

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

Andreia Santoro

**Identificação de *Single Nucleotide Polymorphisms* (SNPs)
no gene Nove-cis-epoxicarotenóide dioxigenase (NCED)
em *Eucalyptus***

Botucatu-SP

2010

Andreia Santoro

**Identificação de *Single Nucleotide Polymorphisms* (SNPs)
no gene Nove-cis-epoxicarotenóide dioxigenase (NCED)
em *Eucalyptus***

Dissertação apresentada ao Instituto de
Biotecnologia, Universidade Estadual Paulista
"Julio de Mesquita Filho" - Campus de
Botucatu para a obtenção do título de Mestre
em Ciências Biológicas - Genética.

Orientador: Celso Luís Marino

Botucatu-SP

2010

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

Santoro, Andreia.

Identificação de *Single Nucleotide Polymorphisms* no gene Nove-cis-epoxicarotenóide dioxigenase (NCED) em *Eucalyptus* / Andréia Santoro. – Botucatu, 2010

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

Orientador: Celso Luis Marino

Capes: 20203004

1. Eucalipto. 2. Código genético. 3. Plantas – Melhoramento genético.

Palavras chave: Ácido abscísico; *Eucalyptus*; Haplótipos; NCED; SNPs.

Aos meus amados pais, Isabel e Antonio, minha fortaleza de amor e paz.

Ao meu irmão querido Daniel, sei que posso contar sempre!

À Pietra, sobrinha linda, alegria em minha vida.

Ao Felipe, pessoa especial, pelo incentivo e amor.

DEDICO

AGRADECIMENTOS

Agradeço em primeiro a **Deus** e minha **Mãezinha do Céu**, que me iluminaram e abençoaram com essa graça.

Ao **Prof. Dr. Celso Luis Marino**, pela oportunidade, orientação e confiança.

Aos meus amados pais, **Isabel e Antonio**, exemplos da minha vida, fundamentais para a realização desse trabalho. Não há palavras para expressar o quanto sou grata à vocês. OBRIGADA!

Ao meu irmãozinho **Daniel**, você é muito importante para mim, amo demais. E, à minha cunhada **Alessandra** e sobrinha **Pietra**, pelo carinho e momentos felizes.

Ao **Felipe**, “mai Love”, pelo companheirismo, paciência e fonte de inspiração... “nós somos feitos um pro outro...”.

A todos do laboratório CAGEN, que estiveram dispostos a ajudar e contribuíram de alguma forma nesse trabalho: **Leonardo, Tânia, Suzana, Vanusa, Júlio, Cecília, Cintia, Helena, Karine e Juliana**.

À minhas amigas de república: **Flávia (Kalose); Gabriela; Michele e Taísa**, pelos momentos de muita alegria e felicidade na casa. Em especial, agradeço a minha irmã de coração, **Kalose**, “kabeçudinha”, pela grande amizade.

À minha avózinha querida **Olga**, tenho saudades, mas sei que está comigo sempre. E, ao meu avô **Mauro**, avó **Aparecida** e avô **Domingo** (*in memorium*), exemplos na minha vida de esforço e dignidade.

Aos meus tios, tias, primos e primas. Em especial, à minha madrinha **Elaine** e prima **Sarah**, que sempre me incentivaram, com muita alegria.

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

À minhas amigas da Barra Bonita, **Ana, Débora e Gabrielle**, pela alegria e incentivo. Adoro vocês meninas!

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

A **CAPES**, pela bolsa concedida.

Em especial, à minhas duas companheiras, **Fly e Baby**, que ficaram perto de mim e foram minhas “amigas” em momentos de distração e alegria.

E, a todos que apoiaram e de alguma forma contribuíram para a realização deste trabalho.

OBRIGADA!!!

“Tudo posso Naquele que me fortalece” (F1 4, 13).

RESUMO

Florestas de eucalipto no Brasil estão entre os ecossistemas mais produtivos do mundo. A madeira do eucalipto é principalmente utilizada pelas empresas de papel e celulose. Contudo, fatores abióticos e bióticos afetam a expansão dessas florestas. A disponibilidade de água é uma das maiores limitações para o desenvolvimento das espécies, afetando a produção de biomassa. O hormônio ácido abscísico (ABA) desencadeia respostas de resistência a estresses abióticos, em particular a seca. O objetivo deste trabalho foi identificar *Single Nucleotide Polymorphisms* (SNPs) no gene NCED (nove-cis-epoxicarotenóide dioxigenase), da rota biossintética do ácido abscísico (ABA), em *Eucalyptus*. Através da utilização da sequência de aminoácido da enzima de *Arabidopsis* (AtNCED3) e do banco de dados de *Expressed Sequence Tags* (ESTs) de *Eucalyptus* foi possível o desenvolvimento de três oligonucleotídeos, os quais amplificaram a região desejada do gene. Então, *primers* específicos foram desenhados e os produtos de amplificação de diferentes indivíduos submetidos ao sequenciamento direto. O alinhamento e as análises das sequências revelaram a ocorrência de sete SNPs no gene NCED, em uma região de 1230 pares de bases. A razão transição/transversão foi 1.33. Após a predição da proteína, no *site Softberry* e *Expasy*, foram verificados cinco SNPs presentes na região codificadora, os quais geraram mutações sinônimas. Através do *software* DnaSP, sete haplótipos foram encontrados na amostra de 65 indivíduos gerando 15 genótipos, distribuídos entre as espécies *E. grandis*, *E. urophylla* e no híbrido Urograndis. Para os sete sítios polimórficos foram desenhados conjuntos de *primers* específicos que permitirão a genotipagem em larga escala para estudos de genética de população e em programas de melhoramento assistido.

Palavras-chave: *Eucalyptus*, ácido abscísico, SNPs, haplótipos, NCED.

ABSTRACT

Eucalyptus plantations in Brazil are among the most productive ecosystems in the world. The eucalyptus wood has its principal use in the paper and cellulose industry. However, abiotic and biotic factors affect the expansion of forest plantation. Water availability is the major limitation to development of species affecting biomass production. The plant hormone abscisic acid (ABA) triggers resistance responses to abiotic stresses, in particular to drought. The objective of this study was to identify Single Nucleotide Polymorphism (SNPs) in nine-cis-epoxy-carotenoid dioxygenase (NCED), enzyme of ABA biosynthetic pathway, in *Eucalyptus*. Through the utilization of amino acid sequence of this enzyme of *Arabidopsis* (AtNCED3) and using availability Expressed Sequence Tags (ESTs) database of *Eucalyptus*, as possible the development of three PCR primers that amplified the desirable regions of the gene. Thus, specific primers were designed and amplification products of different individuals submitted to direct sequencing. The alignment and analysis of sequences revealed the occurrence of seven SNPs in NCED gene, in a region of 1230 base pairs. The rate of transition/transversion was 1.33. After the prediction of protein, by the sites Softberry and Expasy, it was verified five SNPs, in coding region, that generated synonymous substitutions. Using the DnaSP program, seven haplotypes were found in a sample of 65 individuals, consisting of the species *E. grandis*, *E. urophylla* and the hybrid Urograndis. For these seven polymorphic sites were designed SNPs markers sets that will allow large-scale genotyping for population genetic studies and assisted breeding programs.

Key words: *Eucalyptus*, abscisic acid, SNPs, haplotypes, NCED.

LISTA DE FIGURAS

- FIGURA 1:** ROTA BIOSINTÉTICA DO ABA, INICIANDO COM O SUBSTRATO ZEAXANTINA, TRANSFORMADO PELA ZEAXANTINA EPOXIDASE (ZEP) EM, ANTERAXANTINA E VIOLAXANTINA. A NEOXANTINA SINTASE (NSY) JUNTO A UMA POSSÍVEL ISOMERASE, FORMA 9'-CIS-NEOXANTINA. UMA PROVÁVEL ISOMERASE CONVERTE A VIOLAXANTINA EM 9'-CIS-VIOLAXANTINA. ESSES DOIS 9'-CIS-ISÔMEROS SÃO TRANSFORMADOS EM XANTOXAL PELA 9-CIS-EPOXICAROTENÓIDE (NCED) E, UMA ÁLCOOL DEIDROGENASE (ABA2) CONVERTE ESSA SUBSTÂNCIA EM ALDEÍDO ABCÍSICO. ENFIM, A ENZIMA ABCÍSICO ALDEÍDO (AAO3), JUNTO AO COFATOR DE MOLIBDÊNIO, PRODUZEM O ABA. FIGURA EDITADA DE NAMBARA E MARION-POLL, 2005. **24**
- FIGURA 2:** ALINHAMENTO NO CHROMAS PRO DE TRÊS SEQUENCIAS DE DNA DE UM GENE, REFERENTES A ORGANISMOS DE UMA MESMA ESPÉCIE, NO QUAL ESTÃO MARCADOS NO RETÂNGULO EM VERMELHO O POLIMORFISMO (G-A). AS SEQUENCIAS UM E DOIS SÃO HOMIZIGOTAS NO SÍTIO EM DESTAQUE, APRESENTANDO, RESPECTIVAMENTE, GUANINA E ADENINA, ENQUANTO A SEQUENCIA TRÊS É HETEROZIGOTA. **27**
- FIGURA 3:** ORGANOGRAMA COM AS PRINCIPAIS ETAPAS DESENVOLVIDAS AO LONGO DO TRABALHO, NO QUAL ESTÃO ENVOLVIDAS A DEFINIÇÃO DA REGIÃO ALVO, ALINHAMENTO DE SEQUENCIAS DE ESTs PARA DESENHOS DE *PRIMERS LOCUS* ESPECÍFICOS. EM SEGUIDA, OS *AMPLICONS* FORAM PURIFICADOS COM EXOSAP E SEQUENCIADOS. ATRAVÉS DO ALINHAMENTO DAS SEQUENCIAS OBTIDAS FORAM IDENTIFICADOS OS SNPs E OS DADOS ANALISADOS. COM A DEFINIÇÃO DESSES SÍTIOS POLIMÓRFICOS, FORAM DESENHADOS *PRIMERS* DE EXTENSÃO ÚNICA, UTILIZADOS NA REAÇÃO DE SNAPSHOT. A REAÇÃO FOI PURIFICADA COM A ENZIMA SAP E SUBMETIDA À ELETROFORESE DE CAPILARIDADE, PARA ANÁLISE DOS DADOS. **32**
- FIGURA 4:** REPRESENTAÇÃO DO PADRÃO DE QUALIDADE DA SEQUENCIA DE EST (EFFOSTH02AB082G07.G) ENCONTRADA NO BANCO DE DADOS DE *EUCALYPTUS* FOREST-FAPESP. REGIÕES NA SEQUENCIA REPRESENTADAS NA COR PRÓXIMA A TONALIDADE VERMELHA INDICA QUALIDADE BAIXA, ENQUANTO NA VERDE CORRESPONDE AO PADRÃO MAIS ALTO. NA FIGURA ESTÃO DETERMINADAS AS POSIÇÕES INICIAIS E FINAIS, COM BAIXA QUALIDADE, PARA AS TERMINAÇÕES 5' E 3'. **41**

FIGURA 5: ELETROFORESE EM GEL DE AGAROSE 1% COM PRODUTOS DE AMPLIFICAÇÃO DOS *PRIMERS* A_2; A_2.1 E A_2.2. INDIVÍDUOS: V1, V2, V3, V4. B: BRANCO. L: PADRÃO MOLECULAR. 43

FIGURA 6: ESQUEMA ILUSTRANDO A DISTRIBUIÇÃO DOS SNPS NA SEQUENCIA *EUNCED* (BARRA VERDE). OS POLIMORFISMOS ESTÃO NAS REGIÕES AMPLIFICADAS PELOS *PRIMERS* A_2 (SETA AZUL) E A_2.2 (SETA LARANJA). NA REGIÃO AMPLIFICADA PELO *PRIMER* A_2.1 (SETA ROXA) NÃO FOI ENCONTRADO SNPS. 47

FIGURA 7: ALINHAMENTO NO CLUSTALX ENTRE A SEQUENCIA DA PROTEÍNA PREDITA DE *EUNCED* E A SEQUENCIA DA ENZIMA DE *NCED* EM *ARABIDOPSIS THALIANA*. 49

FIGURA 8: SNPS DA SEQUENCIA *EUNCED*. ATRAVÉS DA PREDIÇÃO DA PROTEINA FOI POSSÍVEL ANALISAR A VARIAÇÃO DO CÓDON E A MUTAÇÃO DO AMINOÁCIDO. EM AMARELO ESTÃO MARCADAS AS POSIÇÕES DOS SNPS NO CÓDON. 51

LISTA DE TABELAS

TABELA 1: ENZIMA 9-CIS-EPOXICAROTENOIDE DIOXIGENASE SELECIONADA ATRAVÉS DO BANCO DE DADOS GENBANK, ESTÃO APRESENTADOS SIGLA, <i>E.C. NUMBER</i> , NÚMERO DE ACESSO NO GENBANK, TAMANHO E ORGANISMO ONDE FOI ENCONTRADA.	39
TABELA 2: <i>PRIMERS</i> DA ENZIMA NCED QUE AMPLIFICARAM BANDA ÚNICA E SEUS RESPECTIVOS TAMANHOS ESPERADOS (PB), TEMPERATURAS DE PAREAMENTO (°C) E SEQUENCIAS.	42
TABELA 3: RESULTADO DO BLASTX NO GENBANK DA SEQUENCIA TOTAL OBTIDA PARA NCED EM EUCALIPTO, ESTÃO APRESENTADOS O ACESSO, NOME DA ENZIMA, ORGANISMO ENCONTRADO, SCORE, E-VALUE E MÁXIMA IDENTIDADE (%).	44
TABELA 4: RESULTADO DO BLASTX NO GENBANK DA SEQUENCIA <i>EUNCED</i> EM EUCALIPTO, ESTÃO APRESENTADOS O ACESSO, NOME DA ENZIMA, ORGANISMO ENCONTRADO, SCORE, E-VALUE E MÁXIMA IDENTIDADE (%).	45
TABELA 5: DISTRIBUIÇÃO DOS SÍTIOS DE POLIMORFISMOS ENTRE OS INDIVÍDUOS. LEGENDA: SNPs: SNP1; SNP2; SNP3; SNP4; SNP5; SNP6; SNP7. INDIVÍDUOS: <i>E. GRANDIS</i> – GRANDIS; <i>E. UROPHYLLA</i> – UROPHYLLA; E UROGRANDIS – UROGRANDIS E CLONE.	52
TABELA 6: HAPLÓTIPOS ENCONTRADOS NA AMOSTRA, ONDE ESTÃO OS NOMES DADOS, SEQUENCIAS, DISTRIBUIÇÃO NA POPULAÇÃO, PORCENTAGEM NA AMOSTRA E ESPÉCIES NAS QUAIS FORAM ENCONTRADOS OS HAPLÓTIPOS.	54
TABELA 7: GENÓTIPOS ENCONTRADOS NA AMOSTRA, ONDE ESTÃO OS NOMES DADOS, DISTRIBUIÇÃO NA POPULAÇÃO, PORCENTAGEM NA AMOSTRA E ESPÉCIES NAS QUAIS FORAM ENCONTRADOS OS GENÓTIPOS.	55
TABELA 8: <i>PRIMERS</i> UTILIZADOS NA REAÇÃO DE SNAPSHOT, COM TAMANHO (PB), O POLIMORFISMO E AS CONCENTRAÇÕES (µM) TESTADAS.	58

SUMÁRIO

RESUMO

ABSTRACT

LISTA DE FIGURAS

LISTA DE TABELAS

1. INTRODUÇÃO.....	14
2. REVISÃO DA LITERATURA	18
2.1. O gênero <i>Eucalyptus</i>	18
2.2. Estresses abióticos	20
2.3. Ácido abscísico (ABA)	22
2.4. Nove-cis-epoxicarotenóide dioxigenase (NCED)	25
2.5. <i>Single Nucleotide Polymorphism</i> (SNP)	26
3. OBJETIVO	30
3.1. Objetivo geral	30
3.2. Objetivos específicos.....	30
4. MATERIAL E MÉTODOS	31
4.1. Material utilizado.....	31
4.2. Delineamento experimental	31
4.3. Extração do DNA	33
4.4. Seleção dos oligonucleotídeos iniciadores	34
4.5. Validação dos oligonucleotídeos e condições de amplificação.....	34
4.6. Purificação dos produtos de PCR.....	35
4.7. Reação de sequenciamento.....	35

4.8.	Alinhamento das sequencias	36
4.9.	Análise <i>in silico</i> da sequencia.....	36
4.10.	Análises dos SNPs	36
4.11.	Seleção dos oligonucleotídeos flanqueadores	37
4.12.	Reação e purificação do minisequenciamento.....	37
4.13.	Análise da estrutura genética populacional	38
5.	RESULTADOS E DISCUSSÃO	39
5.1.	Busca de Homólogos de NCED	39
5.2.	Desenho dos oligonucleotídeos iniciadores	41
5.3.	Validação dos oligonucleotídeos iniciadores.....	43
5.4.	Identificação dos SNPs	46
5.5.	Genotipagem de SNPs.....	48
5.6.	Classificação dos SNPs	48
5.7.	Análise populacional	52
5.8.	Testes de extensão dos oligonucleotídeos adjacentes ao polimorfismo...	57
6.	CONSIDERAÇÕES FINAIS	59
7.	CONCLUSÕES	60
8.	REFERÊNCIAS	61

1. INTRODUÇÃO

O gênero *Eucalyptus* é o mais produtivo no Brasil e constitui uma das melhores alternativas para atender às diversas demandas da sociedade referente ao consumo de produtos de base florestal (Stape et al., 2004). Diante dessa necessidade e aumento de exploração, as áreas de produção estão em expansão, com o estabelecimento em regiões de reflorestamento no cerrado, onde os solos apresentam baixa fertilidade e períodos longos de déficit hídrico.

