



RICARDO DIAS CANESCHI

**ANÁLISE COMPARATIVA DE PERFIS DE EXPRESSÃO DE
C. ARABICA SUBMETIDOS A ESTRESSES ABIÓTICOS COM
DADOS DE BIBLIOTECAS GENÔMICAS DE NEW
C. CANEPHORA E *C. EUGENIODES***

**LAVRAS-MG
2022**

RICARDO DIAS CANESCHI

**ANÁLISE COMPARATIVA DE PERFIS DE EXPRESSÃO DE *C. ARABICA*
SUBMETIDOS A ESTRESSES ABIÓTICOS COM DADOS DE BIBLIOTECAS
GENÔMICAS DE NEW *C. CANEPHORA* E *C. EUGENIODES***

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.

Prof. Dr. Alan Carvalho Andrade
Orientador

**LAVRAS-MG
2022**

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Caneschi, Ricardo Dias.

Análise comparativa de perfis de expressão de *C. arabica* submetidos a estresses abióticos com dados de bibliotecas genômicas de New *C.canephora* e *C. eugeniodes* / Ricardo Dias Caneschi. - 2022.

95 p.

Orientador(a): Alan Carvalho Andrade.

Dissertação (mestrado acadêmico) - Universidade Federal de Lavras, 2022.

Bibliografia.

1. Sequenciamento de RNA. 2. Cafeeiro. 3. Resistência a fatores abióticos. I. Andrade, Alan Carvalho. II. Título.

RICARDO DIAS CANESCHI

**ANÁLISE COMPARATIVA DE PERFIS DE EXPRESSÃO DE *C. ARABICA*
SUBMETIDOS A ESTRESSES ABIÓTICOS COM DADOS DE BIBLIOTECAS
GENÔMICAS DE NEW *C. CANEPHORA* E *C. EUGENIODES***

**COMPARATIVE ANALYSIS OF EXPRESSION PROFILES OF *C. ARABICA*
SUBMITTED TO ABIOTIC STRESS WITH DATA FROM NEW *C. CANEPHORA*
AND *C. EUGENIODES* GENOMIC LIBRARIES**

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.

APROVADA em 29 de Abril de 2022.
Dr. Alan Carvalho Andrade, UFLA
Dr. Luciano Vilela Paiva, UFLA
Dr. Pierre Roger René Marracini, UPSUD

Prof. Dr. Alan Carvalho Andrade
Orientador

**LAVRAS-MG
2022**

AGRADECIMENTOS

A Universidade Federal de Lavras e ao Programa de Pós-Graduação em Biotecnologia Vegetal pela oportunidade de realização do mestrado.

Agradeço aos integrantes do laboratório LCBM, principalmente a Renan e Tatiane pela atenção e dedicação ao me auxiliar nos experimentos.

Ao Prof. Dr. Alan Carvalho Andrade pelos momentos de paciência, persistência e engenhosidade como orientador.

A minha mãe e meu padrasto, Guilhermina e Idemar, por me auxiliarem financeiramente e emocionalmente durante momentos diversos no meu mestrado e na vida. Ao meu namorado, companheiro e amigo, Jay, o qual me apoiou incondicionalmente ao final da escrita da minha tese. Aos meus amigos lavrenses, especialmente, a Kauanne, Iasminy e Isadora, que me prestaram conselhos valiosos de escrita e resiliência. À minha tia, Elisabeth, pelas infindáveis horas no telefone, com palavras de apoio e carinho.

E a todos que de alguma maneira me auxiliaram nessa jornada, direta e indiretamente!

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

RESUMO

A análise da expressão gênica tem sido amplamente utilizada como um método para estudar a complexa sinalização e as vias metabólicas subjacentes aos processos celulares e de desenvolvimento em organismos biológicos, incluindo plantas. Um número crescente de estudos de níveis de expressão de vários genes em plantas vem sendo realizado para se entender os mecanismos celulares e moleculares envolvidos no desenvolvimento e crescimento de plantas, bem como em respostas a estresses bióticos (infecção por patógenos) e abióticos (ambientais). Altas e baixas temperaturas, seca, alta salinidade, exposição luminosa inadequada são fatores abióticos que afetam o crescimento da planta e causam perdas nas lavouras. Dessa maneira, entender como as plantas percebem esses sinais ambientais e como esses são traduzidos pela maquinaria genética das mesmas com o propósito de minimizar os danos e se adaptarem é de suma importância. Muitos estudos investigam os efeitos de estresses singulares e combinados e possibilitaram a elucidação da interação molecular complexa que rege respostas ao estresse vegetal. As plantas ativam tanto respostas específicas e não específicas aos estresses causados por fatores relacionados a eventos climáticos adversos, permitindo uma maior eficiência na resposta e conservando recursos para crescimento. A especificidade do sinal é alcançada por meio da interação entre os hormônios de diferentes rotas metabólicas, como o ABA (ácido abscísico), AS (ácido salicílico) e o AJ (ácido jasmônico). Fatores de transcrição, de choque de calor, espécies reativas de oxigênio e pequenos RNAs também estão envolvidos na coordenação de tais respostas. Frisando, os fatores de transcrição são um dos principais responsáveis pela especificidade das respostas aos estresses. A manipulação desses provém uma das maiores oportunidades para conferir tolerância por meio da engenharia genética, uma vez que esses proporcionam uma série de eventos em downstream (XU et al., 2011). Esses fatores interagem com elementos das regiões promotoras dos genes relacionados aos estresses, promovendo uma alta da expressão de vários genes, permitindo então que haja a tolerância ao estresse. (AGARWAL; JHAa, 2010). Uma família de genes conhecidos por fatores ligantes de C-repetitivos (CBFs) ou fatores de elementos ligantes responsivos a desidratação (DREBs) são componentes chaves que implicam na adaptação à seca, salinidade e frio e regulam genes responsivos ao estresse os quais modulam adaptações fisiológicas aos estresses abióticos vegetais. A relação desses fatores compõe uma importante maneira de regulação gênica e rotas metabólicas dependentes e não dependentes ao ácido abscísico. A alta expressão dos genes DREB, ocasionam a regulação positiva de genes regulados pelo frio (CORE) e genes responsivos ao estresse osmótico, resultando em melhora da tolerância a estresses abióticos (CHEN et al., 2009). Dessa maneira, o intuito desse trabalho foi o de se avaliar o uso de diferentes genomas de referência (*C.arabica*, *C.canephora* e *C. eugenioides*) no mapeamento de reads obtidos por RNA-seq e observar os efeitos na comparação do transcriptoma de cafeeiros quando submetidas a diferentes estresses abióticos.

Palavras-chaves: Bioinformática. RNA-seq. mapeamento de genes. *Coffea* spp.

ABSTRACT

Gene expression analysis has been widely used as a method to study the complex signaling and metabolic pathways underlying cellular and developmental processes in biological organisms, including plants. An increasing number of studies of expression levels of various genes in plants are being carried out to understand the cellular and molecular mechanisms involved in plant development and growth, as well as in responses to biotic (pathogen infection) and abiotic (environmental) stresses. High and low temperatures, drought, high salinity, inadequate light exposure are abiotic factors that affect plant growth and cause crop losses. In this way, understanding how plants perceive these environmental signals and how these are translated by their genetic machinery in order to minimize damage and adapt is of paramount importance. Many studies investigate the effects of singular and combined stresses and have enabled the elucidation of the complex molecular interaction that governs responses to plant stress. Plants activate both specific and non-specific responses to stresses caused by factors related to adverse climatic events, allowing greater efficiency in the response and conserving resources for growth. Signal specificity is achieved through the interaction between hormones of different metabolic pathways, such as ABA (abscisic acid), SA (salicylic acid) and JA (jasmonic acid). Transcription factors, heat shock, reactive oxygen species and small RNAs are also involved in the coordination of such responses. Emphasizing, transcription factors are one of the main responsible for the specificity of responses to stress. The manipulation of these provides one of the greatest opportunities to confer tolerance through genetic engineering, as they provide a series of downstream events (XU et al., 2011). These factors interact with elements of the promoter regions of stress-related genes, promoting an increase in the expression of several genes, thus allowing stress tolerance. (AGARWAL; JHA, 2010). A family of genes known as C-repeat binding factors (CBFs) or dehydration-responsive binding element factors (DREBs) are key components that implicate adaptation to drought, salinity and cold and regulate stress-responsive genes which modulate physiological adaptations to plant abiotic stresses. The relationship of these factors makes up an important way of gene regulation and metabolic pathways dependent and not dependent on abscisic acid. The high expression of DREB genes causes the upregulation of genes regulated by cold (CORE) and genes responsive to osmotic stress, resulting in improved tolerance to abiotic stresses (CHEN et al., 2009). Thus, the aim of this work was to evaluate the use of different reference genomes (*C.arabica*, *C.canephora* and *C. eugenioides*) in the mapping of reads obtained by RNA-seq and to observe the effects in the comparison of the transcriptome of coffee plants when subjected to different abiotic stresses.

Keywords: Bioinformatics. RNA-seq. gene mapping. Coffee spp.

SUMÁRIO

1	INTRODUÇÃO	10
2	REFERENCIAL TEÓRICO	11
2.1	Características gerais da produção e origem cafeeira	11
2.2	Estresses ambientais e mecanismos de defesa das plantas	13
2.3	Fatores de transcrição	15
2.4	Vias de transdução de sinal envolvidas na resposta ao estresse da planta	16
2.5	DREB/CBF (fator de ligação de elemento de resposta à desidratação/ fator de ligação de repetição C)	18
2.5.1	Genes DREBs em <i>Coffea canephora</i>	20
2.6	Fisiologia do café e os principais estresses abióticos que afetam as plantas	22
2.6.1	Ácido abscísico – ABA, hormônio do estresse	23
2.6.2	Baixa temperatura	25
2.6.3	Alta temperatura	27
2.6.4	Estresse foto-oxidativo	29
2.6.5	Seca	32
2.6.5.1	Respostas a longo prazo	34
2.6.5.2	Resposta de curto prazo	35
2.7	Transcriptoma	36
2.8	Recursos para análise de genoma usando dados de RNA-seq	37
2.9	Análise de expressão de dados de RNA - seq usando ferramentas de Bioinformática e Estatística	39
3	OBJETIVOS	41
3.1	Objetivo Geral	41
3.2	Objetivos específicos	41
4	MATERIAIS E MÉTODOS	41
4.1	Local do experimento e material para análise	41
4.2	Pipeline de análises de Bioinformática	41
4.3	Delineamento experimental e estatístico	42
5	RESULTADOS E DISCUSSÕES	44
5.1	Padrões de genes diferencialmente expressos resultantes das 3 análises realizadas com diferentes genomas de referência (<i>C. arabica</i> , <i>C. canephora</i> e <i>C. eugenioides</i>), responsivos à exposição ao estresse da aplicação exógena de ABA	45

5.2	Padrões de genes diferencialmente expressos resultantes das 3 análises realizadas com diferentes genomas de referência (<i>C. arabica</i>, <i>C. canephora</i> e <i>C. eugenioides</i>), responsivos ao estresse causado pelo calor.....	56
5.3	Padrões de genes diferencialmente expressos resultantes das 3 análises realizadas com diferentes genomas de referência (<i>C. arabica</i>, <i>C. canephora</i> e <i>C. eugenioides</i>), responsivos ao estresse causado pelo frio.....	67
5.4	Padrões de genes diferencialmente expressos resultantes das 3 análises realizadas com diferentes genomas de referência (<i>C. arabica</i>, <i>C. canephora</i> e <i>C. eugenioides</i>), responsivos ao estresse luminoso.....	76
5.5	Padrões de genes diferencialmente expressos resultantes das 3 análises realizadas com diferentes genomas de referência (<i>C. arabica</i>, <i>C. canephora</i> e <i>C. eugenioides</i>), responsivos ao estresse causado pela restrição hídrica.....	89
5.6	Análise da similaridade das proteínas responsivas identificadas em cada pipeline (ref. <i>C.arabica</i>, <i>C.canephora</i> e <i>C.eugenioides</i>), nos diferentes tratamentos estudados.....	101
6	CONCLUSÃO	106
	REFERÊNCIAS	108

1 INTRODUÇÃO

Mudanças climáticas, causadoras de diferentes estresses nas lavouras de café e consequente perdas, têm trazido à luz a necessidade do uso de ferramentas de manipulação genética e biotecnologia para a indução e compreensão da resistência da cultura. Essas técnicas consideram os mecanismos de defesa já existentes nas espécies estudadas e tentam realçá-los em novas variedades produzidas. Entretanto, o que irá determinar o nível de resistência e elucidar a causa dessa tolerância, é uma análise cuidadosa do perfil de expressão de cada espécie quando exposta a estresses luminosos, hídricos, de temperatura e ao hormônio ABA.

A análise da expressão gênica tem sido amplamente utilizada como um método para estudar a complexa sinalização e as vias metabólicas subjacentes aos processos celulares e de desenvolvimento em organismos biológicos, incluindo plantas. Um número crescente de estudos de níveis de expressão de vários genes em plantas foi realizado para entender os mecanismos celulares e moleculares envolvidos no desenvolvimento e crescimento de plantas, bem como em respostas a estresses bióticos (infecção por patógenos) e abióticos (ambientais).

Altas e baixas temperaturas, seca, alta salinidade, exposição luminosa inadequada são fatores abióticos estressantes que afetam o crescimento da planta e causam perdas nas lavouras. Dessa maneira, entender como as plantas percebem esses sinais ambientais e como esses são traduzidos pela maquinaria genética das mesmas com o propósito de minimizar os danos e se adaptarem é de suma importância para criar cruzamentos favoráveis que favoreçam a lavoura.

Muitos estudos investigam os efeitos de estresses singulares e combinados e possibilitaram a elucidação da interação molecular complexa que rege respostas ao estresse vegetal. As plantas ativam tanto respostas específicas e não específicas aos estresses causados por fatores relacionados a eventos climáticos adversos, permitindo uma maior eficiência na resposta e conservando recursos para crescimento. A especificidade do sinal é alcançada por meio da interação entre os hormônios de diferentes rotas metabólicas, como o ABA (ácido absísico), AS (ácido salicílico) e o AJ (ácido jasmônico). Fatores de transcrição, de choque de calor, espécies reativas de oxigênio e pequenos RNAs também estão envolvidos na coordenação de tais respostas.

Frisando, os fatores de transcrição são um dos principais responsáveis pela especificidade das respostas aos estresses. A manipulação desses provém uma das maiores oportunidades para conferir tolerância transgênicamente, uma vez que esses proporcionam uma série de eventos em downstream (XU et al., 2011). Esses fatores interagem com elementos das regiões promotoras dos genes relacionados aos estresses, promovendo uma alta da expressão

de vários genes repressores, permitindo então que haja a tolerância ao estresse. (AGARWAL; JHAa, 2010).

Uma família de genes conhecidos por fatores ligantes de C-repetitivos (CBFs) ou fatores de elementos ligantes responsivos a desidratação (DREBs) são componentes chaves que implicam na adaptação à seca, salinidade e frio e regulam genes responsivos ao estresse os quais modulam adaptações fisiológicas aos estresses abióticos vegetais. A relação desses fatores compõe uma importante regulação gênica e de rotas metabólicas dependentes e não dependentes ao ácido abscísico. A alta expressão dos genes DREB, ocasionam a regulação positiva de genes regulados pelo frio (CORE) e genes responsivos ao estresse osmótico, resultando em melhora da tolerância a estresses abióticos (CHEN et al., 2009).

Dessa maneira, o intuito desse trabalho foi utilizar dados de RNA-seq, obtidos a partir de folhas de plântulas de *C. arabica* em condições controle e submetidas a diferentes estresses abióticos, e realizar uma análise comparativa dos resultados obtidos, após o mapeamento e análises estatísticas, utilizando-se diferentes genomas como referência (*C. arabica*, *C. eugenioides* e uma nova montagem de *C. canephora*)

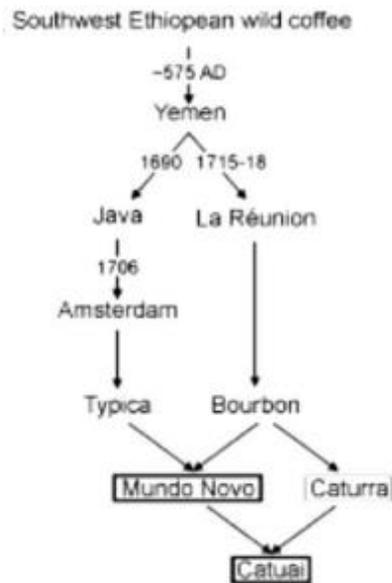
2 REFERENCIAL TEÓRICO

2.1 Características gerais da produção e origem cafeeira

O café pertence à ordem Gentianales, família Rubiaceae e subfamília Ixoreideae, e está dividido em dois gêneros, *Coffea* L. e *Coffea* Hook f., que compreendem mais de 100 espécies que diferem em morfologia, tamanho e adaptação ecológica (DAVIS et al., 2011). O gênero *Coffea* L., pertencente ao grupo das dicotiledôneas, caracteriza-se por folhas persistentes, porte arbustivo, caules e flores hermafroditas (FAZUOLI et al., 1986).

O café arábica é nativo das florestas tropicais da Etiópia, Quênia e Sudão, em altitudes de 1500-2800 m, entre as latitudes de 4°N e 9°N (ANTHONY et al., 1987). Historicamente, a herança genética das cultivares modernas de *Coffea arabica* surgem de duas populações principais, conhecidas como *Typica* e *Bourbon*, que se difundiram mundialmente no século XVIII (ANTHONY et al., 2002). É o caso das cultivares mais comerciais de *C. arabica*, *Mundo Novo*, *Catuai* e *Caturra*. A cultivar *caturra* é um mutante anão do grupo *Bourbon*. Já *Mundo Novo* é um híbrido entre *Bourbon* e *Typica*, enquanto a cultivar *Catuai* que derivou de um cruzamento entre *Mundo Novo* e *Caturra*. Cada cultivar apresenta arquitetura de planta e propriedades fisiológicas distintas (Figura 1).

Figura 1- Origem das cultivares comerciais de *C. arabica*.



Fonte: Vidal et al., (2010)

C. arabica é uma espécie autógama e alotetraploide ($2n = 4X = 44$ cromossomos) (CLARINDO; CARVALHO, 2011). Originou-se nas terras altas centrais da Etiópia a cerca de um milhão de anos como resultado de um cruzamento natural entre as espécies diplóides *C. canephora* e *C. eugenoides* (LASHERMES et al., 1999).

Como consequência de sua história evolutiva e autogamia, a diversidade genética de *C. arabica* é baixa e de irrestrito acesso aos seus recursos genéticos (e também outros *Coffea spp.*), sendo crucial para o desenvolvimento avançado de variedades (VAN DER VOSSEN; BERTRAND, 2015).

Já a espécie *C. canephora* é alógama e diploide ($2n = 2X = 22$ cromossomos), com aproximadamente 568,6 Mb em seu genoma (DENOEUDE et al., 2014). É nativa das terras baixas da floresta da bacia do rio Congo, onde se estende até o Lago Vitória, Uganda. Apresenta alta genética variabilidade genética e grande adaptação às mais variadas condições climáticas, o que possibilitou reunir esta espécie em dois grupos distintos, denominados congolezes e guineenses, de acordo com sua origem geográfica (MONTAGNON; CUBRY; LEROY, 2012). O grupo congolês tem vários subgrupos dos quais dois são de grande importância: o SG1 subgrupo (tolerante à seca), típico de uma região de clima árido, localizado em uma área mais continental do continente, e o subgrupo SG2 (sensível à seca), característico de uma região com altos índices pluviométricos e bem distribuídos ao longo do ano, o que lhe confere suscetibilidade à seca (FAZUOLI, 2007).

A produção comercial de café é baseada principalmente em duas espécies: *C. arabica* e

C. canephora, que respondem por cerca de 65% e 35% da produção mundial, respectivamente. Em 2016, o maior produtor foi o Brasil, seguido pelo Vietnã e Colômbia, com aproximadamente 55.000; 25.500 e 14,5 mil sacas de 60kg, respectivamente produzidas. Entretanto, as mudanças climáticas vêm impactando de maneira avassaladora os cafezais, uma vez que essas têm sido drásticas e impossibilitam a migração e adaptação através de novas mutações. (BUNN ET AL., 2015; DAVIS ET AL., 2012, 2019; MOAT ET AL., 2017, 2019; OVALLE-RIVERA ET AL., 2015) As estratégias adaptativas para mitigar esses efeitos dependem em grande parte de como a espécie responde à variabilidade climática e da disponibilidade de recursos genéticos nas populações selvagens que podem ser aproveitados para aumentar a tolerância à seca e ao calor.

2.2 Estresses ambientais e mecanismos de defesa das plantas

O estresse ambiental é exercido na planta por fatores bióticos ou abióticos. Os fatores bióticos envolvem a interação entre organismos vivos, como patógenos, algas ou ervas daninhas, enquanto estresse é causado por altas ou baixas temperaturas, excesso ou falta de água, alta concentração de sal e também por componentes químicos (QURESHI et al., 2007). Esses fatores ambientais limitam severamente o crescimento e reprodução das plantas e, conseqüentemente, a produtividade agrícola no caso de plantas de cultivo intenso.

Dentre todos os tipos de estresses abióticos, a desidratação e os extremos de temperatura são os que mais afetam o desenvolvimento das plantas. As plantas respondem ao déficit hídrico, bem como às altas e baixas temperaturas, com diversas respostas fisiológicas e moleculares (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2000). Um determinado estresse ambiental que afeta uma planta pode não ser estressante para outra, pois algumas plantas desenvolveram características adaptativas contra esses estresses por meio da ativação de mecanismos de defesa. A resposta da planta ao estresse é a função mais crucial da célula vegetal e ocorre através da alteração no modelo de expressão gênica (QURESHI et al., 2007), cujos produtos (proteínas) podem estar envolvidos em diversas funções adaptativas (THOMASHOW, 1999). Os genes induzidos durante estresses abióticos codificam proteínas que atuam no controle da expressão gênica e transdução de sinal, assim como as proteínas envolvidas, por exemplo, na proteção e desintoxicação das células (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2000).

Um determinado estresse ambiental que afeta uma planta pode não ser estressante para outra, pois algumas plantas desenvolveram características adaptativas contra esses estresses por meio da ativação de mecanismos de defesa. A resposta da planta ao estresse é a função mais

crucial da célula vegetal e ocorre através da alteração no modelo de expressão gênica (QURESHI et al., 2007), cujos produtos (proteínas) podem estar envolvidos em diversas funções adaptativas (THOMASHOW, 1999). Os genes induzidos durante estresses abióticos codificam proteínas que atuam no controle da expressão gênica e transdução de sinal, assim como as proteínas envolvidas, por exemplo, na proteção e desintoxicação das células (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2000). 15 Múltiplas vias de sinalização regulam as respostas ao estresse das plantas (KNIGHT; KNIGHT, 2001) e há uma sobreposição entre os padrões de expressão gênica que são induzidos em resposta a diferentes estressores (CHEN et al., 2002). As respostas moleculares a estresses abióticos incluem percepção de estresse, transdução de sinal para componentes celulares, expressão gênica e, finalmente, alterações metabólicas que conferem tolerância ao estresse. Os genes assim induzidos pelo estresse funcionam não apenas na proteção das células do estresse pela produção de importantes proteínas metabólicas, mas também na regulação dos genes a jusante para a transdução de sinal (AGARWAL et al., 2006). As respostas das plantas a estresses abióticos são altamente complexas e envolvem a expressão de um grande número de genes que codificam proteínas e enzimas relacionadas ao estresse que atuam em vias biossintéticas de osmoprotetores e outros metabólitos relacionados ao estresse (VINOCUR; ALTMAN, 2005). As proteínas funcionais são caracterizadas por proteger a célula contra a desidratação. Nesse grupo estão as aquaporinas (movimento da água através da membrana plasmática), as enzimas osmoprotetoras (relacionadas ao acúmulo de solutos no citosol, com a finalidade de promover a manutenção do turgor celular), as proteínas protetoras de macromoléculas e membranas celulares, como Embiogênese Tardia de Proteínas abundantes (LEA), chaperonas moleculares como HSP 70 BiP (Proteína Heat Shock Binding Protein de 70 kDa) e proteínas anticongelantes, enzimas desintoxicantes (neutralizando radicais livres do estresse oxidativo) e proteases (CUSHMAN; BOHNERT, 2000). As proteínas reguladoras são caracterizadas pela regulação da transdução de sinal e expressão gênica. No grupo das proteínas reguladoras estão presentes proteínas quinases e fatores de transcrição (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 1997). Eles incluem vários fatores de transcrição (TFs), como oncogene de mielocitomatose (MYC), oncogene de mieloblastose (MYB), zíper de leucina básico (bZIP), NAM, ATAF e CUC (NAC), ligação de elemento responsivo à desidratação (DREB), etc. sugerindo o papel de vários mecanismos reguladores da transcrição nas vias de transdução do sinal de estresse (AGARWAL et al., 2006).

2.3 Fatores de transcrição

Os fatores de transcrição (TFs) se ligam aos elementos de ação cis presentes na região promotora de vários genes que são regulados por vias de sinalização que desencadeiam sua ativação. Os elementos regulatórios participam do controle de vários processos biológicos, incluindo respostas de estresse abiótico, respostas hormonais e processos de desenvolvimento (YAMAGUCHI-SHINOZAKI; SHINOZAKI, 2006).

Os fatores de transcrição se enquadram na categoria de genes iniciais e são induzidos em poucos minutos de estresse (HUANG et al., 2012). Eles são de importância fundamental na geração de especificidade nas respostas. Sua manipulação oferece uma das maiores oportunidades para conferir múltiplas tolerâncias aos estresses de forma transgênica, pois controlam uma ampla gama de eventos a jusante (XU et al., 2011).

Eles interagem com elementos cis nas regiões promotoras de vários genes relacionados ao estresse para regular a expressão de muitos genes a jusante, conferindo assim tolerância ao estresse (AGARWAL; JHA, 2010). Uma lista de TFs que podem ser cruciais no controle da resposta a estresses bióticos e abióticos é dada em Atkinson e Urwin (2012).

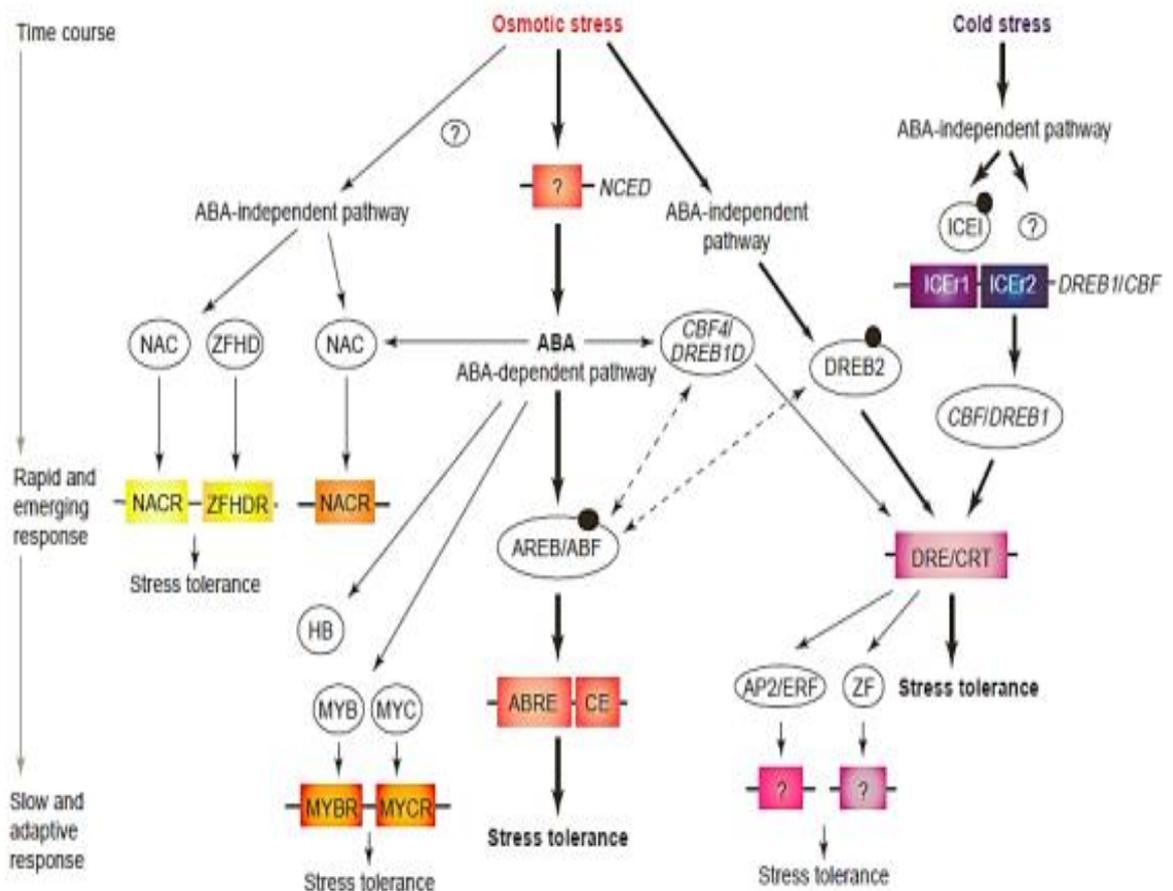
Os TFs não se ligam indiscriminadamente a todos os genes com um elemento de resposta correspondente, pois pode haver milhares de tais sítios em todo o genoma. Em vez disso, a especificidade pode ser gerada por TFs formando homo ou heterodímeros, que então se ligam a um par de elementos de resposta em uma distância apropriada um do outro, ou pela cooperação de proteínas em ponte ou suporte que direcionam os TFs para os elementos de resposta corretos (VAAHTERA; BROSCHE, 2011).

Os genomas de plantas atribuem aproximadamente 7% de sua sequência de codificação aos TFs, o que comprova a complexidade da regulação transcricional (AGARWAL et al., 2006b). Dados de transcriptoma em *Arabidopsis* e em várias outras plantas sugerem que existem vários caminhos que respondem independentemente a estresses ambientais (tanto de maneira dependente quanto independente de ABA), sugerindo que a tolerância ou suscetibilidade ao estresse é controlada no nível transcricional por uma rede reguladora de genes extremamente intrincada (AGARWAL; JHA, 2010).

2.4 Vias de transdução de sinal envolvidas na resposta ao estresse da planta

Quatro vias de transdução de sinal envolvidas na resposta da planta ao estresse foram descritas: vias I e II ABA dependentes e vias III e IV não dependentes de ABA (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2000). Os sistemas de sinalização dependentes de ABA foram descritos como vias que mediam a adaptação ao estresse pela ativação de pelo menos dois TFs diferentes que podem ser identificados: (I) o regulador AREB/ABF (proteína de ligação ao elemento responsivo a ABA/fator de ligação a ABA) ; e (II) o regulador MYC/MYB. Por outro lado, TFs independentes de ABA são: (III) os reguladores CBF/DREB (fator de ligação a frio/ligação de elemento responsivo à desidratação); e (IV) o regulador NAC e ZF-HD (homeodomínio dedo de zinco) (SAIBO et al., 2009). Muitos outros sistemas de regulação transcricional estão envolvidos na expressão de genes responsivos ao estresse. Em paralelo, essas quatro vias principais se relacionam e convergem para a ativação de genes envolvidos na resposta ao estresse. A Figura 2 mostra um modelo de vias de transdução de sinal de estresses bióticos e abióticos da percepção à expressão gênica. Há conversas cruzadas entre essas vias de sinalização, implicando fatores de transcrição comuns para estresses bióticos e abióticos (YAMAGUCHI-SHINOZAKI; SHINOZAKI, 2005; CIARMIELLO et al., 2011).

