

UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
INSTITUTO DE BIOCÊNCIAS
CAMPUS DE BOTUCATU

Leonardo Curi Martin

**Identificação de SNPs (Single Nucleotide
Polymorphisms) no gene colina
monooxigenase relacionado ao metabolismo
da glicina betáina em *Eucalyptus***

Botucatu-SP
Dezembro - 2010

UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
INSTITUTO DE BIOCÊNCIAS
CAMPUS DE BOTUCATU

**Identificação de SNPs (Single Nucleotide
Polymorphisms) no gene colina
monooxigenase relacionado ao metabolismo
da glicina betaína em *Eucalyptus***

Leonardo Curi Martin

Dissertação apresentada ao Instituto de
Biotecnologia da UNESP - Campus de Botucatu, para
a obtenção do título de Mestre em Ciências
Biológicas - Genética.

Orientador: Celso Luís Marino

Botucatu-SP
Dezembro - 2010

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. E TRAT. DA INFORMAÇÃO
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CAMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: **ROSEMEIRE APARECIDA VICENTE**

Martin, Leonardo Curi.

Identificação de SNPs (Single Nucleotide Polymorphisms) no gene colina monooxigenase relacionado ao metabolismo da glicina betaína em Eucalyptus / Leonardo Curi Martin. – Botucatu, 2010

Dissertação (mestrado) – Instituto de Biociências de Botucatu, Universidade Estadual Paulista, 2010

Orientador: Celso Luis Marino

Capes: 20203004

1. Genética vegetal. 2. Eucalipto. 3. Código genético. 4. Plantas – Melhoramento genético.

Palavras chave: Déficit hídrico; Eucalyptus; Glicina betaína; Haplótipo; SNPs.

Agradecimentos

Agradeço ao **Prof. Dr. Celso Luis Marino** pela orientação, incentivo e confiança;

Aos meus pais, **Gilberto** e **Marlene**, meu irmão, **Gil**, e minha tia, **Mirva**, pelo apoio incondicional e por me ajudar a seguir com segurança por esta vida;

À minha namorada, **Priscila**, pelo apoio, carinho e compreensão;

À todos do laboratório **CAGEN**, que estiveram dispostos a ajudar e contribuíram de alguma forma nesse trabalho: **Tânia, Suzana, Vanusa, Cecília, Cíntia, Helena, Karine, Juliana**, e de maneira especial ao **Júlio** e **Andréia** pelas horas de dedicação e empenho;

Aos colegas que dedicaram um tempo para discussão do projeto: **Bonsai, Babito, Godiness, Pituta e Douglas**;

À **FIBRIA Celulose**, pelo fornecimento do material vegetal e colaboração;

À **CNPq**, pela bolsa concedida;

À todos que de forma direta ou indireta, contribuíram para realização deste trabalho.

Resumo

Estresses abióticos, como a seca, podem reduzir significativamente os rendimentos no cultivo de espécies comercialmente importantes, tais como o eucalipto na produção de celulose e papel. Quando submetidas às condições de déficit hídrico, as plantas desenvolvem alguns mecanismos de defesa. As betaínas participam destes mecanismos, sendo seu soluto mais comum a glicina betaína. Esta é uma amina quaternária distribuída extensamente em diversas espécies de plantas superiores, sintetizada em elevadas taxas, tendo como função, manter a turgescência celular. A identificação e estudo de genes relacionados à tolerância à seca são importantes para os programas de melhoramento florestal. Este estudo teve por objetivo identificar SNPs (Single Nucleotide Polymorphisms) no gene colina monooxigenase relacionado ao metabolismo da glicina betaína em *Eucalyptus*. A sequência do gene em estudo foi encontrada em genomas de plantas modelos através do banco de dados GenBank, sendo as mesmas, utilizadas para a procura de similaridades no banco de dados de ESTs de eucalipto FORESTs – FAPESP. Após a constatação de homologia no banco de dados de eucalipto, foram confeccionados oito pares de “primers” para flanquear essas regiões, sendo estes, avaliados, amplificando fragmentos únicos. Depois de realizado o seqüenciamento do gene colina monooxigenase, foi utilizado a ferramenta BLAST no GenBank, confirmando com sucesso a identidade da sequência. Em seguida, as sequências foram alinhadas e os SNPs encontrados. No total, foram identificados 49 SNPs, sendo 12 em regiões codificantes e 37 em regiões UTRs e íntrons. Somente os SNPs localizados nas regiões codificantes foram utilizados neste trabalho, sendo 83,3% deles possuindo mutações sinônimas e 16,7% não-sinônimas. Em seguida, foi utilizada uma população mais abrangente composta de *E. grandis*, *E. urophylla* e o híbrido “Urograndis” para realização da genotipagem dos SNPs, constatando-se a formação de 18 haplótipos e 16 possíveis combinações haplotípicas, as quais, revelaram que alguns genótipos foram exclusivos para as espécies *E. grandis* e *E. urophylla*. Com o intuito de diminuir custos e permitir a genotipagem em larga escala destas mutações, em populações de eucalipto, dez “primers” de SNP foram desenhados e alguns já utilizados com sucesso nas genotipagens de indivíduos através da técnica de extensão alelo-específica, podendo então, serem usadas mais adiante para estudos de genotipagem de SNPs em populações contrastantes ao déficit hídrico e estudos populacionais em eucalipto.

Palavras-chave: *Eucalyptus*, SNPs, déficit hídrico, glicina betaína, haplótipo

Abstract

Abiotic stress, such as drought, can reduce significantly the yield in the important commercially species crop, such as the eucalyptus in the cellulose and paper. When submitted to the water stress conditions, the plants develop some defense mechanisms. The betaines are part of these mechanisms, being their solute more common to the glycine betaine. This is a quaternary amine extensively distributed in several species of higher plants synthesized in high taxes with the function of maintaining the cellular turgescence. Identifying and studying the genes related to the drought tolerance are important for the forest improvement programs. This work aimed at identifying SNPs (Single Nucleotide Polymorphisms) in choline monoxygenase related to the glycine betaine in *Eucalyptus*. The gene sequence was found in model plants genomes through the databank GenBank, being used for searching similarities in the databank ESTs of eucalyptus FORESTs – FAPESP. After noticing homology in the eucalyptus databank, eight pairs of primers were made to amplify these regions, being evaluated, amplifying unique fragments. After the choline monoxygenase sequencing was performed, it was used the BLAST in the GenBank, confirming successfully the sequence identity. Then, the sequences were aligned and the SNPs were found. In the total, 49 SNPs were identified, being 12 in coding regions and 37 in UTRs and intron regions. Only the SNPs located in coding regions were used in this work, being 83.3% of the SNPs with synonymous mutation and 16.7% non-synonymous. Following, it was used a wider population made up of de *E. grandis*, *E. Urophylla* and the hybrid “Urograndis” to perform the SNPs genotyping, establishing the formation of 18 haplotypes and 16 possible haplotypal combinations which revealed that some genotypes were exclusive for the species *E. grandis* and *E. urophylla*. With the objective of decreasing costs and permitting the genotyping of these mutations in large scale in population of eucalyptus, ten primers of SNP were designed and some were already used with success in genotyping of individuals by the allele-specific extension technique, being able to be further used for studies of SNPs genotyping in populations contrasting to the water stress and populational studies in eucalyptus.

Key words: *Eucalyptus*, SNPs, water stress, glycine betaine, haplotype

LISTA DE FIGURAS

- Figura 1- Levantamento estatístico feito pela ABRAF (Associação Brasileira de Produtores de Florestas Plantadas), mostrando a expansão no plantio de eucalipto, por hectare, nos estados brasileiros, em 2009. O estado de Minas Gerais liderou o plantio, seguido do estado de São Paulo..... 17
- Figura 2 - Levantamento estatístico feito pela ABRAF no ano de 2009 mostrando o destino da madeira de eucalipto plantada no Brasil, para as demais utilidades..... 18
- Figura 3 - Rota da biossíntese da glicina betaína em plantas superiores. A colina é oxidada através da reação de dois passos: (1) É catalizada pela CMO e a reação (2) é catalizada pela BADH (Rhodes e Hanson, 1993; Ashraf e Foolad, 2007)..... 28
- Figura 4 - Ilustração mostrando, em destaque, o polimorfismo (CG→TA) de único nucleotídeo, denominado (SNP). 30
- Figura 5 - Ilustração retirada do software ChromasPro versão 1.42 mostrando a metodologia aplicada para o desenho dos “primers” utilizados nas reações de minisequenciamento para genotipagem dos SNPs. A sequência complementar, da figura destacada em preto, foi utilizada como “primer” 44
- Figura 6 - Ilustração retirada e editada do site de rotas metabólicas de plantas Plantcyc, mostrando a atuação das 2 enzimas (colina monooxigenase e betaína aldeído desidrogenase) na formação do osmoprotetor glicina betaína.. 47
- Figura 7 - Ilustração esquemática dos dois “reads” pertencentes ao gene colina monooxigenase (CMO) obtidos através do resultado da pesquisa no banco de dados de ESTs de eucalipto FORESTs – FAPESP, mostrando a localização do desenho de cada “primer” a partir de seus respectivos “reads” 50
- Figura 8 - Teste de amplificação do “primer” GB 2-2 em gel de agarose 1%, corado com brometo de etídio. Nas extremidades estão os marcadores de peso molecular conhecidos (Ladder 1Kb – Fermentas®). Para este teste foram utilizados, aleatoriamente, dez indivíduos divergentes, os quais, apresentaram amplificação de banda única.....51
- Figura 9: Cromatograma parcial, mostrando a ótima qualidade das sequências amplificadas pelas pelos “primers” (A) GB 2-2, (B) GB 2-3, (C) GB 2-4. Picos altos e bem espaçados são apresentados na figura.....52

Figura 10 - Ilustração esquemática comparando a distribuição de íntrons e éxons de CMO de eucalipto e arabidopsis. O alinhamento do seqüenciamento com os “reads” de CMO de ESTs do banco de dados de ESTs de eucalipto FORESTS - FAPESP resultou no esquema do gene em eucalipto. O gene de CMO em arabidopsis foi montado a partir da seqüência obtida no banco de dados de Arabidopsis (TAIR).....55

Figura 11 - Alinhamento realizado pelo programa ChromasPro versão 1.42, comparando indivíduos divergentes na procura de SNPs. A região destacada em vermelho mostra a localização dos SNPs nos indivíduos que possuem alelos em homozigose e heterozigose (A), e indivíduos que possuem alelos diferentes em homozigose (B).....57

Figura 12 - Genotipagem dos SNP utilizando a técnica SNaPshot™. As análises foram realizadas pelo software Peak Scanner versão 1.0. Em (A), os dois alelos gerados pelo marcador L 25, e em (B), os dois alelos gerados pelo marcador L 26. Os picos em laranja, são os marcadores moleculares de pesos conhecidos (GS 120 LIZ – Applied Biosystems).....70

Figura 13 - Genotipagem dos SNPs utilizando a técnica SNaPshot™ em uma reação multiplex. O primeiro pico em preto, evidencia o alelo gerado através do uso do marcador L 25, e o segundo pico em preto, mostra o alelo gerado pelo uso do marcador L 26. Os picos em laranja, são os marcadores moleculares de pesos conhecidos (GS 120 LIZ – Applied Biosystems).....71

LISTA DE TABELAS

Tabela 1 - Enzima selecionada através do banco de dados do GenBank. São apresentados o EC-Number da enzima, nome da enzima, espécie obtida, número de identificação no GenBank, acesso no Tair e referência atual sobre a enzima..... 49

Tabela 2 – “Primers”, forward e reverse, utilizados para amplificar o gene colina monooxigenase, em eucalipto. São apresentados a identidade dos “primers”; temperatura de pareamento, em graus Celsius ($^{\circ}\text{C}$); tamanho do fragmento amplificado, em pares de base (pb); e as sequências de cada “primer”..... 51

Tabela 3 - Resultado do BlastX com a sequência alinhada, obtida na amplificação dos oito pares de “primers” construídos para cobrir todo o gene colina monooxigenase. São apresentados: Número de acesso no Genbank, nome da enzima e organismo possuidor, score e máxima identidade.. 53

Tabela 4 - Determinação das regiões de éxon e íntron da sequência gênica de CMO de arabidopsis e eucalipto. Através de comparações, observa-se que a região exônica das duas espécies analisadas possuem tamanhos parecidos, diferindo-se apenas no tamanho dos íntrons.. 56

Tabela 5 – SNPs localizados em regiões codificantes e selecionados para desenho dos “primers” para SnapShot TM(Applied Biosystems). Houve alteração de aminoácidos em dois dos doze sítios de SNP. São mostrados na tabela a identidade do SNP, os dois ou 3 possíveis códons (1),(2) e (3) respectivos aminoácidos e o tipo de mutação.. 61

Tabela 6 – Genotipagem dos SNPs em 59 indivíduos pertencentes ao gênero *Eucalyptus*. Indivíduos com iniciais (CL, Ug = *Urograndis*), (G = *E. grandis*) e (U = *E. urophylla*). Oito sítios polimórficos foram identificados somente em regiões exônicas amplificadas pelos “primers” GB 2-3 e GB 2-4. 62

Tabela 7 – Formação dos possíveis haplótipos gerados pelos indivíduos utilizados neste estudo. São apresentadas as identidades dos haplótipos, o número de ocorrências (levando-se em conta o par de cromossomos) e porcentagem ocorrida nas amostras.. 64

Tabela 8 - Reconstrução haplotípica com as respectivas frequências para as espécies de eucalipto em estudo. São apresentados os 18 possíveis haplótipos estimados através da genotipagem do indivíduos. Somente os haplótipos 1, 11 e 13 foram comum para os três grupos.....65

Tabela 9 - Genótipos formados a partir da combinação dos possíveis haplótipos. São apresentados as combinações haplotípicas, número de indivíduos onde ocorreram as combinações, porcentagem dos genótipos nas amostras e porcentagem de ocorrência por grupo.....66

Tabela 10 - “Primers” que serão utilizados para reação de extensão alelo-específica dos polimorfismos de único nucleotídeo. São mostrados a identidade dos “primers”, as respectivas sequências e as possíveis mutações. Para facilitar na discriminação dos “primers”, foi adicionada uma cauda poli A ou C de diferentes tamanhos com o intuito de utilizar estes oligos em reação multiplex. Alguns destes marcadores já foram utilizados para este trabalho.....69

Sumário

Resumo

Abstract

1 Introdução.....	12
2 Revisão da Literatura.....	15
2.1 A importância do eucalipto.....	15
2.2 Estresses Abióticos.....	19
2.3 O fenômeno das secas e importância da água.....	20
2.4 Mecanismos de resposta da planta ao déficit hídrico.....	22
2.5 Osmoprotetores.....	24
2.6 Glicina betaína (GB).....	25
2.7 Colina monooxigenase.....	27
2.8 Biossíntese da glicina betaína.....	27
2.9 Marcadores Moleculares.....	28
2.10 SNP – Single Nucleotide Polymorphisms.....	30
2.11 Diversidade nucleotídica e haplótipos de SNP.....	32
3 Objetivos.....	35
4 Material e Métodos.....	36
4.1 Material.....	36
4.2 Métodos.....	37
4.2.1 Extração de DNA.....	37
4.2.2 Quantificação e diluição do DNA.....	39
4.2.3 Procura das enzimas relacionadas ao metabolismo da glicina betaína.....	39
4.2.4 Seleção dos iniciadores específicos e condições de amplificação.....	40

4.2.5 Purificação dos produtos de PCR.....	41
4.2.6 Reação de seqüenciamento.....	41
4.2.7 Precipitação das reações de seqüenciamento.....	42
4.2.8 Alinhamento das sequências.....	43
4.2.9 Análises dos SNPs.....	43
4.2.10 Caracterização dos SNPs por um método semi-automático.....	43
4.2.11 Seleção dos oligonucleotídeos flanqueadores.....	44
4.2.12 Reação de miniseqüenciamento.....	45
4.2.13 Purificação das reações de miniseqüenciamento.....	45
4.2.14 Análises da estrutura genética populacional.....	46
4.2.15 Análise da diversidade.....	46
5 Resultados e Discussão	47
5.1 Enzimas encontradas e relacionadas ao metabolismo da glicina betaína.....	47
5.2 Busca de homólogos para a enzima colina monooxigenase em banco de dados de eucalipto.....	49
5.3 Desenho dos oligonucleotídeos específicos.....	49
5.4 Amplificação dos oligonucleotídeos específicos.....	50
5.5 Sequenciamento do gene colina monooxigenase e análise <i>in silico</i> no GenBank.....	52
5.6 Análise <i>in silico</i> das sequências correspondentes ao gene colina monooxigenase nos bancos de dados de eucalipto.....	54
5.7 Identificação dos SNPs no gene colina monooxigenase em eucalipto.....	57
5.8 Classificação dos SNPs localizados nas regiões codificantes.....	59
5.9 Genotipagem dos SNPs.....	62
5.10 Análises estatísticas da genotipagem.....	64
5.11 Desenho e teste de extensão dos oligonucleotídeos específicos para genotipagem dos SNPs.....	69
6 Considerações finais.....	72

7 Conclusões.....	73
8 Referências.....	74

1 Introdução

Introduzido comercialmente no país no início do século XX, o eucalipto adaptou-se bem às condições climáticas do Brasil, sendo sua cultura, de grande importância econômica, ambiental e social. Este gênero florestal destina-se à produção de celulose, papel, carvão, tábuas, chapas aglomeradas entre outros, sendo o mais propagado no mundo (Scarpinella, 2002).

Atualmente, o plantio de *Eucalyptus* no Brasil já ocupa mais de 4,5 milhões de hectares (0,49% do território nacional), com uma taxa de crescimento anual de 7,4% (ABRAF, 2010), movimentando na economia brasileira algo ao redor dos US\$ 37,3 bilhões, que equivale a aproximadamente 3,5% do PIB nacional sendo responsável por cerca de 6,9 milhões de empregos diretos e indiretos no país (SBS, 2008).

Com essa grande expansão na cultura do eucalipto em nosso país, está havendo a necessidade da ocupação de novas áreas para plantio, que muitas vezes acabam por ser regiões que predominam longos períodos de seca. Estresses abióticos, como a seca, podem reduzir significativamente os rendimentos das lavouras e restringir as latitudes e os solos onde espécies comercialmente importantes podem ser cultivadas.

Quando submetidas ao déficit hídrico, as plantas desenvolvem alguns mecanismos de resposta a esse tipo de estresse, tais como os mecanismos de percepção de sinais. As betaínas fazem parte de um grupo de osmoprotetores que participam destes mecanismos, sendo o soluto mais comum, a glicina betaína (GB). Estes solutos compatíveis são moléculas (amina quaternária) ou íons atóxicos que não interferem no metabolismo e se acumulam geralmente no citoplasma, onde têm função de manter a turgescência celular, além de estabilizar proteínas e estruturas celulares nas condições ambientais (Bray et al., 2001). A glicina betaína é requerida para proteger a planta em

resposta aos vários tipos de estresse ambientais, mantendo o contrapeso da água entre a célula vegetal e o ambiente, estabilizando as macromoléculas (Chen e Murata, 2002; Rontein et al., 2002). As plantas sintetizam a glicina betaína por duas reações de oxidação da colina (colina \rightarrow betaína aldeído \rightarrow glicina betaína) (Rhodes e Hanson, 1993). O acúmulo de GB está associado a uma regulação positiva da expressão do gene para colina monooxigenase (CMO) e para betaína aldeído desidrogenase (BADH) concomitantemente, assim conduzindo a uma atividade enzimática elevada.

Existe um forte investimento no país para o melhoramento genético florestal, porém, o longo tempo necessário para a realização de um ciclo de melhoramento o torna um dos fatores limitantes. Daí a importância da incorporação de novas ferramentas que possibilitem reduzir esse tempo, como é o caso da aplicação da biotecnologia para melhorar a qualidade e aumentar a produtividade da madeira a cada ciclo de seleção.

A utilização de marcadores moleculares em programas de melhoramento genético, dentre outros fatores, permite uma avaliação prévia de indivíduos superiores, otimizando o processo de seleção dentro dessas espécies, além de possibilitar o estudo de caracterização de genótipo de indivíduos a partir de células e tecidos em qualquer estágio de desenvolvimento da planta (Ferreira e Grattapaglia, 1998). Vários marcadores moleculares já foram desenvolvidos para serem aplicados em espécies florestais de interesse (Junghans, 2000; Moretzsohn et al., 2000; Martins Filho et al., 2002; Benko-Iseppon et al., 2003). Em *Eucalyptus grandis*, foi identificado com sucesso o gene de resistência à *Puccinia psidii* (Junghans et al., 2003).

Outra ferramenta importante e de grande utilidade são os bancos de dados de ESTs (*Expressed Sequence Tags*), onde inicialmente, os genes de interesse podem ser identificados, e uma avaliação prévia dos SNPs efetuada. O banco de dados de ESTs de eucalipto FORESTs – FAPESP (Vicentini et al., 2005) é composto por 123.889 “reads”

obtidos de EST de bibliotecas de cDNA derivadas de tecidos de diferentes organismos de *E. grandis*, *E. globulus*, *E. saligna*, *E. urophylla* e *E. camaldulensis* (Guerrini et al., 2005). Santos (2005) estudou a diversidade nucleotídica (SNPs) em dois genes chaves na via de biossíntese de lignina utilizando com sucesso o banco de dados de ESTs gerado no projeto Genolyptus para detecção de SNPs.