A deficiência hídrica é um dos principais fatores de estresse ambiental, que restringe as latitudes e os solos de produção das espécies, além de influenciar negativamente o crescimento e o desenvolvimento das plantas (Lecoeur e Sinclair, 1996; Santos e Carlesso, 1998).

O estresse hídrico pode afetar o crescimento e desenvolvimento de espécies florestais em qualquer fase de seu ciclo. A planta retira a água do solo quando o potencial hídrico de suas raízes for mais negativo que aquele da solução do solo. O gradiente do potencial hídrico do contínuo solo-planta-atmosfera instituiu-se na força motriz para o transporte de água através da planta (Reis e Reis, 1997).

Quando em situações de déficit hídrico, as plantas elaboram um conjunto de respostas fisiológicas, metabólicas e de desenvolvimento, após a percepção do sinal externo, via cascata de eventos moleculares (Bray, 1993). Deste modo, a elucidação da ativação/desativação desses eventos e da interação entre eles é essencial para o desenvolvimento de novas variedades tolerantes à seca (Nepomuceno et al., 2001).

Os hormônios são mensageiros químicos, responsáveis por efeitos marcantes no desenvolvimento em pequenas concentrações, os quais funcionam

como mediadores da comunicação intercelular ao interagirem com proteínas específicas denominadas receptores (Taiz e Zeiger, 2004).

O hormônio ácido abscísico (ABA) das plantas possui um importante papel em muitos aspectos de respostas ao estresse hídrico, além de estar envolvido em outros processos fisiológicos. Este é um hormônio vegetal encontrado em todas as plantas vasculares, bem como nos musgos, mas parece estar ausente nas hepáticas (Taiz e Zeiger, 2004).

A principal função do ABA na resposta ao estresse hídrico parece estar relacionada a regulação do contrapeso de água na planta e a tolerância ao estresse osmótico. As plantas mutantes deficientes na produção do ABA, que apresentam fenótipos anormais, são importantes fontes de estudo no esclarecimento das etapas específicas da via do metabolismo e catabolismo do ácido.

O conhecimento sobre as funções e a biossíntese do ABA foi possível pela identificação e caracterização desses mutantes, que exibiam os seguintes fenótipos: germinação precoce, suscetibilidade a murchar, aumento da condutância dos estômatos e habilidade de germinar ou crescer em condições de alta concentração solutos (Schwartz et al., 2003). Estes mutantes também foram importantes na clonagem de genes que codificam as enzimas da biossíntese do ABA.

O conhecimento sobre as bases fisiológicas, moleculares e genéticas das respostas das plantas ao déficit hídrico ainda é escasso. Há necessidade de um enfoque sistemático com destaque nas diferenças entre os genótipos em relação à tolerância ao déficit hídrico, pois estas devem ser qualitativas e quantitativas na determinação da expressão gênica (Guerrini et al., 2005).

Marcadores moleculares tem sido uma ferramenta poderosa para estimar variação genética dentro e entre as populações de plantas. Até o momento, várias técnicas moleculares estão disponíveis para estimar a diversidade genética em plantas (izoenzimas, RAPD – *Random Amplification of Polymorphic DNA*, RFLP –

Restriction Fragment Length Polymorphism, microssatélites entre outras). Contudo, as limitações dessas técnicas incluem baixo número de *loci* polimórfico, necessidade de grande quantidade de DNA, baixa reprodução e intenso trabalho (Osman et al., 2003).

Polimorfismos de base única ou SNPs representam uma fonte abundante de variação genética. SNPs são variações na sequência do DNA, quando um único nucleotídeo difere no genoma entre os indivíduos de uma espécie ou entre pares de cromossomos, podendo ser polimorfismos bi-, tri- ou tetra-alélicos. Estas são as variações genéticas mais frequentes nos genomas de vertebrados e invertebrados; são codominantes e de fácil detecção quando comparados, por exemplo, aos microssatélites (Taillon-Miller et al., 1998).

Embora a natureza bialélica dos SNPs faça deles menos informativos por *locus* examinado quando comparados a outros marcadores (Jehan e Lakhanpaul, 2006), sua grande vantagem reside na abundância de polimorfismos entre alelos de um determinado gene e ampla distribuição, podendo estar presentes em, praticamente, todos os *loci* gênicos.

Outra utilidade dos SNPs está relacionada às análises de ligação, apresentando como vantagens que estas podem ser executadas no estágio precoce de crescimento das plantas, necessitando de pouca quantidade de DNA, além de menos tempo e trabalho, porque a maioria dos tipos de genotipagem não envolve eletroforese (Jehan e Lakhanpaul, 2006).

Há dois métodos principais para detectar estes polimorfismos: *in silico* e experimentalmente. O método *in silico* envolve o uso de ferramentas de bioinformática como os bancos de dados de *Expressed Sequence Tags* (ESTs). O projeto de ESTs de *Eucalyptus* FORESTs-FAPESP (<https://forest.esalq.usp.br>), por exemplo, apresenta mais de 123.889 *reads* agrupados em 33.080 *clusters*, adquiridos de bibliotecas cDNA de diferentes tecidos da planta, submetidos à

diversas condições fisiológicas. Enquanto o projeto *Genolyptus* gerou um banco de dados com mais de 120.000 leituras de ESTs, visando, principalmente, identificar áreas do genoma relacionadas, sobretudo, à qualidade da madeira e resistência a praga. Esses dados de acesso a informações genômicas da espécie são um excelente recurso para o processo de procura e identificação de marcadores SNPs em genes candidatos. 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 pelo projeto *Genolyptus* para detecção de SNPs.

Apesar da indiscutível utilidade da bioinformática na identificação de SNPs, a necessidade de validação experimental dos dados é indispensável para a avaliação do potencial informativo de cada polimorfismo. Para tanto existem diversos métodos de identificação e genotipagem de SNPs (Vignal et al., 2002).

Diante do exposto, o objetivo desse trabalho foi identificar SNPs no gene que codifica a enzima NCED, proteína envolvida na rota biossintética do ácido abscísico, em *Eucalyptus*.

2. REVISÃO DA LITERATURA

2.1. O gênero *Eucalyptus*

O gênero *Eucalyptus* pertence à divisão *Angiospermae*, classe *Dicotyledonea*, ordem *Myrtales*, família *Myrtaceae* e apresenta cerca de 600 espécies (Fonseca et al., 2010). O principal subgênero, *Symphomyrtus*, possui mais de 300 espécies, dentre as quais estão as de importância comercial - *E. grandis*, *E. globulus*, *E. urophylla*, *E. calmadulensis*, *E. saligna* e *E. tereticornis* (Eldridge, 1994; Mora e Garcia, 2000).

As espécies de *Eucalyptus* são, principalmente, originárias da Austrália, exceto *E. urophylla* e *E. deglupta*, as quais são naturais do Timor e Papua Nova Guiné (Pryor, 1985). O gênero *Eucalyptus*, por ser originário de um país com dimensões continentais, apresenta grande diversidade de clima e solo. E, em função dessa variação ambiental e do isolamento geográfico, a grande diversidade genética culminou em separação de inúmeras espécies (Mayr, 1977).

A necessidade de adaptação aos diferentes nichos ecológicos existentes na Austrália gerou subpopulações dentro de uma mesma espécie, também denominadas procedências. Por esse motivo, são encontradas procedências com diferenças em diversas características de importância econômica, tais como: taxa de crescimento; formato do tronco; resistência a geadas, secas, insetos, fungos; tolerância à salinidade, alcalinidade, acidez e ao alagamento do solo (Eldridge, 1994). Desta maneira, para introduzir o eucalipto em uma nova região, a avaliação de espécies e procedências é fundamental para se iniciar um programa de melhoramento.

Dentre as espécies de *Eucalyptus*, há arbustivas e arbóreas, com algumas espécies alcançando até 100 metros de altura, todas possuem troncos lenhosos. Em geral, são espécies perenes, poucas perdem suas folhas durante as monções em seu hábitat. Suas flores são agrupadas em corimbos, panículas ou, mais frequentemente, em umbelas. Elas são monóicas, hermafroditas e protândricas. A planta é, preferencialmente, alógama (fecundação cruzada), com alguma autogamia (na ordem de 10%), sendo que os principais polinizadores são os insetos (a polinização anemófila não ocorre). 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 um milímetro até dois centímetros (Cavalcanti, 1963; Fonseca et al., 2010).

No Brasil, as espécies de eucaliptos foram introduzidas por Edmundo Navarro de Andrade para ser usado pela Companhia Railroad Paulista no estado de São Paulo (hoje conhecida como Ferrovia Paulista S.A.), no início do século XIX (Andrade, 1939).

O gênero se adaptou muito bem às condições de clima e de solo do maior país da América do Sul, sendo plantado atualmente em várias regiões brasileiras. Em termos climáticos para o cultivo do eucalipto, o Brasil possui duas regiões - tropical e subtropical. A região sudeste, predominantemente tropical, não sujeita à geada de forte intensidade, concentra a maior área de plantio. Para áreas de climas predominantemente tropicais, as espécies mais indicadas são *E. grandis*, *E. urophylla*, *E. saligna*, e *E. cloeziana*. As espécies indicadas para a região subtropical são *E. benthamii* (comprovadamente resistente às geadas) e *E. dunnii* (resistência parcial às geadas) (Embrapa Florestas, 2003).

As plantações de eucalipto têm alcançado um grande avanço no Brasil, tornando-se de grande importância econômica para o país. O gênero possui uma ampla e diversificada utilização na indústria, sendo a matéria prima utilizada na

produção de celulose, papel, chapas de fibra, aglomerados, madeira serrada, casas, estrutura e móveis, além da utilização em outros benefícios, tais como na proteção de solos contra erosões, na capacidade de captação de CO₂, na geração de energia (lenha e carvão), mel (pólen), cápsulas de remédios, óleos essenciais, entre outros.

Apesar da sua crescente importância no cenário florestal mundial, estudos genômicos envolvendo espécies do gênero *Eucalyptus* ainda são incipientes. Eucaliptos são plantas diplóides, possuem $2n = 22$ cromossomos (Eldridge, 1994) e tamanho do genoma variando desde cerca de 380 milhões de pares de bases (Mpb) em *C. citriodora* (anteriormente classificada como *E. citriodora*) até por volta de 650 Mpb em *E. grandis* (Grattapaglia e Bradshaw, 1994) e 644 Mpb em *E. globulus* (Pinto et al., 2004). No entanto, pouco se sabe sobre informações tais como: número de genes, características de éxons e íntrons, regiões promotoras, tamanho médio de genes, conteúdo GC, dentre outras.

2.2. Estresses abióticos

O estresse é considerado um desvio significativo das condições ótimas para a vida, o qual induz mudanças e respostas em todos os níveis funcionais do organismo, podendo ser estas reversíveis a princípio, mas com chances de se tornarem permanentes (Larcher, 2004).

As plantas estão expostas a vários tipos de estresses ambientais que afetam de maneira negativa seu crescimento, metabolismo e produtividade. Os fatores de estresses mais importantes que limitam a produtividade das culturas são: seca; salinidade; inundação; poluentes; radiação; baixas e altas temperaturas (Lawlor, 2002).

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 à morte) (Cambraia, 2005).

O estresse hídrico resulta da deficiência hídrica ou do excesso de água no ambiente, sendo o primeiro mais comum. Todos os processos vitais são afetados de alguma forma pelo declínio do potencial hídrico (Taylor, 1996; Larcher, 2004), interferindo diretamente na eficiência do uso de água pelas plantas, as quais necessitam então desenvolver um potencial hídrico mais negativo que o do solo, de modo a manter o gradiente entre o solo e a planta (Taiz e Zeiger, 2004).

O alto grau de interação entre genótipos e ambiente dificulta a seleção de fenótipos que refletem as diferenças genotípicas (Leung, 2008). Genótipos que diferem em tolerância ao déficit hídrico devem apresentar diferenças quantitativas e qualitativas na expressão gênica quando submetidos a períodos de seca. Essas considerações reforçam a necessidade de um enfoque sistemático nos estudos de estresse hídrico, com ênfase nas diferenças entre os genótipos (Casagrande et al., 2001; Nepomuceno et al., 2001).

A percepção do estresse pela raiz induz a produção de compostos químicos que serão transportados para a parte aérea onde se tornam ativadores de resposta. O conhecimento da transdução do sinal requer a coordenação espacial e temporal apropriada de todas as moléculas sinalizadoras envolvidas. Segundo Jia et al., (2002), a sinalização celular de deficiência hídrica envolve diversas moléculas e íons, em especial, o hormônio ácido abscísico (ABA), com duplo papel em regulação fisiológica.

O déficit hídrico, em plantas, inicia um complexo de respostas, começando com a percepção do estresse, o qual desencadeia uma sequência de

eventos moleculares que é finalizada em vários níveis de respostas fisiológicas, metabólicas e de desenvolvimento (Bray, 1993).

2.3. Ácido abscísico (ABA)

O ácido abscísico (ABA) foi descoberto em 1960, através de cromatografia em papel para separar os extratos vegetais e bioensaios, esses experimentos levaram ao conhecimento do inibidor de crescimento *dormina* em folhas de falso plátano, quando as árvores estavam entrando em dormência no início do outono. Essa substância era quimicamente idêntica àquela descoberta em frutos de algodoeiro, *abscissina II*, que promovia abscisão dos mesmos. Pela identidade, tais compostos foram renomeados como ácido abscísico (ABA), sendo o nome não muito apropriado, porque abscisão de frutas e folhas não é a função principal do ABA (etileno é o hormônio que desencadeia abscisão e o ABA induz a abscisão em algodoeiro pela capacidade de estimular a produção de etileno) (Taiz e Zeiger, 2004).

