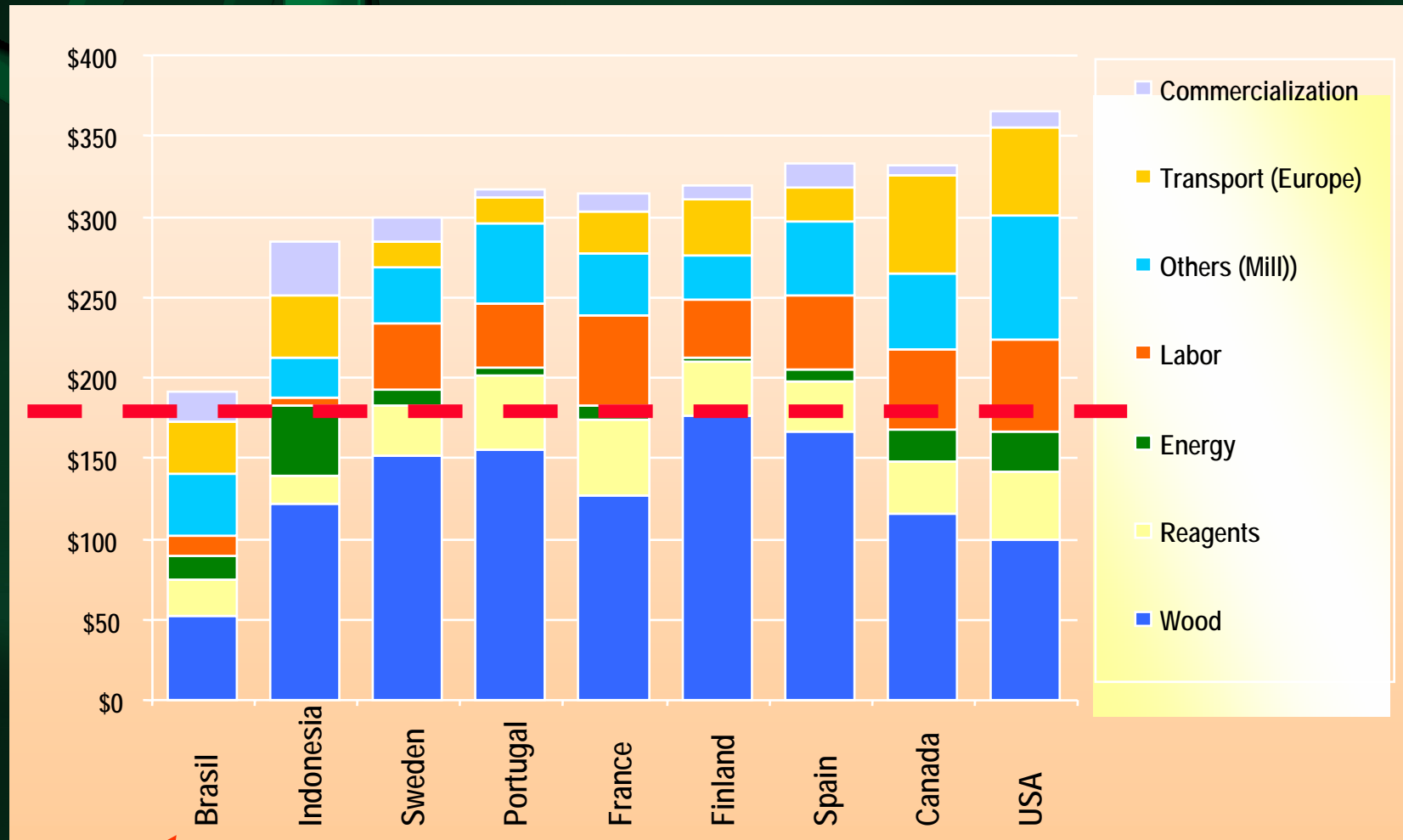
25 m³/ha/yr (Portugal) 50-60 m³/ha/yr (Brazil)