Os SNPs *Single Nucleotide Polymorphism*, ou seja, polimorfismo de base individual ou única é a definição utilizada hoje para definir a variação da seqüência observada em uma posição individual do DNA. Os SNPs são distribuídos por todo o genoma, porém um SNP localizado na região codificadora pode ter impacto relevante na formação da proteína. Um SNP intrônico pode influenciar o *splicing* do mRNA (Krawezak et al., 1992), assim como, um SNP no promotor pode influenciar a expressão gênica (Drazen et al., 1999). Além dos estudos de diversidade populacional, os SNPs podem ser utilizados para identificação de variantes alélicas em regiões codificadoras levando a associação entre o genótipo observado e o fenótipo. Com a utilização do seqüenciamento direto de fragmentos específicos do genoma amplificados por PCR, e subsequente alinhamento e comparação das seqüências, permitiu-se a procura de SNPs em regiões de interesse. Vários estudos utilizaram essa metodologia para buscar SNPs em genes de interesse (Lai et al., 1998; Calvo et al., 2006; Abatepaulo et al., 2008; Ganal et al., 2009; Liao e Lee, 2010).

Portanto, uma melhor compreensão sobre a fisiologia das plantas e a aplicação de metodologias de biologia molecular, poderão desempenhar um papel chave nesse processo, contribuindo assim, com os programas de melhoramento florestal.

2 Revisão da Literatura

2.1 A importância do eucalipto

O eucalipto pertence à divisão Angiospermae, classe Dicotyledonea, ordem Myrtales, família Myrtaceae e gênero *Eucalyptus*. Possui mais de 900 espécies, variedades e híbridos, distribuídos em 7 subgêneros, sendo o principal deles o *Symphyomyrtus*, com mais de 300 espécies, dentre as quais estão as espécies mais plantadas para fins comerciais, *E. grandis*, *E. globulus*, *E. urophylla*, *E. calmadulensis*, *E. saligna* e *E. tereticornis* (Fao, 1981; Eldridge, 1993; Mora e Garcia, 2000). O gênero é originário da Austrália, com exceção de *E. urophylla* e *E. deglupta*, que são naturais do Timor e Papua Nova Guiné, respectivamente (Pryor, 1985).

Este gênero se reproduz predominantemente por alogamia, ou seja, por fecundação cruzada, em sua maioria com a ajuda de insetos, particularmente por abelhas (Quintin, 2001). A autofecundação acaba sendo comum, ocorrendo em 10% a 30% neste gênero.

Uma característica marcante do gênero é a presença de opérculo. Os frutos são capsulares e deiscentes, liberando sementes de cor preta a amarelada e de tamanho variável, de menos de 1 mm até 2 cm (Cavalcanti, 1963). As folhas variam de acordo com a maturidade da planta, mas normalmente são coriáceas, lanceoladas, com elevada quantidade de cutina e ricas em esclerênquima. Normalmente as folhas são alternadas, com algumas espécies mostrando folhas opostas. Os troncos apresentam pouca ramificação e normalmente são lisos e com casca descídua. Possuem, em geral, alborno delgado e claro, e cerne de amarelo a avermelhado (Fao, 1981).

Dentre as espécies de *Eucalyptus* há arbustivas e arbóreas, com algumas espécies alcançando até 100 metros de altura, e todas possuem troncos lenhosos.

Em termos genéticos, todas as espécies plantadas no mundo possuem $2n = 22$ cromossomos (Eldridge et al., 1993), e o tamanho do genoma varia desde cerca de 380 milhões de pares de base (Mpb) em *C. citriodora* até por volta de 650 Mpb em *E. grandis* (Grattapaglia e Bradshaw, 1994) e 644 Mpb em *E. globulus* (Pinto et al., 2004).

A hibridização entre espécies de diferentes subgêneros em geral não ocorre. Porém, é possível entre espécies dentro de subgênero, especialmente se pertencem à mesma secção (Griffin et al., 1988; Assis, 2000).

A formação de povoamentos florestais para fins econômicos originou-se no Brasil, no início do século XX, com a introdução do gênero *Eucalyptus*, na região de Rio Claro-SP, pela Companhia Paulista de Estrada de Ferro, visando à produção de dormentes (Silva, 1994).

O plantio de *Eucalyptus* em nosso país já ocupa mais de 4,5 milhões de hectares (0,49% do território nacional), com uma taxa de crescimento anual de 7,4%. A maior parte destes hectares cobertos por *Eucalyptus* pertence a empresas produtoras de papel e celulose, cultivada, em sua maioria, com clones ou híbridos interespecíficos superiores provenientes de programas de melhoramento genético. Graças a estes programas florestais, nas últimas décadas, a produtividade deste gênero foi elevada, a área de plantio foi aumentada e a qualidade da madeira foi melhorada, de forma a atender a demanda das indústrias de papel e celulose (ABRAF, 2010; Embrapa, 2009) (Figura 1).

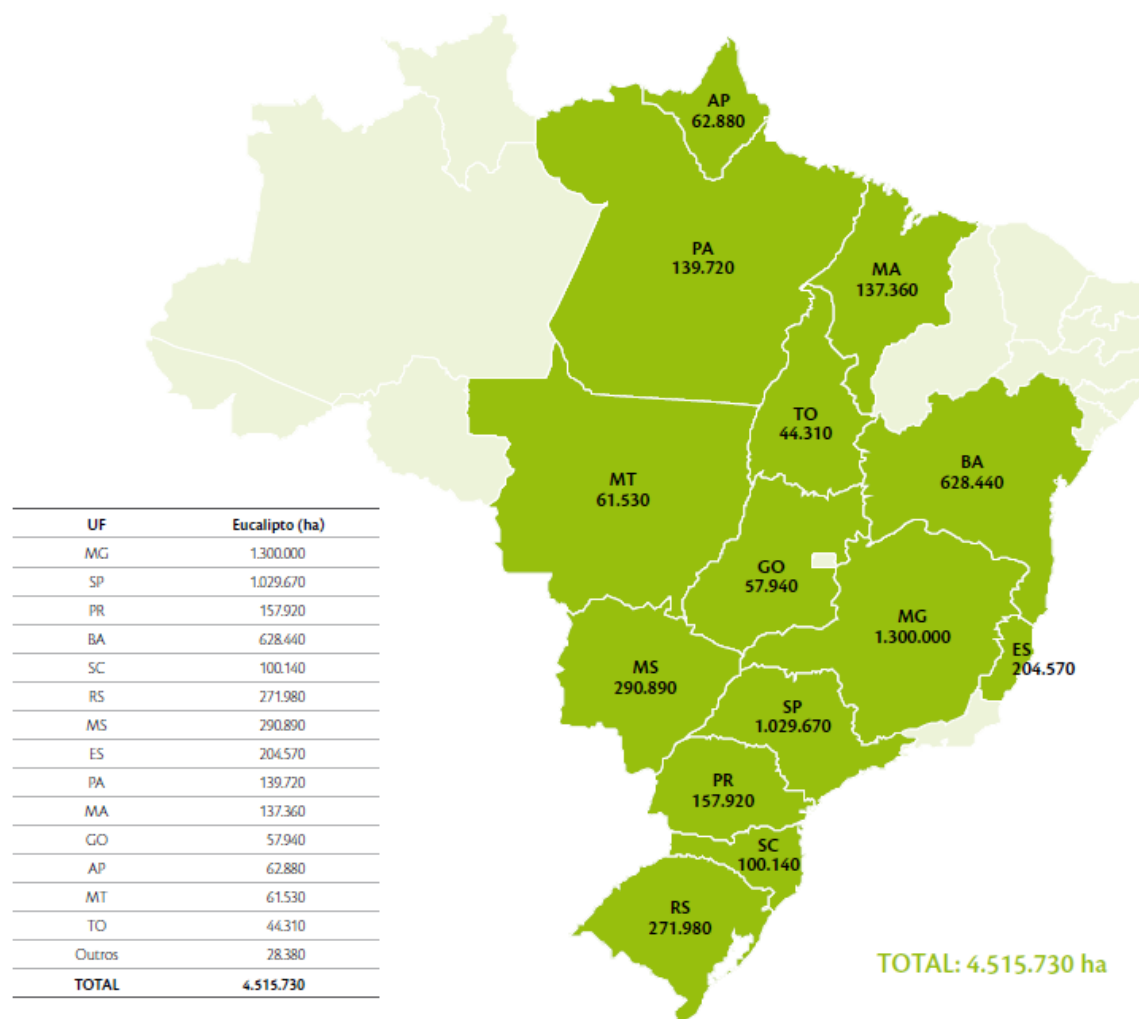


Figura 1: Levantamento estatístico feito pela ABRAF (Associação Brasileira de Produtores de Florestas Plantadas), mostrando a expansão no plantio de eucalipto, por hectare, nos estados brasileiros, em 2009. O estado de Minas Gerais liderou o plantio, seguido do estado de São Paulo.

O eucalipto movimenta no Brasil algo ao redor dos US\$ 37,3 bilhões, que equivale a aproximadamente 3,5% do PIB nacional e foi responsável por cerca de 6,9 milhões de empregos diretos e indiretos no país (SBS, 2008)

O Brasil já figura como o maior exportador mundial de celulose de fibra curta de eucalipto (Figura 2). A produção brasileira de celulose foi de 10 milhões de toneladas em 2007. Em 2008, o Brasil subiu de sexto para quarto lugar entre os produtores mundiais de celulose, com 12,7 milhões de toneladas produzidas, e ganhou posições

entre os principais fabricantes de papel do mundo, com a marca de 9,4 milhões de toneladas (SBS, 2008; Bracelpa, 2009).

O Mercado do Eucalipto no Brasil

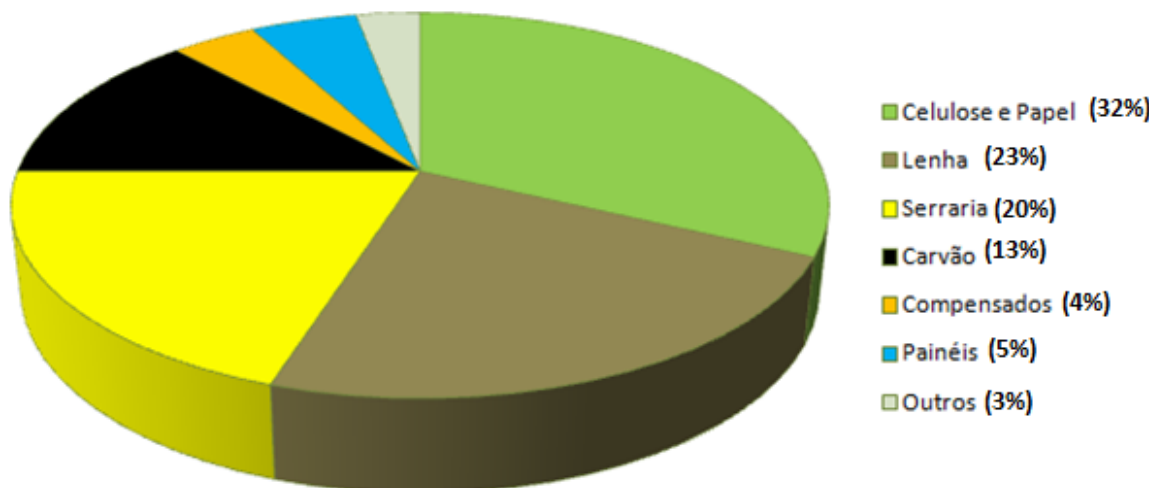


Figura 2: Levantamento estatístico feito pela ABRAF no ano de 2009 mostrando o destino da madeira de eucalipto plantada no Brasil, para as demais utilidades.

Por sua larga empregabilidade econômica e rápido crescimento com idade média de corte de 6 a 8 anos, quando comparada ao ciclo das árvores de mata nativa que levam em média 25 anos, o eucalipto passou a ser uma alternativa racional contra a devastação das florestas nativas em diversas regiões do planeta, propiciando a preservação do meio ambiente (Scarpinella, 2002).

Analisando as espécies deste gênero, *Eucalyptus urophylla* e *E. grandis* representam as espécies de maior interesse para a eucaliptocultura no país. O *E. urophylla* é utilizado para a fabricação de celulose, chapas duras, produção de carvão e outros fins. Esta espécie possui ampla capacidade de adaptação a diversas condições ambientais tropicais, aliada à tolerância ao fungo patogênico *Cryphonectria cubensis*

causador do cancro do eucalipto (Santos et al., 1990). Já o *E. grandis* ocorre em clima quente e úmido nas regiões sudeste e nordeste da Austrália, tendo a sua madeira utilizada principalmente para produção de celulose (Boland et al., 1984, Brooker e Kleining, 1990).

2.2 Estresses Abióticos

O estresse é em geral definido como um fator externo, que exerce uma influência desvantajosa sobre a planta, induzindo-a a mudanças e respostas em todos os níveis funcionais do organismo, sendo reversíveis a princípio, mas podendo, em alguns casos, se tornar permanente (Larcher, 2004).

Em condições naturais e agricultáveis, as plantas estão frequentemente expostas às condições de múltiplos estresses, tais como a seca, a elevada salinidade, a alta temperatura, as inundações e a intensa luminosidade (Chen e Murata, 2002; Rizhsky et al., 2004), as quais limitarão seu desenvolvimento e suas chances de sobrevivência. Estes estresses, por sua vez, desempenham um papel importante na determinação de como o solo e o clima limitam a distribuição de espécies vegetais (Bressan et al., 2004; Verslues et al., 2006).

O estresse pode se manifestar em vários graus de severidade. No que diz respeito às plantas, a intensidade do estresse vai depender do órgão ou do tecido alvo, do estágio do desenvolvimento da planta e do genótipo em questão. Em consideração a estes fatores, o genótipo pode manifestar resistência ou tolerância (sobrevivendo e, às vezes, até crescendo, mesmo que em menores taxas), ou suscetibilidade (sofrendo redução do crescimento e, dependendo da intensidade do estresse, podendo chegar a morte) (Cabraia, 2005).

Dos fatores ambientais que podem ocasionar estresse em uma planta, a falta de água é o mais limitante para a sustentabilidade dos sistemas agrícolas (Paiva et al., 2005; Souza et al., 2006), pois afeta o metabolismo nas plantas, causando, principalmente, grandes perdas em produtividade agrônômica e florestal (Jones e Corllet, 1992; Nogueira et al., 2001; Flexas et al., 2002).

2.3 O fenômeno das secas e importância da água

A água é uma das substâncias essenciais à vida de um vegetal, pois está intrinsecamente relacionada com várias funções como estrutura, crescimento (expansão celular), transporte e metabolismo. É a substância mais abundante em plantas em crescimento ativo, podendo constituir cerca de 90% do peso fresco de muitos órgãos, sendo que seu volume também atua como absorvente de calor e permite, em parte, a tamponização da temperatura interna. Na água, encontram-se dissolvidos o oxigênio e o dióxido de carbono necessários à respiração e à fotossíntese, sendo o meio onde se processam várias reações bioquímicas (Kramer e Boyer, 1995).

O problema da seca tem sua origem no ciclo do carbono no planeta, que vem sendo constantemente acelerado com a queima de petróleo e carvão, além das queimadas de florestas e das emissões de metano pelos grandes rebanhos. A seca é um fenômeno freqüente e característico do nosso país, com intensidade e efeitos variáveis no espaço e no tempo. A situação geográfica do território brasileiro é favorável à ocorrência de episódios de seca, pois este fenômeno não constitui propriamente uma surpresa, devendo antes ser encarado como um elemento climático de determinada freqüência no sentido de que já ocorreu no passado e ocorrerá no futuro (Cushman e Bohnert, 2000).

O estresse por déficit hídrico, diminui o potencial hídrico do solo interferindo diretamente na eficiência do uso da água pelas plantas, as quais necessitam então desenvolver um potencial hídrico mais negativo que o do solo, de modo a manter um gradiente de potencial hídrico entre o solo e a planta (Taiz e Zeiger, 2004). O decréscimo de água no solo diminui o potencial de água na folha e sua condutância estomática, promovendo o fechamento dos estômatos (Yordanov et al., 2003). Esse fechamento bloqueia o influxo de CO₂ para as folhas, diminuindo o acúmulo de fotoassimilados, o que implica na redução da produtividade (Santos e Carlesso, 1998; Ober e Luterbacher, 2002).

Ademais, as plantas sob déficit hídrico apresentam uma diminuição do volume celular, um aumento na concentração de solutos e uma progressiva desidratação do protoplasma, sendo que, a primeira e mais sensível resposta ao déficit hídrico é a diminuição da turgescência e, associada a esse evento, a diminuição do processo de crescimento, particularmente, do crescimento em extensão (Larcher, 2004). Sendo assim, a desidratação causa mudanças fisiológicas e bioquímicas nas plantas, tais como diminuição nas atividades fotoquímicas, redução da fixação de CO₂, acúmulo de osmólitos e osmoprotetores e alteração no metabolismo de carboidratos (Liu e Baird, 2003). Uma perda acima de 30% de água causa um aumento gradual da taxa de respiração, que poderá exceder a taxa de fotossíntese desencadeando um processo de inanição na planta, devido ao decréscimo das reservas de carboidratos (Stefanuto, 2002).

Portanto, não há um processo vital que não seja afetado de alguma forma pelo declínio do potencial hídrico (Taylor, 1996; Larcher, 2004), sendo assim, a identificação e a compreensão dos mecanismos de resistência à seca fundamentais no

desenvolvimento de cultivares comerciais mais resistentes ao estresse hídrico (Nepomuceno et al., 2001).

2.4 Mecanismos de resposta da planta ao déficit hídrico

A cultura do *Eucalyptus* se encontra espalhada por quase todo o território nacional, sendo que grande parte da área de cultivo apresenta limitações ao desenvolvimento das plantas como diversos níveis de estresse hídrico. Desta forma é difícil a escolha de indivíduos adaptados, tornando-se necessário identificar os materiais mais promissores para cada situação (Stape et al., 2004).

Como resposta a deficiência hídrica, as plantas passam por mudanças fundamentais na relação da célula com a água, influenciando a sua capacidade de tolerar as condições adversas do meio. O estresse hídrico por escassez de água ocorre quando a baixa disponibilidade de água afeta os processos fisiológicos (Silva, 1998), como transpiração, fotossíntese, abertura estomática (Wright et al., 2004; Merchant et al., 2007; Tatagiba et al., 2008; Vellini et al., 2008), produção de ácido abscísico (Taylor et al., 2000; Seki et al., 2007) e ajuste osmótico (Bray, 1997; Zhu et al., 1997; Hasegawa et al., 2000 ; Callister et al., 2006; Merchant et al., 2007).

Muitas plantas podem se aclimatar à falta de água através de ajustamento osmótico, onde a resposta comum a diferentes tipos de estresses abióticos é o acúmulo de solutos compatíveis, os quais, além de atuarem no ajustamento osmótico, protegem as estruturas celulares contra os danos induzidos pela desidratação e oxidação que acaba limitando o crescimento vegetal e a produtividade das culturas (Nogueira et al., 2001; Arriel, 2004; Carvalho, 2005; Koryo, 2006). Por essa razão, o nome soluto compatível utilizado algumas vezes para designar estes osmólitos, não é o mais apropriado devido

estes compostos não serem apenas compatíveis, ou seja, não prejudiciais quando em elevada concentração, mas também protetores.

Vários trabalhos têm discutido a osmoproteção em plantas e sua potencial aplicação na tolerância à seca e a salinidade (Rathinasabapathi, 2000; Lv et al., 2007; Ahmad et al., 2008). Tem sido hipotetizado que esses compostos beneficiam células estressadas de duas formas: (1) pela ação com osmólitos citoplasmáticos, desse modo facilitando a absorção e retenção de água (Hare et al., 1998) e (2) proteção e estabilização de estruturas e macromoléculas (proteínas, membranas, cloroplastos e lipossomos) dos danos induzidos por condições adversas (Bohnert e Jensen, 1996).

Os solutos compatíveis que podem ser acumulados diferem dentre as espécies vegetais e podem incluir betaínas e compostos relacionados, açúcares, tais como manitol, sorbitol e trealose e aminoácidos, tais como prolina e hidroxiprolina (Chen e Murata, 2002).

O acúmulo de glicina betaína tem sido frequentemente registrada sob condições de estresse hídrico e salino, tendo sido usualmente considerado ter um efetivo papel fisiológico na osmorregulação de compartimentos citoplasmáticos celulares, na proteção de proteínas e na estabilização de membranas (Sakamoto e Murata, 2000).

Em termos moleculares, o sinal de estresse, ocorrido pela perda do volume e da turgescência celular ou a concentração de solutos altera a conformação de proteínas da parede celular e da membrana plasmática da célula vegetal, ativando rotas de transdução de sinais que dão lugar à expressão de determinados genes, transformando assim o fenômeno físico do déficit hídrico em uma resposta bioquímica (Seki et al., 2003; Yamaguchi-Shinozaki et al., 2005; Willadino e Camara, 2005).

A maior limitação para o melhoramento genético visando à tolerância ao déficit hídrico em plantas é o conhecimento insuficiente sobre as bases fisiológicas, moleculares e genéticas das respostas das plantas ao déficit (Casagrande et al., 2001).