A biossíntese desse hormônio ocorre em quase todas as células que possuem cloroplastos ou amiloplastos. A melhor elucidação dos detalhes dessa via tem sido obtida pela utilização de mutantes que exibem falhas em etapas específicas (Taiz e Zeiger, 2004).

A via inicia-se com o isopentinildifosfato e leva a síntese da xantofila (violaxantina), catalisada pela enzima zeaxantina epoxidase (ZEP), via anteraxantina (Figura 1). A partir disso, há formação de dois *cis*-isômeros de violaxantina e neoxantina (*cis*-xantofilas), as quais necessitam das enzimas neoxantina sintase (NSY) e a isomerase. As *cis*-xantofilas são catalisadas pela família de 9-*cis*-epoxicarotenoide dioxigenase (NCED), para formar o composto xantoxal. Por último, o xantoxal é convertido a ABA por meio de etapas oxidativas. A enzima álcool desidrogenase (ABA2) converte xantoxal em aldeído abscísico, o qual é oxidado em

ABA pela enzima abscísico aldeído oxidase (AAO3), sendo necessária a participação do cofator de molibdênio nessa etapa final (Taiz e Zeiger, 2004; Nambara e Marion-Poll, 2005).

O catabolismo do ABA é categorizado em dois tipos de reações, oxidação e conjugação. A maior causa da inativação do ABA livre é a oxidação, produzindo ácido faséico (PA) e ácido diidrofaséico (DPA). Recentemente, foi identificada a enzima 8'-hidroxilase (gene CYP707A), responsável pela conversão do ABA em ácido faséico (PA) *in vitro* (Taiz e Zeiger, 2004; Nambara e Marion-Poll, 2005).

As funções mais documentadas do ABA estão nos processos de maturação de semente, aquisição de tolerância a dessecação e dormência. Durante o crescimento vegetativo, ABA é o principal hormônio que confere tolerância aos estresses ambientais, notavelmente, à seca e à alta salinidade, permitindo, então, plantas colonizarem nichos ecológicos onde a disponibilidade de água é limitada ou esporádica (Wasilewska et al., 2008).

Quando em situações de déficit hídrico, ocorre alcalinização do xilema, o ABA é transportado e absorvido pelas células-guarda induzindo o fechamento estomático (Davies et al., 2002; Becker et al., 2003). Os movimentos estomáticos são induzidos por mudanças na turgescência das células-guarda, que são mediados por seus conteúdos de íons e solutos orgânicos (Israelsson et al., 2006). O controle estomático impede que a planta perca altos índices de água, mantendo seu turgor e contrabalanceando a inibição de seu crescimento.

O acúmulo de ABA também é requerido para a manutenção da taxa de alongamento da raiz em condições de baixo potencial de água no solo. Segundo Sharp e Lenoble (2002), a interferência do ABA no crescimento da raiz e da parte aérea é indireta, através do seu efeito inibitório na síntese do etileno. Desta forma, o crescimento da parte aérea é interrompido devido à ação do etileno, uma vez que a concentração do ABA nessa área é insuficiente para impedir a síntese do etileno.

Enquanto nas raízes, a alta concentração de ABA previne a inibição do crescimento mediada pelo etileno.

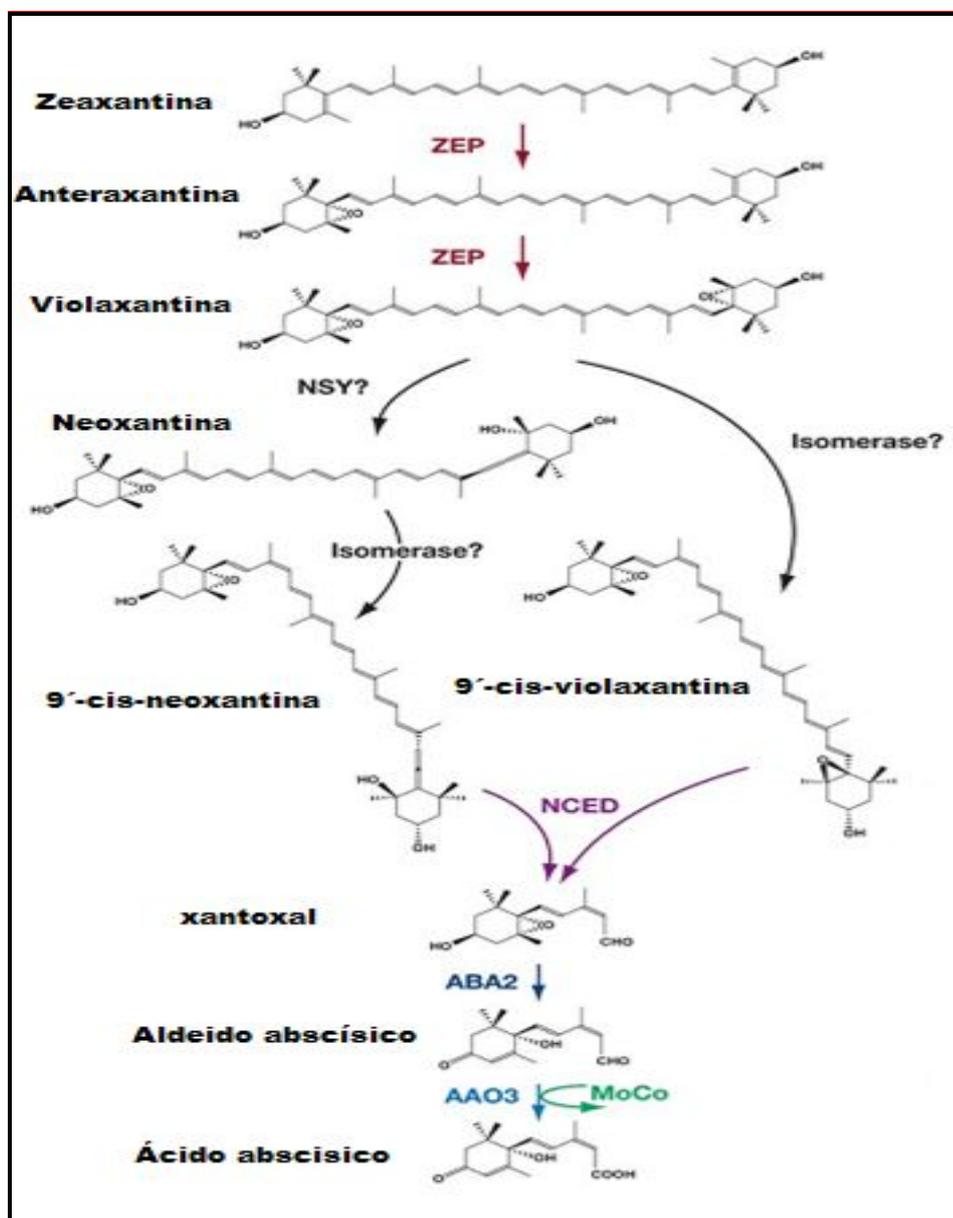


Figura 1: Rota biossintética do ABA, iniciando com o substrato zeaxantina, transformado pela zeaxantina epoxidase (ZEP) em anteraxantina e violaxantina. A neoxantina sintase (NSY), junto a uma possível isomerase, forma 9'-cis-neoxantina. Uma provável isomerase converte a violaxantina em 9'-cis-violaxantina. Esses dois 9'-cis-isômeros são transformados em xantoxal pela 9-cis-epoxicarotenóide dioxigenase (NCED) e, uma álcool desidrogenase (ABA2) converte essa substância em aldeído abscísico. Enfim, a enzima abscísico aldeído (AAO3), junto ao cofator de molibdênio, produzem o ABA. Figura editada de Nambara e Marion-Poll, 2005.

2.4. Nove-cis-epoxicarotenóide dioxigenase (NCED)

O ácido abscísico é um composto apocarotenóide gerado pela clivagem oxidativa dos carotenóides, esta reação ocorre, na maioria das vezes, com a atividade de enzimas da família dioxigenase que clivam o carotenóide (Auldrige et al., 2006, Tan et al., 2003).

Evidências bioquímicas (Zeevaart e Creelman, 1988) e genéticas (Koorneef et al., 1998) indicaram que a clivagem de 9-cis-xantofilas, realizada pelas dioxigenases, é provavelmente a principal etapa na rota biossintética do ABA.

A confirmação dessa natureza regulatória da reação de clivagem foi comprovada pela caracterização do mutante *viviparous* de milho (*Zea mays*), *vp14*, o qual catalisa a reação de clivagem dos compostos 9-cis-epoxicarotenóide-violaxantina ou 9-cis-epoxicarotenóide-neoxantina (Tan et al., 1997). O mutante *vp14* exibiu um defeito na biossíntese do ABA, por inserção de *transposon*, esse gene foi clonado e a sequência de aminoácido derivada foi muito similar a *lignostilbene dioxygenase* (LSDs) de *Pseudomonas* (*Sphingomonas*) *paucimobilis* (Tan et al., 1997).

Após isso, sequência homóloga ao gene *vp14* foi obtida em *Lycopersicon esculentum* (Burbidge et al., 1997), denominado *LeNCED1*. E, atualmente, muitos genes NCEDs têm sido clonados e caracterizados em diferentes espécies como *Stylosanthes guianensis* (Yang e Guo, 2007), *Arabidopsis thaliana* (Iuchi et al., 2001), *Vigna unguiculata* (Iuchi et al., 2000), *Persea americana* (Chernys e Zeevaart, 2000) e *Phaseolus vulgaris* (Qin e Zeevaart, 1999).

No genoma de *Arabidopsis thaliana* existem nove proteínas hipotéticas que compartilham similaridade a sequência NCED e, através de análises filogenéticas, foi verificado que cinco dessas sequências agruparam-se com NCEDs caracterizadas em outras espécies (Schwartz et al., 2003). Em Iuchi et al., (2001), foi relatado que

AtNCED2, 3, 6 e 9 foram capazes de catalisar a reação de clivagem para síntese do ABA, enquanto que *AtNCED5*, apesar da similaridade da sequência, não foi demonstrada a reação bioquímica experimentalmente. As quatro proteínas restantes de *Arabidopsis thaliana*, mais distantes relacionadas, foram nomeadas *AtCCD* (*Carotenoid Cleavage Dioxygenase*), não apresentando envolvimento com a biossíntese do ABA (Schwartz et al., 2001).

2.5. Single Nucleotide Polymorphism (SNP)

Recentemente, os marcadores SNPs tem ganhado muito interesse na comunidade científica e de melhoramento, por ocorrerem em um número ilimitado sendo, portanto, um marcador de grande potencial (Ganal et al., 2009).

Single Nucleotide Polymorphism, ou seja, polimorfismo de base individual ou única, é uma variação nas sequências do DNA (Figura 2) que ocorre em uma posição individual e, difere no genoma entre membros da mesma espécie ou entre pares de cromossomos de um indivíduo, com frequência de pelo menos 1% (Brookes, 1999).

A maioria dos polimorfismos observados ocorre na forma bialélica, pois a probabilidade de alteração de duas bases independentes, presentes em uma mesma posição, é pequena. No entanto, polimorfismos do tipo inserções, deleções, tri- ou tetra-alélicos são também denominados SNPs, quando encontrados em pelo menos 1% da população (Liao e Lee, 2010).

Esses polimorfismos são gerados por substituição de única base ou pequenos eventos de deleção/inserção ocorrendo a uma taxa de aproximadamente 10^{-7} alterações por sítio por gerações, ou menor. As substituições de única base têm sido bem caracterizadas desde o advento da tecnologia do sequenciamento em 1977 e, indiretamente, os SNPs e *indels* têm sido as bases para marcadores genéticos,

como por exemplo, RFLPs (*Restriction Fragment Length Polymorphisms*), AFLPs (*Amplified Fragment Length Polymorphisms*) e RAPD (*Random Amplification of Polymorphic DNA*), entre outros (Cordeiro et al., 2006).

Há muitas técnicas de identificação de SNPs que são usadas para encontrar tal variação nos organismos, baseada em: bancos de dados de sequencias de ESTs; análises de *array*; resequenciamento de *amplicons*; genomas resequenciados, tecnologia de sequenciamento *next generation* (Vignal et al., 2002). A escolha da técnica de identificação é dependente da aplicabilidade, necessidades e algumas limitações dos métodos (Ganal et al., 2009).

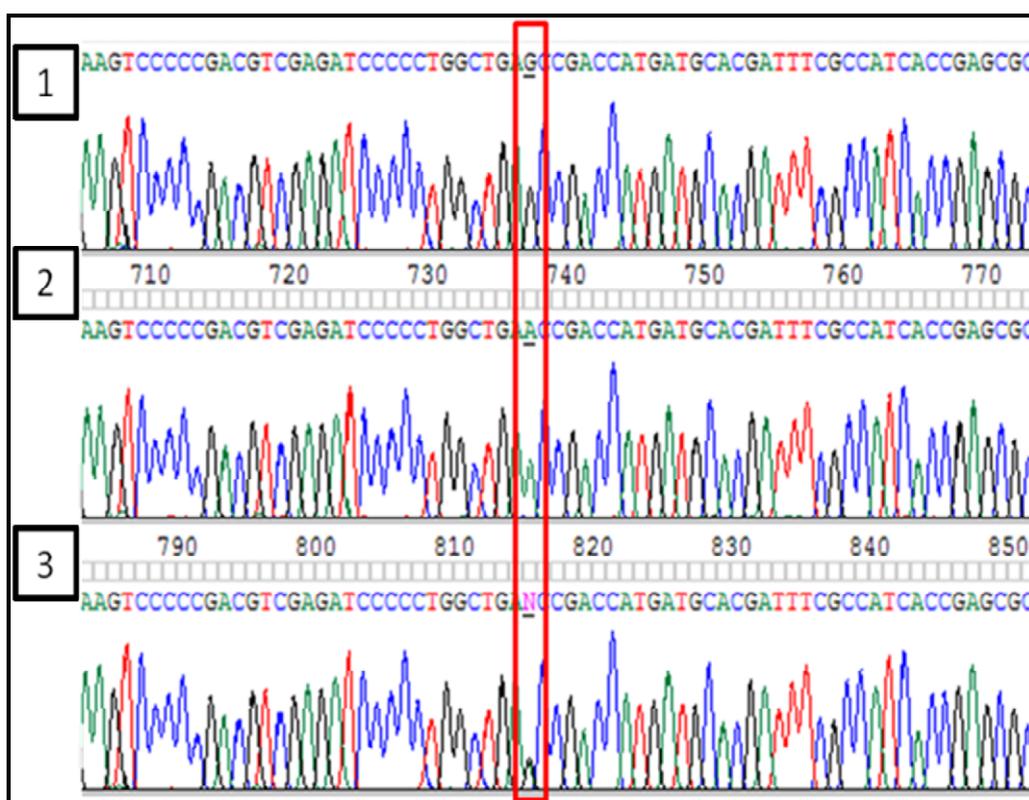


Figura 2: Alinhamento no Chromas Pro de três sequencias de DNA de um mesmo gene, referentes a organismos de uma mesma espécie, no qual estão marcados no retângulo em vermelho o polimorfismo (G-A). As sequencias um e dois são homocigotas no sítio em destaque, apresentando, respectivamente, guanina e adenina, enquanto a sequencia três é heterocigota.

Nos últimos anos, o interesse nas mutações de ponto tem aumentado como resultado da sua numerosa aplicação na genética médica, evolucionária e *forense*.