- ✓ CO₂ sequestration of Eucalyptus: 10 tons of CO₂/hectare/year from planting to harvesting
- ✓ Up to 14 CO₂/hectare/year in fast-growing tropical plantations
- ✓ Positive net carbon balance even when computing production of CO₂ used for energy from charcoal or as pulp and paper
- ✓ Present estimates of sequestration costs are \$100 to \$300/ton of CO₂ emissions avoided



Production and commercialization costs of short fiber pulp in several countries in 2002 (US\$/ton)





- ✓ Eucalyptus can be managed for high quality wood
- ✓ Certified , renewable wood
- ✓ Substitution to native tropical species
- ✓ Great way to fix carbon

"Green" steel with *Eucalyptus* charcoal



MINERAL COAL X EUCALYPT CHARCOAL IN STEEL PRODUCTION

- ✓ Mineral coal: generates 1.65 ton of CO₂, sequesters 1.54 ton of O₂ and generates 7 kg of SO₂ per ton of steel
- ✓ Eucalypt charcoal: sequesters 16.3 tons of CO₂ and regenerates 11.9 tons of O₂ from plantation through harvesting to steel production



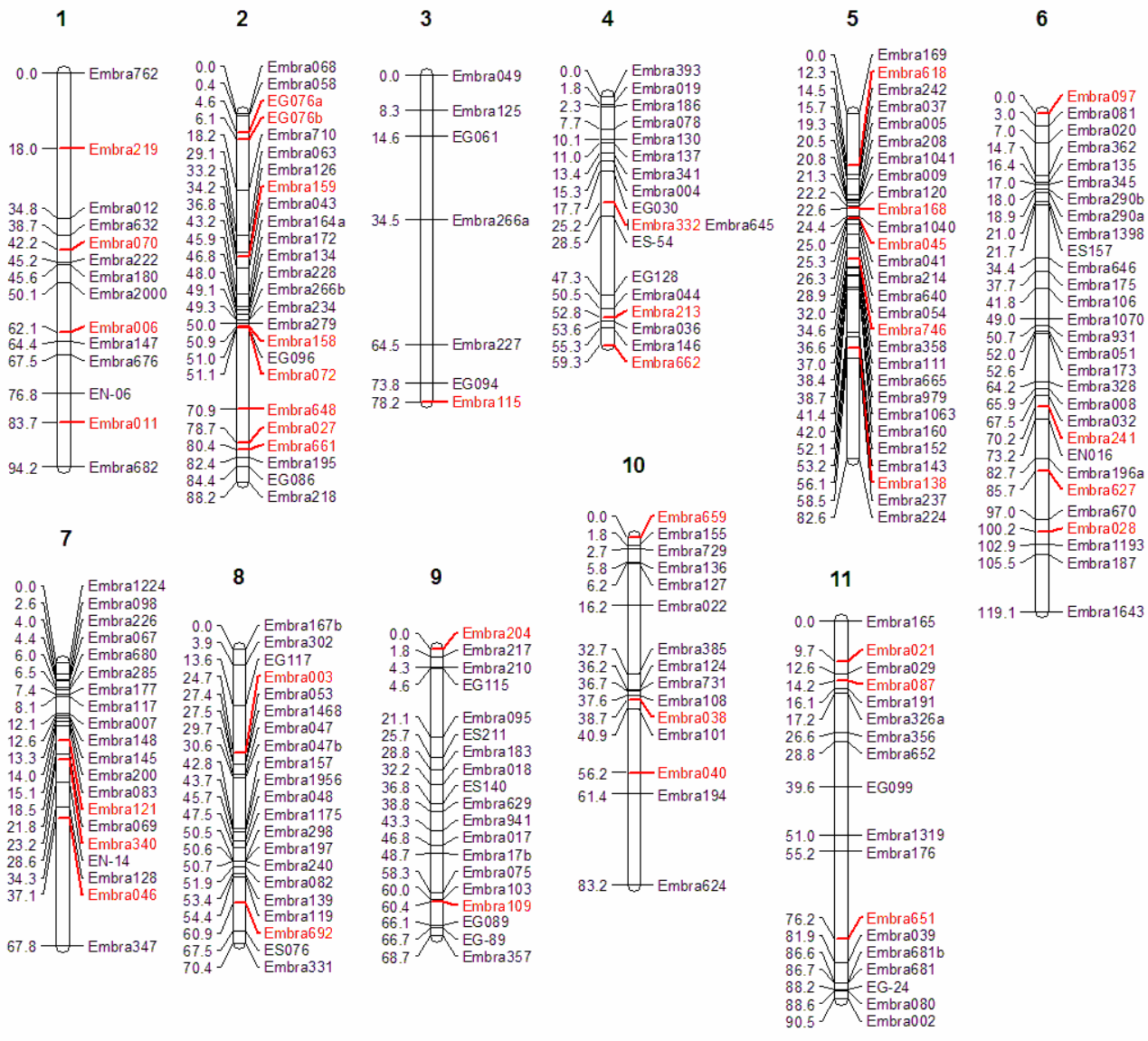
Pomar de sementes por mudas onde BRASUZ1 está localizado em Itapetininga, SP