Figura 2 - Uma representação esquemática da percepção do sinal de estresse e expressão gênica via vias dependentes e independentes de ABA em nível celular em plantas.



Legenda: Os fatores de transcrição que controlam a expressão gênica induzida por estresse são representados como elipses. Os elementos de cisacção envolvidos na transcrição responsiva ao estresse são representados como caixas coloridas. Círculos pequenos, pretos e preenchidos revelam modificação de fatores de transcrição em resposta a sinais de estresse para sua ativação, como a fosforilação. A cascata reguladora da expressão gênica responsiva ao estresse é mostrada de cima para baixo. As respostas precoces e de emergência da expressão gênica são mostradas na parte superior, e as respostas tardias e adaptativas na parte inferior. Setas pretas grossas indicam as principais vias de sinalização; essas vias regulam muitos genes de maneira downstream. As setas quebradas indicam interações proteína-proteína. Abreviaturas: ABA, ácido abscísico; AREB, proteínas de ligação ABRE; ABRE, elemento responsivo ao ABA; CBF, fator de ligação de repetição C; DRE/CRT, elemento responsivo à desidratação/repetição C; DREB, proteína de ligação a DRE; ICE, Indutor de Expressão de CBF; MYBR, sitio de reconhecimento por MYB; MYCR, sitio de reconhecimento por MYC; NACR, local de reconhecimento de NAC; ZFHDR, sitio de reconhecimento de homeodomínio de dedo de zinco. Fonte: Yamaguchi-Shinozaki e Shinozaki (2005)

2.5 DREB/CBF (fator de ligação de elemento de resposta à desidratação/fator de ligação de repetição C)

Os DREB são importantes fatores de transcrição de plantas (FT) AP2 (APETALA2)/ERF (fator responsivo ao etileno) que induzem um conjunto de genes relacionados ao estresse abiótico. Os 145 membros desta grande família AP2/ERF de FTs em *Arabidopsis* foram classificados em quatro subfamílias : DREB/CBF, ERF, AP2, RAV (SAKUMA et al., 2002; LATA; PRASAD, 2011), com base nas semelhanças de suas sequências e nos números de domínios de ligação AP2/ERF (NAKANO et al., 2006).

O primeiro trabalho sobre esta família de fatores de transcrição foi realizado por YamaguchiShinozaki et al. (1993; 1994), através da identificação de cDNAs responsivos à dessecação (rd) de *A. thaliana*, que se expressaram não só em resposta ao déficit hídrico, mas também ao estresse salino, baixa temperatura e ácido abscísico.

Os membros da subfamília AP2/ERF contêm um domínio altamente conservado que consiste em um sequência de 60 a 70 aminoácidos. As proteínas AP2/ERF desempenham papéis importantes na regulação de genes de transcrição em uma variedade de processos biológicos relacionados ao crescimento e desenvolvimento, bem como várias respostas a estímulos ambientais que regulam a expressão de genes vegetais responsivos a estresses bióticos e abióticos (AGARWAL et al., 2006b). Os FTs AP2/ERF têm sido associados há muito tempo com sinalização de estresse e descobertas recentes sugerem que eles podem ser excelentes alvos para melhorar a tolerância de amplo espectro em culturas por meio de engenharia genética (XU et al., 2011).

Os fatores de transcrição DREB são divididos em duas classes, DREB1 e DREB2 (CHINI et al., 2004). Três genes DREB1/CBF estão envolvidos em respostas e aclimatação ao estresse por frio (NOVILLO et al., 2007; FANG et al., 2015), enquanto os genes DREB2 são induzidos por desidratação, alta salinidade e calor (SAKUMA et al., 2006; MIZOI et al., 2012). A subclasse DREB1/CBF tem as seguintes variações: DREB1B/CBF1, DREB1C/CBF2 e DREB1A/CBF3 (STOCKINGER et al., 1997; LIU et al., 1998; MEDINA et al., 1999). O DREB desempenha papéis importantes na regulação da expressão de genes (rd29A, kin1 e eerd10) em resposta a uma variedade de estresses abióticos e bióticos (YAMAGUCHI-SHINOZAKI; SHINOZAKI, 20056; AGARWAL et al., 2006).

Foi estabelecido que eles são os principais fatores de transcrição envolvidos na atividade abiótica das plantas em resposta aos estresses regulando a expressão gênica via

DRE/CRT de ação cis (elemento responsivo à desidratação/repetição C) (YAMAGUCHI-SHINOZAKI; SHINOZAKI, 2005). Esses FTs foram isolados de *Arabidopsis* e designados como DREB1A, DREB1B, DREB1C, DREB2A e DREB2B.

Uma vez que DREB1A, DREB1B, DREB1C são induzidos pelo estresse por frio, e DREB2A e DREB2B por condições de estresse hídrico e alta concentração de sal (DUBOUZET et al., 2003). Os genes DREB têm sido estudados e caracterizados em diversas espécies vegetais, como *Arabidopsis*, trigo, centeio, tomate, milho, arroz, cevada, trigo sarraceno, uva e colza e sob diversas condições ambientais adversas. O DRE/CRT foi identificado como um instrumento essencial de ação do elemento cis no promotor *Arabidopsis thaliana* RD29A para expressão em resposta à desidratação e frio, independentemente da presença de ABA. O gene RD29A (Responsivo à seca) foi muito caracterizado e consiste em uma sequência de nove pares de bases - TACCGACAT – que regula sua indução em condições de déficit hídrico, baixas temperaturas e estresse salino (YAMAGUCHI-SHINOZAKI; SHINOZAKI, 1994; SAKUMA et al., 2006; FANG et al., 2015; ZANDKARIMI et al., 2015).

Os genes DREB desempenham um papel importante nas vias de tolerância ao estresse independente de ABA, induzindo a expressão de vários genes responsivos ao estresse em plantas. Os primeiros cDNAs isolados codificando proteínas de ligação DRE, CBF1 (fator de ligação CRT1), DREB1A e DREB2A foram isolados usando uma triagem de híbrido de levedura (STOCKINGER et al., 1997; LIU et al., 1998) de *Arabidopsis*. Desde então, vários genes DREB foram isolados de várias plantas (LATA; PRASAD, 2011; FANG et al., 2015; ZANDKARIMI et al., 2015).

Uma região de 125 pb do promotor DREB1C demonstrou ser suficiente para conduzir a transcrição induzida pelo frio, e dois segmentos dentro do promotor, designados por ICer1 e ICer2 (indutor da região de expressão de CBF 1 e 2), contribuem para as respostas ao frio (ZARKA et al., 2003). ICE1 e MYB15 também foram identificados como fatores de transcrição de maneira Upstream da indução de genes DREB em resposta ao estresse pelo frio (AGARWAL et al., 2006). No entanto, pouco se sabe sobre a regulação positiva dos genes DREB2 (CHEN et al., 2012).

Essas observações sugerem que as proteínas DREB são importantes FTs na regulação abiótica de genes relacionados ao estresse e desempenham um papel crucial na transmissão de tolerância ao estresse às plantas. O DREB1 e os reguladores DREB2 podem assim ser usados para melhorar a tolerância de vários tipos de

plantas cultivadas importantes para estresses de seca, alta salinidade e congelamento através da transferência de genes (LATA; PRASAD, 2011).

Uma diversidade de genes DREBs foi funcionalmente caracterizada em plantas modelo por transformação homóloga. Em *Arabidopsis* e arroz, por exemplo, membros de subfamílias dos grupos DREB1 e DREB2 têm mostrado resultados promissores na indução de tolerância ao frio, sal e seca (SAKUMA et al., 2002; DUBOUZET et al., 2003; ITO et al., 2006; LIM et al., 2007; TAKASAKI et al., 2010; KIM et al., 2011; CHEN et al., 2009; CHEN et al., 2012).

2.5.1 Genes DREBs em *Coffea canephora*

Para a identificação dos genes candidatos envolvidos na tolerância à seca em cafeeiros diferentes estratégias foram utilizadas, entre elas a identificação desses genes candidatos através do Projeto de sequenciamento do genoma do café, que já desvendou mais de 200 mil sequências de cDNA e identificaram cerca de 30 mil genes do cafeeiro (MONDEGO et al., 2011). A maioria das sequências obtidas foram depositadas no banco de dados de informações do laboratório internacional de biotecnologia do Centro Nacional de Informações sobre Biotecnologia (NCBI). Eles disponibilizaram as sequências dos genes que foram expressos - EST (Expressed Sequence Tags) - removendo os tecidos do café em seus estágios de desenvolvimento ou no momento em que esses tecidos responderam a estresses bióticos ou abióticos. É possível remontar a molécula de RNA, ou seja, a cópia do DNA (cDNA) da planta que se expressa no momento dos estresses (VIEIRA et al., 2006).

Entre os genes candidatos que apresentaram expressão diferencial em folhas de *C. canephora* conilon tolerante (14 e 120) e sensível (clone 22) a clones de seca cultivados em casa de vegetação com irrigação (I) ou sem (NI), foram identificados vários genes, principalmente o gene CcDREB1D (MARRACCINI et al., 2012). A atividade transcricional de formas alélicas e homólogas do promotor CcDREB1D de *C. canephora* foi estudado e isolado do clone 14 (tolerante) e clone 22 (sensível) por transformação genética de *C. arabica* (ALVES et al., 2015).

O gene CcDREB1D e outros membros da subfamília DREB de *C. canephora* foram identificados e divididos em quatro subgrupos: I, II, III e IV (Figura 3) por Alves (2015). Esta identificação foi possível devido ao sequenciamento completo do genoma de *C. canephora* (DENOEUDE et al., 2014), cujas informações foram disponibilizadas em um sistema integrador de informações genômicas, o Hub do Banco de Dados do Genoma do Café (DEREEPER et al.,

2.6 Fisiologia do café e os principais estresses abióticos que afetam as plantas

Sob condições de campo, o desempenho da planta em termos de crescimento, desenvolvimento, acumulação de biomassa e rendimento depende da capacidade de aclimação às mudanças e estresses ambientais, exercendo mecanismos de tolerância específicos que envolvem uma complexa rede de fatores bioquímicos e processos moleculares (WANG et al., 2003). No geral, a seca e as temperaturas desfavoráveis são as grandes limitações climáticas para a produção de café. Espera-se que essas limitações se tornem cada vez mais importante em várias regiões cafeeiras devido às reconhecidas mudanças na clima, e também porque o cultivo do café se espalhou para terras marginais, onde a escassez de água e as temperaturas desfavoráveis constituem grandes restrições à produtividade do café (DAMATTA; RAMALHO, 2006)

Efeitos primários de estresses, como frio e seca, muitas vezes estão relacionados, causando danos e estresses secundários, como por exemplo, o estresse osmótico e oxidativo, resultando em ruptura da homeostase iônica e danos a proteínas e membranas. Esses desequilíbrios vão desencadear sinalização downstream e controle de transcrição, que ativam mecanismos anticongelantes e funções relacionadas as proteínas chaperonas (por exemplo, quitinases, proteínas de choque térmico, proteínas de embriogênese tardia abundante e genes relacionados ao frio), desintoxicação [por exemplo, superóxido dismutase (SOD), ascorbato peroxidase (APX)], osmoproteção (por exemplo, prolina, glicina betaína, açúcar polióis), etc., que restabelecem a homeostase celular (Wang et al., 2003).

Em relação ao cafeeiro arábica, a peroxidação dos lipídios da membrana desempenha um papel decisivo na alta sensibilidade ao frio das raízes, provavelmente ligada a uma limitação no funcionamento dos sistemas antioxidantes provocados por uma diminuição da atividade respiratória. Este último traz uma baixa disponibilidade de poder redutor que, por sua vez, leva à falha das enzimas antioxidantes em proteger membranas lipídicas, causando maior dano ao tecido radicular e rigidez da membrana (QUEIROZ et al., 1998).

Um conjunto de estudos com vasos de plantas de diversos genótipos de café, que foram submetidos a baixas temperaturas positivas (CAMPOS et al., 2003; RAMALHO et al., 2003) mostraram que Icatu (uma híbrido de *C. Arabica* x *C. Canephora*) apresentou melhor capacidade de aclimação ao frio do que os demais genótipos estudados, devido à pequena perda de seletividade da membrana e degradação.

Isso estaria relacionado com a manutenção (ou aumento) na saturação de ácidos graxos de lipídios de membrana e o reforço do sistema antioxidante (observado também na cv.

Catuaí). Tais mudanças permitem que os cafeeiros mantenham maior atividade fotossintética devido a uma estrutura de membrana funcional (ainda mais rígida) menos sensível a EROs. Tais alterações nos lipídios da membrana também parecem desempenhar um papel importante na aclimação do cafeeiro a outras condições que promovem o estresse oxidativo, como a exposição à alta luminosidade (RAMALHO et al., 1998), tornando-se importante em pouco tempo (8 h) (GASCÓ et al., 2004).

Espécies ou cultivares mais tolerantes à seca geralmente diferem morfológicamente e/ou fisiologicamente, com mecanismos que permitem maior produção sob restrição de água. Compreender tais mecanismos em genótipos naturalmente adaptados à seca pode ajudar a melhorar seu desempenho agrônomo. No café, algumas características fisiológicas têm demonstrado potencial de contribuição para a produtividade em condições de seca (DAMATTA, 2004b), mas o desenvolvimento de um método de melhoramento eficiente para tolerância à seca ainda é um objetivo de longa data. Uma importante componente da adaptação diferencial à seca entre genótipos de café arábica e robusta parece ser comportamental e pode ser governado por taxas de uso de água e/ou eficiência de extração de água do solo (PINHEIRO et al., 2005)

Estudos fisiológicos sobre a relação entre tolerância à seca e expressão gênica levaram à identificação de mais de 80 Genes candidatos tanto em *C. CANEPHORA* (MARRACCINI ET AL., 2011, 2012; VIEIRA ET AL., 2013 ☺) e *C. arabica* (FREIRE et al., 2013; MOFATTO ET AL., 2016). Genes adicionais que supostamente desempenham um papel fundamental nas respostas das plantas ao estresse abiótico também foram identificados (MARRACCINI, 2020), como aqueles envolvidos na via biossintética do ABA (COSTA, 2014; COTTA, 2017; SIMKIN ET AL., 2008), proteção e desintoxicação celular (HINNIGER ET AL., 2006; SANTOS & MAZZAFERA, 2012; THIOUNE ET AL., 2017, 2020) e biossíntese de aquaporinas (MINIUSSI ET AL., 2015; SANTOS & MAZZAFERA, 2013). Outros genes estão envolvidos nas vias biossintéticas de carotenóides/fenilpropanóides (LEPELLEY ET AL., 2012; SIMKIN ET AL., 2010), cafeína (DENOEUDE ET AL., 2014) e açúcar (GEROMEL ET AL., 2006), ou codificam fatores de transcrição (ALVES ET AL., 2017, 2018; THIOUNE ET AL., 2017, 2020; TORRES ET AL., 2019).

2.6.1 Ácido abscísico – ABA, hormônio do estresse

O ABA é um importante hormônio vegetal que desempenha um papel regulador em muitos processos fisiológicos em plantas, como maturação embrionária, germinação e

desenvolvimento de sementes, dormência de sementes e gemas, crescimento de raízes, amadurecimento de frutos, regulação da abertura estomática e ativação de genes responsivos ao estresse. (AGARWAL; JHA, 2010). Níveis elevados de ABA são desencadeados por uma variedade de estresses ambientais, como seca, salinidade, frio, dessecação, calor e ferimentos. Além disso, também está provado que o ABA é um importante sinal fisiológico que induz respostas à seca e à alta salinidade (FAROOQ et al., 2009).

O ABA é sintetizado em resposta à redução do potencial hídrico (CHRISTMANN et al., 2007) e a biossíntese é dividida entre os plastídios e o citosol. A clivagem oxidativa do carotenóide precursor 99-cis-neoxantina em xantoxina, catalisada pela enzima plastidial 99-cis-epoxicarotenoide dioxigenase (NCED), é a etapa comprometida para a biossíntese do ABA (NAMBARA; MARION-POLL, 2005). As respostas ao estresse abiótico são amplamente controladas pelo hormônio ABA, enquanto a defesa contra diferentes agressores bióticos é especificada pelo antagonismo entre as vias de sinalização do ácido salicílico (SA) e do ácido jasmônico (JA)/etileno. No entanto, descobertas recentes sugerem que o ABA atua sinergicamente e antagonicamente com a sinalização de estresse biótico, criando uma rede complexa de vias de interação com cross-talk em diferentes níveis (FUJITA et al., 2006; ASSELBERGH et al., 2008b).

Plantas sob estresse hídrico geralmente apresentam um maior acúmulo de ácido abscísico, que exerce diversos efeitos fisiológicos no desenvolvimento das plantas e tem sido identificado como um mensageiro nas vias de resposta perceptiva de estresse hídrico e outros estresses ambientais, como baixas temperaturas e alta salinidade (WAN ; LI, 2006; CLEMENT et al., 2008). Alguns estudos mostraram que a aplicação de ABA em plantas intactas pode aumentar sua tolerância ao estresse (YIN et al, 2004).

O aumento do conteúdo de ABA durante o estresse hídrico foi encontrado em todos os organismos fotossintéticos. Acreditava-se que a biossíntese do ABA ocorresse apenas nas raízes, mas estudos mais recentes mostram que o ABA também é sintetizado nas células do mesófilo, tecido vascular e estômatos. Verificou-se que os genes que regulam pelo menos as últimas etapas da biossíntese do ABA (NCED e AAO) são os mais importantes e são fortemente regulados durante o estresse hídrico, mostrando o importante papel do ABA como uma resposta rápida ao estresse (SEO et al. , 2000).

Muitos genes responsivos ao estresse são regulados pelo ABA. Os primeiros estudos de ABA na transdução de sinal de estresse osmótico foram a indução de estresse de vários desses genes na *Arabidopsis*. Uma conclusão geral desses estudos foi que, enquanto a expressão

gênica regulada por baixa temperatura é relativamente independente do ABA, os genes regulados pelo estresse osmótico podem ser ativados através de vias dependentes e independentes de ABA (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2000).

2.6.2 Baixa temperatura

A ocorrência de geadas, mesmo que esporádicas, pode limitar fortemente o sucesso econômico da cultura. Nesse caso, baixas temperaturas provocam a descoloração das folhas, já exposição a temperaturas abaixo de 1°C pode resultar em lesões graves tanto nas folhas quanto nos frutos do café e abaixo de -2°C por períodos de 6 horas ou mais podem causar a morte das plantas (DESCROIX ; SNOECK, 2004).

Os cafeeiros arábicas são severamente danificados pela geada e, portanto, não são adequados para regiões que experimentam temperaturas abaixo de zero, mesmo por curtos períodos de tempo (GAY et al., 2006). Grandes mudanças nas temperaturas diurnas podem afetar o rendimento e a qualidade, e a tolerância máxima está na faixa de 19°C (DESCROIX; SNOECK, 2004).

O estresse por frio, que inclui temperaturas de resfriamento (<20°C) e/ou congelamento (<0°C), afeta negativamente o crescimento e o desenvolvimento das plantas. Resfriamento e congelamento são estresses que apresentam diferentes efeitos nas plantas: o primeiro leva a reações bioquímicas lentas, como atividades de transporte de enzimas e membranas; a segunda leva à formação de cristais de gelo que podem causar a ruptura do sistema de membrana celular (CHINNUSAMY et al., 2007).

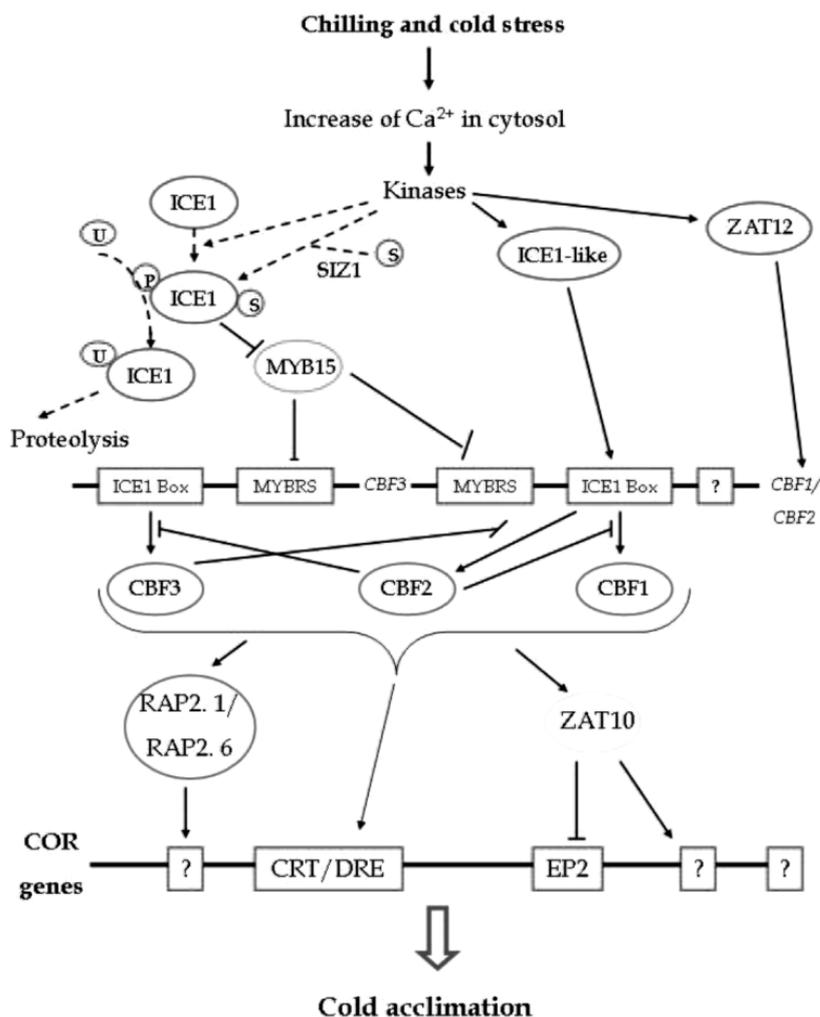
A baixa temperatura é um dos fatores abióticos mais importantes que limitam o crescimento, a produtividade e a distribuição geográfica das culturas agrícolas. Muitas plantas aumentam a tolerância ao congelamento em resposta à baixa temperatura, um fenômeno conhecido como aclimação ao frio. A aclimação ao frio é o processo que permite que plantas resistentes desenvolvam tolerância essencial para a sobrevivência ao estresse por frio através de múltiplos níveis de alterações bioquímicas e biológicas celulares. Essas respostas são devidas à reprogramação da expressão gênica que resulta nas alterações metabólicas ajustadas. O primeiro passo para ativar tais respostas moleculares é perceber o estresse à medida que ele ocorre e retransmitir informações sobre ele por meio de uma via de transdução de sinal (HEIDARVAND; MAALI AMIRI, 2010).

O estresse por frio induz a expressão de FTs da família AP2/ERF, ou seja, CBFs, que podem se ligar a elementos cis nos promotores dos genes COR e ativar sua expressão

(CIARMIELLO et al., 2011) (Figura 4). Os CBFs regulam a expressão de genes envolvidos no metabolismo de fosfoinosítídeos, transcrição, biossíntese de osmólitos, desintoxicação de ROS, transporte de membrana, hormônio metabolismo e sinalização e muitos outros com funções de proteção celular conhecidas ou presumidas (LEE et al., 2005).

Um grande número de estudos usou uma abordagem de perfil transcricional para identificar genes em *Arabidopsis* que respondem a temperaturas muito frias (4°C) e frias (13°C). Os resultados mostraram que as plantas respondem a baixas temperaturas alterando os níveis de mRNA de um grande número de genes pertencentes a vias diferentes e independentes. A indução a frio de genes envolvidos na sinalização de cálcio, sinalização lipídica ou codificação de proteínas quinases semelhantes a receptores também são afetadas pela mutação do gene *ice1* (LEE et al., 2005).

Figura 4 - Rede transcripcional responsiva ao frio em *Arabidopsis*.



Legenda: CBFs regulam a expressão de genes COR que conferem tolerância ao frio. Os CBFs podem regular a transcrição um do outro. CBFs induzem a expressão de ZAT10 que pode regular negativamente a expressão de genes COR. O ICE1 expresso constitutivo é ativado através de sumoilação e fosforilação induzida por estresse por frio. ICE1 ativado induz a transcrição de CBFs e reprime MYB15. A expressão de CBFs é regulada negativamente por MYB15 e ZAT12. HOS1 medeia a ubiquitinação e proteólise de ICE1, regula assim negativamente os reguladores CBF. As linhas que terminam com barra indicam regulação negativa; ponto de interrogação (?) indica elementos cis desconhecidos; setas quebradas indicam regulação pós-traducional; setas sólidas indicam ativação; linhas que terminam com barra indicam regulação negativa.

Fonte: Ciarmiello et al, (2011).

Em *Arabidopsis*, ICE1 (Inducer of CBF Expression1), um bHLH TF do tipo MYC, pode se ligar a elementos de reconhecimento de MYC no promotor CBF3 e é importante para a expressão de CBF3 durante a aclimatação ao frio. ICE1 é constitutivamente expresso e localizado no núcleo, mas induz a expressão de CBFs apenas sob estresse por frio. Isso sugere que a modificação pós-traducional induzida pelo estresse por frio é necessária para que o ICE1 ative os genes de maneira downstream nas plantas (CHINNUSAMY et al., 2003).

A expressão de genes regulados a frio é regulada por vias independentes e dependentes de ABA (CHINNUSAMY et al., 2004). A análise do promotor dos genes regulados pelo frio mostrou que eles contêm elementos de sequência que medeiam a indução de estresse dos genes. Existem alguns fatores de transcrição que identificam esses elementos e os ligam (HEIDARVAND; MAALI AMIRI, 2010).

2.6.3 Alta temperatura

Para *C. arabica*, a temperatura média anual ótima varia de 18 a 23°C. Acima de 23°C, o desenvolvimento e o amadurecimento dos frutos são acelerados, muitas vezes levando à perda de qualidade (CAMARGO, 1985). A exposição contínua a temperaturas diárias de até 30°C pode resultar não apenas em diminuição do crescimento, mas também em anormalidades como o amarelamento das folhas (DAMATTA; RAMALHO, 2006). A temperatura ótima dia/noite é de 18/22°C, com extremos tolerados para o café arábica chegando a uma mínima de 15°C durante a noite, e entre 25°C - 30°C durante o dia (DESCROIX; SNOECK, 2004).

As altas temperaturas podem prejudicar significativamente a produtividade das plantas de café de várias maneiras. A exposição contínua a temperaturas de até 30°C leva a um crescimento deprimido e anormalidades, como amarelamento das folhas, crescimento de tumores no caule e durante a floração pode causar aborto de flores, reduzindo a produtividade do café. Além disso, favorecem o desenvolvimento da Ferrugem da Folha do Café (*Hemileia*

vastatrix) e da praga dos frutos, enquanto a Doença da Baga do Café é mais proeminente em regiões frias (VAN DER VOSSSEN; BERTRAND, 2015).

A alta temperatura é agora uma grande preocupação para a produção agrícola e as abordagens para sustentar altos rendimentos de plantas cultivadas sob esse estresse são importantes objetivos agrícolas (HASANUZZAMAN et al., 2013a). As condições que excedem a temperatura ideal para o crescimento são reconhecidas como estresse térmico (HS) em organismos vivos. As células vegetais respondem ao HS por mecanismos de percepção de temperatura, vias de transdução de sinalização e acúmulo de fatores de transcrição de choque térmico (Hsfs) conservados, bem como proteínas relacionadas ao HS, como chaperonas moleculares ou proteínas de choque térmico (Hsps) (KOTAK et al., 2007).). Uma das principais consequências do estresse térmico é o excesso de geração de espécies reativas de oxigênio (ROS), que leva ao estresse oxidativo (HASANUZZAMAN et al, 2013a).

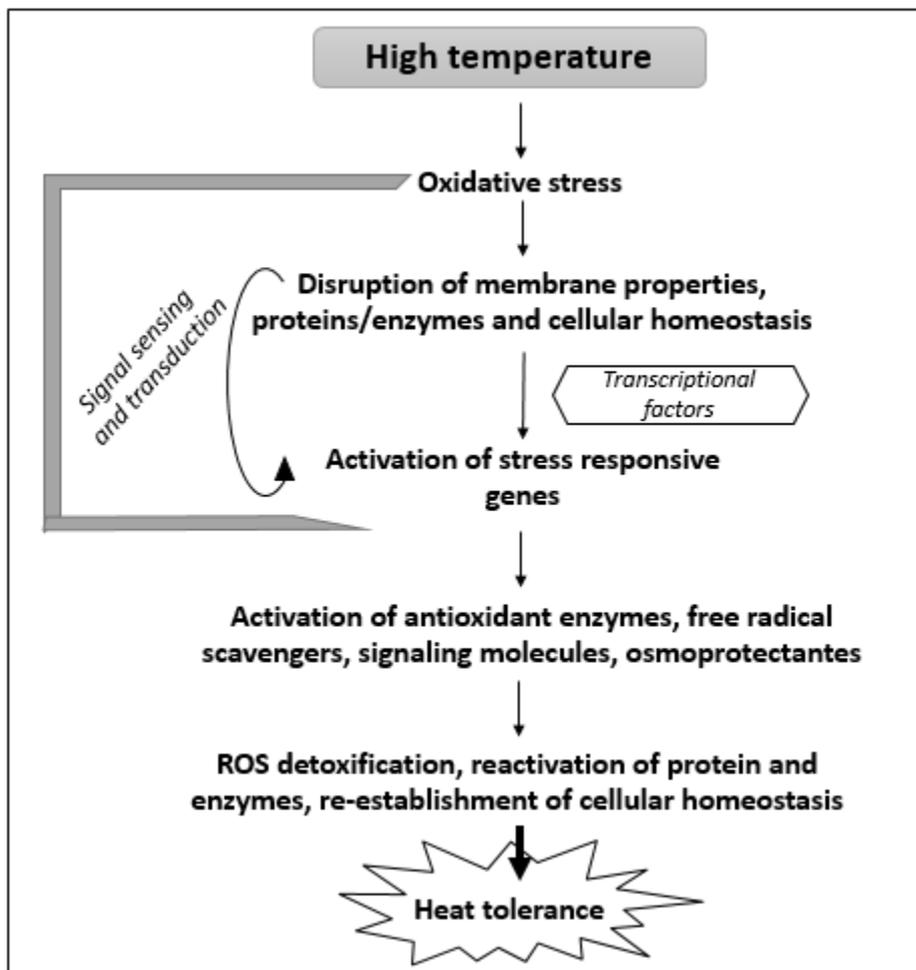
Vários trabalhos na literatura citaram tanto em nível estrutural quanto em nível molecular, os processos que ocorrem nas células vegetais durante o estresse térmico. As plantas alteram seu metabolismo de várias maneiras em resposta à temperatura do calor, principalmente produzindo solutos compatíveis que são capazes de organizar proteínas e estruturas celulares, manter o turgor celular por ajuste osmótico e modificar o sistema antioxidante para restabelecer o equilíbrio redox e a homeostase celular. No nível molecular, o estresse térmico causa alterações na expressão de genes envolvidos na proteção direta ao estresse térmico. Estes incluem genes responsáveis pela expressão de osmoprotetores, enzimas desintoxicantes, transportadores e proteínas reguladoras (KRASENSKY; JONAK, 2012).