Grandes avanços têm sido realizados com o uso de SNPs no estudo de doenças em humanos, além da utilização dessa variação na farmacogenética, para melhores resultados nas respostas a drogas (Cordeiro et al., 2006). Os indivíduos podem, por exemplo, ser distinguidos por um conjunto de SNPs, sendo assim útil, na separação de diferentes origens de amostras heterogêneas (plasma de mulher grávida e pacientes cancerígenos), na capacitação de diagnóstico pré-natal não invasivo, diagnóstico de câncer e monitoramento da doença (Ding, 2007).

Os polimorfismos de única base são menos mutáveis, o que faz deles marcadores estáveis, sendo excelentes para estudo de características genéticas complexas e para entendimento da evolução genômica. Logo, são marcadores adequados e mais fáceis de suceder em estudos populacionais (Jehan e Lakhanpaul, 2006). Devido à capacidade de gerar haplótipos (padrão de associação de SNPs que são herdados conjuntamente), o uso dos SNPs permite a comparação estatística entre grupos de estudos ou controle.

SNPs localizados dentro da parte não recombinante do cromossomo Y e no genoma mitocondrial são particularmente valiosos por possuírem herança uniparental (paterna e materna, respectivamente), e não sofrerem recombinação, sendo, assim, herdados como blocos haplótipos. Este aspecto permite identificar como as mutações foram transferidas de geração para geração, para analisar a história evolutiva de cada *locus* gênico e deduzir padrões de migrações por estudo presentes nas linhagens (Quintáns et al., 2004).

A distribuição dos SNPs varia no genoma, sendo menos frequente em regiões codificadoras que nas não-codificadoras (Liao e Lee, 2010). Nas regiões codificantes, as mutações são categorizadas em duas classes: não sinônimas, as

quais alteram a sequência do produto proteico e sinônimas que não afetam a estrutura primária dos produtos. No entanto, um SNP sinônimo pode modificar a estrutura e a estabilidade do RNA mensageiro e, conseqüentemente, afetar a quantidade de proteína produzida (Griffiths et al., 2001). Além do mais, SNPs podem afetar o *splicing* alternativo, gerar ou suprimir códons de terminação ou poliadenilação na molécula de RNA mensageiro e alterar códons de iniciação de tradução (Liao e Lee, 2010).

A informação gerada dos SNPs poderá, posteriormente, ser utilizada em vários níveis nas análises 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).

3. OBJETIVOS

3.1. Objetivo geral

O objetivo desse trabalho foi identificar SNPs no gene que codifica a enzima NCED, proteína envolvida na rota biossintética do ácido abscísico, em *Eucalyptus*.

3.2. Objetivos específicos

- Identificar número e distribuição de haplótipos para os polimorfismos encontrados entre os indivíduos das espécies *E. urophylla*, *E. grandis* e híbrido Urograndis;
- Disponibilizar *primers* específicos para genotipagem em estudos de populações e programas de melhoramento genético assistido.

4. MATERIAL E MÉTODOS

4.1. Material utilizado

O material utilizado neste trabalho pertence ao programa de melhoramento da empresa FIBRIA Celulose, consistindo em um total de 65 indivíduos divergentes geneticamente, sendo 22 desses da espécie *Eucalyptus grandis*, 12 *E. urophylla* e 31 híbridos Urograndis.

Entre essa amostra, dez indivíduos Urograndis constituem um ensaio biológico de caracterização de eucalipto exposto a estresse hídrico, sendo que cinco indivíduos mostraram-se tolerantes e cinco suscetíveis. O ensaio foi realizado pelo Professor Doutor Rinaldo César de Paula, do Departamento de Produção Vegetal - UNESP Jaboticabal, baseado em características fisiológicas (principalmente: fotossíntese, condutância estomática, transpiração, potencial hídrico foliar, conteúdo relativo de água) e de crescimento (secundárias: área foliar, número de folhas, massa seca). As plantas foram 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).

4.2. Delineamento experimental

O organograma abaixo ilustra os procedimentos realizados ao longo do trabalho (Figura 3).

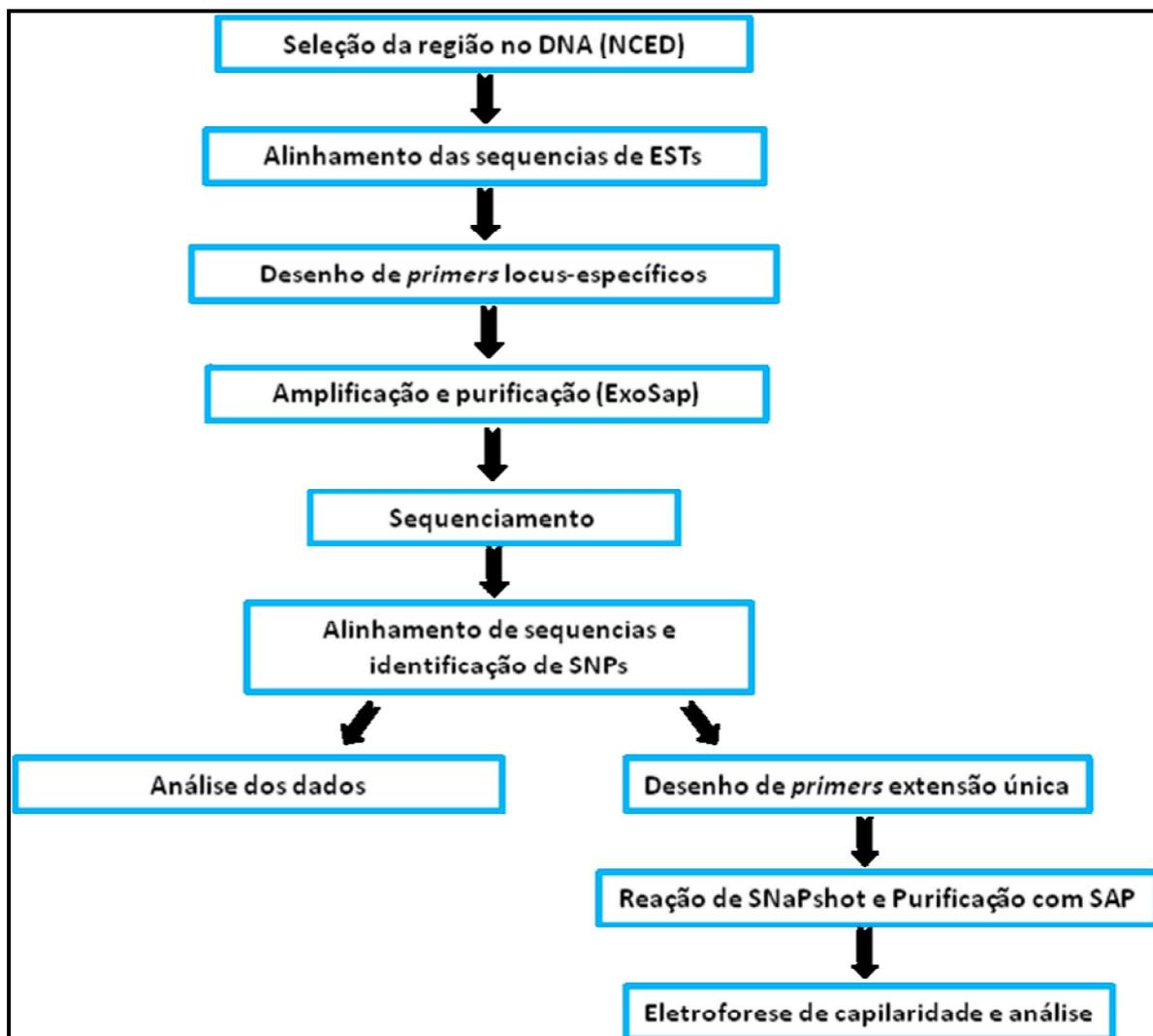


Figura 3: Organograma com as principais etapas desenvolvidas ao longo do trabalho, no qual estão, inicialmente, a definição da região alvo e alinhamento de sequências de ESTs, para desenhos de *primers locus* específicos. Em seguida, os *amplicons* foram purificados com ExoSap e sequenciados. Através do alinhamento das sequências obtidas foram identificados os SNPs e os dados analisados. Com a definição desses sítios polimórficos, foram desenhados *primers* de extensão única, utilizados na reação de SNaPshot. A reação foi purificada com a enzima SAP e submetida à eletroforese de capilaridade para, posteriormente, análise dos dados.

4.3. Extração do DNA

O DNA foi extraído utilizando o protocolo de extração descrito por Ferreira e Grattapaglia (1998).

Em um tubo foram adicionados 150 mg de folhas maceradas na presença de nitrogênio líquido, 700 µl de solução de extração com PVP e 2-mercaptoetanol (2 µl/ml de solução de extração). Esta mistura foi aquecida à temperatura de 65°C, por 45 a 60 minutos.

Para a primeira extração foram adicionados 600 µl de clorofórmio-álcool isoamílico (CIA 24:1), os tubos levados à centrífuga refrigerada a 10°C, por 5 minutos, na rotação de 12000 rpm. A fase superior foi transferida para um novo tubo, a esta solução foram adicionados 100 µl de CTAB 5% e homogeneizada levemente. A segunda extração foi realizada adicionando 600 µl de CIA 24:1 e a solução centrifugada nas mesmas condições anteriores.

O sobrenadante foi transferido para um novo tubo, no qual foram adicionados NaCl 1,5 M e centrifugado nas condições anteriores. O sobrenadante foi, então, transferido para um novo tubo, no qual foram adicionados 400 µl de álcool isopropílico gelado, a solução foi armazenada à temperatura de -20°C até o dia seguinte e, centrifugada nas mesmas condições anteriores.

O sobrenadante foi descartado, o *pellet* lavado com etanol, seco em centrífuga a vácuo e, posteriormente, suspenso em 80 µl de tampão TE contendo 10 ng/µl de RNase. Em seguida, foi incubado em banho-maria a 37°C para a digestão do RNA, por no mínimo 40 minutos ou até completa dissolução do *pellet*.

O material extraído teve seu DNA quantificado por espectrofotômetro Nano Drop® ND-1000.

4.4. Seleção dos oligonucleotídeos iniciadores

A sequência de aminoácido da enzima NCED foi obtida no banco de dados GenBank e, utilizando a ferramenta BLAST (Altschul et al., 1997), foi realizada procura por sequências similares no banco de dados de ESTs de *Eucalyptus* FORESTs-FAPESP (Vicentini et al., 2005). Os *reads* selecionados foram aqueles que possuíam maior similaridade à sequência de referência da procura inicial. Em seguida, pares de *primers* foram construídos através do *software Perl Primer* v1.1.14 (Marshall, 2004), para amplificarem região com tamanho de aproximadamente 600 pb. A melhor temperatura de reação para cada par de oligonucleotídeo foi determinada e avaliada a qualidade de amplificação.

4.5. Validação dos oligonucleotídeos e condições de amplificação

Reações de amplificação foram confeccionadas para um volume final de 10 μ l contendo: 0,3 μ M de cada *primer*, 10-20 ng DNA genômico e *PCR Master Mix*, da Promega Corporation, o qual contém 50 unidades/ml de Taq DNA polimerase suprido de um tampão para reação (pH 8,5), 400 μ M dATP, 400 μ M dGTP, 400 μ M dCTP, 400 μ M dTTP e 3 mM MgCl. As reações foram realizadas no termociclador.

Os ciclos de amplificação foram os seguintes: desnaturação inicial durante 5 minutos a 96°C, seguida de 32 ciclos compreendendo 30 segundos a 96°C, 45 segundos a temperatura de pareamento dos *primers* e 1 minuto a 72°C, com extensão final de 15 minutos a 72°C.

Os produtos de PCR (6 μ L) foram verificados utilizando gel de agarose a 1%, preparado com tampão Tris-borato (TBE 1x pH 8,0), corado com brometo de etídio e comparado ao DNA *ladder* de 250 pb (Invitrogen). Os restantes dos

produtos de PCR foram sequenciados, utilizando cada oligonucleotídeo iniciador específico *forward* e *reverse*.

4.6. Purificação dos produtos de PCR

De acordo com as especificações do produto, deve-se adicionar 2 µl da enzima *ExoSAP IT* à 5 µl do produto de PCR e, posteriormente, incubar a solução à 37°C durante 15 minutos (para ocorrer a degradação do excesso de *primers* e nucleotídeos) e à 80°C por 15 minutos (para a inativação da enzima).

Nesse trabalho foi realizado tratamento com a enzima *ExoSAP IT* adicionando 1 µl da enzima à 4 µl do produto de PCR, seguido de incubação como recomendado pelo fabricante. Esses produtos de PCR purificados foram submetidos ao sequenciamento ou minissequenciamento.

4.7. Reação de sequenciamento

As reações de sequenciamento foram realizadas utilizando cerca de 100 ng de produto de PCR purificado com a enzima *ExoSAP IT*; tampão de sequenciamento 1x; 1,8 µl de *Big Dye*; 1 µM dos oligonucleotídeos *forward* e *reverse*; a reação foi completada com água deionizada para um volume final com 10 µl.

As condições de amplificação foram as seguintes: extensão inicial de 96°C; 40 ciclos de 96°C por 10 segundos, 50°C por 10 segundos, 60°C por 4 minutos. Para cada amostra foram realizadas duas reações, sendo uma para o oligonucleotídeo *forward* e outra para o *reverse*.

Antes do sequenciamento, a reação foi purificada com 80 µL de isopropanol 75%, por 15 minutos, a placa foi centrifugada por 45 minutos a 4000

rpm, o álcool descartado e foi adicionado etanol 70% para lavagem dos *pellets*, em seguida, realizada uma nova centrifugação nas condições anteriores, por 10 minutos. O álcool foi descartado e, após os *pellets* estarem secos, adicionou-se 40 µL de formamida. Em seguida, foi realizada a desnaturação a 95°C por 5 minutos e choque térmico em gelo. O sequenciamento foi realizado em sequenciador automático 3100 *Applied Biotechnology*.

4.8. Alinhamento das sequencias

Os eletroferogramas das sequencias *forward* e *reverse*, geradas automaticamente, foram inspecionadas visualmente e os heterozigotos detectados. O alinhamento das sequencias de diversos individuos foi realizado através do *software* ChromasPro v.1.42, sendo as sequencias identificadas utilizando o BLAST e comparadas à outras disponíveis no GenBank.

4.9. Análise *in silico* da sequencia

A sequencia obtida foi traduzida no *site* *Expasy* (Gasteiger et al., 2003) e no *site* *Softberry* (www.softberry.com). A proteína predita foi alinhada no *software* ClustalX 2.0.11 com sequencia de proteína de *Arabidopsis thaliana* a fim de comparar a região sequenciada com a proteína de outro organismo.

4.10. Análises dos SNPs

O alinhamento das sequencias foi realizado para possibilitar a identificação visual dos SNPs e a classificação em transições ou transversões. Em regiões codificantes e não-codificantes, conforme predição da proteína, as mutações

foram classificadas em sinônimas e não sinônimas, além da determinação da posição do SNP no códon.

4.11. Seleção dos oligonucleotídeos flanqueadores

No presente estudo, cada oligonucleotídeo flanqueador foi construído adjacente ao sítio da mutação, com até 18 pb. Para facilitar a discriminação dos SNPs de um mesmo gene, em algumas sequências de oligonucleotídeos foram inseridas cauda não homóloga na extremidade 5' dos *primers*, para a diferenciação por tamanho.

4.12. Reação e purificação do minisequenciamento

A reação de SNaPshot é baseada na extensão de base única de um oligonucleotídeo não marcado na posição 3', a qual está imediatamente adjacente ao SNP, por um dideoxi (ddNTP), na ausência de dNTPs. Cada ddNTP é marcado com corante fluorescente diferente e a quinta cor é usada para marcar o tamanho interno do marcador LIZ-120.