Portanto, compreender como tais eventos interagem entre si representa um importante passo no desenvolvimento de plantas com maior tolerância à seca. Dessa maneira, a biologia molecular assume papel-chave na identificação pontual de genes e de rotas metabólicas envolvidos nas respostas fisiológicas à seca (Zhang et al., 2003; Rathinasabapathi, 2000 ; Casagrande et al., 2001; Nepomuceno et al, 2001).

2.5 Osmoprotetores

A classe de pequenas moléculas conhecidas como osmólitos compatíveis inclui aminoácidos como a prolina, compostos quaternários de amônio (glicina betaína, prolina betaína, B-alanina betaína, e colina-0-sulfato) e o composto terciário de sulfato, 3-dimetilsulfoniopropionato (DMSP). Os compostos quaternários de amônio e DMSP são derivados de precursores de aminoácidos (Yancey, 1994; Sakamoto e Murata, 2000). Estes compostos compartilham a propriedade de permanecerem invariáveis em pH neutro e serem altamente solúveis em água (Ballantyne e Chamberlin, 1994). Além disso, em altas concentrações, têm pequeno ou nenhum efeito sobre a interação de macromoléculas solventes (Low, 1985; Somero, 1986; Timasheff, 1993; Yancey, 1994), diferentemente, dos íons inorgânicos que entram prontamente na esfera de hidratação de proteínas, favorecendo seu desdobramento. Os osmólitos compatíveis tendem a ser excluídos da esfera de hidratação das proteínas e estabilizam a estrutura de proteínas (Low, 1985). Estes compostos têm um papel fundamental no ajustamento osmótico do citoplasma da planta durante a resposta ao estresse osmótico (Wyn Jones et al., 1977).

O acúmulo de osmólitos em células de plantas resulta em um decréscimo no potencial osmótico e também na manutenção da absorção de água e pressão de turgor da célula, o que contribui para a manutenção dos processos fisiológicos, como abertura estomática, fotossíntese e crescimento da planta (Morgan, 1984; Ludlow e Muchow 1990; Blum, 1996).

O acúmulo de osmólitos em plantas tem sido utilizado com ênfase em critérios de seleção nos programas tradicionais de melhoramento de plantas para aumento da produtividade em ambiente com deficiência hídrica (Ludlow e Muchow, 1990; Belhassem et al., 1995; Zhang et al., 2003; Claussen, 2005).

2.6 Glicina betaina (GB)

A glicina betaína é uma amina quaternária, sendo um composto eletricamente neutro sobre uma gama enorme de valores de pH fisiológicos, além de ser extremamente solúvel em água. As características moleculares da GB permitem interagir com domínios hidrófilos e hidrofóbicos de macromoléculas, assim como enzimas e proteínas complexas (Sakamoto e Murata, 2002).

Entre os vários compostos de amina quaternária conhecida em plantas, a GB é a que ocorre com maior abundância, em resposta ao estresse por desidratação (Mansour, 2000; Mohanty et al., 2002; Yang et al., 2003).

Estudos *in vitro* tem indicado que a GB sofre aumento da osmolaridade intracelular quando esta célula é exposta a condições hiperosmótica de estresse-induzido (Le Rudulier et al., 1984; Gorham, 1998; Bray et al., 2001; Meloni et al., 2004), agindo na estabilização tanto da estrutura quaternária das proteínas quanto da complexa estrutura de membrana celular (Le Rudulier et al., 1984).

O principal papel da glicina betaína seria, provavelmente, o de proteção das células vegetais por meio da preservação do balanço osmótico; da estabilização da estrutura de proteínas-chave, tal como a Rubisco; da proteção do aparato fotossintético e de sua atuação como eliminadora de espécies reativas de oxigênio (ROS) (Cha-um et al., 2006).

Este osmoprotetor é acumulado em várias espécies sob condições de estresse, principalmente o estresse salino e a seca. O acúmulo de glicina betaina ocorre principalmente nos cloroplastos, e tem sido registrado em beterraba, espinafre, ervilha, trigo, sorgo, amendoim, cevada, andiroba e cana-de-açúcar sendo mostrado nessas espécies, que os genótipos mais tolerantes acumulam maiores teores de glicina betaína (Weimberg et al., 1984; Fallon e Phillips, 1989; Mccue e Hanson, 1990; Rhodes e Hanson, 1993; Girija et al., 2002; Yang et al., 2003; Silva, 2009; Carlin, 2009).

Por outro lado, a relação entre tolerância a estresse e acúmulo de glicina betaína nem sempre é existente e há espécies que não produzem esse composto, como é o caso do tomate (Mccue e Hanson, 1990), *Arabidopsis*, arroz e do tabaco (Rhodes e Hanson, 1993); porém, plantas transgênicas dessas espécies que expressem os genes responsáveis pela síntese de glicina betaína passam a exibir maior tolerância à seca que as linhagens selvagens não transformadas, tais como: tabaco (Rhodes e Hanson, 1993); *arabidopsis* (Huang et al., 2000) arroz (Takabe et al., 2006;) tomate (Park et al., 2007).

2.7 Colina Monooxigenase (CMO)

Esta enzima é existente somente em plantas (Yamada, et al., 1979; Landfald e Strom, 1986; Rozwadowski et al., 1991), localizando-se no cloroplasto (Brinkmann e Martin, 1995), fazendo parte do primeiro passo da biossíntese da glicina betaina (Burnet et al., 1995; Rathinasabapathi et al., 1997). Em bactérias e animais, o primeiro passo para produção de glicina betaína é dado pelas enzimas colina desidrogenase (CDH) ou pela colina oxidase (COD).

A CMO é uma monooxigenase incomum, pertencente a família das oxigenases (Burnet et al., 1995), dependente de ferroxidina fazendo parte do grupo RIESKE, ao qual faz transferência de elétron coordenados por dois ligantes cis e dois trans (Mason & Cammack, 1992).

O primeiro cDNA da enzima CMO caracterizado, foi de espinafre (*Spinacia oleracea*), contendo no total 1622 pb e uma fase de leitura aberta (ORF) que codifica 440 aminoácidos (Cline e Henry, 1996).

2.8 Biossíntese da glicina betaina

Em sistemas biológicos conhecidos, a GB é sintetizada por duas vias distintas: colina ou glicina. A conversão da colina para GB foi estudada em vários organismos e a rota envolve uma ou duas enzimas, dependendo da forma de oxidação da colina. A segunda enzima da rota é difundida, acontecendo naturalmente em várias plantas, animais e microorganismos (Rhodes e Hanson, 1993).

Em plantas superiores, as reações são catalizadas pela enzima colina monooxigenase (CMO) a qual é dependente de ferroxidina; e a enzima betaína aldeído

desidrogenase (BADH) a qual é dependente de NAD⁺ (Chen e Murata, 2002; Rontein et al., 2002), ambos localizados no estroma do cloroplasto (Rhodes e Hanson, 1993)(Figura 3).

O acúmulo de glicina betaína é associado a uma regulação positiva da expressão do gene para CMO e para BADH concomitantemente, assim conduzindo a uma atividade enzimática elevada (Agboma et al., 1997).

A biossíntese da GB é de estresse induzido e a concentração da GB *in vivo* varia entre as espécies acumuladoras naturais (Rhodes e Hanson, 1993).

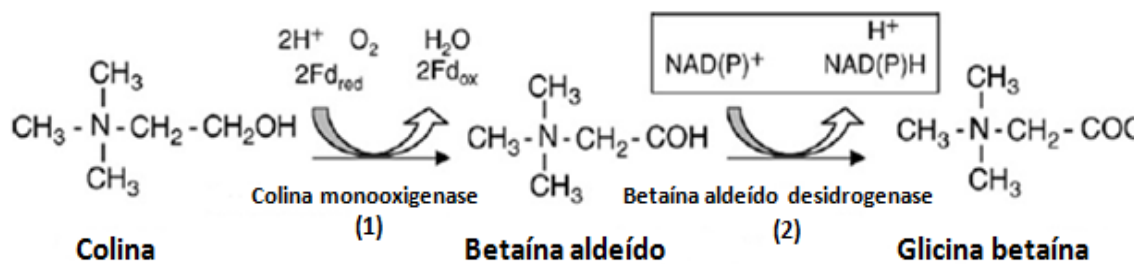


Figura 3: Rota da biossíntese da glicina betaína em plantas superiores. A colina é oxidada através da reação de dois passos: (1) É catalizada pela CMO e a reação (2) é catalizada pela BADH (Rhodes e Hanson, 1993; Ashraf e Foolad, 2007).

2.9 Marcadores Moleculares

Os diferentes tipos de marcadores moleculares disponíveis variam conforme a tecnologia utilizada para revelar variabilidade em nível de DNA, habilidade para detectar diferenças entre indivíduos, custo, facilidade de uso, consistência e repetibilidade (Oliveira et al., 2007).

O surgimento de técnicas de manipulação de DNA levou ao desenvolvimento de marcadores RFLP. Na década de 80, a técnica de PCR permitiu a amplificação de

fragmentos de DNA, e levou ao aparecimento de marcadores como RAPD, microssatélites e AFLP. Estes dois últimos marcadores em sido amplamente utilizados em análises de diversidade genética de plantas.

Apesar da disponibilidade de diferentes marcadores, grande parte da informação obtida é oriunda de polimorfismos em regiões do genoma detectadas por marcadores aleatórios. Com a popularização das técnicas de seqüenciamento foi possível o desenvolvimento de marcadores funcionais e a análise de polimorfismos de seqüências.

Esses marcadores, baseados em genes candidatos ou genes-alvo, aumentam a probabilidade de se estabelecer correlações entre fenótipo e genótipo, pois a variabilidade observada reflete diferenças fenotípicas (Kwok e Gu, 1999).

Já foram desenvolvidos vários marcadores moleculares de interesse em diversas espécies vegetais (Junghans, 2000; Moretzsohn et al., 2000; Martins Filho et al., 2002; Benko-Iseppon et al., 2003). Em *Eucalyptus grandis* foi identificado com sucesso gene de resistência a *Puccinia psidii* Winter (Junghans et al., 2003). Estes marcadores podem ser utilizados em programas de melhoramento possibilitando a seleção precoce de indivíduos portadores do gene de interesse.

Seqüenciamento de DNA e marcadores moleculares começaram a convergir quando os dados acumulados passaram a mostrar variações entre seqüências de indivíduos de uma mesma espécie. Logo se percebeu que Single-Nucleotide Polymorphisms (SNP) eram os marcadores mais abundantes nos genomas (Kwok & Gu, 1999; Ganal et al., 2009).

Esses marcadores podem ser utilizados como uma nova estratégia de identificação de polimorfismo em programas de melhoramento de plantas (Nepomuceno et al., 2001).

2.10 SNP - *Single Nucleotide Polymorphism*

Single Nucleotide Polymorphism, ou seja, polimorfismo de base individual ou única (SNP) é o acrônimo amplamente utilizado hoje para definir a variação de seqüência observada em uma posição individual da seqüência de DNA (Figura 4). O SNP ocorre quando um único nucleotídeo – A, T, C, ou G – no genoma difere entre membros de uma espécie com freqüência de pelo menos 1%. Por exemplo, duas seqüências de um fragmento de DNA de diferentes indivíduos contém uma diferença em um único nucleotídeo, GGGG(C)CG e GGGG(T)CG. Neste caso diz-se que existem dois alelos: C ou T. (Brookes, 1999). Assim, na prática, SNPs são marcadores bi-alelicos, podendo ocorrer tri-alelicos em uma proporção menor, de forma que o conteúdo informativo em um único SNP é limitado, em comparação com os marcadores microssatélites (SSR) que são polialélicos (Griffin e Smith, 2000; Gupta et al., 2001; Schlotterer et al., 2004; Oraguzie et al., 2007).

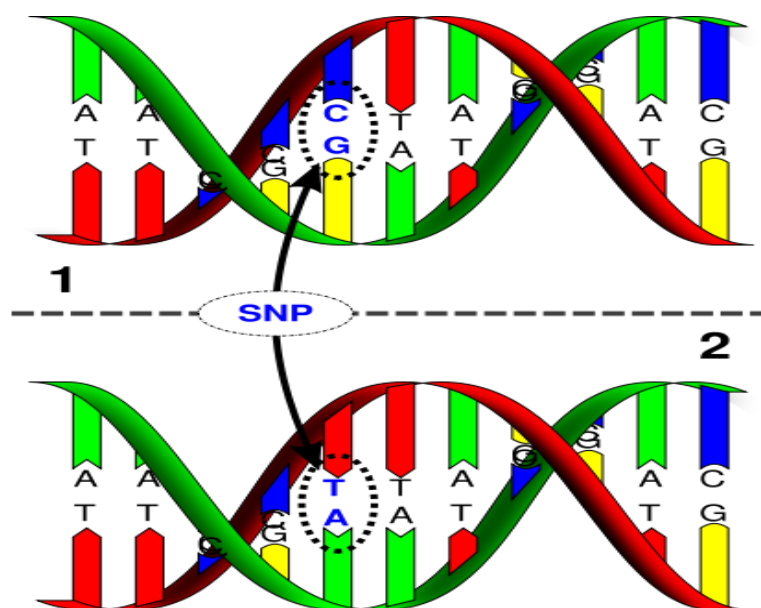


Figura 4: Ilustração mostrando, em destaque, o polimorfismo (CG→TA) de único nucleotídeo, denominado (SNP).

A existência de SNPs no genoma não é novidade. Os primeiros trabalhos de seqüenciamento de fragmentos específicos de DNA detectaram esse tipo de polimorfismo (Orita et al., 1989).

Este marcador mostra várias vantagens sobre os outros tipos de marcadores tal como a sua alta frequência, facilidade de automatização e estabilidade do loco SNP comparado com o loco do SSR, por exemplo.

Os SNPs são o tipo mais freqüente de variação encontrada no DNA (Brookes, 1999), sendo distribuídos por todo o genoma como na região de íntrons, éxons, regiões intergênicas, promotores ou *enhancers*. Entretanto, a localização dos SNPs pode ser de relevância funcional e fisiológica para o organismo. Um SNP localizado na região codificadora por exemplo, pode ter impactos profundos na formação da proteína. Um SNP intrônico pode influenciar na região de *splicing* do mRNA (Krawezak et al., 1992), assim como, um SNP no promotor pode influenciar na expressão gênica (Drazen et al., 1999).

Há também um grande potencial para o uso dos SNPs na detecção de associações entre forma alélicas de um gene e fenótipos, especialmente doenças ou características comuns que possuem genética multifatorial (Rafalski, 2002).

A descoberta de SNPs que afetam a atividade biológica tem se tornado cada vez mais importante e vem sendo facilitada pela disponibilidade de um vasto mapeamento do genoma. Os métodos de detecção de sítios únicos são, potencialmente, mais fáceis de automatizar e de aplicar na análise genética em larga escala. Deste modo, a informação gerada sobre os SNPs pode ser, posteriormente, utilizada em vários níveis nas análises do DNA como: marcadores no mapeamento genético; material de estudo na identificação de SNPs funcionais e seu respectivo fenótipo; ferramenta para estudos

populacionais, por comparação do genoma de indivíduos de diferentes populações (Zhang e Hewitt, 2003).

A frequência habitual de SNPs relatados para genomas de plantas é de cerca de 1 SNP a cada 100-300 pb (Gupta et al., 2001).

Nas plantas, os SNPs passaram a ser mais utilizados a partir da crescente disponibilidade de coleções de ESTs nos bancos de dados públicos, o que reduziu os custos para descoberta dos SNPs. Essas coleções tem sido fonte de identificação de SNPs em algumas plantas como beterraba (Schneider et al., 2001), milho (Ching et al., 2002), arroz (Nasu et al., 2002), soja (Zhu et al., 2003), e de cana de açúcar (Grivet et al., 2003). A disponibilidade de bancos de dados de EST torna possível identificar os polimorfismos para regiões funcionais do genoma e até mesmo de genes específicos (Kota et al., 2001).

2.11 Diversidade nucleotídica e haplótipos de SNP

A variabilidade genética, sendo ela de ocorrência neutra ou seletiva, pode gerar variações na composição dos organismos. A fonte desta variabilidade é uma alteração súbita e herdável no material genético, caracterizada como mutação. Quando estas são localizadas e alteram apenas poucas bases, são chamadas de mutações de ponto, ou mutações gênicas. Por outro lado, se as mesmas alterarem muitas bases ou até cromossomos inteiros, são chamadas de mutações ou aberrações cromossômicas.

Segundo Nei e Li (1979), uma das estimativas levando em conta a diversidade nucleotídica (π) é denominada:

$$\pi = \sum_{ij} x_i x_j d_{ij}$$

onde “q” é o número total de alelos, “ x_i ” é a frequência na população (em seqüências) do alelo “i”, “ x_j ” é a frequência na população (em seqüências) do alelo “j” e “ d_{ij} ” é o número de diferenças nucleotídicas ou substituições por sítio entre os alelos “i” e “j”. Resumindo, é o somatório das diferenças de pares de nucleotídeos entre seqüências de uma amostra e depende tanto das frequências quanto do número de sítios polimórficos.

Para um melhor entendimento dos diferentes parâmetros de diversidade estimados, consideremos um conjunto de 10 seqüências de 11 bases, conforme está indicado abaixo:

```

AGCTTAATTAG
AGCTTAATTG
AGCTTAATTAG
AGTTAATTAG
AGCTTAATTAG
AGCTTAATTAG
CGCTTAATTAG
CGCTTAATTAG
CGCTTAATTAG
AGCGTATTAG

```

Observando estas seqüências temos que:

n = número de nucleotídeos , $n = 11$

S = número total de sítios segregantes = 6

p_s = número total de sítios segregantes por nucleotídeo = $6/11 = 0,545$

Considerando apenas as primeiras quatro sequencias:

AGCTTAATTAG
 AGCTTAATTG
 AGCTTAATTAG
 AGTTAATTAG

A diversidade nucleotídica π é igual a:

$$\pi = (1 / a) \times \Sigma \text{diferenças} / (a \times n) =$$

$$\pi = \frac{1}{4} \times (12 / (4 \times 11)) = 0,06818$$

onde

n = número de nucleotídeos , n = 11

a = tamanho da amostra (número de seqüências) = 4

O conceito de haplótipo está intimamente relacionado à diversidade nucleotídica, e tem como significado, o padrão de associação de SNPs que são herdados conjuntamente, isto é, segregam conjuntamente e caracterizam um grupo, espécie ou população.

Vários trabalhos em plantas utilizando a diversidade nucleotídica e haplótipos já estão descritos, em milho (Ching et al., 2002), tomate (Simko et al., 2004), *Pinus* (Dvornyk et al., 2002; Gonzalez-Martinez et al., 2004), cevada (Russel et al., 2004), soja (Cregan et al., 2002; Zhu et al., 2003) e eucalipto (Novaes et al., 2008; Külheim et al., 2009).

3 Objetivos

Os objetivos deste trabalho foram:

Geral:

- Identificar SNPs no gene colina monooxigenase relacionado ao metabolismo da glicina betaína em *Eucalyptus*.

Específico:

- Definir o número de haplótipos e a distribuição destes;
- Disponibilizar “primers” de SNP para genotipagem em estudos populacionais e programas de melhoramento genético assistido.

4 Material e Métodos

4.1 Material

O material vegetal utilizado neste trabalho foi dividido em 2 grupos:

- Material Vegetal utilizado na procura de SNPs:

Para os testes de amplificação dos “primers” e procura dos SNPs, o material vegetal utilizado no estudo foi composto por 12 indivíduos híbridos, denominados “Urograndis”, formados pelo cruzamento de duas espécies de eucalipto (*Eucalyptus grandis* x *Eucalyptus urophylla*), os quais pertencem ao programa de melhoramento florestal da empresa Fábria Celulose. Este material consiste em indivíduos divergentes geneticamente e contrastantes quanto à seca. Todos os ensaios biológicos de estresses abióticos, foram realizados pela equipe do Prof. Dr. Rinaldo César de Paula do Departamento de Produção Vegetal, da Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal na Universidade Estadual Paulista Júlio de Mesquita Filho. Para a seleção deste material vegetal, os parâmetros utilizadas foram: características fisiológicas (fotossíntese, condutância estomática, transpiração, potencial hídrico foliar e conteúdo relativo de água); e de crescimento (área foliar, número de folhas, massa seca). As plantas eram submetidas a duas condições de disponibilidade de água: Testemunha (sem limitação de água) e estressadas (suspensão da irrigação até o aparecimento de sintomas de estresse hídrico: murcha).

- Material Vegetal utilizado para a genotipagem dos SNPs:

Para os testes de análise da diversidade nucleotídica e haplótipos de SNP, foram utilizados, no total, 60 indivíduos do gênero *Eucalyptus* formados pelas espécies *Eucalyptus grandis*, *Eucalyptus urophylla*, e o híbrido “Urograndis” fornecidos pela empresa Fíbria Celulose.

4.2 Métodos

4.2.1 Extração de DNA

Para a extração do DNA foi utilizado o protocolo descrito por Ferreira e Grattapaglia (1998) com adaptação na etapa de CTAB 10%, que foi substituída pela utilização de CTAB 5%. O CTAB 5% apresenta uma manipulação mais fácil, tendo menor viscosidade, sua utilização em trabalhos anteriores não mostrou alteração na qualidade do DNA final (Domingues, 2003; Bortoloto, 2003).