A regulação transcricional da expressão de genes nucleares responsivos a HS é um dos principais alvos das cascatas de transdução de sinal relacionadas a HS e requer conjuntos de fatores de ação trans e elementos cis-regulatórios (CREs) nas regiões promotoras de genes induzíveis por HS (CHEN et al., 2012). Os HSFs servem como componentes terminais da transdução de sinal, mediando a expressão de HSPs e outros transcritos induzidos por HS. As plantas possuem múltiplos genes codificadores de HSF, com 21 membros definidos em *Arabidopsis* (KOTAK et al., 2007).

A regulação positiva de muitos genes tem sido relatada para ajudar a planta a suportar as condições de estresse que levam à adaptação da planta (TUTEJA, 2009). Sob estresse as plantas percebem os sinais externos e internos através de diferentes vias independentes ou interligadas que são usadas para regular várias respostas para o desenvolvimento de sua tolerância. A adaptação das plantas ao estresse térmico inclui mecanismos de prevenção e

tolerância que empregam uma série de estratégias (HASANUZZAMAN et al., 2013a) (Figura 5).

Figura 5 - Ilustração esquemática do mecanismo de transdução de sinal induzido pelo calor e desenvolvimento de tolerância ao calor em plantas.



Fonte: Adaptado de Hasanuzzaman et al, (2013a).

2.6.4 Estresse foto-oxidativo

Para as plantas, a luz é indiscutivelmente o parâmetro ambiental mais importante, pois fornece a melhor fonte de energia biológica. Desempenha um papel crítico na regulação do crescimento e desenvolvimento das plantas através da modulação dos níveis de expressão de genes responsivos à luz que regulam os processos metabólicos e de desenvolvimento (QUAIL, 2002b).

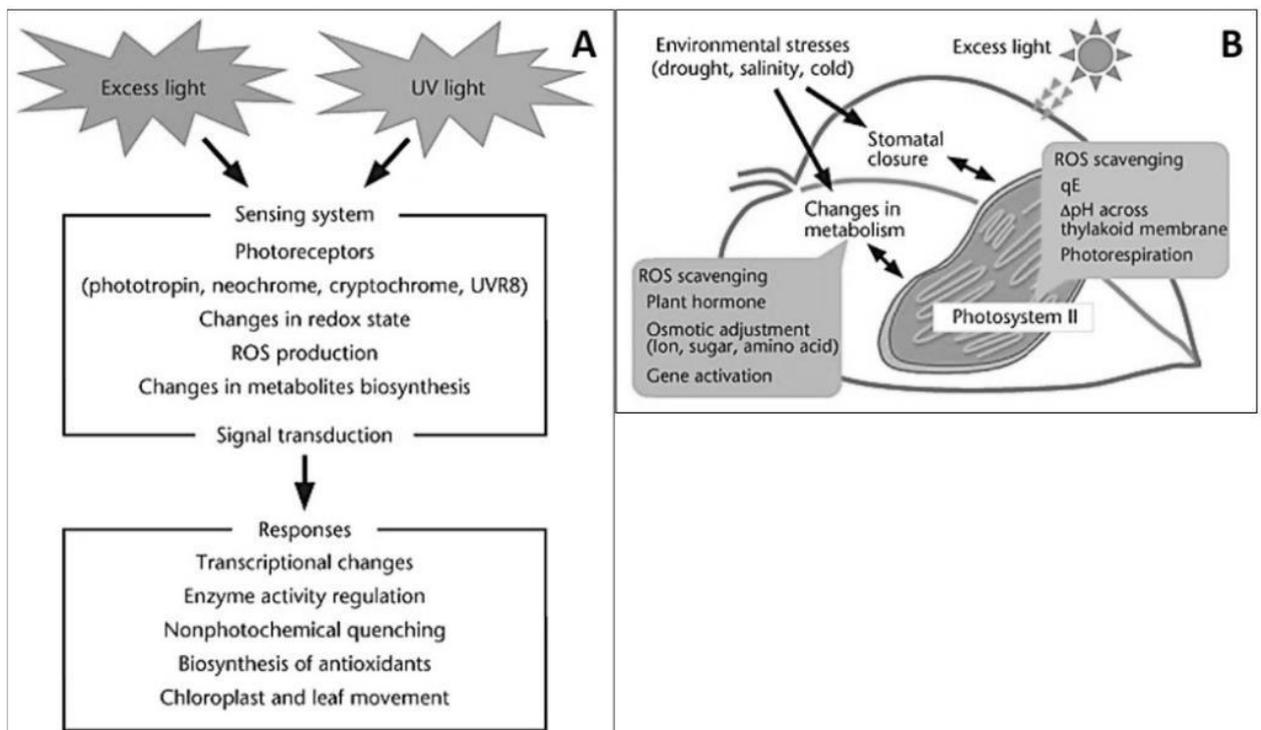
Embora a luz seja essencial para a fotossíntese e, portanto, crucial para a sobrevivência das plantas, ela também pode causar estresse oxidativo. A exposição de uma

planta à luz que excede o que é utilizado na fotoquímica leva à inativação das funções fotossintéticas e à produção de espécies reativas de oxigênio (ROS), como peróxido de hidrogênio (H_2O_2), superóxido (O_2^-), radicais hidroxila e oxigênio singlete (1O_2) (NIYOGI, 1999).

Os efeitos dessas EROs causam distúrbios no estado redox normal das células podendo causar efeitos tóxicos através da produção de peróxidos e radicais livres que danificam todos os componentes da célula, incluindo proteínas, lipídios e DNA. O estresse oxidativo do metabolismo oxidativo causa danos à base, bem como quebras de fitas no DNA (KALA et al, 2015). Além do excesso de luz, uma série de condições ambientais abióticas como O_3 , sal, metais tóxicos e temperatura podem induzir o aumento da produção de EROs, limitando a capacidade de uma planta de utilizar a energia luminosa através da fotossíntese (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2000).

O estresse oxidativo é descrito como um desequilíbrio entre espécies reativas de oxigênio (EROs) e antioxidantes em sistemas biológicos. Ela pode ser desencadeada pelo aumento da formação de EROs e/ou redução das defesas antioxidantes, e uma variedade de condições estressantes, tanto bióticas quanto abióticas, desencadeia um aumento de EROs em células animais e vegetais (MAHALINGAM; FEDOROFF, 2003) (Figura 6).

Figura 6 – Percepção do estresse luminoso por ROS.



Legenda: (A) Esquema de detecção e respostas ao excesso de luz. Estresses de luz são percebidos por fotorreceptores e causam alterações na biossíntese de metabólitos, produção de espécies reativas de oxigênio (ERO) e estado redox. As vias a jusante são ativadas e controlam as respostas da planta, como ativação de genes, atividade enzimática, extinção não fotoquímica, biossíntese de antioxidantes e movimentos de cloroplastos e folhas. (B) Fotoproteção das plantas em excesso de estresse por luz e as respostas ao estresse ambiental nas plantas. O PSII fotodanificado é reparado por eliminação de ERO, qE dependente de ΔpH e fotorrespiração em cloroplastos. O fechamento estomático, a produção de EROs e as alterações metabólicas sob estresse ambiental afetam esses mecanismos de reparo. Fonte: Osakabe et al. (2012).

Embora tóxicas em altas concentrações, as espécies reativas de oxigênio atuam como moléculas sinalizadoras durante o desenvolvimento da planta, bem como sinalizando vias responsivas ao estresse (MØLLER; SWEETLOVE, 2010). Para evitar a superprodução de EROs no interior das células, as plantas possuem mecanismos que visam diminuir a produção dessas moléculas em situações de estresse (MITTLER, 2002). Os mecanismos utilizados podem ser via adaptações anatômicas, fisiológicas, moleculares (JOHNSON et al., 2011) e via enzimas oxidases alternativas, que retiram elétrons da cadeia fotossintética para utilizá-los na redução do oxigênio molecular intracelular, reduzindo assim o fluxo de elétrons e concentração molecular de oxigênio, ambos levando a uma diminuição na produção de EROs (MITTLER, 2002).

Além dos mecanismos que visam diminuir a produção de EROs na célula vegetal, eles também possuem um complexo sistema antioxidante. Este sistema constitui uma gama complexa de antioxidantes enzimáticos e não enzimáticos que regulam as EROs, protegendo as células do dano oxidativo (VRANOVÁ et al., 2002). Enzimas, incluindo catalase (CAT), superóxido dismutase (SOD), peroxidase (POX), ascorbato peroxidase (APX) e glutathione reductase (GR) e antioxidantes não enzimáticos como tocoferóis, ácido ascórbico, flavonóides e glutathione trabalham juntos para regular as quantidades de ROS (HUNG et al., 2005). Dentre estas, destacam-se as enzimas SoD, Cat e APx na detoxificação do peróxido de hidrogênio e do íon superóxido. O processo de detoxificação do H₂O₂ catalisado pela enzima ascorbato-peroxidase é acompanhado por uma série de reações que juntas formam uma das mais importantes vias antioxidantes presentes nas plantas, o ciclo ascorbato-glutathione (GILL; TUTEJA, 2010).

As abordagens genéticas e moleculares clássicas identificaram vários reguladores a jusante dos fotorreceptores. Muitos destes codificam FTs, bem como quinases, fosfatases e proteínas da via de degradação. Embora alguns desses reguladores sejam específicos para a qualidade da luz, outros regulam as redes de transdução de sinal em resposta a vários sinais de luz, representando potenciais pontos de integração do sinal. Vários mecanismos pós-traducionais básicos estão

envolvidos na regulação das atividades de FT e na localização subcelular em resposta à luz. A fosforilação de FTs é uma modificação comum que pode influenciar sua capacidade de se ligar a promotores (KLIMCZAK et al., 1995).

O HY5, um fator de transcrição bZIP que é um dos principais reguladores da fotomorfogênese controlada por criptocromo e fitocromo, é um importante componente da rede de sinalização induzida por UVB. UVB promove rápida ativação transcricional de HY5 independentemente de todos os fotorreceptores conhecidos, e a perda de HY5 resulta no comprometimento da indução transcricional de um subconjunto de genes responsivos a UVB (ULM et al., 2004). Dado que HY5 parece regular a expressão de vários genes de *Arabidopsis* conhecidos por responder a condições de estresse abiótico (por exemplo, CBF1, DREB2A, RD20 e MYB59). Infere-se que o HY5 também poderia estar envolvido na regulação da fotossíntese por condições ambientais adversas (LEE et al., 2007).

As folhas de *Arabidopsis thaliana* expostas à luz intensa acumulam peróxido de hidrogênio (H₂O₂) nos cloroplastos das células da bainha do feixe como parte de uma rede de sinalização retrógrada que induziu ascorbato peroxidase 2 (APX2) (GALVEZ-VALDIVIESO et al., 2009).

Alterações no nível de mRNA de genes de *Arabidopsis* após exposição a condições de alta luminosidade foram relatadas por Rossel et al. (2002). Os autores identificaram um aumento na expressão de genes antioxidantes conhecidos como ascorbato peroxidase (APX1) e dehidroascorbato redutase (DHAR), bem como genes desconhecidos com homologias com genes reguladores e enzimas metabólicas. Além disso, eles descobriram que vários genes da proteína de choque térmico foram regulados positivamente, implicando-os na resposta antioxidante, além de sua função de chaperonas.

Em resposta a alta luz, as plantas de *Arabidopsis* de tipo selvagem mostraram expressão muito melhorada de CSD1 e CSD2 e maior atividade enzimática de MKK5 (proteína quinase quinase 5 ativada por mitógeno), que mediou um sinal da alta expressão induzida por luz dos genes CSD1 e CSD2. Manipular MKK5 e, assim, regular os níveis de transcritos de CSD1 e CSD2 pode melhorar a tolerância da planta ao alto estresse por luz (XING et al., 2013).

2.6.5 Seca

Diversas definições para déficit hídrico estão disponíveis na literatura. Taiz e Zeiger (2004) definem como um dos principais estresses que restringem a capacidade e eficiência da planta em realizar importantes processos bioquímicos, causados pelo déficit hídrico. Isso pode

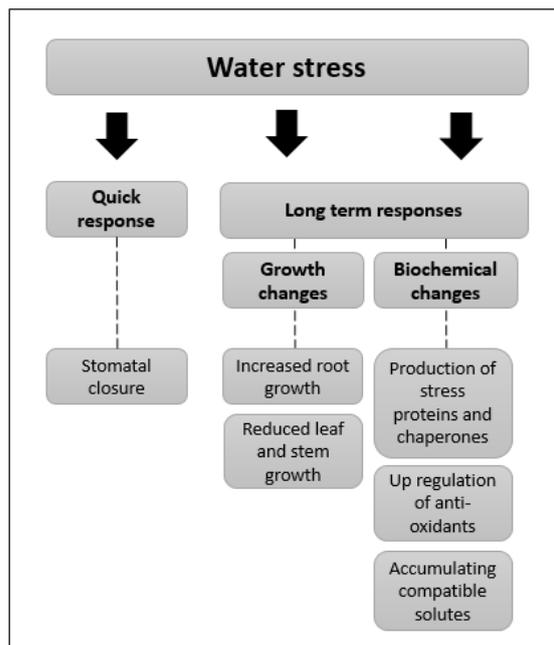
ser definido como "qualquer teor de água de um tecido ou célula que esteja abaixo do teor de água mais alto exibido no estado de maior hidratação".

Outra definição seria a redução da disponibilidade de água; seja por escassez de água (seca) ou estresse osmótico (altas concentrações de sal) ou encharcamento; muita água. Na natureza o estresse hídrico é comum por longos ou curtos períodos de tempo, dependendo do clima local. A maioria das plantas, portanto, tem alguma adaptação ou resposta para aumentar a taxa de crescimento e sobrevivência durante o estresse hídrico e posterior recuperação (ARVE et al., 2011).

O estresse hídrico pode reduzir a fotossíntese, a respiração e a absorção de íons, alterar os padrões metabólicos e de crescimento da planta e, em casos graves, resultar na morte da planta. Os ambientes que proporcionam esse déficit hídrico são condições de seca, alta salinidade e baixa temperatura (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 1997; SEKI et al., 2003). Na agricultura o estresse hídrico é um dos maiores problemas, causando grandes perdas de safra a cada ano, bem como perda de valor estético em ornamentais.

Com a população humana global aumentando rapidamente, simultaneamente com o aumento da escassez de água, a perda de colheita será ainda mais grave do que antes. A descoberta e o desenvolvimento de culturas tolerantes ao estresse para evitar a perda de rendimento durante o estresse hídrico é, portanto, muito importante. Em resposta ao estresse hídrico, as plantas desenvolveram vários diferentes mecanismos que aumentam a tolerância à dessecação e a retenção de água. Essas respostas podem ser divididas em respostas de curto e longo prazo (Figura 7) (ARVE et al., 2011).

Figura 7 – Respostas das plantas ao estresse hídrico.



Fonte: Adaptado de ARVE et al. (2011).

2.6.5.1 Respostas a longo prazo

Durante o estresse hídrico prolongado, as plantas devem ser capazes de sobreviver com baixo teor de água e manter uma quantidade mínima de água, através da absorção e retenção de água. Para lidar com o estresse hídrico prolongado, as plantas respondem com processos demandantes de energia que alteram o padrão de crescimento, o conteúdo químico das plantas e a regulação para cima ou para baixo de genes (ARVE et al., 2011).

Quando a disponibilidade de água é reduzida, as plantas alteram a bioquímica para serem capazes de reter o máximo de água possível e absorver a água que puderem. Durante o estresse hídrico, as plantas produzem e acumulam solutos compatíveis, como açúcares, polióis e aminoácidos, para diminuir o potencial osmótico nas células e facilitar a absorção e retenção de água. Alguns dos solutos compatíveis também contribuem para manter a conformação das macromoléculas, evitando o dobramento incorreto ou a desnaturação (XIONG; ZHU, 2002). As plantas também produzem níveis mais altos do hormônio do estresse da planta ABA durante o estresse hídrico e isso afeta seu padrão de crescimento e tolerância ao estresse (ARVE et al., 2011).

Um grupo de proteínas chamadas de proteínas abundantes na embriogênese tardia (LEA) também são produzidas durante o estresse hídrico. Essas proteínas semelhantes a LEA são altamente hidrofílicas, ricas em glicina e altamente solúveis e foram identificadas como reguladas pelo ABA (XIONG; ZHU, 2002). Acredita-se que as proteínas do tipo LEA atuem

como chaperonas, protegendo as atividades enzimáticas e prevenindo o dobramento incorreto e a desnaturação de proteínas importantes (XIONG; ZHU, 2002).

Alguns estudos mostram que o ABA pode funcionar como um sinal para reduzir a taxa de crescimento foliar, tanto quando o ABA é aplicado exogenamente quanto gerado por estresse hídrico. Ao reduzir a expansão foliar, as folhas ficam menores e, portanto, transpiram menos. Em alguns casos, o estresse hídrico pode até levar à abscisão das folhas. A redução do volume celular também concentra os solutos nas células e comprime as membranas plasmáticas fazendo com que elas aumentem de espessura (WILKINSON; DAVIES, 2010).

2.6.5.2 Resposta de curto prazo

Quando as plantas de repente encontram a seca, é importante responder o mais rápido possível. Uma resposta mais rápida à seca significa que menos água é perdida e a taxa de sobrevivência das plantas é aumentada. A resposta rápida mais importante é o fechamento estomático (FAROOQ et al., 2009).

O desenvolvimento dos estômatos é frequentemente considerado um dos desenvolvimentos mais importantes na evolução das plantas. Grande parte da perda de água ocorre via estoma através da transpiração, devido à exposição foliar à luz solar, e consequente aquisição de luz solar e dióxido de carbono (CO₂), que são fundamentais para o processo fotossintético (BRODRIBB; McADAM, 2011).

Os estômatos são pequenos poros presentes na epiderme dos tecidos verdes que medeiam as trocas entre a planta e a atmosfera: o CO₂ entra pelo poro estomático como fonte de carbono para a fotossíntese, enquanto o vapor de água é liberado pela transpiração. Os estômatos são circundados por um par de células altamente especializadas, denominadas células-guarda, cujas alterações da pressão de turgescência controlam a abertura e o fechamento dos estômatos (SIRICHANDRA et al., 2009). Curiosamente, o transcriptoma da célula-guarda é particularmente rico em genes codificadores de TF (BATES et al., 2012).

Durante o estresse hídrico, o poro estomático pode ser fechado para reduzir a perda de água. Ao fechar o poro estomático aumenta-se a eficiência do uso da água (FAROOQ et al., 2009), reduzindo a quantidade de água perdida por molécula de CO₂ assimilada. Vários mecanismos trabalham juntos para fechar os estômatos, como o fechamento hidropassivo e os sinais químicos do hormônio do estresse da planta ABA. O aumento dos níveis de ABA também causa aumento da condutividade hidráulica nas raízes e no xilema, permitindo que as plantas

transportem mais água e assim se recuperem mais rapidamente após o estresse hídrico (KUDOYAROVA et al., 2011).

2.7 Transcriptoma

O transcriptoma compreende o conjunto completo de transcritos de uma célula, bem como sua quantidade, para um determinado estágio específico do desenvolvimento ou condição fisiológica de um organismo. Sua compreensão é essencial para a interpretação dos elementos funcionais do genoma, para a descoberta dos componentes moleculares de células e tecidos e para a compreensão dos estágios de desenvolvimento e doença. Os principais objetivos da transcriptômica são catalogar todas as espécies de transcritos (mRNAs, RNAs não codificantes, RNAs pequenos, etc.); determinar a estrutura transcricional do gene (seus locais de iniciação; extremidades 5' e 3'; padrões de emenda e outras modificações pós-transcricionais) e quantificar as mudanças nos níveis de expressão gênica de cada transcrito durante o desenvolvimento do organismo e em outras condições diferentes (WANG; GERSTEIN; SNYDER, 2009).

A investigação das respostas moleculares de plantas a múltiplos estresses muitas vezes se concentra na sobreposição de padrões de transcrição. Para tanto, diversos estudos têm sido realizados nos quais diferentes grupos de plantas são expostos a um ou outro estresse em paralelo, e seus padrões de expressão gênica comparados. Conjuntos de genes sobrepostos que são regulados por ambos os estresses são então identificados e propostos para representar uma resposta generalizada ao estresse ou pontos de cross-talk entre vias de sinalização (HUANG et al., 2008).

Uma sequência completa do genoma e mapa genético de alta densidade de *Coffea canephora* foi publicado recentemente (DENOEUDE et al., 2014). Isso oferecerá oportunidades para melhorar o progresso do melhoramento genético para aumentar a qualidade e a produtividade da lavoura, bem como proteger a lavoura de café de grandes perdas causadas por doenças, pragas de insetos e estresses abióticos relacionados às mudanças climáticas (GONGORA et al., 2015).

Outros estudos de transcriptoma em várias espécies identificaram genes, hormônios e processos importantes no controle da resposta da planta a múltiplos estresses abióticos ou bióticos e forneceram alvos para a melhoria da tolerância ao estresse (WANG et al., 2010; GRIGOROVA et al., 2011) .

A maioria dos dados de sequenciamento do transcriptoma do café se baseia em duas principais espécies cultivadas (*C. arabica* e *C. canephora*). Projetos de sequenciamento Sanger

EST foram desenvolvidos para *C. arabica* (VIEIRA et al., 2006; VIDAL et al., 2010; MONDEGO et al., 2011) e *C. canephora* (LIN et al., 2005; PONCET et al., 2006). A análise do transcriptoma usando sequenciamento de próxima geração – NGS tem sido realizada em estudos que investigam interações bióticas e abióticas (FERNANDEZ et al., 2004; COMBES et al., 2013). Os dois subgenomas contidos no genoma alotetraplóide de *C. arabica* (subgenoma *C. canephora*—CaCc e subgenoma *C. eugenioides*—CaCe) não contribuem igualmente para o transcriptoma (VIDAL et al., 2010; COTTA et al., 2014).

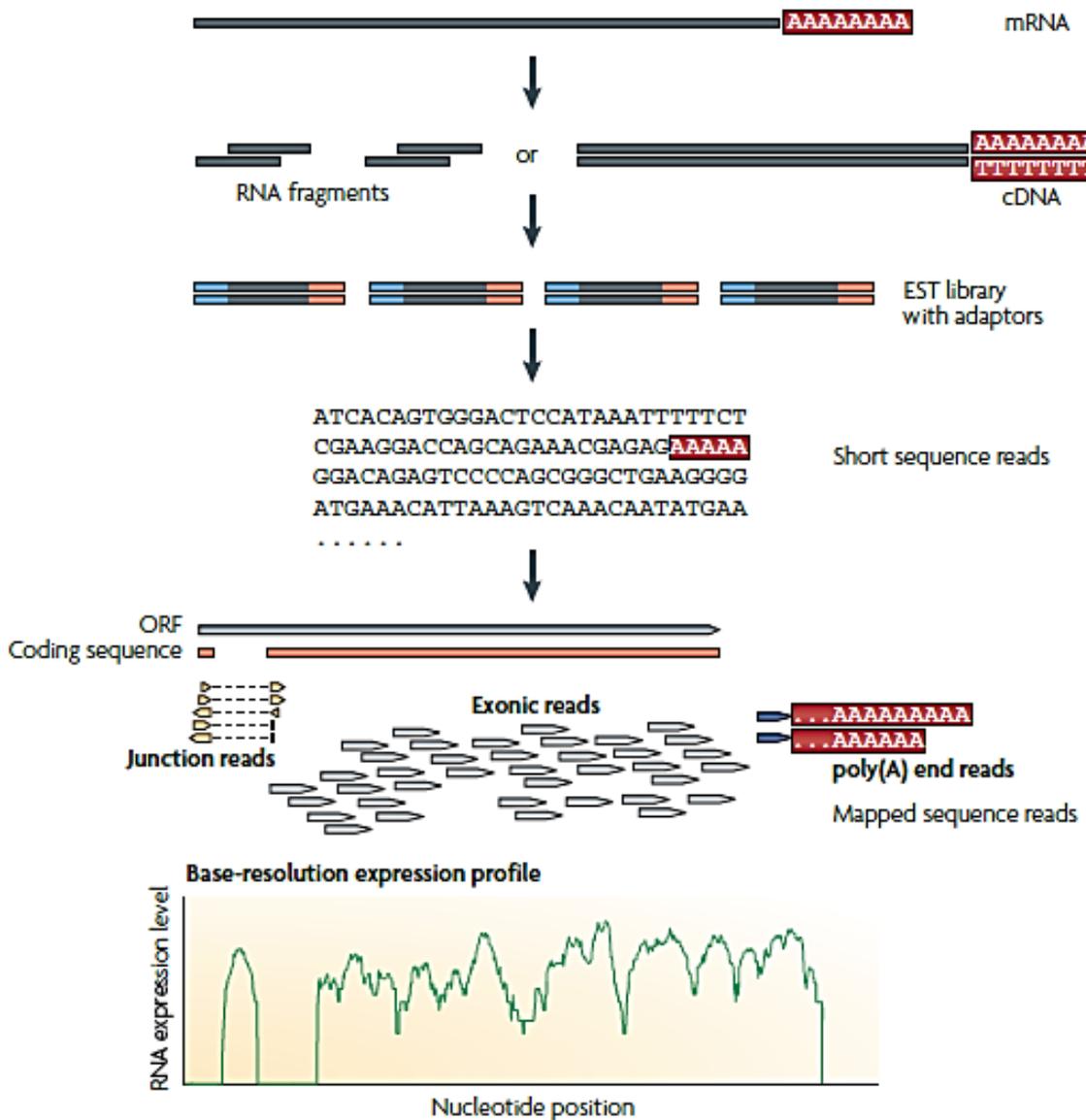
Em *Coffea sp.*, recursos de EST foram desenvolvidos para várias espécies e tecidos, incluindo raízes, folhas e frutos (LIN et al., 2005). No entanto, não há recursos genômicos disponíveis para os ápices caulinares, que são considerados órgãos-chave para o desenvolvimento das plantas, integrando diversos sinais, como estímulos ambientais, além de hormônios (ácido abscísico [ABA], auxinas, citocininas) e transcrição (TRAAS; VERNOUX, 2002). Por outro lado, o sequenciamento de próxima geração (NGS) oferece novas oportunidades para estudar as respostas transcriptômicas e combinar o sequenciamento de alto rendimento com a capacidade de anotação funcional dos ESTs gerados (TORRES et al., 2008).

2.8 Recursos para análise de genoma usando dados de RNA-seq

Ao contrário do genoma, o transcriptoma celular é dinâmico e específico para um determinado estágio de desenvolvimento e condição fisiológica. A compreensão do transcriptoma é essencial na interpretação dos elementos funcionais do genoma e no desenvolvimento dos constituintes moleculares celulares. O recente desenvolvimento de novos métodos de alto rendimento para sequenciamento de DNA levou ao surgimento de uma nova metodologia que permite o sequenciamento de moléculas de RNA em resoluções sem precedentes. Esse método é chamado de RNA-Seq e surgiu como a tecnologia preferida na caracterização e quantificação de transcritos celulares (CRUZ, 2013). RNA-Seq é o primeiro método baseado em sequenciamento que permite que um transcriptoma completo seja pesquisado em larga escala e quantitativamente, oferecendo resolução de até uma única base para anotação e níveis de expressão gênica “digital” em escala genômica, geralmente a um custo bem menor quando comparado com técnicas de microarray de alta densidade ou sequenciamento EST. Destaca-se como uma ferramenta valiosa para a compreensão da dinâmica transcricional, não apenas durante as alterações fisiológicas normais associadas ao desenvolvimento do organismo, mas também na comparação de tecidos normais e doentes (WANG; GERSTEIN; SNYDER, 2009).

Ao longo dos anos, um grande número de metodologias foram desenvolvidas com o objetivo de traçar o perfil do transcriptoma. Um desses métodos revolucionou o perfil de expressão ao permitir a medição de milhares de genes simultaneamente é a moderna tecnologia RNA-Seq baseada em sequenciamento de alto rendimento (Figura 8) (WANG; GERSTEIN; SNYDER, 2009).

Figura 8 - Um experimento típico de RNA-seq. Resumidamente, RNAs longos são primeiro convertidos em uma biblioteca de fragmentos de cDNA através de fragmentação de RNA ou fragmentação de DNA.



Legenda: Adaptadores de sequenciamento (azul) são posteriormente adicionados a cada fragmento de cDNA e uma sequência curta é obtida de cada cDNA usando tecnologia de sequenciamento de alto rendimento. As leituras de sequência resultantes são alinhadas com o genoma de referência ou transcriptoma e classificadas em três tipos: leituras exônicas, leituras de junção e leituras finais poli(A).

Esses três tipos são usados para gerar um perfil de expressão de resolução de base para cada gene, conforme ilustrado na parte inferior; uma ORF de levedura com um íntron é mostrada.

Fonte: Wang, Gerstein e Snyder, (2009).

Uma vantagem adicional é que a tecnologia RNA-Seq pode ser usada não apenas para perfis de expressão gênica, mas também para detecção de eventos de fusão gênica (MAHER et al., 2009), descoberta de polimorfismos de nucleotídeo único (LALONDE et al., 2011), investigação de mutações de RNA pós-transcricionais (GARCION et al., 2004), estudo de eventos de splicing alternativos (PAN et al., 2008), descoberta de novos transcritos (GUTTMAN et al., 2010) ou investigação de expressão específica de alelos (DEGNER et al. ., 2009).

Existem várias tecnologias disponíveis para realizar o sequenciamento de alto rendimento de moléculas de DNA. Atualmente, os que dominam o campo são: 454 GS-FLX da Roche Applied Science, Genome Analyzer II da Illumina, Inc. e AB SOLiD da Applied Biosystems. No entanto, outras tecnologias estão sendo desenvolvidas com rendimentos de qualidade potencialmente ainda mais altos (por exemplo, Pacific Biosciences, Helicos). As diferentes tecnologias requerem protocolos experimentais distintos. Porém, a essência desses sistemas é a mesma: miniaturizar reações de sequenciamento individuais (BENTLEY et al., 2008).

2.9 Análise de expressão de dados de RNA - seq usando ferramentas de Bioinformática e Estatística

Nesse contexto, a bioinformática e a biologia computacional são campos interdisciplinares que combinam tecnologia computacional, matemática e biologia molecular para responder questões fundamentais nas ciências da vida (HOGEWEG, 2011). Particularmente, eles apresentam ferramentas poderosas que são adequadas para analisar dados gerados a partir de plataformas de sequenciamento de alto rendimento. Conseqüentemente, o sucesso das tecnologias de sequenciamento de próxima geração (NGS) está intimamente relacionado com a criação de ferramentas computacionais eficientes que sejam capazes de processar os dados. De fato, sem os meios fornecidos pelas ferramentas computacionais, o processamento dos dados NGS seria quase impossível (POP; SALZBERG, 2008).