Tissue collection for RNA:

JGI has planned to generate a large EST collection sepecifically for BRASUZ1 using 454 technology

Selected microsat marker set to cover the genome based on existing maps picking markers at evenly spaced distances in cM

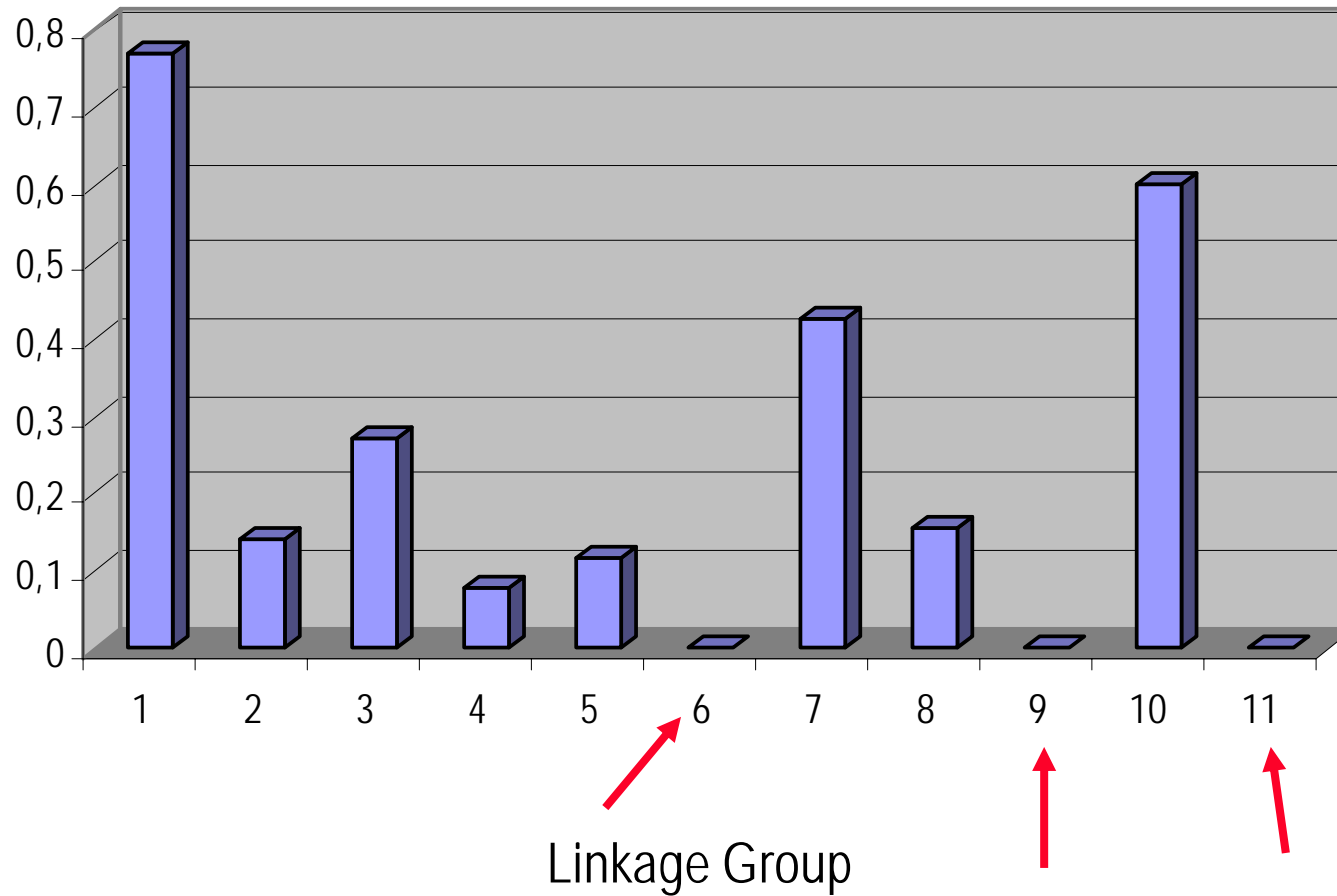


✓ Genotyped 212 microsatellites at an average spacing of 5 to 10 cM for a total of 900 cM

✓ Genotyped BRASUZ1 plus six other S1 full sibs available from the same orchard

RESULTS: putative homozygosity by linkage group

Significant variation across linkage groups suggests variable tolerance to homozygosity due to variable distribution of genetic load





General Information

- [Orders Home](#)
- [Ordering Help](#)
- [Contact Information](#)
- [AGI Home](#)

View BAC/EST Resources

- [EST, cDNA - Libraries & Clones](#)
- [Genomic - Libraries, Clones, & Filters](#)
- [Vectors](#)
- [Search for Genus or AGI library Name](#)
- [Library Manifest Search](#)

Shopping Cart

Arizona Genomics Institute
Department of Plant Sciences

[Back to Search](#)

Search Results

EG__Ba [Eucalyptus grandis cv.](#)