Em um tubo foram macerados, aproximadamente, 150mg de folhas com nitrogênio líquido, em seguida, foram adicionados 700µl de solução de extração 29 com polivinilpinolidona (PVP) e 2-mercapto-etanol (2µl de mercapto-etanol para cada ml de solução de extração), aplicado em capela de exaustão e misturado até que todo o material macerado entrasse em contato com a solução. Esta mistura foi incubada por 45 a 60 minutos à temperatura de 65°C.

Em seguida, foi realizada a primeira extração, adicionando-se 600µl de clorofórmio-álcool isoamílico (CIA 24:1). Esta solução foi misturada com o auxílio de vortex e os tubos levados à centrífuga refrigerada a 10°C, numa rotação de 12000 rotações por minuto (rpm), durante 5 minutos. Após a centrifugação, a solução apresentou três fases, a superior foi transferida para um novo tubo (±500µl) e a esta

solução foram adicionados 100µl de CTAB 5%. Foi realizada então a segunda extração adicionando-se 600µl de CIA 24:1. Essa solução foi homogeneizada em vortex e centrifugada nas mesmas condições anteriores.

Após a centrifugação a solução apresentou duas fases separadas por uma película de proteínas. O sobrenadante (\pm 500µl) foi transferido para um novo tubo onde foi adicionado 400µl de NaCl 5M e centrifugado nas mesmas condições anteriores. O sobrenadante (\pm 500µl) foi novamente transferido para um novo tubo onde foram adicionados 400µl de álcool isopropílico gelado. O material foi mantido à temperatura de -20°C por um período de aproximadamente 10 horas. As amostras foram novamente centrifugadas nas mesmas condições anteriores.

Após a centrifugação pode-se visualizar um pellet no fundo dos tubos, o sobrenadante foi descartado e o pellet lavado com 500µl de etanol 95% por 2 minutos, o etanol foi descartado e o pellet após secagem em centrífuga a vácuo foi suspenso em tampão TE contendo 10ng/µl de RNase (volume variou de 100 a 50µl, dependendo do tamanho do “pellet”). As amostras foram incubadas a 37°C para a digestão do RNA por no mínimo 40 minutos ou até completa dissolução do “pellet”.

As amostras de DNA que após a digestão do seu RNA pela RNase, que ainda apresentaram resíduo de RNA, verificado após quantificação, foram tratadas com 1 µl de RNase a 10ng/µl.

As amostras, que não apresentaram completa dissolução do pellet após 2 horas incubadas a 37°C, foram deixadas over-night à temperatura ambiente e/ou foram levadas ao aquecedor a 65°C por 30 minutos.

4.2.2 Quantificação e diluição do DNA

A quantificação foi realizada em espectrofotômetro Nano Drop[®] - ND1000. Este equipamento, que além de avaliar o DNA com precisão, utiliza volumes pequenos (1 µl) sem necessidade de diluição, e fornece parâmetros para avaliar sua pureza, quanto a proteínas e compostos fenólicos. O DNA utilizado nas reações de amplificação padrão foram diluídos à concentração de 15 ng/µl em água ultra pura autoclavada.

4.2.3 Procura das enzimas relacionadas ao metabolismo da glicina betaína

As enzimas relacionadas a biossíntese da glicina betaína foram obtidas em “*Plantcyc*” (Plant metabolic pathway database) (<http://www.plantcyc.org/>) e “*Kegg*” (Kyoto Encyclopedia of Genes and Genomes) (<http://www.genome.jp/kegg/>). Foi utilizado o E. C. number como palavra chave para procura das sequências no GenBank (<http://www.ncbi.nlm.nih.gov>). Utilizando as seqüências de aminoácidos de referência foi feita uma prévia procura por sequências similares no banco de dados de EST de eucalipto FORESTs – FAPESP (Vicentini et al., 2005) utilizando tBlastx (Altschul et al., 1997), e posteriormente, foram comparadas no banco de dados público de eucalipto (eucalyptusDB) (<http://eucalyptusdb.bi.up.ac.za/>).

4.2.4 Seleção dos iniciadores específicos e condições de amplificação

Os pares de ‘primers’ foram desenhados com a utilização do programa PerlPrimer, versão 1.1.14, para amplificarem produtos de 400-1400 pb (pares de base).

A ressuspensão destes ‘primers’ foi feita com tampão Tris-EDTA (T.E.) a uma concentração final de 100 μM , sendo os mesmos diluídos a uma concentração final de 10 μM para serem utilizados nas reações de PCR.

A melhor temperatura de pareamento para amplificação de cada ‘primer’ foi determinada via gradiente de temperatura utilizando-se um termociclador PTC-100 (MJ Research, Inc., Watertown, MA, USA).

As reações foram realizadas em cadeia da polimerase (PCR) em 32 ciclos de 1 minuto a 95 °C, 1 minuto a X °C e 1 minutos a 72 °C, onde X (temperatura de pareamento de cada primer). Cada reação conteve um volume final de 10 μl , onde foram utilizados 0,3 μM de cada ‘primer’ forward e reverse, 15 ng de DNA genômico e PCR Master Mix, da Promega Corporation[®], o qual contém 50 unidades/ml de Taq DNA polimerase, 400 μM dATP, 400 μM dGTP, 400 μM dCTP, 400 μM dTTP, 3mM MgCl₂, estocados em tampão para reação com pH 8.5.

A não amplificação ou amplificação de bandas fracas indicaram a necessidade de reduzir a temperatura de anelamento do ‘primer’. Por outro lado, quando houve a presença de várias bandas, indicou-se a amplificação de produtos inespecíficos, tornando-se necessário o aumento da temperatura de anelamento para aumentar a especificidade da reação. Todos estes procedimentos, de aumento ou diminuição da temperatura, foram utilizados de acordo com as necessidades de cada ‘primer’.

Os fragmentos resultantes foram visualizados em gel de agarose a 1%, corado com brometo de etídeo e comparado com um padrão de peso molecular conhecido

(Ladder – Fermentas[®]) de 100 pb e 1Kb para confirmação da especificidade da amplificação.

4.2.5 Purificação dos produtos de PCR

Os fragmentos de PCR de cada indivíduo foram purificados com tratamento da enzima ExoSAP IT (GE Healthcare). Este reagente de purificação, é formado por duas enzimas hidrolíticas – Exonuclease I e a Fosfatase Alcalina de Camarão (SAP).

A enzima exonuclease I tem como função remover os resíduos de “primers” e alguns DNAs de fita simples “estranhos” produzidos durante a reação de PCR com terminal 3'. Já a função da enzima SAP, é o de remover os dNTPs não incorporados na reação de PCR.

Foi adicionado 1 µl da enzima ExoSAP IT à 5 µl do produto de PCR e, posteriormente, incubar a solução a 37°C durante 15 minutos (ocorrer a degradação do excesso de “primers” e nucleotídeos) e a 80 °C por 15 minutos (inativação da enzima).

4.2.6 Reação de sequenciamento

As seqüências de DNA foram determinadas em seqüenciador automático ABI 3100 (Applied Biosystems[®]) utilizando-se 4 µl de 2,5x Save Money (400mM Tris-HCl pH 9,0, 10mM MgCl₂), 4µl BigDyeTM Terminator Cycle Sequencing Ready Reaction Kit versão 3,1 (Applied Biosystems, Foster City, CA), 3,2 pmol dos oligonucleotídeos forward e reverse, 4 µL do produto de PCR purificado e H₂O q.s.p.. As reações foram realizadas em termociclador Eppendorf MasterCycler Gradient (Eppendorf, Hamburg,

Germany) com os ciclos de temperatura programados para: 25 ciclos de 95°C por 10 seg, 50°C por 5 seg, 60°C por 4 seg, com rampa de 1°C/seg, como recomendado pelo fabricante. Após a amplificação as amostras serão mantidas a 4°C até a precipitação.

Para cada amostra serão realizadas duas reações, sendo uma para o ‘‘primer’’ forward e outra para o reverse.

4.2.7 Precipitação das reações de sequenciamento

Nas reações de sequenciamento foram adicionados 80µl de isopropanol 65%, incubando a temperatura ambiente (T.A.) por 20 min. Em seguida, as amostras foram centrifugadas a velocidade de 10.000xg por 25 min a T.A.. O isopropanol foi removido invertendo os tubos e, em seguida, foram adicionados 200µl de etanol 60% e centrifugado a velocidade de 10.000xg por 5 min a T.A.. Todo o etanol foi removido com o auxílio de uma micropipeta.

As amostras foram secas em T.A. e o produto ressuspenso em 2µl de tampão de amostra contendo Formamida Hi-DiTM (Applied Biosystems) + Loading Buffer (25mM EDTA pH 8,0 contendo 50 mg/ml de Blue Dextran) (5:1). No momento da aplicação em sequenciador automático ABI PRISM® 3100 (Applied Biosystems, USA) as amostras foram aquecidas a 95 °C por 5 min e rapidamente transferidas para o gelo.

4.2.8 Alinhamento das Sequências

Os eletroferogramas das seqüências “forward” e “reverse”, gerados automaticamente, foram inspecionados visualmente através do programa ChromasPro versão 1.42 (Technelysium Pty Ltd), no qual consensos de seqüências podem ser formados para um mesmo gene. Após obtenção das seqüências, foi realizado o BLAST, para comparações com os “reads” originais e com outras seqüências disponíveis no GenBank.

4.2.9 Análises dos SNPs

O alinhamento das seqüências foi realizado para possibilitar a identificação visual dos SNPs como transições ou transversões em regiões de íntrons e éxons. SNPs localizados dentro de regiões codificantes foram classificados como sinônimos e não-sinônimos e a posição de cada um deles no códon também foi determinada.

4.2.10 Caracterização dos SNPs por um método semi-automático

O método de genotipagem deve fornecer estimativas acuradas da frequência de alelos, além de ter tempo e custo baixos para sua realização. Atualmente existem diferentes métodos disponíveis para genotipagem por SNPs em amostras individuais (Syvanen, 2001). Um dos métodos de escolha para este trabalho se baseia na extensão alelo-específica ou minisequenciamento de um oligonucleotídeo adjacente ao sítio SNP utilizando o kit SNaPshot (Applied Biosystems). A utilização do kit SNaPshot™ permite a detecção de mutações, pelo método dideoxi marcado com fluorescência e

técnica de extensão de uma única base que tem se mostrado altamente eficiente (Le Hellard et al., 2002).

4.2.11 Seleção dos oligonucleotídeos flanqueadores

No presente estudo, cada oligonucleotídeo flanqueador foi construído adjacente aos sítios mutacionais, com 18 pb (Figura 5) de acordo com as recomendações da técnica de minisequenciamento (Applied Biosystems). Para facilitar a discriminação dos SNPs de um mesmo gene, alguns “primers” foram construídos com diferentes tamanhos, pela inserção de cauda poli C na extremidade 5’. A melhor temperatura de reação para cada oligonucleotídeo foi determinada utilizando o termociclador Eppendorf MastewrCycle Gradient (Eppendorf, Hamburg, Germany).

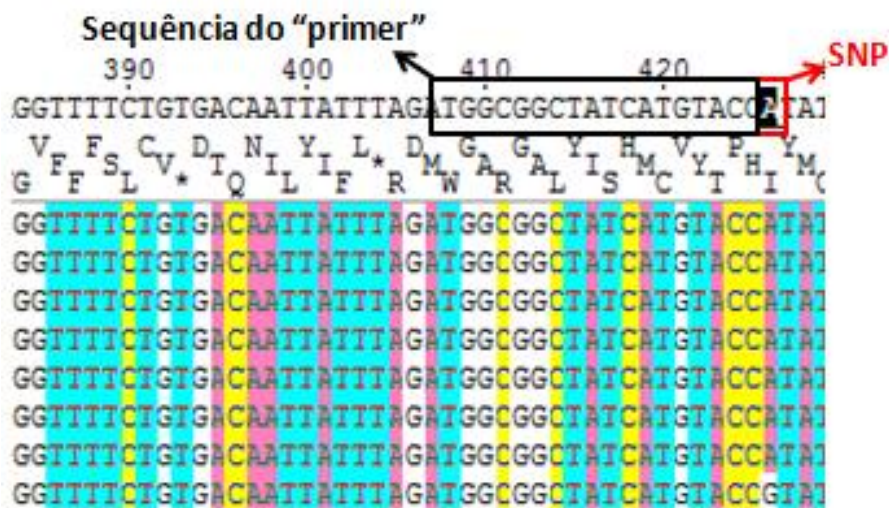


Figura 5: Ilustração retirada do software ChromasPro versão 1.42 mostrando a metodologia aplicada para o desenho dos “primers” utilizados nas reações de minisequenciamento para genotipagem dos SNPs. A sequência complementar, da figura destacada em preto, foi utilizada como “primer”.

4.2.12 Reação de minisequenciamento

Produtos de PCR foram purificados por reação enzimática com ExoSAP IT (GE Healthcare) de acordo com as recomendações do fabricante. Reações de minisequenciamento foram montadas com volume final de 10 µl, contendo 2 µl do produto de PCR purificado, 1 µl de 2,5x save money (400 mM Tris-HCl pH 9,0, 10 mM MgCl₂), 1 µl de Snapshot™ Multiplex mix, 0,4 mol de cada oligonucleotídeo e água ultra pura. As reações foram realizadas em termociclador Eppendorf MasterCycler Gradient (Eppendorf, Hamburg, Germany) programado para: 40 ciclos de 95°C por 10 seg, 58°C por 5 seg, 60°C por 30 seg, com rampa de 1°C/seg.

4.2.13 Purificação das reações de minisequenciamento

Reações de minisequenciamento foram purificadas com 1 µl de SAP (GE Helthcare) conforme as recomendações do fabricante. A seguir, as reações foram diluídas 1:1 com Formamida Hi-Di (Applied Biosystems) + Loading buffer (25 Mm EDTA pH 8,0 contendo 50 mg/ml de Blue Dextran). No momento da aplicação em sequenciador automático ABI PRISM 3100 (Applied Biosystems) as amostras foram aquecidas a 95°C por 3 min e rapidamente transferidas para o gelo.

4.2.14 Análises da Estrutura Genética Populacional

Apos a corrida eletroforética no sequenciador, os SNPs foram identificados com base no tamanho do produto amplificado e na fluorescência. Planilhas com dados genotípicos diplóides foram obtidas por contagem direta e montadas no formato Excel.

4.2.15 Análise da diversidade

Utilizando o software DnaSP 5.10 (www.ub.es/dnasp) (Rozas et al., 2010) foram obtidas as seguintes estimativas:

- a) Número total de sítios segregantes (S);
- b) Diversidade nucleotídica π (Nei e Li, 1979; Nei, 1987), o número médio de diferenças nucleotídicas (SNPs, polimorfismos de bases individuais); e haplótipos de SNP.

5 Resultados e discussão

5.1 Enzimas encontradas e relacionadas ao metabolismo da glicina betaína

Como resultado, após a aplicação da metodologia para a procura das enzimas, foram selecionadas 2 enzimas relacionadas a biossíntese da glicina betaína (colina monooxigenase - CMO e betaína aldeído desidrogenase - BADH) (Figura 6). Segundo Rhodes e Hanson (1993), somente as enzimas mencionadas acima participam diretamente da biossíntese da glicina betaína em plantas. Trabalhos recentes de outros autores continuam mencionando as duas enzimas como as únicas participantes desta rota (Bhuiyan et al., 2007; Zhang et al., 2008; Missihoun et al., 2010).

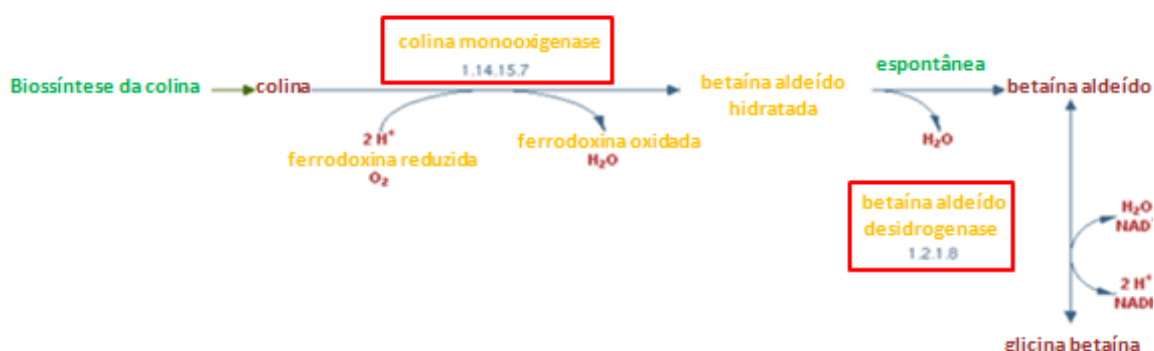


Figura 6: Ilustração retirada e editada do site de rotas metabólicas de plantas *Plantcyc*, mostrando a atuação das 2 enzimas (colina monooxigenase e betaína aldeído desidrogenase) na formação do osmoprotetor glicina betaína.

Somente a colina monooxigenase foi utilizada neste trabalho. Bhuiyan et al (2007) consideraram-na a enzima-chave na biossíntese da glicina betaína, pois através de ensaios laboratoriais, de expressão e supressão da CMO, com a planta *Amaranthus tricolor*, pode-se constatar a ocorrência ou não na produção de glicina betaína.

Já o gene que codifica a enzima betaína aldeído desidrogenase não foi utilizado neste trabalho, pois apesar de sua importância na biossíntese deste soluto, possui no mesmo genoma, outras regiões idênticas ou muito similares em várias espécies de plantas, como já foi demonstrado em *Arabidopsis thaliana*, onde 14 sequências foram consideradas homólogas (Kirch e Bartels, 2004), e em *Oryza sativa*, com um total de 20 sequências homólogas (Gao e Han, 2009). Várias cópias de genes no genoma poderia dificultar o andamento deste trabalho, pois os oligos específicos construídos para amplificar o gene candidato teriam grande probabilidade de se parear em regiões inespecíficas no genoma de eucalipto.

Após a escolha da enzima CMO para dar continuidade a este trabalho, foi efetuada a busca da sequência de aminoácidos da mesma, no banco de dados de sequências genéticas, GenBank, e a espécie de planta escolhida para a obtenção destes dados foi a *Arabidopsis thaliana* (Tabela 1), sendo esta, utilizada com frequência como modelo em pesquisas laboratoriais para estudo de genes de interesse, pois além de rápida reprodução e desenvolvimento, possui seus genes conservados em relação a outras espécies de plantas (Lewin, 2009). Novaes et al. (2008) realizaram um seqüenciamento em larga escala do genoma de *Eucalyptus grandis*, onde foram seqüenciados 148 mega pares de base (Mbp), sendo estes, possuindo uma homologia maior com a espécie de planta *Arabidopsis thaliana*, seguido de *Populus trichocarpa* e *Oryza sativa*.

A espécie *A. thaliana* não produz naturalmente o soluto glicina betaína (Rhodes e Hanson, 1993), mas possui o gene CMO não funcional em seu genoma (Hayashi et al., 1997; Hibino et al., 2002). Hayashi et al. (1997) chegaram a esta conclusão quando inseriram o gene colina oxidase (COD) da bactéria *Arthrobacter globiformis* na planta *Arabidopsis thaliana*, passando assim, a espécie modelo, a produzir o osmoprotetor

glicina betaína, além de exibir maior tolerância ao estresse hídrico e salino. Segundo Ikuta et al (1977), a enzima colina oxidase, em bactérias, participa do primeiro passo na biossíntese da glicina betaína, o mesmo ocorrendo com a enzima colina monooxigenase, neste caso em plantas (Burnet et al., 1995; Rathinasabapathi et al., 1997).

Tabela 1 - Enzima selecionada através do banco de dados do GenBank. São apresentados o EC- Number da enzima, nome da enzima, espécie obtida, número de identificação no GenBank, acesso no Tair e referência atual sobre a enzima.

E.C NUMBER	ENZIMA	ESPÉCIE	GenBank	TAIR	Referência sobre a enzima
1.14.15.7	Colina Monooxigenase (CMO)	<i>Arabidopsis thaliana</i>	Q9SZR0	AT4g29890	Missihoun et al., 2010)

5.2 Busca de homólogos para a enzima colina monooxigenase em banco de dados de eucalipto

As buscas foram realizadas no banco de dados de ESTs de eucalipto FORESTs-FAPESP, onde a sequência modelo de aminoácidos da enzima candidata colina monooxigenase de *Arabidopsis thaliana* apresentou maior identidade a um consenso de 1338 pares de base (pb) formado por dois “reads”. Estes “reads” foram utilizados como moldes para o desenho dos oligonucleotídeos específicos “primers”.

5.3 Desenho dos oligonucleotídeos específicos

Após a constatação da sequência homóloga no banco de dados de ESTs de eucalipto FORESTs-FAPESP, oito pares de “primers” foram desenhados a partir dos dois “reads” obtidos neste banco. Além disso, o banco de dados de sequências

genômicas (EucalyptusDB), o qual se tornou disponível a partir de 2010, também auxiliou no fornecimento de informações importantes para o desenho de dois primers, GB 2-5 e GB 2-8. Graças as informações desses dois bancos de dados, pode-se obter “primers” que pudessem amplificar todo o gene (Figura 7).

