



GABRIEL SÉRGIO COSTA ALVES

**IDENTIFICAÇÃO E ANÁLISE DE
POLIMORFISMOS NA REGIÃO PROMOTORA
DO GENE *DREB2A* EM DIFERENTES
GENÓTIPOS DO GÊNERO *COFFEA***

LAVRAS - MG

2011

GABRIEL SÉRGIO COSTA ALVES

**IDENTIFICAÇÃO E ANÁLISE DE POLIMORFISMOS NA REGIÃO
PROMOTORA DO GENE *DREB2A* EM DIFERENTES GENÓTIPOS DO
GÊNERO *COFFEA***

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Biotecnologia Vegetal, área de concentração em Biotecnologia Vegetal, para obtenção do título de Mestre.

Orientador
Dr. Alan Carvalho Andrade

Coorientador
Dr. Pierre Roger René Marraccini

LAVRAS-MG

2011

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca da UFLA**

Alves, Gabriel Sérgio Costa.

Identificação e análise de polimorfismos na região promotora do gene *DREB2A* em diferentes genótipos do gênero *Coffea* / Gabriel Sérgio Costa Alves. – Lavras : UFLA, 2011.

103 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2011.

Orientador: Alan Carvalho Andrade.

Bibliografia.

1. Marcadores moleculares. 2. SNPs. 3. Tolerância à seca. 4. Promotor. 5. Engenharia genética. I. Universidade Federal de Lavras. II. Título.

CDD – 583.520487328

GABRIEL SÉRGIO COSTA ALVES

**IDENTIFICAÇÃO E ANÁLISE DE POLIMORFISMOS NA REGIÃO
PROMOTORA DO GENE *DREB2A* EM DIFERENTES GENÓTIPOS DO
GÊNERO *COFFEA***

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Biotecnologia Vegetal, área de concentração em Biotecnologia Vegetal para a obtenção do título de Mestre.

Aprovada em 25 de Fevereiro de 2011.

PhD.Luciano Vilela Paiva

UFLA

PhD. Antônio Chalfun Júnior

EMBRAPA

Dr. Alan Carvalho Andrade
Orientador

LAVRAS - MG

2011

Dedico

A todos que
contribuíram das mais variadas
formas para a concretização
deste trabalho.

“Cada um tem de mim exatamente o que cativou, e cada um é responsável pelo que cativou, não suporto falsidade e mentira, a verdade pode machucar, mas é sempre mais digna. Bom mesmo é ir a luta com determinação, abraçar a vida e viver com paixão. Perder com classe e vencer com ousadia, pois o triunfo pertence a quem se atreve e a vida é muito para ser insignificante. Eu faço e abuso da felicidade e não desisto dos meus sonhos. O mundo está nas mãos daqueles que tem coragem de sonhar e correr o risco de viver seus sonhos." Coragem..Coragem..Coragem é não buscar desculpas para ser feliz !

(Charles Chaplin)

RESUMO

Apesar de alguns estudos em fisiologia vegetal resultarem em uma melhor compreensão dos mecanismos envolvidos na tolerância à seca em cafeeiro, ainda é escasso o conhecimento acerca das alterações metabólicas e moleculares envolvidos na resposta do cafeeiro às condições de estresse hídrico. Neste sentido, estudos recentes já resultaram na identificação de vários genes candidatos apresentando expressão diferencial entre genótipos contrastantes para essa característica. Dentre esses, se encontram genes integrantes da via de resposta ao estresse hídrico, previamente identificados em outras espécies vegetais, tais como os fatores de transcrição *DREB*. Resultados prévios obtidos, analisando-se a expressão relativa do gene *DREB2A* em genótipos de *C. canephora*, tolerante (Clone 14) e sensível (Clone 22) demonstraram expressão diferencial destes genótipos, em plantas submetidas ao estresse hídrico. Desta forma, o objetivo deste trabalho consistiu na identificação e análise de polimorfismos na região promotora do gene *DREB2A* destes genótipos (Clone 14 e 22) de *C. canephora*, além de outros diferentes genótipos do gênero *Coffea* (*C. eugenioides*, e *C. arabica*). Com os resultados obtidos foi possível identificar os elementos *cis* de regulação, e avaliar a ocorrência de polimorfismos na região promotora do gene *DREB* de cafeeiro. Os polimorfismos encontrados também permitiram a identificação de diferentes haplótipos nos genótipos estudados. A distribuição dos alelos deste locus nos genótipos sugere a segregação dos alelos de *C. canephora* e *C. eugenioides* nos genótipos de *C. arabica*, corroborando com resultados obtidos em estudos genealógicos anteriores. A presença de elementos *cis* de regulação no promotor *DREB2A* para fatores de transcrição da via dependente de ABA demonstram a possibilidade de atuação conjunta das duas vias (dependente e independente de ABA), na regulação do gene *DREB2A*. A engenharia de construções utilizando o vetor binário pBI121, com diferentes segmentos da região promotora do gene *DREB2A* identificados, visa a uma melhor caracterização *in vivo*, da participação dos elementos *cis* e da influência dos polimorfismos observados na funcionalidade deste promotor em relação às respostas do cafeeiro ao estresse hídrico.

Palavras-chave: Marcadores moleculares. SNPs. Tolerância à seca. Promotor. Engenharia genética.

ABSTRACT

Although some studies in plant physiology have resulted in a better understanding of the mechanisms involved in drought tolerance in coffee, there are still lack of information about the metabolic and molecular mechanisms involved in the response of coffee to water stress conditions. In this regard, recent studies have identified several candidate genes showing differential expression between contrasting genotypes for this trait. Previous results obtained by analyzing the relative expression of the gene *DREB2A* in genotypes of *C. canephora*, tolerant (clone 14) and sensitive (clone 22) showed differential expression of these genotypes in plants subjected to water stress. Thus, the aim of this study was to identify and analyze the polymorphisms within the promoter region of the gene *DREB2A* in those two clones of *C. canephora*, and other genotypes of *Coffea* genus (*C. eugenoides* and *C. arabica*). With the results achieved it was possible to identify the *cis* regulatory elements, and evaluate the occurrence of polymorphisms in the promoter region of the coffee gene *DREB2A*. The polymorphisms found have also allowed the identification of different haplotypes among those genotypes and the allelic distribution of these loci suggests the segregation of *C. canephora* and *C. eugenoides* alleles in the *C. arabica* genotype. The presence of *cis* regulatory elements from TFs belonging to ABA – dependent pathway in the promoter *DREB2A* suggest the possibility of cross-talk between the two network (ABA – independent and – dependent) response in the regulation process of *DREB2A* gene. The engineering of gene cassettes using the binary vector pBI121 combined with different segments of the gene promoter *DREB2A*, isolated from the clones 14 and 22 of *C. canephora* aim at characterizing *in vivo* the involvement of *cis* elements and the influence of the polymorphism found in the functionality of this promoter upon the coffee response to water stress.

Keywords: Molecular markers. SNPs. Drought tolerance. Promoter. Genetic engineering.

LISTA DE ILUSTRAÇÕES

- Figura 1. Média da produção cafeeira, no Brasil, em milhões de sacas/ano durante os últimos 10 anos. (Fonte: MAPA - SPAE/CONAB).....21
- Figura 2. Demonstrativo da participação das produções de *C. arabica* e *C. canephora* na produção total de café.....22
- Figura 3. (A) Heredograma genealógico do *Coffea arabica* anfidiplóide. (B) Origem e diversificação da espécie *C. arabica*.....24
- Figura 4. Grupos de distribuição da diversidade de *Coffea canephora*. (Fonte: Montagnon 2000)25
- Figura 5. Fluxograma genealógico das variedades e cultivares de *C. arabica* mais conhecidas e/ou comercializadas.26
- Figura 6. Cadeia de eventos moleculares na resposta das plantas aos estresses abióticos: Percepção e transferência de sinal, controle da transcrição e mecanismos de respostas aos estresses abióticos, adaptado de Wang, Vinocur e Altman (2003).28
- Figura 7. Redes transcricionais reguladoras de sinais de estresse abiótico e da expressão gênica, adaptado de Shinozaki e Yamaguchi-Shinozaki (2007)..30
- Figura 8. Esquema representativo da estratégia utilizada para isolar e amplificar a região promotora do gene *DREB2A*..43
- Figura 9. Mapa do vetor pCR®II TOPO TA Cloning (Invitrogen). (Fonte: Manual Invitrogen).....45
- Figura 10. Amplificação da região promotora do gene *DREB2A* dos clones 14 e 22 de *C. canephora* var. Conilon com os *primers* DPProx, DPMed, DPDist e DPR. Esquema ilustrativo das regiões do promotor do gene *DREB2A* amplificadas visando à construção do cassete gênico.....51

- Figura 11. Perfis de expressão do gene *DREB2A* em folhas dos clones 14,73,120 e 22 de *C. canephora* var. Conilon com Irrigação (I) e sem Irrigação (NI), obtidos por meio da técnica de qPCR.....56
- Figura 12. Representação da estratégia de alinhamento dos *hits* obtidos por Blast da sequência EST do gene *DREB2A* no banco de dados de Shotgun do DNAG de *C. canephora*.....57
- Figura 13. Amplificação da região promotora nos diferentes genótipos da tabela 4 (Apêndice C). Estão representados, de cima para baixo, os amplicons S1, S2 e S1S2 usados para identificação da região promotora por sequenciamento. O marcador molecular de 1KB permite a confirmação do tamanho esperado dos fragmentos 837, 952 e 1547 bp, respectivamente.58
- Figura 14. Extração plasmidial dos clones selecionados para o sequenciamento. Confirmação da presença das formas multiméricas do vetor clonado. .59
- Figura 15. A estratégia de sequenciamento adotada para identificação das diferentes regiões que integram o promotor *DREB2A*. Foram sequenciados 16 clones para cada genótipo e região amplificada. Representação gráfica das 48 sequências obtidas para o genótipo Obatã.60
- Figura 16. Alinhamento da região S1 do promotor *DREB2A*. Regiões polimórficas entre os genótipos estão enquadrados em vermelho. Motivos de regulação *cis* do DNA, região consenso promotora (GATA-BOX e CATA-BOX), elemento inicial de resposta a desidratação (ERD1) e elemento de resposta ao etileno (ERE).63
- Figura 17. Alinhamento da região S1 do promotor *DREB2A*. Regiões polimórficas entre os genótipos estão enquadrados em vermelho. Motivos de regulação *cis* do DNA, elemento de resposta ao ácido abscísico (ABRE) e elemento de resposta ao TF MYC (MYCRS).....64

Figura 18. Alinhamento da região S1 do promotor <i>DREB2A</i> . Regiões polimórficas entre os genótipos estão enquadrados em vermelho. Motivos de regulação <i>cis</i> do DNA, região consenso promotora (CGCG-BOX) e elemento de resposta ao TF MYB (MYBRS).	65
Figura 19. Alinhamento da região S2 do promotor <i>DREB2A</i> . Motivos de regulação <i>cis</i> do DNA, elemento de resposta ao TF MYC (MYCRS) e elemento de resposta aos FTs ERF (família de fatores de transcrição do DREB).....	67
Figura 20. Alinhamento da região S2 do promotor <i>DREB2A</i> . Regiões polimórficas entre os genótipos estão enquadrados em vermelho. Motivos de regulação <i>cis</i> do DNA dos fatores de transcrição da família ERF, a qual está incluído o DREB (ERF)..	68
Figura 21. Alinhamento da região S2 do promotor <i>DREB2A</i> . Início da região codante +1 da proteína DREB.....	69
Figura 22. Alinhamento da região S2 do promotor <i>DREB2A</i> . Região codante da proteína DREB.	70
Figura 23. Amplificação da região promotora nos genótipos clone 14 e 22 com os <i>primers</i> DPProx, DPMed, DPDist e DPR. Estão representados os amplicons Dreb Proximal, Dreb Mediano, Dreb Distal de cada um dos clones escolhidos (14A06, 14C06 e 22A08). O marcador molecular de 1KB permite a confirmação do tamanho esperado dos fragmentos 765, 1118 e 1469 bp, respectivamente	71
Figura 24. Esquema ilustrativo da estratégia adotada para construção dos cassetes gênicos utilizando o vetor binário pBI121.	72
Tabela 1. Lista dos diferentes genótipos usados para obtenção do material genético visando à identificação da região promotora do gene <i>DREB2A</i>	40

Tabela 2. Primers utilizados para as ampliações dos DNAs dos 12 genótipos estudados. Onde (a), (b) e (c) correspondem aos pares de primers utilizados para amplificação do promotor <i>DREB2A</i>	43
Tabela 3 Primers utilizados para as ampliações dos DNAs dos genótipos Clone 14 e 22. Onde (a), (b) e (c) correspondem aos pares de primers utilizados para amplificação do promotor <i>DREB2A</i>	50
Tabela 4. Lista de genótipos utilizados para verificar a conservação da região complementar ao primers.	10

SUMÁRIO

RESUMO.....	7
1 INTRODUÇÃO.....	17
2 REFERENCIAL TEÓRICO	20
2.1 Importância Econômica da Produção Cafeeira	20
2.2 Origem da Cultura Cafeeira.....	22
2.3. Estresse Hídrico (aspectos gerais)	26
2.3.1 Respostas ao estresse abiótico em plantas	27
2.3.2 Respostas ao estresse hídrico em <i>Coffea</i>	32
2.3.3 Estudos moleculares da tolerância à seca em cafeeiro.....	33
2.4 Marcadores moleculares e SNPs (Single Nucleotide Polymorphisms, INDELS – Insertion/Deletions).....	34
3 MATERIAL E MÉTODO.....	38
3.1 Material vegetal	38
3.2 PCR quantitativo em tempo real (qPCR):.....	39
3.3 Análises <i>in silico</i>	40
3.3.1 Extração de DNA	41
3.3.2 Análise quantitativa e qualitativa das alíquotas de DNA	42
3.4 Identificação da Região Promotora do gene <i>DREB2A</i>	42
3.4.1 Arquitetura dos primers para amplificação da região promotora do gene <i>DREB2A</i> nos diferentes genótipos.	42
3.4.2 Reação em cadeia da polimerase (PCR) e amplificação do DNAg	44
3.4.3 Ligação dos produtos de PCR e clonagem no vetor.....	45
3.4.4 Transfecção bacteriana	46
3.4.5 Extração de DNA plasmidial	46
3.4.6 Sequenciamento	48
3.4.7 Análise das sequências.....	49
3.5 Caracterizações da Região Promotora do gene <i>DREB2A</i>	49
3.5.1 Arquitetura dos primers para amplificação da região promotora do gene <i>DREB2A</i> nos genótipos de <i>C. canephora</i> clone 14 e clone 22, visando à construção do vetor binário.	49
3.5.2 Reação em cadeia da polimerase (PCR) e amplificação do DNAg	51
3.5.3 Subclonagem dos produtos de PCR e digestão.....	52
3.5.4 Digestão do vetor binário	53
3.5.5 Ligação dos produtos de PCR ao vetor binário	53
3.5.6 Sequenciamento	54
4 RESULTADOS E DISCUSSÃO.....	55
4.1 Expressão do gene <i>DREB2A</i> por PCR quantitativo em tempo real q(PCR)	55
4.2 Isolamento e identificação da região promotora	56
4.3 Análises dos polimorfismos na região promotora do gene <i>DREB2A</i>	61
4.4 Análise dos elementos <i>cis</i> -reguladores do DNA	62
4.5 Caracterização da região promotora.	70
5. CONCLUSÕES.....	73
REFERÊNCIA BIBLIOGRÁFICA	74
APÊNDICE	1

LISTA DE ABREVIATURAS

ABA	Fitohormônio ácido abscísico
ABI3/VP1	proteínas com domínios ABI3/VP1
ABRE	<u>A</u> BA <u>R</u> esponsive <u>E</u> lement
AFLP	Amplified fragment length polymorphism
AP2/EREBP/ERF	proteínas com domínios AP2/EREBP/ERF
ASAPs	Allele-specific associated primers
ATHB	HD-ZIP Responsive Element
bHLH	proteínas básicas de hélice-volta-hélice
bZIP	proteínas região básica/zíper de leucina
CAPS	Cleaved amplified polymorphic sequences
DAMD-PCR	Direct amplification of minisatellite DNA markers
DNAg.	DNA genômico
DREB	Dehydration responsive element binding
ERD1	<u>E</u> arly <u>R</u> esponsive to <u>D</u> ehydration 1
ERE	<u>E</u> thylene <u>R</u> esponsive <u>E</u> lement
ERF/DRE	<u>D</u> ehydration <u>R</u> esponsive <u>E</u> lement
EST	Expressed sequence tag markers
FTs.	Fatores de transcrição
GUS.	Gene da beta-glucuronidase
HD-Zip	proteínas homeodomínio-zíper de leucina
ISSR	Inter simple sequence repeat markers
M.	molar
min.	minutos
mL.	mililitros
µg.	micrograma
µl.	microlitros
MPa.	Megapascal
MYB	proteínas com domínio do oncogene Myb
MYBRS	<u>MYB</u> <u>R</u> esponsive <u>B</u> inding <u>S</u> ite
MYCRS	<u>MYC</u> <u>R</u> esponsive <u>B</u> inding <u>S</u> ite
ng.	nanogramas
NOS	Gene da nopalina sintase
pb.	pares de base
PCR.	reação em cadeia da polimerase
pd.	<i>pre-dawn water potential</i> (potencial hídrico de antemanhã)
PLACE	Banco de Dados Plant <i>Cis</i> -acting Regulatory DNA elements
PlantCARE	Banco de Dados Plant <i>Cis</i> -Acting Regulatory Elements
RAMPO	Randomly amplified microsatellite polymorphisms
RAPD	Randomly-amplified polymorphic DNA markers
RFLP	Restriction Fragment Length Polymorphism

RLGS	Restriction Landmark Genomic Scanning
rpm.	rotações por minuto
SCAR	Sequence characterized amplified regions for amplification of specific band
seg.	segundos
SNPs	Single nucleotide polymorphism (polimorfismo de base única)
SSCP	Single strand conformation polymorphism
SSRs	Simple sequence repeats (microsatélite ou sequência de repetição simples)
STMS	Sequence-tagged microsatellite site markers
STS	Sequence-tagged sites

1 INTRODUÇÃO

Entre os fatores ambientais, a demanda por água é a maior limitação para a qualidade e produtividade vegetal, influenciando diretamente a produção econômica e o suprimento alimentar. Eficientes tecnologias de irrigação têm reduzido a disparidade entre o potencial e o rendimento na produção e qualidade, entretanto, devido à escassez de água em várias regiões e países, o melhoramento genético visando ao aumento da produção em cultivares com baixa disponibilidade de água faz-se necessário como uma solução sustentável e economicamente viável (Gentzbittel et al., 2009).

Historicamente, climas amenos e índices pluviométricos bem distribuídos ao longo do ano são característicos das regiões de cultivo da cultura cafeeira (Sylvain, 1955). Condições climáticas desfavoráveis e a crescente expansão da cafeicultura para regiões marginais, onde tais condições intensificam-se, contribuem significativamente para a redução da produtividade, tornando a seca o principal estresse abiótico que afeta a produção de café nos países cafeicultores. Em regiões de cultivo de café sem irrigação, períodos de seca intensa podem representar uma redução de até 80% da área de plantio (DaMatta and Ramalho, 2006). A difícil reprodutibilidade das condições ideais de cultivo requer algumas vezes o desenvolvimento de métodos biotecnológicos para garantir a sustentabilidade da produção cafeeira (Andrade, 1998).

Atualmente, muitas pesquisas têm objetivado a compreensão dos mecanismos de tolerância dos vegetais ao estresse hídrico. Diferentemente da resistência aos estresses bióticos, cuja maioria é dependente de características monogênicas, as repostas genéticas aos estresses abióticos são multigênicas, e assim mais complexas e difíceis de controlar (Vinocur and Altman, 2005). O déficit hídrico desencadeia respostas variadas nas plantas, tais como alterações na expressão gênica, acúmulo de compostos de baixo peso molecular

(metabólitos ou componentes osmóticos) e a síntese de proteínas específicas (proteínas hidrofílicas, chaperonas, etc.). A estratégia básica de engenharia genética para tolerância à seca visa à introdução de genes funcionais envolvidos diretamente nesses eventos (Umezawa et al., 2006). Dessa forma, a escolha de promotores estresse induzidos, têm sido relevante para o sucesso da transferência gênica e expressão do transgene (Dale et al., 2002). Os efeitos transcricionais quantitativos e qualitativos ocasionados por promotores podem ser ajustados benéficamente com a seleção do promotor mais apropriado considerando-se as características da planta e o tipo de transgene. Genes induzidos por estresse têm contribuído para a caracterização de promotores valiosos. Tais promotores são potencialmente úteis na engenharia genética para a expressão direcionada sob condições controladas (Cruces, 2004), evitando os efeitos secundários indesejáveis na ausência do estresse e o silenciamento gênico às vezes resultante da utilização de promotores constitutivos.

A transformação genética em plantas modelos, tais como tabaco e *Arabidopsis thaliana*, demonstraram o envolvimento principalmente de fosfatases, proteínas quinases e fatores de transcrição na resposta à desidratação (Bray, 2004; Rizhsky et al., 2004; Chen and Zhu, 2004). Estudos recentes em *Coffea canephora* sugeriram a participação de alguns genes no aumento da tolerância dessa espécie ao estresse abiótico (Marraccini et al., 2009). Contudo, a manipulação desses genes em plantas perenes ainda é rara devido à grande dificuldade de se estabelecer metodologias eficientes e rápidas para transformação.

O objetivo deste trabalho foi (i) identificar e isolar a região promotora do gene *DREB2A* em diferentes genótipos do gênero *Coffea* visando analisar as características deste promotor, e assim, tentar elucidar a influência da sua regulação na adaptação ao déficit hídrico, e (i) engenheirar um cassete gênico

que auxiliará na avaliação da participação das regiões deste promotor na adaptação fenotípica.

2 REFERENCIAL TEÓRICO

2.1 Importância Econômica da Produção Cafeeira

A comercialização cafeeira mundial movimenta mais de 91 bilhões de reais por ano, sendo que a produção cafeeira mundial foi de 120 milhões de sacas em 2009. No mercado ativo, o café ocupa o segundo lugar entre as *commodities* comercializadas, gerando cerca de meio bilhão de empregos. Maior produtor mundial de café, o Brasil é responsável por 32,88 % da produção de grãos de café, com 48 milhões de sacas, e por 31,7 % da exportação mundial, movimentando aproximadamente 7 bilhões de reais por ano (fonte: ABIC ano 2010) (Figura 1). As estimativas deste ano projetam uma colheita entre 41,89 e 44,73 milhões de sacas de 60 quilos do produto beneficiado, o que representaria a maior produção dentre os anos da baixa bienalidade (Oliveira et al., 2011) (Figura 2). Além de maior produtor mundial, o Brasil possui o segundo mercado consumidor do mundo, apresentando um crescimento de 4,03% neste último ano alcançando o maior consumo *per capita* já registrado (fonte: ABIC ano 2010).

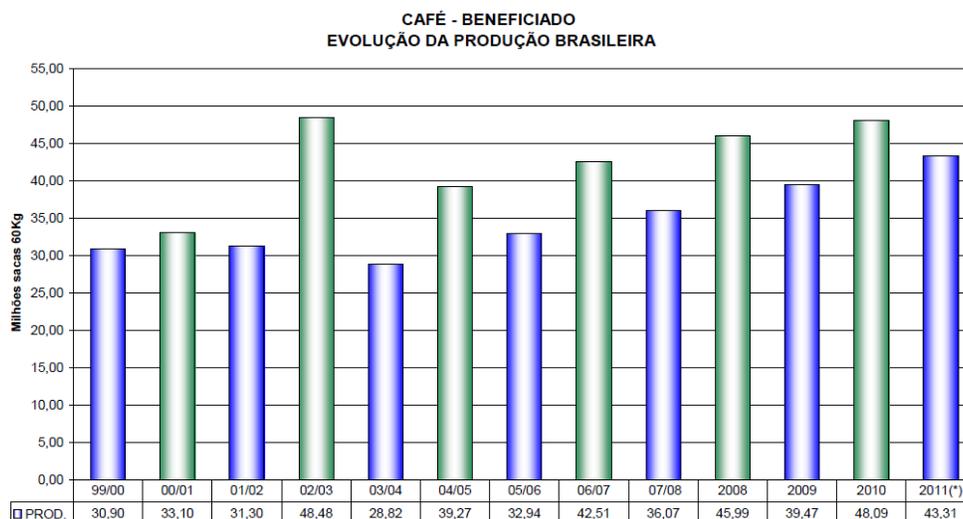


Figura 1. Média da produção cafeeira, no Brasil, em milhões de sacas/ano durante os últimos 10 anos. A estimativa da produção de 2011 projeta este ano como a maior produção entre os anos de baixa bienalidade.

(Fonte: MAPA - SPAE/CONAB)

Apenas duas espécies cafeeiras, dentre as cultivadas, apresentam importância econômica, *Coffea arabica* (Arábica) e *Coffea canephora* (Robusta) e correspondem a 76% e 24% da produção mundial, respectivamente (COMPANHIA NACIONAL DE ABASTECIMENTO - CONAB, 2010). O baixo teor de cafeína presente no grão Arábica propicia uma bebida de qualidade superior, mais apreciada no mercado mundial. O café Robusta ou Conilon, amplamente distribuído na África e Ásia, produz uma bebida de qualidade inferior ao Arábica. Entretanto, sua extensa base genética contribui para melhor adaptação às condições climáticas e maior resistência às doenças, minimizando as oscilações e diminuindo os custos da produção. Os grãos desta espécie apresentam maior teor de cafeína e ácidos, tendo como mercado principal as indústrias de café solúvel e, sendo também utilizados na composição dos *blends* (Leroy et al., 2006).

**CAFÉ BENEFICIADO
COMPARATIVO DE PRODUÇÃO
(Em milhões de sacas)**

SAFRA	2010	2011	VARIAÇÃO %
ARÁBICA	36,82	30,96 a 33,17	-15,9 a -9,9
ROBUSTA	11,27	10,93 a 11,56	-3,0 a 2,6
TOTAL	48,09	41,89 a 44,73	-12,9 a -7,0

Figura 2. Demonstrativo da participação das produções de *C. arabica* e *C. canephora* na produção total de café. A produção do *C. arabica* corresponde a aproximadamente 76% da produção cafeeira no Brasil, enquanto a participação do *C. canephora* é de aproximadamente 24%.