Essa estratégia é baseada na combinação de cores (usando diferentes fluorocromos para analisar polimorfismos A/G e C/T) e tamanho de extensão dos *primers* (o tamanho do *primer* foi modificado pela adição de uma cauda não homóloga a terminação 5'). O tamanho final dos *primers* abrangeu entre 18 a 30 pb.

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 q.s.p.. As amostras foram submetidas a uma

desnaturação a 96°C por 10 segundos, em seguida, 30 ciclos a 96°C por 10 segundos, 50°C por 5 segundos e 60°C por 30 segundos.

As reações de minisequenciamento (volume de 10 µl) foram tratadas com 1 U de *Shrimp Alkaline Phosphatase* (USB), por 45 minutos à 37°C e 20 minutos à 80°C. O produto purificado (1 µl) foi misturado a 10 µl de formamida e padrão GenScan – 120 Liz (*Applied Biosystems*). No momento da aplicação no sequenciador, as amostras foram aquecidas a 95°C por 5 minutos e transferidas para o gelo durante 5 minutos.

4.13. Análise da estrutura genética populacional

Os dados foram obtidos após seqüenciamento ou minisequenciamento. Os SNPs foram identificados, através da técnica de SNaPshot, com base no tamanho do produto amplificado e na fluorescência emitida, com a utilização do *software* Gene Scan™ v. 3.7 (*Applied Biosystems*).

A identidade de nucleotídeos e aminoácidos foi obtida pela contagem direta dos sítios conservados das sequencias de DNA de eucalipto, quando comparadas par a par com outras de *Arabidopsis thaliana* depositadas no GenBank. As substituições sinônimas e não-sinônimas foram, também, determinadas por contagem direta, com a comparação das sequencias disponíveis no GenBank.

O *software* DnaSP v.5 (Librado e Rozas, 2009) foi utilizado para determinar o número de sítios polimórficos, o número de haplótipos e o número médio de substituições nucleotídicas por sítio (π) das sequencias obtida para o gene. As combinações de dois haplótipos geraram um genótipo, esses dados foram desenvolvidos em planilhas, para a realização de estimativas de frequencias e distribuição dos genótipos na amostra.

5. RESULTADOS E DISCUSSÃO

5.1. Busca de Homólogos de NCED

A sequência de aminoácido da enzima NCED foi obtida da espécie *Arabidopsis thaliana*, com tamanho de 599 aminoácidos (*E.C. number* - 1.13.11.51 e número de acesso no GenBank – NP_188062) (Tabela 1).

A sequência escolhida foi referente ao gene NCED3, o qual apresenta os seguintes sinônimos na literatura: ATNCED3, *nine-cis-epoxycarotenoid dioxygenase 3*, *salt tolerant 1*, *sis 7*, *sto1*, *sugar insensitive 7*. Esse gene está envolvido na resposta a privação de água, processo biossintético do ABA, resposta a salinidade e estresse osmótico (<http://www.ncbi.nlm.nih.gov/protein/15231856>).

O gene NCED3 é formado por uma região codificante de 1800 pb, não apresentando região intrônica. As regiões 5' e 3' UTRs (*Untranslated Region*) apresentam tamanho de 122 e 390 pb, respectivamente.

Tabela 1: Enzima 9-cis-epoxicarotenoide dioxigenase selecionada através do banco de dados GenBank, estão apresentados sigla, *E.C. number*, número de acesso no GenBank, tamanho e organismo onde foi encontrada.

Sigla	Acesso	E.C. number	Tamanho	Organismo
NCED3	NP_188062	1.13.11.51	599	<i>Arabidopsis thaliana</i>

Entre as angiospermas, a espécie *Arabidopsis thaliana* é utilizada como planta modelo em diferentes ensaios de biologia molecular para o estudo de genes de interesse. O mesmo ocorre com as espécies *Oryza sativa* e *Populus trichocarpa*, as quais tiveram seu genoma sequenciado e, características nelas encontradas, podem ser utilizadas em espécies menos caracterizadas, como as do gênero *Eucalyptus* (Poke et al., 2005).

Os genes de florescimento, por exemplo, muito estudados em espécies como *Arabidopsis* e *Antirrhinum* (quanto à morfogênese floral, estrutura e desenvolvimento das flores), foram utilizados na busca de homólogos em outras espécies, sendo isolados e caracterizados em *E. globulus* e *E. grandis*. Homólogos de *Eucalyptus* foram encontrados em *Arabidopsis* (*LEAFY* e *AP1*) e em *Antirrhinum* (*FLORICAULA*), os quais são genes envolvidos na primeira etapa do desenvolvimento das flores, o meristema floral (Kyoizuka et al., 1997; Southerton et al., 1998).

A partir da obtenção da sequência da enzima (NCED3), foi realizada a busca por sequências similares no banco de dados de ESTs de *Eucalyptus*, através da ferramenta BLAST.

A comparação da sequência quanto a *cluster*, apresentou similaridade a um conjunto de 20 ESTs (19 ESTs de caule de mudas em condições de déficit hídrico e um EST de folha adulta em condições de ferrugem, cancro, déficit de fósforo e boro). Essas sequências apresentavam, geralmente, a região das extremidades com padrão de qualidade reduzido, segundo dados do próprio banco de dados, no qual regiões com melhores padrões estavam indicadas na cor verde e a representação da qualidade decaía reflexo de um gradiente até a cor vermelha (Figura 4). Um dos problemas no uso de banco de dados está, justamente, na baixa qualidade de algumas sequências, isso se deve as dificuldades em sequenciar determinadas regiões do genoma.

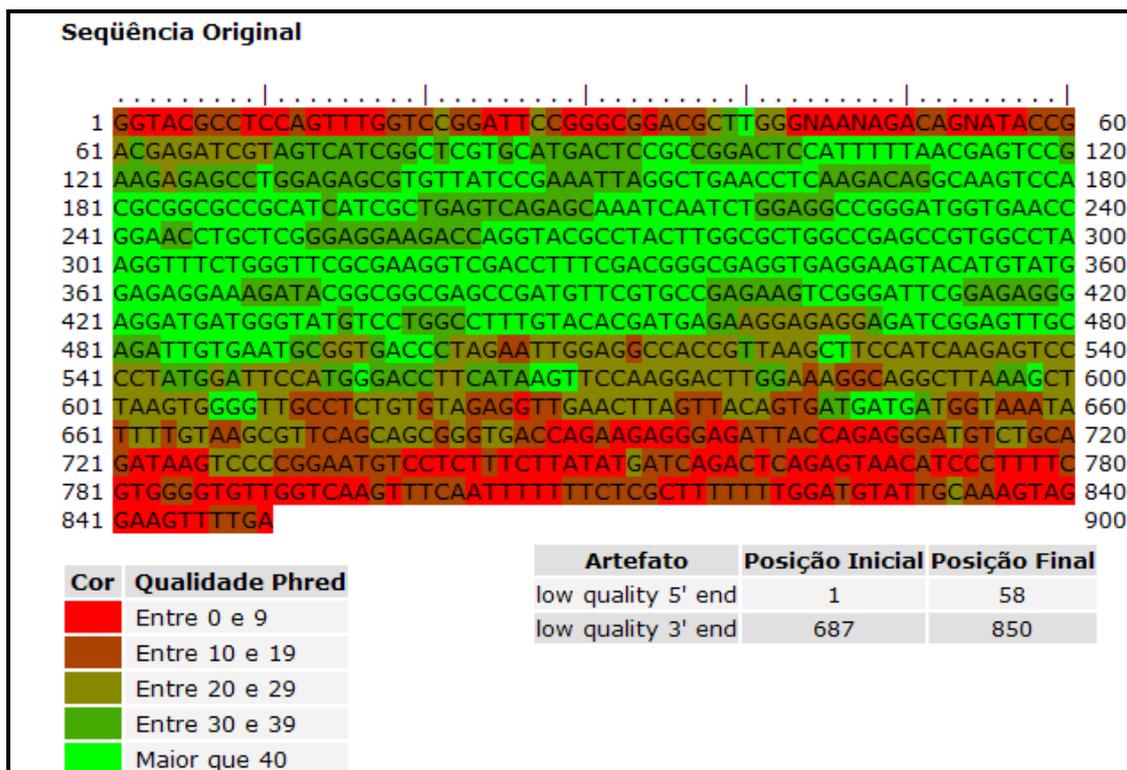


Figura 4: Representação do padrão de qualidade da sequência de EST (EFFOSTH02AB082G07.g) encontrada no banco de dados de *Eucalyptus* FORESTS-FAPESP. Regiões na sequência representadas na cor próxima a tonalidade vermelha indicam qualidades baixas, enquanto na verde correspondem a padrões mais altos. Na figura estão determinadas as posições iniciais e finais, com baixa qualidade, nas terminações 5' e 3'.

5.2. Desenho dos oligonucleotídeos iniciadores

As sequências de ESTs foram, então, alinhadas no Chromas Pro e, levando em consideração as regiões com melhor padrão de sequenciamento e os locais de alinhamento de mais de uma sequência, foram desenhados os três pares de oligonucleotídeos iniciadores (A_2; A_2.1, A_2.2) (Tabela 2), para uma região de amplificação de aproximadamente 1200 pb. Vale ressaltar que o desenho desses *primers* foi baseado em sequências únicas de ESTs, na região de melhor qualidade.

Recentemente, com a disponibilidade do banco de dados do genoma do eucalipto, foi possível o desenho de *primers* que englobam tamanho maior do gene que codifica NCED nessa espécie. O *Eucalyptus Genome Database* (<http://eucalyptusdb.bi.up.ac.za/>) é um banco de dados gerenciado pelo

Eucalyptus Genome Network (EUCAGEN) em colaboração com *US Department of Energy* (DOE) *Joint Genome Institute* (JGI). O genoma do *E. grandis* está atualmente sendo sequenciado, mas já se encontra disponível para o acesso em sua montagem preliminar (8 vezes sequenciada) de 600 Mpb (690 Mpb em 6043 *scaffolds*).

Essa fonte de informação foi útil para a construção dos oligonucleotídeos (ABA_1, ABA_2, ABA_3, ABA_4, ABA_5) a fim de amplificar uma região do gene NCED, *upstream* (964 pb) e *downstream* (263 pb) da sequência obtida pelos *primers* A_2; A_2.1 e A_2.2.

Tabela 2: *Primers* da enzima NCED que amplificaram banda única e seus respectivos tamanhos esperados (pb), temperaturas de pareamento (°C) e sequências.

Primer	Tamanho	Temperatura (°C)	Sequências
A_2	597	55	5' CTCATGCTCTTCTACGCCC 3' 3' GATCTCGTCGGTCTCTGTC 5'
A_2.1	626	53	5' ATTGCTTCTGCTTCCATCTC 3' 3' ACTAAGTTCAACCTCTACACAG5'
A_2.2	596	53	5' AGAGCAAATCAATCTGGAGG3' 3' TTGAAACTTGAACAACACCC 5'
ABA_1	320	57	5' AAGAAGAACAATCCGACACTCC3' 3' CCACCTCCAGAAGAAACGAC 5'
ABA_2	497	57	5' TCTCTTCTGCTTCAGCTCAC3' 3' CTCCACGACATCTAATGCGA 5'
ABA_3	292	57	5' GTCAAGTCTAGTTTCCTCTCG 3' 3' GACTCCACGACATCTAATGC 5'
ABA_4	677	57	5' GAAACAGAGCTCCACCTCC 3' 3' GATCATCGGAGAGTCGAGC 5'
ABA_5	562	57	5' TACACGATGAGAAGGAGAGG 3' 3' AAGCAGAAGATACACAGATCAC5'

5.3. Validação dos oligonucleotídeos iniciadores

Após testes de amplificação, os oito *primers* apresentaram banda única, conforme figura 5. As condições das reações, tamanho e sequencia dos *primers* estão na tabela 2.



Figura 5: Eletroforese em gel de agarose 1% com produtos de amplificação dos *primers*. **Legenda** – **A_2**, **A_2.1** e **A_2.2:** *primers*. **V1**, **V2**, **V3**, **V4:** indivíduos. **B:** branco. **L:** padrão molecular.

A região sequenciada pelos *primers* A_2; A_2.1 e A_2.2 apresentou tamanho aproximado de 1230 pb. Essa região recebeu o nome de *EuNCED*, para facilitar a citação no decorrer do texto.

As regiões amplificadas pelos oligonucleotídeos ABA_2; ABA_3; ABA_4; ABA_5 foram sequenciadas e apresentaram similaridade as regiões correspondentes, segundo banco de dados do genoma de *Eucalyptus*. Exceto a região amplificada pelo *primer* ABA_1, com sequenciamento de péssima qualidade, não foi possível realizar análise de similaridade.

O tamanho total obtido para esta sequência foi, aproximadamente, 2.400 pb, somando regiões codificantes e UTRs. Mas, devido à recente disponibilidade do banco de dados do genoma de *Eucalyptus*, os *primers* dessa região foram amplificados e sequenciados em dois indivíduos, para a confirmação da região

através do BLAST, sendo parte de um plano futuro, a procura de SNPs também nessa região do gene.

Na tabela 3 está o resultado do BLAST da região total obtida pelo sequenciamento das regiões amplificadas pelos *primers*: A_2, A_2.1, A_2.2, ABA_2, ABA_3, ABA_4, ABA_5. Essa região apresentou identidade maior que 80% entre as regiões de NCED nos outros organismos, através do BLASTX.

Tabela 3: Resultado do BLASTX no GenBank da sequencia total obtida para NCED em eucalipto, estão apresentados o acesso, nome da enzima, organismo encontrado, *score*, *E-value* e máxima identidade (%).

Acesso	Enzima/ Organismo	Score	E-value	Max. Identidade
XP_002316871	NCED/ <i>Populus trichocarpa</i>	877	0.0	81
XP_002330729.1	NCED/ <i>P. trichocarpa</i>	864	0.0	81
XP_0022773541	Proteína hipotética/ <i>Vitis vinifera</i>	860	0.0	81
AAAY98512.2	NCED/ <i>Stylosanthis guianensis</i>	859	0.0	79
AAR11193.1	NCED/ <i>V. vinifera</i>	859	0.0	81
CAE00459.2	NCED/ <i>Arachis hypogaea</i>	853	0.0	79

A região EuNCED apresentou identidade maior que 80% entre as regiões no BLASTX: o par de *primer* A_2 apresentou um fragmento com 85% de identidade em 161 aminoácidos, o A_2.1 obteve identidade de 81% em 183 aminoácidos e A_2.2 com 81% de identidade em 118 aminoácidos. Na tabela 4 estão os valores obtidos com o BLASTX da sequencia *EuNCED*.

Tabela 4: Resultado do BLASTX no GenBank da sequência *EuNCED*, estão apresentados o acesso, nome da enzima, organismo encontrado, *score*, *E-value* e máxima identidade (%).

Acesso	Enzima/ organismo	Score	E-value	Max. Identidade
XP_002316871.1	NCED/ <i>P. trichocarpa</i>	613	2e-173	83
XP_002330729.1	NCED/ <i>P. trichocarpa</i>	607	1e-171	82
AAAY98512.2	NCED/ <i>S. guianensis</i>	598	6e-169	81
XP_002277354.1	Proteína hipotética/ <i>V. vinifera</i>	597	1e-168	82
CAN81383.1	Proteína hipotética/ <i>V. vinifera</i>	595	3e-168	82
AAR11193.1	NCED1/ <i>V. vinifera</i>	595	4e-168	82
CAE00459.2	NCED/ <i>A. hypogaea</i>	593	2e-167	81
ABC26013.1	NCED3/ <i>CitrusClementina</i>	589	3e-166	81
XP_002529528.1	NCED/ <i>Ricinus communis</i>	582	3e-164	80
AAAY89370.1	NCED1/ <i>C. sinensis</i>	581	5e-164	80
ABA43901.1	NCED3/ <i>Coffea canephora</i>	579	2e-163	80
AAS47838.1	NCED/ <i>Gentiana lútea</i>	573	1e-161	78
AAT75151.1	NCED1/ <i>Solanum tuberosum</i>	573	2e-161	78

A partir de análises de similaridade de sequências homólogas no banco de dados foi possível verificar a existência de uma grande diversidade de NCED, sendo plausível que, segundo Chernys e Zeevaart (2000), genes com 60% ou mais de identidade exerçam a mesma função, enquanto que aqueles com menor identidade catalisam reações distintas. Essas regiões conservadas entre as sequências de proteínas estão envolvidas na ligação de substratos e cofatores.