A plataforma RNA-seq aborda uma infinidade de aplicações, incluindo análises de expressão relativa, splicing alternativo, descoberta de novos transcritos e isoformas, edição de RNA, expressão alelo-específica e a exploração de transcriptomas de organismos não-modelo

(WANG; GERSTEIN; SNYDER, 2009). Um dos principais objetivos do software de análise de sequenciamento de RNA é reconstruir o conjunto completo de transcrições (isoformas) de genes que estavam presentes nas células originais. Além das estruturas de transcrição, os experimentadores precisam estimar os níveis de expressão para todas as transcrições.

O primeiro passo no processo de análise é mapear as leituras de RNAseq em relação ao genoma de referência, que fornece o local de origem das leituras (MORTAZAVI et al., 2008). Em contraste com o alinhamento da sequência de DNA, os algoritmos de mapeamento de sequência de RNA têm dois desafios adicionais. Primeiro, porque os genes em genomas eucarióticos contêm íntrons, e porque as leituras sequenciadas de transcritos de mRNA maduros não incluem esses íntrons, qualquer programa de alinhamento de RNA-seq deve ser capaz de lidar com alinhamentos gapped (ou spliced) com gaps muito grandes. Em segundo lugar, a presença de pseudogenes processados, dos quais alguns ou todos os íntrons foram removidos, pode fazer com que muitas leituras de extensão de exon sejam mapeadas incorretamente (PEI et al., 2012).

Para lidar com este problema, o software TopHat2 utiliza um procedimento de duas etapas. TopHat é um alinhador de spliced popular para experimentos de RNA-seq que pode alinhar leituras de vários comprimentos produzidas pelas mais recentes tecnologias de sequenciamento, permitindo indels de comprimento variável em relação ao genoma de referência (KIM et al., 2013).

É interessante a descoberta de genes expressos diferencialmente em diferentes condições (por exemplo, tecidos, perturbações) enquanto opcionalmente se ajustam a outros fatores sistemáticos que afetam o processo de coleta de dados. Há vários aspectos sutis, mas cruciais, dessas análises, como contagem de leituras, tratamento adequado da variabilidade biológica, verificações de controle de qualidade e configuração apropriada de modelagem estatística. Diversas variações foram apresentadas na literatura, e há necessidade de orientação sobre as melhores práticas atuais (YU et al., 2013; CHU, 2014).

Um componente crucial de tal análise é o procedimento estatístico usado para chamar genes diferencialmente expressos. Existem dois protocolos amplamente utilizados para esta tarefa: DESeq (ANDERS; HUBER, 2010) e edgeR (ROBINSON; McCARTHY; SMYTH, 2010), ambos disponíveis como pacotes do projeto de desenvolvimento de software Bioconductor. Para dados de RNA-seq, a estratégia adotada é contar o número de reads que se enquadram em genes anotados e realizar análises estatísticas na tabela de contagens para

descobrir mudanças quantitativas nos níveis de expressão entre os grupos experimentais (ANDERS; HUBER, 2010).

3 OBJETIVOS

3.1 Objetivo Geral

Realizar uma análise comparativa dos resultados obtidos após o mapeamento de reads de RNA-seq e análises estatísticas, utilizando-se diferentes genomas de referência no pipeline das análises (*C.canephora*, *C. arabica*, *C. eugenoides*).

3.2 Objetivos específicos

Identificar potenciais homeólogos de *C. canephora* e *C. eugenoides*, diferencialmente expressos, após a exposição aos estresses abióticos estudados. Verificar os efeitos da utilização de diferentes genomas de referência nos resultados finais em análises de RNA-seq.

4 MATERIAIS E MÉTODOS

4.1 Local do experimento e material para análise

Os experimentos foram conduzidos no Laboratório de Genética Molecular (Inovacafé) da Universidade Federal de Lavras (UFLA) utilizando dados provenientes de estudos realizados anteriormente pelo grupo de pesquisa. Os dados coletados foram obtidos a partir do sequenciamento de RNA, obtidos de folhas de plântulas de (*C. arabica*) em condições controle e submetidas a diferentes estresses abióticos (restrição hídrica, calor, frio, calor, luminosidade e tratamento com ABA) (TORRES, 2017).

4.2 Pipeline de análises de Bioinformática

O alinhamento dos reads obtidos de RNA-seq, consistiu em: 1. Qualidade – eliminar sequencias de baixa qualidade (FASTQC); 2. Eliminar sequencias de adaptadores (cutAdaptor); 3. Mapeamento dos reads (Software STAR) e 4. Análises estatísticas (DESeq2). O mesmo

pipeline foi seguido para as análises com a alteração do genoma de referência, utilizado no mapeamento com o STAR. Para tanto, foram realizadas 3 análises distintas: 1. Mapeamento utilizando-se o genoma de referência de *C. arabica* (unpublished); 2. Mapeamento utilizando-se o genoma de referência de *C. eugenioides* (unpublished); 3. Mapeamento utilizando-se uma nova montagem do genoma de referência de *C. canephora* (New canephora -unpublished);

4.3 Delineamento experimental e estatístico

Por meio do programa R (versão 4.1.3) com o pacote Deseq2, foi possível realizar uma análise estatística paramétrica, para se identificar os genes diferencialmente expressos em cada condição analisada. Em cada um dos cinco tratamentos das três espécies, um script foi utilizado (figura 9) no qual ao seu final, gerava um gráfico com o total de genes diferencialmente expressos. Para cada análise (*C.arabica*, *C.canephora* e *C. eugenioides*) foram selecionados um total de 10 genes com base no valor de “log2foldchange”, sendo os 5 mais expressos e os 5 mais reprimidos, utilizando-se sempre, as condições controle, como parâmetro de comparação.

Figura 9 - Script utilizado para comparação de cada tratamento com o controle. Ao final do mesmo, um gráfico de dispersão foi gerado demonstrando a quantidade de genes diferencialmente expressos ($p < 0.05$).

```

ABAvsT0 - Bloco de Notas
Arquivo Editar Formatar Exibir Ajuda
library("DESeq2")

countData <- read.table ("C:/Program Files/R/R-4.1.3/library/Ricardo/Arabica/ABA_T0.txt")
colData <- read.csv ("C:/Program Files/R/R-4.1.3/library/Ricardo/Arabica/ABAvsT0_info.csv")

colData$condition <- factor(colData$condition, levels = c("ABA","Control"))

dds <- DESeqDataSetFromMatrix(countData = countData,
colData = colData,
design = ~ condition)

dds$condition <- relevel(dds$condition, ref="Control")

keep <- rowSums(counts(dds)) >= 10

ddsF <- dds[keep,]

dds1 <- DESeq(ddsF)

res <- results(dds1)

summary (res)

resultsNames(dds1)

resOrdered <- res[order(res$pvalue),]
write.csv(as.data.frame(resOrdered),
file="C:/Program Files/R/R-4.1.3/library/Ricardo/Arabica/Alan/resABAvsT0.csv")

resSig <- subset(resOrdered, padj < 0.05)
write.csv(as.data.frame(resSig),
file="C:/Program Files/R/R-4.1.3/library/Ricardo/Arabica/Alan/resSigABAvsT0.csv")

resNorm <- lfcShrink(dds1, coef=2, type="normal")
resOrderedNorm <- resNorm[order(res$pvalue),]
write.csv(as.data.frame(resOrderedNorm),
file="C:/Program Files/R/R-4.1.3/library/Ricardo/Arabica/Alan/resABAvsT0_Norm.csv")

resSigNorm <- subset(resOrderedNorm, padj < 0.05)
write.csv(as.data.frame(resSigNorm),
file="C:/Program Files/R/R-4.1.3/library/Ricardo/Arabica/Alan/resSigABAvsT0_Norm.csv")

plotMA(resNorm, ylim=c(-10,10))

#UP
par(mfrow=c(1,5), mar=c(4,4,2,1))
plotCounts(dds1, gene="Cara017g014900", intgroup="condition", returnData=FALSE)
plotCounts(dds1, gene="Cara018g007530", intgroup="condition", returnData=FALSE)
plotCounts(dds1, gene="Cara009g021370", intgroup="condition", returnData=FALSE)
plotCounts(dds1, gene="Cara003g037460", intgroup="condition", returnData=FALSE)
plotCounts(dds1, gene="Cara010g022480", intgroup="condition", returnData=FALSE)

*SALVAR IMAGEMSc

#DOWN
par(mfrow=c(1,5), mar=c(4,4,2,1))
plotCounts(dds1, gene="Cara003g017520", intgroup="condition", returnData=FALSE)
plotCounts(dds1, gene="Cara014g008300", intgroup="condition", returnData=FALSE)
plotCounts(dds1, gene="Cara001g017020", intgroup="condition", returnData=FALSE)
plotCounts(dds1, gene="Cara003g044960", intgroup="condition", returnData=FALSE)
plotCounts(dds1, gene="Cara012g015180", intgroup="condition", returnData=FALSE)
*SALVAR IMAGEM

```

Fonte: Do autor

As análises foram realizadas por meio da ferramenta de BlastP - Blast2sequences,

disponível no site do NCBI https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&BLAST_SPEC=blast2seq&LINK_LOC=blasttab (Figura 10)

Figura 10 – Print do site do National Library of Medicine que contém a ferramenta Blast P utilizada para a comparação entre as estruturas proteicas.

The image shows the NCBI BLAST web interface. At the top, the NIH logo and 'National Library of Medicine' are displayed. Below this, the 'BLAST® » blastp suite' is highlighted. The interface is divided into several sections:

- blastn** | **blastp** | blastx | tblastn | tblastx
- Enter Query Sequence**: A text input field for 'Enter accession number(s), gi(s), or FASTA sequence(s)', a 'Query subrange' section with 'From' and 'To' fields, and an 'Or, upload file' section with a file selector showing 'Ccan_RH9.txt'. There is also a 'Job Title' field and a checkbox for 'Align two or more sequences'.
- Enter Subject Sequence**: A text input field for 'Enter accession number(s), gi(s), or FASTA sequence(s)', a 'Subject subrange' section with 'From' and 'To' fields, and an 'Or, upload file' section with a file selector showing 'Ceug_RH9.txt'.
- Program Selection**: A radio button selected for 'blastp (protein-protein BLAST)' and a 'Choose a BLAST algorithm' link.
- BLAST** button: A blue button with the text 'BLAST'.
- Search protein sequence using Blastp (protein-protein BLAST)**: A checkbox checked for 'Show results in a new window'.
- + Algorithm parameters**: A blue bar at the bottom with a plus sign and the text 'Algorithm parameters'.

Fonte: Do autor

5 RESULTADOS E DISCUSSÕES

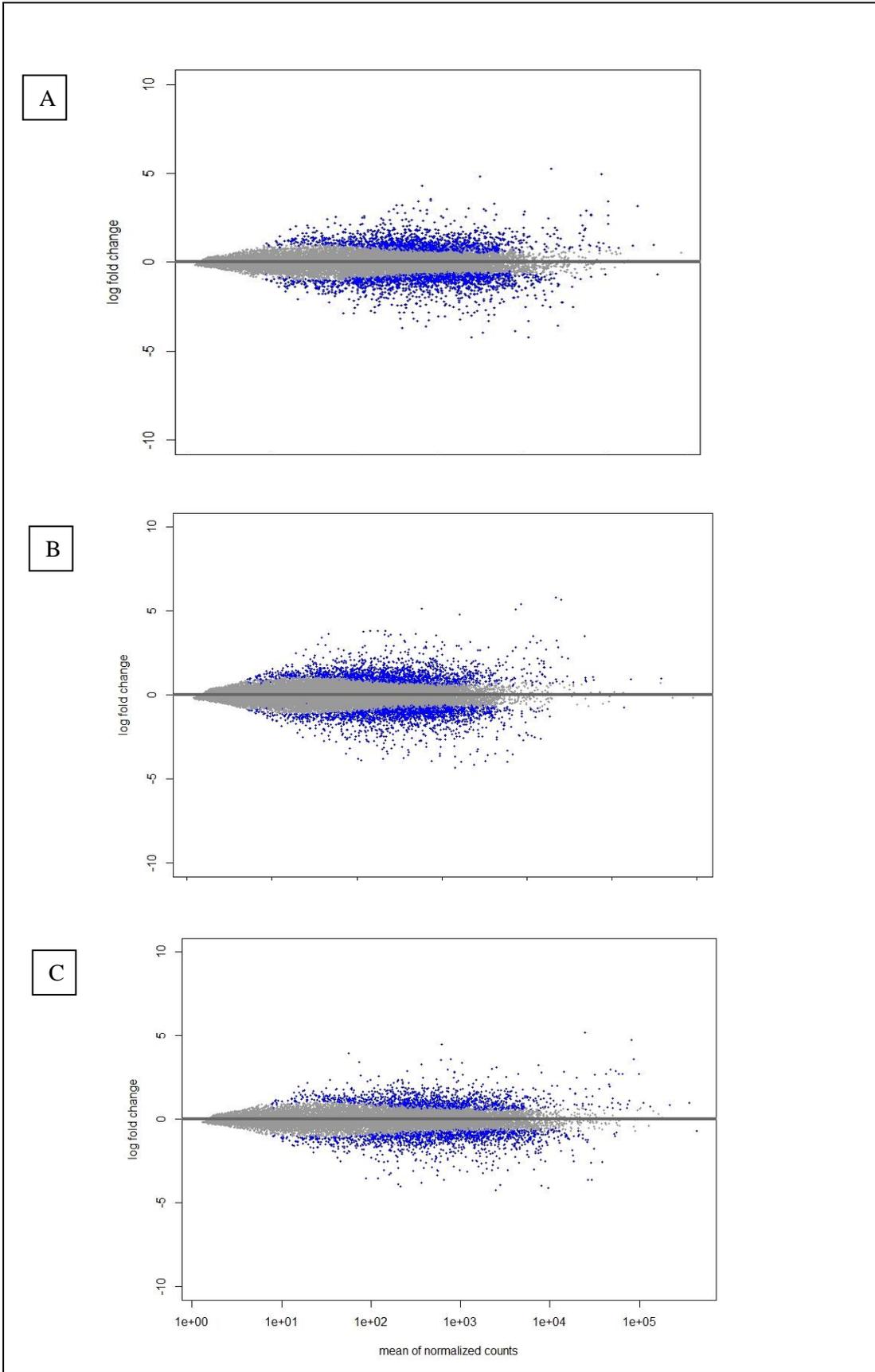
Gráficos de dispersão para as 3 análises realizadas (*C.arabica*, *C.canephora* e

C.eugenoides) foram gerados ao final de cada comparação dos tratamentos estudados e a condição controle. É possível notar padrões de expressão diferencial, com genes mais expressos e mais reprimidos. A seguir, resultados para cada tratamento serão expostos de maneira com que se possa observar e comparar esses padrões entre as espécies.

5.1 Padrões de genes diferencialmente expressos resultantes das 3 análises realizadas com diferentes genomas de referência (*C. arabica*, *C. canephora* e *C. eugenoides*), responsivos à exposição ao estresse da aplicação exógena de ABA

Na figura 11, estão apresentados os gráficos de dispersão para as 3 análises realizadas (Referência: *C.arabica*, *C.canephora* e *C.eugenoides*) gerados ao final da comparação do tratamento de ABA e a condição controle. É possível notar padrões de expressão diferencial, com genes mais expressos (*up_regulated*) e com maior repressão (*down_regulated*).

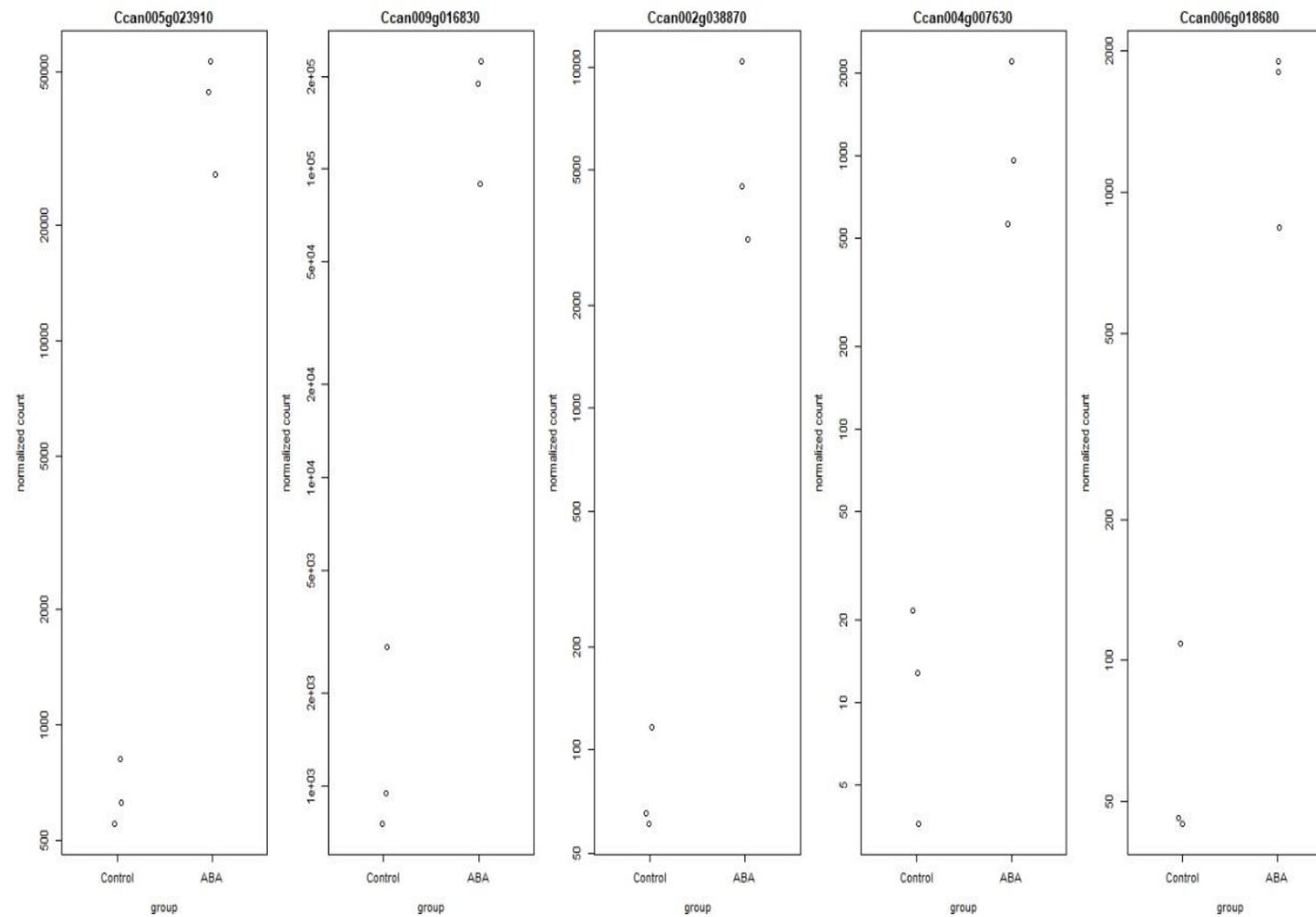
Figura 11 - Gráficos de dispersão contendo os genes de *Coffea new canephora* (A), *Coffea arabica* (B), *Coffea eugenoides* (C) com expressão diferencial significativa ($p < 0.05$; pontos em azul) quando expostos ao estresse da aplicação de ABA exógeno. Pontos na parte superior do eixo X representam genes *up-regulated* e na parte inferior, genes *down-regulated*.



Fonte: Do autor

Posteriormente foram selecionados os 5 genes mais expressos e os 5 genes com maior repressão. Na figura 12, é possível observar que os genes mais expressos com a análise utilizando-se o genoma de referência de *C. canephora* foram os Ccan005g023910, Ccan009g016830, Ccan002g038870, Ccan004g007630, Ccan006g018680.

Figura 12 - 5 genes mais expressos na condição de estresse induzido pela aplicação de ABA externo. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente menor. Esses são os Ccan005g023910, Ccan009g016830, Ccan002g038870, Ccan004g007630, Ccan006g018680

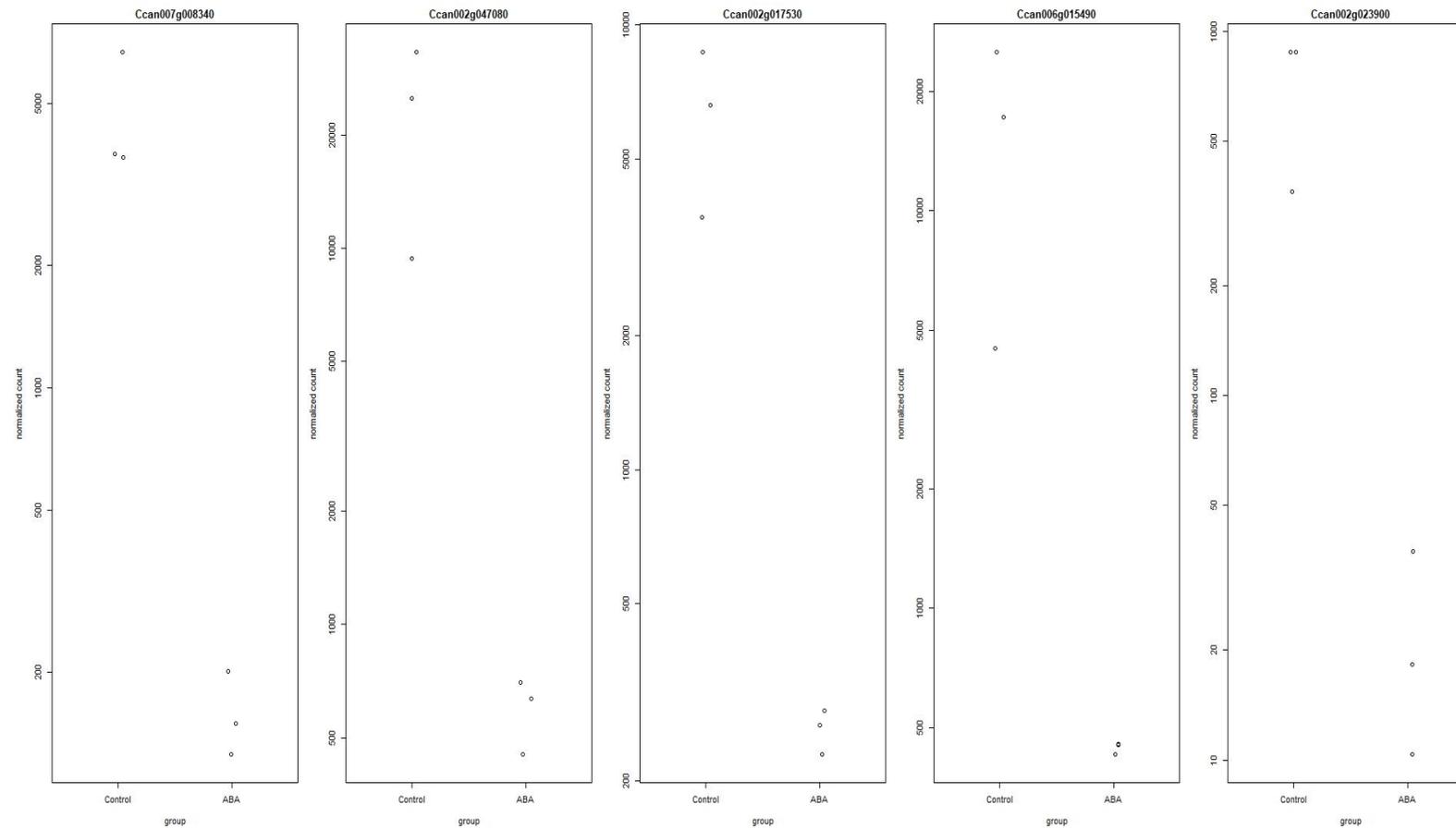


Fonte: Do autor

Os dois genes mais expressos em *canephora* (Ccan005g023910, Ccan009g016830) são responsáveis por expressão de proteínas ligantes a clorofila a-b associadas à resposta ao estresse oxidativo (Nott *et al.*, 2006; Staneloni *et al.*, 2008) e o fitormônio ABA (Bartholomew *et al.*, 1991; Chang and Walling, 1991; Weatherwax *et al.*, 1996; Staneloni *et al.*, 2008). Em estudos anteriores, foi apresentado que o ABA exógeno regula genes relacionados às proteínas ligadas à clorofila foto sequestrante (LHCB) de maneira negativa de maneira contrastante com o obtido nesse estudo, como pode ser visto nos trabalhos com aplicação exógena em folhas de tomate de Bartholomew *et al.*, 1991 e células de *Lemna gibba* crescidas em meio líquido (Weatherwax *et al.*, 1996). Entretanto, sementes de soja em fase de desenvolvimento com uma pequena dose do hormônio, obtiveram um aumento do mRNA ligado ao LHCB (Voigt *et al.*, 2010). Além das proteínas LHCB, proteínas similares não caracterizadas e coativadores transcricionais como a proteína isoforma LNK2-like X4, sendo essa relacionada aos genes ligados ao ciclo circadiano da planta.

Adicionalmente, foram avaliados os cinco genes mais reprimidos comparados com o controle da mesma espécie, esses são: Ccan007g008340, Ccan002g047080, Ccan002g017530, Ccan006g015490, Ccan002g023900. Os dois primeiros genes estão associados a proteínas responsáveis pela mediação de estresse na planta, como a proteína ligante ao cálcio CML45 e a invertase alcalina/neutra A, sendo que a primeira, está ligada à uma cascata de fosforilação responsável pela regulação genica (Mahajan *et al.*, 2005)

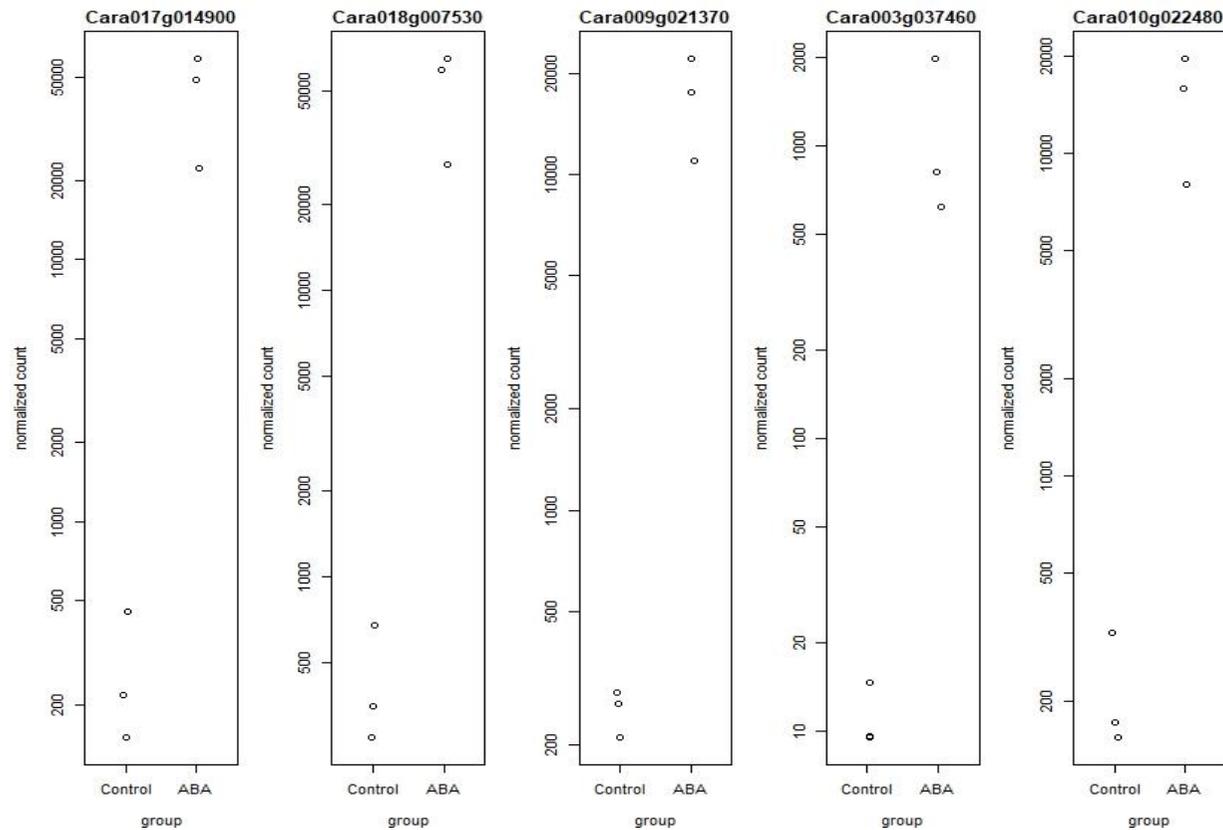
Figura 13 - 5 genes mais reprimidos na condição de estresse induzido pela aplicação de ABA externo em *Coffea canephora*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Os genes são os Ccan007g008340, Ccan002g047080, Ccan002g017530, Ccan006g015490, Ccan002g023900



Fonte: Do autor

O resultado das análises com a utilização do genoma de referência de *C. arabica*, indicam que os genes mais expressos estão apresentados na figura 14.

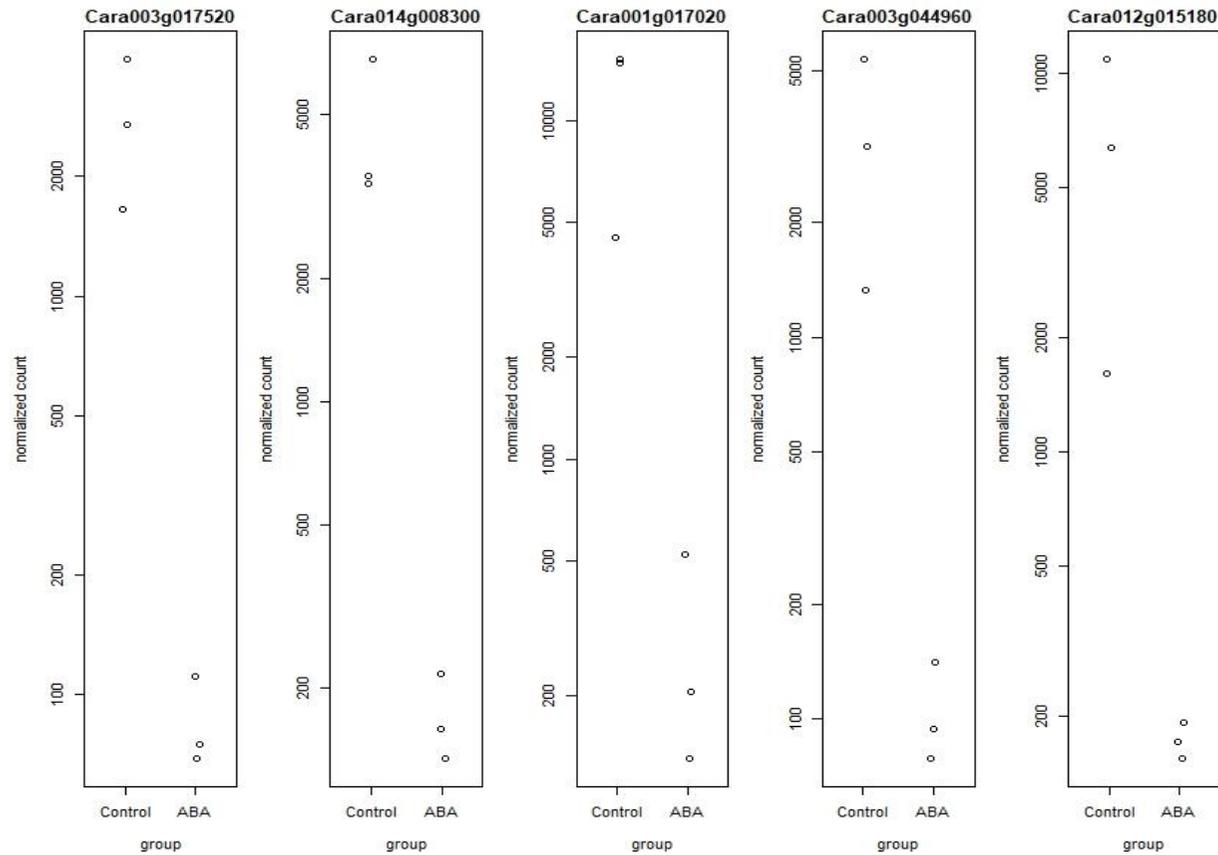
Figura 14 - 5 genes mais expressos na condição de estresse induzido pela aplicação de ABA externo em *Coffea arabica*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente menor. Esses são Cara017g014900, Cara018g007530, Cara009g021370, Cara003g037460, Cara010g022480.