(Fonte: MAPA - SPAE/CONAB)

2.2 Origem da Cultura Cafeeira

Originário da África ou Ásia, a tribo *Coffeae* possui atualmente 11 gêneros. Entre estes se inclui o gênero *Coffea* classificado em dois subgêneros, *Coffea* subgênero *Coffea* e *Coffea* subgênero *Baracoffea* (Davis, 2003; Davis et al., 2005; Bridson, 1994; Bridson, 2003). O subgênero *Coffea* abrange 95 espécies com diferentes variedades incluindo as de maior importância econômica, *C. arabica* L. e *C. canephora* (Davis, 2003). Nativo das regiões montanhosas do sudoeste da Etiópia (Sylvain, 1955), do Platô Boma no Sudão (Thomas, 1942) e do Monte Marsabit no Quênia (Anthony et al., 1987), o *C. arabica* vem sendo cultivado no Iêmen pelos últimos cinco séculos e disseminou-se para o sudeste asiático e Europa, segundo registros que datam de 1700. No começo do século XVIII, a primeira plantação cafeeira na América Latina, em Suriname, foi introduzida pelos holandeses por progênes de uma única planta da Indonésia cultivada em Amsterdam e Paris (Carvalho, 1946; Chevalier and Dagrón, 1928). Essa primeira população cafeeira foi denominada

Typica, e teve sua inserção no Brasil pelo Pará. Outras plantas introduzidas provenientes do Iémen, passando pela Ilha de Réunion (antiga Ilha de Bourbon), originaram uma segunda população, chamada Bourbon (Haarer, 1956). Essas populações formaram a base para o surgimento de vários cultivares e foram classificadas em duas variedades distintas, *C. arabica* variedade *arabica*, conhecida como *C. arabica* variedade *typica* Cramer, e *C. arabica* variedade *bourbon* Choussy (Figura 3B).

O *C. arabica*, híbrido natural do *C. canephora* e *C. eugenioides*, é a única espécie cafeeira anfidiplóide ($2n=4x=44$) e autógama, com taxa de autofecundação de aproximadamente 90% (Carvalho et al., 1991; Lashermes et al., 1999) (Figura 3A). Originalmente, os eventos de autofecundação dessa variedade promoveram a homogeneização das estruturas genéticas (Charrier and Berthaud, 1985), porém novos cultivares estão sendo obtidos por programas de melhoramento diversificando a base genética (Etienne et al., 2002). Em meados da década de 30, no Brasil, um programa de melhoramento nas populações “Bourbon” produziu cultivares altamente produtivos que, posteriormente, substituiriam as plantações de Typica nas décadas de 50 e 60. Ainda na década de 30, foi descoberta uma variedade anã em uma plantação de Bourbon, cuja característica gênica dominante, chamada Caturra, produzia plantas de pequeno porte (Carvalho et al., 1991). Outras variedades foram obtidas ao longo dos anos, como o Mundo Novo e Catuaí, ambas obtidas do cruzamento entre as linhagens Typica e Bourbon. No Brasil, cerca de 90% do parque cafeeiro é formado pelos cultivares Mundo Novo (Bourbon x Typica) e Catuaí (Caturra x Mundo Novo) (Figura 5).

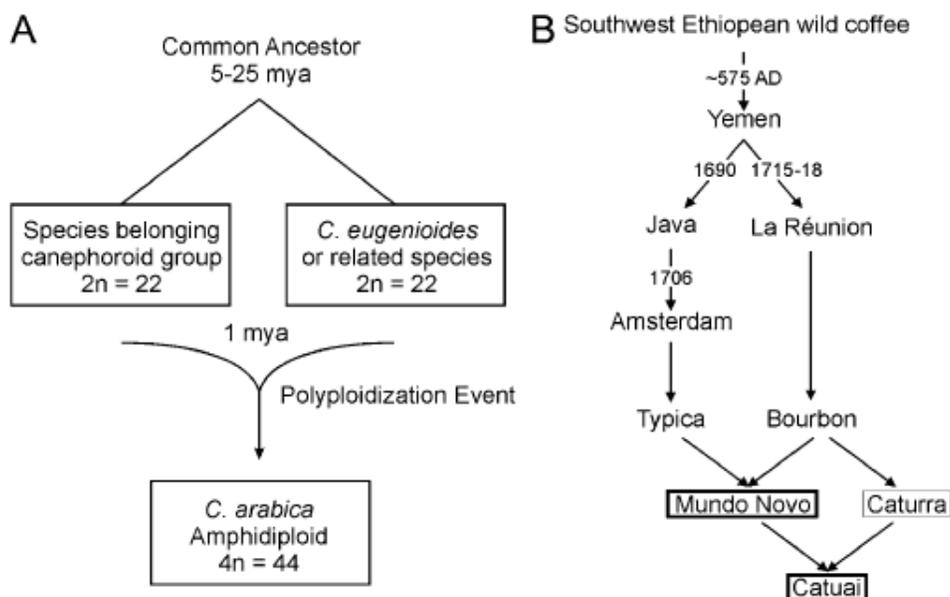


Figura 3. (A) Heredograma genealógico do *Coffea arabica* anfidiplóide. O cruzamento natural de duas espécies diplóides *C. canephora* e *C. eugenioides* originou o *C. arabica*. Os genomas parentais apresentam 97,5 % de similaridade. (B) Origem e diversificação da espécie *C. arabica*.

Estima-se que 80% das variedades cultivadas mundialmente derivam das bases genéticas tradicionais (Etienne et al., 2002). A introgressão de características genéticas pelo cruzamento entre espécies motivou a busca de híbridos naturais necessários para o retrocruzamento. O Híbrido de Timor (HdT), híbrido natural entre *C. arabica* ($2n=4x=44$) e *C. canephora* ($2n=2x=22$), é a principal fonte de variabilidade genética usada em programas de melhoramento genético visando a resistência às doenças (Carvalho, 2008). Ao contrário do Arábica, o *C. canephora* possui uma base genética mais diversificada decorrente da própria incompatibilidade de autofecundação. Estudos de diversidade genética em populações de *C. canephora* (Montagnon et al., 1992) demonstraram a ocorrência em dois grupos distintos classificados em função de suas origens geográficas: Congolês e Guineano.

O primeiro grupo, Congolês, é dividido em vários subgrupos dentre os quais dois de maior importância, um tolerante e outro sensível à seca: o subgrupo 1 (SG1) tolerante à seca, e o subgrupo 2 (SG2) sensível à seca. Originário de uma região com índices pluviométricos elevados e bem distribuídos ao longo do ano, os integrantes do grupo Congolês possuem susceptibilidade à seca (Figura 4). Outras características compõem este grupo, tais como resistência à ferrugem e menor teor de cafeína nos grãos, elevando a qualidade da sua bebida (Montagnon et al., 1992; Fazuoli, 2007; Dussert, 1999).

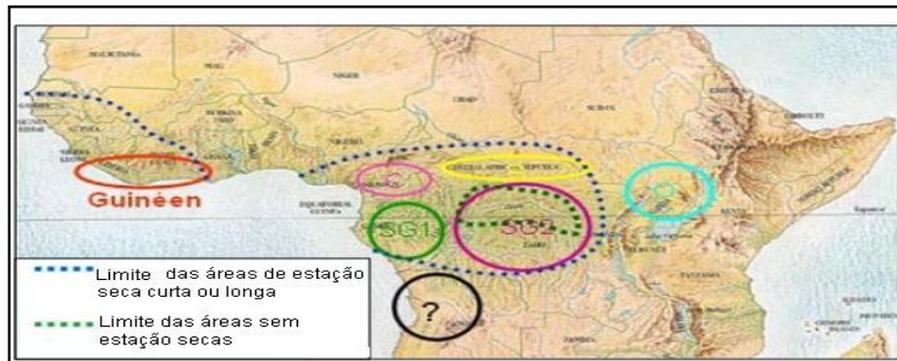


Figura 4. Grupos de distribuição da diversidade de *Coffea canephora*. As linhas tracejadas, azul e verde, delimitam regiões com distribuição pluviométrica diferente. O subgrupo SG1 originou-se em uma região de clima mais árido, já o subgrupo SG2 teve origem em uma região de clima tropical. (Fonte: Montagnon 2000)

As populações Guineanas estão localizadas na Costa do Marfim, região de baixo índice pluviométrico com ocorrência de estiagem, e apresentam como principais características a suscetibilidade à ferrugem, tolerância ao déficit hídrico e elevado teor de cafeína nos grãos produzindo uma bebida inferior ao do grupo Congolês (Montagnon et al., 1992; Fazuoli, 2007; Dussert, 1999).

Estresses bióticos e abióticos representam o maior desafio para a produção agrícola sustentável podendo reduzir em até 70% a produção. Dentre os estresses abióticos, o estresse hídrico decorrente de secas prolongadas e elevadas temperaturas é o maior responsável pela redução da produção agrícola (Agarwal et al., 2006). Regiões áridas, atualmente inadequadas para plantio, correspondem a 28% da superfície do planeta (Bray, 2004).

O déficit hídrico tem grande impacto no desenvolvimento e crescimento vegetal. As respostas das plantas a esse estresse podem ser percebidas em diferentes níveis, compreendendo mudanças fisiológicas até alterações em nível molecular. Considerando-se que as respostas são controladas pelo genoma, uma melhor compreensão dos genes expressos em resposta ao déficit hídrico é necessária para uma caracterização completa do mecanismo adaptativo à condição limitada de água.

Pesquisas científicas têm objetivado a elucidação dos mecanismos de sinalização nas células ativadas pelo estresse hídrico (Shinozaki et al., 2003; Xiong et al., 2002). De fato, o estresse abiótico é um estímulo complexo que altera diferentes, porém correlacionados atributos, tais como o balanço iônico e o estresse osmótico, dando às células informações indispensáveis a sua sobrevivência.

2.3.1 Respostas ao estresse abiótico em plantas

As reações desencadeadas pelo estresse hídrico nas plantas são inúmeras, dentre as quais estão: alteração da expressão gênica; acúmulo de metabólitos como o fitohormônio ácido abscísico (ABA) ou compostos osmóticos ativos; e a síntese de proteínas específicas (grandes proteínas hidrofílicas, proteínas antioxidantes, chaperonas, etc.) (Ramachandra-Reddy et al., 2004).

As respostas moleculares e celulares ao estresse abiótico obedecem a uma cascata de eventos que incluem a percepção do estímulo, transmissão do sinal ao citoplasma e núcleo, expressão gênica e por fim alterações metabólicas que resultam na manifestação da tolerância (Figura 6).

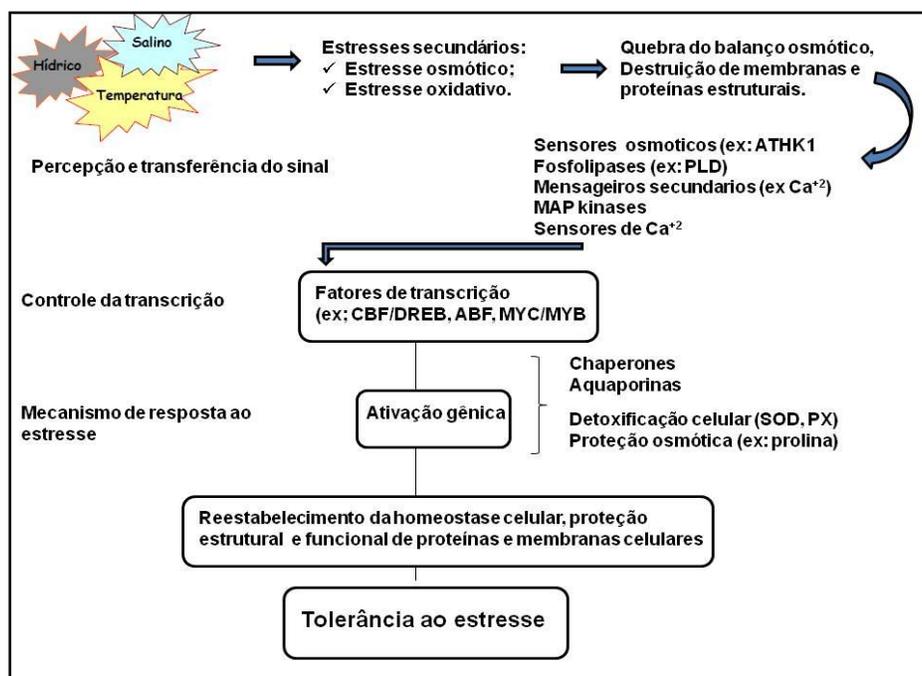


Figura 6. Cadeia de eventos moleculares na resposta das plantas aos estresses abióticos: Percepção e transferência de sinal, controle da transcrição e mecanismos de respostas aos estresses abióticos, adaptado de Wang, Vinocur e Altman (2003).

Análises de transcriptomas por microarranjo (Bohnert et al., 2001; Seki et al., 2001; Zhu et al., 2001) demonstraram que genes induzidos por estresse abiótico podem ser agrupados conforme a função dos seus produtos. Fazem parte do primeiro grupo as proteínas funcionais, como as proteínas de membrana que permitem o fluxo de água (proteínas de canais de água e transportadores de membrana); as enzimas fundamentais para biossíntese dos compostos osmóticos (prolina, betaína e açúcares); enzimas detoxificantes, que auxiliam no equilíbrio

do metabolismo bioquímico, fisiológico e celular (glutathione S-transferase, hidrolase, catalase, superóxido desmutase, ascorbato peroxidase, etc.); e outras proteínas de proteção a macromoléculas (proteínas LEA, osmotina, proteínas anticongelantes, chaperonas e proteínas de ligação ao mRNA, etc.). Já no segundo grupo estão incluídas as proteínas regulatórias ou fatores de transcrição (bZIP, MYC, MYB e DREB, etc.), as proteínas quinases (MAP quinase, CDP quinase, e proteínas quinases receptoras, ribossomais e reguladoras da transcrição, etc.) e proteinases (fosfotransferases e fosfolipases C, etc.) envolvidas na regulação da transmissão de sinais e expressão gênica.

O elevado número de genes regulatórios somados à expressão independente ou em alguns casos integrada ilustra a complexidade da regulação do processo de transcrição na resposta aos estresses abióticos. Entre tantas proteínas, os fatores de transcrição (FTs) compõem uma classe importante para a compreensão dos mecanismos de resposta ao estresse abiótico. Isso se deve ao fato dos FTs interagirem com elementos *cis*-regulatórios integrantes das regiões promotoras dos vários genes induzidos por esse estresse, e por consequência, corrigirem os genes envolvidos na manifestação da tolerância.

Na resposta ao estresse abiótico, os FTs podem ser classificados em vias de regulação, ABA – independente e ABA – dependente (Shinozaki and Yamaguchi-Shinozaki, 2000; Xiong et al., 2002; Thomashow, 1999) (Figura 7).

Estudos em *Arabidopsis* (Fowler and Thomashow, 2002) demonstraram a existência de algumas vias ABA – independente, e entre estas, uma de grande importância envolve o complexo de regulação DREB/CBF.

Os fatores de transcrição estão classificados em famílias de proteínas que compartilham um mesmo domínio de ligação ao DNA. Atualmente, os FTs associados à tolerância por desidratação são classificados em seis famílias: bZIP (proteínas região básica/zíper de leucina); AP2/EREBP/ERF (proteínas com domínios AP2/EREBP/ERF); HD-Zip (proteínas homeodomínio-zíper de

leucina); MYB (proteínas com domínio do oncogene Myb); bHLH (proteínas básicas de hélice-volta-hélice); e ABI3/VP1 (proteínas com domínios ABI3/VP1) (Kirch et al., 2003).

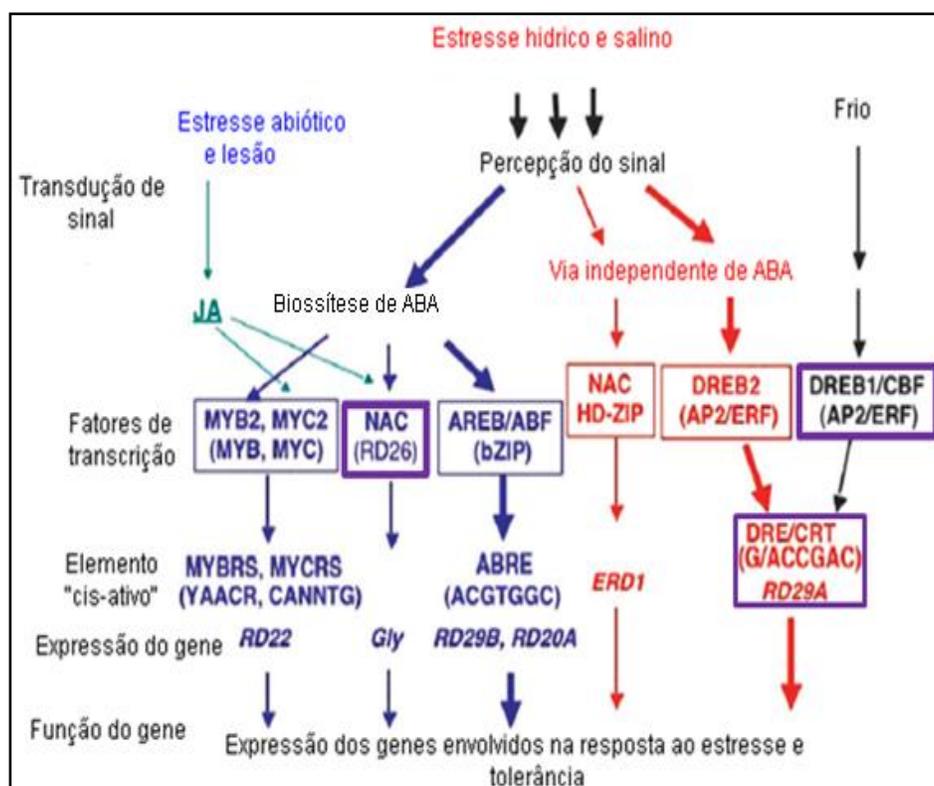


Figura 7. Redes transcricionais reguladoras de sinais de estresse abiótico e da expressão gênica, adaptado de Shinozaki e Yamaguchi-Shinozaki (2007). Pelo menos seis vias de transdução de sinal existem e estão relacionadas ao estresse hídrico, alta salinidade e ao frio, fornecendo respostas ao estresse: três são dependentes de ABA (importantes na adaptação à desidratação) e três são independentes de ABA. MYB2 e MYC2 são fatores de transcrição (FT) que agem na via de ABA-dependente e possuem a função de ativar a expressão do gene RD22. O fator de transcrição NAC RD26 está envolvido na via dependente de ABA. A sequência DRE é importante para a regulação da expressão do gene RD29A induzido na resposta ABA-independente ao frio e a desidratação.

Em condições de desidratação, ocorre um aumento da biossíntese do ABA, ativando a expressão de dois genes regulatórios ABA – dependentes, os

sistemas bZIP/ABRE e MYC/MYB (Abe et al., 1997; Uno et al., 2000). Os elementos de regulação *cis* do DNA, ABRE, apresentam motivos de ligação ao domínio básico/zíper de leucina, estrutura presente nos FTs da família bZIP. As proteínas bZIP fazem parte da sinalização das vias de defesa à luz UV, salinidade, seca (Choi et al.; Jakoby et al., 2002; Uno et al., 2000) e ao ácido salicílico (Zhang et al., 1999).

As proteínas MYB desempenham funções importantes no metabolismo de fenilpropanoides, no controle da estrutura celular, nas vias de sinalização dos reguladores de crescimento e na resposta ao estresse abiótico (Martin and Paz-Ares, 1997; Abe et al., 2003).

A classe dos fatores de ligação aos elementos de resposta ao etileno (ERF), exclusiva em plantas, desempenha um papel fundamental na resposta aos estresses ambientais. As ERFs apresentam um domínio conservado de 58-59 resíduos de aminoácido (domínio ERF) capaz de se ligar a dois elementos *cis*, GCC Box e ao motivo de repetição C CRT/elemento de resposta à desidratação (DRE). O primeiro geralmente é encontrado em promotores de genes induzidos por patógenos e apresenta estimulação por etileno, o segundo participa da expressão de genes de resposta à desidratação e ao frio.

Incluem-se nessa família de FTs, os DREB (proteína de ligação ao elemento de resposta à desidratação) que integram a via de tradução dos sinais ABA–independente. Os fatores de transcrição DREB se dividem em duas classes: DREB1 participa da resposta ao frio; e DREB2, envolvido na resposta à desidratação. Entretanto, relatos de interações entre as vias já foram descritos (Chini et al., 2004).

Os primeiros cDNAs codificantes para proteínas de ligação ao DRE foram isolados de *Arabidopsis*, CBF1, DREB1A e DREB2A . A partir de então, a expressão dos genes *DREB* vem sendo investigada em várias espécies vegetais.

2.3.2 Respostas ao estresse hídrico em *Coffea*

A desidratação desencadeia uma série de alterações fisiológicas e metabólicas no cafeeiro. A fim de evitar perdas adicionais de água, a planta promove o fechamento dos estômatos resultando no bloqueio de dióxido de carbono (CO₂) e conseqüente alteração na taxa fotossintética. Nas espécies mais susceptíveis, o estresse oxidativo ocasiona a desintegração das membranas e promove a redução do conteúdo de clorofila (DaMatta, 2003). A redução do fluxo de vapor e da absorção de água e nutrientes pelo sistema radicular resulta na diminuição da produção (Matiello and Dantas, 1987; Kumar, 1999). A escassez de água no ambiente leva a perda do turgor celular, ao aumento da concentração do citossol e a conseqüente desidratação do protoplasma (Pinheiro, 2005).

Estudos prévios, no gênero *Coffea*, sugerem que os mecanismos fisiológicos responsáveis pela manifestação da tolerância em cafeeiro estão diretamente relacionados à sensibilidade dos estômatos ao déficit hídrico no solo ou na atmosfera (Pinheiro, 2005). A sensibilidade da condutância estomática em *C. arabica*, apresentando oscilações à variação de 1/3 do aporte de água no solo, faz dessa propriedade fisiológica um ótimo indicador do nível de água acessível à planta. Plantas de *C. canephora*, por sua vez, apresentam controle estomático menos responsivo às variações no nível de água do solo (Fahl, 2001). Tais diferenças podem ser explicadas considerando-se a origem das espécies. A Etiópia, região nativa do *C. arabica*, está situada entre 1600 a 2800 m de altura, possui umidade atmosférica baixa e com um período de estiagem de 4 a 5 meses por ano. Já o *C. canephora* é proveniente da bacia do Rio Congo, região de baixa altitude, umidade perto da saturação, clima tropical equatorial e elevado índice pluviométrico com distribuição homogênea das chuvas ao longo de 9-10 meses.

A tolerância à dissecação pode também estar relacionada ao controle do mecanismo de ajuste osmótico e estudos recentes têm investigado a participação do sistema antioxidante em cafeeiros. Segundo Lima (2002) a tolerância à seca em *C. canephora* poderia estar parcialmente associada ao aumento da atividade de enzimas antioxidantes.

2.3.3 Estudos moleculares da tolerância à seca em cafeeiro

Devido à característica perene da cultura cafeeira (Bouharmont and Awemo, 1979), os programas de melhoramento convencional demandam aproximadamente 30 anos para desenvolvimento de novos cultivares (Fazuoli, 2000).

Com o objetivo de se estudar a variabilidade genética e fornecer informações para programas de melhoramento genético, o Projeto Genoma Café (Vieira et al., 2006) viabilizou a criação de um banco de dados genômicos formado por sequências expressas (ESTs). Utilizando materiais genéticos provenientes de folhas, raízes, sementes e frutos, foram construídas várias bibliotecas, incluindo as representativas de condições específicas como estresse hídrico, embriogênese somática e outras. Acessos de vários genes expressos – mais de 200 mil sequências – foram submetidos ao banco de dados cafEST (Andrade, 2007; Vieira et al., 2006).

Ao todo, o Projeto Genoma Café EST, gerou 214.964 sequências provenientes de 37 bibliotecas de cDNA das espécies *C. arabica*, *C. canephora*, *C. racemosa*. Após processadas e analisadas, as sequências foram agrupadas em 17,982 clusters e 32,155 sequências únicas. Dessas, 22% não apresentaram similaridade significativa após o alinhamento com banco de dados do NCBI. Posteriores análises comparativas entre os bancos de dados de ESTs disponíveis para café, o banco NCBI e o cafEST demonstraram haver apenas 13% de

redundância nos unigenes das diferentes bases de dados (Lashermes et al., 2008).

Atualmente, o banco de ESTs vem sendo utilizado em vários projetos de pesquisa que visam: desenvolver novos e eficientes marcadores moleculares, que auxiliem na análise da diversidade genética disponível nos programas de melhoramento; identificar e caracterizar genes de interesse agrônômico; compreender a regulação gênica e influência dos genes nas vias metabólicas; estudar os mecanismos de resposta a diferentes estresses; e finalmente acelera o melhoramento genético do cafeeiro (Caixeta, 2008; Lashermes et al., 2008).

Estudos, em andamento, têm objetivado a identificação e validação de genes candidatos (GC) selecionados a partir das informações extraídas do banco de dados cafEST. Análises de macroarranjos (Vinecky, 2009) identificaram 28 clusters diferencialmente expressos, nas bibliotecas representativas de estresse hídrico (SH2 e SH3) do banco de dados cafEST. Análises por PCR quantitativo demonstraram diferenças no perfil de expressão gênica, em café, de genes candidatos à resposta ao déficit hídrico (Alves, 2009).

2.4 Marcadores moleculares e SNPs (Single Nucleotide Polymorphisms, INDELs – Insertion/Deletions)

Todos os DNAs genômicos apresentam padrões para fragmentos específicos que conferem aos mesmos, identidades exclusivas. A combinação de vários sistemas de detecção de locus singulares para atribuir identidade exclusiva ao genoma chama-se “*DNA fingerprinting*”. Os marcadores moleculares, um desses sistemas de detecção, podem ser constituintes bioquímicos (metabólitos secundários em plantas) e macromoléculas, proteínas, e ácido desoxirribonucléico (DNA).

As aplicações dos marcadores moleculares de DNA são inúmeras: conservação de bancos de germoplasma; caracterização de germoplasmas (genotipagem); determinação da variabilidade genética, mapeamento e identificação de genes; estudos filogenéticos e evolutivos; entre outros. Os marcadores moleculares diferenciam-se pelo princípio da tecnologia aplicada para se determinar a variabilidade, por sua capacidade de indicar regiões polimórficas, reprodutibilidade, praticidade e custo.