Estes resultados indicam que genomas de uma diversidade de angiospermas apresentam muita similaridade em relação a enzimas dioxigenases de

carotenóides, sugerindo que o caminho biossintético do apocarotenóide, ABA, está altamente conservado (Tan et al., 2003).

5.4. Identificação dos SNPs

Através do alinhamento das 20 sequências de ESTs (anteriormente citadas), referente à enzima NCED, foi realizada a procura prévia de SNPs *in silico*. De acordo com Schmid et al., (2003), deve-se adotar um critério de qualidade para evitar erros de sequenciamento, sendo considerado como polimorfismo, quando o alelo menos frequente ocorrer em pelo menos duas sequências confiáveis da lista de sequências de ESTs. A partir desse pressuposto, não foi identificado SNP, pois quando feito o alinhamento, cerca de 30 nucleotídeos de ambas as extremidades das sequências (regiões 5' e 3') foram descartados e o alelo menor, quando ocorria, apresentava-se em apenas sequências individuais.

Apesar da valiosa contribuição da bioinformática na descoberta de SNPs, a validação experimental dos dados se faz necessária para avaliar o real potencial informativo de cada polimorfismo.

Mesmo com a disponibilidade de inúmeras técnicas modernas de genotipagem, o sequenciamento de fragmentos específicos foi utilizado nesse trabalho, por ser empregado no laboratório no qual foi realizado, mostrando-se eficiente na identificação de polimorfismos e não se necessitava da compra de novos equipamentos. Este é um método confiável com taxa de erro inferior a 5%, no entanto caro e requer maior trabalho (Liao e Lee, 2010; Jehan e Lakhanpaul, 2006; Vignal et al., 2002).

Para a identificação de SNPs, inicialmente, foram sequenciados 20 indivíduos para as regiões amplificadas pelos *primers* A_2; A_2.1 e A_2.2. Sete polimorfismos de única base foram encontrados na sequência *EuNCED* (Figura 6). No alinhamento dessas sequências, foi observado que os polimorfismos estavam

nas regiões amplificadas pelos *primers* A_2 e A_2.2, diante desse resultado, foram estes os dois *primers locus* específicos utilizados no sequenciamento para maior número de indivíduos.

Os sete SNPs considerados estão de acordo com Brookes (1999), pois o alelo polimórfico ocorre em pelo menos 1% da população testada. Importante ressaltar que os picos adjacentes (15 a 20 pb) aos SNPs foram verificados quanto ao padrão de qualidade para evitar a classificação incorreta do polimorfismo, devido a artefatos de sequenciamento.

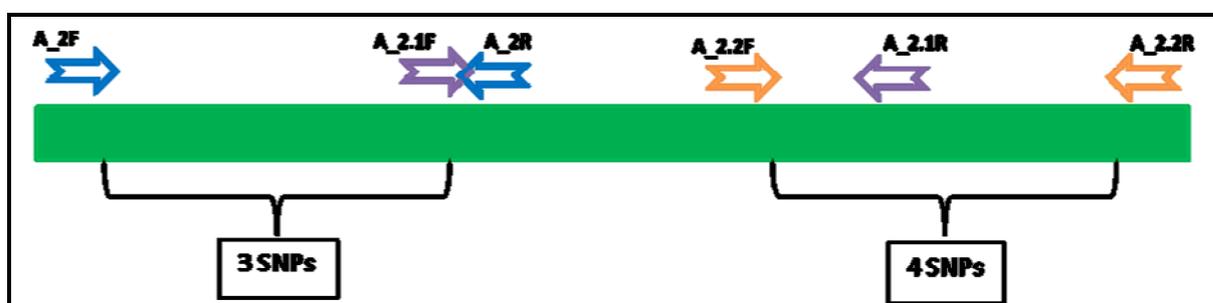


Figura 6: Esquema ilustrando a distribuição dos SNPs na sequência EuNCED (barra verde). Os polimorfismos estão nas regiões amplificadas pelos *primers* A_2 (seta azul) e A_2.2 (seta laranja). Na região amplificada pelo *primer* A_2.1 (seta roxa) não foi encontrado SNP.

No alinhamento das sequências, foram encontradas algumas mutações pontuais que ocorrem em determinados indivíduos. Essas alterações, algumas vezes, podem ser devido a erros de incorporação de bases, durante a replicação do DNA ou causadas por lesões no DNA por agentes ambientais. No entanto, se estas mutações ocorrerem em células germinativas e, dessa forma, forem transmitidas às gerações seguintes, se fixando na população na frequência mínima de 1%, passam a ser consideradas como polimorfismos (Kwok e Gu, 1999).

5.5. Genotipagem de SNPs

Na sequência *EuNCED* (1230 pb) a frequência de SNPs encontrada foi um a cada 176 pb, que se encontra dentro da média encontrada nas espécies de plantas, a qual é cerca de um SNP a cada 100 - 300 pb (Gupta et al., 2001). Entretanto, em Hao et al., (2009), a frequência de SNPs para a enzima NCED, em *Arabidopsis thaliana*, foi de um SNP a cada 87.4 pb, abrangendo, nesse estudo, o gene total (2011 pb).

Outros exemplos de distribuição de SNPs, que denotam a abundância desses polimorfismos em espécies vegetais são encontrados em: milho - um SNP a cada 104 pb (Tenaillon et al., 2001); trigo - um SNP a cada 540 pb (Somers et al., 2003); *Arabidopsis* - um SNP a cada 336 pb (Schmid et al., 2003); soja - um SNP a cada 336 pb (Zhu et al., 2003); arroz - um SNP a cada 232 pb (Nasu et al., 2002); *Populus* - um SNP a cada 60 pb (Ingvarsson, 2005).

Em estudos relacionados ao gênero *Eucalyptus* foi observada frequência de um SNP a cada 94 pb no gene CCR (*cinnamoyl coA reductase*), em *Eucalyptus nitens* (Thumma et al., 2005). Enquanto em uma análise mais ampla, envolvendo quatro espécies de *Eucalyptus*, foram sequenciados 23 genes em 1.764 indivíduos e descobertas diferentes frequências de SNPs: um a cada 33 pb em *E. nitens*; um a cada 31 pb em *E. globulus*; um a cada 16 pb para *E. camaldulensis* e um a cada 17 pb em *E. loxophleba* (Külheim et al., 2009).

5.6. Classificação dos SNPs

A sequência *EuNCED*, de 1230 pb, foi traduzida no site *Softberry* (www.softberry.com) gerando uma proteína hipotética de 356 aminoácidos, a qual foi utilizada na comparação com a sequência de NCED de *Arabidopsis*.

O alinhamento no ClustalX da sequência da proteína predita de *EuNCED* e a sequência da enzima de NCED em *Arabidopsis*, apresentou similaridade entre a região dos aminoácidos 245 ao 600, conforme figura 7.

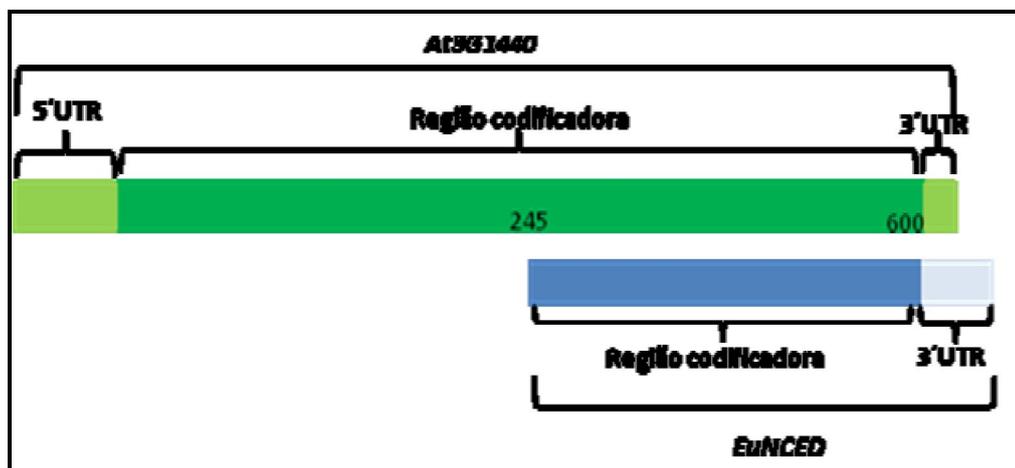


Figura 7: Alinhamento no ClustalX entre a sequência da proteína predita de *EuNCED* e a sequência da enzima de NCED em *Arabidopsis thaliana*.

A sequência *EuNCED* não apresentou indícios de íntrons quando realizado o alinhamento entre as sequências das proteínas (*EuNCED* e *AtNCED3*). Já que essa é uma das características importante desta enzima, relatada em Tan et al., (2003), que denota a ausência de íntrons nos cinco genes da subfamília *AtNCED* e no gene *AtCCD4*, enquanto *AtCCD1*, *AtCCD7* e *AtCCD8* apresentam estruturas de genes mais típicas a de eucariotos, incluindo múltiplos íntrons.

Enquanto a importância regulatória de poucos ou nenhum íntron na estrutura do gene não foi esclarecida, tal propriedade também se mostrou conservada em genes que codificam a NCED em outras espécies, como milho *Vp14* (Tan et al., 1997), *Notabilis* do tomate (Burbidge et al., 1999) e feijão *PvNCED1* (Qin e Zeevaart, 1999).

Há evidências de que a eficiência do *splicing* de íntrons possa sofrer degradação quando sob severos estresses em plantas (Bournay et al., 1996, Marrs e Walbot, 1997), logo a ausência de íntrons, neste caso, pode significar um

mecanismo para aumentar a expressão de NCED e, conseqüentemente, a síntese do ABA sob estresse.

Os SNPs encontrados na sequência *EuNCED* foram classificados de acordo com o tipo de variação do nucleotídeo: transição, purina-purina ou pirimidina-pirimidina (duas possibilidades) ou transversões, purina-pirimidina (quatro possibilidades).

Na sequência *EuNCED*, dos sete SNPs encontrados, quatro foram transições (57%) e três transversões (43%), com frequências de substituições G/C (28%), G/A (28%), C/T (28%) e G/T (14%). A razão transição/transversão foi de 1.33, próximo ao encontrado em uva (1,46 (Lijavetzky et al., 2007)), batata (1,48 (Simko et al., 2006)) e beterraba (1,63 (Schneider et al., 2007)), no entanto maior que aquela encontrada em soja (0,93 (Zhu et al., 2003)).

Outra classificação realizada foi de acordo com a mudança na codificação de aminoácidos em mutações sinônimas e não-sinônimas. Para a realização desta análise foi definido o quadro de leitura da sequência, através de um alinhamento BLASTX contra os peptídeos de *Arabidopsis thaliana* e, em seguida, os códons contendo os SNPs foram isolados e os aminoácidos traduzidos para cada alelo.

Dentre os sete SNPs encontrados, cinco estavam situados na região codificadora do gene (1071 pb), apresentando a frequência de um SNP a cada 214 pb, enquanto dois polimorfismos estavam localizados na região 3' UTR da sequência, ocorrendo um SNP a cada 79 pb.

Em *Arabidopsis thaliana* (AtNCED3), a frequência de SNP na região não codificadora foi 2,59 maior que na região codificadora (Hao et al., 2009). Na sequência *EuNCED*, a qual representou uma parcela do gene, a região não codificadora apresentou distribuição de SNPs três vezes maior que na região codificadora. Esta distribuição ocorreu de modo similar em milho (Ching et al., 2002) e soja (Zhu et al., 2003).

Na região codificadora, quatro das mutações ocorreram na terceira posição do códon, enquanto uma delas na primeira posição. No entanto, todas as mutações foram sinônimas (Figura 8), contrastando com o resultado encontrado por Hao et al., (2009), em *AtNCED3*, no qual a maioria foi mutações não-sinônimas (onze não sinônimas e oito sinônimas).

As mutações sinônimas não acarretam em troca de aminoácidos e, com isso, mudança da proteína. Entretanto, há perspectivas de que existam aminoácidos preferenciais (*códon usage bias*) em *Eucalyptus* (Mattioli et al., 2008), sendo, assim, algumas trincas de códons ocorrem em maior frequência e, dessa forma, tais diferenças sinônimas seriam fatores limitantes para expressão das proteínas.

SNPs	Códon		Mutação
C/G	GCC	GCG	Alanina
G/A	GAG	GAA	Ácido Glutâmico
G/A	GAG	GAA	Ácido Glutâmico
C/T	GTC	GTT	Valina
C/T	CTG	TTG	Leucina

Figura 8: SNPs da sequência *EuNCED*. Através da predição da proteína foi possível analisar a variação do códon e o aminoácido correspondente. Em amarelo estão marcadas as posições dos SNPs no códon.

5.7. Análise populacional

A sequência *EuNCED* (regiões codificantes e não codificantes), com 1230 pb, apresentaram sete sítios polimórficos entre 65 indivíduos analisados (Tabela 5).

Tabela 5: Distribuição dos sítios de polimorfismos entre os indivíduos.
Legenda - SNPs: SNP1; SNP2; SNP3; SNP4; SNP5; SNP6; SNP7.
Grandis: *E. grandis*. **Urophylla:** *E. urophylla*. **Urograndis:** *Urograndis*.
Clone: Clone.