Fonte: Do autor

Já os mais reprimidos foram os seguintes quando considerando a espécie *arabica*, Cara003g017520, Cara014g008300, Cara001g017020, Cara003g044960, Cara012g015180. (Figura 15).

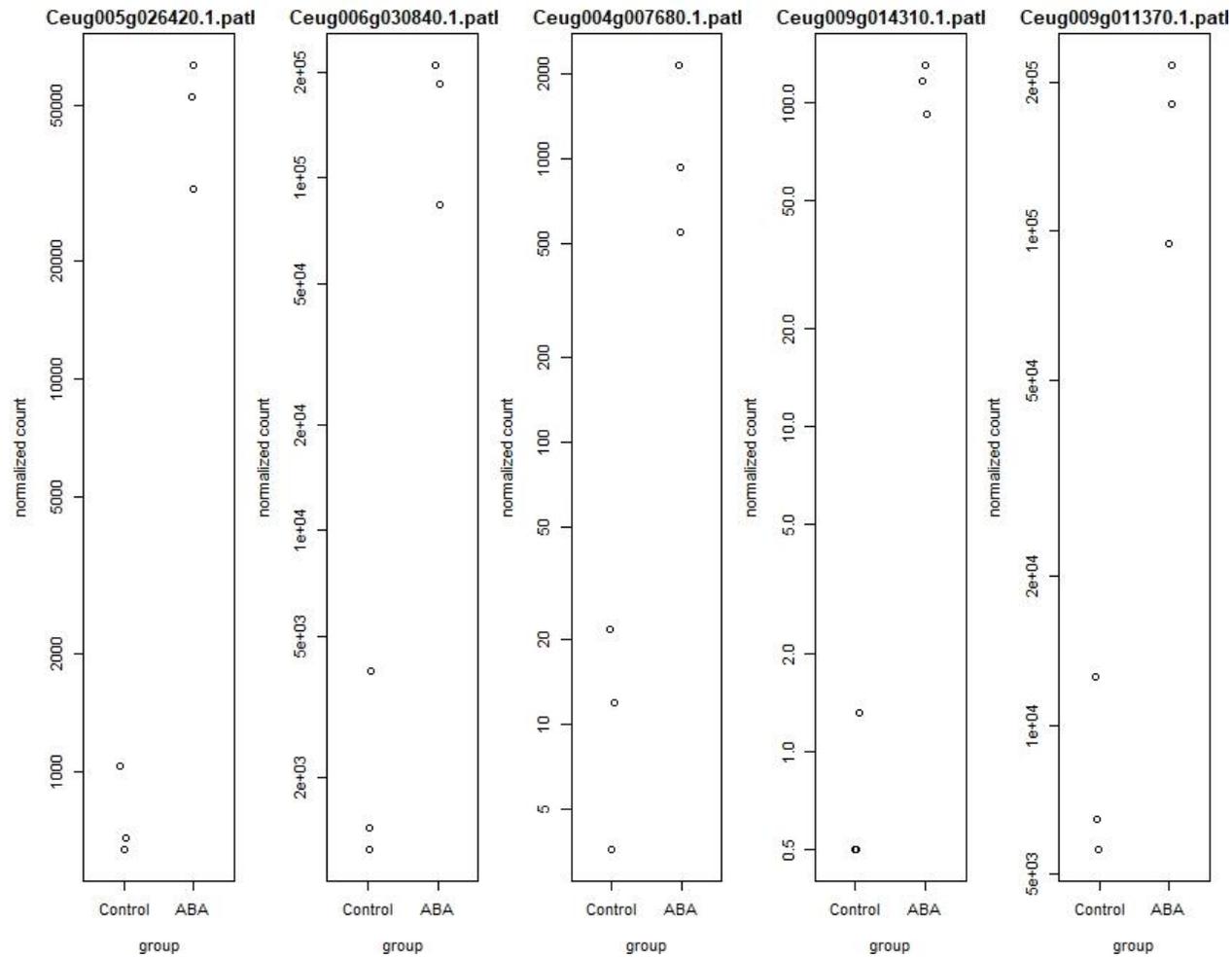
Figura 15 - 5 genes mais reprimidos na condição de estresse induzido pela aplicação de ABA externo em *Coffea arabica*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são Cara003g017520, Cara014g008300, Cara001g017020, Cara003g044960, Cara012g015180.



Fonte: Do autor.

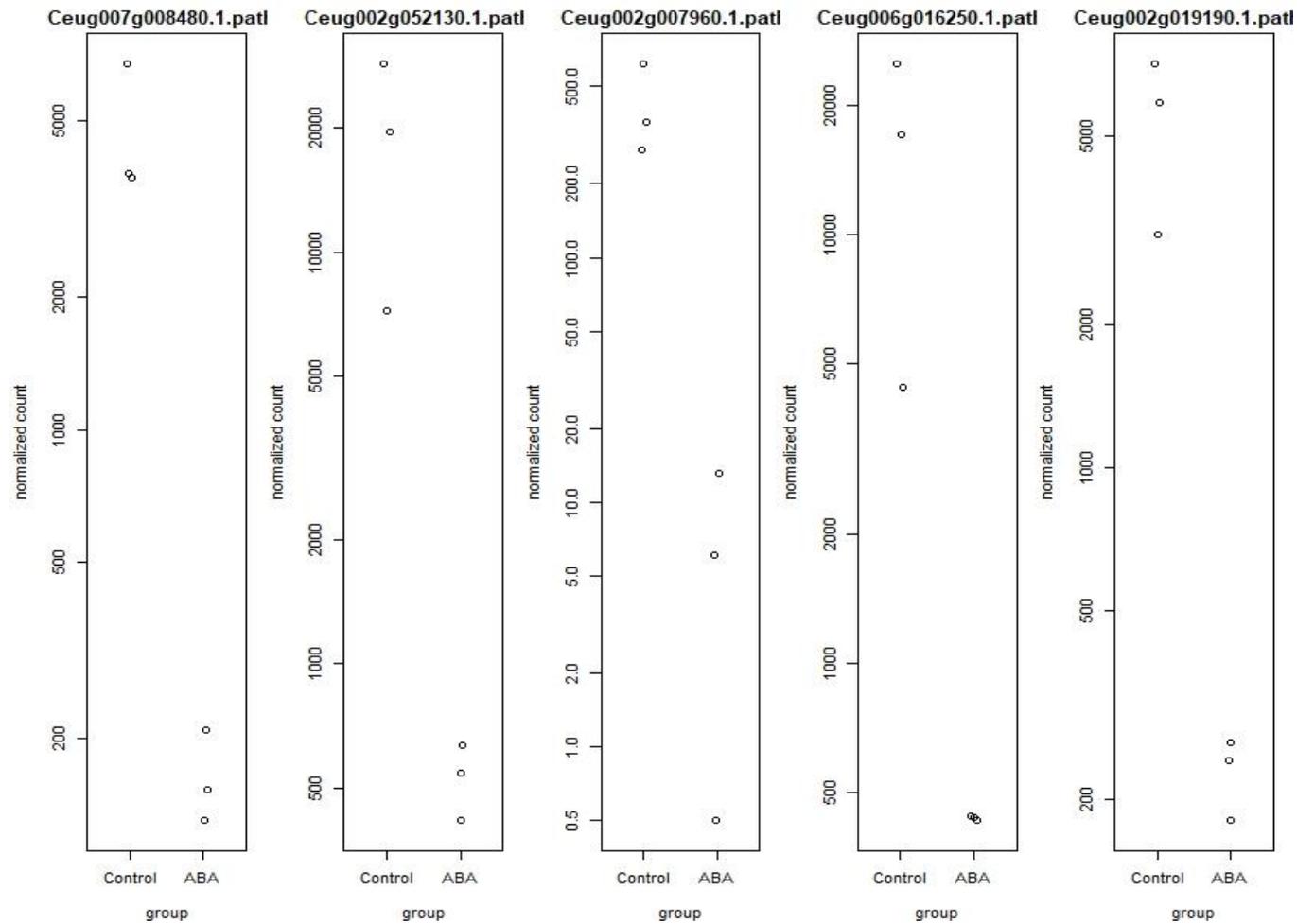
E por fim, quando a análise foi realizada com a referência do genoma de *C. eugenioides*, os 5 genes mais expressos foram Ceug005g026420.1.path1, Ceug006g030840.1.path1, Ceug004g007680.1.path1, Ceug009g014310.1.path1, Ceug009g011370.1.path1 (figura 16) e os 5 mais reprimidos da espécie *eugenioides* que são Ceug007g008480.1.path1, Ceug002g052130.1.path1, Ceug002g007960.1.path1, Ceug006g016250.1.path1, Ceug002g019190.1.path1 (figura 17).

Figura 16 - 5 genes mais expressos na condição de estresse induzido pela aplicação de ABA externo em *Coffea eugenioides*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são os Ceug005g026420.1.pat1, Ceug006g030840.1.pat1, Ceug004g007680.1.pat1, Ceug009g014310.1.pat1, Ceug009g011370.1.pat1



Fonte: Do autor.

Figura 17 - 5 genes mais reprimidos na condição de estresse induzido pela aplicação de ABA externo em *Coffea eugenioides*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Ceug007g008480.1.path1, Ceug002g052130.1.path1, Ceug002g007960.1.path1, Ceug006g016250.1.path1, Ceug002g019190.1.path1.

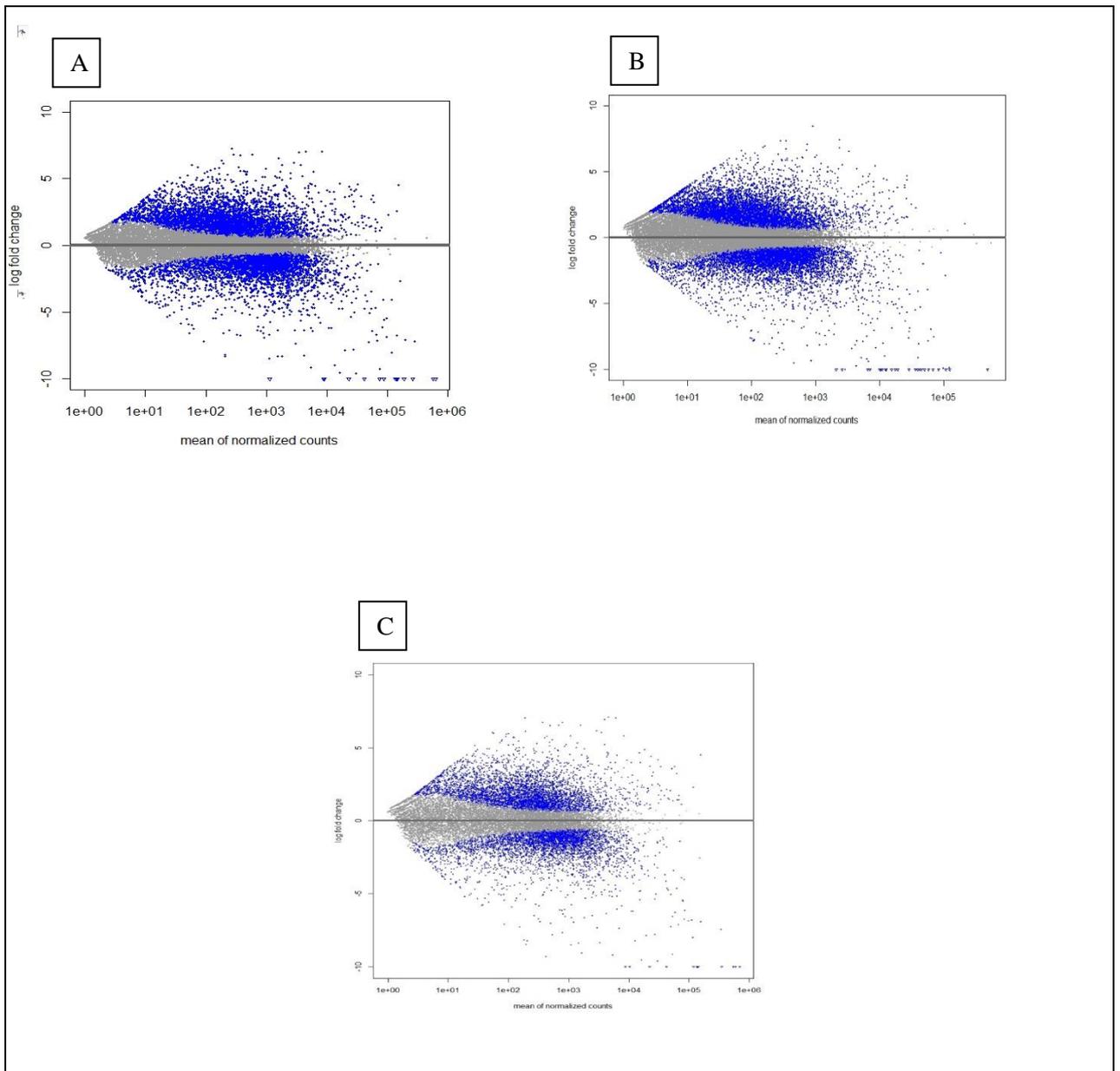


Fonte: Do autor

5.2 Padrões de genes diferencialmente expressos resultantes das 3 análises realizadas com diferentes genomas de referência (*C. arabica*, *C. canephora* e *C. eugenoides*), responsivos ao estresse causado pelo calor

Na figura 20, estão apresentados os gráficos de dispersão para as 3 análises realizadas (Referência: *C.arabica*, *C.canephora* e *C.eugenoides*) gerados ao final da comparação do tratamento de calor e a condição controle. É possível notar padrões de expressão diferencial, com genes mais expressos (*up_regulated*) e com maior repressão (*down_regulated*). Nota-se claramente que os dados apresentados na figura 18, indicam um maior número de genes diferencialmente expressos, independente do genoma de referência utilizado na análise, responsivos ao estresse de calor, quando comparados com os dados apresentados na figura 11 (responsivos ao tratamento de ABA. Pode-se observar também, pelos dados apresentados na figura 18, uma pequena redução no número de genes diferencialmente expressos quando a análise foi realizada utilizando-se o genoma de referência de *C. eugenoides*.

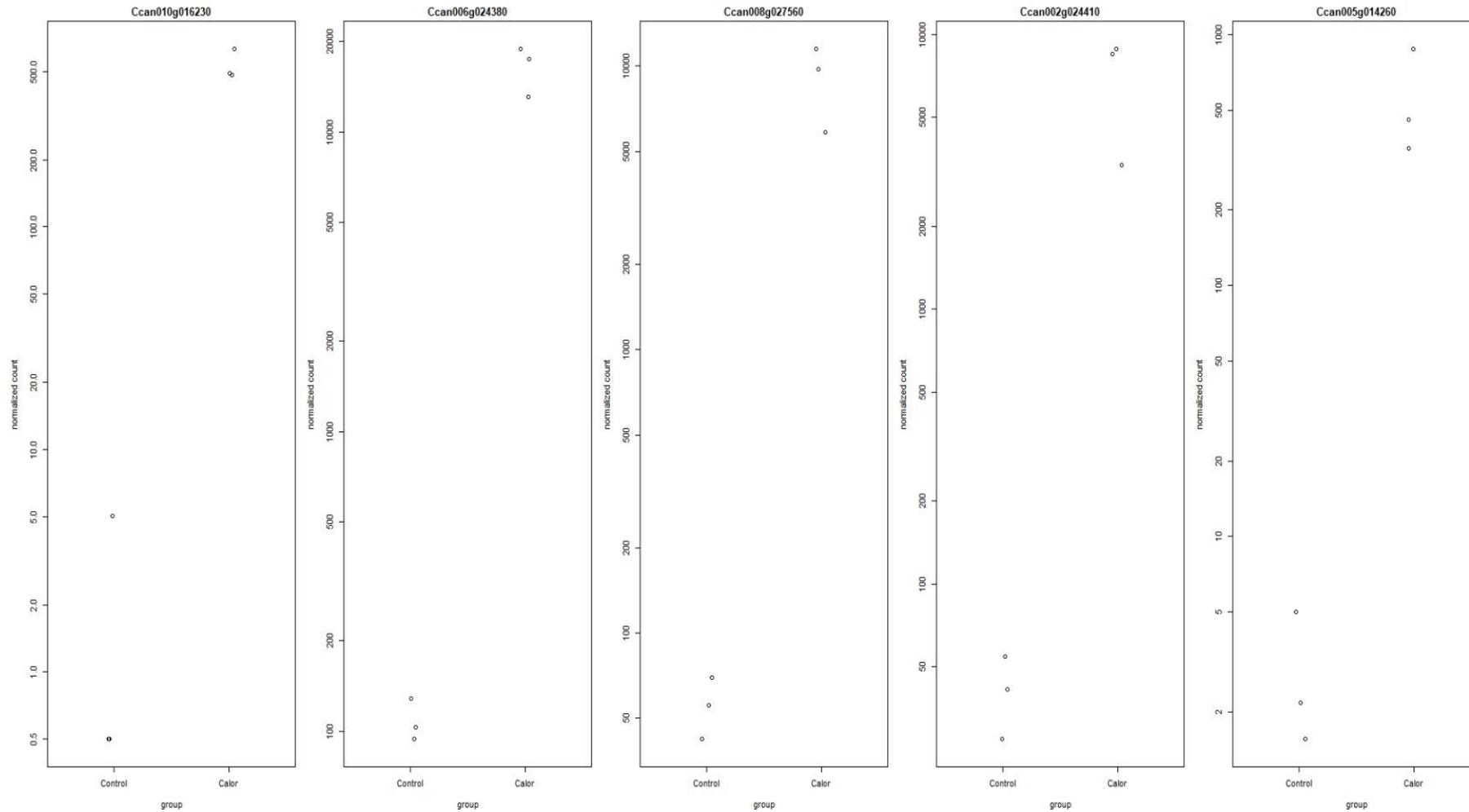
Figura 18 - Gráficos de dispersão contendo os genes de *Coffea new canephora* (A), *Coffea arabica* (B), *Coffea eugenoides* (C) com expressão diferencial significativa ($p < 0.05$; pontos em azul) quando expostos ao estresse térmico (calor). Pontos na parte superior do eixo X representam genes *up-regulated* e na parte inferior, genes *down-regulated*.



Fonte: Do autor.

Na figura 19 é possível observar que os genes mais expressos com a análise utilizando-se o genoma de referência de *C. canephora* foram: Ccan010g016230, Ccan006g024380, Ccan008g027560, Ccan002g024410, Ccan005g014260.

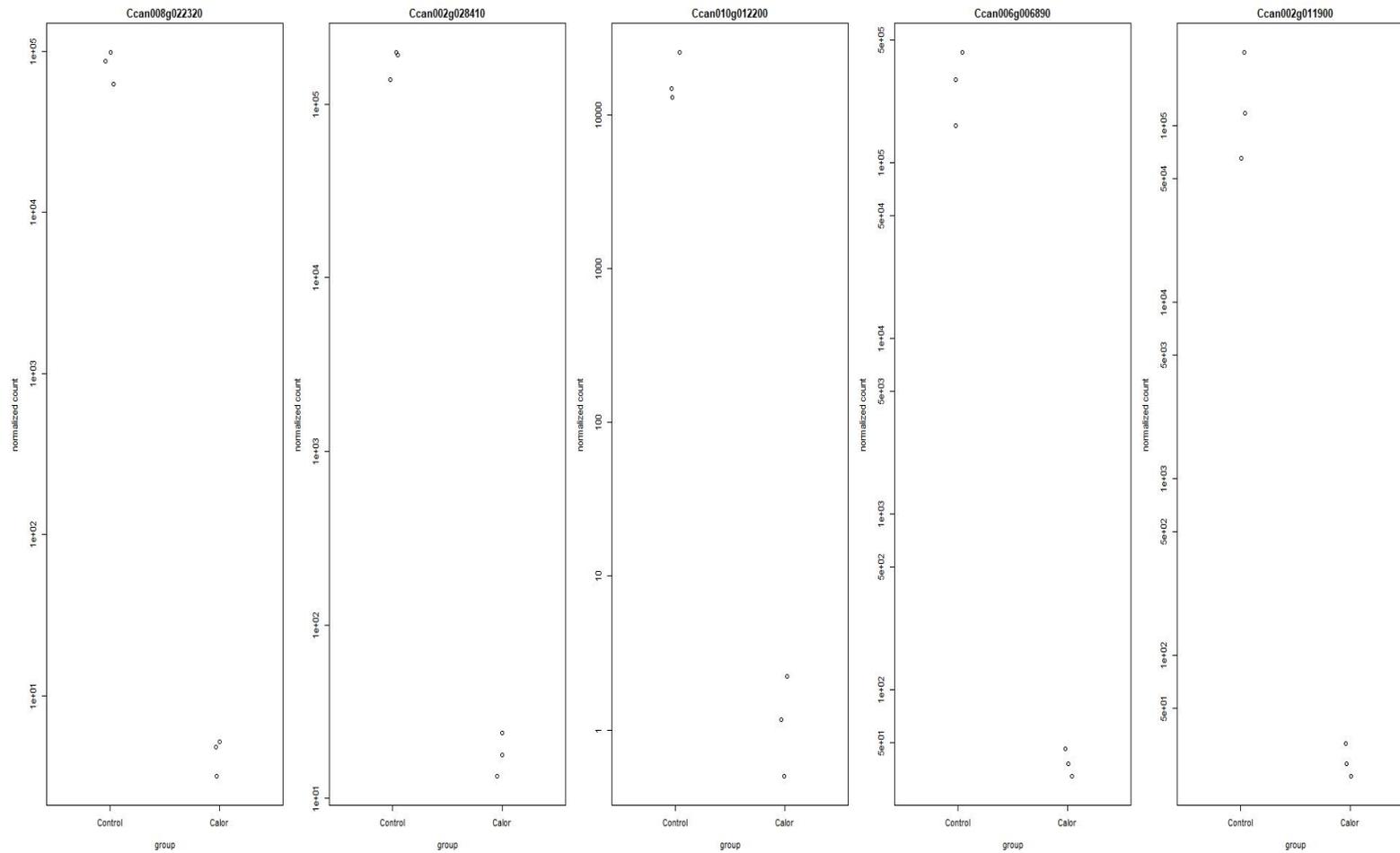
Figura 19 - 5 genes mais expressos na condição de estresse térmico induzido pelo calor. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente menor. Esses são os Ccan010g016230, Ccan006g024380, Ccan008g027560, Ccan002g024410, Ccan005g014260



Fonte: Do autor

Adicionalmente, foram avaliados os cinco genes mais reprimidos quando comparado às condições controle, esses foram: Ccan008g022320, Ccan002g028410, Ccan010g012200, Ccan006g006890, Ccan002g011900.

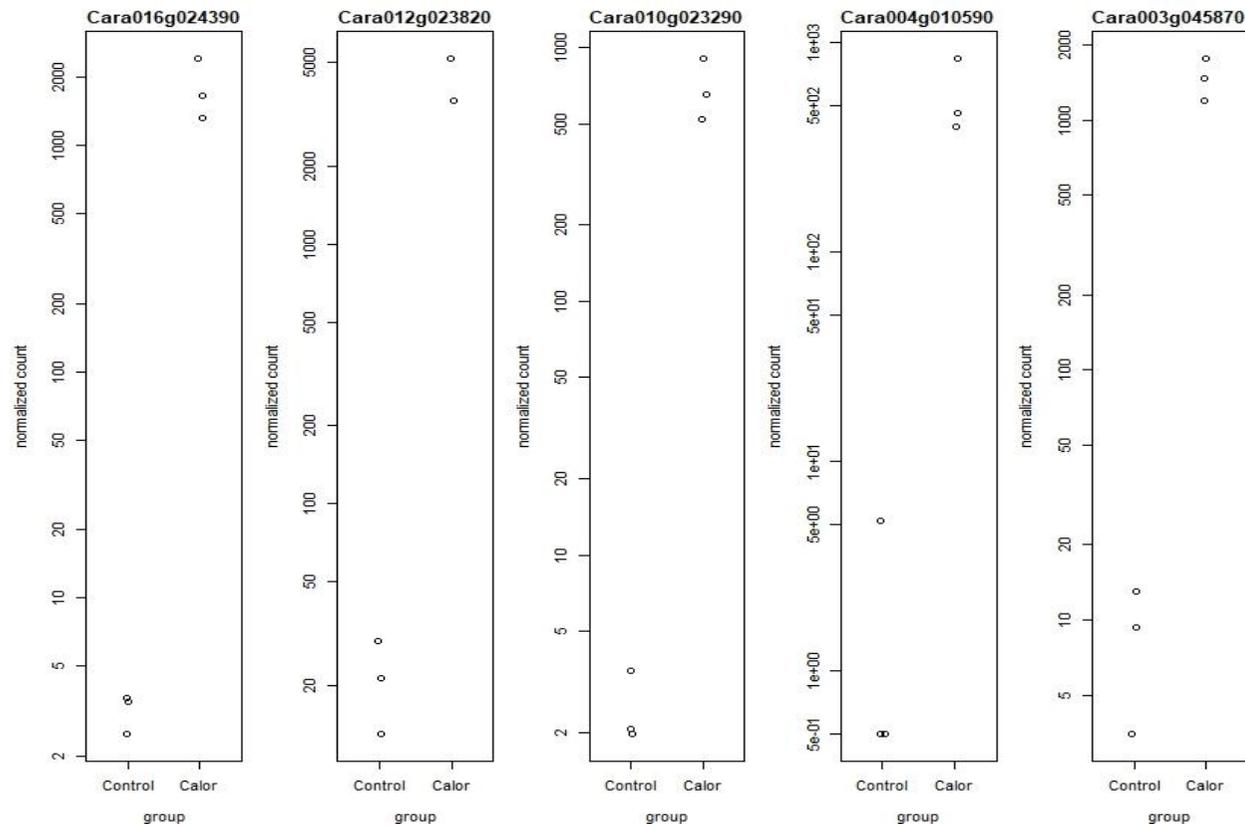
Figura 20 - 5 genes mais reprimidos na condição de estresse térmico induzido pelo calor em *Coffea canephora*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são os Ccan008g022320, Ccan002g028410, Ccan010g012200, Ccan006g006890, Ccan002g011900.



Fonte: Do autor

O resultado das análises com a utilização do genoma de referência de *C. arabica*, indicam que os genes mais expressos em foram os seguintes:

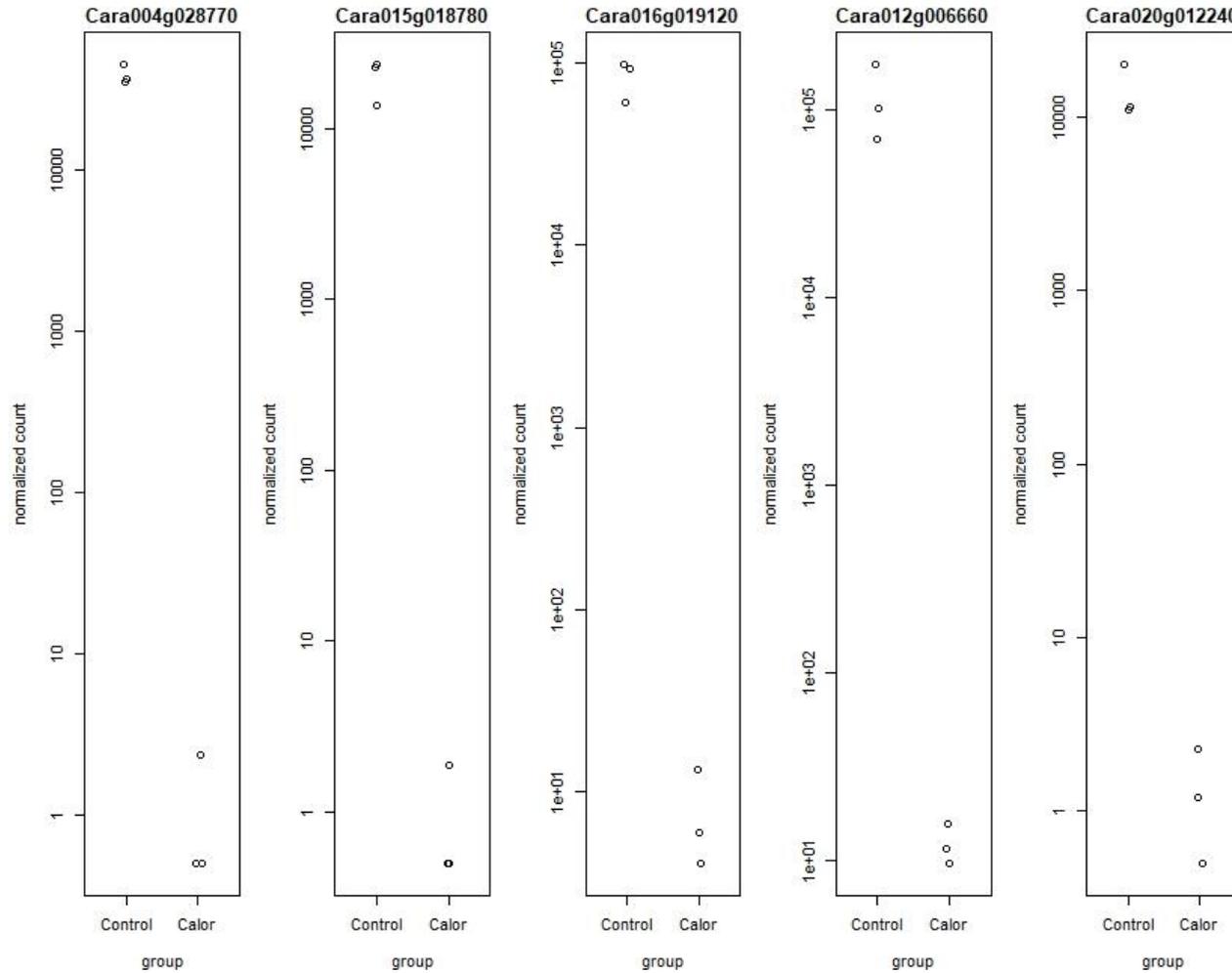
Figura 21 - 5 genes mais expressos na condição de estresse térmico induzido pelo calor em *Coffea arabica*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente menor. Esses são Cara016g024390, Cara012g023820, Cara010g023290, Cara004g010590, Cara003g045870.