Existem diferentes tipos de marcadores moleculares, tais como RFLP (Restriction Fragment Length Polymorphism), RLGS (Restriction Landmark Genomic Scanning), STS (Sequence-tagged sites), ASAPs (Allele-specific associated primers), EST (Expressed sequence tag markers), SSCP (Single strand conformation polymorphism), STMS (Sequence-tagged microsatellite site markers), DAMD-PCR (Direct amplification of minisatellite DNA markers), ISSR (Inter simple sequence repeat markers), RAPD (Randomly-amplified polymorphic DNA markers), SCAR (Sequence characterized amplified regions for amplification of specific band), CAPS (Cleaved amplified polymorphic sequences), RAMPO (Randomly amplified microsatellite polymorphisms), AFLP (Amplified fragment length polymorphism), e etc. Atualmente, os marcadores moleculares mais empregados em estudos de polimorfismo são os SNPs (polimorfismo de base única) e os SSRs (microsatélite ou sequência de repetição simples). A preferência pelos polimorfismos de modificações nucleotídicas é atribuída à alta frequência destes no genoma.

Os polimorfismos de base única são as variações genômicas mais ocorrentes. A análise de SNPs pode ser dividida em três etapas. A primeira etapa compreende na identificação e mapeamento dos polimorfismos em sequências gênicas ou genomas conhecidos. O segundo passo consiste na verificação de SNPs conhecidos na sequência de interesse. Na terceira etapa, procura-se estabelecer correlações entre os polimorfismos do genoma e manifestações

fenotípicas embasadas por diferentes estudos experimentais e análises computacionais (Liao and Lee, 2010). Os polimorfismos podem ser decorrentes de alterações, inserções ou deleções nucleotídicas. Nas alterações nucleotídicas por transição, ocorre substituição de uma base por outra de mesma natureza, ou seja, purina por purina (A – G) ou pirimidina por pirimidina (C – T). Já as alterações por transversão ocorrem quando há substituição de uma base por outra de natureza diferente, purina por pirimidina ou vice-versa (A – C/ A – T/ G – C/ G – T). Os polimorfismos por inserção ou deleção são denominados INDELS (“Inserção/Deleção”). A quantidade e distribuição dos polimorfismos de base única variam conforme a pressão de seleção. As mutações ocorrem tanto na região gênica codante, podendo alterar o produto da sua expressão, como na região promotora, tendo o potencial de afetar a regulação da expressão gênica (Liao and Lee, 2010).

Os polimorfismos regulatórios são classificados como *cis*-atuantes, sendo aqueles que afetam genes dentro ou próximo do loco polimórfico e *trans*-atuantes, aqueles polimorfismos que ocorrem em um gene, mas influencia na expressão de outro gene em um loco distante. Esses polimorfismos na região regulatória podem alterar a afinidade pelas sequências regulatórias ocasionando, assim, diferenças na expressão gênica.

Quando bem caracterizados e mapeados, os SNPs auxiliam na avaliação da diversidade genotípica das culturas. Nos últimos anos, a grande quantidade de informação genética obtida com as novas tecnologias de sequenciamento aliada às ferramentas bioinformáticas vem facilitando a identificação de SNPs e novos microssatélites (SSRs).

Em café, um estudo pioneiro de marcadores moleculares, realizado com a técnica RAPD no DNA de cloroplastos e mitocôndrias das espécies *C. arabica*, *C. canephora*, *C. congensis*, *C. liberica*, *C. eugenioides*, *C. excelsa*, *C. arabusta*, *Paracoffea ebracteolata*, consolidou as relações filogenéticas

atribuídas pelos critérios convencionais (Berthou et al., 1983). Desde então, várias técnicas vêm sendo utilizadas para a detecção de polimorfismos em nível de DNA. Recentes trabalhos demonstraram a importância das informações contidas em bancos de dados ESTs na identificação dos SNPs e na elucidação dos mecanismos de regulação gênica (Yanagui, 2009; Freire, 2010; Vidal et al., 2010). Freire (2010) descreveu a ocorrência de polimorfismos em genes envolvidos na resposta ao estresse hídrico em *C. arabica*, e também evidenciou diferenças no padrão de expressão de alguns genes integrantes da via de resposta à desidratação, quando expostos a diferentes níveis de estresse hídrico.

3 MATERIAL E MÉTODO

3.1 Material vegetal

Os clones tolerantes (14, 73 e 120) e sensível (22) à seca da variedade Conilon de *C. canephora* foram selecionados pelo INCAPER (Ferrão et al., 2000) e cultivados em casa de vegetação, com irrigação (I) ou sem irrigação (NI: condição de estresse hídrico nas folhas com $pd = -3.0$ MPa). Os cultivares Rubi e Iapar59, cultivados em condições de campo (Embrapa Cerrados, Planaltina-DF) com irrigação (I: $pd = -0.5$ MPa) ou sem irrigação (NI: $H_2O - 1.7$ MPa) no sistema de pivô central. Em todos os casos, o potencial de base (pd : “pre-dawn water potential”) foi medido usando uma bomba de Scholander. Para todas as plantas, as folhas foram coletadas, congeladas em nitrogênio líquido e conservadas na temperatura de $-80^{\circ}C$ até a extração dos RNA totais.

O material vegetal utilizado para isolar a região promotora do gene *DREB2A* foi obtido a partir de uma seleção de genótipos do gênero *Coffea*. Os genótipos foram escolhidos com o intuito de obter-se a maior cobertura possível da variabilidade genotípica existente entre as variedades e cultivares de café. A escolha dos genótipos baseou-se em alguns critérios como: a resistência e susceptibilidade ao estresse hídrico; a genealogia das variedades e cultivares; e a importância econômica. O IAPAR (Instituto Agrônomo do Paraná) juntamente com o CIRAD (Centre de coopération Internationale en Recherche Agronomique pour le Développement) cederam a Embrapa Recursos Genéticos e Biotecnologia – Cenargen, os DNAs genômicos (DNAg) isolados de folhas de *C. arabica* cv. Mundo Novo e das variedades Bourbon e Typica (coleção Etiópia do IAPAR), além de outras espécies como *C. eugenioides*, e *C. racemosa*. Os materiais biológicos vegetais, folhas, dos genótipos *C. canephora*, clones 14 e 22, e *C. arabica*, Catuaí 144, Iapar 59, Obatã, Rubi e Tupi, foram coletados no

banco de germoplasma localizado no campo experimental da Embrapa Cerrados. Desta forma, foram incluídos exemplares de cultivares de *C. arabica* sem introgressão recente de *C. canephora* (Bourbon, Catuaí, Mundo Novo, Rubi), acessos de *C. canephora* tolerante (clone 14) e sensível (clone 22) que apresentaram diferença no perfil de expressão para o gene *DREB2A*, cultivares comerciais de *C. arabica* com introgressão genética de *C. canephora* (Tupi e Obatã), e espécies distantes geneticamente, *C. eugenioides* e *C. racemosa*.

3.2 PCR quantitativo em tempo real (qPCR):

As amostras de RNA total extraído de folhas foram submetidos a um tratamento enzimático com DNaseI RNase “free” (Promega), e 1µg de cada amostra foi utilizado para a transcrição reversa, com a enzima ImPromII (Promega), conforme recomendações do fabricante. O cDNA fita-simples sintetizado, foi diluído (1/50) e testado por qPCR usando pares de *primers* específicos do gene *DREB2A*. Esses *primers* foram desenhados usando as sequências consenso de contigs disponíveis, na Base de Dados do Genoma Café (<https://alanine.cenargen.embrapa.br/cafEST>). As reações de amplificação foram realizadas com 1µl de cDNA fita-simples em um volume final de 10µl. O fluoróforo de detecção da amplificação em tempo real utilizado foi o SYBR-green (SYBRGreen qPCR Mix-UDG/ROX, Invitrogen) de acordo com as instruções do fabricante e a plataforma de detecção e análise das reações foi o aparelho ABI 7500-FAST (Applied Biosystems). Para cada amostra, os níveis da expressão foram normalizados com a expressão do gene *UBQ10* (www.sgn.cornell.edu SGN-U347154), o qual codifica para a ubiquitina, como controle endógeno. Os resultados obtidos foram analisados com o auxílio do programa SDS v. 2.0.5 (Applied Biosystems). Os níveis da expressão gênica foram apresentados na forma de quantificação relativa.

3.3 Análises *in silico*

As sequências de DNA genômico de um genótipo duplo haplóide de *Coffea canephora* obtidas por shotgun e sequenciadas pelo sistema de pirosequenciamento 454 foram organizadas em um banco de dados. Os primeiros 150 pares de bases da região codante do gene *DREB2A* serviram como sequência referência para as análises de BlastN no banco de dados 454. Foram realizados vários BLASTs objetivando a obtenção de contigs que cobrissem até aproximadamente 1500 pares de bases upstream a região codante do gene *DREB2A*. Análises *in silico* da sequência consenso dos contigs da região promotora, nos bancos de dados de elementos regulatórios Plant CARE (Plant *cis*-acting regulatory element) (Rombauts et al., 1999) e PLACE (Plant *cis*-acting regulatory DNA elements) (Higo et al., 1999), permitiu a identificação de motivos *cis*-regulatórios no DNA. A partir da sequência consenso de 1451 pares de base foram arquitetados *primers* para o isolamento dos promotores por amplificação do DNA genômico nos diferentes genótipos da tabela (1).

Tabela 1. Lista dos diferentes genótipos usados para obtenção do material genético visando à identificação da região promotora do gene *DREB2A*.

Genótipos	Espécie	Descrição
1. Bourbon	<i>C. arabica</i>	Acesso
2. Catuaí 144	<i>C. arabica</i>	Melhoramento
3. Clone 14	<i>C. canephora</i>	Melhoramento
4. Clone 22	<i>C. canephora</i>	Melhoramento
5. Eugenioides	<i>C. eugenioides</i>	Acesso
6. IAPAR 59	<i>C. arabica</i>	Melhoramento
7. Mundo novo	<i>C. arabica</i>	Melhoramento
8. Obatã	<i>C. arabica</i>	Cruzamento Natural
9. Racemosa	<i>C. racemosa</i>	Acesso
10. Rubi	<i>C. arabica</i>	Melhoramento
11. Tupi	<i>C. arabica</i>	Cruzamento Natural
12. Typica	<i>C. arabica</i>	Acesso

3.3.1 Extração de DNA

Folhas de cafeeiros dos cultivares de *C. arabica* cv. Catuaí 144, Iapar 59, Obatã, Rubi e Tupi foram coletadas, acondicionadas em tubos falcon, congeladas em N₂ líquido e armazenadas a -80 °C, para posterior extração de DNA genômico (DNAg). Os materiais utilizados para extração (cadinhos e pistilos de porcelana, microtubos, ponteiros e água destilada) foram esterilizados. As folhas de cafeeiro foram pulverizadas com N₂ líquido e 20 mg do material pulverizado foi transferido para um microtubo (1,5 ml).

A extração de DNAg foi realizada pelo método CTAB (SAGHAI-MAROOF et al., 1984) modificado de acordo com o protocolo do Laboratório Central de Biologia Molecular (LCBM) da Universidade Federal de Lavras (UFLA). Para a extração, foi adicionado 1 ml do tampão de extração CTAB pré aquecido a 65 °C [CTAB 2% (p/v); NaCl 2 M; Tris-HCl 100 mM, pH 8,0; EDTA 25 mM; b-mercaptoetanol 0,2% (v/v)] ao microtubo contendo o tecido pulverizado e, em seguida, a mistura foi homogeneizada em vórtex e incubados por 30 min. a 65 °C. Então, os microtubos foram centrifugados por 10 min. a 11000 rpm. Após a transferência da fase líquida para novos microtubos foi adicionado um volume igual de clorofórmio: álcool isoamílico 24:1 (p/v) para permitir a desnaturação e insolubilização protéica. A solução foi homogeneizada e centrifugada por 5 min. a 11000 rpm. A fase líquida foi novamente transferida para novos microtubos e 4 µL de RNase (100 µg/mL) foram adicionados. As amostras foram incubadas a 37 °C por 30 min. para permitir a digestão completa dos RNAs presentes. Em seguida, o DNA foi precipitado com a adição de 300 µL de álcool isopropílico e centrifugação a 13000 rpm durante 15 minutos. Após descartar o sobrenadante, o excesso de sais foi removido com a adição de 400 µL de etanol absoluto e centrifugação dos microtubos em rotação máxima por 3 minutos. As amostras foram secadas durante 3 min. e ressuspensas em 30 µL

de água Milli-Q autoclavada. A extração de DNA e as análises, quantitativa e qualitativa, foram realizadas no LCBM – UFLA.

3.3.2 Análise quantitativa e qualitativa das alíquotas de DNA

A integridade das amostras de DNA extraídas foi avaliada por eletroforese em gel de agarose 0,8% corado com 1 µL de Brometo de Etídeo (0,5 µg.mL⁻¹). O gel de agarose foi visualizado sob luz ultravioleta e a imagem foi captada pelo fotodocumentador da Loccus Biotecnologia. As amostras foram quantificadas em espectrofotômetro (Nanodrop[®] Spectrophotometer ND-1000). A qualidade foi avaliada pelo espectro (220-600nm) com a razão DO₂₆₀/DO₂₈₀.

3.4 Identificação da Região Promotora do gene *DREB2A*.

3.4.1 Arquitetura dos *primers* para amplificação da região promotora do gene *DREB2A* nos diferentes genótipos.

Com base na sequência consenso das análises *in silico*, dois pares de *primers* foram desenhados com o auxílio do programa PrimerQuestSM® (Integrated DNA Technologies - IDT) para a amplificação da sequência promotora do gene *DREB2A*. Os *primers* arquitetados para a amplificação da região promotora foram:

DrebS1F	forward	5'–
ACTCCTAGTAAGCGGCACGTTGTT–3'	DrebS1R	reverse 5'–
CCTTTCGTGGTTGTCTCTTGACCT–3'	DrebS2F	forward 5'–
TCGTGCATTCAACAGCACCGTCA–3'	DrebS2R	reverse 5'–
TGGCTTTGCAGGCATTGACTACG–3'		

Os *primers* foram utilizados na reação de PCR objetivando a amplificação de três fragmentos distintos da região promotora, da seguinte

forma: amplicon S1 (837pb), DrebS1F e DrebS1R, amplicon S2 (952pb), DrebS2F e DrebS2R, e amplicon S1/S2 (1560pb), DrebS1F e DrebS2R (Tabela 2).

Tabela 2. *Primers* utilizados para as ampliações dos DNAs dos 12 genótipos estudados. Onde ^(a), ^(b) e ^(c) correspondem aos pares de *primers* utilizados para amplificação do promotor *DREB2A*.

<i>Primer</i>	Sequências dos <i>Primers</i>	Amplicon (bp)
<i>DrebS1F</i> ^(a)	5'-ACTCCTAGTAAGCGGCACGTTGTT-3'	837
<i>DrebS1R</i> ^(a)	5'-CCTTTCGTGGTTGTCTCTTGACCT-3'	
<i>DrebS2F</i> ^(b)	5'-TCGTGCATTCAACAGCACCGTCA-3'	952
<i>DrebS2R</i> ^(b)	5'-TGGCTTTGCAGGCATTGACTACG-3'	
<i>DrebS1F</i> ^(c)	5'-ACTCCTAGTAAGCGGCACGTTGTT-3'	1560
<i>DrebS2R</i> ^(c)	5'-TGGCTTTGCAGGCATTGACTACG-3'	

O amplicon S1 abrange os pares de base -1438 ao -601 da região promotora. Já o amplicon S2 amplifica a região -843 a +109, abrangendo, assim, 843pb upstream à região codante e mais 109pb downstream ao códon inicial. O amplicon S1/S2 corresponde à amplificação da região -1438 a +109 do gene *DREB2A* (Figura 8).

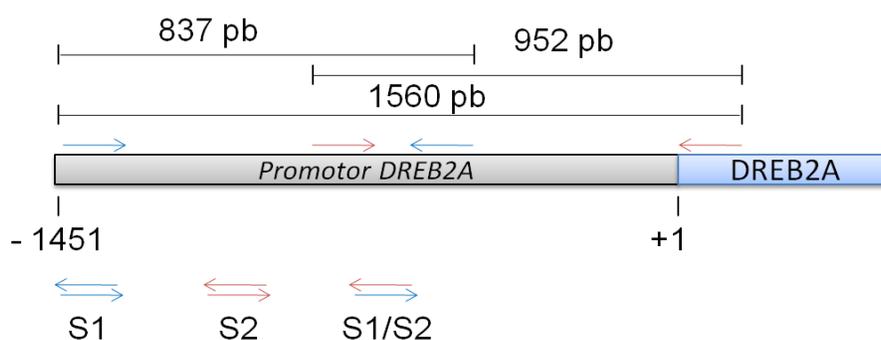


Figura 8. Esquema representativo da estratégia utilizada para isolar e amplificar a região promotora do gene *DREB2A*. Os amplicons obtidos S1, S2 e S1S2 e o tamanho dos fragmentos 837, 952 e 1547 pb são indicados.

A formação de potenciais estruturas secundárias e a especificidade dos *primers* foram avaliadas pelos programas OligoAnalyzer 3.1 (IDT Technologies) e FastPCR Professional 6.1 (Primer Digital Ltda) (Kalendar et al., 2009).

3.4.2 Reação em cadeia da polimerase (PCR) e amplificação do DNAg

Para a amplificação dos fragmentos foi utilizado 6.25 ng de DNAg de cada genótipo. As reações de PCR foram realizadas com um volume final de 50 μ L de solução usando a enzima Platinum® *Taq* DNA Polimerase (Invitrogen), segundo as especificações do fabricante. Essa enzima, assim como as demais *Taq* Polimerases, possui atividade transferásica terminal independente de fita molde que adiciona uma deoxiadenosina na extremidade 3' dos produtos de PCR, auxiliando no processo de clonagem em vetores. As reações foram incubadas no termociclador modelo Mastercycler® (Eppendorf) com o programa de termociclagem descrito a seguir: desnaturação inicial de 94 °C por 2 min., seguido de 35 ciclos de amplificação: 94 °C (desnaturação), por 30 seg.; 68 °C (anelamento), por 30 seg.; 72 °C (extensão), por 2 min. e de uma etapa final de polimerização por 10 min. a 72 °C. A temperatura de anelamento foi ajustada conforme a necessidade dos pares de *primers*. A análise qualitativa dos amplicons foi realizada por eletroforese em gel de agarose, como descrito anteriormente, e o tamanho dos fragmentos foi verificado com o marcador 1 kb DNA Ladder (Amresco). Em seguida, todos os produtos de PCR foram clonados.

3.4.3 Ligação dos produtos de PCR e clonagem no vetor

Os produtos de PCR foram clonados no vetor pCR[®]2.1 - TOPO pertencente ao TOPO TA Cloning[®] Kit Dual Promoter (Invitrogen) (Figura 9). As reações de ligação foram preparadas adicionando-se 0,5 µL de solução salina (salt solution 1.2M NaCl e 0.06M MgCl₂) e 0,5 µL de vetor a 2 µL do produto da PCR. As misturas foram incubadas à temperatura ambiente (22-23 °C) por 30 minutos e, em seguida, acondicionadas no gelo até a transfecção bacteriana.

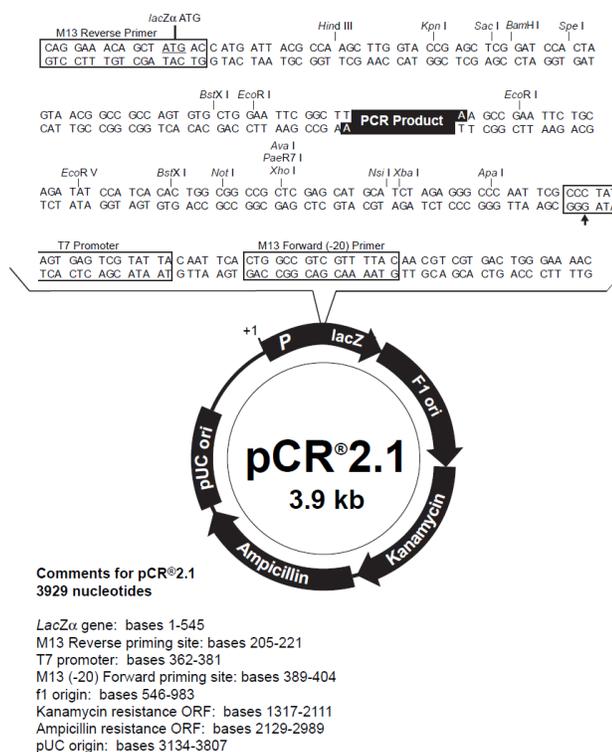


Figura 9. Mapa do vetor pCR[®]II TOPO TA Cloning (Invitrogen).
(Fonte: Manual Invitrogen)

3.4.4 Transfecção bacteriana

Células quimiocompetentes de *E. coli* DH5 α foram preparadas e utilizadas para transfecção bacteriana. Foram adicionados 1,5 μL de ligação a uma alíquota de 100 μL da preparação de célula competente, pré-descongeladas em gelo. A seguir, as células foram resfriadas em gelo por 30 min. e expostas a um choque térmico a 42 $^{\circ}\text{C}$ por 1 min. em banho seco seguido de um resfriamento a 4 $^{\circ}\text{C}$ por 2 minutos. A mistura DNA/célula foi ressuspensa em 250 μL de meio SOC, e incubada durante 1 hora a 37 $^{\circ}\text{C}$ com agitação constante (300 rpm, modelo do agitador horizontal). A seleção dos clones transformados foi realizada em placas de petri (J.Prolab) com meio LB sólido (acrescido de Ágar), contendo ampicilina (100 $\mu\text{g.mL}^{-1}$), X-Gal (40 $\mu\text{g.mL}^{-1}$) e IPTG (24 $\mu\text{g.mL}^{-1}$). As placas de petri contendo 100 μL da solução de cultura bacteriana foram incubadas por 16 horas em estufa a 37 $^{\circ}\text{C}$. Para cada transfecção, 16 colônias brancas selecionadas foram estriadas em uma nova placa de meio LB sólido, contendo ampicilina, X-Gal e IPTG nas concentrações indicadas previamente e, em seguida, foram inoculadas em placas de 96 poços de cultura tipo “*deep well*” (Axygen) contendo meio LB líquido acrescido de ampicilina (100 $\mu\text{g.mL}^{-1}$). As placas estriadas foram incubadas em estufa a 37 $^{\circ}\text{C}$ por 16 horas e, posteriormente, armazenadas a 4 $^{\circ}\text{C}$. As placas *deep well* foram incubadas a 37 $^{\circ}\text{C}$ com agitação constante a 300 rpm.

3.4.5 Extração de DNA plasmidial

As extrações do DNA plasmidial das colônias selecionadas foram realizadas em placas de 96 poços “*Deep well*” (Axygen). A placa “*Deep well*” foi centrifugada a 4000 rpm por 6 min. para precipitar as células bacterianas. O sobrenadante foi descartado, e uma lise inicial das células foi realizada

adicionando-se 240 μL de solução I (GET) em cada poço da placa e homogeneizando em vórtex até que as células fossem ressuspensas. A placa foi centrifugada a 4000 rpm durante 6 minutos. O sobrenadante foi novamente descartado, e adicionou-se mais 80 μL de solução I em cada poço da placa para garantir o rompimento completo das células. A solução foi homogeneizada e 60 μL da suspensão de células foram transferidos para uma placa com fundo U (Axygen). Em cada poço, foram adicionados 1 μL de RNase ($10 \text{ mg}\cdot\text{mL}^{-1}$) e 60 μL de solução II. A placa foi selada e homogeneizada por inversão (20 vezes). Após um rápido “*spin*” realizou-se a adição de 60 μL de solução III em cada poço, a mistura foi homogeneizada por inversão (10 vezes) e centrifugada como indicado acima. A placa foi incubada na estufa ($90 \text{ }^\circ\text{C}$ por 30 min.), depois colocada no gelo por 10 min. e centrifugada novamente. O lisado foi, então, filtrado por centrifugação com auxílio de uma placa de filtro (PALL) para uma placa fundo V. Em seguida, adicionou-se 90 μL de isopropanol absoluto no filtrado em cada poço da placa de fundo V. A placa foi selada e homogeneizada por inversão (10 vezes). Após homogeneização, a placa foi centrifugada por 45 minutos a 4000 rpm para a precipitação do DNA plasmídeo, e o sobrenadante foi descartado. O excesso de sais presente na extração foi removido pela adição de 200 μL de etanol 70% ($4 \text{ }^\circ\text{C}$) seguido de centrifugação a 4000 rpm durante 5 minutos. Após o descarte do sobrenadante, a placa foi incubada em estufa a $37 \text{ }^\circ\text{C}$ por 15 min. para evaporação completa etanol e o DNA foi ressuspensado em 20 μL de água Milli-Q.

Para a avaliação da integridade do DNA obtido, amostras dos plasmídeos extraídos foram analisados qualitativamente por eletroforese em gel de agarose e quantificados em espectrofotômetro (Nanodrop® Spectrophotometer ND-1000) como descrito previamente.

3.4.6 Sequenciamento

Foram sequenciados os clones contendo os diferentes amplicons da região promotora dos diferentes genótipos de café. Todos os 16 clones, previamente selecionados, foram sequenciados em ambas as direções. Assim sendo, foram utilizados os *primers* universais M13 F (forward) e M13 R (reverse). Nas reações de sequenciamento foi utilizado o kit “Big Dye Terminator” (Applied Biosystems) de acordo com as especificações do fabricante. As reações de sequenciamento foram realizadas no termociclador modelo PTC-100 (MJ Research), com o seguinte programa: temperatura de desnaturação inicial de 96 °C, por 1min. seguidos de 35 ciclos de amplificação compostos de 3 etapas: 96 °C (desnaturação), por 10 seg.; 55 °C (anelamento), por 5 seg.; 60 °C (extensão), por 4 min. e 4 °C etapa final. Os produtos amplificados da reação de sequenciamento foram precipitados com a adição de 40 µL de isopropanol 75% em cada amostra. As placas contendo as amostras foram seladas com adesivos e centrifugadas a 3400 rpm por 45 min. a 20 °C. A remoção do isopropanol foi realizada por inversão seguido de um rápido “*spin*” (até 1.000 rpm). Então, foram adicionados 200 µL de etanol 70% em cada amostra, a placa foi novamente selada e centrifugada por 10 min. a 20 °C. A remoção do etanol foi realizada por inversão e um rápido “*spin*” (1000 rpm). Para garantir a remoção completa do etanol, as placas foram secadas por 30 min. a 37 °C, em estufa.