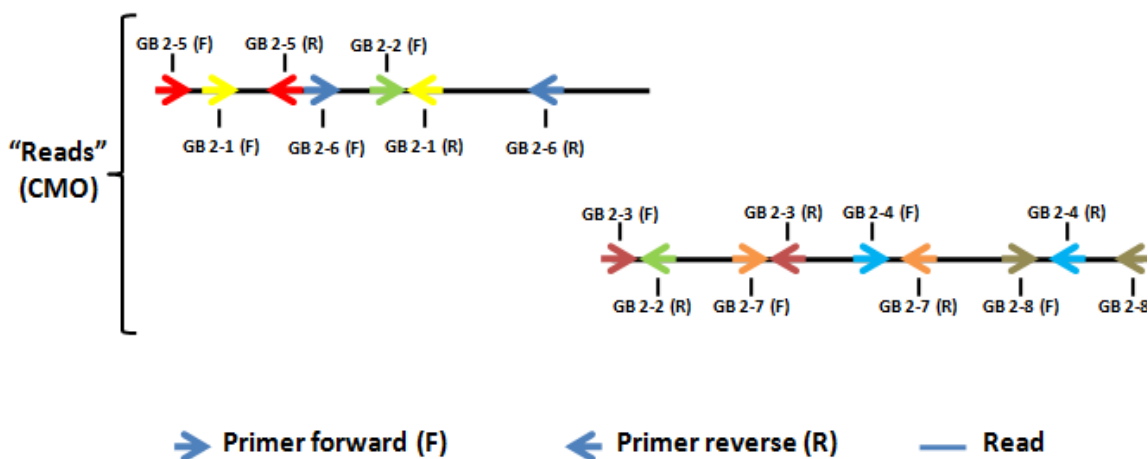


Figura 7 – Ilustração esquemática dos dois “reads” pertencentes ao gene colina monooxigenase (CMO) obtidos através do resultado da pesquisa no banco de dados de ESTs de eucalipto FORESTs – FAPESP, mostrando a localização do desenho de cada “primer” a partir de seus respectivos “reads”.

5.4 Amplificação dos oligonucleotídeos específicos

Dos oito pares de “primers” construídos para amplificar o gene CMO, todos (100%) apresentaram amplificação de banda única (Figura 8), variando a temperatura de pareamento entre 55 e 59 °C, e os fragmentos variando em tamanhos de 500 a 1400 pares de base (pb) (Tabela 2). Nenhum fragmento apresentou tamanho amplificado esperado, sendo este resultado, discutido com maiores detalhes mais adiante.

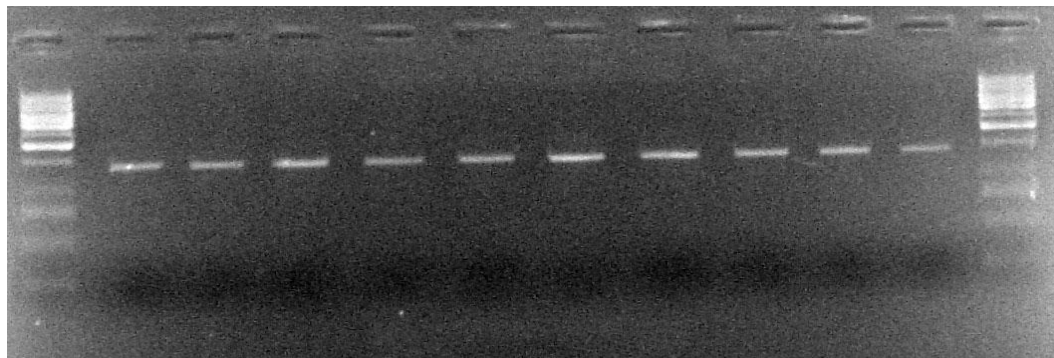


Figura 8 - Teste de amplificação do “primer” GB 2-2 em gel de agarose 1%, corado com brometo de etídio. Nas extremidades estão os marcadores de peso molecular conhecidos (Ladder 1Kb – Fermentas®). Para este teste foram utilizados, aleatoriamente, dez indivíduos divergentes, os quais, apresentaram amplificação de banda única.

Tabela 2: “Primers”, forward e reverse, utilizados para amplificar o gene colina monooxigenase, em eucalipto. São apresentados a identidade dos “primers”; temperatura de pareamento, em graus Celsius ($^{\circ}\text{C}$); tamanho do fragmento amplificado, em pares de base (pb); e as sequencias de cada “primer”.

Primer	Temperatura pareamento	Fragmento amplificado	Sequencia forward/reverse
GB 2-1	59 $^{\circ}\text{C}$	500 pb	5'-CGCTCGACTCCATCATCAC-3' 3'-GGCGATTTAACCTGATCCGT-5'
GB 2-2	55 $^{\circ}\text{C}$	1400 pb	5'-CGCCTCAGGACTTCTTCAC-3' 3'-CCATTCCTTTGCTGCTATGTC-5'
GB 2-3	57 $^{\circ}\text{C}$	800 pb	5'-AGATGTTTTCTCCCGATCAG-3' 3'-CCACCTTCACATATTTGTAGAC-5'
GB 2-4	55 $^{\circ}\text{C}$	1300 pb	5'-GAGTCTACAAATATGTGAAGGTGG-3' 3'-CAATAGGCTGGAGATTCAAGG-5'
GB 2-5	55 $^{\circ}\text{C}$	700 pb	5'-CACCATTTGACTTGTGGAAGA-3' 3'-GATCGAAGTCGTCCACCAG-5'
GB 2-6	55 $^{\circ}\text{C}$	600 pb	5'-GGACTTACGCACTGAATGG-3' 3'-TTACACTATACAGACCAGAACC-5'
GB 2-7	55 $^{\circ}\text{C}$	500 pb	5'-CTTGACTCTTACTCCACCAC-3' 3'-ATAAAGCAGCAAACCTCTCAG-5'
GB 2-8	55 $^{\circ}\text{C}$	400 pb	5'-GCAGGATAGTGAAAGAGTACAG-3' 3'-ACACGATAGATTCAATAAGCCC-5'

5.5 Sequenciamento do gene colina monooxigenase e análise *in silico* no GenBank

Inicialmente, para confirmação dos fragmentos amplificados e posterior localização dos SNPs, efetuou-se um primeiro seqüenciamento utilizando 12 indivíduos divergentes geneticamente, indicados pela empresa Fíbria Celulose. A escolha destes indivíduos não foi por acaso, uma vez que a divergência entre eles aumentam as chances de polimorfismos, conseqüentemente, o aparecimento de SNPs.

Os 12 indivíduos foram amplificados pelos oito pares de “primers”, a fim de que todo o gene fosse seqüenciado. As seqüências resultantes apresentaram ótima qualidade, como mostrado no cromatograma (Figura 9), sendo posteriormente, alinhadas no software ChromasPro versão 1.42. Em seguida, foi realizado um BlastX (Altschul, 1997), no GenBank, com os fragmentos seqüenciados, mostrando, que os mesmos, correspondem as enzimas em estudo (Tabela 3).

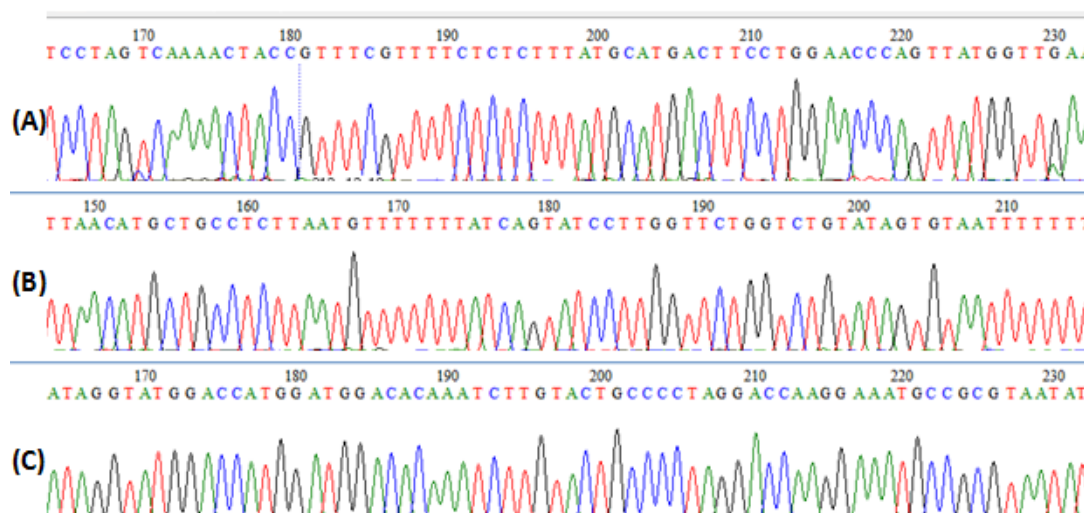


Figura 9: Cromatograma parcial, mostrando a ótima qualidade das seqüências amplificadas pelos “primers” (A) GB 2-2, (B) GB 2-3, (C) GB 2-4. Picos altos e bem espaçados são apresentados na figura.

Com o resultado do Blast, também foi possível observar que os éxons de eucalipto são similares em relação ao das outras plantas como *Arabidopsis thaliana*, *Vitis vinifera*, *Populus trichocarpa*, entre outras. Porém, os íntrons se diferenciam entre os organismos. A explicação mais plausível seria o fato deles evoluírem mais rapidamente que os éxons, envolvendo alterações no seu tamanho (deleções e inserções). As mutações ocorrem com a mesma frequência nas regiões de íntrons e éxons, mas são mais facilmente removidas dos éxons devido a pressão seletiva (Lewin, 2009).

Tabela 3: Resultado do BlastX com a sequência alinhada, obtida na amplificação dos oito pares de “primers” construídos para cobrir todo o gene colina monooxigenase. São apresentados: Número de acesso no Genbank, nome da enzima e organismo possuidor, score e máxima identidade.

Número de Acesso	Nome da enzima/organismo	Score	Máxima identidade
ABS71853.1	Colina monooxigenase/ <i>Eucalyptus camaldulensis</i>	466	100%
T08550	Colina monooxigenase/ <i>Arabidopsis thaliana</i>	511	75%
XP_002264997.1	Proteína hipotética/ <i>Vitis vinifera</i>	637	96%
XP_002308100.1	Proteína hipotética/ <i>Populus trichocarpa</i>	618	96%
ABX57826.1	Colina monooxigenase/ <i>Chrysanthemum lavandulifolium</i>	356	93%
XP_002518256.1	Colina monooxigenase/ <i>Ricinus communis</i>	595	90%
NP_001174977.1	Colina monooxigenase/ <i>Oryza sativa Japonica</i>	515	80%

5.6 Análise *in silico* das sequências correspondentes ao gene colina monooxigenase nos bancos de dados de eucalipto

As análises *in silico* realizadas no banco de dados de ESTs de eucalipto FORESTs-FAPESP e no banco de dados de sequências genômicas (EucalyptusDB), mostraram que não houve a ocorrência de cópias homólogas no genoma, sendo possível seguir a diante com os objetivos deste trabalho na procura de SNPs, no gene CMO.

Utilizando o programa ClustalX (Thompson et al., 1997), realizou-se um alinhamento das sequências obtidas no sequenciamento, com a sequência expressa obtida no banco de dados de ESTs de eucalipto FORESTs-FAPESP, localizando-se assim, a posição dos íntrons e éxons deste gene.

Chegou-se a conclusão que, o gene CMO de eucalipto, possui dez éxons e nove íntrons distribuídos em 3885 pb. Os mesmos números de éxons e íntrons são encontrados em *A. thaliana* (acesso no TAIR – AT4g29890), mostrando desta forma, que o número de regiões é conservado entre essas plantas, diferindo no tamanho, 2166 pb, pois os íntrons em eucalipto são maiores em pares de base (pb) (Figura 8). Até o momento, nada foi demonstrado e registrado na literatura mostrando alguma relação entre o tamanho do genoma com a complexidade do organismo (Lewin, 2009).

A localização da região 5' UTR e 3' UTR do gene no eucalipto, foi determinada através do alinhamento utilizando a ferramenta Blastx, no GenBank, comparando a sequência de aminoácidos de CMO de eucalipto com a sequência de aminoácidos de CMO de outros organismos (Figura 10 e Tabela 4).

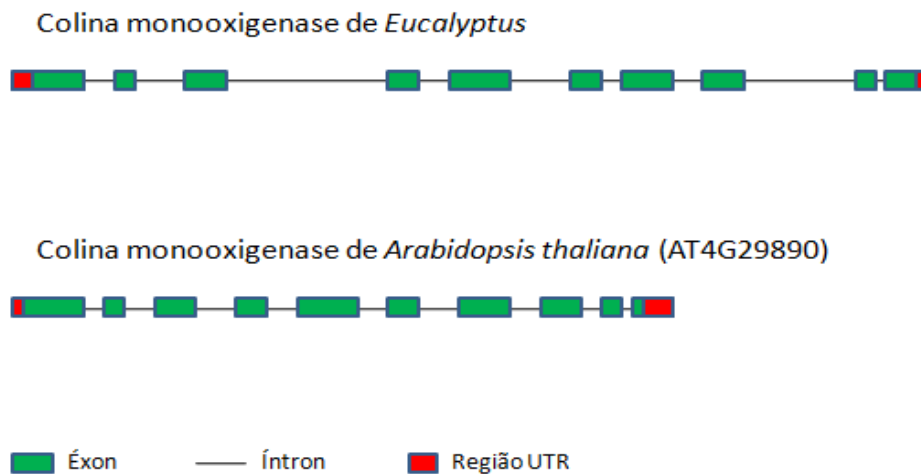


Figura 10 - Ilustração esquemática comparando a distribuição de íntrons e éxons de CMO de eucalipto e arabidopsis. O alinhamento do seqüenciamento com os “reads” de CMO de ESTs do banco de dados de ESTs de eucalipto FORESTs - FAPESP resultou no esquema do gene em eucalipto. O gene de CMO em arabidopsis foi montado a partir da seqüência obtida no banco de dados de *Arabidopsis* (TAIR).

Tabela 4 – Determinação das regiões de éxon e íntron da sequência gênica de CMO de *arabidopsis* e *eucalypto*. Através de comparações, observa-se que a região exônica das duas espécies analisadas possuem tamanhos parecidos, diferindo-se apenas no tamanho dos íntrons.

Região	CMO de <i>eucalyptus</i> (posição)	CMO de <i>arabidopsis</i> (posição)
5' UTR	1 a 60 (60pb)	1 a 35 (35pb)
Éxon (1)	61 a 374 (313pb)	36 a 334 (298pb)
Íntron (1)	375 a 558 (183pb)	335 a 409 (74pb)
Éxon (2)	559 a 610 (51pb)	410 a 458 (48pb)
Íntron (2)	611 a 841 (230pb)	459 a 532 (73pb)
Éxon (3)	842 a 982 (140pb)	533 a 671 (138pb)
Íntron (3)	983 a 1728 (745pb)	672 a 763 (91pb)
Éxon (4)	1729 a 1804 (75pb)	764 a 839 (75pb)
Íntron (4)	1805 a 1892 (87pb)	840 a 916 (76pb)
Éxon (5)	1893 a 2109 (216pb)	917 a 1135 (218pb)
Íntron (5)	2110 a 2362 (252pb)	1136 a 1214 (78pb)
Éxon (6)	2363 a 2460 (97pb)	1215 a 1316 (101pb)
Íntron (6)	2461 a 2639 (178pb)	1317 a 1422 (105pb)
Éxon (7)	2640 a 2764 (124pb)	1423 a 1548 (125pb)
Íntron (7)	2765 a 2842 (77pb)	1549 a 1611 (62pb)
Éxon (8)	2843 a 2939 (96pb)	1612 a 1720 (108pb)
Íntron (8)	2940 a 3623 (683pb)	1721 a 1802 (81pb)
Éxon (9)	3624 a 3676 (52pb)	1803 a 1855 (52pb)
Íntron (9)	3677 a 3754 (77pb)	1856 a 1934 (78pb)
Éxon (10)	3755 a 3872 (117pb)	1935 a 2076 (141pb)
3' UTR	3873 a 3885 (12pb)	2077 a 2166 (89pb)

5.7 Identificação dos SNPs no gene colina monoxygenase (CMO) em eucalipto

Para a procura dos SNPs, as sequências dos 12 indivíduos divergentes geneticamente foram alinhadas pelo programa ChromasPro versão 1.42, sendo considerados, na análise, tanto os SNPs que possuem alelos diferentes em homozigose quanto os que possuem alelos em heterozigose. A procura foi realizada através dos resultados emitidos pelo cromatograma, comparando-se os picos em cada sequência, além da comparação das sequências nucleotídicas, como mostrado na Figura 11.

Os SNPs foram identificados nos fragmentos amplificados pelos “primers” GB 2-2, GB 2-3 e GB 2-4.

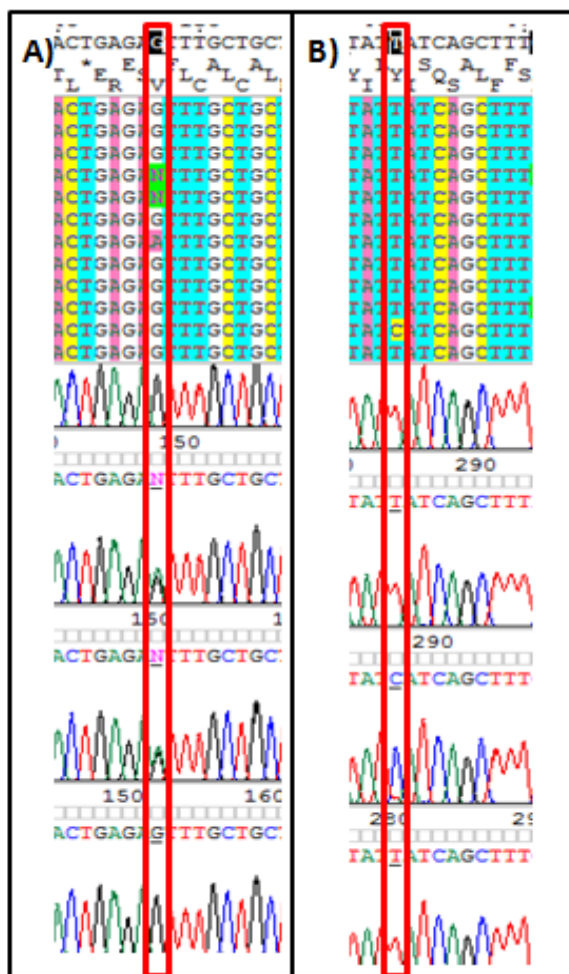


Figura 11: Alinhamento realizado pelo programa ChromasPro versão 1.42, comparando indivíduos divergentes na procura de SNPs. A região destacada em vermelho mostra a localização dos SNPs nos indivíduos que possuem alelos em homozigose e heterozigose (A), e indivíduos que possuem alelos diferentes em homozigose (B).

Esta metodologia, que consiste em seqüenciamento direto de indivíduos, subsequente alinhamento utilizando ferramentas de bioinformática, para posterior comparação das seqüências na procura de SNPs, foi utilizada com sucesso em diversos tipos de estudos, como em humanos (Lai et al., 1998; Levy et al., 2009), bovinos (Calvo et al., 2006; Abatepaulo et al., 2008), plantas (Van et al., 2004; Novaes et al., 2008; Külheim et al., 2009; Ganal et al., 2009), além de outros organismos (Liao e Lee, 2010).

Segundo Brookes et al (1999), o alelo polimórfico deverá ocorrer em, pelo menos, 1% na população testada. Neste trabalho, foram identificados 49 SNPs em uma seqüência de 3885 pb do gene CMO, em eucalipto, sendo a média de SNPs de 1 para cada 79,2 pb. Estudos também relacionados ao gênero *Eucalyptus* com outros autores, revelam uma freqüência relativamente semelhante dependendo do gene em estudo. Thumma et al (2005) mostraram que a freqüência observada no gene CCR (cinnamoyl coA redutase), em *Eucalyptus nitens* é de 1 SNP para cada 94 pares de base. Neves et al (2008) também trabalharam com um gene, o F5H, em *E. globulus*, *E. grandis* e *E. urophylla*, onde foi apresentado uma freqüência, respectivamente de 1 SNP para cada 155, 77,5 e 52 pb. Em um outro estudo, considerado mais abrangente, utilizando 4 espécies de *Eucalyptus* com 23 genes seqüenciados, revelou que no total, para as espécies *E. globulus*, *E. nitens*, *E.camaldulenses* e *E.loxophleba*, o número médio de SNPs é, respectivamente, 38.3, 35.5, 79.3 e 73.0 para cada 1000 pb (Külheim et al., 2009). No geral, para genoma de plantas, estes polimorfismos de base única variam entre 1 SNP para cada 100 a 300 pb (Gupta et al., 2001).