This Eucalyptus grandis (eucalyptus) BAC library has an average insert size of 157kb (Genome size: 660Mb) covering 17 genome equivalents. It was constructed by Xiang Song and Ammiraju Jetty at AGI. It was constructed in the HindIII site of pAGIBAC1 and contains 73728 clones. (192 plates) (4 filters per set) The library can also be referred to as Eucalyptus HindIII.

EG__Bb [Eucalyptus grandis cv.](#)
This Eucalyptus grandis (eucalyptus) BAC library has an average insert size of 135kb (Genome size: 660Mb) covering 15 genome equivalents. It was constructed by Xiang Song and Ammiraju Jetty at AGI . It was constructed in the BstYI site of pAGIBAC1 and contains 73728 clones. (192 plates) (4 filters per set) The library can also be referred to as Eucalyptus BstYI.

General Information

- [Orders Home](#)
- [Ordering Help](#)
- [Contact Information](#)
- [AGI Home](#)

View BAC/EST Resources

- [EST, cDNA - Libraries & Clones](#)
- [Genomic - Libraries, Clones, & Filters](#)
- [Vectors](#)
- [Search for Genus or AGI library Name](#)
- [Library Manifest Search](#)

Shopping Cart

Number of Products = 2
Subtotal = \$ 4608.00

[View Shopping Cart](#)

AGI Shopping Cart

[back](#)

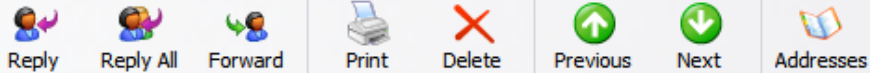
Please make sure that the information below is correct before going to "Checkout".