As amostras foram ressuspendidas em 10 µL de Formamida Hi-Di (Applied Biosystems) e as cadeias de DNA foram desnaturadas em termociclador por 5 min. a 95 °C. Em seguida, as amostras foram sequenciadas na plataforma ABI 3130xl Genetic Analyser (Applied Biosystems).

3.4.7 Análise das sequências

As sequências nucleotídicas obtidas foram analisadas a partir do uso do programa “Sequencing analysis” versão 5.2 da Applied Biosystems. Após a análise das sequências, as mesmas foram alinhadas com o auxílio do programa Seqman (DNASTAR). A partir dos alinhamentos obtidos e análises dos eletroferogramas, foi possível a detecção dos polimorfismos.

3.5 Caracterizações da Região Promotora do gene *DREB2A*

3.5.1 Arquitetura dos *primers* para amplificação da região promotora do gene *DREB2A* nos genótipos de *C. canephora* clone 14 e clone 22, visando à construção do vetor binário.

Primers foram arquitetados visando amplificar a região promotora para clonagem no vetor binário para transformação de citrus. O objetivo das construções é caracterizar funcionalmente o núcleo promotor e avaliar a influência dos motivos *cis* de regulação na manifestação da tolerância à desidratação. Com base na sequência consenso das análises *in silico*, quatro *primers* foram desenhados no programa PrimerQuestSM® (Integrated DNA Technologies - IDT) para a amplificação de três sub-regiões, de extensões diferentes, da sequência promotora do gene *DREB2A* de dois genótipos de *C. canephora* contrastantes para a tolerância à seca. Os *primers* arquitetados para a amplificação das sub-regiões do promotor foram: DrebPProxF forward 5’-TAATTCC**AAGCTT**TGTCTGAAGT-3’, DrebPMedF forward 5’-AAGAGAACAAC**AAGCTT**CTTGT-3’, DrebPDistF forward 5’-TCCTAGT**AAGCTT**CACGTTGT-3’, DrebPR reverse 5’-TGTTGAGAAATGGTT**AGATCTT**GAA-3’.

As sequências oligonucleotídicas dos *primers* foram alteradas para a inserção de sítios de restrição das enzimas *Hind* III (vermelho), nos *primers* forward, e *Bgl* II (verde), no primer reverse. Essas alterações foram concebidas por uma estratégia de clonagem para possibilitar a ligação dos amplicons ao vetor binário.

Os *primers* foram utilizados na PCR objetivando a amplificação de três fragmentos distintos da região promotora, da seguinte forma: Amplicon Dreb Proximal (765bp), DrebPProxF e DrebPR, amplicon Dreb Mediano (1.118bp), DrebPMedF e DrebPR, e amplicon Dreb Distal (1.469bp), DrebPDistF e DrebPR (Tabela 3).

Tabela 3. *Primers* utilizados para as amplificações dos DNAs dos genótipos Clone 14 e 22. Onde ^(a), ^(b) e ^(c) correspondem aos pares de *primers* utilizados para amplificação do promotor *DREB2A*.

<i>Primers</i>	Sequências dos <i>Primers</i>	Adaptadores (Sítios de Restrição)
<i>DrebPProx</i> ^(a)	5'– TAATTCC AAGCTT TGTCTGAAGT–3'	<i>Hind</i> III
<i>DrebPMed</i> ^(b)	5'– AAGAGAACAAC AAGCTT CCTTGT–3'	<i>Hind</i> III
<i>DrebPDist</i> ^(c)	5'– TCCTAGT AAGCTT CACGTTGT–3'	<i>Hind</i> III
<i>DrebPR</i> ^(a,b,c)	5'– TGTTGAGAAATGGTT AGATCTT GAA– 3'	<i>Bgl</i> II

O amplicon Dreb Proximal abrange os pares de base + 1 ao - 765 da região promotora. Já o amplicon Dreb Mediano amplifica a região + 1 a - 1.118. O amplicon Dreb Distal corresponde à amplificação da região + 1 a - 1469 do gene *DREB2A* (Figura 10).

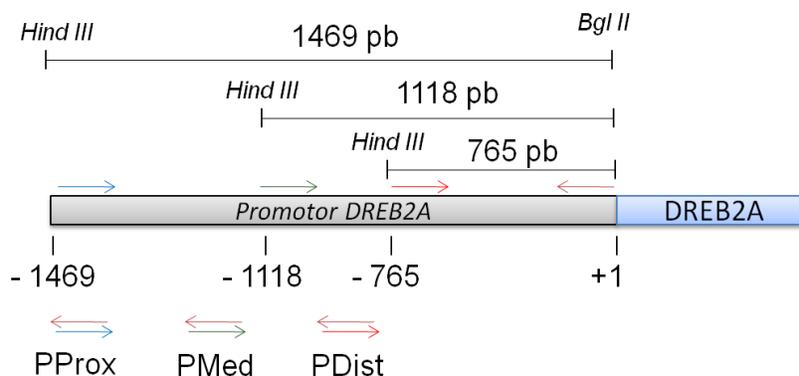


Figura 10. Amplificação da região promotora do gene *DREB2A* dos clones 14 e 22 de *C. canephora* var. Conilon com os *primers* DPProx, DPMed, DPDist e DPR. Esquema ilustrativo das regiões do promotor do gene *DREB2A* amplificadas visando à construção do cassete gênico

A formação de potenciais estruturas secundárias, a especificidade dos *primers*, e a diminuição da temperatura de anelamento pelo pareamento inadequado (*mismatch*) foram avaliadas pelos programas OligoAnalyzer 3.1 (IDT Technologies) e FastPCR Professional 6.1 (Primer Digital Ltda) (Kalendar et al., 2009).

3.5.2 Reação em cadeia da polimerase (PCR) e amplificação do DNAg

Para a amplificação dos fragmentos foi utilizado 6.25 ng de DNAg de cada genótipo. As reações de PCR foram realizadas com um volume final de 50 μ L de solução usando a enzima Platinum® *Taq* DNA Polimerase (Invitrogen), segundo as especificações do fabricante. As reações foram incubadas no termociclador modelo Mastercycler® (Eppendorf) com o programa de termociclagem descrito a seguir: desnaturação inicial de 94 °C por 2 min., seguido de 35 ciclos de amplificação: 94 °C (desnaturação), por 30 seg.; 68 °C (anelamento), por 30 seg.; 72 °C (extensão), por 2 min. e de uma etapa final de polimerização por 10 min. a 72 °C. A temperatura de anelamento foi ajustada

conforme a necessidade dos pares de *primers*. Em todas as reações de amplificação foram inseridos 10 ciclos iniciais com valores mais baixos de temperatura de anelamento para garantir o pareamento dos *primers* na sequência-alvo. O reajuste na temperatura de anelamento foi calculado pelo valor de temperatura, em graus, perdido com a presença dos *mismatches*. A análise qualitativa dos amplicons foi realizada por eletroforese em gel de agarose, como descrito anteriormente, e o tamanho dos fragmentos foi verificado com o marcador 1 kb DNA Ladder (Amresco). Em seguida, todos os produtos de PCR foram clonados.

3.5.3 Subclonagem dos produtos de PCR e digestão

Os produtos de PCR, com adaptadores de sítio de restrição, foram ligados ao vetor de transição pGEM-T Easy Vector (Promega), conforme as especificações do fabricante, para estabilizar as extremidades 5' e 3' dos amplicons e facilitar a digestão endonucleásica das enzimas de restrição. Os plasmídeos foram transfectados em células bacterianas e extraídos, posteriormente, como descrito nos subitens anteriores.

Os insertos, com adaptadores de sítio de restrição, foram digeridos com as enzimas *Hind* III e *Bgl* II. A reação de digestão foi realizada em um volume final de 20 μL , sendo 1 μL de DNA plasmidial, 2 μL de tampão, 0,5 μL de cada enzima, 0,2 μL de BSA e 15,8 μL de água-miliq. A reação foi incubada a 37 °C por 2 horas. O produto da digestão foi purificado utilizando-se o kit “Wizard SV Gel and PCR Clean-Up System” (Promega), segundo as normas do fabricante.

3.5.4 Digestão do vetor binário

O vetor binário utilizado para a construção foi o pBI121. Este vetor possui um cassete gênico para expressão da enzima β -glucuronidase. O cassete é composto por um promotor constitutivo, 35S, regulando a região codante do gene GUS e a região terminadora NOS. Formas circularizadas do vetor foram digeridas, com as enzimas de restrição *Hind* III e *Bam*H I, para excisão do promotor 35S e posterior ligação com os amplicons do gene *DREB2A* amplificados. Na reação de digestão foram utilizados 40 μ L do vetor pBI121, 20 μ L do tampão de digestão, e 5 μ L das enzimas *Hind* III e *Bam*H I, o volume final foi completado com água para 100 μ L. A reação foi incubada a 37 °C por 2 horas. O produto de digestão foi analisado em gel de agarose 0,7%, o fragmento do vetor linearizado (9kb), sem o promotor 35S, foi isolado e purificado utilizando-se o kit “Wizard SV Gel and PCR Clean-Up System” (Promega), segundo as recomendações do fabricante.

3.5.5 Ligação dos produtos de PCR ao vetor binário

Os fragmentos digeridos do vetor pGEM-T Easy Vector foram ligados a formas linearizadas do vetor pBI121. As reações de ligação foram preparadas adicionando-se 6 μ L do DNA digerido, 1 μ L do vetor pBI 121, 1 μ L do tampão, 1 μ L da enzima T4 DNA ligase e 1 μ L de água. As misturas foram incubadas à temperatura ambiente (22-23 °C) por 1 hora e, em seguida, acondicionadas no gelo até a transfecção bacteriana. A transfecção foi realizada conforme descrito no item 3.4.4.

3.5.6 Sequenciamento

Os clones positivos confirmados por digestão foram sequenciados para a confirmação da sequência promotora. Assim sendo, foram utilizados os *primers* universal M13 R (reverse), DrebS1R (reverse) e DrebS2F (forward). Nas reações de sequenciamento das construções Dreb Proximal foi utilizado somente o *primer* M13R, nas construções Dreb Mediano, os *primers* M13 R e DrebS1R, e nas construções Dreb Distal, os *primers* M13 R, DrebS1R e DrebS2F. A reação de sequenciamento foi realizada conforme descrito no item 3.4.6.

4 RESULTADOS E DISCUSSÃO

4.1 Expressões do gene *DREB2A* por PCR quantitativo em tempo real q(PCR)

Os resultados da expressão do gene *DREB2A* (Contig 14421) nos genótipos de *C. canephora* clone 14, 22, 73 e 120 mostraram uma expressão diferencial entre os genótipos tolerantes (14, 73 e 120) e o genótipo sensível (22), corroborando com os resultados obtidos por Marraccini et al. (2009) (Figura 11). Todos os clones foram submetidos a dois regimes hídricos, irrigado (I) e não-irrigado (NI). Ambos os genótipos, tolerante e sensível, apresentaram aumento da expressão do gene *DREB2A* em condições de déficit hídrico. Também, foi possível observar que o nível de expressão foi maior (3 a 10x) no grupo tolerante em comparação ao genótipo sensível. Esses resultados confirmam a influência do estresse hídrico na expressão do gene *DREB2A* e permitem inferir que a participação desse gene no mecanismo de resposta à seca pode contribuir para o desenvolvimento da tolerância ao estresse hídrico, consistente com os resultados de Dubouzet et al. (2003), Kasuga et al. (2001) e Liu et al. (1998).

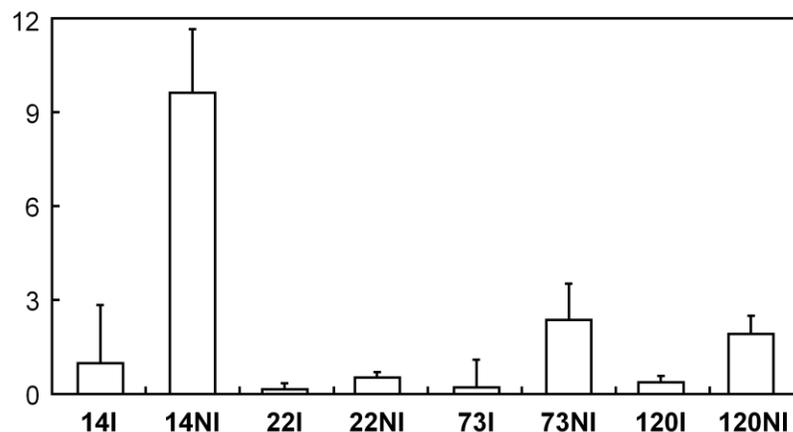


Figura 11. Perfis de expressão do gene *DREB2A* em folhas dos clones 14,73,120 e 22 de *C. canephora* var. Conilon com Irrigação (I) e sem Irrigação (NI), obtidos por meio da técnica de qPCR. O gene *DREB2A* apresentou aumento de expressão sob limitação de água e diferença significativa no perfil de expressão entre os genótipos tolerantes 14, 73, 120 e o genótipo sensível 22.

4.2 Isolamento e identificação da região promotora

A determinação da sequência promotora do gene *DREB2A* foi realizada com a análise das sequências do banco de dados de *shotgun* do DNA genômico de *C. canephora*. O banco de dados foi obtido com o pirosequenciamento por 454 (Roche) de fragmentos produzidos pela técnica *shotgun* do DNA genômico de um *C. canephora* duplo haplóide. Análises por BlastN da sequência de EST do gene *DREB2A* neste banco de dados permitiram a identificação de sequências similares (Hits). Os *hits* foram montados em uma sequência consenso objetivando a identificação dos motivos *cis* de regulação na região promotora. A sequência consenso de 1588 pares de bases abrange a região da posição -1451 a +137 do gene *DREB2A*, incluindo os primeiros 45 resíduos de aminoácidos da proteína.

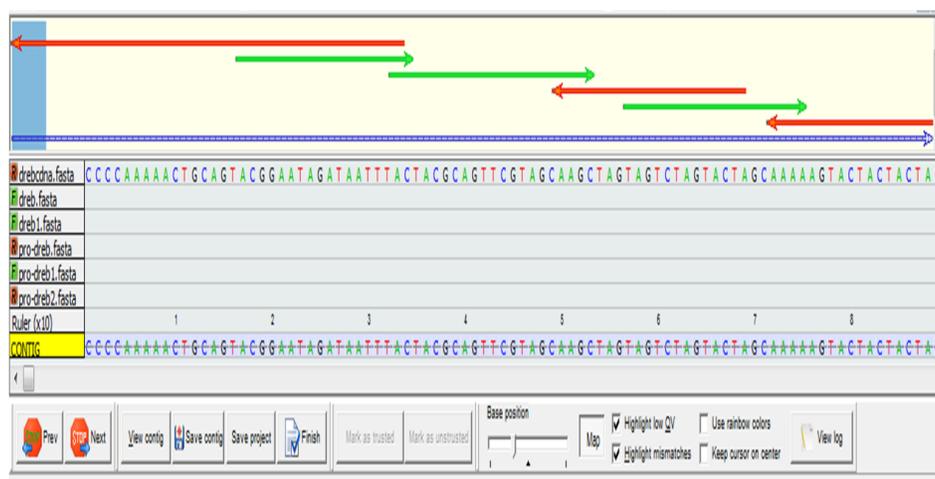


Figura 12. Representação da estratégia de alinhamento dos *hits* obtidos por Blast da sequência EST do gene *DREB2A* no banco de dados de Shotgun do DNAg de *C. canephora*. Foi obtida uma sequência consenso de 1558 pb.

As análises da sequência consenso nos bancos de dados de motivos *cis* de regulação PLACE (Plant *Cis*-acting Regulatory DNA elements) e PlantCARE (Plant *Cis*-Acting Regulatory Elements) possibilitou a identificação dos elementos de regulação característicos da resposta ao estresse hídrico, como ERF/DRE (Dehydration Responsive Element), ERE (Ethylene Responsive Element), ABRE (ABA Responsive Element), MYCRS (MYC Responsive Binding Site), MYBRS (MYB Responsive Binding Site), ATHB (HD-ZIP Responsive Element), ERD1 (Early Responsive to Dehydration 1), entre outros (anexo). A base dos bancos PLACE e PlantCARE é constituída apenas de referências comprovadas experimentalmente da participação desses elementos de regulação no controle da expressão gênica.

Com base na sequência consenso obtida de *C. canephora*, foram arquitetados *primers* visando à amplificação dos diferentes genótipos do gênero *Coffea* (APÊNDICE - Tabela 4). Como característico das regiões promotoras, a sequência consenso apresentou elevado número de sequências repetitivas diminuindo as opções de regiões adequadas para o anelamento dos *primers*.

Entretanto, os *primers* demonstraram-se eficientes na amplificação dos diferentes genótipos demonstrando certa conservação na sequência complementar aos oligonucleotídeos. Os pares de *primers* S1F/R, S2F/R e S1F/S2R amplificaram sequências de aproximadamente 800, 900 e 1500 bp, respectivamente, condizentes com os amplicons esperados para as sequências-alvo, S1 (837bp), S2 (952bp), e S1/S2 (1560bp).

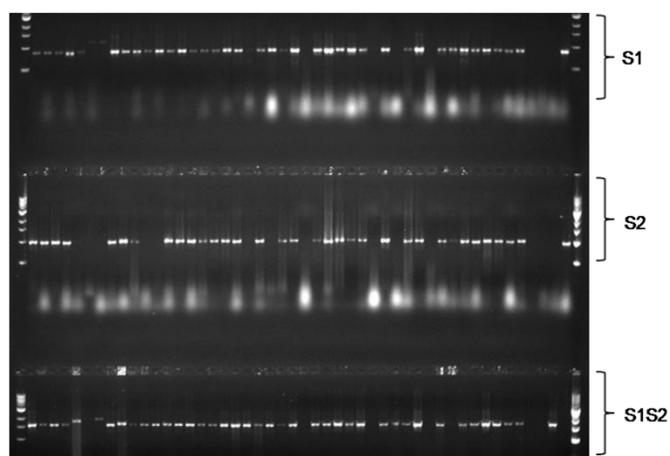


Figura 13. Amplificação da região promotora nos diferentes genótipos da tabela 4 (Apêndice C). Estão representados, de cima para baixo, os amplicons S1, S2 e S1S2 usados para identificação da região promotora por sequenciamento. O marcador molecular de 1KB permite a confirmação do tamanho esperado dos fragmentos 837, 952 e 1547 bp, respectivamente.

Foi realizada a clonagem dos produtos de amplificação apenas dos genótipos de *C. canephora* (Clone 14 e 22), *C. arabica* (Catuaí 144, Iapar 59, Mundo Novo, Obatã, Rubi, Tupi e Typica), *C. eugenioides* e *C. racemosa*. Os amplicons, S1, S2 e S1S2 foram clonados no vetor pCR 2.1 TOPO com o objetivo de individualizar os alelos para posterior sequenciamento. A clonagem foi igualmente eficiente, produzindo várias colônias de coloração branca e com eficiência calculada em 10^7 . Exceção para a espécie *C. racemosa* que apresentou um número inferior de colônias na clonagem dos amplicons S1 e S2

devido à menor eficiência de amplificação destes fragmentos. Essa diferença de eficiência na amplificação pode ser explicada por sua maior distância genealógica dentro do grupo de genótipos. Entretanto, o amplicon S1S2 apresentou eficiência de amplificação semelhante a dos demais genótipos, o que sugere diferenças apenas nas sequências-alvo dos *primers* Drebs1R e Drebs2F. Para o sequenciamento, foram selecionados 16 clones por genótipo para a (i) clonagem S1, (ii) clonagem S2, e (iii) clonagem S1S2, além de um clone controle (colônia azul). A presença das formas multiméricas do vetor nos clones selecionados foi confirmada por extração plasmideal (Figura 14).

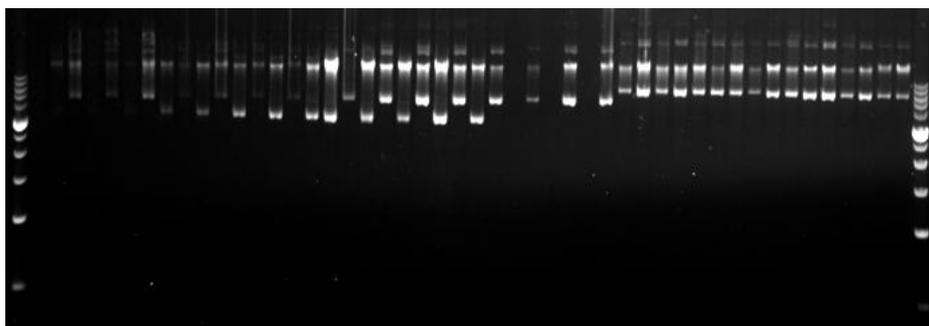


Figura 14. Extração plasmideal dos clones selecionados para o sequenciamento. Confirmação da presença das formas multiméricas do vetor clonado.

Os 48 clones de cada genótipo (16 clones x 3 clonagens) foram sequenciados na plataforma de sequenciamento 3730 xl da Applied Biosystem. As sequências foram analisadas no programa Sequencing Analises (Applied Biosystem) e apresentaram boa qualidade gerando fragmentos de aproximadamente 600 pb com Phred superior ou igual a 20. Com o auxílio do programa Seqman (DNASTAR), as estruturas do vetor foram removidas e as sequências alinhadas. Primeiramente foi realizado um alinhamento completo por genótipo com todas as 48 sequências obtidas. O resultado do alinhamento permitiu visualizar as estratégias de sequenciamento e a orientação dos

fragmentos clonados. Pode ser observado que a distribuição e orientação dos fragmentos são proporcionais, confirmando a eficiência do sequenciamento e da estratégia de clonagem.

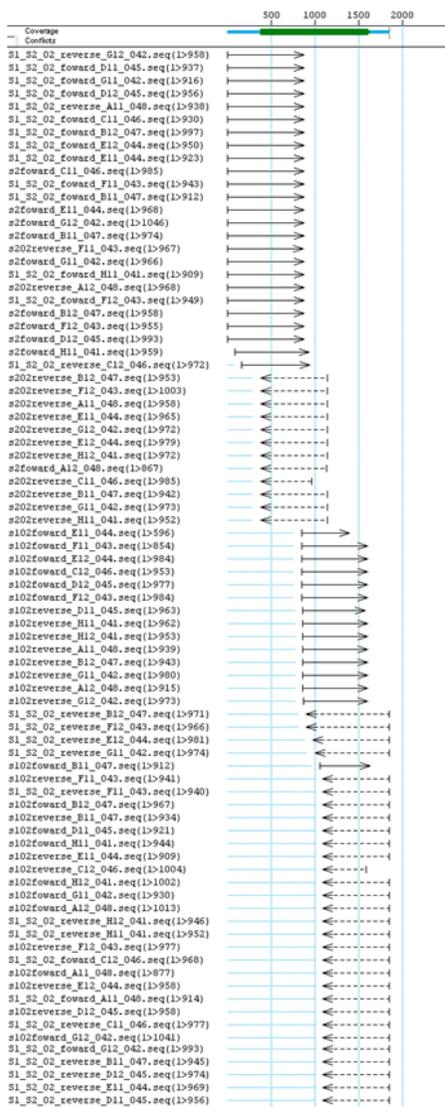


Figura 15. A estratégia de sequenciamento adotada para identificação das diferentes regiões que integram o promotor *DREB2A*. Foram sequenciados 16 clones para cada genótipo e região amplificada. Representação gráfica das 48 sequências obtidas para o genótipo Obatã.

Posteriormente, visando facilitar a montagem dos contigs e a identificação dos polimorfismos, o alinhamento por genótipo foi iniciado pelo amplicon S1. A escolha baseou-se no maior número de sequências obtidas para este fragmento nos diferentes genótipos.

4.3 Análises dos polimorfismos na região promotora do gene *DREB2A*

As sequências geradas para a região promotora do gene *DREB2A* permitiram a avaliação de polimorfismos em todos os 12 genótipos. A análise dos polimorfismos permitiu a identificação dos diferentes haplótipos nos diferentes genótipos analisados. Dois alelos de cada genótipo foram selecionados e suas sequências alinhadas no programa de alinhamento múltiplo de sequências ClustalW2 (EMBL-EBI), para ilustração. Os genótipos de *C. canephora*, clone 14 e 22 apresentaram distribuição alélica diferente para o locus do gene *DREB2A* (Figura 16). Ambos os clones 14 e 22 apresentaram heterozigose para o locus analisado. As alterações encontradas nos genes alelos adquiridos na domesticação são mais comuns em regiões promotoras do que em regiões codantes (Zhu, 2003, Borevitz, 2004). Desta forma, o estudo dessa diferença observada poderá contribuir para identificar as causas da variação na adaptação ao estresse hídrico entre os genótipos clones 14 e 22. A espécie *C. eugenioides* também demonstrou heterozigose para o locus, assim como a maioria dos demais genótipos de *C. arabica*, tais como Bourbon, Catuaí, Iapar 59, Mundo Novo, Rubi e Typica. A diferença de haplótipos observada nos genótipos heterozigóticos de *C. arabica* corrobora com estudos anteriores da diversidade genética no gênero *Coffea*. (Lashermes et al., 1996; Lashermes, 1999). Lashermes et al. descreve a origem do *C. arabica* como resultado de uma hibridização natural entre o *C. canephora* e o *C. eugenioides*. As análises da

distribuição dos haplótipos permitem a identificação de dois grupos de alelos homozigóticos diferentes, um pertencente ao genótipo *C. canephora* e outro ao *C. eugenoides*. Os genótipos de *C. arabica* heterozigotos apresentam um haplótipo de cada grupo alélico, indicando um comportamento de segregação dos alelos parentais na progênie.

4.4 Análise dos elementos *cis*-reguladores do DNA

As sequências obtidas pela amplificação da região promotora nos diferentes genótipos foram analisadas quanto à ocorrência dos motivos *cis* de regulação do DNA. Várias regiões consenso características de promotores foram encontradas como CATA-, TATA-, CGCG-BOX (Figura 16). Essas regiões são receptoras do complexo de transcrição da RNA polimerase II (TATA- e CATA-BOX) ou de fatores ativadores da transcrição (CGCG-BOX).