Dos 49 SNPs encontrados neste trabalho, para o gene CMO, 12 estão localizados em regiões codificantes (éxons) e 37 em região não codificantes (íntrons e regiões UTRs). Estas diferenças entre SNPs, em regiões codificantes e não codificantes, também ocorre no trabalho de Külheim et al (2009), onde as 4 espécies de eucalipto,

principalmente o *E. camaldulensis* e *E. loxophleba*, possuem uma proporção significativa de SNPs nas regiões onde não ocorre a tradução. Portanto, esta explicação pode ser dada por diversos autores, onde afirmam que o maior número de SNPs ocorre em regiões não codificantes, pois lá, existe pouca ou nenhuma pressão seletiva (Brookes, 1999; Rafalski, 2002; Ching et al., 2002; Gonzalez-Martinez et al., 2006).

Segundo Hammer (1995) e Gonzales-Martinez (2006), SNPs localizados em regiões codificantes tem particular importância por ser mais provável que estas variações nucleotídicas tenham algum tipo de significado funcional, em nível fenotípico, nos organismos. Por este motivo, somente os 12 SNPs identificados nas regiões codificantes participaram das próximas etapas deste trabalho.

5.8 Classificação dos SNPs localizados nas regiões codificantes

Quanto a classificação de mutações de substituição por transição e transversão, dez SNPs bi-alélicos, foram classificados como substituições por transição. Essas substituições, ocorrem quando uma pirimidina é substituída por outra pirimidina, ou, uma purina é substituída por outra purina. Neste caso, seis SNPs (60,0%) foram substituídos pelas bases (A→G) e quatro SNPs (40,0%) foram substituídos pelas bases (T→C). Somente dois sítios de SNP (L26 e L38) apresentaram-se tri-alélicos, isto é, obtiveram três tipos diferentes de base (A→G→C) e (A→C→T), respectivamente. Portanto, possuem os dois tipos diferentes de substituição (transição e transversão). No geral, essas mutações por substituição, seguiram a normalidade, pois as substituições por transição, costumam ocorrer com maior frequência no genoma de diversos organismos (Lewin, 2009).

Quanto a classificação por mutações sinônimas e não sinônimas dos 12 SNPs, inicialmente efetuou-se a predição da proteína gerada pelo gene CMO, onde foram traduzidas as sequências exônicas no site Expasy (Gasteiger et al., 2003), o que gerou uma sequência de 445 aminoácidos, além disso, quando analisada no PFAM (Finn et al., 2008), apresentou identidade ao domínio do tipo Rieske (2Fe – 2S). Mason e Cammack (1992) foram os primeiros autores a demonstrar este domínio da proteína utilizando CMO da planta *Spinacea oleracea*.

A tradução da região codificante do gene CMO, gerada no site Expasy, foi utilizada para análise do posicionamento de cada SNP referente ao respectivo códon. As análises mostraram que dez das mutações ocorreram no terceiro nucleotídeo de cada códon, e duas, ocorreram no segundo nucleotídeo de seus respectivos códons. Deste modo, dez SNPs (83,3%) referentes a esta pesquisa foram classificados como sinônimos, isto é, a mudança ocorrida no códon não gerou mudança no aminoácido, e dois SNPs (16.7%) foram classificados como sendo não-sinonimos, portanto, esta alteração nucleotídica acabou gerando mudança de aminoácido (Tabela 5).

Tabela 5 – SNPs localizados em regiões codificantes e selecionados para desenho dos “primers” para SnapShot™ (Applied Biosystems). Houve alteração de aminoácidos em dois dos doze sítios de SNP. São mostrados na tabela a identidade do SNP, os dois ou 3 possíveis códons (1),(2) e (3) respectivos aminoácidos e o tipo de mutação.

SNP	Códon 1	Códon 2 ou 3	Tipo de Mutação
L24	CC(A) - Prolina	CC(G) - Prolina	Sinônima
L25	GC(G) - Alanina	GC(A) - Alanina	Sinônima
L26	GC(G) - Alanina	GC(C) - Alanina GC(A) - Alanina	Sinônima Sinônima
L31	TC(G) - Serina	TC(A) - Serina	Sinônima
L32	CT(G) - Leucina	CT(A) - Leucina	Sinônima
L33	TT(C) - Fenilalanina	TT(T) - Fenilalanina	Sinônima
L34	CC(A) - Prolina	CC(G) - Prolina	Sinônima
L35	TT(C) - Fenilalanina	TT(T) - Fenilalanina	Sinônima
L36	GA(C) - Aspartato	GA (T) - Aspartato	Sinônima
L37	TT(C) - Fenilalanina	TT(T) - Fenilalanina	Sinônima
L38	C(T)A - Leucina	C(C)A - Prolina C(A)A - Glutamato	Não-sinônima Não-sinônima
L39	C(G)T - Arginina	C(A)T - Histidina	Não-sinônima

Apesar da maioria das mutações serem sinônimas, estudos recentes mostram que este tipo de mutação, algumas vezes, participa de alterações em expressões de genes e dobramentos de proteínas (Chamary et al., 2006; Gupta e Lee, 2008). Nos últimos anos, estudos com códons sinônimos demonstraram resultados diferentes na expressão de genes, mostrando, desta forma, que SNPs sinônimos também podem produzir diferentes fenótipos (Kudla et al., 2009).

5.9 Genotipagem dos SNPs

Para a genotipagem dos SNPs, foi utilizado uma população mais abrangente, constituída, no total, por 59 indivíduos de *E. grandis*, *E. urophylla* e o híbrido “Urograndis”. A região amplificada para esta genotipagem abrange 2100 pb do gene CMO, entre íntrons e éxons, amplificados pelos “primers” GB 2-3 e GB 2-4. Os sítios polimórficos detectados em regiões de íntron foram desconsideradas para esta etapa do trabalho. A região de éxon amplificada possui oito regiões de SNP. A Tabela 6 mostra a identificação de cada nucleotídeo em seus respectivos sítios polimórficos. Para esta etapa, a identificação foi visual.

Tabela 6: Genotipagem dos SNPs em 59 indivíduos pertencentes ao gênero *Eucalyptus*. Indivíduos com iniciais (CL, Ug = Urograndis), (G = *E. grandis*) e (U = *E. urophylla*). Oito sítios polimórficos foram identificados somente em regiões exônicas amplificadas pelos “primers” GB 2-3 e GB 2-4.

Indivíduos	Sítios polimórficos							
	L24	L25	L26	L38	L39	L34	L35	L36
CL1	A	G	G	T	G	G	T	A
CL2	A	G	G	T	G	G	T	A
CL3	G	G	A	T	A	G	T	A/G
CL4	A	A	G	T	G	G	T	A/G
CL5	A	G	G	T	G	G	T	A
CL6	A	A	G	T	G	G	T	A/G
CL7	A	G	G	T	G	G	T	A
CL8	A	A	C	A	G	G	T	A
CL9	A	G	G	T	G	G	T	A
CL10	A	G	C	T	G	G	T	A/G
CL13	A	G	G	T	G	G	T	A
CL14	A	G	G	T	G	G	T	A
Ug1028	A	G	G	A	G	G	T	A/G
Ug1163	A	G	G	T	G	G	T	A/G
Ug1288	A	G	G	T	G	G	T	A
Ug1313	A	G	G	T	G	G	T	A/G
Ug1315	A	G	G	T	G	G	T	A
Ug1340	A	G	G	T	G	G	T	A/G
Ug1399	A	G	G	T	G	A	T	A

Continua

Ug1406	A	G	G	T	G	G	T	A/G
Ug1417	A	G	G	T	G	G	T	A
Ug1510	A	G	G	T	G	G	T	A/G
Ug1623	A	G	G	T	G	G	T	A
Ug1229	A	A	C	T	G	G	T	A
Ug1348	A	G	G	T	G	G	T	A
Ug1372	A	G	G	T	G	G	T	A
Ug1401	A	G	G	T	G	G	T	A
Ug1462	A	G	G	T	G	G	T	A
Ug1395	A	G	G	T	G	G	C	A
Ug3748	A	G	G	T	G	G	T	A
Ug3709	A	G	G	T	G	G	T	A
Ug30	A	G	G	T	G	G/A	T	A
Ug219	A	G	G	T	G	G/A	T	A
Ug-V1	G	G	A	T	A	G	T	A
Ug-V2	A	G	G	T	G	A	T	A
(G)53	A	G	G	T	G	G	T	A/G
(G)55	A	G	G	T	G	G	T	A/G
(G)56	A	A	C	T	G	G	T	A
(G)57	A	G	G	T	G	G	T	A
(G)61	A	G	G	T	G	G	T	A/G
(G)62	A	G	G	T	G	G	T	A
(G)85	A	G	G	T	G	G	T	A
(G)86	A	A	C	T	G	G	T	A
(G)87	A	A	C	T	G	G	T	A/G
(G)107	A	G	G	T	G	G	T	A
(G)71	A	G	G	T	G/A	G	T	A
(G)81	A	G	G	T	G	G	T	A/G
(U)46	A	A	C	T	G	G	T	A
(U)59	A	G	G	T	G	G	T	A/G
(U)60	A	G	G	C	G	G	T	A/G
(U)63	A	G	G	T	G	G	T	A/G
(U)68	A	G	G	C	G	G	T	A/G
(U)75	A	G	G	C	G	G	T	A/G
(U)77	A	G	G	T	G	G	T	A
(U)83	A	G	G	C	G	G	T	A
(U)84	A	G	G	T	G	G	T	A
(U)85	A	G	G	C	G	G	T	A
(U)87	A	G	G	C	G	G	T	A
(U)88	A	G	G	C	G	G	T	A

5.10 Análises estatísticas da genotipagem

Nas análises realizadas nos 59 indivíduos genotipados, constatou-se a formação de 18 haplótipos (Tabela 7), os quais são formados por uma sequência de oito sítios polimórficos cada (Tabela 8). A frequência destes haplótipos na população em estudo variou bastante (Tabela 7 e 8), tendo o H1 ocorrido em mais de 50% das amostras, abrangendo os três grupos de população em estudo, o mesmo ocorrendo com os haplótipos H11 e H13, mas com frequências de 9,32% e 7,63, respectivamente. Outros 15 haplótipos foram específicos para determinados grupos, ou seja, 11 haplótipos comuns para “Urograndis”, dois para *E. grandis* e dois para *E. urophylla* (Tabela8).

Dentre os 18 haplótipos, alguns contiveram baixa frequência nas amostras (H7, H8, H9, H10, H15 e H16, ocorrendo em apenas 1 indivíduo.

Tabela 7: Formação dos possíveis haplótipos gerados pelos indivíduos utilizados neste estudo. São apresentadas as identidades dos haplótipos, o número de ocorrências (levando-se em conta o par de cromossomos) e porcentagem ocorrida nas amostras.

Haplótipo	Ocorrência	(%) na amostra
H1	60	50,85
H2	3	2,54
H3	1	0,85
H4	2	1,69
H5	2	1,69
H6	2	1,69
H7	1	0,85
H8	1	0,85
H9	1	0,85
H10	1	0,85
H11	11	9,32
H12	6	5,08
H13	9	7,63
H14	2	1,69
H15	1	0,85
H16	1	0,85
H17	11	9,32
H18	3	2,34

Tabela 8: Reconstrução haplotípica com as respectivas frequências para as espécies de eucalipto em estudo. São apresentados os 18 possíveis haplótipos estimados através da genotipagem do indivíduos. Somente os haplótipos 1, 11 e 13 foram comum para os três grupos.

(H)	Sítios polimórficos								Espécie
	L24	L25	L26	L38	L39	L34	L35	L36	
H1	A	G	G	T	G	G	T	A	<i>E. grandis</i> , <i>E. urophylla</i> e "Urograndis"
H2	G	G	A	T	A	G	T	A	"Urograndis"
H3	G	G	A	T	A	G	T	G	"Urograndis"
H4	A	A	G	T	G	G	T	A	"Urograndis"
H5	A	A	G	T	G	G	T	G	"Urograndis"
H6	A	A	C	A	G	G	T	A	"Urograndis"
H7	A	G	C	T	G	G	T	A	"Urograndis"
H8	A	G	C	T	G	G	T	G	"Urograndis"
H9	A	G	G	A	G	G	T	A	"Urograndis"
H10	A	G	G	A	G	G	T	G	"Urograndis"
H11	A	G	G	T	G	G	T	G	<i>E. grandis</i> , <i>E. urophylla</i> e "Urograndis"
H12	A	G	G	T	G	A	T	A	"Urograndis"
H13	A	A	C	T	G	G	T	A	<i>E. grandis</i> , <i>E. urophylla</i> e "Urograndis"
H14	A	G	G	T	G	G	C	A	"Urograndis"
H15	A	A	C	T	G	G	T	G	<i>E. grandis</i>
H16	A	G	G	T	A	G	T	A	<i>E. grandis</i>
H17	A	G	G	C	G	G	T	A	<i>E. urophylla</i>
H18	A	G	G	C	G	G	T	G	<i>E. urophylla</i>

Contudo, os 18 haplótipos formados, geraram 16 combinações diferentes, as quais, foram denominadas genótipo, como mostrado na tabela 9.

Tabela 9: Genótipos formados a partir da combinação dos possíveis haplótipos. São apresentados as combinações haplotípicas, número de indivíduos onde ocorreram as combinação, porcentagem dos genótipos nas amostras e porcentagem de ocorrência por grupo.

Genótipo	Número indivíduos	(%) na amostra	(%) por grupo
H1/H1	23	38,98 %	<i>E. grandis</i> (17,4%) <i>E. urophylla</i> (8,7%) “Urograndis” (73,90%)
H1/H11	11	18,64 %	<i>E. grandis</i> (36,36%) <i>E. urophylla</i> (18,19%) “Urograndis” (45,45%)
H1/H12	2	3,39 %	“Urograndis” (100,0%)
H1/H16	1	1,69 %	<i>E. grandis</i> (100,0%)
H12/H12	2	3,39 %	“Urograndis” (100,0%)
H13/H13	4	6,78 %	<i>E. grandis</i> (50,0%) <i>E. urophylla</i> (25,0%) “Urograndis” (25,0%)
H13/H15	1	1,69 %	<i>E. grandis</i> (100,0%)
H14/H14	1	1,69 %	“Urograndis” (100,0%)
H17/H17	4	6,78	<i>E. urophylla</i> (100,0%)
H17/H18	3	5,08	<i>E. urophylla</i> (100,0%)
H2/H2	1	1,69	“Urograndis” (100,0%)
H2/H3	1	1,69	“Urograndis” (100,0%)
H4/H5	2	3,39	“Urograndis” (100,0%)
H6/H6	1	1,69	“Urograndis” (100,0%)
H7/H8	1	1,69	“Urograndis” (100,0%)
H9/H10	1	1,69	“Urograndis” (100,0%)

Após as análises das frequências dos genótipos identificados na população, pode-se constatar o aparecimento em maior número (38,98%) da combinação haplotípica H1/H1, nas amostras, seguido do genótipo H1/H11, aparecendo em 18,64%. Alguns

genótipos foram raros nesta população em estudo, aparecendo somente em 1,69% das amostras (H13/H15, H14/H14, H2/H2, H2/H3, H6/H6, H7/H8 e H9/H10) em “Urograndis”.

Uma observação importante nas análises, foi o aparecimento de dois genótipos (H17/H17 e H17/H18) em praticamente 12% das amostras, os quais, foram exclusivos para alguns indivíduos *E. urophylla*, e o genótipo H13/H15, exclusivo para *E. grandis*.

Nas análises de diversidade nucleotídica, “Urograndis” foi o que apresentou o maior nível de diversidade, $\pi = 0,15694$, seguido das espécies *E. grandis* e *E. urophylla*, que apresentaram, respectivamente, valores de $\pi = 0,15127$ e $\pi = 0,1262$.

Relacionando a maior diversidade nucleotídica, número de haplótipos e frequência haplotípica, entre os outros grupos desse estudo, pode-se constatar que “Urograndis” obteve um resultado que já era esperado, pois se trata de um híbrido formado a partir de duas espécies (*E. grandis* x *E. urophylla*), podendo desta forma, ocorrer maior variabilidade genética, e conseqüentemente um aumento na diversidade nucleotídica e número de haplótipos.

Já na espécie *E. urophylla*, a diversidade nucleotídica e frequência haplotípica foram as mais baixas quando relacionadas aos outros dois grupos em questão. Mckinnon et al (2004) sugerem que a melhor explicação para a diversidade nucleotídica ser menor em *E. urophylla*, é o fato desta espécie, possuir um baixo fluxo gênico em relação a outras espécies estudadas, por causa de seu isolamento geográfico nas ilhas do arquipélago indonésio. O oposto já ocorre com *E. grandis*, a qual possui ampla distribuição geográfica possibilitando um maior fluxo gênico.

É importante observar, entretanto, que as estimativas de diversidade nucleotídica podem variar em ordens de magnitude, dependendo de diversos fatores, sendo os

principais o sistema reprodutivo, efeitos de amostragem, seja em termos de indivíduos e populações, bem como os genes que foram analisados. A exemplo disso, outros estudos relacionados a eucalipto, demonstram resultados diferentes de π . Santos (2005) estudando um segmento do gene *ccoamt* pode observar que os níveis de diversidade nucleotídica para as espécies *E. grandis*, *E. globulus* e *E. urophylla*, foram respectivamente, 0,00356, 0,00168 e 0,00254. Já em um outro estudo mais recente, Neves et al. (2008), trabalhando com os genes CCR e F5H, observaram que os valores médios para π nas espécies *E. grandis*, *E. globulus* e *E. urophylla* foram respectivamente 0,01604; 0,00391 e 0,00305.

Em outras espécies relacionadas a silvicultura, Pot et al (2006) avaliando a diversidade nucleotídica em oito genes candidatos envolvidos na formação da madeira, encontraram valores médios de 0,00241 para *Pinus pinaster* e 0,00186 para *Pinus radiata*, estando este gênero mais domesticado que o *Eucalyptus*.

Em linhas gerais, devido à polinização cruzada, elevada heterozigosidade, baixo nível de domesticação e incipiente seleção artificial, espécies florestais tendem a apresentar elevada diversidade nucleotídica quando comparada com outras espécies de plantas, como o milho, que já se apresenta altamente domesticado, possuindo uma diversidade nucleotídica baixa (Savolainen e Pyhajarvi, 2007).

5.11 Desenho e teste de extensão dos oligonucleotídeos específicos para a genotipagem dos SNPs

Baseado em cada um dos pontos de mutação nas regiões codificantes do gene CMO, foram desenhados adjacentes aos sítios de SNP, doze “primers” com o objetivo de diminuir custos e permitir a genotipagem em larga escala destas mutações de única base em populações de *Eucalyptus* (Tabela 10).

Tabela 10: “Primers” que serão utilizados para reação de extensão alelo-específica dos polimorfismos de único nucleotídeo. São mostrados a identidade dos “primers”, as respectivas sequencias e as possíveis mutações. Para facilitar na discriminação dos “primers”, foi adicionada uma cauda poli A ou C de diferentes tamanhos com o intuito de utilizar estes oligos em reação multiplex. Alguns destes marcadores já foram utilizados para este trabalho.

Sequencia dos “primers”		
Primer	(5' - 3')	Mutação
L24	ACCGCCGATAGTACATGG	A ou G
L25	(AAAA)AGACGCTAGGCCCTTGTG	A ou G
L26	(AAAAAAAA)GTCAAGCTTTAGACCAGA	A, G ou C
L31	ACGAAGAGAAGAACGTAG	A ou G
L32	(CCCC)TCTGAGAAGGGCTCCATT	A ou G
L33	(CCCCCCCC)AGACGTCCCCACCTGAAT	T ou C
L34	(CCCCCCCCCCCC)TGAAAAGAGAATGACTCT	A ou G
L35	CCATGGTCCATACCTATA	T ou C
L36	(CCCC)GGGCAGTACAAGATTTGT	T ou C
L37	(CCCCCCCC)TGTCCTACTGTTCTTAAA	T ou C
L38	TGACGTAAGTAGGTAGTG	T ou C
L39	(CCCC)ATTGAATAGAAACATCTG	G ou A

Até o momento, três “primers” para SNP, identificados por L24, L25 e L26, foram testados para genotipar cinco indivíduos. Estes indivíduos genotipados, foram utilizados propositalmente, pois fazem parte do grupo de divergentes seqüenciados anteriormente, tendo suas mutações de ponto já identificadas visualmente.

Esta técnica mostrou-se eficiente em dois dos três “primers” testados, sendo observado que todos os genótipos obtidos utilizando esta técnica foram idênticos aos resultados obtidos anteriormente pelo seqüenciamento (Figura 12).