Part No.	Description	Clone Addresses	Quantity	Price
EG__Bb- Library	Eucalyptus grandis cv.		<input type="text" value="1"/> <input type="button" value="+"/> <input type="button" value="-"/> <input type="button" value="Remove"/>	US\$2304.00 each
EG__Ba- Library	Eucalyptus grandis cv.		<input type="text" value="1"/> <input type="button" value="+"/> <input type="button" value="-"/> <input type="button" value="Remove"/>	US\$2304.00 each
Totals				

Subtotal: US\$4608.00

Re: Eucalyptus update

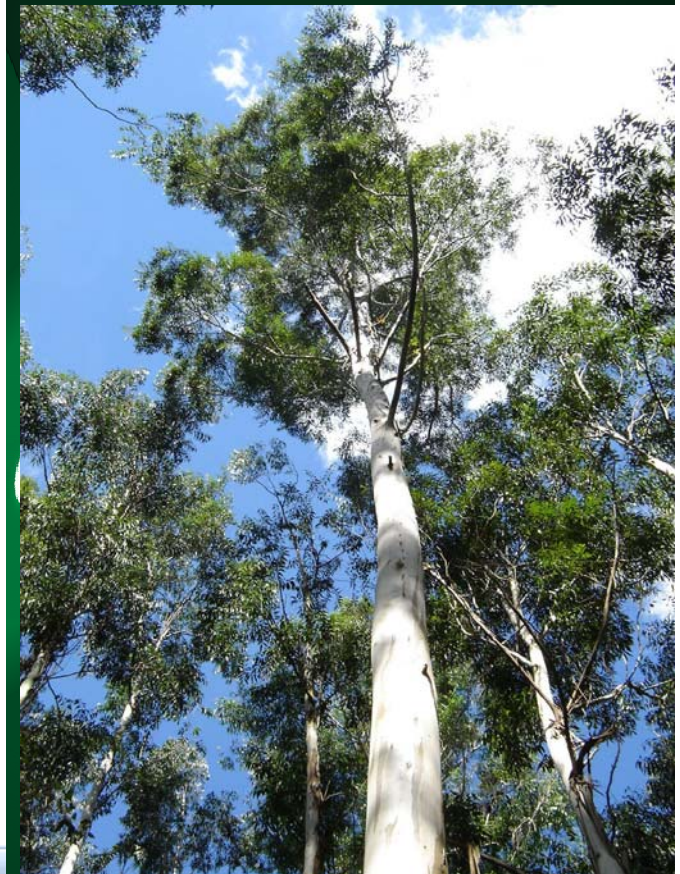
File Edit View Tools Message Help



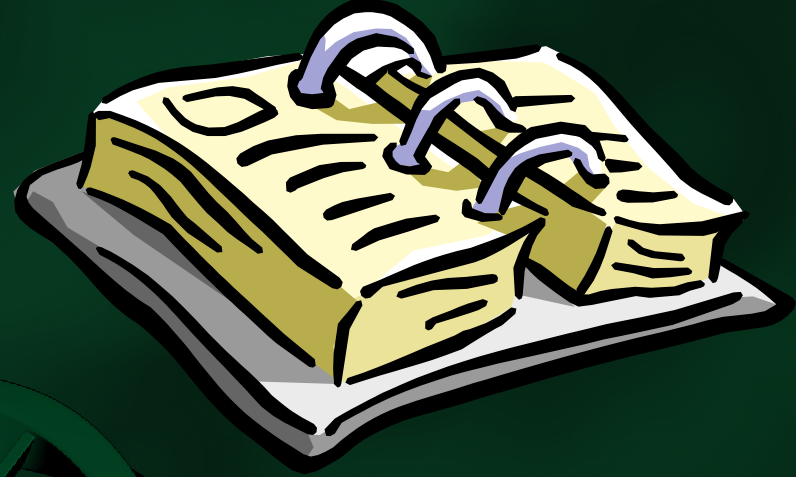
From: Jeremy Schmutz
Date: quinta-feira, 6 de novembro de 2008 15:35
To: Marcia McGowan
Cc: dario@cenargen.embrapa.br; gtk@ornl.gov; mahinch@arborgen.com; whrottm@arborgen.com; zander.myburg@fabi.up.ac
Subject: Re: Eucalyptus update
Attach: T2wallscafg1000.png (325 KB)

```
>
> Lower End Scaffolds Contigs Scaffold Size Basepairs Bp Coverage HQ Coverage GC %
> -----
> 0 14,662 48,483 530,484,220 450,579,352 84.94% 4.26x 39%
> 1,000 14,649 48,470 530,475,979 450,571,111 84.94% 4.26x 39%
> 2,500 14,619 48,440 530,415,064 450,510,196 84.94% 4.26x 39%
> 5,000 9,219 43,033 509,518,326 429,614,494 84.32% 4.32x 39%
> 10,000 4,819 38,371 478,081,583 398,323,354 83.32% 4.42x 39%
> 25,000 1,660 32,721 430,620,917 352,346,146 81.82% 4.53x 39%
> 50,000 1,123 30,929 412,530,814 335,989,730 81.45% 4.56x 39%
> 100,000 911 29,639 398,483,314 324,117,618 81.34% 4.58x 39%
> 250,000 526 25,404 337,702,344 284,814,243 84.34% 4.60x 39%
> 500,000 254 18,015 242,833,507 210,691,983 86.76% 4.73x 39%