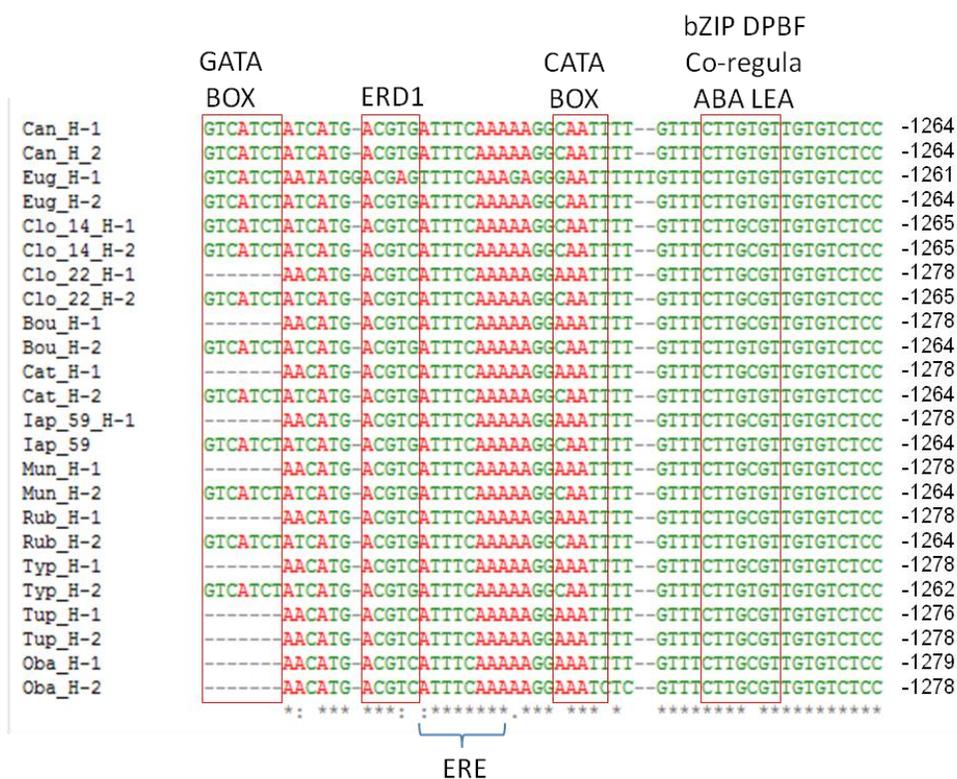


Figura 16. Alinhamento da região S1 do promotor *DREB2A*. Regiões polimórficas entre os genótipos estão enquadrados em vermelho. Motivos de regulação *cis* do DNA, região consenso promotora (GATA-BOX e CATA-BOX), elemento inicial de resposta a desidratação (ERD1) e elemento de resposta ao etileno (ERE). (Can=Canephora duplo diplóide; Eug=Eugenióides; Clo_14=Clone 14; Clo_22=Clone 22; Bou=Bourbon; Cat=Catuaí; Iap_59=Iapar 59; Mun=Mundo Novo; Rub=Rubi; Typ=Typica; Oba=Obatã; H_1= Haplótipo 1; H_2=Haplótipo 2)

A presença dos CATA e TATA-BOX ao longo de toda a região promotora permite uma maior combinação dos motivos reguladores ativos durante o processo transcricional. Vários motivos conservados atuantes no mecanismo de resposta à seca, previamente encontrados na sequência consenso do *C. canephora* duplo diplóide, foram confirmados no promotor dos demais genótipos. Entre os quais os motivos ABRE, MYBRS e MYCRS, regulados

pelos fatores de transcrição pertencentes às famílias b-ZIP, MYB e MYC, respectivamente (Figuras 17 e 18).

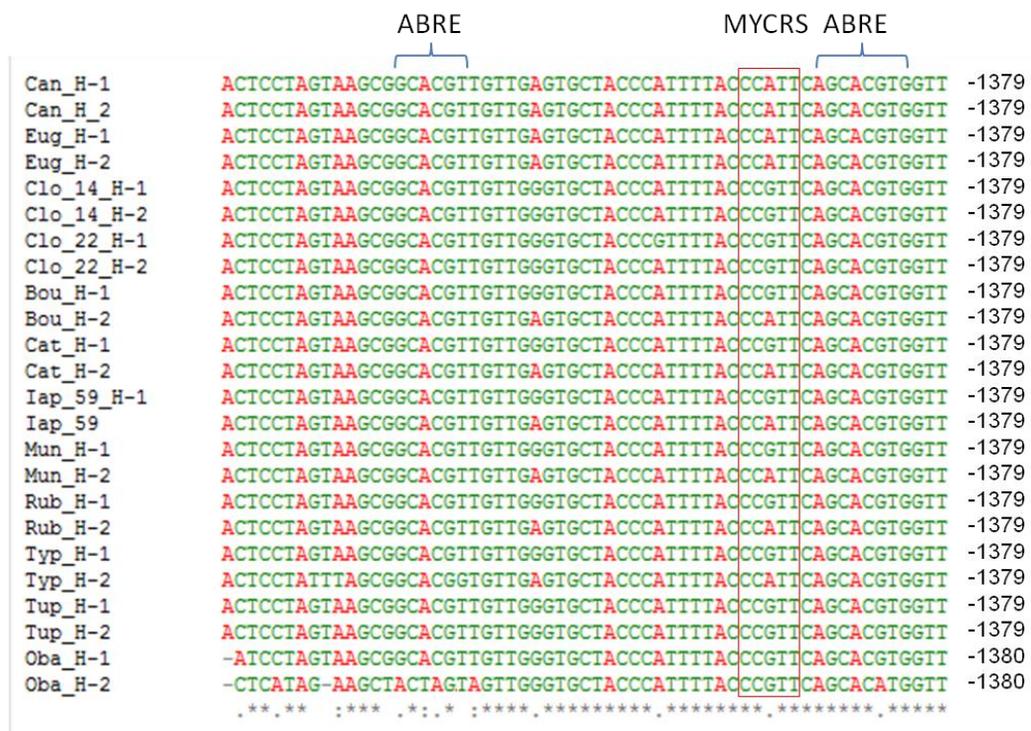


Figura 17. Alinhamento da região S1 do promotor *DREB2A*. Regiões polimórficas entre os genótipos estão enquadrados em vermelho. Motivos de regulação *cis* do DNA, elemento de resposta ao ácido abscísico (ABRE) e elemento de resposta ao TF MYC (MYCRS). (Can=Canephora duplo diplóide; Eug=Eugenióides; Clo_14=Clone 14; Clo_22=Clone 22; Bou=Bourbon; Cat=Catuai; Iap_59=Iapar 59; Mun=Mundo Novo; Rub=Rubi; Typ=Typica; Oba=Obatã; H_1= Haplótipo 1; H_2=Haplótipo 2).

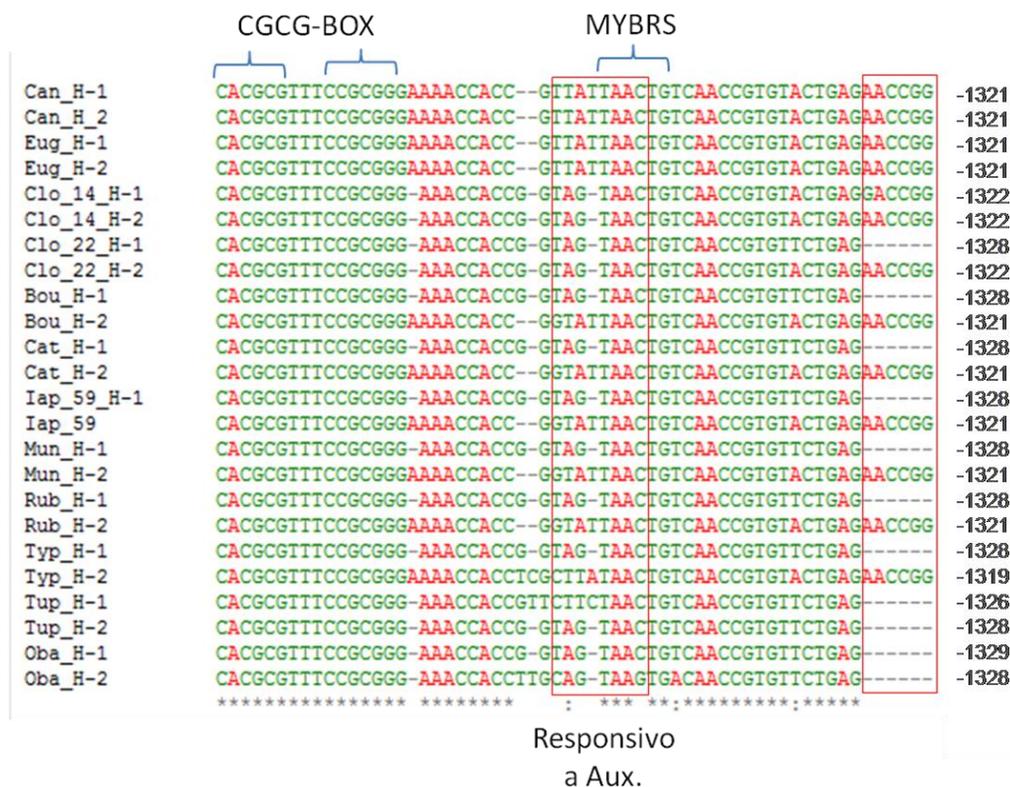


Figura 18. Alinhamento da região S1 do promotor *DREB2A*. Regiões polimórficas entre os genótipos estão enquadradas em vermelho. Motivos de regulação *cis* do DNA, região consenso promotora (CGCG-BOX) e elemento de resposta ao TF MYB (MYBRS). (Can=Canephora duplo diplóide; Eug=Eugenióides; Clo_14=Clone 14; Clo_22=Clone 22; Bou=Bourbon; Cat=Catuaí; Iap_59=Iapar 59; Mun=Mundo Novo; Rub=Rubi; Typ=Typica; Oba=Obatã; H_1= Haplótipo 1; H_2=Haplótipo 2).

Essas proteínas FTs participam da via de resposta à desidratação dependente de ABA. Entretanto, a família ERF, na qual está incluído o fator de transcrição DREB, participa da via independente de ABA (Liu et al., 1998). Essa aparente contradição pode ser explicada por estudos que sugerem a interação e correção das vias dependente e independente de ABA (Agarwal et al., 2006). Dubouzet et al.(2003) e Shen et al.(2003) comprovaram que os genes,

isolados de arroz (*OsDREB2A*) e trigo (*TaDREB1*), codificantes de FTs tipo DREB2 sofrem uma fraca indução ao ácido abscísico.

Foi possível a identificação de motivos de regulação responsivos ao estresse abiótico, como o GT-1. A presença de elementos *cis* de regulação a estresse biótico corroboram com os resultados anteriores que demonstram haver uma correção do gene *adr1* (activated disease resistance 1) pelo FTs DREB2A (Chini et al., 2004).

Os elementos *cis* de ligação do fator de transcrição DREB (ERF), também foram identificados na sequência promotora. A presença dos motivos ERF na região mais próxima ao início do códon de transcrição sugere uma participação mais direta do TF DREB2A na regulação da expressão do gene *DREB2A*, sem muita interferência dos demais elementos de regulação *cis* (Figuras 19, 20, 21 e 22).

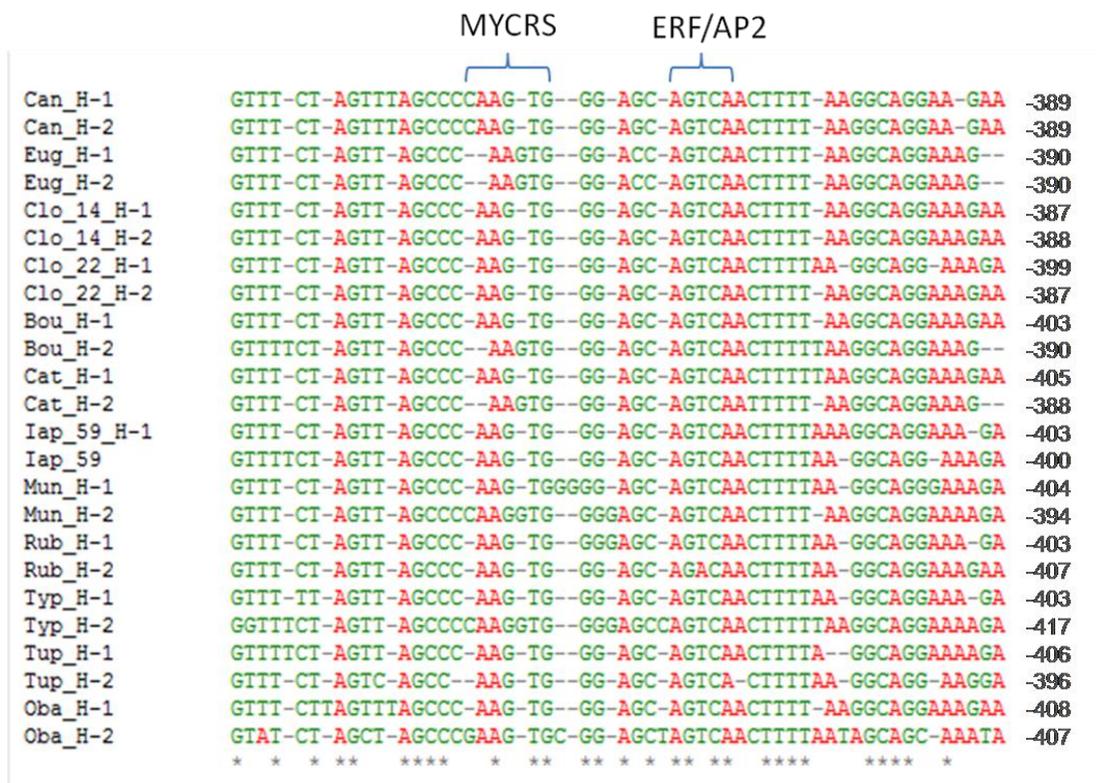


Figura 19. Alinhamento da região S2 do promotor *DREB2A*. Motivos de regulação *cis* do DNA, elemento de resposta ao TF MYC (MYCRS) e elemento de resposta aos FTs ERF (família de fatores de transcrição do DREB). (Can=Canephora duplo diplóide; Eug=Eugenióides; Clo_14=Clone 14; Clo_22=Clone 22; Bou=Bourbon; Cat=Catuai; Iap_59=Iapar 59; Mun=Mundo Novo; Rub=Rubi; Typ=Typica; Oba=Obatã; H_1= Haplótipo 1; H_2=Haplótipo 2).

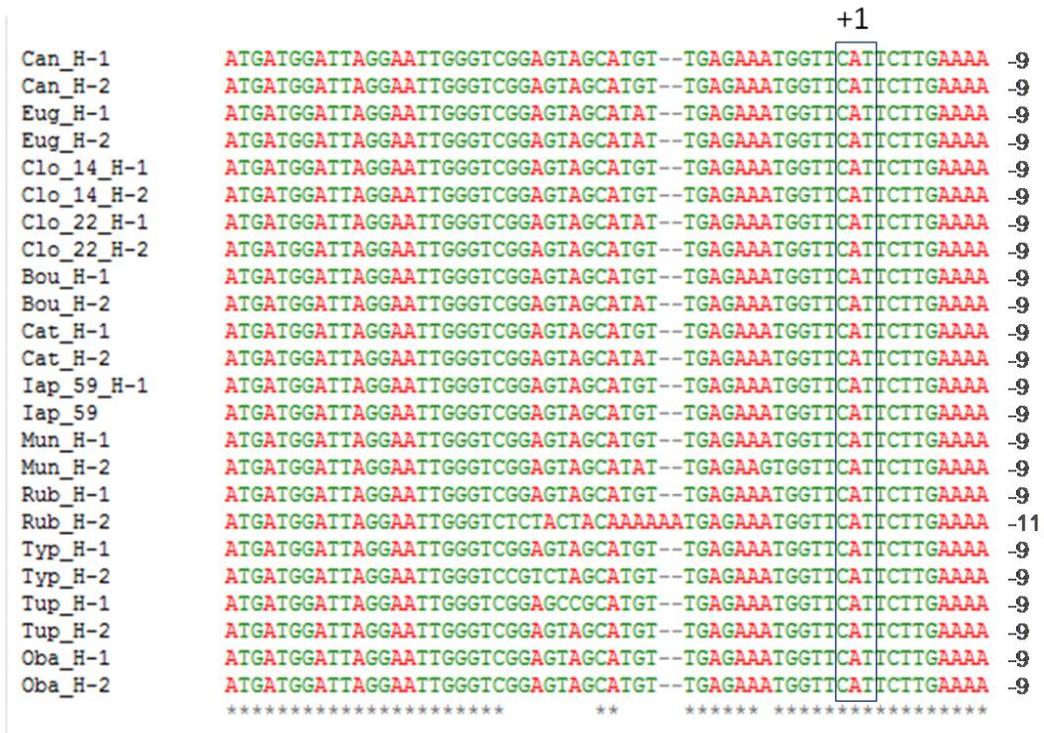


Figura 21. Alinhamento da região S2 do promotor *DREB2A*. Início da região codante +1 da proteína DREB. (Can=Canephora duplo diplóide; Eug=Eugenióides; Clo_14=Clone 14; Clo_22=Clone 22; Bou=Bourbon; Cat=Catuaí; Iap_59=Iapar 59; Mun=Mundo Novo; Rub=Rubi; Typ=Typica; Oba=Obatã; H_1= Haplótipo 1; H_2=Haplótipo 2).

```

Can_H-1      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Can_H-2      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Eug_H-1      TGGCTTTGTAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAATCTGGAAAATCAGGGG +49
Eug_H-2      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAATCTGGAAAATCAGGGG +49
Clo_14_H-1   TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Clo_14_H-2   TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Clo_22_H-1   TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Clo_22_H-2   TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Bou_H-1      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Bou_H-2      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAATCTGGAAAATCAGGGG +49
Cat_H-1      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Cat_H-2      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAATCTGGAAAATCAGGGG +49
Iap_59_H-1   TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Iap_59       TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Mun_H-1      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Mun_H-2      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAATCTGGAAAATCAGGGG +49
Rub_H-1      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Rub_H-2      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Typ_H-1      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Typ_H-2      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Tup_H-1      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Tup_H-2      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Oba_H-1      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
Oba_H-2      TGGCTTTGCAGGCATTGACTACGTTGTCCGAAGTTGATGAAGAGTCTGGAAAATCAGGGG +49
*****

```

Figura 22. Alinhamento da região S2 do promotor *DREB2A*. Região codante da proteína DREB. (Can=Canephora duplo diplóide; Eug=Eugenióides; Clo_14=Clone 14; Clo_22=Clone 22; Bou=Bourbon; Cat=Catuaí; Iap_59=Iapar 59; Mun=Mundo Novo; Rub=Rubi; Typ=Typica; Oba=Obatã; H_1=Haplótipo 1; H_2=Haplótipo 2).

4.5 Caracterizações da região promotora.

Objetivando a caracterização da região promotora do gene *DREB2A*, considerando a expressão diferencial da proteína DREB2A observada nos genótipos de *C. canephora*, clones 14 e 22, foram construídos 8 cassetes para expressão heteróloga objetivando avaliar (i) o núcleo promotor necessário para o controle da expressão gênica, e (ii) a influência dos polimorfismos no promotor para o controle da expressão gênica. Para amplificação dos alelos selecionados – clones 14A06, 14C06 e 22A08 – foram desenhados *primers* com adaptadores contendo sítio de restrição. O sítio de restrição foi adicionado no interior dos

primers para garantir uma melhor eficiência da atividade endonucleásica das enzimas de restrição. Apesar da instabilidade no pareamento devido à modificação de algumas bases nucleotídicas para a criação do sítio de restrição, os pares de primer foram eficientes na amplificação da sequência-alvo dos clones selecionados. Os amplicons obtidos apresentaram tamanhos correspondentes ao esperado, Dreb Proximal (765bp), Dreb Mediano (1.118bp), e Dreb Distal (1.469bp) (Figura 23).

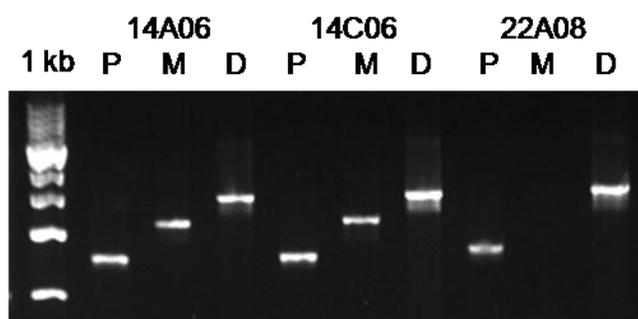


Figura 23. Amplificação da região promotora nos genótipos clone 14 e 22 com os *primers* DPProx, DPMed, DPDist e DPR. Estão representados os amplicons Dreb Proximal, Dreb Mediano, Dreb Distal de cada um dos clones escolhidos (14A06, 14C06 e 22A08). O marcador molecular de 1KB permite a confirmação do tamanho esperado dos fragmentos 765, 1118 e 1469 bp, respectivamente

O vetor binário escolhido para a transformação por *Agrobacterium tumefaciens* foi o pBI121. Esse vetor binário foi escolhido por possuir uma construção com o gene GUS controlado por um promotor constitutivo 35S flanqueado por sítios de restrição que permitem a excisão do promotor constitutivo e a inserção do promotor de interesse. A atividade regulatória do promotor pode ser determinada por análise da expressão do gene repórter *GUS* (Figura 24).

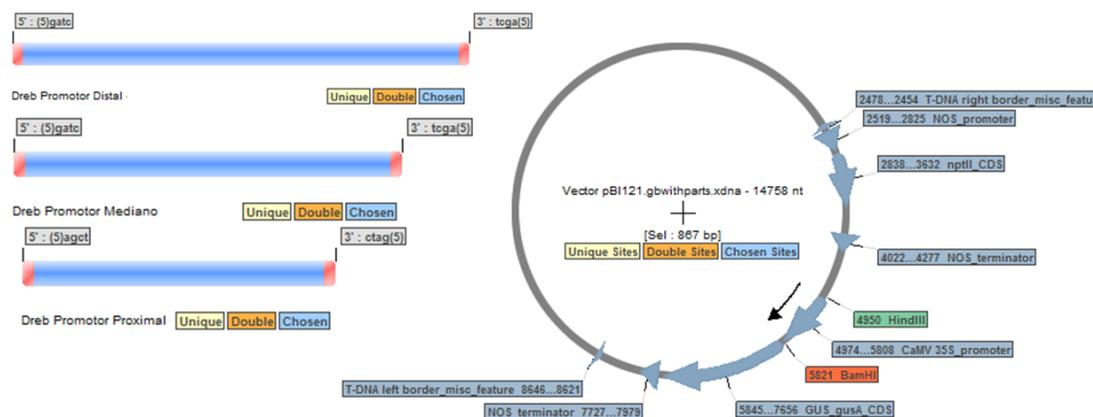


Figura 24. Esquema ilustrativo da estratégia adotada para construção dos cassetes gênicos utilizando o vetor binário pBI121. O promotor constitutivo indicado pela seta é excisado pela digestão com as enzimas de restrição *Hind* III e *Bam*H I e novos promotores são incorporados ao vetor. Os novos promotores passam a mediar à expressão da proteína GUS.

Foram escolhidas enzimas de restrição que possuísem sítios únicos de no vetor prevenindo a inversão da sequência promotora durante ligação ao vetor. Apesar de terem sido utilizados dois pares diferentes de enzimas para a digestão do vetor binário – *Hind* III/*Bam*H I – e dos produtos de PCR – *Hind* III/*Bgl* II, as extremidades 3' resultantes da digestão por *Bgl* II ou *Bam*H I produzem extremidades adesivas. As construções geradas foram pD14–1 (pDREB2A 14A06 + pBI121), pD14–2 (pDREB2A 14C06 + pBI121) e pD22 (pDREB2A 22A08 + pBI121). Nas construções pD14–1 e pD14-2 foram utilizadas três regiões do promotor pDREB2A 14A06 e 14C06, a região proximal, mediana e distal, sendo as construções geradas por estes fragmentos pD14–1 P, pD14–1 M, pD14–1 D e pD14–2 P, pD14–2 M, pD14–2 D, respectivamente. Apenas duas construções foram obtidas para o clone 22A08, uma da região proximal, pD22 P e outra com a região distal, pD22 D. A clonagem foi eficiente e os clones foram confirmados por sequenciamento.

5. CONCLUSÕES

Os resultados do qPCR obtidos, avaliando-se os genótipos de *C. canephora*, dos Clones 14, 73, 120 e 22 demonstraram realmente a participação do gene *DREB2A* no mecanismo de resposta ao estresse hídrico em cafeeiro. A presença de polimorfismos na região promotora do gene *DREB2A* entre os genótipos tolerante (clone 14) e sensível (clone 22) pode ser a causa da diferença significativa na expressão gênica observada nas análises quantitativas de PCR em tempo real. No entanto, para confirmar essa hipótese estudos experimentais adicionais necessitam ser realizados. A transformação genética com cassetes gênicos contendo um gene repórter sob a regulação da região promotora do gene *DREB2A* isolada de ambos os genótipos auxiliará na elucidação da influência dos polimorfismos no mecanismo de resposta ao déficit hídrico. As análises dos elementos *cis* de regulação da região promotora permitem inferir a existência de uma regulação interligada das vias dependente e independente de ABA na resposta ao estresse hídrico. Além disso, as análises demonstram a ocorrência de elementos de regulação característicos de genes responsivos ao estresse biótico. A presença desses diversos elementos regulatórios na região promotora do gene *DREB2A* confirma a grande complexidade da cascata de resposta aos estresses bióticos e abióticos. É notável também, a diferença no padrão alélico entre os genótipos analisados. A distribuição alélica evidencia a presença de pelo menos dois alelos diferentes para o locus DREB. A presença dos alelos de *C. canephora* e *C. eugenioides* em heterozigose no *C. arabica* confirma os dados já observados de que *C. arabica* é um híbrido de *C. canephora* e *C. eugenioides*. A maioria dos cultivares de *C. arabica* apresenta um padrão alélico heterozigótico para o locus DREB.

REFERÊNCIA BIBLIOGRÁFICA

ABE, H.; URAO, T.; ITO, T.; SEKI, M.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. 2003. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. **Plant Cell Reports** 15: 63-78.

ABE, H.; YAMAGUCHI-SHINOZAKI, K.; URAO, T.; IWASAKI, T.; HOSOKAWA, D.; SHINOZAKI, K. 1997. Role of Arabidopsis MYC and MYB homologs in drought- and abscisic acid-regulated gene expression. **Plant Cell Reports** 9: 1859-1868.