Fonte: Do autor.

Já os mais reprimidos foram os seguintes quando considerando a espécie *arabica*, Cara004g028770, Cara015g018780, Cara016g019120, Cara012g006660, Cara020g012240.

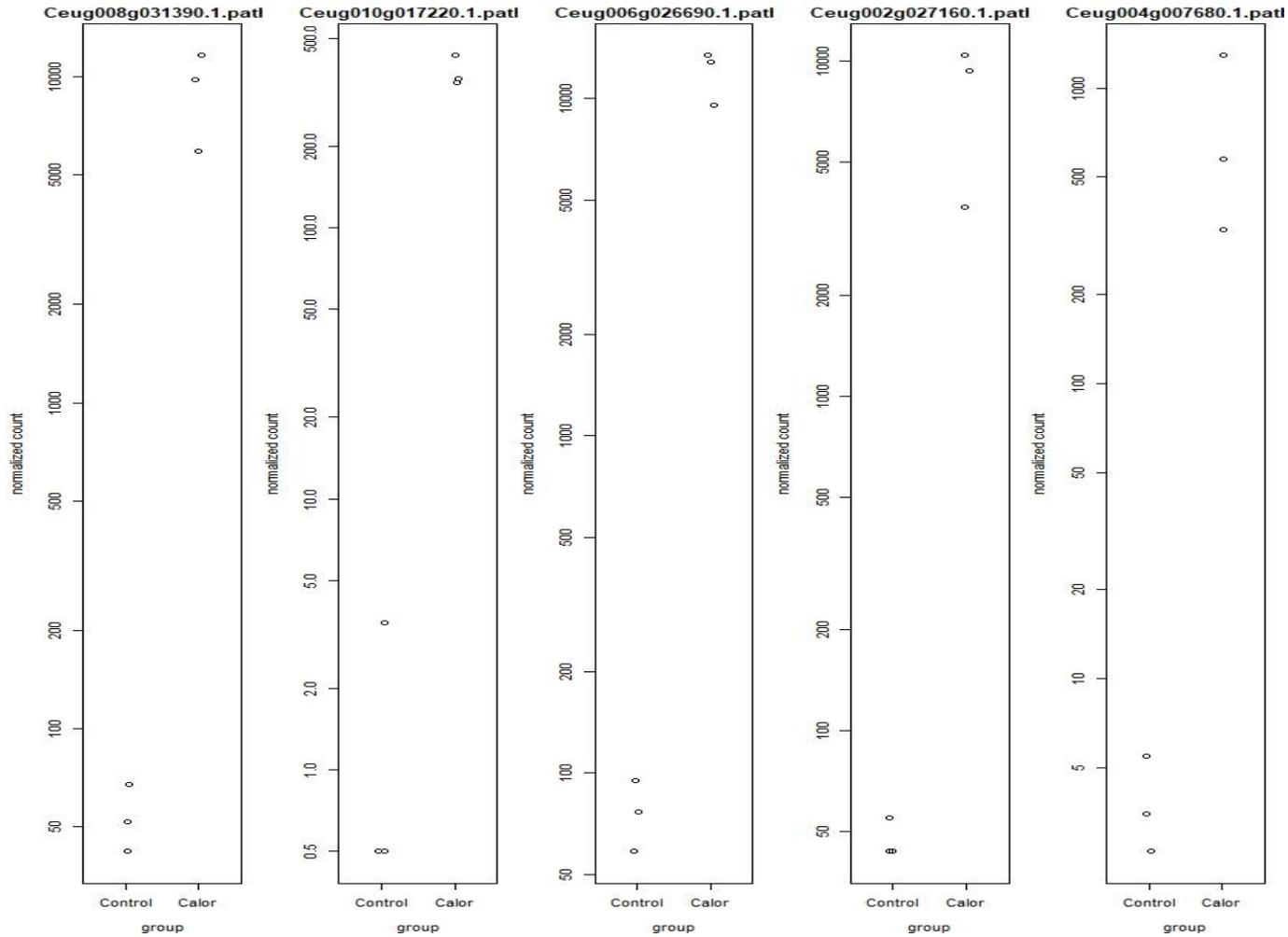
Figura 22 - 5 genes mais reprimidos na condição de estresse térmico induzido pelo calor em *Coffea arabica*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são Cara003g017520, Cara014g008300, Cara001g017020, Cara003g044960, Cara012g015180.



Fonte: Do autor

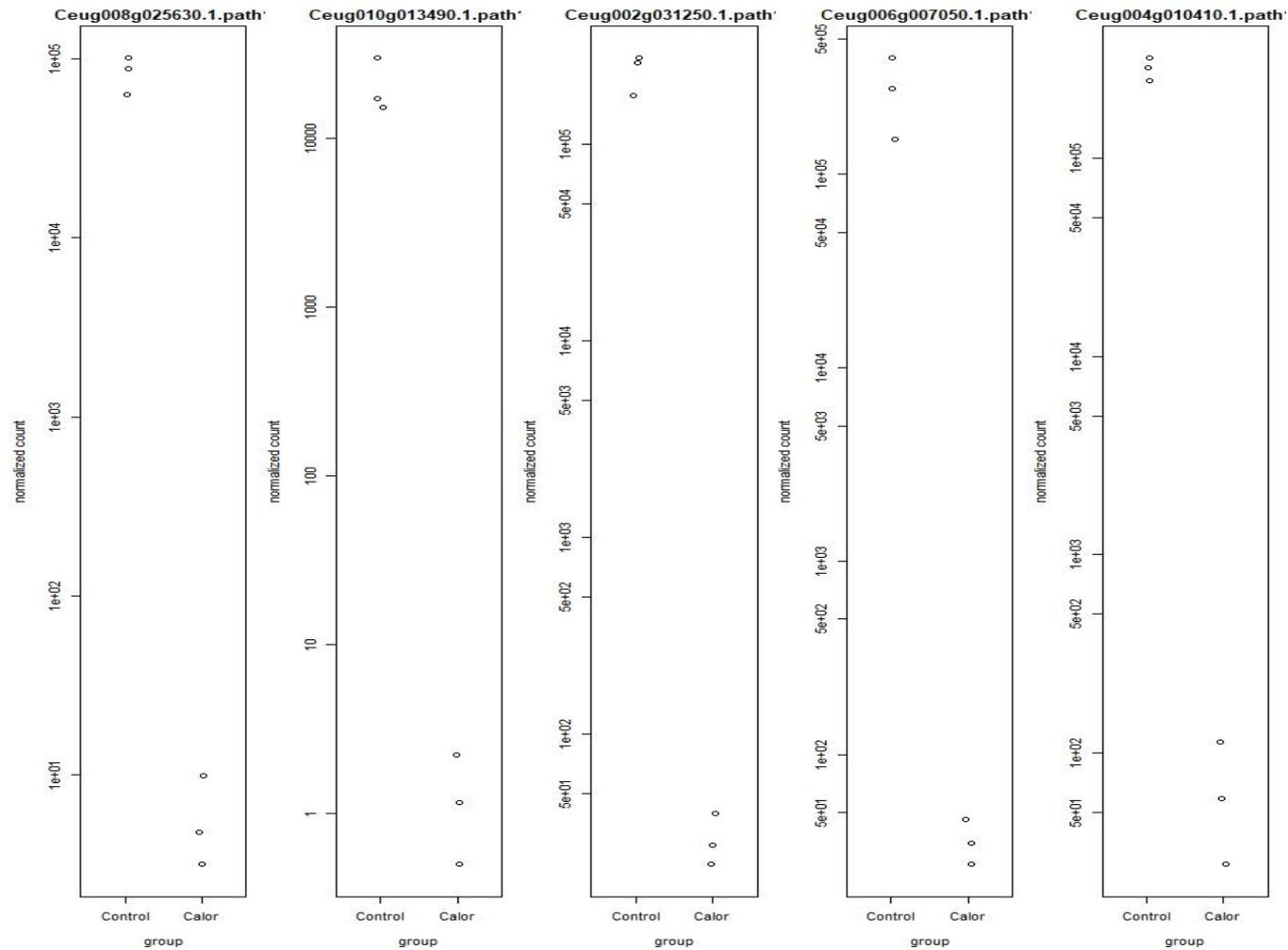
E por fim, quando a análise foi realizada com a referência do genoma de *C. eugenioides*, os 5 genes mais expressos foram Ceug008g031390.1.path1, Ceug010g017220.1.path1, Ceug006g026690.1.path1, Ceug002g027160.1.path1, Ceug004g007680.1.path1 (figura 23) e os 5 mais reprimidos (figura 24) da espécie *eugenioides* que são Ceug004g007680.1.path1, Ceug010g013490.1.path1, Ceug002g031250.1.path1, Ceug006g007050.1.path1, Ceug004g010410.1.path1.

Figura 23 - 5 genes mais expressos na condição de estresse térmico induzido pelo calor em *Coffea eugenioides*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são os Ceug008g031390.1.path1, Ceug010g017220.1.path1, Ceug006g026690.1.path1, Ceug002g027160.1.path1, Ceug004g007680.1.path1



Fonte: Do autor

Figura 24 - 5 genes mais reprimidos na condição de estresse térmico induzido pelo calor em *Coffea eugenioides*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Ceug008g025630.1.path1, Ceug010g013490.1.path1, Ceug002g031250.1.path1, Ceug006g007050.1.path1, Ceug004g010410.1.path1.

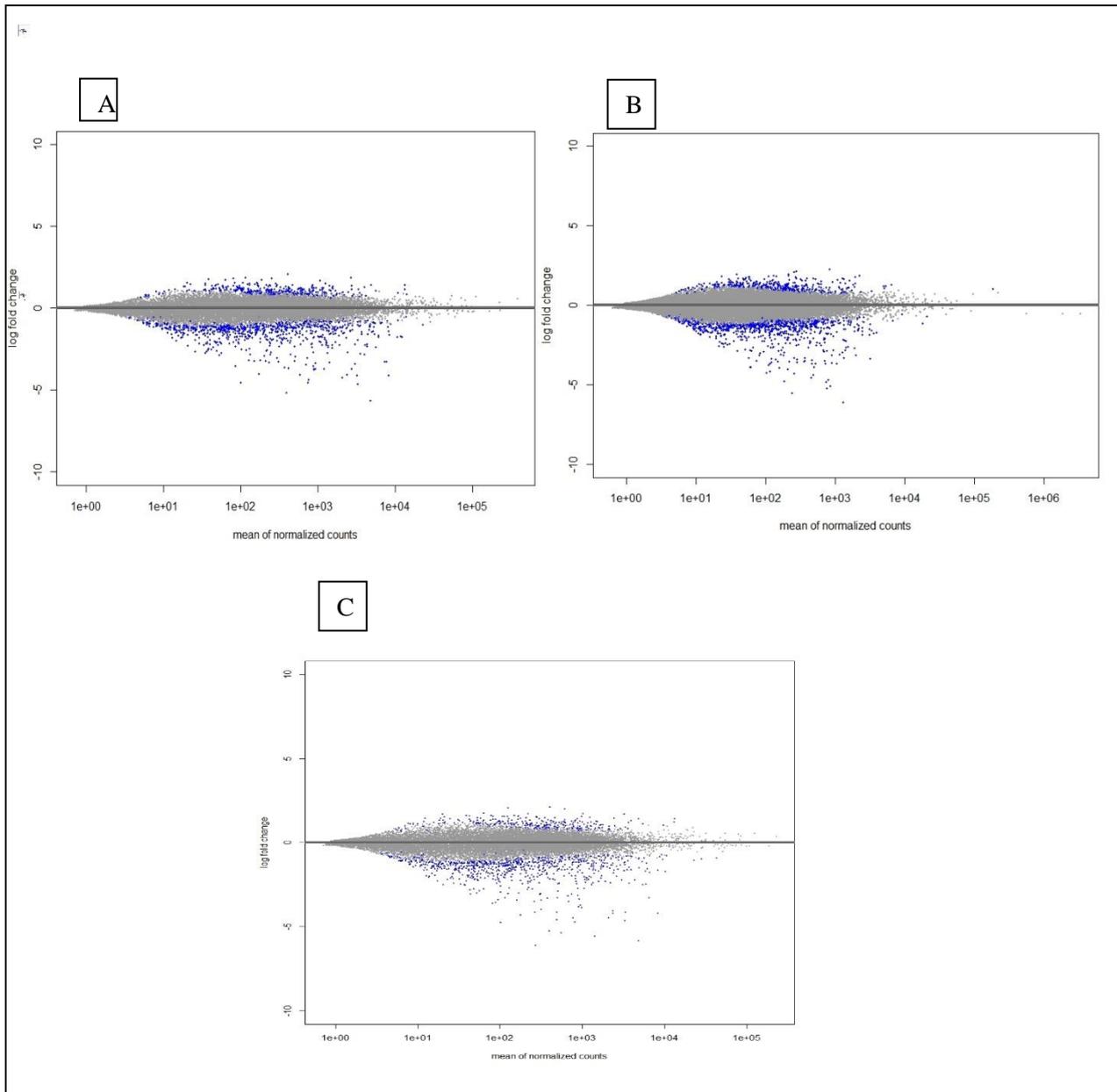


Fonte: Do autor

5.3 Padrões de genes diferencialmente expressos resultantes das 3 análises realizadas com diferentes genomas de referência (*C. arabica*, *C. canephora* e *C. eugenoides*), responsivos ao estresse causado pelo frio

Na figura 25, estão apresentados os gráficos de dispersão para as 3 análises realizadas (Referência: *C.arabica*, *C.canephora* e *C.eugenoides*) gerados ao final da comparação do tratamento de exposição ao frio e a condição controle. É possível notar padrões de expressão diferencial, com genes mais expressos (*up_regulated*) e com maior repressão (*down_regulated*). Como comentado anteriormente, nota-se claramente que os dados apresentados na figura 25, indicam um maior número de genes diferencialmente expressos, independente do genoma de referência utilizado na análise, responsivos ao estresse de frio, quando comparados com os tratamentos abordados anteriormente (ABA – figura 11; Calor – figura 18). Da mesma forma, pode-se observar também, pelos dados apresentados na figura 25, uma pequena redução no número de genes diferencialmente expressos quando a análise foi realizada utilizando-se o genoma de referência de *C. eugenoides*.

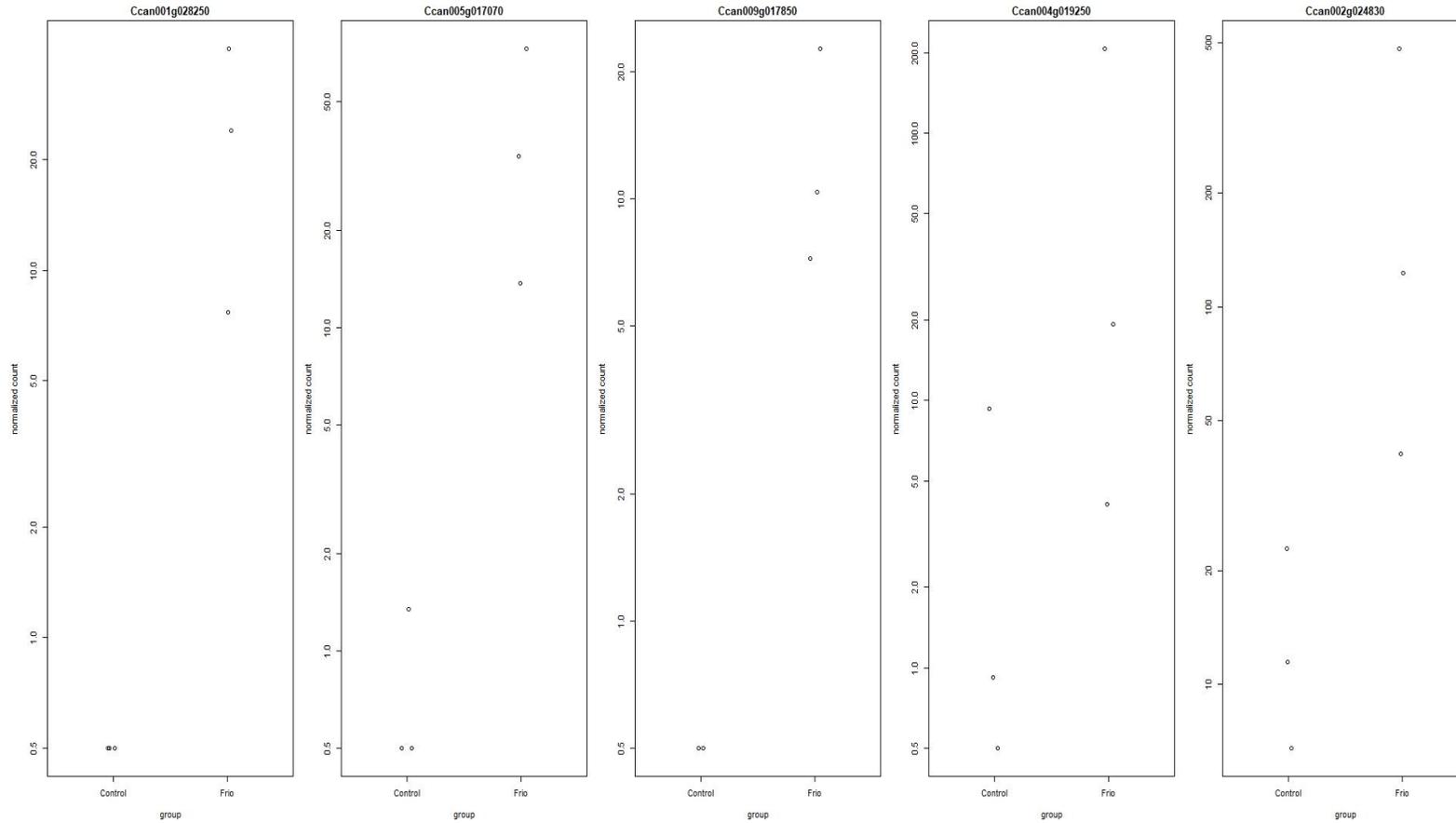
Figura 25 - Gráficos de dispersão contendo os genes de *Coffea new canephora* (A), *Coffea arabica* (B), *Coffea eugenoides* (C) com expressão diferencial significativa ($p < 0.05$; pontos em azul) quando expostos ao estresse térmico (frio). Pontos na parte superior do eixo X representam genes *up-regulated* e na parte inferior, genes *down-regulated*.



Fonte: Do autor.

Na figura 26 é possível observar que os genes mais expressos com a análise utilizando-se o genoma de referência de *C. canephora* foram os Ccan001g028250, Ccan005g017070, Ccan009g017850, Ccan004g019250, Ccan002g024830:

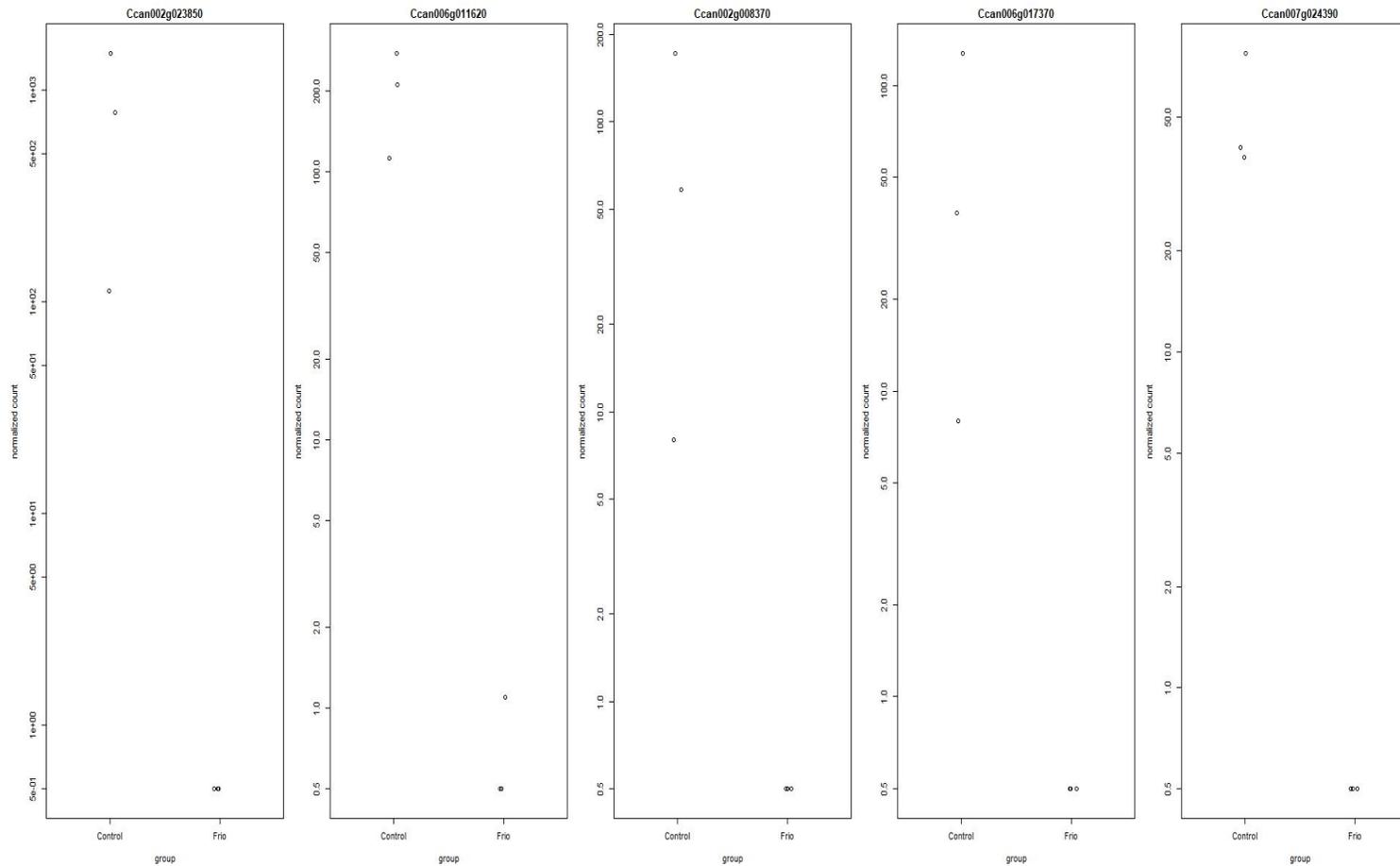
Figura 26 - 5 genes mais expressos na condição de estresse térmico induzido pelo frio. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente menor. Esses são os Ccan001g028250, Ccan005g017070, Ccan009g017850, Ccan004g019250, Ccan002g024830.



Fonte: Do autor.

Adicionalmente, foram avaliados os cinco genes mais reprimidos quando levando em consideração o controle da mesma espécie, esses foram: Ccan002g023850, Ccan006g011620, Ccan002g008370, Ccan006g017370, Ccan007g024390.

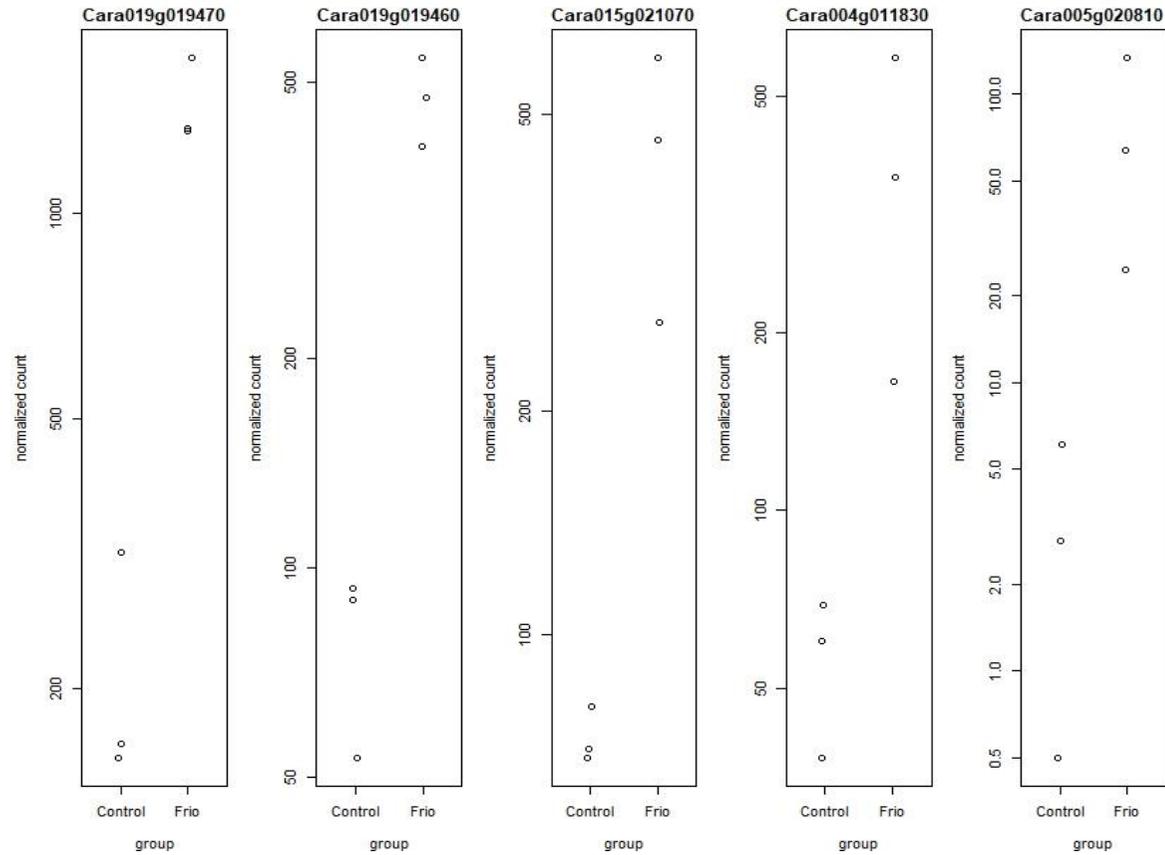
Figura 27 - 5 genes mais reprimidos na condição de estresse térmico induzido pelo frio em *Coffea canephora*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são os Ccan002g023850, Ccan006g011620, Ccan002g008370, Ccan006g017370, Ccan007g024390.



Fonte: Do autor

O resultado das análises com a utilização do genoma de referência de *C. arabica*, indicam que os genes mais expressos em foram os seguintes (figura 28):

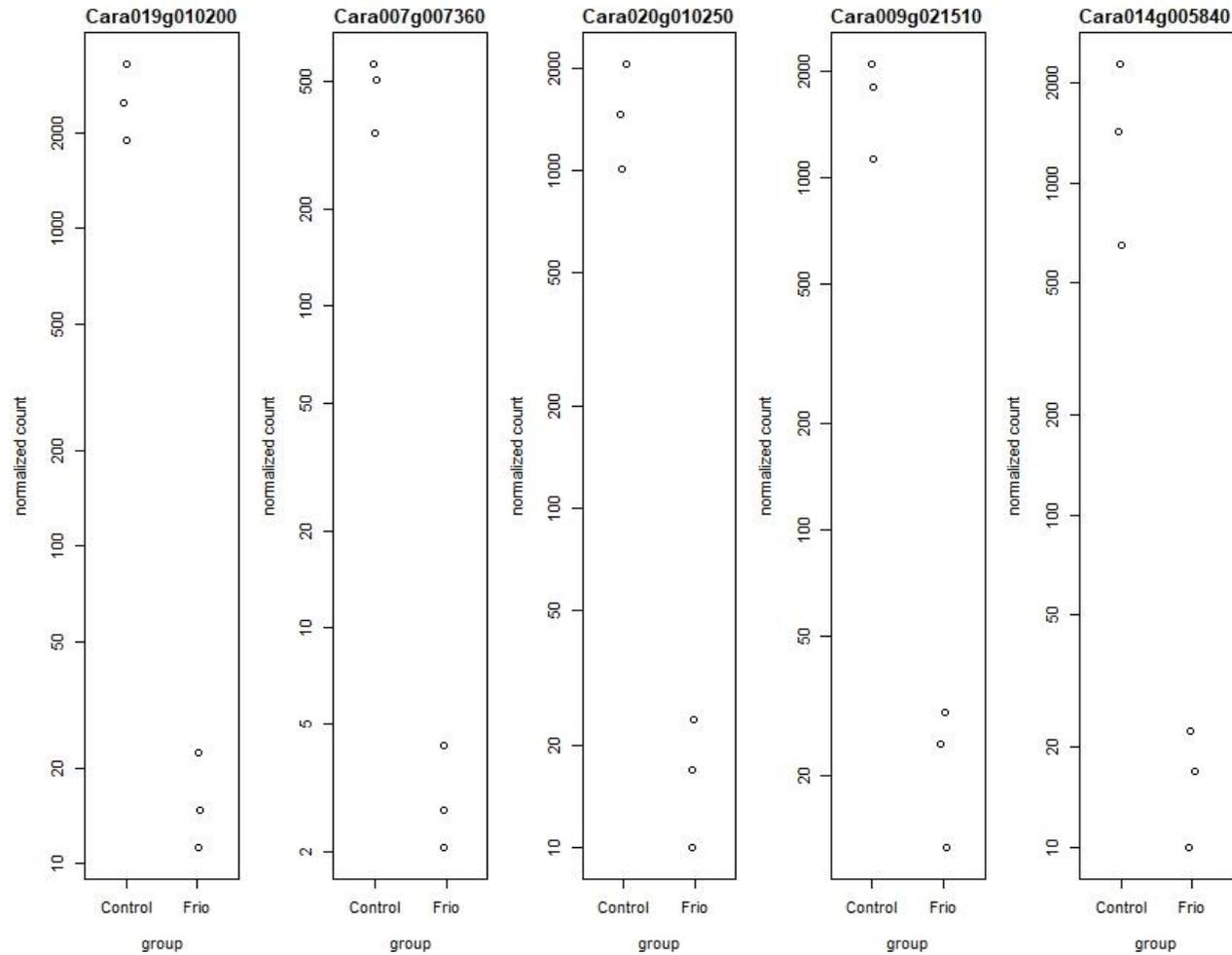
Figura 28 - 5 genes mais expressos na condição de estresse térmico induzido pelo frio em *Coffea arabica*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente menor. Esses são Cara019g019470, Cara019g019460, Cara015g021070, Cara004g011830, Cara005g020810



Fonte: Do autor

Já os mais reprimidos foram os seguintes quando considerando a espécie *arabica*, Cara019g010200, Cara007g007360, Cara020g010250, Cara009g021510, Cara014g005840 (figura 29).

Figura 29 - 5 genes mais reprimidos na condição de estresse térmico induzido pelo frio em *Coffea arabica*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são Cara019g010200, Cara007g007360, Cara020g010250, Cara009g021510, Cara014g005840.