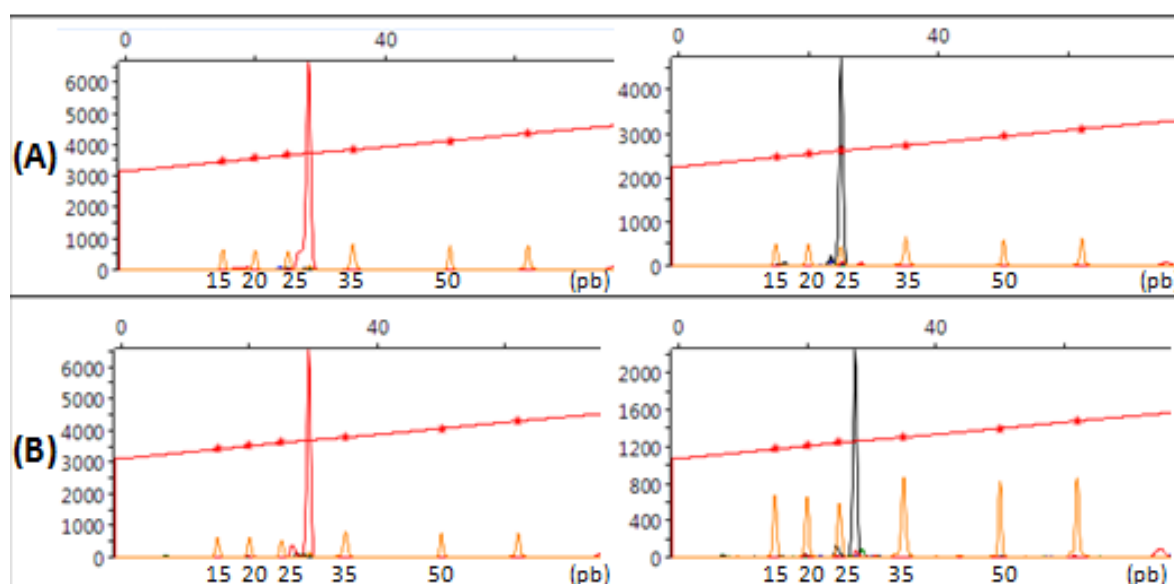


Figura 12: Genotipagem dos SNP utilizando a técnica SNaPshot™. As análises foram realizadas pelo software Peak Scanner versão 1.0. Em (A), os dois alelos gerados pelo marcador L 25, e em (B), os dois alelos gerados pelo marcador L 26. Os picos em laranja, são os marcadores moleculares de pesos conhecidos (GS 120 LIZ – Applied Biosystems).

Após a validação, com sucesso, do miniseqüenciamento utilizando os marcadores L25 e L26, foi realizada uma reação multiplex para genotipagem utilizando em um mesmo produto de PCR os dois marcadores em questão. Estes dois marcadores, podem ser utilizados em uma mesma reação, pois possuem tamanhos diferentes pela inserção da

cauda poli A na extremidade 5', gerando fragmentos de tamanhos diferentes, permitindo a genotipagem em multiplex.

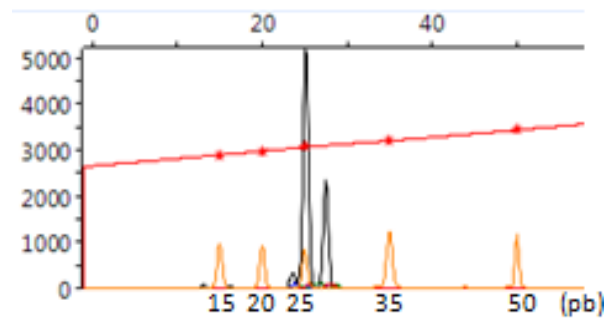


Figura 13: Genotipagem dos SNPs utilizando a técnica SNaPshot™ em uma reação multiplex. O primeiro pico em preto, evidencia o alelo gerado através do uso do marcador L 25, e o segundo pico em preto, mostra o alelo gerado pelo uso do marcador L 26. Os picos em laranja, são os marcadores moleculares de pesos conhecidos (GS 120 LIZ – Applied Biosystems).

A reação multiplex, mostrou-se eficiente (Figura 13) quando utilizada com os marcadores L25 e L26, podendo então, serem usadas mais adiante para estudos de genotipagem de SNPs em populações contrastante ao déficit hídrico e em estudos populacionais com indivíduos pertencentes as espécies *E. grandis*, *E. urophylla* e o híbrido “Urograndis”.

6 Considerações finais

A identificação e validação de novos marcadores moleculares nas espécies de *Eucalyptus* em genes expressos, é uma estratégia que vem sendo utilizada por diferentes grupos de pesquisa. A disponibilidade destes novos marcadores, permite a utilização, dos mesmos, em programas de melhoramento, genotipagem em programas clonais e estudos de genética de população.

Por se tratar de genes que, segundo a literatura, apresentam um grande impacto à resposta da planta ao estresse hídrico, os marcadores tipo SNP desenvolvidos e validados, neste trabalho, poderão ser avaliados em progênies obtidas de cruzamentos controlados, segregantes para características em estudo, onde poderá ser avaliada a real participação deste polimorfismo na detecção do estresse.

7 Conclusões

- 1) O gene CMO de eucalipto, apresenta a mesma estrutura, em relação a número de íntrons e éxons, além de alta similaridade da região codificante, em comparação com a espécie modelo *Arabidopsis thaliana*.
- 2) Nos 3885 pb seqüenciados do gene CMO, foram encontrado 49 SNPs, sendo 12 localizados em regiões codificantes e 37 em regiões não-codificantes (íntrons e regiões UTRs).
- 3) Não houve relação entre indivíduos tolerantes e suscetíveis através dos SNPs.
- 4) As três populações de eucalipto analisadas, formaram 18 haplótipos e 16 combinações haplotípicas possíveis.
- 5) Os genótipos H17/H17 e H17/H18, foram exclusivos para alguns indivíduos de *E. urophylla*, e o genótipo H13/H15 para *E. grandis*.
- 6) A técnica de genotipagem através da extensão alelo-específica mostrou-se eficiente nas amostras, podendo ser utilizadas, mais adiante para estudos de genotipagem de SNPs em populações contrastantes ao déficit hídrico e em estudos populacionais em eucalipto.

8 Referências

ABATEPAULO, AR, CAETANO, AR, MENDES JR, CT, Detection of SNPs in bovine immune-response genes that may mediate resistance to the cattle tick *Rhipicephalus (Boophilus) microplus*. **Animal Genetics**, v.39, n.3, p.328-329, 2008.

ABRAF. 2009. Anuário estatístico da ABRAF: ano base 2008. **ABRAF**, Brasília. 120p.

AGBOMA, PC, SINCLAIR, TR, JOKINEN, K, PELTONEN-SAINIO, P, PEHU, E, An evaluation of the effect of exogenous glycinebetaine on the growth and yield of soybean: timing of application, watering regimes and cultivars. **Field Crops Research**. 54, 1: 51-64, 1997.

AHMAD, R et al. Stress-induced expression of choline oxidase in potato plant chloroplasts confers enhanced tolerance to oxidative, salt, and drought stresses. **Plant Cell Rep**. 27, 687–698, 2008.

ALTSCHUL, SF, GISH, W, MILLER, W, MYERS, EW, LIPMAN, DJ, Basic local alignment search tool. **Journal Molecular Biology**, v.215, n.3, p.403-410, 1997.

ARRIEL, EF, **Divergência genética em *Cnidocolus phyllacanthus* (Mart.) Pax. EtK. Hoffm.** 2004. 89f. Tese (Doutorado em Produção Vegetal) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, 2004.

ASHRAF, M, FOOLAD, MR, Roles of glycine betaine and proline in improving plant abiotic stress resistance. **Environmental and Experimental Botany**. 59: 206-216, 2007.

ASSIS, TF, Production and use of Eucalyptus hybrids for industrial purposes. In: DUNGEY, HS, DIETERS, MJ, NIKLES DG, comp. **Híbrid Breeding and Genetics of Forest trees: Proceedings of QFRI/CRC-SPF Symposium**, 9-14, Noosa. Brisbane: Department Industries, 63-74, 2000.

BALLANTYNE, JS, CHAMBERLIN, ME, Regulation of cellular amino acid levels. In: **Cellular and Molecular Physiology of Cell Volume Regulation** (K Strange ed), CRC Press, Boca Raton, p.111-122, 1994.

BELHASSEN, E, THIS, D, e MONNEVEUX, P, L'adaptation génétique face aux contraintes de sécheresse. **Cahiers Agricultures**, v.4, p.251-261. 1995.

BENKO-ISEPPON, AN, WINTER, P, HUETTEL, B, STAGINNUS, C, MUEHLBAUER, FJ KAHL, G, Molecular markers closely linked to fusarium resistance genes in chickpea show significant alignments to pathogenesis-related genes located on Arabidopsis chromosomes 1 and 5. **Theor Appl Genet.** 107 (2): 379-86, 2003.

BHUIYAN, NH, HAMADA, A, YAMADA, N et al, Regulation of betaine synthesis by precursor supply and choline monooxygenase expression in *Amaranthus tricolor*. **Journal of Experimental Botany.** 58, 4203-4212, 2007.

BLUM, A, Crop responses to drought and the interpretation of adaptation. **Plant Growth Regulation.** v.20, p.135-148. 1996.

BOHNERT, H.J.; SU, H.; SHEN, B. Molecular mechanisms of salinity tolerance. In: SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Molecular responses to cold, drought, heat, and salt stress in higher plants. Austin, Texas: R.G. Landes , 1999. p. 29-62.

BOLAND, DJ, BROOKER, MIH, CHIPPENDALE, GM, HALL, N, KLEINIG, DA, TURNER, HD, **Forest Tree of Australia**, 4 ed., Australia, CSIRO, 1984. 687 p.

BORTOLOTO, T.M. **Construção de mapa molecular de uma população de Eucalyptus grandis segregante para o caráter florescimento precoce utilizando marcadores RAPD.** Botucatu, 2006, Dissertação (para a obtenção do título de Mestre em Ciências Biológicas – Área de concentração em Genética)– Instituto de Biociências, Universidade Estadual Paulista.

BRACELPA- Associação Brasileira de Celulose e Papel. Relatório anual 2008/2009. Disponível em: <<http://www.bracelpa.org.br>>. Acesso em: 2 mar. 2010

BRAY, EA, BAILEY-SERRES, J, WERETILNYK, E, Responses to abiotic stress. In: Buchanan, B.; Gruissem, W.; Jones, R. (eds). **Biochemistry & Molecular Biology of Plants.** 3a Impressão. 2001. American Society of Plant Physiologists, Rockville, Cap.22, p. 1167-1168.

BRAY, EA, Plant responses to water deficit. **Trends Plant Science**, v.2, p. 48-54, 1997.

BRESSAN, RA, HASEGAWA, PM, LOCY, RD, **Fisiologia do estresse.** In: TAIZ, L.; ZEIGER, E. **Fisiologia Vegetal.** Porto Alegre: Artmed, Trad. Eliane Romanato Santarém et al., p.613-643, 2004.

BRINKMANN, H.; e MARTIN, W. **Plant Mol. Biol.** 30, 65–75, 1995.

BROOKES, AJ, The essence of SNPs. **Gene**, v.234, n.2, p.177-186, 1999.

BROOKER, MIH, KLEINING, DA, **Field guide to eucalypts Melbourne and Sydney**. Melbourne: Inkata Press, 428 p. 1990.

CALLISTER, AN, ARNDT, SK, ADAMS, MA, Comparison of four methods for measuring osmotic potential of tree leaves. **Physiologia Plantarum**, v.127, p.383-392, 2006.

CALVO JH, MARTÍNEZ-ROYO A, SILVERI L, et al, Isolation, mapping and identification of SNPs for four genes (ACP6, CGN, ANXA9, SLC27A3) from a bovine QTL region on BTA3. **Cytogenetics and Genome Research**, v.114, n.1, p.39-43, 2006.

CAMBRAIA, J. Aspectos bioquímicos, celulares e fisiológicos dos estresses nutricionais em plantas. In: Eds. NOGUEIRA, R.J.M.C.; ARAÚJO, E.L.;

WILLADINO, LG, CAVALCANTE, UMT, **Estresses ambientais: danos e benefícios em plantas**, p.95-105, 2005.

CARLIN, SD, Indicadores fisiológicos da interação entre déficit hídrico e acidez do solo em cana-de-açúcar. **Pesq. agropec. Bras.** V. 44, 9: 1106-1113, 2009.

CARVALHO, CJR, Respostas de plantas de *Schizolobium amazonicum* [*S. parahyba* var. *amazonicum*] e *Schizolobium parahyba* [*Schizolobium parahybum*] à deficiência hídrica. **Revista Árvore**, v.29, p.907-914, 2005.

CASAGRANDE, EC, FARIAS, JRB, NEUMAIER, N, OYA, T, PEDROSO, J, MARTINS, PK, BRETON, M, C, NEPOMUCENO, AL, Expressão gênica diferencial durante déficit hídrico em soja. **Revista Brasileira de Fisiologia Vegetal**, v. 13, n. 2, p. 168-184, 2001.

CHA-UM, S, SUPAIBULWATANA, K, KIRDMANEE, C, Water relation, photosynthetic ability and growth of Thai jasmine rice (*Oryza sativa* L. ssp. *indica* cv. KDML 105) to salt stress by application of exogenous glycinebetaine and choline. **Journal Agronomy & Crop Science**, v. 192, p. 25-36, 2006.

CHAMARY, JV, PARMLEY, JL, HURST, LD, Hearing silence: non-neutral evolution at synonymous sites in mammals, **Nat. Rev. Genet.** 7:98–108, 2006.

CHEN, THH, MURATA, N, Enhancement of tolerance of abiotic stress by metabolic engineering of betaines and other compatible solutes. *Current Opinion in Plant Biology*, v.5, p. 250-257, 2002.

CHING, A, et al, SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. **BMC Genetics**, v.3, p.19, 2002.

CLAUSSEN, W, Proline as a measure of stress in tomato plants. **Plant Science**, v. 168, p. 241-248, 2005.

CLINE, K. e HENRY, R. (1996) **Annu. Rev. Cell. Dev. Biol.** 12, 1–26.

CREGAN, P, RANDALL, N, YOULIN, Z, Sequence variation, haplotype diversity and linkage disequilibrium in cultivated and wild soybean. In: **First International Conference on Legume Genomics and Genetics**. Plenary Session IV: Structural Genomics. Minneapolis-St. Paul, 2002.

CUSHMAN, JC, BOHNERT, HJ, Genomic approaches to plant stress tolerance. **Current Opinion in Plant Biology**, v.3, p. 117-124, 2000.

DOMINGUES, DS. **Identificação de marcador RAPD e SCAR relacionado ao caráter florescimento precoce em *Eucalyptu grandis***. Botucatu, 2003, 24p. Trabalho científico (exigência parcial para obtenção do título de Bacharel em Ciências Biológicas) – Instituto de Biociências, Universidade Estadual Paulista.

DRAZEN, JM, YANDAVA, CN, DUBE, L, et al. Pharmacogenetic association between ALOX5 promoter genotype and the response to anti-asthma treatment. **Nat Genet**, v. 22, p. 168-170, 1999.

DVORNYK, V, SIRVIÖ, A, MIKKONEN, M, *et al*, Low Nucleotide Diversity at the pal1 locus in the widely distributed *Pinus sylvestris*. **Mol Biol Evol**, v. 19, n. 2, p. 179-188, 2002.

ELDRIDGE, K, **Eucalypt domestication and breeding**.ed. Oxford, New York: Clarendon Press ;Oxford University Press, 1993.

EMBRAPA FLORESTAS. Sistemas de Produção, 4 ago. 2003. Disponível em:<http://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Eucalipto/CultivodoEucalipto/02_indicacao_de_especies.htm> Acesso em: 2 mar. 2010.

FALLON, KM, PHILLIPS, R, Responses to water stress in adapted and unadapted carrot cell suspension cultures. **Journal of Experimental Botany**, v. 40, p. 681-687, 1989.

FAO. **Eucalypts for Planting**: Forestry Ser.ed.: Fao, 1981.

FERREIRA, ME, GRATTAPAGLIA, D, Introdução ao uso de marcadores moleculares em análises genéticas. Brasília: **Embrapa-Cenargen**, 3 ed. 220p. 1998.

FINN, R.D.; TATE, J.; MISTRY, J.; COGGILL, P.C.; SAMMUT, J.S.; HOTZ, H.R.; CERIC, G.; FORSLUND, K.; EDDY, S.R.; SONNHAMMER, E.L. ; BATEMAN, A. The Pfam protein families database. **Nucleic Acids Research**, v.36, p.281-288, 2008.

FLEXAS, J, BOTA, J, ESCALONA, JM, SAMPOL, B, MEDRANO, H. Effects of drought on photosynthesis in grapevines under field conditions. **Functional Plant Biology**, v.29, p.461-471, 2002.

GANAL, MW, ALTMANN, T, RODER, MS, SNP identification in crop plants. **Plant Biol.** 12(2):211-7. 2009.

GAO, C.; HAN, B. Evolutionary and expression study of the aldehyde dehydrogenase (ALDH) gene superfamily in rice (*Oryza sativa*). **Gene**, V.431, P. 86-94, 2009.

GASTEIGER, E.; GATTIKER, A. ; HOOGLAND, C. ; IVANYI, I. ; APPEL, R.D.; BAIROCH, A. ExpASY: the proteomics server for in-depth protein knowledge and analysis. **Nucleic Acids Research**, v.31, p.3784-3788, 2003.

GIRIJA, C, SMITH, BN, SWAMY, PM, Interactive effects of sodium chloride and calcium chloride on the accumulation of praline and glycinebetaine in peanut (*Arachis hypogaea* L.). **Environmental Experimental Botany**, v.47, p.1–10, 2002.

GONZÁLEZ-MARTÍNEZ, SC, BROWN, GR, ERSOZ, E, *et al*, Nucleotide diversity, linkage disequilibrium and adaptive variation in natural populations of loblolly pine. In: **12nd Conference International Plant and Animal Genomes**, San Diego, 2004. W3.

GONZALEZ-MARTINEZ SC, ERSOZ E, BROWN GR, WHEELER NC, NEALE DB: DNA sequence variation and selection of tag SNPs at candidate genes for drought-stress response in *Pinus taeda* L. **Genetics**. 172:1915-1926, 2006.

GORHAM, J, K, JOKINEN, **Glycinebetaine treatment improves cotton yields in field trials in Pakistan. Abstracts.** World Cotton Research Conference II, Athens, Greece. August, p. 329, 1998.

GRATTAPAGLIA, D.; H. J. BRADSHAW. Nuclear DNA content of commercially important *Eucalyptus* species and hybrids. **Canadian Journal of Forest Research**, v.24,n.(5): p. 1074-1078, 1994.

GRIFFIN, TJ, SMITH, LM, Single-nucleotide polymorphism analysis by MALDI-TOF mass spectrometry. **Trends in Biotechnology**, v.18, n.2, p.77-84, 2000.

GRIVET, L et al, ESTs as a source for sequence polymorphism discovery in sugarcane: example of the *Adh* genes. **Theoretical and Applied Genetics**, v.106, n.2, p.190-197, 2003.

GUPTA, PK, ROY, JK, PRASAD, M, Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. **Current Science**, v.80, n.4, p.4524-535, 2001.

GUPTA, P, LEE, KH, Silent mutations result in HlyA hypersecretion by reducing intracellular HlyA protein aggregates, **Biotechnol. Bioeng.** 101:967–974, 2008.

HAYASHI, H, ALIA, MUSTARDY, L, DESHNIUM, P, IDA, MURATA, N, Transformation of *Arabidopsis thaliana* with the coda gene for choline oxidase; accumulation of glycine betaine and enhanced tolerance to salt and cold stress. **Plant J.** 12, 133–142, 1997.

HAMMER, MF, A recent common ancestry for human Y chromosomes. **Nature**, v.378, n.6555, p.376-378, 1995.

HARE, PD, CRESS, WA, VAN STADEN, J, Dissecting the roles of compatible osmolyte accumulation during stress. **Plant Cell Environ**, v.21, p.535–553, 1998.

HASEGAWA, PM, BRESSAN, RA, ZHU, JK, BOHNERT, HJ, Plant cellular and molecular responses to high salinity. **Annual Review of Plant Physiology and Plant Molecular Biology**, v.51, p. 463-499, 2000.

HIBINO T, WADITEE R, ARAKI E, ISHIKAWA H, AOKI K, TANAKA Y, TAKABE T, Functional characterization of choline monooxygenase, an enzyme for betaine synthesis in plants. **J Biol Chem** 277:41352–41360, 2002.

HUANG, J, et al, Genetic engineering of glycinebetaine production toward enhancing stress tolerance in plants: metabolic limitations. **Plant Physiol.** 122, 747–756, 2000.

IKUTA, S, IMAMURA, S, MISAKI, H, HORIUTI, Y, Purification and Characterization of Choline Oxidase from *Arthrobacter globiformis*. **Journal of Biochemistry.** 82: 1741-1749, 1977.

JONES, HG, CORLETT, JE, Current topics in drought physiology. **Journal of Agricultural Science**, v.119, p. 291-296, 1992.

JUNGHANS, DT, ALFENAS, AC, BROMMONSCHENKEL, SH, ODA, S, MELLO EJ, GRATTAPAGLIA, D, Resistance to rust (*Puccinia psidii* Winter) in eucalyptus: mode of inheritance and mapping of a major gene with RAPD markers. **Theor Appl Genet.** 175-80. 2003.

JUNGHANS, DT. **Quantificação da severidade, herança de resistência e identificação de marcadores RAPD ligados à resistência à ferrugem (*Puccinia psidii*) em *Eucalyptus grandis***. Viçosa, 2000. 53p. Tese (doutorado em Fitopatologia) – Universidade Federal de Viçosa.

KIRCH, HH, BARTELS, Y, WEI, PS, SCHNABLE, WOOD, AJ, The ALDH gene superfamily of Arabidopsis, **Trends Plant Sci.** 9, 371–377, 2004.

KORYO, HW, Effect of salinity on growth, photosynthesis, water relations and solute composition of the potential cash crop halophyte *Plantago coronopus* (L.). **Environmental and Experimental Botany**, v.56, n.2p.136-146, 2006.