> 1,000,000 78 9,117 123,720,307 111,634,029 90.23% 4.42x 39%
> 2,500,000 6 1,256 16,900,138 15,719,334 93.01% 4.44x 40%
> 5,000,000 0 0 0 0 -0- -0- -0-
>
> Reads
> -----
> Included 2,413,616 74.12%
> Low Quality 139,257 4.28%
> Multiple Instance 1,906 0.06%
> Other 92 0.00%
> Unplaced 701,011 21.53%
> Vector Or Host 672 0.02%
>
> -----
> Total 3,256,554
```

Sequenciamento do genoma completo de *Eucalyptus grandis* Status em 6/11/2008
2,9 Gigabases sequenciadas
4,5 coberturas
~480 Megabass montadas



O genoma é como uma lista telefônica



LISTA TELEFÔNICA

- ✓ Nomes de pessoas
- ✓ Sobrenomes comuns entre pessoas podem indicar relacionamento, função
- ✓ Endereço, bairro, vizinhança
- ✓ **Profissão, função ??**
- ✓ **Interações com outras pessoas e redes de relacionamento ??**

SEQUÊNCIA DO GENOMA

- ✓ Nomes de genes
- ✓ Sequências comuns entre genes indicam relacionamento de função
- ✓ Local no genoma, vizinhanças etc.
- ✓ **Função ??**
- ✓ **Interações com outros genes e cadeias metabólicas ??**

O GRANDE DESAFIO PARA O PRÓXIMO SÉCULO É DESVENDAR A FUNÇÃO DOS GENES E SUAS INTERAÇÕES NA DEFINIÇÃO FINAL DA COMPLEXIDADE DO FUNCIONAMENTO DO ORGANISMO

Agradecimentos

- ✓ *Este é o trabalho de MUITA GENTE!!!!*
- ✓ Todos os participantes do GENOLYPTUS, professores, pesquisadores, técnicas das empresas, bolsistas e estudantes
- ✓ EMPRESAS PARTICIPANTES DO GENOLYPTUS
- ✓ MINISTÉRIO DE CIÊNCIA E TECNOLOGIA
 - ✓ FINEP - Fundo Verde Amarelo
 - ✓ PADCT (1995-1999)
 - ✓ CNPq
 - ✓ CAPES



OBRIGADO PELA ATENÇÃO

dario@cenargen.embrapa.br