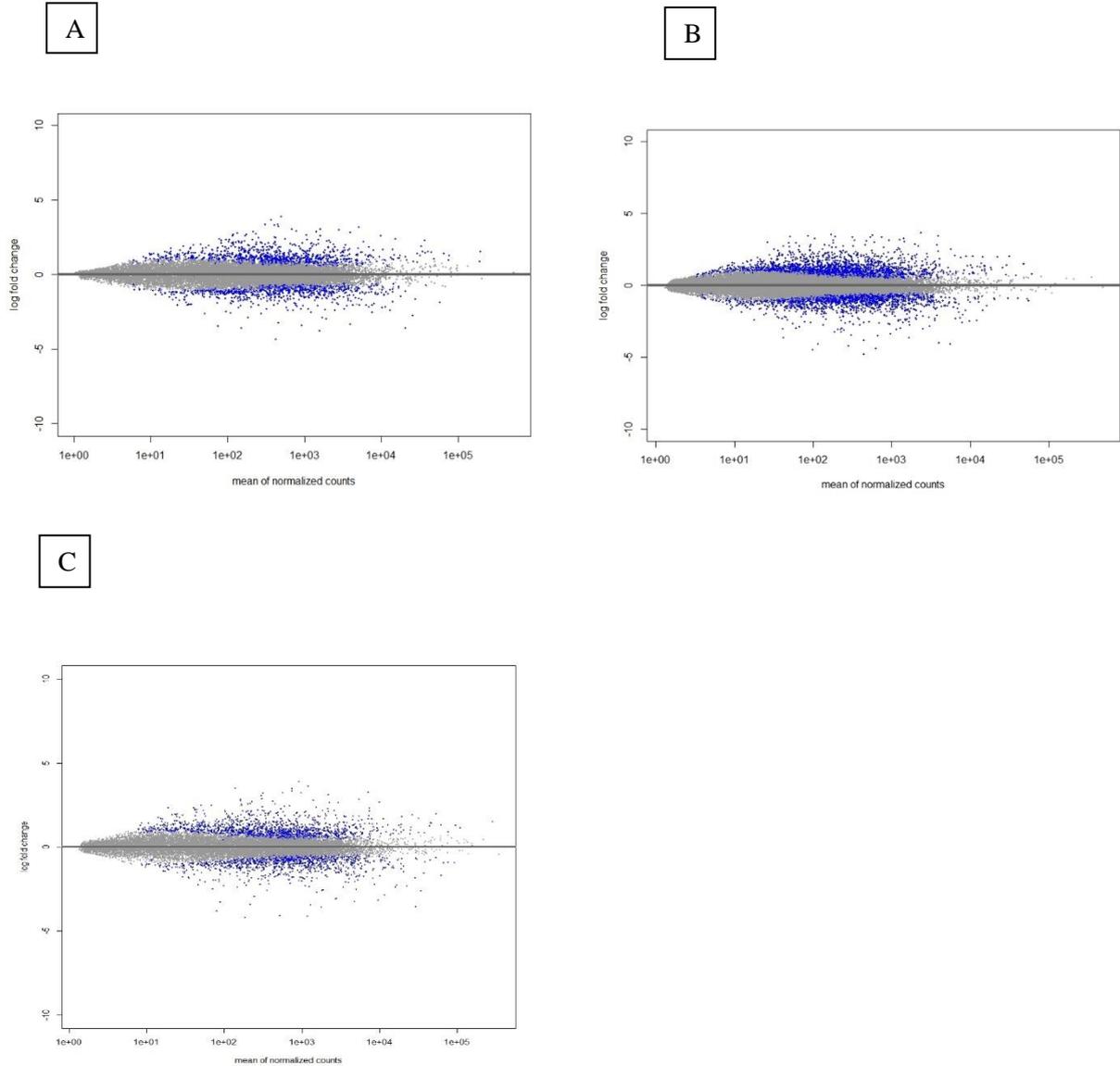


Fonte: Do autor

5.4. Padrões de genes diferencialmente expressos resultantes das 3 análises realizadas com diferentes genomas de referência (*C. arabica*, *C. canephora* e *C. eugenioides*), responsivos ao estresse luminoso

Na figura 30, estão apresentados os gráficos de dispersão para as 3 análises realizadas (Referência: *C.arabica*, *C.canephora* e *C.eugenioides*) gerados ao final da comparação do tratamento de exposição ao estresse luminoso e a condição controle. É possível notar padrões de expressão diferencial, com genes mais expressos (*up_regulated*) e com maior repressão (*down_regulated*). Nota-se, neste caso (figura 30) um número intermediário de genes diferencialmente expressos, independente do genoma de referência utilizado na análise, responsivos ao estresse luminoso, quando comparados com os tratamentos abordados anteriormente (ABA – figura 11; Calor – figura 18 e Frio – Figura 25). Da mesma forma, pode-se observar também, pelos dados apresentados na figura 30, uma pequena redução no número de genes diferencialmente expressos quando a análise foi realizada utilizando-se o genoma de referência de *C. eugenioides*.

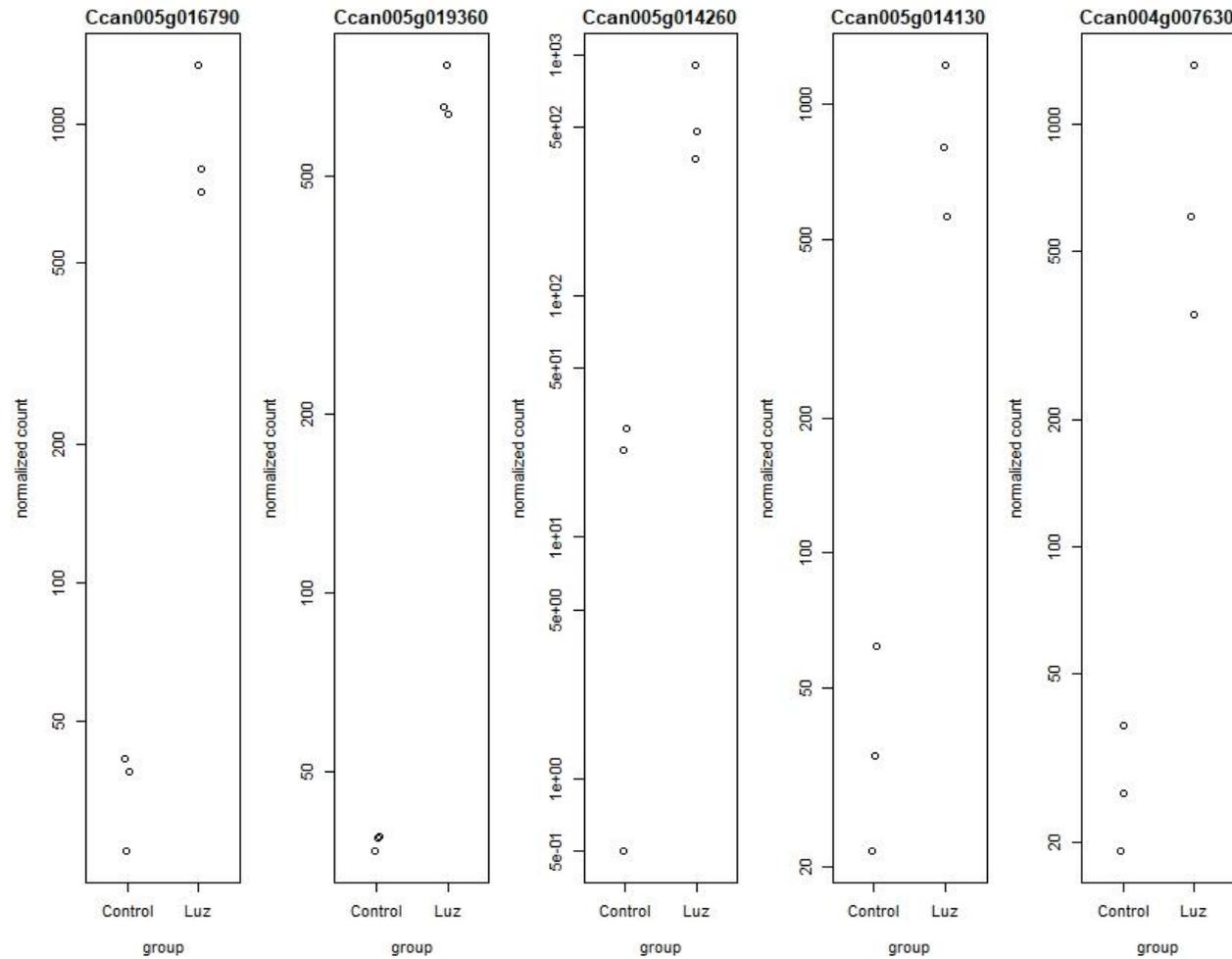
Figura 30 - Gráficos de dispersão contendo os genes de *Coffea new canephora* (A), *Coffea arabica* (B), *Coffea eugenioides* (C) com expressão diferencial significativa ($p < 0.05$; pontos em azul) quando expostos ao estresse luminoso. Pontos na parte superior do eixo X representam genes *up-regulated* e na parte inferior, genes *down-regulated*.



Fonte: Do autor

Na figura 31 é possível observar que os genes mais expressos com a análise utilizando-se o genoma de referência de *C. canephora*, foram: Ccan005g016790, can005g019360, Ccan005g014260, Ccan005g014130, Ccan004g007630:

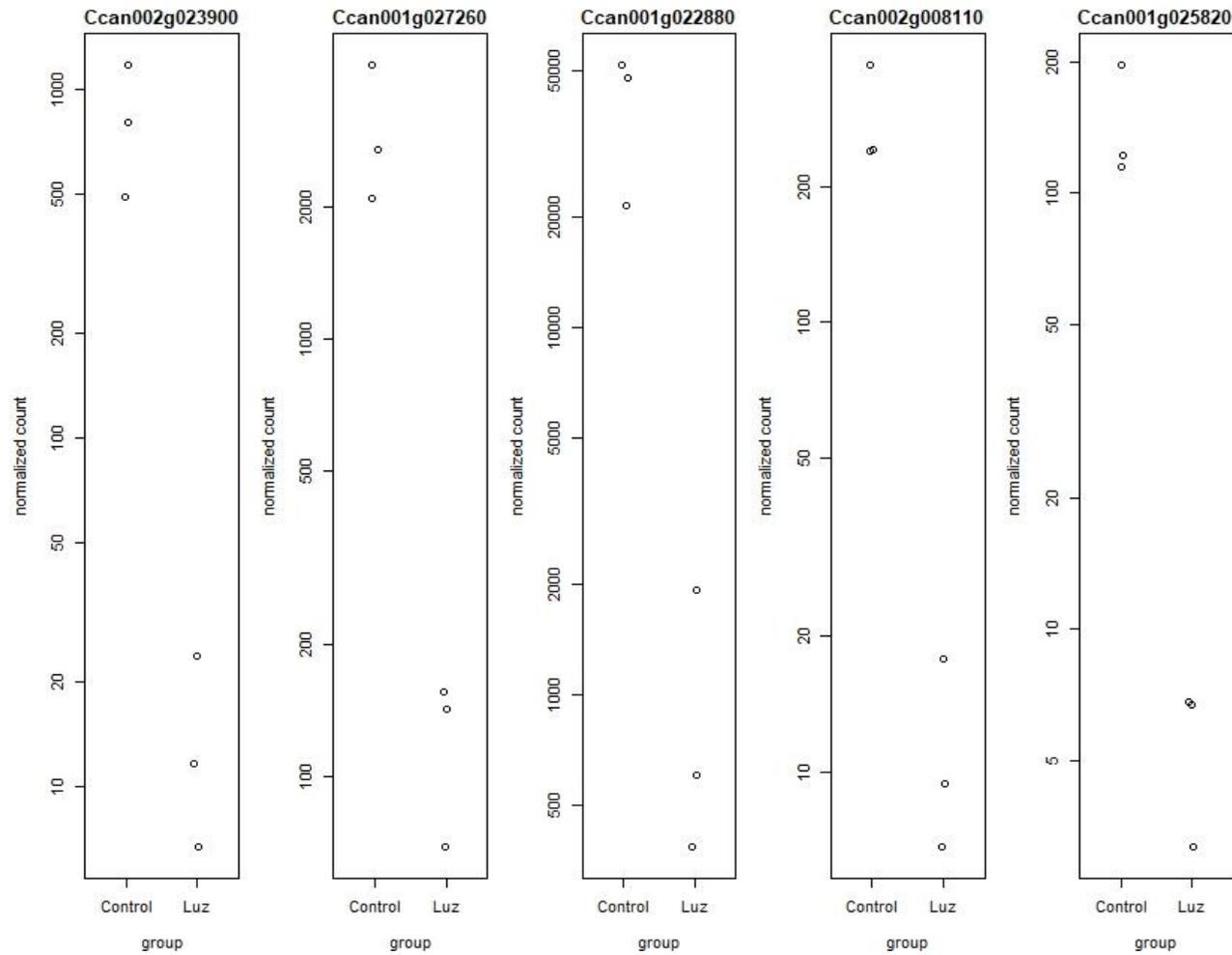
Figura 31 - 5 genes mais expressos na condição de estresse luminoso. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente menor. Esses são os Ccan005g016790, Ccan005g019360, Ccan005g014260, Ccan005g014130, Ccan004g007630.



Fonte: Do autor

Adicionalmente, foram avaliados os cinco genes mais reprimidos quando levando em consideração o controle da mesma espécie, esses foram: Ccan002g023900, Ccan001g027260, Ccan001g022880, Ccan002g008110, Ccan001g025820.

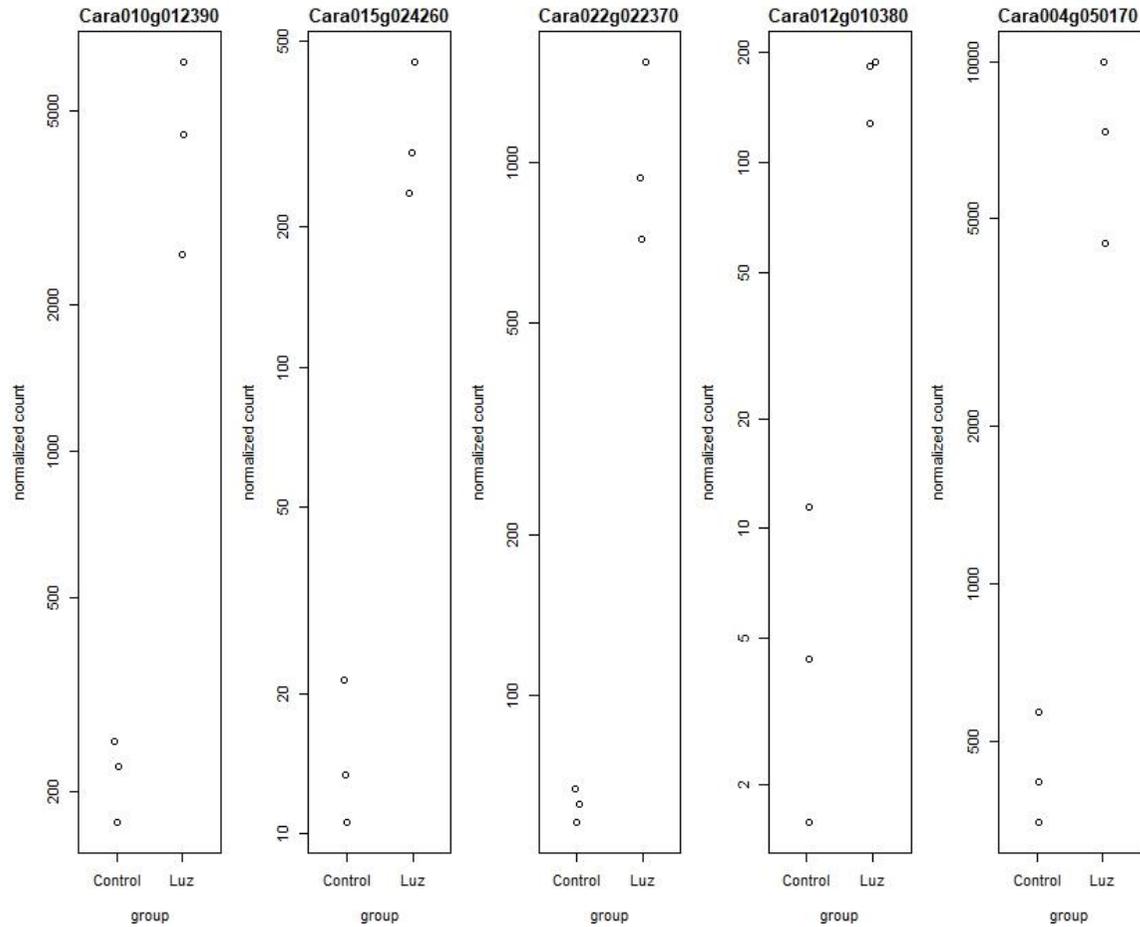
Figura 32 - 5 genes mais reprimidos na condição de estresse luminoso em *Coffea canephora*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são os Ccan002g023900, Ccan001g027260, Ccan001g022880, Ccan002g008110, Ccan001g025820.



Fonte: Do autor

O resultado das análises com a utilização do genoma de referência de *C. arabica*, indicam que os genes mais expressos foram os seguintes (figura 33):

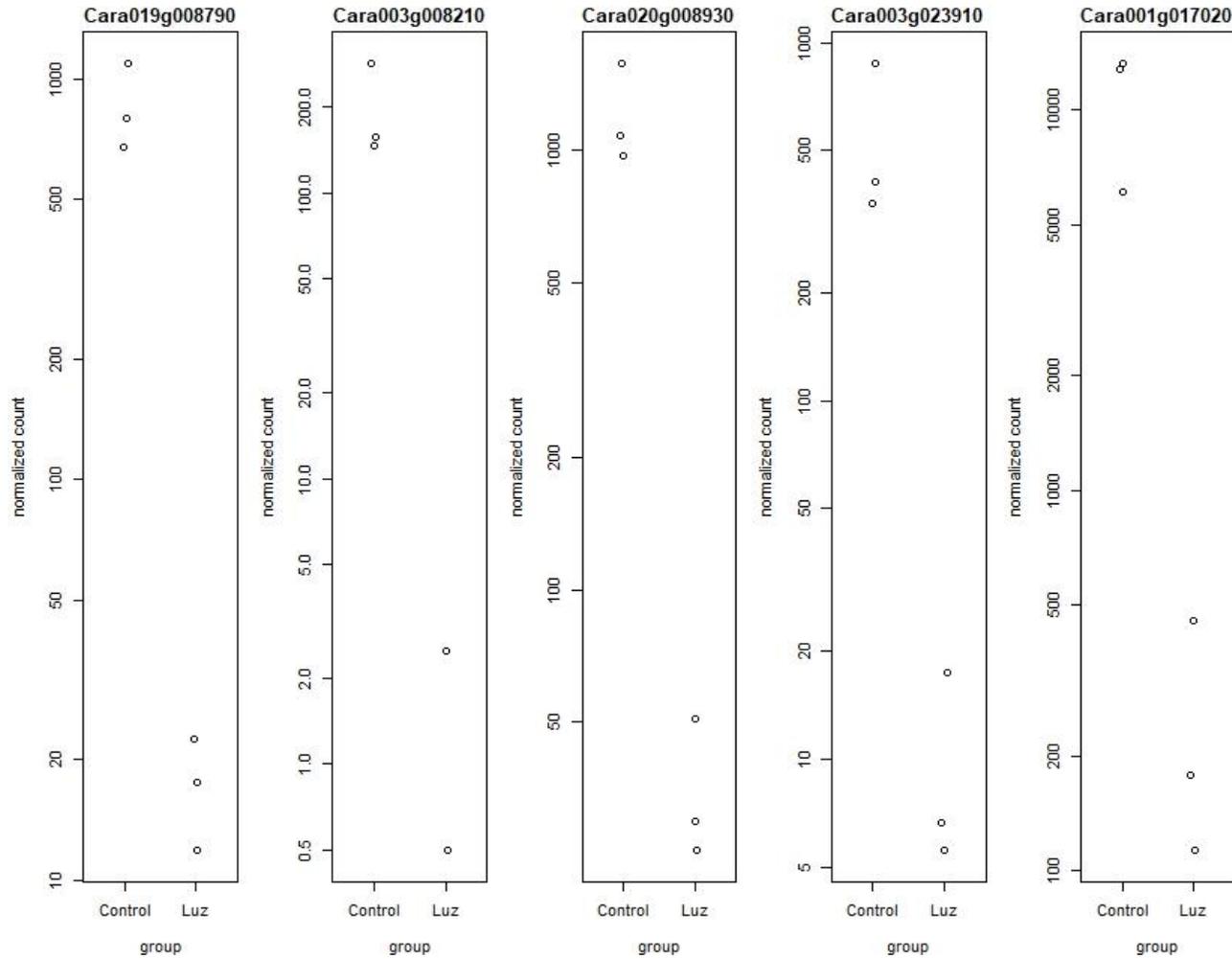
Figura 33 - 5 genes mais expressos na condição de estresse luminoso em *Coffea arabica*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente menor. Esses são Cara010g012390, Cara015g024260, Cara022g022370, Cara012g010380, Cara004g050170.



Fonte: Do autor.

Já os mais reprimidos foram Cara019g008790, Cara003g008210, Cara020g008930, Cara003g023910, Cara001g017020 (figura 34).

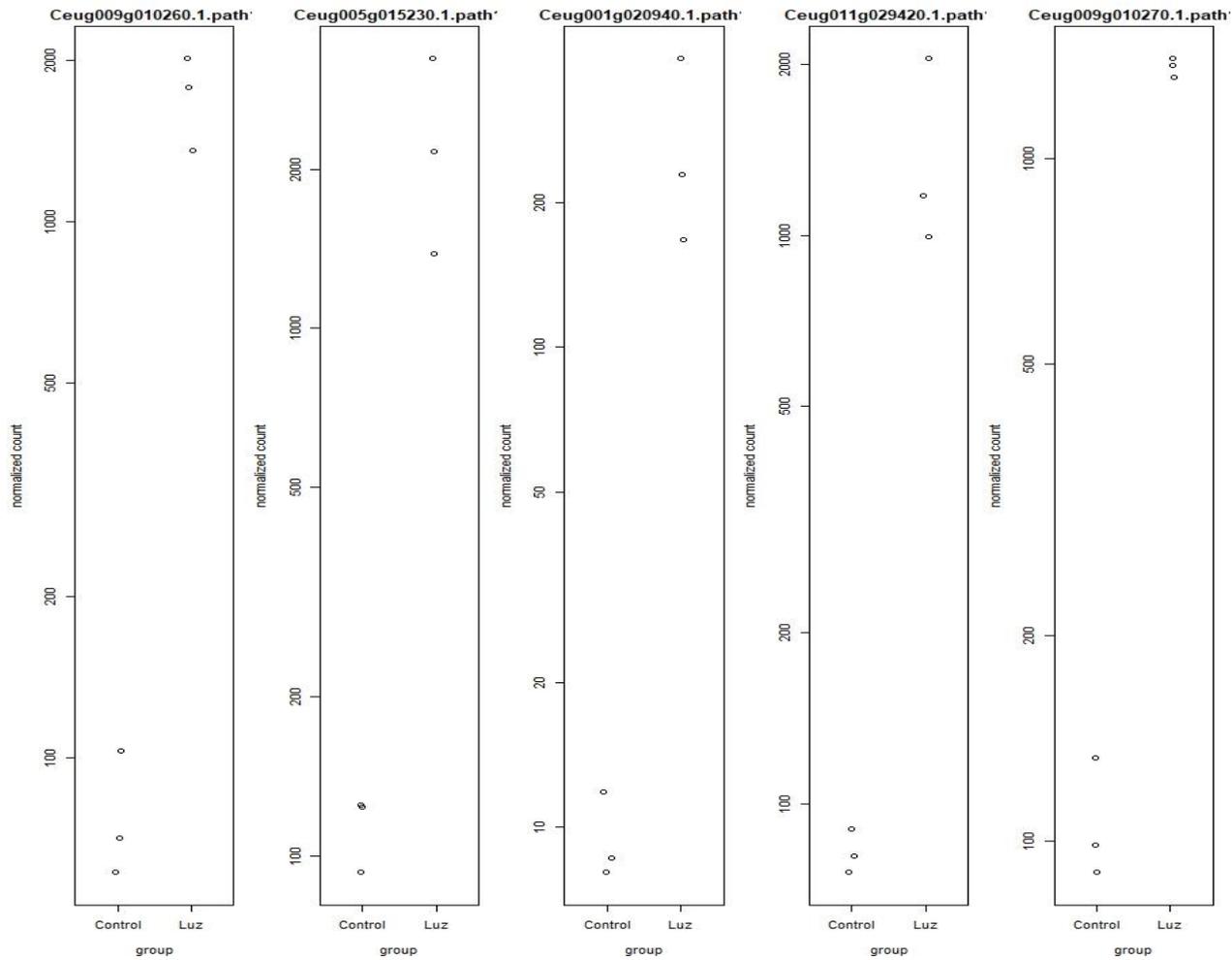
Figura 34 - 5 genes mais reprimidos na condição de estresse luminoso em *Coffea arabica*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são Cara019g008790, Cara003g008210, Cara020g008930, Cara003g023910, Cara001g017020.



Fonte: Do autor

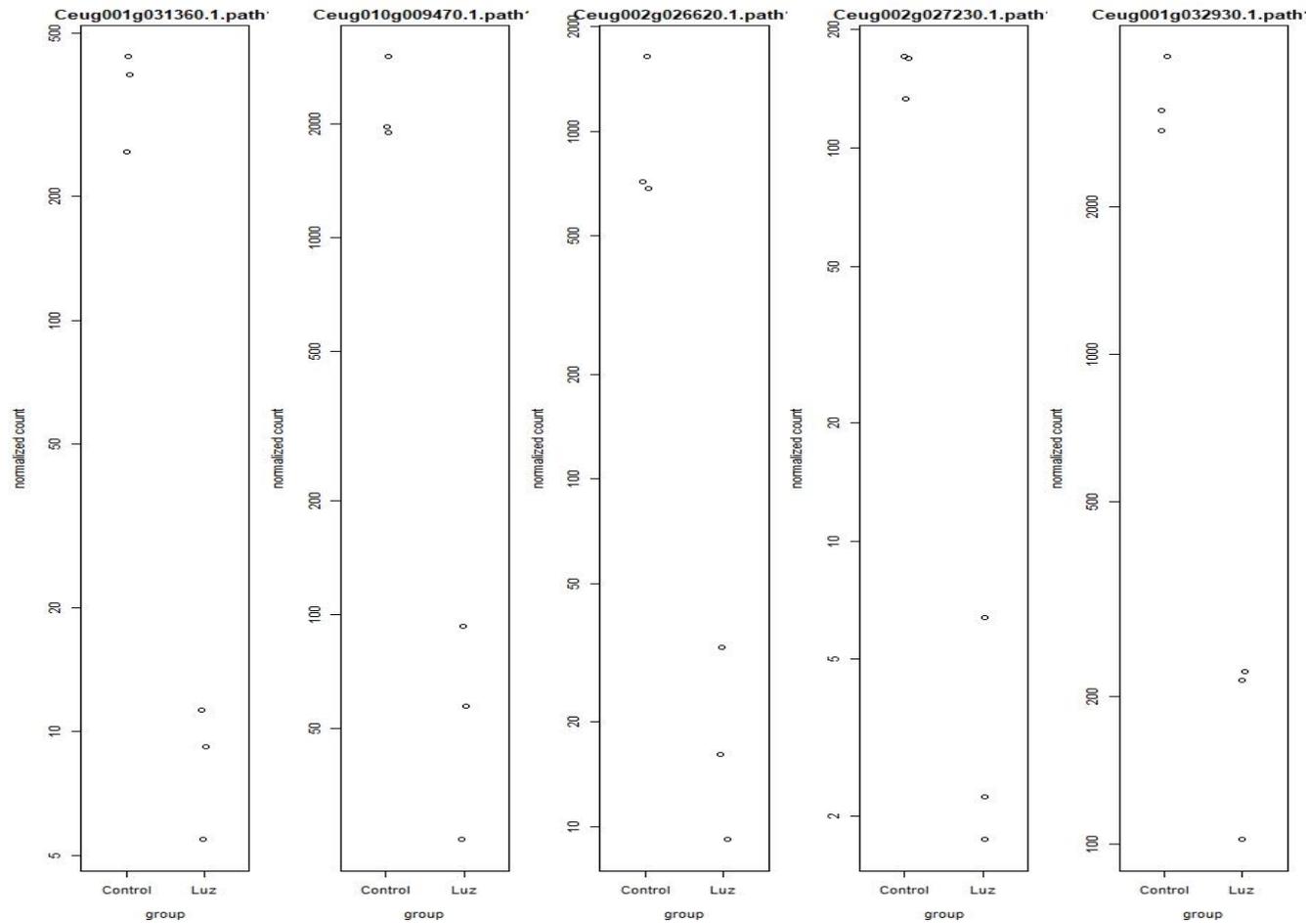
E por fim, quando a análise foi realizada com a referência do genoma de *C. eugenioides*, os 5 genes mais expressos que são Ceug009g010260.1.path1, Ceug005g015230.1.path1, Ceug001g020940.1.path1, Ceug011g029420.1.path1, Ceug009g010270.1.path1 (figura 35) e os 5 mais reprimidos da espécie *eugenioides* que são Ceug001g031360.1.path1, Ceug010g009470.1.path1, Ceug002g026620.1.path1, Ceug002g027230.1.path1, Ceug001g032930.1.path1 (figura 36).

Figura 35 - 5 genes mais expressos na condição de estresse luminoso em *Coffea eugenioides*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são os Ceug009g010260.1.path1, Ceug005g015230.1.path1, Ceug001g020940.1.path1, Ceug011g029420.1.path1, Ceug009g010270.1.path1



Fonte: Do autor.

Figura 36 - 5 genes mais reprimidos na condição de estresse luminoso em *Coffea eugenioides*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são os Ceug001g031360.1.path1, Ceug010g009470.1.path1, Ceug002g026620.1.path1, Ceug002g027230.1.path1, Ceug001g032930.1.path1.