Espécies	SNP1	SNP2	SNP3	SNP4	SNP5	SNP6	SNP7
Clone 1	C	G	A	C	T	T	G
Clone 2	G	G	A	T	C	K	C
Clone 3	C	G	A	C	T	T	G
Clone 4	C	G	R	C	T	T	G
Clone 5	S	G	A	Y	Y	K	S
Clone 6	G	G	A	T	C	K	C
Clone 13	C	R	R	C	T	T	G
Clone 14	G	G	A	T	C	K	C
Clone 15	C	G	G	C	T	T	G
Clone 16	C	G	R	C	T	T	G
Clone 17	C	G	A	C	T	T	G
Clone 18	C	R	R	C	T	T	G
Urograndis 1399	S	R	R	Y	Y	K	S
Urograndis 1460	G	G	A	T	C	G	C
Urograndis 1406	G	G	A	T	C	G	C
Urograndis 1462	S	G	A	Y	Y	K	S
Urograndis TC30	S	G	A	Y	Y	K	S
Urograndis 1348	S	G	R	Y	Y	K	S
Urograndis VR3709	S	G	R	Y	Y	K	S
Urograndis C219	S	G	A	Y	Y	K	S
Urograndis VR3748	S	G	R	Y	Y	K	S
Urograndis 1401	S	G	R	Y	Y	K	S
Urograndis 1288	C	R	G	C	T	T	G
Urograndis 1395	S	G	R	C	Y	T	S
Urograndis 1313	S	R	R	C	Y	K	S
Urograndis 1315	C	A	G	C	T	T	G
Urograndis 1372	C	R	R	C	T	T	G
Urograndis 1229	S	G	A	Y	Y	K	S
Urograndis 1163	C	R	G	C	T	T	G
Urograndis V1	G	G	A	Y	C	G	C
Urograndis V2	S	G	A	Y	Y	K	S
Grandis 62	C	G	R	C	T	T	G
Grandis 81	C	G	R	C	T	T	G

Continua

Continuação Tabela 5

Grandis 67	C	G	A	C	T	T	G
Grandis 75	C	G	R	C	T	T	G
Grandis 87	C	G	R	C	T	T	G
Grandis 107	C	G	R	C	T	T	G
Grandis 57	C	G	A	C	T	T	G
Grandis 76	C	G	A	C	T	T	G
Grandis 124	C	G	G	C	T	T	G
Grandis 121	C	G	R	C	T	T	G
Grandis 85	C	G	A	C	T	T	G
Grandis 86	C	G	G	C	T	T	G
Grandis 71	C	G	R	C	T	T	G
Grandis 50	C	G	A	C	T	T	G
Grandis 59	C	G	R	C	T	T	G
Grandis 93	C	G	R	C	T	T	G
Grandis 56	C	G	G	C	T	T	G
Grandis 55	C	G	R	C	T	T	G
Grandis 53	C	G	A	C	T	T	G
Grandis 52	C	G	A	C	T	T	G
Grandis 47	C	G	A	C	T	T	G
Grandis 46	C	G	R	C	T	T	G
Urophylla 54	G	G	A	Y	C	K	C
Urophylla 47	S	R	R	Y	Y	K	S
Urophylla 86	S	R	R	Y	Y	K	S
Urophylla 91	G	G	A	T	C	K	C
Urophylla 84	S	G	A	Y	Y	K	S
Urophylla 59	S	G	A	Y	Y	K	S
Urophylla 60	G	G	A	T	C	G	C
Urophylla 46	C	G	G	C	T	T	G
Urophylla 77	G	G	A	Y	C	G	C
Urophylla 75	S	G	A	Y	Y	K	S
Urophylla 68	G	G	A	Y	C	K	C
Urophylla 85	S	G	A	Y	Y	K	S

O número de haplótipos encontrados nessa amostra foi sete (Tabela 6), sendo o Hap_1 mais frequente (36,9%), seguido de Hap_3 (23,8%) e Hap_4 (23%). Os haplótipos Hap_2 e Hap_4 estavam presentes nas três espécies. Enquanto Hap_1, Hap_3, Hap_5 e Hap_7 apareceram nas espécies *E. urophylla* e no híbrido Urograndis. Já o Hap_6 foi encontrado unicamente no organismo Urograndis.

Tabela 6: Haplótipos encontrados na amostra, onde estão os nomes dados, sequencias, distribuição na população, porcentagem na amostra e espécies nas quais foram encontrados os haplótipos.

Nome	Sequencia	Distribuição	% na amostra	Espécies
Hap_1	CGACTTG	48	36,9	<i>E.urophylla</i> e <i>Urograndis</i>
Hap_2	GGATCTC	4	3,1	<i>E.grandis</i> , <i>E.urophylla</i> e <i>Urograndis</i>
Hap_3	GGATCGC	31	23,8	<i>E.urophylla</i> e <i>Urograndis</i>
Hap_4	CGGCTTG	30	23	<i>E. grandis</i> , <i>E.urophylla</i> e <i>Urograndis</i>
Hap_5	CAGCTTG	11	8,5	<i>E.urophylla</i> e <i>Urograndis</i>
Hap_6	GGACCTG	3	2,3	<i>Urograndis</i>
Hap_7	GGACCGC	3	2,3	<i>E.urophylla</i> e <i>Urograndis</i>

Com a combinação dos sete haplótipos encontrados foi possível identificar 15 genótipos. Os genótipos mais frequentes foram I (16,92%), II (15,38%) e III (20%), distribuídos entre os três tipos de eucaliptos. Enquanto os menos frequentes foram XIII, XIV, XV (1,54%). A distribuição desses genótipos na amostra demonstrou que os genótipos IV, VII, IX, XII, XIII, XIV e XV foram únicos para o híbrido *Urograndis*. Enquanto que os demais genótipos estavam distribuídos entre as espécies, como mostrado na tabela 7.

Para esses polimorfismos, os indivíduos caracterizados como tolerantes ou susceptíveis ao déficit hídrico não apresentaram padrões genéticos que os diferenciasssem em haplótipos ou genótipos específicos.

Como, por exemplo, Hao et al., (2009) sugeriu em seu trabalho, que um acesso, o 163av de *Arabidopsis*, quando comparado a outras sequencias da mesma espécie, apresentou variações de aminoácidos específicas que o distinguiu em um haplótipo. As variações estavam no sítio conservado da proteína NCED, sendo que uma variação nessa região, segundo o autor, poderia alterar a atividade enzimática,

levando, assim, ao aumento da síntese de ABA, quando em condições de estresse hídrico, tornando esse acesso mais tolerante a seca.

Tabela 7: Genótipos encontrados na amostra, onde estão os nomes dados, distribuição na população, porcentagem na amostra e espécies nas quais foram encontrados os genótipos.

Nome	Genótipos	Distribuição	% na amostra	Espécies
I	1/1	11	16,92	72,7 % <i>E. grandis</i> ; 27,3 % <i>Urograndis</i>
II	1/3	10	15,38	60 % <i>E. urophylla</i> 40 % <i>Urograndis</i>
III	1/4	13	20	84,6 % <i>E. grandis</i> 15,4 % <i>Urograndis</i>
IV	1/5	3	4,62	100 % <i>Urograndis</i>
V	2/3	4	6,15	25 % <i>E. urophylla</i> 75 % <i>Urograndis</i>
VI	3/3	3	4,62	33,3 % <i>E. urophylla</i> 66,7 % <i>Urograndis</i>
VII	3/4	4	6,15	100 % <i>Urograndis</i>
VIII	3/5	3	4,62	66,7 % <i>E. urophylla</i> 33,3 % <i>Urograndis</i>
IX	3/6	2	3,08	100 % <i>Urograndis</i>
X	3/7	2	3,08	50 % <i>Urograndis</i> 50 % <i>E. urophylla</i>
XI	4/4	5	7,69	60 % <i>E. grandis</i> 20 % <i>E. urophylla</i> 20 % <i>Urograndis</i>
XII	4/5	2	3,08	100 % <i>Urograndis</i>
XIII	4/6	1	1,54	100 % <i>Urograndis</i>
XIV	5/5	1	1,54	100 % <i>Urograndis</i>
XV	5/7	1	1,54	100 % <i>Urograndis</i>

Como esses dados foram obtidos a partir de um número reduzido de amostra, é preciso destacar a necessidade de validação desses SNPs em um número maior de indivíduos. Mas, essas informações já são indicativas da presença de alelos específicos entre as espécies estudadas. Uma possível utilização desses dados seria o desenvolvimento de marcadores espécies-específicos para o monitoramento de cruzamentos interespecíficos em programas de melhoramento de eucalipto.

Outro parâmetro utilizado nas análises populacionais foi a diversidade nucleotídica, ou seja, a taxa de diferenças nucleotídicas que ocorre entre duas sequências escolhidas aleatoriamente. No entanto, essas estimativas podem variar dependendo de diversos fatores, sendo os principais: sistema reprodutivo, efeitos de amostragem, tanto em termos de indivíduos e população, como em quais ou quantos genes foram analisados.

As diversidades nucleotídicas (π) de *E. urophylla* e *Urograndis* foram 0,41 e 0,44, respectivamente. Enquanto que o valor encontrado em *E. grandis* foi bem menor, 0,06. Em Santos (2005), em um trecho de 440 pb do gene *coaomt*, *E. grandis* ($\pi = 0,00356$) apresentou diversidade nucleotídica duas vezes maior que *E. globulus* ($\pi = 0,00168$) e cerca de 1,4 vezes maior que em *E. urophylla* ($\pi = 0,00254$). Esses resultados foram muito contrastantes em relação ao encontrado no gene NCED.

Resultado diferente do obtido em NCED foi o encontrado no estudo com *Pinus sylvestris*, no qual a diversidade nucleotídica na região do fitocromo *O* variou entre 0,0020 a 0,0037 na região intrônica e 0,0005 a 0,0020 na região do éxon (Garcia-Gil et al., 2003).

No entanto, vale destacar, que não foram realizados testes estatísticos para verificar se existiam diferenças significativas entre as estimativas obtidas para as espécies. Além de que, fazem-se necessárias análises com maior extensão da

sequencia do gene, para revelar estimativas mais próximas do parâmetro real e, possivelmente, confirmar esses dados.

5.8. Testes de extensão dos oligonucleotídeos adjacentes aos polimorfismos

Nos sítios polimórficos encontrados foram desenhados *primers* adjacentes aos pontos de mutação, com o objetivo de reduzir os custos e permitir a realização de genotipagem em larga escala desses polimorfismos de única base em populações de *Eucalyptus*.

Os testes de reação SNaPshot (Tabela 8) mostraram que: quatro *primers* falharam (A_S2, A_S3, A_S5 e A_S6), mesmo após ajustes nas concentrações dos mesmos, indicando a ausência de pareamento com a sequencia de DNA; o *primer* A_S1 apresentou monomorfismo quando testado em uma população de 96 indivíduos, fato que se deve a provável formação de ligações inespecíficas; A_S4 teve tamanho maior que o esperado, devido também a possível formação de pareamentos inespecíficos; e, A_S7 apresentou polimorfismo, quando testado em cinco indivíduos, mas necessita de padronização, em uma população maior.

No entanto, no último teste realizado ocorreu erro na genotipagem com problemas relacionado a reagentes, supostamente, a enzima SAP não estava purificando corretamente. Dessa forma, a padronização em uma maior população será realizada posteriormente.

Tabela 8: *Primers* utilizados na reação de SNaPshot, com tamanho (pb), o polimorfismo e as concentrações (μM) testadas.

Primer	Sequencia 5' → 3'	Tamanho	Polimorfismo	Concentrações
A_S1	GTACCCGCAGCGCTTG CG	18	C/G	0.2; 0.4; 0.8
A_S2	AAAAGCTCTAGGGGGA CCGACT	22	G/A	0.2; 0.4; 0.8; 1.5
A_S3	AAAAAAAAGCCTTAGTT CACCTATCT	26	G/A	0.2; 0.4; 0.8; 1.5
A_S4	AAGACCCAAGCGCTTC CA	18	C/T	0.2; 0.4; 0.8; 1.5
A_S5	AAAATTCCTCTCCTCTA GCCTC	22	C/T	0.2; 0.4; 0.8; 1.5
A_S6	AAAAAAAATACTACTACT ACCATTTAT	26	G/T	0.2; 0.4; 0.8; 1.5
A_S7	AAAAAAAAAAAAACCCG CTGCTGAACGCTTA	30	G/C	0.2

Novos *primers* para esses sítios de polimorfismos que falharam foram confeccionados e serão testados, a fim de realizar genotipagem com maior número de indivíduos através da técnica de SNaPshot, um método mais rápido e automatizado.

6. CONSIDERAÇÕES FINAIS

A identificação e validação de novos marcadores moleculares em *Eucalyptus*, em genes expressos, é uma estratégia utilizada por diferentes grupos de pesquisas. Essa tática foi facilitada com a disponibilidade de diferentes bancos públicos de sequências de DNA.

Nesse trabalho foram disponibilizados novos marcadores presentes em uma cópia do gene que sintetiza a proteína NCED, em eucalipto. Esses marcadores foram caracterizados em duas espécies de eucalipto e em um híbrido, demonstrando, assim, a aplicabilidade em estudos de genética de população e em pesquisas, as quais possam inferir sobre a participação desses polimorfismos na resposta da planta a diferentes tipos de estresse abiótico.

Por se tratar de um gene que, segundo a literatura, apresenta grande impacto na resposta das plantas ao estresse hídrico, esses marcadores poderão ser avaliados em progênies segregantes para o estresse hídrico. Com isso, diferenças alélicas desse gene poderão ser validadas com o uso dessas progênies.

7. CONCLUSÕES

- O gene que codifica a enzima NCED, em *Eucalyptus*, apresenta a mesma estrutura descrita em *Arabidopsis*, sem região intrônica.
- Na região de 1230 pb avaliada do gene NCED foram identificados sete sítios polimórficos, cinco na região codificante e dois na não-codificante.
- As mutações encontradas na região codificante são sinônimas.
- Os sete polimorfismos encontrados na região formaram sete haplótipos, os quais, por sua vez, geraram 15 combinações haplotípicas.

8. REFERÊNCIAS

- ALTSCHUL, S.F.; MADDEN, T.L.; SCHÄFFER, A.A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, v.25, n.17, p.3389-3402, 1997.
- ANDRADE, E.D. O eucalipto. 1.ed. S. Paulo: *Chácaras e quintais*, 1939, p.8-13.
- AULDRIDGE, M.E.; BLOCK, A.; VOGEL, J.T.; DABNEY-SMITH, C.; MILA, I.; BOUZAYEN, M.; MAGALLANES-LUNDBACK, M.; DELLAPENNA, D.; MCCARTY, D.R.; KLEE, H.J. Characterization of three members of the Arabidopsis carotenoid cleavage dioxygenase family demonstrates the divergent roles of this multifunctional enzyme family. *The Plant Journal*, v.45, p.982-993, 2006.
- BECKER, D.; HOTH, S.; ACHE, P.; WENKEL, S.; ROELFSEMA, M.R.G.; MEYERHOFF, O.; HARTUNG, W.; HEDRICH, R. Regulation of the ABA-sensitive Arabidopsis potassium channel gene GORK in response to water stress. *FEBS Letters*, v.554, p.119-126, 2003.
- BOURNAY, A.S.; HEDLEY, P.E.; MADDISON, A.; WAUGH, R.; MACHRAY, G.C. Exon skipping induced by cold stress in a potato invertase gene transcript. *Nucleic Acids Research*, v.24, p.2347-2351, 1996.
- BRAY, E.A. Molecular responses to water deficit. *Plant Physiology*, v.103, p.1035-1040, 1993.
- BROOKES, A.J. The essence of SNPs. *Gene*, v.234, n.2, p.177-186, 1999.
- BURBIDGE, A.; GRIEVE, T.M.; JACKSON, A.; THOMPSON, A.; MCCARTY, D.R.; TAYLOR, I.B. Characterization of the ABA-deficient tomato mutant notabilis and its relationship with maize Vp14. *Plant Journal*, v.17, p.427-431, 1999.

BURBIDGE, A.; GRIEVE, T.; JACKSON, A.; THOMPSON, A.; TAYLOR, I. Structure and expression of a cDNA encoding a putative neoxanthin cleavage enzyme (NCE) isolated from a wilt-related tomato (*Lycopersicon esculentum* Mill.) library. *Journal of Experimental Botany*, v.48, p.2111–2112, 1997.

CAMBRAIA, J. Aspectos bioquímicos, celulares e fisiológicos dos estresses nutricionais em plantas. In: Eds. NOGUEIRA, R.J.M.C.; ARAÚJO, E.L.; WILLADINO, L.G.; CAVALCANTE, U.M.T. *Estresses ambientais: danos e benefícios em plantas*, p.95-105, 2005.

CASAGRANDE, E.C.; FARIAS, J.R.B.; NEUMAIER, N.; OYA, T.; PEDROSO, J.; MARTINS, P.K.; BRETON, M.C.; NEPOMUCENO, A.L. Expressão gênica diferencial durante déficit hídrico em soja. *Revista Brasileira de Fisiologia Vegetal*, v.13, n.2, p.168-184, 2001.

CAVALCANTI, G. Identificação das principais espécies de *Eucalyptus* existentes no Brasil. *Silvicultura em São Paulo*, v.1, n.2, 1963.