KOTA, R, et al, A. Application of denaturing high-performance liquid chromatography for mapping single nucleotide polymorphisms in barley (*Hordeum vulgare* L.). **Genome**, v.44, n.4, p.523–528, 2001.

KRAMER, PJ, BOYER, N, **Water relations of plants and soils**. New York: MacGraw-Hill, 196p, 1995.

KRAWEZAK, M, REISS, J, COOPER, DN, The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. **HumGenet**, v. 90, p. 41-54, 1992.

KUDLA, G, et al, Coding-sequence determinants of gene expression in *Escherichia coli*. **Science** 324: 255–258, 2009.

KÜLHEIM C, YEOH, SH, MAINTZ J, FOLEY WJ, MORAN GF, Comparative SNP diversity among four *Eucalyptus* species for genes from secondary metabolite biosynthetic pathways. **BMC Genomics** 10:452-463, 2009.

KWOK, PY, GU, Z, Single nucleotide polymorphism libraries: why and how are we building them? **Molecular Medicine Today**, v.5(12), p. 538-543, 1999.

LAI, E, RILEY, J, PURVIS, I. et al. A 4-Mb high-density single nucleotide polymorphism-based map around human APOE. **Genomics**, v.54, n.1, p.31-38, 1998.

LANDFALD, B, STROM, AR, Choline-glycine betaine pathway confers a high level of osmotic tolerance in *Escherichia coli*. **Journal of Bacteriology** 165: 849–855, 1986.

LARCHER, W. A Planta sob Estresse. **Ecofisiologia vegetal**. São Carlos: RiMa Artes e Textos, p.341-478, 2004.

Le HELLARD, S. et al. SNP genotyping on pooled DNAs: comparison of genotyping technologies and a semi automated method for data storage and analysis. **Nucleic Acids. Res.**, v.30, n.15, p.e74, aug. 2002.

Le RUDULIER, D, STROM, AR, DANDEKAR, AM, SMITH, LT, VALENTINE, R C. **Molecular biology of osmoregulation**. Science, v. 224, p. 1064-1068, 1984.

LEVY, D, et al, Genome-wide association study of blood pressure and hypertension. **Nature Genetcs**. 41: 677-687, 2009.

LEWIN, B. **Genes IX**. 9 ed. Porto Alegre. Artmed, 2008, 15-16, 60 p.

LIAO, P, LEE, KH, From SNPs to functional polymorphism: The insight into biotechnology applications. **Biochemical Engineering Journal**. 149-158, 2010.

LOW, PS, Molecular basis of the biological compatibility of nature's osmolytes. In: Transport Processes, Ion and Osmoregulation (R Gilles, M Gilles-Baillien eds), **Springer-Verlag**, Berlin. p.469-477. 1985.

LUDLOW, MM, SANTAMARIA, JM e FUKAI, S, Contribution of osmotic adjustment to grain yield in *Sorghum bicolor* (L.) Moench under water-limited conditions. II. Water stress after anthesis. **Australian Journal of Agricultural Research** v.41, p.67-78. 1990.

LV, S. et al, Increase of glycinebetaine synthesis improves drought tolerance in cotton. **Mol. Breed**. 20, 233–248, 2007.

MANSOUR, MMF, Nitrogen containing compounds and adaptation of plants to salinity stress. **Biol. Plant**. 43, 491–500, 2000.

MARTINS FILHO, S, SEDIYAMA, CS, MOREIRA, MA, BARROS EG, RAPD and SCAR markers linked to resistance to frogeye leaf spot in soybean. **Genetics and Molecular Biology**, v.25, n.3, p.317-321, 2002.

MASON, JR, CAMMACK, R, The electron-transport proteins of hydroxylating bacterial oxygenases. **Annu. Rev. Microbiol**. 46:277–305, 1992.

McCUE, KF, HANSON, AD, Drought and salt tolerance: towards understanding and application. **Trends in Biotechnology**, v. 8, p. 358–362, 1990.

McKINNON, GE, JORDAN, G J, VAILLANCOURT, RE, *et al*, Glacial refugia and reticulate evolution: the case of the Tasmanian eucalypts. **Philos Trans R Soc Lond B Biol Sci**, v. 29, n. 359, p. 275-284, 2004.

MELONI, DA, GULOTTA, MR, MARTINEZ, CA, OLIVA, MA, The effects of salt stress on growth, nitrate reduction and proline and glycinebetaine accumulation in *Prosopis alba*. **Brazilian Journal of Plant Physiology**, v. 16, n. 1, p. 39-46, 2004.

MERCHANT, A, CALLISTER, A, ARNDT, S, TAUSZ M, ADAMS, M, Contrasting physiological responses of six Eucalyptus species to water deficit. **Annals of Botany** 100:1507-1515, 2007.

MISSIHOUN, TD, SCHMITZ, J, KLUG, R, et al, Betaine aldehyde dehydrogenase genes from Arabidopsis with different sub-cellular localization affect stress responses. **Planta**. 2010.

MOHANTY, A, KATHURIA, H, FERJANI, A, SAKAMOTO, A, MOHANTY, P, MURATA, N, TYAGI, AK, Transgenics of an elite indica rice variety Pusa Basmati 1 harbouring the codA gene are highly tolerant to salt stress. **Theor. Appl. Genet.** 106, 51–57, 2002.

MORA, AEC, GARCIA, A Cultura do Eucalipto no Brasil. 1.ed. São Paulo: **Sociedade Brasileira de Silvicultura**, 2000.

MORETZSON, MC, NUNES, CDM, FERREIRA, ME, GRATTAPAGLIA, D. RAPD linkage mapping of the shell thickness locus in oil palm (*Elaeis guineensis* Jacq.) **Theor Appl Genetic.** 100:63-70, 2000.

MORGAN, JM, Osmoregulation and water stress in higher plants. **Annual Review of Plant Physiology** v.35, p.299-319. 1984.

NASU, S, et al, Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa*, *Oryza rufipogon*) and establishment of SNP markers. *DNA Research*, v.9, n.5, p.163–171, 2002.

NEI, M.; LI, W.-H. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA*, v. 76, n. 10, p. 5269-5273, 1979.

NEI, M. *Molecular evolutionary genetics*. New York: Columbia University Press, 512 p. 1987.

NEPOMUCENO, AL, NEUMAIER, N, FARIAS, JRB, OYA, T, Tolerância à seca em plantas. **Embrapa-Londrina**. 10p. 2001.

NEVES, LG, et al, Diversidade Nucleotídica e utilização de SNPs para o mapeamento de genes candidatos em *Eucalyptus* spp. **Embrapa**. Outubro, 2008.

NOGUEIRA, RJMC, MORAES, JAPV, BURITY, HÁ, BEZERRA NETO, E, Alterações na resistência à difusão de vapor das folhas e relações hídricas em aceroleiras submetidas a déficit de água. **Revista Brasileira de Fisiologia Vegetal**, v. 1, p. 75-87, 2001.

NOVAES, E, DROST, DR, FARMERIE, WG, et al, High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. **BMC Genomics**. 9: 320, 2008.

OLIVEIRA, ACB, CAIXETA, ET, ZAMBOLIM, EM, ZAMBOLIM, L, SAKIYAMA, NS, **Aplicação técnica de marcadores moleculares no melhoramento de plantas**:. Campinas: Instituto Agronômico, 17p. (Documentos IAC, 81, ISSN: 1809 – 7693), 2007.

ORAGUZIE, NC, et al, Association Mapping in Plants. New York, USA: **Springer**, v.IX, p.277, 2007.

ORITA, M, IWAHANA, H, KANAZAWA, H, et al. Detection of polymorphisms of human DNA by gel electrophoresis as singlestrand conformation polymorphisms. **Proceedings of the National Academy of Sciences USA**, v.86, n.8, p.2766-70, 1989.

PAIVA, AS, FERNANDES, EJ, RODRIGUES, TJD, TURCO, JEP, Condutância estomática em folhas de feijoeiro submetido a diferentes regimes de irrigação. **Engenharia Agrícola**, v.25, p.161-169, 2005.

PARK, EJ, et al, Glycinebetaine accumulation in chloroplasts is more effective than that in cytosol in protecting transgenic tomato plants against abiotic stress. **Plant Cell Environ**. 30, 994–1005, 2007.

PINTO G, LOUREIRO, J, SANTOS, C, Analises of the genetic stability of *Eucalyptus globulus* Labill. Somatic embryos by flow cytometry. **Theoretical and Applied Genetcs**. 109: 580-587, 2004.

POT, D.; CHANTRE, G.; ROZENBERG, P.; RODRIGUES, J-C.; JONES, G. L.; PEREIRA, H.; HANNRUP, B.; CAHALAN, C.; PLOMION, C. Genetic control of pulp and timber properties in maritime pine (*Pinus pinaster* Ait.). **Annual of forest science**, [S.l.], v. 59, p. 563-575, 2002.

PRYOR, L, *Eucalyptus*. In: **A. Halevy CRC Handbook of Flowering**. 1.ed. Boca Raton: Franklin Book Company, Incorporated. 1985. v.2:p. 476-482.

QUINTIN, K, 2001 Queensland Term Wildlife Field Guide: *Eucalyptus grandis*(Flooded gum). Disponível em: <<http://people.hws.edu/fieldguide/delete.asp?ID=120&dlete=no>>. Acesso em: 24 de maio de 2004.

RAFALSKI, A, Applications of single nucleotide polymorphisms in crop genetics. **Curr Opin Plant Biol**, v. 5, n. 2, p. 94-100, 2002.

RATHINASABAPATHI, B. 2000. Metabolic engineering for stress tolerance: installing osmoprotectant synthesis pathways. **Annals of Botany** 86: 709–716.

RHODES, D, HANSON, AD, Quaternary ammonium and tertiary sulfonium compounds in higher plants. Annual Review in **Plant Physiology and Plant Molecular Biology**, v.44, p. 357-384, 1993.

RIZHISKY, L, LIANG, H, SHUMAN, J, SHULAEV, V, DAVLETOVA, S, MITTLER, R, When defense pathways collide. The response of Arabidopsis to a combination of drought and heat stress. **Plant Physiology**, v. 134, p. 1683-1696, 2004.

ROZAS, J. et al. DnaSP, DNA polymorphism analyses by the coalescent and other methods. **Bioinformatics**, v.19, n.18, p.2496-2497, dec. 2003.

RONTEIN, D, BASSET, G, HANSON, AD, Metabolic engineering of osmoprotectant accumulation in plants. **Metabolic Engineering**, v.4, p. 49-56, 2002.

RONTEIN, D, DIEUAIDE-NOUBHANI, M, DUFOURC, E.J, RAYMOND, P, ROLIN, D, The metabolic architecture of plant cells: stability of central metabolism and flexibility of anabolic pathways during the growth cycle of tomato cells. **Journal of Biological Chemistry** v.277, p.43948-43960, 2002.

ROZWADOWSKI, K. L., KHACHATOURIANS, G. G. e SELVARAJ, G. J. **Bacteriol.** 173, 472–478, 1991.

RUSSELL, J, BOOTH, A, FULLER, J, *et al*, A comparison of sequence-based polymorphism and haplotype content in transcribed and anonymous regions of the barley genome. **Genome**, v. 47, p. 389-398, 2004.

SAKAMOTO, A, MURATA, N, Genetic engineering of glycinebetaine synthesis in plants: current status and implication for enhancement of stress tolerance. **J. Exp. Bot.** 51, 81–88, 2000.

SANTOS, PET, GARCIA, CH, MORI, ES, MORAES, MLT, Potencial para programas de melhoramento, estimativa de parâmetros genéticos, interação progênies x locais em populações de *Eucalyptus urophylla* ST. Blake. **Boletim Técnico do IPEF**, v. 43, n. 44, p 11 - 19, 1990.

SANTOS, RF, CARLESSO, R, Déficit hídrico e os processos morfológicos e fisiológicos das plantas. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 2(3), p. 287-294, 1998.

SANTOS, S.N. **Genes de Lignificação em Eucalyptus: estrutura e diversidade dos genes 4cl e ccoaomt**. 2005. 229 p. Dissertação (Mestrado em Ciências Genômicas e Biotecnologia) - Universidade Católica de Brasília, Brasília, D.F.

SBS- Sociedade Brasileira de Silvicultura. Fatos e Números do Brasil Florestal, dez. 2008. Disponível em: <<http://www.sbs.org.br/estatisticas.htm>> Acesso em: 2 mar. 2010

SCARPINELLA GA, **Reflorestamento no Brasil e o Protocolo de Quioto**. São Paulo, 2002, Dissertação (exigência para obtenção do título de mestre em energia) – Instituto de Eletrotécnica e Energia –IEE, Universidade de São Paulo.

SCHLOTTERER C: The evolution of molecular markers - just a matter of fashion? **Nature Reviews Genetics**, v.5, n.1, p.63-69, 2004.

SCHNEIDER, K. et al. SNP frequency and allelic haplotype structure of Beta vulgaris expressed genes. **Molecular Breeding**, v.8, n.1, p.63–74, 2001.

SAVOLAINEN, O.; PYHAJARVI, T. Genomic diversity in forest trees. **Current opinion in plant biology**, [S.l.], v. 10, p. 162-167, 2007.

SEKI, M, KAMEI, A, YAMAGUCHI-SHINOZAKI, K, SHINOZAKI, K, Molecular responses to drought, salinity and frost: common and different paths for plant protection. **Current Opinion of Biotechnology** v.14, p.194-199, 2003.

SEKI, M, UMEZAWA, T, URANO, K, SHINOZAKI, K, Regulatory metabolic networks in drought stress responses. **Current Opinion in Plant Biology**, v.10, p.296:302, 2007.

SILVA, E, **Avaliação quantitativa de impactos ambientais do reflorestamento no Brasil**. Viçosa, 309p. Tese (Doutorado/ Engenharia Florestal) - Universidade Federal de Viçosa, 1994.

SILVA, JRR, **Comportamento ecofisiológico de plantas jovens de andiroba (carapa guianensis aubl.) sob dois regimes hídricos**. Belém, 2009. Tese. Universidade Federal Rural da Amazônia.

SILVA, MR, **Caracterização morfológica, fisiológica e nutricional de mudas de Eucalyptus grandis Hill ex Maiden submetidas a diferentes níveis de estresse hídrico durante a fase de rustificação**. 1998. 105f. Dissertação (Mestrado e Silvicultura) - Universidade do Federal do Paraná, Curitiba, 1998.

SIMKO, I, HAYNES, KG, EWING, EE, *et al*, Mapping genes for resistance to Verticillium albo-atrum in tetraploid and diploid potato populations using haplotype association tests and genetic analysis. **Mol Genet Genomics**, v. 271, p. 522-531, 2004.

SIRONI, M.; FORNILI, A.; FORNILI, S. L. Water interaction with glycine betaine : A hybrid QM/MM molecular dynamics simulation. **Physical Chemistry Chemical Physics**, Cambridge, v. 3, n. 6, p. 1081-1085, 2001.

SOMERO, GN, Protons, osmolytes, and fitness of internal milieu for protein function. **American Journal of Physiology** v.251, p.197-213. 1986.

SOUZA, MJH, RIBEIRO, A, LEITE, HG, LEITE, FP, MINUZZI, RB, Disponibilidade hídrica do solo e produtividade do eucalipto em três regiões da Bacia do Rio Doce. **Revista Árvore**, v.30, p.399-410, 2006.

STAPE, JL, BINKLEY, D, RYAN, MG, Eucalyptus production and the supply, use and efficiency of use of water, light and nitrogen across a geographic gradient in Brazil. **Forest Ecology and Management**, v.193, p.17-31, 2004.

STEFANUTO, V, A, **Efeito do cálcio na homeostase de brotações de um clone de Eucalyptus grandis Hill (ex Maiden) sob condições de deficiência hídrica induzida in vitro**. 2002. 65 f. Dissertação (Mestrado) - Escola Superior de Agricultura "Luiz de Queiroz" – Universidade de São Paulo, Piracicaba, 2002.

SYVANEN, A.C. Accessing genetic variation: genotyping single nucleotide polymorphisms. **Nat. Rev. Genet.**, v.2, n.12, p.930-942, dec. 2001.

TAIZ, L, ZEIGER, E. **Fisiologia Vegetal**. 3a ed., Porto Alegre: ARTMED Editora S.A., 2004.

TAKABE, et al, Accumulation of Glycinebetaine in Rice Plants that Overexpress Choline Monooxygenase from Spinach and Evaluation of their Tolerance to Abiotic Stress. **Annals of Botany**. April, 2006.

TATAGIBA, SD, PEZZOPANE, JEM, REIS, EF, Relações hídricas e trocas gasosas na seleção precoce de clones de eucalipto para ambientes com diferenciada disponibilidade de água no solo. **Floresta** 38:387-400, 2008.

TAYLOR, IB, BURBIDGE, A, THOMPSON, AJ, Control of abscisic acid synthesis. **Journal of Experimental Botany**, v.51, p.1563-1575, 2000.

TAYLOR, C. B. Proline and water deficit: ups, downs, ins, and outs. **The Plant Cell**, v. 8, p. 1221-1224, 1996.

THOMPSON, J.D. et al. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Res.**, v.25, n.24, p.4876-4882, dec. 1997.

THUMMA, B.R.; NOLAN, M.F.; EVANS, R.; MORAN, G.F. Polymorphisms in Cinnamoyl CoA Reductase (CCR) are associated with variation in microfibril angle in *Eucalyptus* ssp. **Genetics**, 2005.

TIMASHEFF, SN, The control of protein stability and association by weak interactions with water: how do solvents affect these processes? **Annual Review of Biophysical and Biomolecular Structure** v.22, p.67-97. 1993.

VAN, K, et al, Discovery of single nucleotide polymorphisms in soybean using primers designed from ESTs. **Euphytica**. 139: 147-155, 2004.

VELLINI ALTT, PAULA NF, ALVES PLCA, PAVANI LC, BONINE CAV, SCARPINATI EA, PAULA RC, Respostas fisiológicas de diferentes clones de eucalipto sob diferentes regimes de irrigação. **Revista Árvore** 32:651-663, 2008.

VERSLUES, PE, AGARWAL, M, KATIYAR-AGARWAL, S, ZHU, J, ZHU, JK, Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. **The Plant Journal**, v. 45, p. 523–539, 2006.

VICENTINI, R.; SASSAKI, F. T.; GIMENES, M. A.; MAIA,I.G.; ENOSSI, M. In silico evaluation of the *Eucalyptus* transcriptome **Genet. Mol. Biol.** vol.28 no.3 suppl.0 São Paulo 2005.

WEIMBERG, R, LERNER, HR, POLJAKOFF-MAYBER, A. Changes in growth and water soluble concentrations in *Sorghum bicolor* stressed with sodium and potassium salts. **Physiologia Plantarum**, v. 62, p. 472-480, 1984.

WRIGHT, IJ, GROOM, PK, LAMONT, BB, POOT, P, PRIOR, LA, REICH, PB, SCHULZE, ED, VENEKLAAS, EJ, WESTOBY, M. Leaf trait relationships in Australian plant species. **Functional Plant Biology**, v.31, p.551-558, 2004.

WYN JONES, RG, STOREY, R, LEIGH, RA, AHMAD, N, POLLARD, A, A hypothesis on cytoplasmic osmoregulation. In: Regulation of Cell Membrane Activities in Plants (E Marre, O Cifferi eds), **Elsevier**, Amsterdam, p.121-136, 1977.

YAMADA, H, MORI, N, TANI, Y. **Agric. Biol. Chem.** 43, p. 2173- 2177, 1979.

YAMAGUCHI-SHINOZAKI, K, WADA, K, HARADA, Y, SHINOZAKI, K, Correlation between the induction of a gene for delta1-pyrroline -5-carboxylate synthetase and the accumulation of proline in *Arabidopsis thaliana* under osmotic stress. **The Plant Journal** v.7, p.751-760, 2005.

YANCEY, PH, Compatible and counteracting solutes. In: **Cellular and Molecular Physiology of Cell Volume Regulation** (K Strange ed), CRC Press, Boca Raton, p.81-109. 1994.

YANG, CM, WU, JT, SHEN BK, YU, JZ, LO, CT, SHEN,Y. Estimating growth and nitrogen status in plants of *Amaranthus mangostanus* with canopy spectral characteristics. **Journal of Agricultural Research**, v.52, p. 268-290, 2003.

YORDANOV, I, VELIKOVA, V, TSONEV, T, Plant responses to drought and stress tolerance. **Bulgarian Journal of Plant Physiology**, Special Issue, p. 187–206, 2003.

ZHANG, DX, HEWITT, GM, Molecular DNA analyses in genetic studies of populations: practice, problems and prospects. **Mol. Ecol.**, v.12, n.3, p.563-584, jun. 2003.

ZHANG, J, TAN, W, YANG, X, Plastid-expressed choline monooxygenase gene improves salt and drought tolerance through accumulation of glycine betaine in tobacco. **Plant Cell Rep.** 27: 1113-1124, 2008.

ZHU, YL, et al, Single-nucleotide polymorphisms in soybean. **Genetics**, v.163, n.3, p.1123, 2003.

ZHU, J.K, HASEGAWA, PM, BRESSAN, RA, Molecular aspects of osmotic stress in plants. **Critical Review of Plant Science**, v.16, p. 253–277, 1997.