AGARWAL, P.; AGARWAL, P.; REDDY, M.; SOPORY, S. 2006. Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. **Plant Cell Reports** 25: 1263-1274.

ALVES, G.S.C. 2009. Análise quantitativa da expressão de genes candidatos em genótipos de *Coffea arabica* submetidos ao déficit hídrico. **Anais do 18 Congresso de Pós-Graduação da UFLA**. Lavras, CD-Rom.

ANDRADE, A.C. 2007. Projeto genoma café: uma ferramenta para acelerar o melhoramento de variedades. **Irrigação e Tecnologia Moderna** 73: 19-20.

ANDRADE, A.C. 2008. Advanced Biology Applied to Coffee Research - Current Status and Future Perspectives. **Proceedings of 22 International Conference On Coffee Science**. Campinas: 22 International Conference On Coffee Science - ASIC, CD-ROM.

ANTHONY, F.; BERTHAUD, J.; GUILLAUMENT, J.L.; LOURD, M. 1987. Collecting wild *Coffea* species in Kenya and Tanzania. **Plant Genetic Resources Newsletter** 69: 23-29.

BERTHOU, F.; MATHIEU, C.; VEDEL, F. 1983. Chloroplast and mitochondrial DNA variation as indicator of phylogenetic relationships in the genus *Coffea* L. **Theoretical and Applied Genetics** 65: 77-84.

BOHNERT, H.J.; AYOUBI, P.; BORCHERT, C.; BRESSAN, R.A.; BURNAP, R.L.; CUSHMAN, J.C.; CUSHMAN, M.A.; DEYHOLOS, M.; FISHER, R.; GALBRAITH, D.W.; HASEGAWA, P.M.; JENKS, M.; KAWASAKI, S.; KOIWA, H.; KORE-EDA, S.; LEE, B.; MICHALOWSKI, C.B.; MISAWA, E.; NOMURA, M.; OZTURK, N.; POSTIER, B.; PRADE, R.; SONG, C.; TANAKA, Y.; WANG, H.; ZHU, J.K. 2001. A genomic approach towards salt stress tolerance. **Plant Physiology and Biochemistry** 39: 295-311.

BOREVITZ, J. 2004. Genomics tools for QTL analysis and gene discovery. **Current Opinion in Plant Biology** 7: 132-136.

BOUHARMONT, P.; AWEMO, J. 1979. La sélection végétative du caféier au Cameroun. **Café Cacao Thé** 23: 227-254.

BRAY, E.A. 2004. Genes commonly regulated by water-deficit stress in *Arabidopsis thaliana*. **Journal of experimental botany** 55: 2331-41.

BRIDSON, D. 1994. Additional notes on *Coffea* (Rubiaceae) from Tropical East Africa. **Kew Bulletin** 49: 331-342.

BRIDSON, D. 2003. *Coffea*. **Kew: Royal Botanic Gardens** 5: 452-463.

CAIXETA, E.T. 2008. Microsatellite markers from the Brazilian Coffee Genome Project. **Proceedings of 22 International Conference On Coffee Science**. Campinas: ICCS, CD-Rom.

CARVALHO, A. 1946. Distribuição geográfica e classificação botânica do gênero *Coffea* com referência especial à espécie Arábica. **Boletim da Superintendência dos Serviços do café** 21: 174-180.

CARVALHO, C.H. 2008. **Cultivares de café: origem, características e recomendações** (CH Carvalho). Brasília: Embrapa Café.

CARVALHO, A.; FILHO, H.P.M.; FAZUOLI, L.C.; FILHO, O.G.; LIMA, M.M. 1991. Aspectos genéticos do cafeeiro. **Revista Brasileira de Genética** 14: 135-183.

CHARRIER, A.; BERTHAUD, J. 1985. **Botanical classification of coffee** (MN Clifford and KC Willson). London & Sydney: Coffee Botany, Biochemistry and Production of Beans and Beverage.

CHEN, W.J.; ZHU, T. 2004. Networks of transcription factors with roles in environmental stress response. **Trends Plant Sci.** 9: 591-596.

CHEVALIER, A.; DAGRON, M. 1928. Recherches historiques sur les débuts de la culture du caféier en Amérique. **Communications et Actes de l'Académie des Sciences Coloniales.**

CHINI, A.; GRANT, J.J.; SEKI, M.; SHINOZAKI, K.; LOAKE, G.J. 2004. Drought tolerance established by enhanced expression of the CC-NBSLRR gene, ADR1, requires salicylic acid, EDS1 and ABI1. **Plant Journal** 38: 810-822.

CHOI, H.; HONG, J.; HA, J.; KANG, J.; KIM, S.Y. ABFs, a family of ABA-responsive element binding factors. **Journal of Biology and Chemistry** 275: 1723-1730.

CRUCES, L. 2004. Invited review : targeting transgene expression in research, agricultural, and environmental applications: promoters used in plant transformation. **In Vitro:** 1-22.

DAMATTA, F.M. 2003. Drought tolerance of two field-grown clones of *Coffea canephora*. **Plant Science** 164: 111-117.

DAMATTA, F.; RAMALHO, J. 2006. Impacts of drought and temperature stress on coffee physiology and production: a review. **Brazilian Journal of Plant Physiology** 18: 55–81.

DALE, P.J.; CLARKE, B.; FONTES, E.M. 2002. Potential for the environmental impact of transgenic crops. **Nature Biotechnology** 20: 567-574.

DAVIS, A. 2003. A new combination in *Psilanthus* (Rubiaceae) for Australia, and nomenclatural notes on *Paracoffea*. **Novon** 13: 182-184.

DAVIS, A.; BRIDSON, D.; RAKOTNASOLO, F. 2005. A re-examination of *Coffea* subgenus *Baracoffea* and *Psilanthus* (Rubiaceae-Coffeae). **In Keating, R.C., Hollowell, V.C. & Croat, T. (eds) Festschrift for William G. D'Arcy: The Legacy of a Taxonomist (Monogr. Syst. Bot. 104)**. St Louis: Missouri Bot. Gard. Press. 398–420.

DUBOUZET, J.; SAKUMA, Y.; ITO, Y.; KASUGA, M.; DUBOUZET, E.; MIURA, S.; SEKI, M.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. 2003. OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. **Plant Journal** 33: 751-763.

DUSSERT, S. 1999. **Coffee (*Coffea canephora*)**. Montpellier: Genetic diversity of cultivated tropical plants.

ETIENNE, H.; ANTHONY, F.; DUSSERT, S.; FERNANDEZ, D.; LASHERMES, P.; BERTRAND, B. 2002. Biotechnological applications for the improvement of coffee (*Coffea arabica* L.). **In Vitro Cellular & Developmental Biology-Plant** 38: 129–138.

FAHL JI. 2001. Exchange, growth, yield and beverage quality of *Coffea arabica* cultivars grafted on to *C. canephora* and *C. congensis*. **Experimental Agriculture** 37: 241-252.

FAZUOLI, L.C. 2000. Avaliação antecipada de progênies e de plantas matrizes no café Robusta. **Anais do 20 Simpósio de Pesquisa dos Cafés do Brasil**: 462-464.

FAZUOLI, L.C. 2007. Cultivares de café arábica do IAC: um patrimônio da cafeicultura brasileira. **O Agrônomo** 59: 12-15.

FOWLER, S.; THOMASHOW, F. 2002. Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. **Plant Cell Reports** 14: 1675-1690.

FREIRE, L.P. Estudo da expressão e análise de polimorfismos de genes candidatos para a tolerância à seca em cafeeiro. 2010. 117p. Dissertação (Mestrado em Biotecnologia Vegetal) - Universidade Federal de Lavras, Lavras, 2010.

GENTZBITTEL, L.; ROCHE, J.; HEWEZI, T. 2009. Real-time PCR monitoring of signal transduction related genes involved in water stress tolerance mechanism of sunflower. **Plant Physiology and Biochemistry** 47: 139-145.

HAARER, A.E. 1956. **Modern coffee production**. London: Leonard Hill.

HIGO, K.; UGAWA, Y.; IWAMOTO, M.; KORENAGA, T. 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. **Nucleic acids research** 27: 297-300.

JAKOBY, M.; WEISSHAAR, B.; DROGE-LASER, W.; VICENTE-CARBAJOSA, J.; TIEDEMANN, J.; KROJ, T.; PARCY, F. 2002. bZIP transcription factors in Arabidopsis. **Trends Plant Sci.** 7: 106-111.

KALENDAR, R.; LEE, D.; SCHULMAN, A. 2009. FastPCR Software for PCR Primer and Probe Design and Repeat Search. **Genes, Genomes and Genomics** 3: 1-14.

KASUGA, M.; LIU, Q.; MIURA, S.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. 2001. Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor. **Novartis Foundation symposium** 236: 176-86; discussion 186-9.

KIRCH, H.; PHILLIPS, J.; BARTELS, D. 2003. Dehydration-stress signal transduction. In: Scheel, D, Wasternack, C, **Plant Signal Transduction**. Midsomer Norton: Oxford University Press, 140-164.

KUMAR, L.S. 1999. DNA markers in plant improvement: an overview. **Biotechnology Advances** 17: 143-182.

LASHERMES, P. 1996. Inheritance and genetic mapping of self-incompatibility in *Coffea canephora* Pierre. **Theoretical and Applied Genetics** 93: 458-462.

LASHERMES, P.; ANDRADE, A.C.; ETIENNE, H. 2008. Genomics of coffee, one of the world's largest traded commodities. In: Jorgensen, RA, Moore, PH, Ming, R, **Genomics of Tropical Crop Plants**. New York: Springer, 203-226.

LASHERMES, P.; COMBES, M.C.; ROBERT, J.; TROUSLOT, P.; D'HONT, A.; ANTHONY, F.; CHARRIER, A. 1999. Molecular characterization and origin of the *Coffea arabica* L. genome. **Molecular and General Genetics** 261: 259-266.

LEROY, T.; RIBEYRE, F.; BERTRAND, B.; CHARMETANT, P.; DUFOUR, M.; MONTAGNON, C.; MARRACCINI, P.; POT, D. 2006. Genetics of coffee quality. **Brazilian Journal of Plant Physiology** 18: 229-242.

LIAO, P.; LEE, K.H. 2010. From SNPs to functional polymorphism: The insight into biotechnology applications. **Biochemical Engineering Journal**: 10.

LIU, Q.; KASUGA, M.; SAKUMA, Y.; ABE, H.; MIURA, S.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. 1998. Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular

signal transduction pathways in drought-and low-temperature-responsive gene expression, respectively, in Arabidopsis. **The Plant Cell Online** 10: 1391.

MARRACCINI, P.; ALVES, G.S.C.; VINECKY, F.; FREIRE, L.P.; VIEIRA, N.G.; RAMOS, H.; VIEIRA, L.G.E.; ELBELT, S.; MARQUES, T.; RODRIGUES, G.; ANDRADE, A.C. 2009. Identificação de genes com expressão diferencial em folhas de cafeeiro *Coffea canephora* e *Coffea arabica* submetidas à diferentes condições de estresse hídrico. **VI Simpósio de Pesquisa dos Cafés do Brasil**. Vitória: Simpósio de Pesquisa dos Cafés do Brasil.

MARTIN, C.; PAZ-ARES, J. 1997. MYB transcription factors in plants. **Trends Plant Sci.** 13: 67-73.

MATIELLO, J.B.; DANTAS, S.F. 1987. Desenvolvimento do cafeeiro e do sistema radicular, com e sem irrigação, em Brejões, PE. **Congresso Brasileiro de Pesquisas Cafeeiras** 14: 165.

MONTAGNON, C.; LEROY, T.; YAPO, A. 1992. Diversité génotypique et phénotypique de quelques groupes de caféiers (*Coffea canephora* Pierre) en collection. **Café Cacao Thé** 36: 187-197.

OLIVEIRA, E.P.; COSTA, F.C.; NEGREIROS, J.C.; OLIVEIRA, J.B.; ALMEIDA, M.B.; ANDRADE, R.A. 2011. Acompanhamento da Safra Brasileira Café, Safra 2011 primeira estimativa. **CONAB** : 1-25.

PINHEIRO HA. 2005. Drought tolerance is associated with rooting depth and stomatal control of water use in clones of *Coffea canephora*. **Annals of botany** 96: 101-108.

RAMACHANDRA-REDDY, A.; CHAITANYA, K.V.; VIVEKANANDAN, M. 2004. Drought-induced responses of photosynthesis and antioxidant metabolism in higher plants. **Journal of plant physiology** 161: 1189-1202.

RIZHISKY, L.; LIANG, H.; SHUMAN, J.; SHULAEV, V.; DAVLETOVA, S.; MITTLER, R. 2004. When Defense Pathways Collide. The Response of Arabidopsis to a Combination of Drought and Heat Stress. **Stress: The International Journal on the Biology of Stress** 134: 1683-1696.

ROMBAUTS, S.; DÁHAIS, P.; MONTAGU, M.V.; ROUZÉ, P. 1999. PlantCARE, a plant *cis*-acting regulatory element database. **Nucleic acids research** 27: 295-296.

SAGHAI-MAROOF, M.A.; SLOMAN, K.M.; JORGENSEN, R.A.; ALLARD, R.W. 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. **Proceedings of National Academy of Science of USA** 81: 8014-8018.

SEKI, M.; NARUSAKA, M.; ABE, H.; KASUGA, M.; YAMAGUCHI-SHINOZAKI, K.; CARNINCI, P.; HAYASHIZAKI, Y.; SHINOZAKI, K. 2001. Monitoring the expression pattern of 1300 Arabidopsis genes under drought and cold stresses by using a full-length cDNA microarray. **Plant Cell Reports** 13: 71-62.

SHEN, Y.G.; ZHANG, W.K.; HE, S.J.; ZHANG, J.S.; LIU, Q.; CHEN, S.Y. 2003. An EREBP/AP2-type protein in *Triticum aestivum* was a DRE-binding transcription factor induced by cold, dehydration and ABA stress. **Theoretical and Applied Genetics** 106: 923-930.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. 2000. Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. **Current Opinion in Plant Biology** 3: 217-223.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K.; SEKI, M. 2003. Regulatory network of gene expression in the drought and cold stress responses. **Current Opinion in Plant Biology** 6: 410-417.

SYLVAIN, P. 1955. Some observations on *Coffea arabica* L. in Ethiopia. **Turrialba (Costa Rica)** 5: 37-51.

THOMAS, A. 1942. The wild arabica coffee on the Boma Plateau, Anglo-Egyptian Sudan. **Empire Journal of Experimental Agriculture** 10: 207-212.

THOMASHOW, MF. 1999. Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. **Ann Rev Plant Physiology Plant Molecular Biology** 50: 571-599.

UMEZAWA, T.; FUJITA, M.; FUJITA, Y.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. 2006. Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. **Current Opinion in Biotechnology** 17: 113–122.

UNO, Y.; FURIHATA, T.; ABE, H.; YOSHIDA, R.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. 2000. Arabidopsis basic leucine zipper transcription factors involved in abscisic-acid-dependent signal transduction pathway under drought and high salinity conditions. **Proceedings of National Academy of Science of USA** 97: 11632-11637.

VIDAL, R.O.; MAURI, J.; ANDRADE, A.C.; PEREIRA, L.F.P.; COLOMBO, C.A.; VIEIRA, L.G.E.; CARAZZOLLE, M.F. 2010. A High-Throughput Data Mining of Single Nucleotide Polymorphisms in Coffea Species Expressed Sequence Tags Suggests Differential Homeologous Gene Expression in the Allotetraploid *Coffea arabica*. **Plant physiology** 154: 1053-1066.

VIEIRA, L.G.; ANDRADE, A.C.; COLOMBO, C.A.; CARAZZOLLE, M.F.; PEREIRA, G.A. 2006. Brazilian coffee genome project: an EST-based genomic resource. **Brazilian Journal of Plant Physiology** 18: 95-108.

VINECKY, F. Identificação e caracterização de genes candidatos para a tolerância à seca em cafeeiro. 2009. 120p. Dissertação (Mestrado em Biotecnologia Vegetal) - Universidade Federal de Lavras, Lavras, 2009.

VINOCUR, B.; ALTMAN, A. 2005. Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. **Current Opinion in Biotechnology** 16: 123–132.

XIONG, L.; SCHUMAKER, K.S.; ZHU, J. 2002. Cell signalling during cold, drought, and salt stress. **Plant Cell Reports** 14: 165-183.

YANAGUI, K. 2009. Análise *in vivo* da diversidade nucleotídica de Coffea spp. **Anais do 55 Congresso Brasileiro de Genética**. Águas de Lindóia: SBG, CD-Rom.

ZHANG, Y.; FAN, W.; KINKEMA, M.; LI, X.; DONG, X. 1999. Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the PR-1 gene. **Proceedings of National Academy of Science of USA** 96: 6523-6528.

ZHU, T. 2003. Global analysis of gene expression using GeneChip microarrays. **Current Opinion in Plant Biology** 6: 418-425.

ZHU, T.; BUDWORTH, P.; HAN, B.; BROWN, D.; CHANG, H.S.; ZOU, G.; WANG, X. 2001. Toward elucidating the global expression patterns of developing Arabidopsis: parallel analysis of 8300 genes by a high-density oligonucleotide probe array. **Plant Physiology and Biochemistry** 39: 221-242.

APÊNDICE

APÊNDICE A – Sequência utilizada para a arquitetura dos *primers*

>Sequência Promotora *DREB2A Coffea*

GGACAACCATAACCACTCCTAGTAAGCGGCACGTTGTTGGGTGCTACCCATTTTT
 ACCCGTTCAGCACGTGGTTCACGCGTTTTCCGCGGGAAACCACCGGTAGTAACTG
 TCAACCGTGTTCTGAGAACATGACGTCATTTCAAAAAGGCCATTTTGTTTCTTG
 CGTTGTGTCTCCCAAAGGCCAGAACGAAAAATAGGAAATACAAGGACACCTCTA
 GGCTCTGGCTTGTGCTTGTAGCTTTTTGTTGAATACTAGTTCGTAATCAATTA
 TTGCCTTTTCCACCAACAAAAAAAAAAAAAAAAAGAGCAACCGCTGGTAAAAAGCC
 ATAAGAATCATTAGTAGTAGTACTATAAAGAGAACAACCTGGTTCCTTGTTTTTT
 TTCATATTTTTAAAGATAAATAAATTATCGGAGTTATTTTACCAAGTTGTTTGTG
 CCCCATAAATGTCACAATACAATTAAGTTTTTTATTTCAAATGAGCTTCCGAT
 CCAAGAATATTTTCATATTAATTATGAATTAATATCTATACACTGACAGTGTATA
 TAGTTTTATTATTTGATGCATGACACATAATTTAAATTTGAATTTGAAATTCAA
 ATTTTATATATGAATCGTGCATTCAACAGCACCGTCAACTCACTGTATATAAGAT
 TTATTTATTAATTATTTAGCTAAGATAGAGTGCTCCTCAAGATTGTCACCTTCT
 ATTAATTCAGGCTTTGTCTGAAGTTAAAAGTGATCAGTTGGCTCAATTTTCTG
 ACAGTCAAACAGACGAGGAAACAAAAAGAGTGAGGCAACAAGGAATTTGACG
 TCCTCACAGAAAGAAAAGGTCAAGAGACAACCACGAAAGGATCAAAGATTATAA
 AACTAGCGAAATGAGGGGGCCCCTAATCCTTCCAGCTCAGAAGAGTCAAATCC
 CCTCCTCCACCTGGCTCCAGCAGCCAAGTTGCTGCGTTACATCAGAGCACGTGT
 CAAATGCACCATGCCTCGAACCTCAGTACTACTGTGAAACAAAGTACAACCTGG
 GGCCCGAAGACAGCTTCGAGTCGGAAGAAATCCACTTTCTTCTGCCTTAAAAG
 TTGACTGCTCCCACCTGGGGCTAAACTAGAACTATCAAATCCCCGCTCCCCG
 GCCCCAGCCTGGCTGGCTGCTTATCACCGTTCCGCGTTCGCGTCTCTTGAGCG
 TTCAATCATTCACCCACCCCCTCGAAAACCCGCCAAGTTTATATACAAACTAG
 CTCATCATCATCATCAAAACGCTCATCAAATATTTTCATATCACATCATACA
 TTCCAGTCAATTCAACCATCATTTCAATTAATCTGCATTATATATATATACCCNCT
 TTTGAGTTTCACAACACCAAGTCGTACGATTCCCTCGCTTCAGCCTTTAACTTTC
 AGTTGAATTAACCTCCGCACTGTCCACTACTACTACATTTTTTTCAAGAATGAACC
 ATTTCTCAACATGCTACTCCGACCCAATTCCCTAATCCATCATCCCCTGATTTTC

CAGACTCTTCATCAACTTCGGACAA**CGTAGTCAATGCCTGCAAAGCCA**ATTACT
 CGGACGAAGAAGTGCTCTTAGC

Primers para o isolamento da região promotora visando o sequenciamento

1° PAR

- Forward *Primer*

Primer Sequence:	ACTCCTAGTAAGCGGCACGTTGTT		
Primer Start Position:	14	Primer Length:	24
Primer T _M :	60.1 °C	Primer Self Any:	4.0
Primer GC %:	50.0 %	Primer Self End:	0.0
Primer End Stability:	4.89	Primer Penalty:	0.09

- Reverse Primer

Primer Sequence:	CCTTTCGTGGTTGTCTCTTGACCT		
Primer Start Position:	850	Primer Length:	24
Primer T _M :	58.6 °C	Primer Self Any:	3.0
Primer GC %:	50.0 %	Primer Self End:	0.0
Primer End Stability:	5.86	Primer Penalty:	1.36

Primer Pair/Product

Primer Pair Penalty:	1.45	Primer Pair Comp Any:	4.0
Primer Product Size:	837	Primer Pair Comp End:	0.0

2° PAR

- Forward *Primer*

Primer Sequence:	TCGTGCATTCAACAGCACCGTCA		
Primer Start Position:	609	Primer Length:	23
Primer T _M :	61.3 °C	Primer Self Any:	5.0
Primer GC %:	52.2 %	Primer Self End:	1.0
Primer End Stability:	6.36	Primer Penalty:	4.44

- Reverse Primer

Primer Sequence:	TGGCTTTGCAGGCATTGACTACG		
Primer Start Position:	1560	Primer Length:	23
Primer T _M :	59.9 °C	Primer Self Any:	5.0
Primer GC %:	52.2 %	Primer Self End:	3.0
Primer End Stability:	5.47	Primer Penalty:	3.23

Primer Pair/Product

Primer Pair Penalty:	7.67	Primer Pair Comp Any:	6.0
Primer Product Size:	952	Primer Pair Comp End:	2.0

APÊNDICE B – Arquitetura dos *primers* para a construção gênica

>Sequência Promotora *DREB2A Coffea*

GGACAACCATAACCAC **TCCTAGTAAGC** **GGCACGTTGT** **TTGGGTGCTACCCATTTTT**
 ACCCGTTTCAGCACGTGGTTCACGCGTTTTCCGCGGGAAACCACCGGTAGTAACTG
 TCAACCGTGTTCTGAGAACATGACGTCATTTCAAAAAGGCCATTTTTGTTTTCTTG
 CGTTGTGTCTCCCAAAGGCCAGAACGAAAAATAGGAAATACAAGGACACCTCTA
 GGCTCTGGCTTGTGCTTGTAGCTTTTTGTTGAATACTAGTTCGTAATCAATTA
 TTGCCTTTTTCCACCAACAAAAAAAAAAAAAAAAAGAGCAACCGCTGGTAAAAAGCC
 ATAAGAATCATTAGTAGTAGTACTATA **AAGAGAACAAC** **TTGGTTCTTGT** **TTTTT**
 TTCATATTTTTAAAGATAAATAAATTATCGGAGTTATTTTTACCAAGTTGTTTGTC
 CCCAAATAATGTCACAATACAATTAAGTTTTTTATTTCAAATGAGCTTCCGAT
 CCAAGAATATTTTCATATTAATTATGAATTAATATCTATACACTGACAGTGTATA
 TAGTTTTATTATTTGATGCATGACACATAATTTAAATTTGAATTTGAAATTCAA
 ATTTTATATATGAATCGTGCATTCAACAGCACCGTCACTCACTGTATATAAGAT
 TTATTTATTAATTATTTAGCTAAGATAGAGTGCTCCTCAAGATTGTCACTTTCT
 AT **TAATTCAG** **GCTTTGTCTGAAGT** **TAAAAGTGATCAGTTGGCTCAATTTTCTG**
 ACACGTCAAACAGACGAGGAAACAAAAAGAGTGAGGCAACAAGGAATTTGACG
 TCCTCACAGAAAAGAAAGGTCAAGAGACAACCACGAAAGGATCAAAGATTATAA
 AACTAGCGAAATGAGGGGGCCCCTAATCCTTCCAGCTCAGAAGAGTCAAAATCC
 CCTCCTCCACCTGGCTCCAGCAGCCAAGTTGCTGCGTTACATCAGAGCACGTGT
 CAAATGCACCATGCCTCGAACCTCAGTACTACTGTGAAACAAAGTACAACCTGG
 GGCCCGAAGACAGCTTCGAGTCGGAAGAAATCCACTTTCTTCTGCTTAAAAAG
 TTGACTGCTCCCACCTGGGGCTAAACTAGAACTATCAAAATCCCCGCTCCCCG
 GCCCCAGCCTGGCTGGCTGCTTATCACCGTTCCGCGTTCGCGTCTCTTGAGCG
 TTCAATCATTCACCCACCCCCCTCGAAAACCCGCCAAGTTTATATACAAACTAG
 CTCATCATCATCATCAAAACGCTCATCAAATATTTTCATATCACATCATACA
 TTCCAGTCAATTCAACCATCATTTCAATTAATCTGCATTATATATATATACCCNCT
 TTTGAGTTTCAACACCAAGTCGTACGATTCCCTCGCTTCAGCCTTTAACTTTT
 AGTTGAATTAACCTCCGCACTGTCCACTACTACTACATTTTT **TTCAAGAATGAACC**
ATTTCTCAACA TGCTACTCCGACCCAATTCCCTAATCCATCATCCCCTGATTTT
 CAGACTCTTCATCAACTTCGGACAACGTAGTCAATGCCTGCAAAGCCAATTACT
 CGGACGAAGAAGTGCTCTTAGC

Primers para isolar a sequência promotora do gene *DREB2A* visando a clonagem no vetor pBI121

Oligo F

- Proximal

Hind III: 5' TAATTCCAAGCTTTGTCTGAAGT 3'
 Normal: 5' TAATTCCAGGCTTTGTCTGAAGT 3'

LENGTH: 23
 GC CONTENT: 34.8 %
 MELT TEMP: 52.5 °C

MELTING TEMPERATURES
 EXACT MATCH TM: 52.5°C
 MISMATCH TM: 47.8°C
 DELTA TM: 4.7°C

PERCENT BOUND AT 52.5°C
 EXACT MATCH: 50.0%
 MISMATCH: 2.8%

- Mediano

Hind III: 5' AAGAGAACAACAAGCTTCTTGT 3'
 Normal: 5' AAGAGAACAAC TTGGTTCTTGT 3'

LENGTH: 22
 GC CONTENT: 36.4 %
 MELT TEMP: 52.7 °C

MISMATCH TM: 35.8°C

- Distal

Hind III: 5' TCCTAGTAAGCTT CACGTTGT 3'
Normal: 5' TCCTAGTAAGCGGCACGTTGT 3'

LENGTH: 21 MISMATCH TM: 47.4°C
GC CONTENT: 42.9 % DELTA TM: 4.7°C
MELT TEMP: 53.4 °C

Oligo R

- Comum a todos

Bgl II: 5' TGTTGAGAAATGGTTAGATCTTGAA 3'
Normal: 5' TGTTGAGAAATGGTTCATCTTGAA 3'

LENGTH: 25 MISMATCH TM: 48.5°C
GC CONTENT: 32.0 %
MELT TEMP: 52.5 °C

> Sequência promotora *DREB2A Coffea* alterada com o Sítio de Restrição Proximal. Utilizada para verificar a similaridade dos *primers*.