CHERNYS, J.T.; ZEEVAART, J.A.D. Characterization of the 9-cis-epoxycarotenoid dioxygenase gene family and the regulation of abscisic acid biosynthesis in avocado. *Plant Physiology*, v.124, p.343–353, 2000.

CHING, A.; CALDWELL, K.S.; JUNG, M.; DOLAN, M.; SMITH, O.S.; TINGEY, S.; MORGANTE, M.; RAFALSKI, A. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BioMedCentral Genetics*, v.3, n.19, p.1-14, 2002.

CORDEIRO, G.M.; ELIOTT, F.; MCINTYRE, C.L.; CASU, R.E.; HENRY, R.J. Characterisation of single nucleotide polymorphisms in sugarcane ESTs. *Theoretical and Applied Genetics*, v.113, p.331-343, 2006.

DAVIES, W.J.; WILKINSON, S.; LOVEYS, B. Stomatal control by chemical signaling and the exploitation of this mechanism to increase water use efficiency in agriculture. *New Phytologist*, v.153, p.449-460, 2002.

DING, C. 'Other' applications of single nucleotide polymorphisms. *Elsevier*, v.25, n.7, p.1-5, 2007.

ELDRIDGE, K. *Eucalypt domestication and breeding* ed. Oxford, New York: Clarendon Press; Oxford University Press, 1994.

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 mai. 2010.

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

FONSECA, S.M.; RESENDE, M.D.V.; ALFENAS, A.C.; GUIMARÃES, L.M.S.; ASSIS, T.F.; GRATTAPAGLIA, D. *Manual Prático de Melhoramento Genético do Eucalipto*. Viçosa: UFV, 2010, 200p.

GANAL, M.W.; ALTMANN, T.; RÖDER, M.S. SNP identification in crop plants. *Current Opinion in Plant Biology*, v.12, p.211-217, 2009.

GARCIA-GIL, M.R.; MIKKONEN, M.; SAVOLAINEN, O. Nucleotide diversity at two phytochrome loci along a latitudinal cline in *Pinus sylvestris*. *Molecular Ecology*, v.12, p.1195-1206, 2003.

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.

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

GRIFFITHS, A.J.F.; GELBART, W.M.; MILLER, J.H.; LEWONTIN *Genética Moderna*. Rio de Janeiro: Guanabara Koogan, 2001. 589p.

GUERRINI, I.A.; TRIGUEIRO, R.M.; LEITE, R.M.; WILKEN, C.F.; VELINI, E.D.; MORI, E.S.; FURTADO, E.L.; MARINO, C.L.; MAIA, I.G. *Eucalyptus* ESTs involved in the production of 9-cis epoxycarotenoid dioxygenase, a regulatory enzyme of abscisic acid production. *Genetics and Molecular Biology*, v.28, n.3, p.640-643, 2005.

GUPTA, P.K.; ROY, J.K.; 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.25, 2001.

HAO, G.P.; ZHANG, X.H.; WANG, Y.Q.; WU, Z.Y.; HUANG, C.L. Nucleotide variation in the NCED3 region of *Arabidopsis thaliana* and its association study with abscisic acid content under drought stress. *Journal of Integrative Plant Biology*, v.51, n.2, p.175-83, 2009.

INGVARSSON, P.K. Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European aspen (*Populus tremula* L., Salicaceae). *Genetics*, v.169, n.2, p.945-953, 2005.

ISRAELSSON, M.; SIEGEL, R.S.; YOUNG, J.; HASHIMOTO, M.; IBA, K.; SCHROEDER, J.I. Guard cell ABA and CO₂ signaling network updates and Ca²⁺ sensor priming hypothesis. *Current Opinion in Plant Biology*, v.9, p.654-663, 2006.

IUCHI, S.; KOBAYASHI, M.; TAJI, T.; NARAMOTO, M.; SEKI, M.; KATO, T.; TABATA, S.; KAKUBARI, Y.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in *Arabidopsis*. *Plant Journal*, v.27, p.325-333, 2001.

IUCHI, S.; KOBAYASHI, M.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. A stress-inducible gene for 9-cis-epoxycarotenoid dioxygenase involved in abscisic acid biosynthesis under water stress in drought-tolerant cowpea. *Plant Physiology*, v.123, p.553-562, 2000.

JEHAN, T.; LAKHANPAUL, S. Single Nucleotide Polymorphism (SNP) – Methods and application in plant genetics: a review. *Indian Journal of Biotechnology*, v.5, p.435-459, 2006.

JIA, W.; XING, Y.; ZHAO, J.; ZHANG, J.; ZHANG, S.Q. Cellular stress signaling from water deficit-perception to ABA accumulation. *Advanced Plant Science*, v.4, p.205-212, 2002.

KOORNEEF, M.; LEON-KLOOSTERZIEL, K.M.; SCHWARTZ, S.H.; ZEEVAART, J.A.D. The genetic and molecular dissection of abscisic acid biosynthesis and signal transduction in Arabidopsis. *Plant Physiology Biochemical*, v.36, p.83-89, 1998.

KÜLHEIM, C.; YEOH, S.H.; MAINTZ, J.; FOLEY, W.J.; MORAN, G.F. Comparative SNP diversity among four *Eucalyptus* species for genes from secondary metabolite biosynthetic pathways. *BioMed Central Genomics*, v.10, n.452, p.1-11, 2009.

KWOK, P.Y.; GU, Z. Single Nucleotide Polymorphism libraries: why and how are we building them? *Molecular Medicine Today*, v.5; n.12, p.538-543, 1999.

KYOZUKA, J.; HARCOURT, R.; PEACOCK, W.J.; DENNIS, E.S. *Eucalyptus* has functional equivalents of the *Arabidopsis AP1* gene. *Plant Molecular Biology*, v.35, n.5, p.573-584, 1997.

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

LAWLOR, D. W. Limitation of photosynthesis in water-stressed leaves: stomata vs metabolism and the role of ATP. *Annals of Botany*, v.89, p.871-885, 2002.

LECOEUR, J.; SINCLAIR, R.T. Field pea transpiration and leaf growth in response to soil water deficit. *Crop Science*, v.36, p.331-335, 1996.

LEUNG, H. Stressed genomics – bringing relief to rice fields. *Current opinion in Plant Biology*, v.11, p.201-208, 2008.

LIAO, M P.Y.; LEE, K.H. From SNPs to functional polymorphism: The insight into biotechnology applications. *Biochemical Engineering Journal*, v.49, p.149-158, 2010.

LIBRADO, P.; ROZAS, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, v.25, p.1451-1452, 2009.

LIJAVETZKY, D.; CABEZAS, J.A.; IBÁÑEZ, A.; RODRÍGUEZ, V.; MARTÍNEZ-ZAPATER, J.M. High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. *BioMedCentral*, v.8, p.1-11, 2007.

MARRS, K.A.; WALBOT, V. Expression and RNA splicing of the maize glutathione S-transferase Bronze2 gene is regulated by cadmium and others stresses. *Plant Physiology*, v.113, p.93-102, 1997.

MARSHALL, O.J. PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics*, v.20, n.15, p.2471-2472, 2004.

MATTIOLI, M.A.P. *Estudo do padrão de utilização de códons em sequencias gênicas de eucalipto visando a maximização da produção de proteínas através da otimização das sequencias de DNA*. Tese (mestrado) Universidade Estadual de Campinas, 2008.

MAYR, E. *Populações, espécie e evolução*. São Paulo: Nacional, (Biblioteca Universitária, 3), v.5, 1977.

MORA, A.; GARCIA, C. *A Cultura do Eucalipto no Brasil*. 1.ed. São Paulo: Sociedade Brasileira de Silvicultura, 2000. 112p.

NAMBARA, E.; MARION-POLL, A. ABA biosynthesis and catabolism. *Annual Review Plant Biology*, v.56, p.165-185, 2005.

NASU, S.; SUZUKI, J.; OHTA, R.; HASEGAWA, K.; YUI, R.; KITAZAWA, N.; MONNA, L.; MINOBE, Y. Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa*, *Oryza rufipogon*) and establishment of SNP markers. *DNA Research*, v.9, p.163-171, 2002.

NEPOMUCENO, A.L.; NEUMAIER, N.; FARIAS, J.R.B.; OYA, T. Tolerância à seca em plantas. *Biotechnology Ciência e Desenvolvimento*, v.23, p.12-18, 2001.

OSMAN, A.; JORDAN, B.; LESSARD, P.A.; MUHAMMAD, N.; HARON, M.R.; RIFFIN, N.M.; SINSKEY, A.J.; RHA, C.; HOUSMAN, D.E. Genetic diversity of *Eurycoma*

longifolia inferred from Single Nucleotide Polymorphisms. *Plant Physiology*, v.131, p.1294-1301, 2003.

PINTO, G.; LOUREIRO, J.; LOPES, T.; SANTOS, C. Analysis of the genetic stability of *Eucalyptus globulus* Labill. somatic embryos by flow cytometry. *Theoretical and Applied Genetics*, v.109, p.580-587, 2004.

POKE, F.S.; VAILLANCOURT, E.; POTTS, B.M.; REID, J.B. Genomic research in *Eucalyptus*. *Genetica*, v.125, p.79-101, 2005.

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

QIN, X.; ZEEVAART, J.A.D. The 9-cis-epoxycarotenoid cleavage reaction is the key regulatory step of abscisic acid biosynthesis in water-stressed bean. *Proceedings of the National Academy of Sciences USA*, v.96, p.15354-15361, 1999.

QUINTÁNS, B.; ÁLVAREZ-IGLESIAS, V.; SALAS, A.; PHILLIPS, C.; LAREU, M.V.; CARRACEDO, A. Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing. *Elsevier*, v.104, p.251-257, 2004.

REIS, G.G.; REIS M.G.F. Fisiologia da brotação de eucalipto com ênfase nas suas relações hídricas. *Série técnica IPEF*, v.11, n.30, p.9-22, 1997.

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.

SANTOS, R.F.; CARLESSO, R. Deficit hídrico e os processos morfológicos e fisiológicos das plantas. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.2, n.3, p.287-294, 1998.

SCHMID, K.J.; SORENSEN, T.R.; STRACKE, R.; TORJEK, O.; ALTMANN, T.; MITCHAELL-OLDS, T.; WEISSHAAR, B. Large-scale identification and analysis of

genome-wide Single Nucleotide Polymorphisms for mapping in *Arabidopsis thaliana*. v.3, p.1250-1257, 2003.

SCHNEIDER, K.; KULOSA, D.; SOERENSEN, T.R.; MÖHRING, S.; HEINE, M.; DURSTEWITZ, G.; POLLEY, A.; WEBER, E.; LEIN, J.; HOHMANN, U.; TAHIRO, E.; WEISSHAAR, B.; SCHULZ, B.; KOCH, G.; JUNG, C.; GANAL, M. Analysis of DNA polymorphisms in sugar beet (*Beta vulgaris* L.) and development of an SNP-based map of expressed genes. *Theoretical and Applied Genetics*, v.115, p.601-615, 2007.

SCHWARTZ, S.H.; QIN, X.Q.; ZEEVAART, J.A.D. Characterization of a novel carotenoid cleavage dioxygenase from plants. *The Journal of Biological Chemistry*, v.276, p.2520-2528, 2001.

SCHWARTZ, S.H.; QIN, X.; ZEEVAART, J.A.D. Elucidation of the indirect pathway of abscisic acid biosynthesis by mutants, genes, and enzymes. *Plant Physiology*, v.131, p.1591-1601, 2003.

SHARP, R.E.; LENOBLE, M.E. ABA, ethylene and the control of shoot and root growth under water stress. *Journal of Experimental Botany*, v.53, n.366, p.33-37, 2002.

SIMKO, I.; HAYNES, K.G.; JONES, R.W. Assessment of linkage disequilibrium in potato genome with single nucleotide polymorphism markers. *Genetics*, v.173, n.4, p.2237-2245, 2006.

SOMERS, D.J.; KIRKPATRICK, R.; MONIWA, M.; WALSH, A. Mining single nucleotide polymorphisms from hexaploid wheat ESTs. *Genome*, v.46, n.3, p.431-437, 2003.

SOUTHERTON, S.G.; STRAUSS, S.H.; OLIVE, M.R.; HARCOURT, R.L.; DECROOCQ, V.; ZHU, X.M.; LLEWELLYN, D.J.; PEACOCK W.; DENNIS, E.S. *Eucalyptus* has a functional equivalent of the *Arabidopsis* floral meristem identity gene *LEAFY*. *Plant Molecular Biology*, v.37, n.6, p.897-910, 1998.

STAPE, J.L.; RYAN, G.M.; BINKLEY, D. Testing the utility of the 3-PG model for growth of *Eucalyptus grandis* x *urophylla* with natural and manipulated supplies of

water and nutrients. *Forest Ecology and Management*, v.193, n.1/2, p.219-234, 2004.

TAILLON-MILLER, P.; GU, Z.; LI, Q.; HILLER, L.; KWOK, P.Y. Overlapping genomic sequences: a treasure trove of single nucleotide polymorphisms. *Genome Research*, v.8, n.7, p.748-754, 1998.

TAIZ, L.; ZEIGER, E. *Fisiologia Vegetal*, 3° ed., Piracicaba: Artmed Editora S.A., 2004.

TAN, B.C.; JOSEPH, L.M.; DENG, W.T; LIU, L.; LI, Q.B.; CLINE, K.; MCCARTY, D.R. Molecular characterization of the *Arabidopsis* 9-cis-epoxycarotenoid dioxygenase gene family. *The Plant Journal*, v.35, p.44-56, 2003.

TAN, B.C.; SCHWARTZ, S.H.; ZEEVAART, J.A.D.; MCCARTY, D.R. Genetic control of abscisic acid biosynthesis in maize. *Proceedings of the National Academy of Sciences USA*, v.94, p.12235-12240, 1997.

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

TENAILLON, M.I.; SAWKINS, M.C.; LONG, A.D.; GAUT, R.L.; DOEBLEY, J.F.; GAUT, B.S. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Proceedings of the National Academy of Sciences USA*, v.98, n.16, p.9161-9166, 2001.

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*, v.171, p.1257-1265, 2005.

VICENTINI, R.; SASSAKI, F. T.; GIMENES, M. A.; MAIA, I. G.; MENOSSI, M. In silico evaluation of the *Eucalyptus* transcriptome. *Genetics and Molecular Biology*, v.28, n.3 (suppl), p.487-495, 2005.

VIGNAL, A.; MILAN, D.; SANCRISTOBAL, M.; EGGEN, A. A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, v.34, p.274-305, 2002.

WASILEWSKA, A.; FLORINA, V.L.A.D.; SIRICHANDRA, C.; REDKO, Y.; JAMMES F.; VALON C.; DIT FREY, N. F.; LEUNG, J. An update on abscisic acid signaling in plants and more ... *Molecular Plant*, v.1, n.2, p.198–217, 2008.

YANG, J.F.; GUO, Z.F. Cloning of a 9-cis-epoxycarotenoid dioxygenase gene (*SgNCED1*) from *Stylosanthes guianensis* and its expression in response to abiotic stresses. *Plant Cell Reports*, v.26, p.1383-1390, 2007.

ZEEVAART, J.A.D.; CREELMAN, R.A. Metabolism and physiology of abscisic acid. *Annual Review of Plant Physiology and Plant Molecular Biology*, v.39, p.439–473, 1988.

ZHANG, D.X.; HEWITT, G.M. Molecular DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, v.12, n.3, p.563-584, 2003.

ZHU, Y.L.; SONG, Q.J.; HYTEN, D.L.; VAN TASSELL, C.P.; MATUKUMALLI, L.K., GRIMM, DR; HYATT, S.M.; FICKUS, E.W.; YOUNG, N.D.; CREGAN, P.B. Single-nucleotide polymorphisms in soybean. *Genetics*, v.163, p.1123–1134, 2003.