GGACAACCATAACCACTCCTAGTAAGCGGCACGTTGTTGGGTGCTACCCATTTTT
ACCCGTTTCAGCACGTGGTTCACGCGTTTTCCGCGGAAACCACCGGTAGTAACTG
TCAACCGTGTTCTGAGAACATGACGTCATTTCAAAAAGGCCATTTTGTTTCTTG
CGTTGTGTCTCCCAAAGGCCAGAACGAAAAATAGGAAATACAAGGACACCTCTA
GGCTCTGGCTTGTGCTTGTAGCTTTTTGTTGAATACTAGTTCGTAATCAATTA
TTGCCTTTTCCACCAACAAAAAAAAAAAAAAAAAGAGCAACCGCTGGTAAAAAGCC
ATAAGAATCATTAGTAGTAGTACTATAAAGAGAACAACCTGGTTCCTTGTTTTTT
TTCATATTTTTAAAGATAAATAAATTATCGGAGTTATTTTACCAAGTTGTTTGTG
CCCCAATAATGTCACAATACAATTAAGTTTTTTATTTCAAATGAGCTTCCGAT
CCAAGAATATTTTCATATTAATTATGAATTAATATCTATACTGACAGTGTATA
TAGTTTTATTATTTGATGCATGACACATAATTTAAATTTGAATTTGAAATTCAA
ATTTTATATATGAATCGTGCATTCAACAGCACCGTCACTCACTGTATATAAGAT
TTATTTATTAATTATTTAGCTAAGATAGAGTGCTCCTCAAGATTGTCACCTTCT
ATTAATTTCAAGCTTTGTCTGAAGTTAAAGTGATCAGTTGGCTCAATTTTCTG
ACACGTCAAACAGACGAGGAAACAAAAAGAGTGAGGCAACAAGGAATTTGACG
TCCTCACAGAAAAGGTCAAGAGACAACCACGAAAGGATCAAAGATTATAA
AACTAGCGAAATGAGGGGGCCCTAATCCTTCCAGCTCAGAAGAGTCAAAATCC
CCTCCTCCACCTGGCTCCAGCAGCCAAGTTGCTGCGTTACATCAGAGCACGTGT
CAAATGCACCATGCCTCGAACCTCAGTACTACTGTGAAACAAAGTACAACCTGG
GGCCGAAGACAGCTTCGAGTCGGAAGAAATCCACTTTCTTCTCCTGCCTTAAAAG
TTGACTGCTCCCACTTGGGGCTAAACTAGAACTATCAAAATCCCCGCTCCCCG
GCCCCAGCCTGGCTGGCTGCTTATCACCGTTCCGCGTTCGCGTCTCTTGAGCG
TTCAATCATTCACCCACCCCTCGAAAACCCGCCAAGTTTATATACAAACTAG
CTCATCATCATCATCAAAACGCTCATCAAATATTTTCATATCACATCATA
TTCCAGTCAATTCAACCATCATTTCAATTAATCTGCATTATATATATATACCNCT
TTTGAGTTTCACAACACCAAGTCGTACGATTCTCGCTTCAGCCTTTAACTTTC
AGTTGAATTAACCTCGCACTGTCCACTACTACTACATTTTTTCAAGATCTAACCC
ATTTCTCAACAATGCTACTCCGACCCAATTCCTAATCCATCATCCCCTGATTTTC
CAGACTCTTCATCAACTTCGGACAACGTAGTCAATGCCTGCAAAGCCAATTACT
CGGACGAAGAAGTGCTCTTAGC

> Sequência promotora *DREB2A Coffea* alterada com o Sítio de Restrição Mediano. Utilizada para verificar a similaridade dos *primers*.

```

GGACAACCATAACCACTCCTAGTAAGCGGCACGTTGTTGGGTGCTACCCATTTTT
ACCCGTTTCAGCACGTGGTTCACGCGTTTTCCGCGGGAAACCACCGGTAGTAACTG
TCAACCGTGTTCTGAGAACATGACGTCATTTCAAAAAGGCCATTTTGTTTCTTG
CGTTGTGTCTCCCAAAGGCCAGAACGAAAAATAGGAAATACAAGGACACCTCTA
GGCTCTGGCTTGTGCTTGTAGCTTTTTGTTGAATACTAGTTCGTAATCAATTA
TTGCCTTTTCCACCAACAAAAAAAAAAAAAAAAAGAGCAACCGCTGGTAAAAAGCC
ATAAGAATCATTAGTAGTAGTACTATAAAGAGAACAACAAGCTTCTTTGTTTTTT
TTCATATTTTTAAAGATAAATAAATTATCGGAGTTATTTTACCAAGTTGTTTGTC
CCCCAAATAATGTCACAATACAATTAAGTTTTTTATTTCAAATGAGCTTCCGAT
CCAAGAATATTTTCAATTAATTATGAATTAATATCTATACACTGACAGTGTATA
TAGTTTTTATTATTTGATGCATGACACATAATTTAAATTTGAATTTGAAATTCAA
ATTTTATATATGAATCGTGCATTCAACAGCACCGTCACTCACTGTATATAAGAT
TTATTTATTAATTTATTTAGCTAAGATAGAGTGCTCCTCAAGATTGTCACTTTCT
ATTAATTCAGGCTTTGTCTGAAGTTAAAAGTGATCAGTTGGCTCAATTTTCTG
ACACGTCAAACAGACGAGGAAACAAAAAGAGTGAGGCAACAAGGAATTTGACG
TCCTCACAGAAAAGAAAGGTCAAGAGACAACCACGAAAGGATCAAAGATTATAA
AACTAGCGAAATGAGGGGGCCCTAATCCTTCCAGCTCAGAAGAGTCAAAATCC
CCTCCTCCACCTGGCTCCAGCAGCCAAGTTGCTGCGTTACATCAGAGCACGTGT
CAAATGCACCATGCCTCGAACCTCAGTACTACTGTGAAACAAAGTACAACCTGG
GGCCCGAAGACAGCTTCGAGTCGGAAGAAATCCACTTTCTTCTGCCTTAAAAG
TTGACTGCTCCCACCTGGGGCTAAACTAGAACTATCAAAATCCCCGCTCCCCG
GCCCCAGCCTGGCTGGCTGCTTATCACCGTTCCGCGTTCGCGTCTCTTGAGCG
TTCAATCATTACCCCCACCCCTCGAAAACCCGCCAAGTTTATATACAAACTAG
CTCATCATCATCATCAAAACGTCATCAAATATTTTATATCACATCATAACA
TTCCAGTCAATTCAACCATCATTTCAATTAATCTGCATTATATATATATACCCNCT
TTTGAGTTTCAACACCAAGTCGTACGATTCCCTCGCTTCAGCCTTTAACTTTC
AGTTGAATTAACCTCCGCACTGTCCACTACTACTACATTTTTTCAAGATCTAACC
ATTTCTCAACAATGCTACTCCGACCCAATTCCCTAATCCATCATCCCCTGATTTTC
CAGACTCTTCATCAACTTCGGACAACGTAGTCAATGCCTGCAAAGCCAATTACT
CGGACGAAGAAGTGCTCTTAGC

```

> Sequência promotora *DREB2A Coffea* alterada com o Sítio de Restrição Distal. Utilizada para verificar a similaridade dos *primers*.

GGACAACCATACCAC**TCCTAGTAAGCTTCACGTTGT**TGGGTGCTACCCATTTTT
 ACCCGTTCAGCACGTGGTTCACGCGTTTTCCGCGGGAAACCACCGGTAGTAACTG
 TCAACCGTGTTCTGAGAACATGACGTCATTTCAAAAAGGCCATTTTGTTTTCTTG
 CGTTGTGTCTCCCAAAGGCCAGAACGAAAAATAGGAAATACAAGGACACCTCTA
 GGCTCTGGCTTGTGCTTGTAGCTTTTTGTTGAATACTAGTTCGTAATCAATTA
 TTGCCTTTTCCACCAACAAAAAAAAAAAAAAAAAGAGCAACCGCTGGTAAAAAGCC
 ATAAGAATCATTAGTAGTAGTACTATAAAGAGAACAACCTGGTTCCTTGTTTTTT
 TTCATATTTTTAAAGATAAATAAATTATCGGAGTTATTTTACCAAGTTGTTTGTC
 CCCCATAAATGTCACAATACAATTAAGTTTTTTATTTCAAATGAGCTTCCGAT
 CCAAGAATATTTTATATTAATTATGAATTAATATCTATACACTGACAGTGTATA
 TAGTTTTTATTATTTGATGCATGACACATAATTTAAATTTGAATTTGAAATTCAA
 ATTTTATATATGAATCGTGCATTCAACAGCACCGTCACTCACTGTATATAAGAT
 TTATTTATTAATTTATTTAGCTAAGATAGAGTGCTCCTCAAGATTGTCACTTTCT
 ATTAATTCAGGCTTTGTCTGAAGTTAAAAGTGATCAGTTGGCTCAATTTTCTG
 ACACGTCAAACAGACGAGGAAACAAAAAGAGTGAGGCAACAAGGAATTTGACG
 TCCTCACAGAAAAGAAAGGTCAAGAGACAACCACGAAAGGATCAAAGATTATAA
 AACTAGCGAAATGAGGGGGCCCCTAATCCTTCCAGCTCAGAAGAGTCAAAATCC
 CCTCCTCCACCTGGCTCCAGCAGCCAAGTTGCTGCGTTACATCAGAGCACGTGT
 CAAATGCACCATGCCTCGAACCTCAGTACTACTGTGAAACAAAGTACAACCTGG
 GGCCGAAGACAGCTTCGAGTCGGAAGAAATCCACTTTCTTCTGCCTTAAAAG
 TTGACTGCTCCCACTTGGGGCTAAACTAGAACTATCAAAATCCCCGCTCCCCG
 GCCCCAGCCTGGCTGGCTGCTTATCACCGTTCGCGTTCGCGTCTCTTGAGCG
 TTCAATCATTACCCCCACCCCTCGAAAACCCGCCAAGTTTATATACAACTAG
 CTCATCATCATCATCAAAACGTCATCAAATATTTTATATCACATCATA
 TTCCAGTCAATTCAACCATCATTTCAATTAATCTGCATTATATATATATACCCNCT
 TTTGAGTTTCAACACCAAGTCGTACGATTCCCTCGCTTCAGCCTTTAACTTTC
 AGTTGAATTAACCTCGCACTGTCCACTACTACTACATTTTT**TTCAAGATCTAACC**
ATTTCTCAACATGCTACTCCGACCCAATTCCCTAATCCATCATCCCCTGATTTTC
 CAGACTCTTCATCAACTTCGGACAACGTAGTCAATGCCTGCAAAGCCAATTACT
 CGGACGAAGAAGTGCTCTTAGC

APÊNDICE C – Lista dos diferentes genótipos utilizados na amplificação da região promotora do gene *DREB2A*.

Tabela 4. Lista de genótipos utilizados para verificar a conservação da região complementar ao *primers*.

Genótipos	Espécie
1. Acaia 47119	<i>C. arabica</i>
2. Bourbon	<i>C. arabica</i>
3. C1007	<i>C. canephora</i>
4. C2011	<i>C. canephora</i>
5. C3001	<i>C. canephora</i>
6. C4001	<i>C. canephora</i>
7. C4001 D.P.	<i>C. canephora</i>
8. <i>Canephora</i> (G21)	<i>C. canephora</i>
9. Catuai 25	<i>C. arabica</i>
10. Catuai 144	<i>C. arabica</i>
11. Clone 14	<i>C. canephora</i>
12. Clone 22	<i>C. canephora</i>
13. E007	<i>C. arabica</i>
14. E017	<i>C. arabica</i>
15. E123A	<i>C. arabica</i>
16. E123B	<i>C. arabica</i>
17. E238	<i>C. arabica</i>
18. E464	<i>C. arabica</i>
19. E516	<i>C. arabica</i>
20. Eugenioides	<i>C. eugenioides</i>
21. E237	<i>C. arabica</i>
22. G2020	<i>C. canephora</i>
23. Guatemalense Baixo	<i>C. arabica</i>
24. Guatemalense Alto	<i>C. arabica</i>
25. IAPAR 59	<i>C. arabica</i>
26. IAPAR 59	<i>C. arabica</i>
27. Ikatú Colombiano	<i>C. arabica</i>
28. Mundo novo	<i>C. arabica</i>
29. Obatã	<i>C. arabica</i>
30. Palma 02	<i>C. arabica</i>
31. Psilanthus	<i>C. bengalensis</i>
32. Purpurenses	
33. Racemosa	<i>C. racemosa</i>
34. RH3 Rubi	<i>C. arabica</i>
35. RH3 IAPAR 59	<i>C. arabica</i>
36. Rubi	<i>C. arabica</i>
37. Sabiá	<i>C. arabica</i>
38. San Bernardo	
39. San Ramon Baixo	
40. Sequeirio	
41. Tupi	<i>C. arabica</i>
42. Typica	<i>C. arabica</i>
43. UW002	<i>C. canephora</i>
44. UW099	<i>C. canephora</i>

APÊNDICE D – Resultado das análises dos motivos *cis* de regulação do promotor *DREB2A* no PLACE

Web Signal Scan Program

Database searched: PLACE

This is the sequence you submitted
>Contig Promotor do gene *DREB2A*, 1454 bases, 3768F10
checksum.

```

GGACAACCATACCACTCCTAGTAAGCGGCACGTTGTTGGGTGCTACCCAT
TTTTACCCGTTTCAGCACGTGGTTTCACGCGTTTTCCGCGGGAAACCACCGGT
AGTAACTGTCAACCGTGTCTGAGAACATGACGTCATTTCAAAAAGGCCA
TTTTGTTTCTTGCGTTGTGTCTCCCAAAGGCCAGAACGAAAAATAGGAAA
TACAAGGACACCTCTAGGCTCTGGCTTGTTCCTTGTAGCTTTTGTGAA
TACTAGTTCGTAATCAATTATTGCCTTTTCCACCAACAAAAAAAAAAAAA
AAGAGCAACCGCTGGTAAAAAGCCATAAGAATCATTAGTAGTAGTACTAT
AAAGAGAACAACCTGGTTCTTGTTTTTTTTTTCATATTTTAAAGATAAATAA
ATTATCGGAGTTATTTTACCAAGTTGTTTGTCCCCCAAATAATGTCACAA
TACAATTAAGTTTTTTTATTTCAAATGAGCTTCCGATCCAAGAATATTTCA
TATTAATTATGAATTAATATCTATACACTGACAGTGTATATAGTTTTTATT
ATTTGATGCATGACACATAATTTAAATTTGAATTTGAAATTCAAATTTTA
TATATGAATCGTGCATTC AACAGCACCGTCACTCACTGTATATAAGATTT
ATTTATTAATTATTTAGCTAAGATAGAGTGCTCCTCAAGATTGTCACCTT
CTATTAATTCCAGGCTTTGTCTGAAGTTAAAAGTGATCAGTTGGCTCAAT
TTTCTGACACGTCAAACAGACGAGGAAACAAAAAAGAGTGAGGCAACAAG
GAATTTGACGTCCCTCACAGAAAGAAAAGGTCAAGAGACAACCACGAAAGG
ATCAAAGATTATAAAACTAGCGAAATGAGGGGGCCCTAATCCTTCCAGC
TCAGAAGAGTCAAAATCCCCTCCTCCACCTGGCTCCAGCAGCCAAGTTGC
TGCGTTACATCAGAGCACGTGTCAAATGCACCATGCCTCGAACCTCAGTA
CTACTGTGAAACAAAAGTACAACCTTGGGGCCCCGAAGACAGCTTCGAGTCGG
AAGAAATCCACTTTCTTCTGCCTTAAAAGTTGACTGCTCCCACTTGGGG
CTAAACTAGAACTATCAAAATCCCCGCTCCCCGGCCCCAGCCTGGCTG
GCTGCTTATCACCGTTCCGCGTTCGCGTCTCTTGAGCGTTCAATCATTCA
CCCCACCCCTCGAAAACCCGCCAAGTTTATATACAACTAGCTCATCAT
CATCATCATCAAAACGCTCATCAAATATTTTCATATCACATCATAACATTCC
AGTCAATTC AACCATCATTTCAATTAATCTGCATTATATATATATACCCNCT
TTTGAGTTTCACAACACCAAGTCGTACGATTCCCTCGCTTCAGCCTTTAAC
TTTCAGTTGAATTAACCTCCGCACTGTCCACTACTACTACATTTTTTTTCAAG
AATG

```

RESULTS OF YOUR SIGNAL SCAN SEARCH REQUEST

../tmp/sigscan//signalseqdone.7432: 1454 base pairs

Signal database file: user.dat

Factor or Site Name		Loc.(Str.)	Signal Sequence	SITE #
SV40COREENHAN	site	8 (-)	GTGGWWHG	S000123
CACTFTPPCA1	site	13 (+)	YACT	S000449
CACTFTPPCA1	site	20 (-)	YACT	S000449
RHERPATEXPA7	site	28 (+)	KCACGW	S000512
ABRERATCAL	site	28 (-)	MACGYGB	S000507
ABRELATERD1	site	29 (-)	ACGTG	S000414
BP5OSWX	site	29 (-)	CAACGTG	S000436
T/GBOXATPIN2	site	29 (-)	AACGTG	S000458
ACGTATERD1	site	30 (+)	ACGT	S000415
ACGTATERD1	site	30 (-)	ACGT	S000415
RAV1AAT	site	34 (-)	CAACA	S000314
SEF4MOTIFGM7S	site	49 (+)	RTTTTTR	S000103
GT1CONSENSOS	site	52 (-)	GRWAAW	S000198
MYBCOREATCYCB1	site	57 (-)	AACGG	S000502
RHERPATEXPA7	site	64 (+)	KCACGW	S000512
ABRERATCAL	site	64 (-)	MACGYGB	S000507
ABRERATCAL	site	65 (+)	MACGYGB	S000507
CACGTGMOTIF	site	65 (+)	CACGTG	S000042
EBOXBNNAPA	site	65 (+)	CANNTG	S000144
IRO2OS	site	65 (+)	CACGTGG	S000505
MYCCONSOSAT	site	65 (+)	CANNTG	S000407
ABRELATERD1	site	65 (-)	ACGTG	S000414
CACGTGMOTIF	site	65 (-)	CACGTG	S000042
EBOXBNNAPA	site	65 (-)	CANNTG	S000144
MYCCONSOSAT	site	65 (-)	CANNTG	S000407
ABRELATERD1	site	66 (+)	ACGTG	S000414
ACGTATERD1	site	66 (+)	ACGT	S000415
ACGTATERD1	site	66 (-)	ACGT	S000415
GTGANTG10	site	73 (-)	GTGA	S000378
ABRERATCAL	site	74 (+)	MACGYGB	S000507
CGCGBOXAT	site	75 (+)	VCGCGB	S000501
ABRERATCAL	site	75 (-)	MACGYGB	S000507
CGCGBOXAT	site	75 (-)	VCGCGB	S000501
CGCGBOXAT	site	83 (+)	VCGCGB	S000501
CGCGBOXAT	site	83 (-)	VCGCGB	S000501
E2F1OSPCNA	site	85 (+)	GCGGGAAA	S000396
E2FANTRNR	site	85 (-)	TTTCCCGC	S000366
E2FCONSOSOS	site	85 (-)	WTTSSCSS	S000476

MYB1AT	site	90	(+)	WAACCA	S000408
D1GMAUX28	site	100	(-)	ACAGTTACTA	S000328
CACTFTPPCA1	site	101	(-)	YACT	S000449
MYB2AT	site	103	(+)	TAACTG	S000177
MYB2CONSENSOSAT	site	103	(+)	YAACKG	S000409
MYBCORE	site	103	(-)	CNGTTR	S000176
BIHD1OS	site	107	(+)	TGTCA	S000498
WBOXATNPR1	site	108	(-)	TTGAC	S000390
WRKY71OS	site	108	(-)	TGAC	S000447
MYBCORE	site	110	(-)	CNGTTR	S000176
ASF1MOTIFCAMV	site	129	(+)	TGACG	S000024
PALINDROMICBOXGM	site	129	(+)	TGACGTCA	S000255
TGACGTVMAMY	site	129	(+)	TGACGT	S000377
WRKY71OS	site	129	(+)	TGAC	S000447
HEXMOTIFTAH3H4	site	129	(-)	ACGTCA	S000053
PALINDROMICBOXGM	site	129	(-)	TGACGTCA	S000255
ACGTCBOX	site	130	(+)	GACGTC	S000131
ACGTCBOX	site	130	(-)	GACGTC	S000131
ACGTATERD1	site	131	(+)	ACGT	S000415
HEXMOTIFTAH3H4	site	131	(+)	ACGTCA	S000053
ACGTATERD1	site	131	(-)	ACGT	S000415
TGACGTVMAMY	site	131	(-)	TGACGT	S000377
ASF1MOTIFCAMV	site	132	(-)	TGACG	S000024
WRKY71OS	site	133	(-)	TGAC	S000447
ERELEE4	site	136	(+)	AWTTCAAA	S000037
PYRIMIDINEBOXOSRAMY1A	site	142	(-)	CCTTTT	S000259
DOFCOREZM	site	143	(+)	AAAG	S000265
ANAERO1CONSENSOS	site	152	(-)	AAACAAA	S000477
POLLEN1LELAT52	site	156	(-)	AGAAA	S000245
ARFAT	site	168	(+)	TGTCTC	S000270
SURECOREATSULTR11	site	169	(-)	GAGAC	S000499
DOFCOREZM	site	176	(+)	AAAG	S000265
GT1CONSENSOS	site	188	(+)	GRWAAW	S000198
GT1GMSCAM4	site	188	(+)	GAAAAA	S000453
GT1CONSENSOS	site	196	(+)	GRWAAW	S000198
RAV1AAT	site	227	(-)	CAACA	S000314
DOFCOREZM	site	240	(-)	AAAG	S000265
RAV1AAT	site	244	(-)	CAACA	S000314
-10PEHVPSBD	site	247	(-)	TATTCT	S000392
CACTFTPPCA1	site	251	(+)	YACT	S000449
ARR1AT	site	262	(-)	NGATT	S000454
ATHB1ATCONSENSOS	site	265	(+)	CAATWATTG	S000317
ATHB5ATCORE	site	265	(+)	CAATNATTG	S000371
CAATBOX1	site	265	(+)	CAAT	S000028
ATHB1ATCONSENSOS	site	265	(-)	CAATWATTG	S000317
ATHB5ATCORE	site	265	(-)	CAATNATTG	S000371
POLASIG3	site	267	(-)	AATAAT	S000088

CAATBOX1	site	270	(-)	CAAT	S000028
PYRIMIDINEBOXOSRAMY1A	site	274	(+)	CCTTTT	S000259
DOFCOREZM	site	275	(-)	AAAG	S000265
GT1CONSENSOS	site	276	(-)	GRWAAW	S000198
RAV1AAT	site	284	(+)	CAACA	S000314
MARTBOX	site	288	(-)	TTWTWTTWTT	S000067
MARTBOX	site	289	(-)	TTWTWTTWTT	S000067
MARTBOX	site	290	(-)	TTWTWTTWTT	S000067
MARTBOX	site	291	(-)	TTWTWTTWTT	S000067
MARTBOX	site	292	(-)	TTWTWTTWTT	S000067
MARTBOX	site	293	(-)	TTWTWTTWTT	S000067
DOFCOREZM	site	300	(+)	AAAG	S000265
NODCON2GM	site	301	(-)	CTCTT	S000462
OSE2ROOTNODULE	site	301	(-)	CTCTT	S000468
MYBCORE	site	306	(-)	CNGTTR	S000176
GT1CONSENSOS	site	314	(+)	GRWAAW	S000198
DOFCOREZM	site	319	(+)	AAAG	S000265
ARR1AT	site	330	(-)	NGATT	S000454
CACTFTPPCA1	site	337	(-)	YACT	S000449
CACTFTPPCA1	site	340	(-)	YACT	S000449
CACTFTPPCA1	site	343	(-)	YACT	S000449
CURECORECR	site	344	(+)	GTAC	S000493
CURECORECR	site	344	(-)	GTAC	S000493
CACTFTPPCA1	site	345	(+)	YACT	S000449
TAAAGSTKST1	site	350	(+)	TAAAG	S000387
DOFCOREZM	site	351	(+)	AAAG	S000265
NODCON2GM	site	352	(-)	CTCTT	S000462
OSE2ROOTNODULE	site	352	(-)	CTCTT	S000468
REALPHALGLHCB21	site	363	(-)	AACCAA	S000362
GT1CONSENSOS	site	376	(-)	GRWAAW	S000198
GT1GMSCAM4	site	376	(-)	GAAAAA	S000453
ROOTMOTIFTAPOX1	site	382	(+)	ATATT	S000098
LECPLEACS2	site	382	(-)	TAAAATAT	S000465
TAAAGSTKST1	site	388	(+)	TAAAG	S000387
DOFCOREZM	site	389	(+)	AAAG	S000265
NODCON1GM	site	389	(+)	AAAGAT	S000461
OSE1ROOTNODULE	site	389	(+)	AAAGAT	S000467
GATABOX	site	392	(+)	GATA	S000039
GT1CONSENSOS	site	392	(+)	GRWAAW	S000198
IBOXCORE	site	392	(+)	GATAA	S000199
TATABOX5	site	395	(-)	TTATTT	S000203
POLASIG1	site	396	(+)	AATAAA	S000080
GT1CONSENSOS	site	401	(-)	GRWAAW	S000198
IBOXCORE	site	402	(-)	GATAA	S000199
GATABOX	site	403	(-)	GATA	S000039
TATABOX5	site	411	(+)	TTATTT	S000203
GT1CONSENSOS	site	415	(-)	GRWAAW	S000198

TATABOX5	site	437	(-)	TTATTT	S000203
POLASIG3	site	438	(+)	AATAAT	S000088
BIHD1OS	site	443	(+)	TGTCA	S000498
TGTCACACMCUCUMISIN	site	443	(+)	TGTCACA	S000422
WRKY71OS	site	444	(-)	TGAC	S000447
GTGANTG10	site	445	(-)	GTGA	S000378
CAATBOX1	site	448	(+)	CAAT	S000028
CAATBOX1	site	453	(+)	CAAT	S000028
POLASIG1	site	464	(-)	AATAAA	S000080
TATABOX5	site	465	(+)	TTATTT	S000203
ERELEE4	site	467	(+)	AWTTCAAA	S000037
EBOXBNNAPA	site	471	(+)	CANNTG	S000144
MYCCONSENSOSAT	site	471	(+)	CANNTG	S000407
EBOXBNNAPA	site	471	(-)	CANNTG	S000144
MYCCONSENSOSAT	site	471	(-)	CANNTG	S000407
-10PEHVPSBD	site	490	(-)	TATTCT	S000392
ROOTMOTIFTAPOX1	site	492	(-)	ATATT	S000098
ROOTMOTIFTAPOX1	site	493	(+)	ATATT	S000098
ROOTMOTIFTAPOX1	site	500	(+)	ATATT	S000098
TATABOX3	site	501	(+)	TATTAAT	S000110
TATABOX3	site	513	(-)	TATTAAT	S000110
ROOTMOTIFTAPOX1	site	516	(-)	ATATT	S000098
GATABOX	site	518	(-)	GATA	S000039
CACTFTPPCA1	site	526	(+)	YACT	S000449
WRKY71OS	site	529	(+)	TGAC	S000447
BIHD1OS	site	529	(-)	TGTCA	S000498
CACTFTPPCA1	site	533	(-)	YACT	S000449
MARTBOX	site	544	(+)	TTWTWTTWTT	S000067
POLASIG1	site	545	(-)	AATAAA	S000080
POLASIG3	site	548	(-)	AATAAT	S000088
TATABOX5	site	549	(+)	TTATTT	S000203
RYREPEATGMGY2	site	556	(-)	CATGCAT	S000105
RYREPEATLEGUMINBOX	site	556	(-)	CATGCAY	S000100
RYREPEATBNNAPA	site	557	(-)	CATGCA	S000264
WRKY71OS	site	561	(+)	TGAC	S000447
BIHD1OS	site	561	(-)	TGTCA	S000498
ERELEE4	site	577	(-)	AWTTCAAA	S000037
ERELEE4	site	583	(-)	AWTTCAAA	S000037
EECCRCAH1	site	586	(-)	GANTTNC	S000494
ERELEE4	site	588	(+)	AWTTCAAA	S000037
TATAPVTRNALEU	site	597	(+)	TTTATATA	S000340
TATABOX4	site	598	(-)	TATATAA	S000111
ARR1AT	site	607	(-)	NGATT	S000454
RHERPATEXPA7	site	609	(-)	KCACGW	S000512
RAV1AAT	site	618	(+)	CAACA	S000314
MYBCORE	site	618	(-)	CNGTTR	S000176
ARE1	site	623	(-)	RGTGACNNGC	S000022

ASF1MOTIFCAMV	site	627 (-)	TGACG	S000024
WRKY71OS	site	628 (-)	TGAC	S000447
GTGANTG10	site	629 (-)	GTGA	S000378
CACTFTPPCA1	site	630 (+)	YACT	S000449
GTGANTG10	site	633 (-)	GTGA	S000378
CACTFTPPCA1	site	634 (+)	YACT	S000449
TATABOX4	site	639 (+)	TATATAA	S000111
ARR1AT	site	645 (+)	NGATT	S000454
MARABOX1	site	648 (-)	AATAAAAYAAA	S000063
POLASIG1	site	648 (-)	AATAAA	S000080
TATABOX5	site	649 (+)	TTATTT	S000203
POLASIG1	site	652 (-)	AATAAA	S000080
TATABOX3	site	654 (+)	TATTAAT	S000110
POLASIG3	site	659 (-)	AATAAT	S000088
TATABOX5	site	660 (+)	TTATTT	S000203
GATABOX	site	672 (+)	GATA	S000039
CACTFTPPCA1	site	677 (-)	YACT	S000449
ARR1AT	site	688 (+)	NGATT	S000454
CAATBOX1	site	690 (-)	CAAT	S000028
SEBFCONSSTPR10A	site	691 (+)	YTGTCWC	S000391
BIHD1OS	site	692 (+)	TGTCA	S000498
WRKY71OS	site	693 (-)	TGAC	S000447
GTGANTG10	site	694 (-)	GTGA	S000378
CACTFTPPCA1	site	695 (+)	YACT	S000449
DOFCOREZM	site	697 (-)	AAAG	S000265
POLLEN1LELAT52	site	698 (-)	AGAAA	S000245
BOXIINTPATPB	site	699 (-)	ATAGAA	S000296
TATABOX3	site	702 (+)	TATTAAT	S000110
DOFCOREZM	site	715 (-)	AAAG	S000265
DOFCOREZM	site	730 (+)	AAAG	S000265
CACTFTPPCA1	site	732 (-)	YACT	S000449
GTGANTG10	site	733 (+)	GTGA	S000378
EBOXBNNAPA	site	738 (+)	CANNTG	S000144
MYBCORE	site	738 (+)	CNGTTR	S000176
MYCCONSSENSOSAT	site	738 (+)	CANNTG	S000407
EBOXBNNAPA	site	738 (-)	CANNTG	S000144
MYB2CONSSENSOSAT	site	738 (-)	YAACKG	S000409
MYCCONSSENSOSAT	site	738 (-)	CANNTG	S000407
INRNTPSADB	site	745 (+)	YTCANTYY	S000395
CAATBOX1	site	747 (+)	CAAT	S000028
GT1CONSSENSOS	site	749 (-)	GRWAAW	S000198
POLLEN1LELAT52	site	751 (-)	AGAAA	S000245
WRKY71OS	site	755 (+)	TGAC	S000447
BIHD1OS	site	755 (-)	TGTCA	S000498
ACGTABREMOTIFA2OSEM	site	756 (-)	ACGTGKC	S000394
GADOWNAT	site	756 (-)	ACGTGTC	S000438
ABRELATERD1	site	758 (-)	ACGTG	S000414

CGTATERD1	site	759 (+)	ACGT	S000415
HEXMOTIFTAH3H4	site	759 (+)	ACGTCA	S000053
ACGTATERD1	site	759 (-)	ACGT	S000415
TGACGTVMAMY	site	759 (-)	TGACGT	S000377
ASF1MOTIFCAMV	site	760 (-)	TGACG	S000024
WBOXATNPR1	site	761 (-)	TTGAC	S000390
WRKY71OS	site	761 (-)	TGAC	S000447
ANAERO1CONSENSOS	site	776 (+)	AAACAAA	S000477
DOFCOREZM	site	783 (+)	AAAG	S000265
NODCON2GM	site	784 (-)	CTCTT	S000462
OSE2ROOTNODULE	site	784 (-)	CTCTT	S000468
INRNTPSADB	site	785 (-)	YTCANTYY	S000395
SORLIP5AT	site	786 (+)	GAGTGAG	S000486
CACTFTPPCA1	site	787 (-)	YACT	S000449
GTGANTG10	site	788 (+)	GTGA	S000378
RAV1AAT	site	794 (+)	CAACA	S000314
WBOXATNPR1	site	805 (+)	TTGAC	S000390
ASF1MOTIFCAMV	site	806 (+)	TGACG	S000024
TGACGTVMAMY	site	806 (+)	TGACGT	S000377
WRKY71OS	site	806 (+)	TGAC	S000447
HEXMOTIFTAH3H4	site	806 (-)	ACGTCA	S000053
ACGTCBOX	site	807 (+)	GACGTC	S000131
ACGTCBOX	site	807 (-)	GACGTC	S000131
ACGTATERD1	site	808 (+)	ACGT	S000415
ACGTATERD1	site	808 (-)	ACGT	S000415
GTGANTG10	site	814 (-)	GTGA	S000378
POLLEN1LELAT52	site	818 (+)	AGAAA	S000245
DOFCOREZM	site	820 (+)	AAAG	S000265
POLLEN1LELAT52	site	822 (+)	AGAAA	S000245
PYRIMIDINEBOXOSRAMY1A	site	824 (-)	CCTTTT	S000259
DOFCOREZM	site	825 (+)	AAAG	S000265
QELEMENTZMZM13	site	827 (+)	AGGTCA	S000254
ELRECOREPCR1	site	828 (-)	TTGACC	S000142
WBOXNTERF3	site	828 (-)	TGACY	S000457
WBOXATNPR1	site	829 (-)	TTGAC	S000390
WRKY71OS	site	829 (-)	TGAC	S000447
NODCON2GM	site	832 (-)	CTCTT	S000462
OSE2ROOTNODULE	site	832 (-)	CTCTT	S000468
SURECOREATSULTR11	site	834 (+)	GAGAC	S000499
ARFAT	site	834 (-)	TGTCTC	S000270
SEBFCONSSTPR10A	site	834 (-)	YTGTCWC	S000391
DOFCOREZM	site	846 (+)	AAAG	S000265
DOFCOREZM	site	854 (+)	AAAG	S000265
NODCON1GM	site	854 (+)	AAAGAT	S000461
OSE1ROOTNODULE	site	854 (+)	AAAGAT	S000467
ARR1AT	site	856 (+)	NGATT	S000454
INRNTPSADB	site	872 (-)	YTCANTYY	S000395

SORLIP2AT	site	881	(+)	GGGCC	S000483
SORLIP2AT	site	882	(-)	GGGCC	S000483
ARR1AT	site	889	(-)	NGATT	S000454
NODCON2GM	site	905	(-)	CTCTT	S000462
OSE2ROOTNODE	site	905	(-)	CTCTT	S000468
WBOXPCWRKY1	site	908	(-)	TTTGACY	S000310
WBOXHVIS01	site	908	(-)	TGACT	S000442
WBOXNTERF3	site	908	(-)	TGACY	S000457
WBOXATNPR1	site	909	(-)	TTGAC	S000390
WRKY71OS	site	909	(-)	TGAC	S000447
ARR1AT	site	914	(-)	NGATT	S000454
SITEIOSPCNA	site	925	(-)	CCAGGTGG	S000224
EBOXBNNAPA	site	926	(+)	CANNTG	S000144
MYCCONSENSOSAT	site	926	(+)	CANNTG	S000407
RAV1BAT	site	926	(+)	CACCTG	S000315
EBOXBNNAPA	site	926	(-)	CANNTG	S000144
MYCCONSENSOSAT	site	926	(-)	CANNTG	S000407
ANAERO2CONSENSOS	site	937	(+)	AGCAGC	S000478
RHERPATEXPA7	site	965	(+)	KCACGW	S000512
ABRERATCAL	site	965	(-)	MACGYGB	S000507
ABRERATCAL	site	966	(+)	MACGYGB	S000507
CACGTGMOTIF	site	966	(+)	CACGTG	S000042
EBOXBNNAPA	site	966	(+)	CANNTG	S000144
MYCCONSENSOSAT	site	966	(+)	CANNTG	S000407
ABRELATERD1	site	966	(-)	ACGTG	S000414
CACGTGMOTIF	site	966	(-)	CACGTG	S000042
DPBFCOREDCDC3	site	966	(-)	ACACNNG	S000292
EBOXBNNAPA	site	966	(-)	CANNTG	S000144
MYCCONSENSOSAT	site	966	(-)	CANNTG	S000407
ABRELATERD1	site	967	(+)	ACGTG	S000414
ACGTABREMOTIFA2OSEM	site	967	(+)	ACGTGKC	S000394
ACGTATERD1	site	967	(+)	ACGT	S000415
GADOWNAT	site	967	(+)	ACGTGTC	S000438
ACGTATERD1	site	967	(-)	ACGT	S000415
BIHD1OS	site	970	(+)	TGTCA	S000498
WBOXATNPR1	site	971	(-)	TTGAC	S000390
WRKY71OS	site	971	(-)	TGAC	S000447
EBOXBNNAPA	site	973	(+)	CANNTG	S000144
MYCCONSENSOSAT	site	973	(+)	CANNTG	S000407
EBOXBNNAPA	site	973	(-)	CANNTG	S000144
MYCCONSENSOSAT	site	973	(-)	CANNTG	S000407
CACTFTPPCA1	site	997	(-)	YACT	S000449
CURECORECR	site	998	(+)	GTAC	S000493
CURECORECR	site	998	(-)	GTAC	S000493
CACTFTPPCA1	site	999	(+)	YACT	S000449
CACTFTPPCA1	site	1002	(+)	YACT	S000449
GTGANTG10	site	1006	(+)	GTGA	S000378

ANAERO1CONSENSOS	site	1009	(+)	AAACAAA	S000477
TBOXATGAPB	site	1012	(-)	ACTTTG	S000383
DOFCOREZM	site	1013	(+)	AAAG	S000265
CACTFTPPCA1	site	1015	(-)	YACT	S000449
CURECORECR	site	1016	(+)	GTAC	S000493
CURECORECR	site	1016	(-)	GTAC	S000493
SORLIP2AT	site	1026	(+)	GGGCC	S000483
SORLIP2AT	site	1027	(-)	GGGCC	S000483
LTRECOREATCOR15	site	1046	(-)	CCGAC	S000153
POLLEN1LELAT52	site	1052	(+)	AGAAA	S000245
ARR1AT	site	1055	(-)	NGATT	S000454
CACTFTPPCA1	site	1059	(+)	YACT	S000449
DOFCOREZM	site	1061	(-)	AAAG	S000265
POLLEN1LELAT52	site	1062	(-)	AGAAA	S000245
DOFCOREZM	site	1077	(+)	AAAG	S000265
WBOXATNPR1	site	1081	(+)	TTGAC	S000390
WBOXHVIS01	site	1082	(+)	TGACT	S000442
WBOXNTERF3	site	1082	(+)	TGACY	S000457
WRKY71OS	site	1082	(+)	TGAC	S000447
BOXCPSAS1	site	1088	(+)	CTCCCAC	S000226
CACTFTPPCA1	site	1092	(+)	YACT	S000449
EBOXBNNAPA	site	1092	(+)	CANNTG	S000144
MYCCONSENSOSAT	site	1092	(+)	CANNTG	S000407
EBOXBNNAPA	site	1092	(-)	CANNTG	S000144
MYCCONSENSOSAT	site	1092	(-)	CANNTG	S000407
POLLEN1LELAT52	site	1108	(+)	AGAAA	S000245
GATABOX	site	1114	(-)	GATA	S000039
ARR1AT	site	1120	(-)	NGATT	S000454
BS1EGCCR	site	1124	(-)	AGCGGG	S000352
SORLIP2AT	site	1134	(-)	GGGCC	S000483
ANAERO2CONSENSOS	site	1151	(-)	AGCAGC	S000478
IBOX	site	1155	(-)	GATAAG	S000124
IBOXCORE	site	1156	(-)	GATAA	S000199
GATABOX	site	1157	(-)	GATA	S000039
GTGANTG10	site	1159	(-)	GTGA	S000378
MYBCOREATCYCB1	site	1162	(-)	AACGG	S000502
CGCGBOXAT	site	1167	(+)	VCGCGB	S000501
ABRERATCAL	site	1167	(-)	MACGYGB	S000507
CGCGBOXAT	site	1167	(-)	VCGCGB	S000501
SURECOREATSULTR11	site	1177	(-)	GAGAC	S000499
NODCON2GM	site	1179	(+)	CTCTT	S000462
OSE2ROOTNODULE	site	1179	(+)	CTCTT	S000468
CAATBOX1	site	1191	(+)	CAAT	S000028
ARR1AT	site	1192	(-)	NGATT	S000454
GTGANTG10	site	1198	(-)	GTGA	S000378
TATAPVTRNALEU	site	1227	(+)	TTTATATA	S000340
TATABOX4	site	1228	(-)	TATATAA	S000111

CDA1ATCAB2	site	1260	(+)	CAAAACGC	S000440
ROOTMOTIFTAPOX1	site	1274	(-)	ATATT	S000098
ROOTMOTIFTAPOX1	site	1275	(+)	ATATT	S000098
GATABOX	site	1283	(-)	GATA	S000039
GTGANTG10	site	1285	(-)	GTGA	S000378
WBOXHVIS01	site	1301	(-)	TGACT	S000442
WBOXNTERF3	site	1301	(-)	TGACY	S000457
WBOXATNPR1	site	1302	(-)	TTGAC	S000390
WRKY71OS	site	1302	(-)	TGAC	S000447
CAATBOX1	site	1304	(+)	CAAT	S000028
CAATBOX1	site	1321	(+)	CAAT	S000028
ARR1AT	site	1326	(-)	NGATT	S000454
TATABOX4	site	1334	(-)	TATATAA	S000111
PYRIMIDINEBOXOSRAMY1A	site	1348	(+)	CCTTTT	S000259
DOFCOREZM	site	1349	(-)	AAAG	S000265
EECCRCAH1	site	1354	(+)	GANTTNC	S000494
GTGANTG10	site	1359	(-)	GTGA	S000378
RAV1AAT	site	1362	(+)	CAACA	S000314
CURECORECR	site	1374	(+)	GTAC	S000493
CURECORECR	site	1374	(-)	GTAC	S000493
ARR1AT	site	1377	(+)	NGATT	S000454
DOFCOREZM	site	1394	(-)	AAAG	S000265
TAAAGSTKST1	site	1394	(-)	TAAAG	S000387
DOFCOREZM	site	1400	(-)	AAAG	S000265
EBOXBNNAPA	site	1404	(+)	CANNTG	S000144
MYBCORE	site	1404	(+)	CNGTTR	S000176
MYCCONSENSOSAT	site	1404	(+)	CANNTG	S000407
EBOXBNNAPA	site	1404	(-)	CANNTG	S000144
MYB2CONSENSOSAT	site	1404	(-)	YAACKG	S000409
MYCCONSENSOSAT	site	1404	(-)	CANNTG	S000407
CACTFTPPCA1	site	1421	(+)	YACT	S000449
CACTFTPPCA1	site	1428	(+)	YACT	S000449
CACTFTPPCA1	site	1431	(+)	YACT	S000449
CACTFTPPCA1	site	1434	(+)	YACT	S000449
GT1CONSENSOS	site	1442	(-)	GRWAAW	S000198
GT1GMSCAM4	site	1442	(-)	GAAAAA	S000453

APÊNDICE E – Resultado das análises dos motivos *cis* de regulação do promotor *DREB2A* no PlantCARE

>Promotor do gene DREB 1439nt

```

+ CTCTAGTAA GCGGCACGTT GTTGGGTGCT ACCCATTTTT ACCCGTTCAG CACGTGGTTC ACGCGTTTTCC
- GAGGATCATT CGCCGTGCAA CAACCCACGA TGGGTAAAAA TGGGCAAGTC GTGCACCAAG TGCGCAAAGG

+ GCGGGAAACC ACCGGTAGTA ACTGTCAACC GTGTCTGAG AACATGACGT CATTTCAAA AGGCCATTTT
- CGCCCTTTGG TGCCCATCAT TGACAGTTGG CACAAGACTC TTGTACTGCA GTAAAAGTTT TCCGGTAAAA

+ GTTTCITGCG TTGTGTCTCC CAAAGGCCAG AACGAAAAAT AGGAAATACA AGGACACCTC TAGGCTCTGG
- CAAAGAACGC AACACAGAGG GTTTCGGTGC TTGCTTTTAA TCCTTTATGT TCCTGTGGAG ATCCGAGACC

+ CTGTGTGCGT TGTAGCTTTT GTTGAATACT AGTTCGTAAT CAATTATTGC CTTTCCACC AACAAAAAAA
- GAACACGGGA ACATCGAAAA CAACCTATGA TCAAGCATTG GATAATAACG GAAAAGGTGG TTGTTTTTTT

+ AAAAAAAGA GCAACCGCTG GTAAAAAGCC ATAAGAATCA TTAGTAGTAG TACTATAAAG AGAACAACTT
- TTTTTTTTCT CGTTGGCGAC CATTTCGCG TATTCTTAGT AATCATCATC ATGATATTTT TCTTGTGAA

+ GGTTCTGTGT TTTTTTCATA TTTTAAGAT AAATAAATTA TCGGAGTAT TTTACCAAGT TGTTGTCCC
- CCAAGAACAA AAAAAAGTAT AAAAATTTCTA TTTATTTAAT AGCCTCAATA AAATGGTTCA ACAAAACAGGG

+ CCAAAATATG TCACAATACA ATTAAGTTT TTAATTTCAA TGAGCTCCCG ATCCAAGAAT ATTTCAATTT
- GGTTTATTAC AGTGTATGT TAATTCAAAA AATAAAGTTT ACTCGAAGGC TAGGTTCTTA TAAAGTATAA

+ AATTATGAAT TAATTCAT ACACTGACAG TGTATATAGT TTTATTATT GATGCATGAC ACATAATTTA
- TTAATCTTA ATATAGATA TGTACTGTC ACATAATCA AATTAATAA CTACGTACTG TGTATTAAAT

+ AATTGAATT TGAAATCAA ATTTATAATA TGAATCGTGC AITCAACAGC ACCGTCACTC ACTGTATAA
- TTAAGTAA ACTTTAAGTT TAAAAATAAT ACTTAGCAGC TAAGTTGTGC TGGCAGTGAG TGACATAAT

+ AGATTTATTT ATTAATTAIT TAGCTAAGAT AGAGTGTCC TCAAGATTGT CACTTTCTAT TAATCCAGG
- TCTAATAATA TAATTAATA ATCGATTCTA TCTCAGGAG AGTCTAAGC GTGAAAGATA ATTAAGGTCC

+ CTTTGTCTGA AGTTAAAAGT GATCAGTTGG CTCAAITTTT TGACACGTCA AACAGACGAG GAAACAAAA
- GAAACAGACT TCAATTTCA CTAGTCAACC GAGTTAAAAG ACTGTGCGAT TTGTCTGCTC CTTTGTTTT

+ AGAGTGAGGC AACAAAGGAT TTGACGTCTT CACAGAAAGA AAAGGTCAAG AGACAACCAC GAAAGGATCA
- TCTCACTCCG TTGTTCTTA AATGTCAGGA GTGTCTTTCT TTCCAGTTC TCTGTGGTG CTTTCTAGT

+ AAGATTATAA AACTAGCGAA ATGAGGGGGC CCCTAATCTT TCCAGTCTAG AAGAGTCAAA ATCCCCTCTC
- TTTAATAAT TTGATCGCTT TACTCCCCG GGGATTAGGA AGGTCGAGTC TTCTCAGTT TAGGGGAGGA

+ CCACCTGGCT CCAGCAGCCA AGTTGTGCGT TTACATCAGA GCACGTGTCA AATGCACCAT GCCTCGAACC
- GGTGGACCGA GGTCTGCGGT TCAACGACGC AATGTAGTCT CTGTCCAGC TTACGTGGTA CCGAGCTTGG

+ TCAGTACTAC TGTGAAACAA AGTACAACCT GGGGCCCGAA GACAGCTTCG AGTCGGAAGA AATCCACTT
- AGTCATGATG ACACTTTGT TCAATGTTAA CCCCAGGCTT CAGCTGAAAG TCAGCCTTCT TTAGGTGAAA

+ CTTCTGCTT TAAAAGTTGA CTGCTCCCAC TTGGGGCTAA ACTAGAAACT ATCAAAATCC CCGTCCCCG
- GAAGGACCGA ATTTCAACT GACGAGGGTG AACCCGATT TGATCTTTGA TAGTTTTAGG GCGAGGGGGC

+ GCCCCAGCC TGGCTGGCTG CTTATCACCG TTCCGGTTC GCGTCTGTG AGCGTTCAAT CATTACCCC
- CGGGGTCCG ACCGACCGAC GAATAGTGGC AAGGCGAAG CCGCAGAAC TCACAGTGA TAAAGTGGG

+ ACCCCCTCGA AAACCCGCCA AGTTATATA CAAACTAGCT CATCATCATC ATCATAAAA CGCTCATCAA
- TGGGGAGCT TTTGGCGGT TCAATAATAT GTTTGATCGA GTAGTAGTAG TAGTAGTTT GCGAGTAGIT

+ ATATTTTATA TCACATCATA CATTCCAGTC AATTCACCA TCAITTCAT TAACTGCAT TATATAATA
- TATAAGTAT AGTGTAGTAT GTAAGTTCAG TTAAGTTGTT AGTAAAGTTA ATTAGCGTA ATATAATAT

+ CCCCTTTTGA GTTTCACAAC ACCAAGTCTG ACGAATCCTC GCTTCAGCCT TAACTTTCA GTTGAATTA
- GGGGAAACT CAAAGTGTG TGGTTCAGCA TGCTAAGGAG CGAAGTCGGA AATGTAAAGT CAACCTAAT

+ CTCCGCACTG TCCACTACTA CTACATTTT TCAAGAAT
- GAGGCGTGAC AGGTGATGAT GATGAAAAA AGTTCCTA
    
```

Motifs Found

- AAGAA-motif
- ABRE
- AC-II
- AE-box
- ARE
- Box 4
- Box I
- Box-W1
- CAAT-box
- CATT-motif
- CGTCA-motif
- E2Fa
- ERE
- G-Box
- G-box
- GARE-motif
- GATA-motif
- GCN4_motif
- HD-Zip 1
- HD-Zip 2
- MBS
- P-box
- Skn-1_motif
- Sp1
- TATA-box
- TCA-element
- TCCACCT-motif
- TGACG-motif
- Unnamed_1
- Unnamed_13
- Unnamed_2
- Unnamed_3
- Unnamed_4
- W box
- WUN-motif
- as-2-box
- as1
- box II