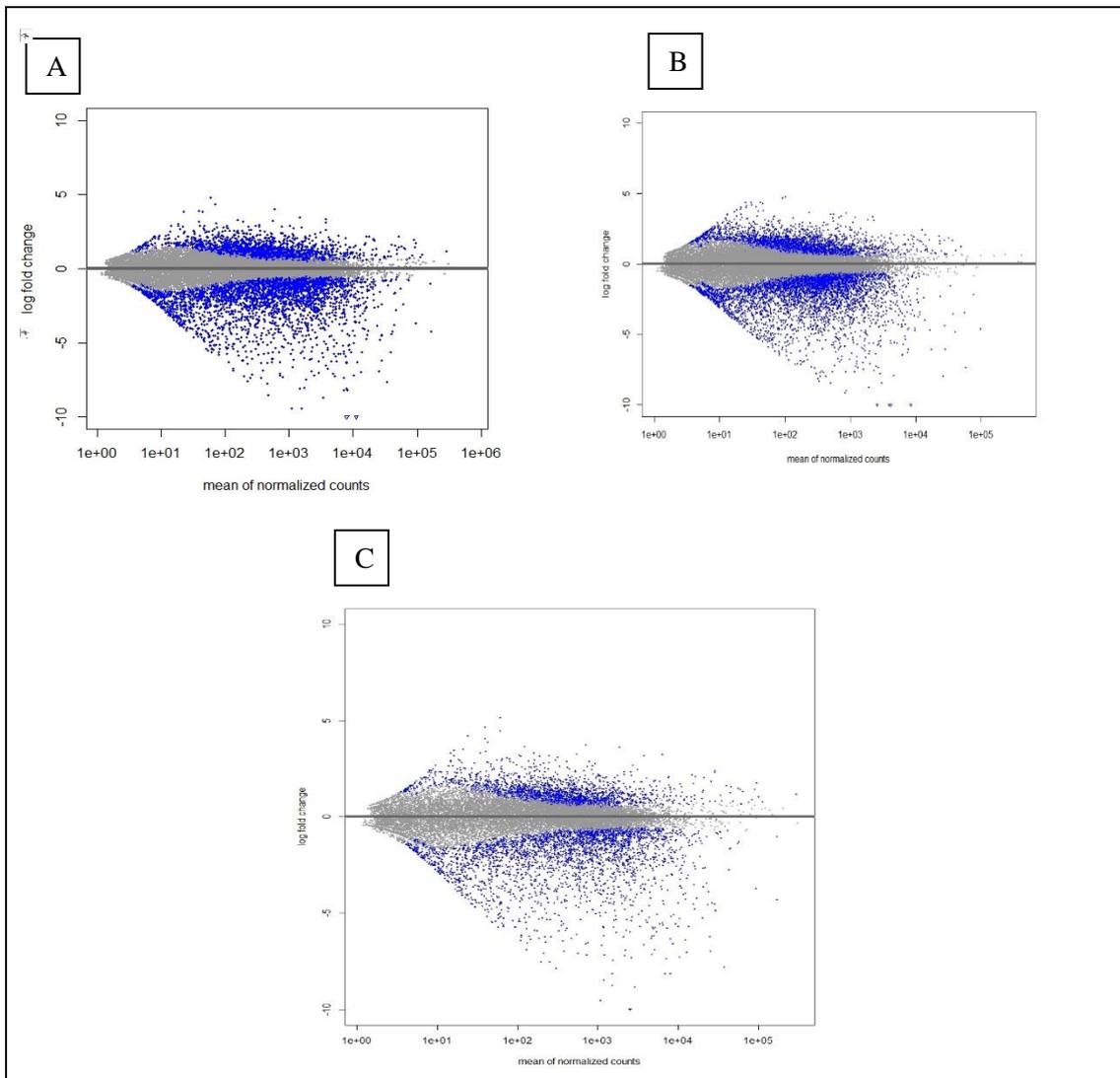


Fonte: Do autor.

5.5. Padrões de genes diferencialmente expressos resultantes das 3 análises realizadas com diferentes genomas de referência (*C. arabica*, *C. canephora* e *C. eugenioides*), responsivos ao estresse causado pela restrição hídrica

Na figura 37, estão apresentados os gráficos de dispersão para as 3 análises realizadas (Referência: *C.arabica*, *C.canephora* e *C.eugenioides*) gerados ao final da comparação do tratamento de exposição ao estresse causado pela restrição hídrica e a condição controle. É possível notar padrões de expressão diferencial, com genes mais expressos (*up_regulated*) e com maior repressão (*down_regulated*). Nota-se, neste caso (figura 35) um número maior de genes diferencialmente expressos, independente do genoma de referência utilizado na análise, responsivos ao estresse causado pela restrição hídrica, de forma similar ao tratamento de calor (Calor – figura 20). Da mesma forma, pode-se observar também, pelos dados apresentados na figura 37, uma pequena redução no número de genes diferencialmente expressos quando a análise foi realizada utilizando-se o genoma de referência de *C. eugenioides*.

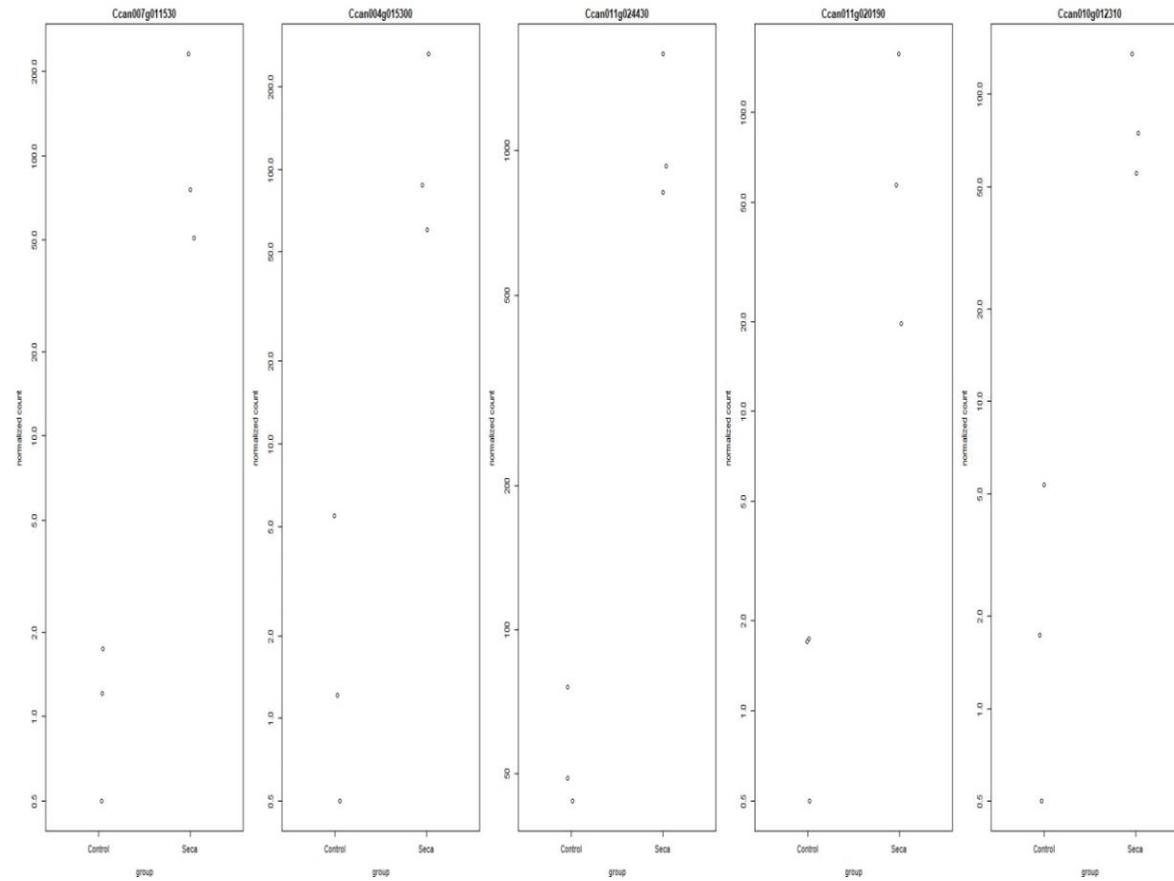
Figura 37 - Gráficos de dispersão contendo os genes de *Coffea new canephora* (A), *Coffea arabica* (B), *Coffea eugenioides* (C) quando expostos à restrição hídrica.



Fonte: Do autor.

Na figura 38 é possível observar que os genes mais expressos com a análise utilizando-se o genoma de referência de *C. canephora*, foram Ccan007g011530, Ccan004g015300, Ccan011g024430, Ccan011g020190, Ccan010g012310.

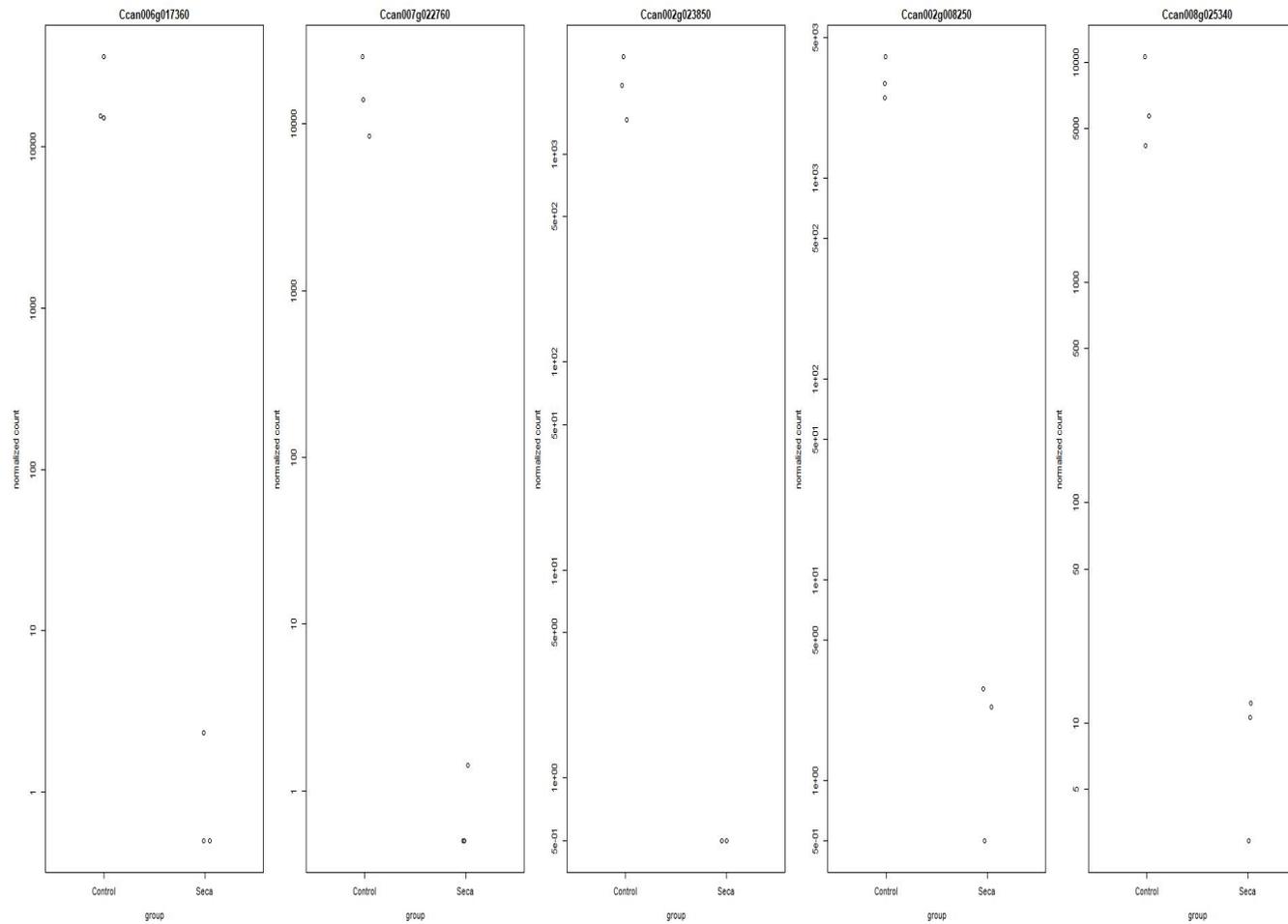
Figura 38 - 5 genes mais expressos na condição de restrição hídrica. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente menor. Esses são os Ccan007g011530, Ccan004g015300, Ccan011g024430, Ccan011g020190, Ccan010g012310.



Fonte: Do autor.

Adicionalmente, foram avaliados os cinco genes mais reprimidos quando levando em consideração o controle da mesma espécie, esses foram: Ccan006g017360, Ccan007g022760, Ccan002g023850, Ccan002g008250, Ccan008g025340.

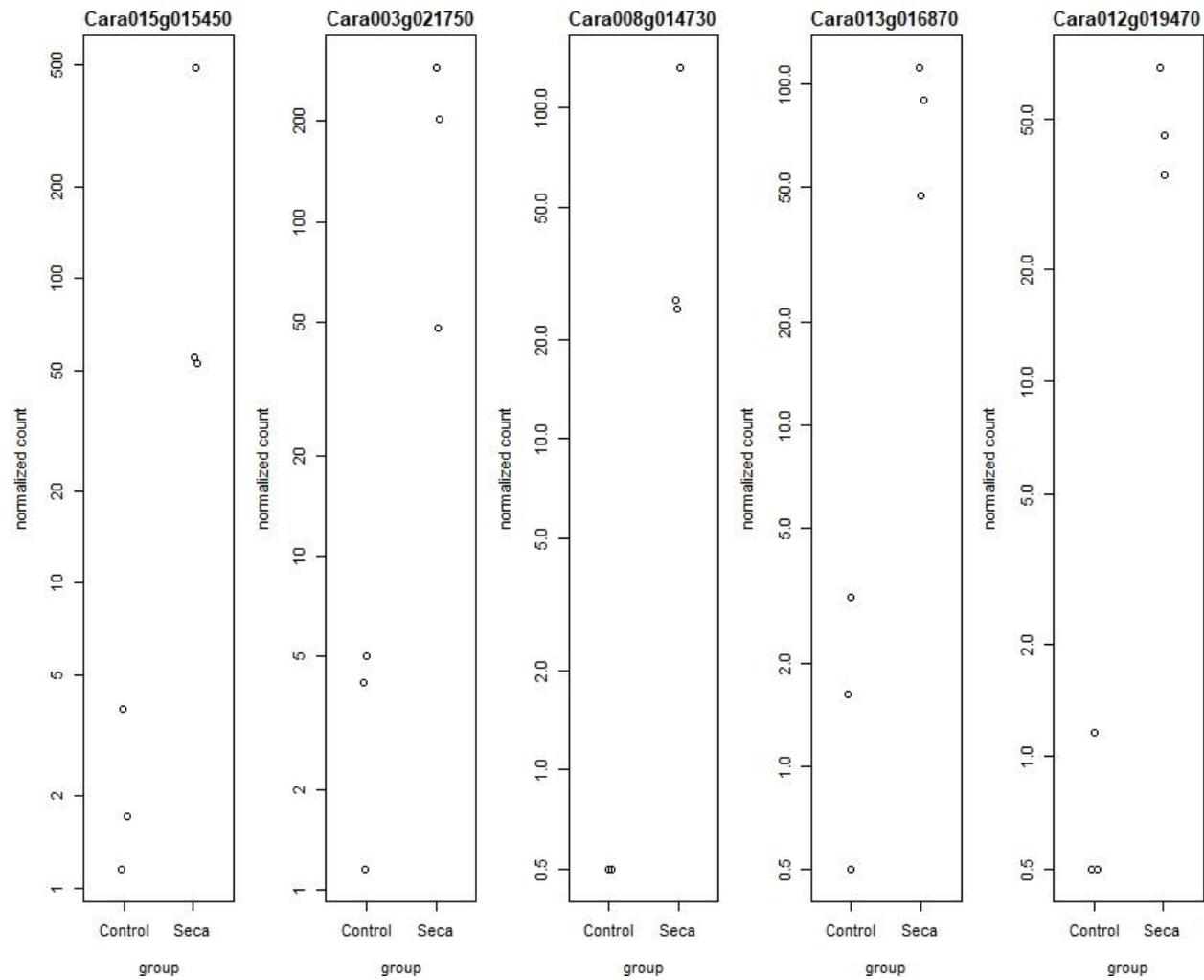
Figura 39 - 5 genes mais reprimidos na condição de restrição hídrica em *Coffea canephora*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são os Ccan006g017360, Ccan007g022760, Ccan002g023850, Ccan002g008250, Ccan008g025340.



Fonte: Do autor

O resultado das análises com a utilização do genoma de referência de *C. arabica*, indicam que os genes mais expressos foram (figura 40).

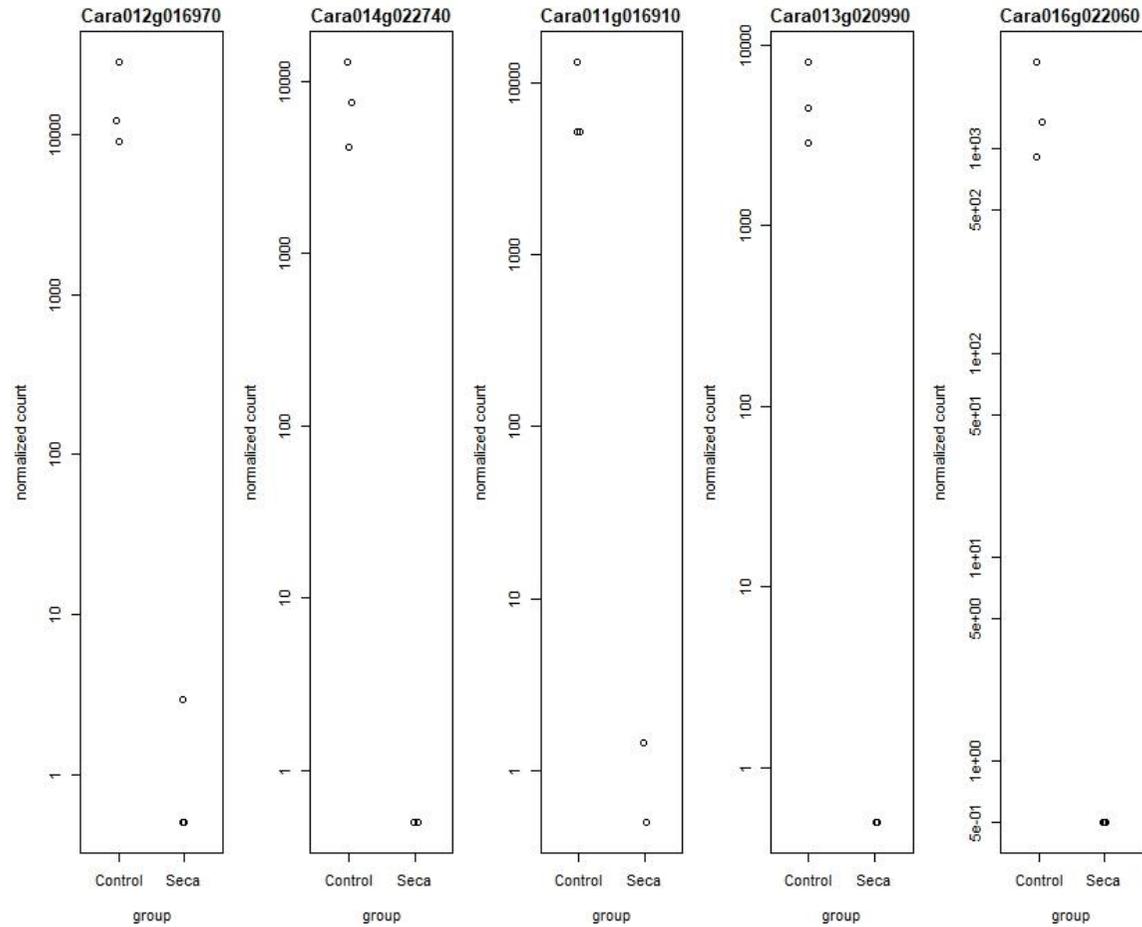
Figura 40 - 5 genes mais expressos na condição de restrição hídrica em *Coffea arabica*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente menor. Esses são Cara015g015450, Cara003g021750, Cara008g014730, Cara013g016870, Cara012g019470.



Fonte: Do autor.

Já os mais reprimidos foram Cara012g016970, Cara014g022740, Cara011g016910, Cara013g020990, Cara016g022060 (figura 45).

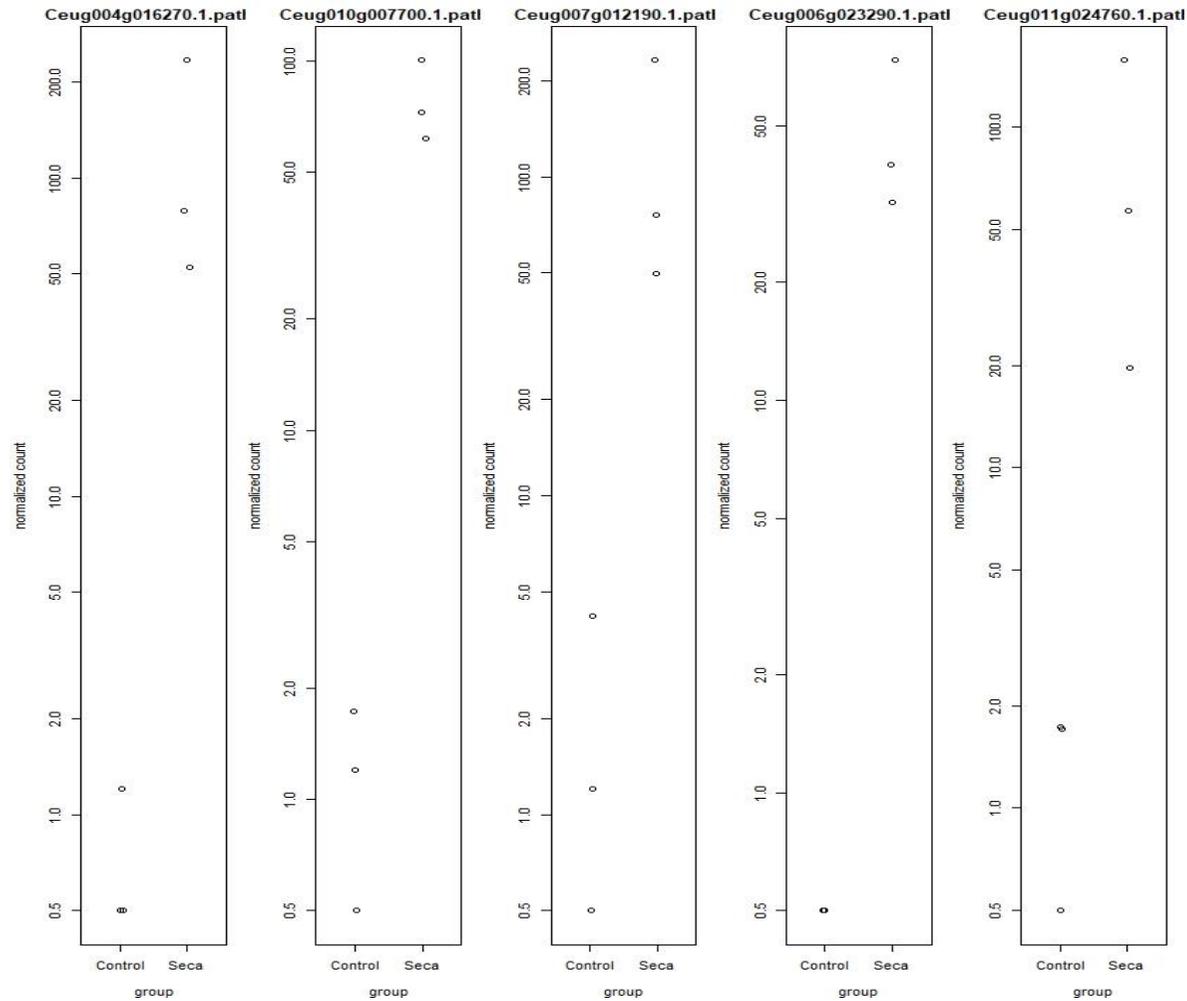
Figura 41 - 5 genes mais reprimidos na condição de restrição hídrica em *Coffea arabica*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são Cara012g016970, Cara014g022740, Cara011g016910, Cara013g020990, Cara016g022060.



Fonte: Do autor.

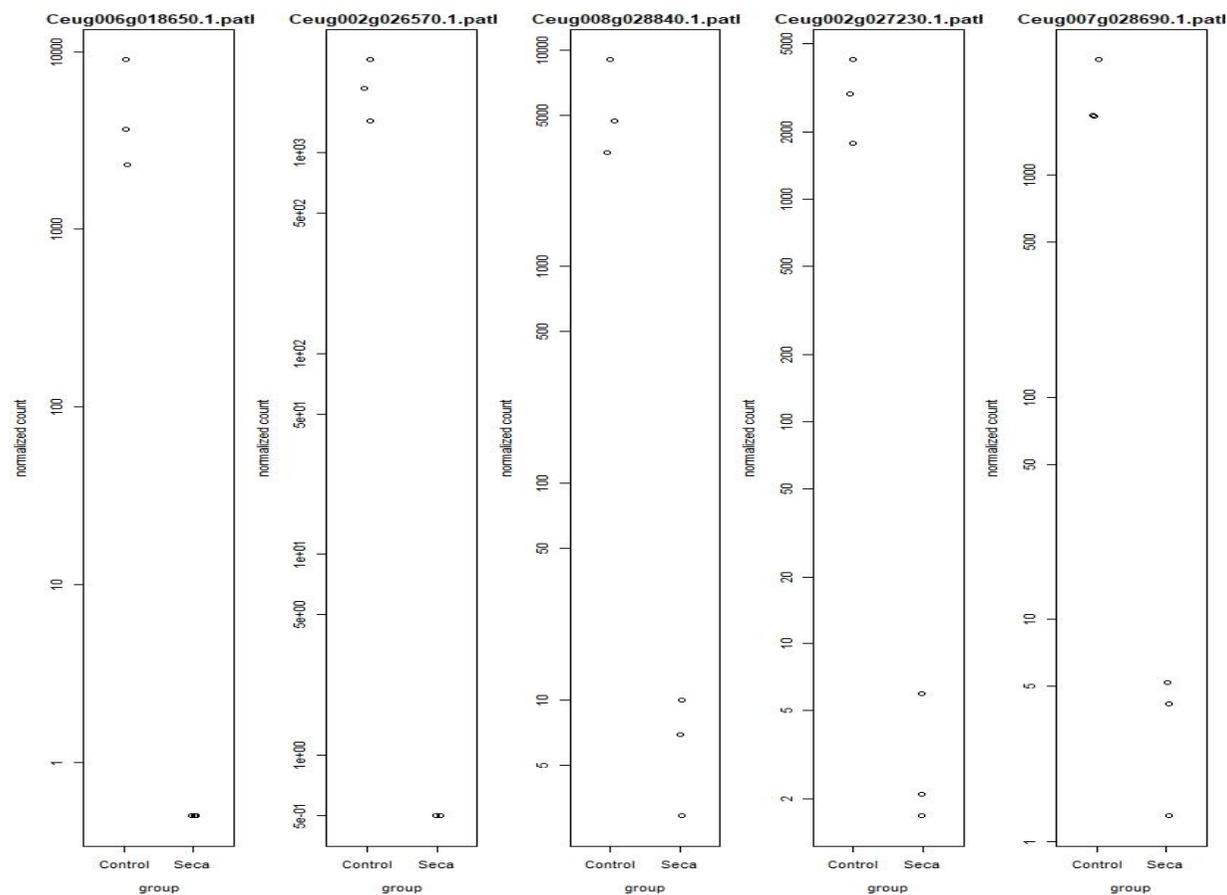
E por fim, quando a análise foi realizada com a referência do genoma de *C. eugenioides*, os 5 genes mais expressos que são Ceug004g016270.1.path1, Ceug010g007700.1.path1, Ceug007g012190.1.path1, Ceug006g023290.1.path1, Ceug011g024760.1.path1 (figura 42) e os 5 mais reprimidos da espécie *eugenioides* que são Ceug006g018650.1.path1, Ceug002g026570.1.path1, Ceug008g028840.1.path1, Ceug002g027230.1.path1, Ceug007g028690.1.path1 (figura 43).

Figura 42 - 5 genes mais expressos na condição de restrição hídrica em *Coffea eugenioides*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são os Ceug004g016270.1.path1, Ceug010g007700.1.path1, Ceug007g012190.1.path1, Ceug006g023290.1.path1, Ceug011g024760.1.path1



Fonte: Do autor.

Figura 43 - 5 genes mais reprimidos na condição de restrição hídrica em *Coffea eugenioides*. Nota-se que os mesmos, quando em condição de controle, geraram um fator de expressão consideravelmente maior. Esses são os Ceug006g018650.1.path1, Ceug002g026570.1.path1, Ceug008g028840.1.path1, Ceug002g027230.1.path1, Ceug007g028690.1.path1.



Fonte: Do autor.

5.6. Análise da similaridade das proteínas responsivas identificadas em cada pipeline (ref. *C.arabica*, *C.canephora* e *C.eugenioides*), nos diferentes tratamentos estudados

Ao comparar a similaridade das proteínas geradas pelos 10 genes identificados (5 *up* e 5 *down-regulated*) nas diferentes análises realizadas, utilizando-se diferentes genomas de referências no mapeamento dos reads de RNA-Seq, buscou-se avaliar os efeitos resultantes da metodologia utilizada nos resultados obtidos e a possibilidade de se identificar potenciais homeólogos de *C.canephora* e *C. eugenioides*, com a estratégia estudada. As tabelas de 1 à 5 apresentam os resultados de concordância e discrepância das análises realizadas, abordando os diferentes tratamentos em estudo.

Os dados apresentados na Tabela 1, indicam uma alta convergência nos resultados de expressão diferencial aqui utilizados no estudo, sendo que das 10 proteínas indenficadas com os maiores valores de *Log2foldchange*, na comparação ABA vs controle, observou-se 90% de concordância nas análises comparativas entre os resultados obtidos com ref: *C.arabica* vs. *C.canephora*, 80% nas análises comparativas entre os resultados obtidos com ref: *C.arabica* vs. *C.eugenioides* e 70% nas análises comparativas entre os resultados obtidos com ref: *C.canephora* vs. *C.eugenioides*.

Tabela 1 - Comparação das sequências proteicas geradas, a partir dos 10 genes com maiores valores de *Log2foldchange* (5 *up* e 5 *down-reguated*), realizadas com diferentes genoma de referência (*Coffea arabica*, *Coffea canephora* e *Coffea eugenioides*) nos estudos de RNA-seq de dados obtidos de plantas submetidas ao estresse de aplicação exógena de ABA. As lacunas preenchidas com N.R significam que nenhuma similaridade proteica foi encontrada entre o conjunto de proteínas comparado.(continua)

Resultados obtidos na análise do tratamento exógeno de ABA

Coffea arabica x *Coffea canephora*

Sequências inqueridas	Sequências alinhamentos com significativos	Porcentagem da abrangência da sequência inquerida	*Porcentagem da identidade
Cara017g014900	Ccan009g016830	96%	100%
Cara018g007530	Ccan009g016830	96%	100%
Cara009g021370	Ccan005g023910	100%	100%
Cara003g037460	Ccan002g038870	59%	99%
Cara010g022480	Ccan005g023910	100%	90%
Cara003g017520	Ccan002g017530	98%	99%
Cara014g008300	Ccan007g008340	93%	86%
Cara001g017020	N.R.	N.R.	N.R

Tabela 1 - Comparação das sequências proteicas geradas, a partir dos 10 genes com maiores valores de *Log2foldchange* (5 up e 5 down-regulated), realizadas com diferentes genoma de referência (*Coffea arabica*, *Coffea canephora* e *Coffea eugenioides*) nos estudos de RNA-seq de dados obtidos de plantas submetidas ao estresse de aplicação exógena de ABA. As lacunas preenchidas com N.R significam que nenhuma similaridade proteica foi encontrada entre o conjunto de proteínas comparado.(conclusão)

Cara003g044960	Ccan002g047080	50%	99%
Cara012g015180	Ccan006g015490	100%	96%
<i>Coffea arabica x Coffea eugenioides</i>			
Cara017g014900	Ceug006g030840	96%	100%
Cara018g007530	Ceug006g030840	96%	100%
Cara009g021370	Ceug005g026420	100%	100%
Cara003g037460	N.R.	N.R.	N.R.
Cara010g022480	Ceug005g026420	100%	90%
Cara003g017520	Ceug002g019190	100%	74%
Cara014g008300	Ceug007g008480	100%	99%
Cara001g017020	N.R.	N.R.	N.R.
Cara003g044960	Ceug002g052130	50%	97%
Cara012g015180	Ceug006g016250	100%	100%
<i>Coffea canephora x Coffea eugenioides</i>			
Ccan005g023910	Ceug005g026420	100%	100%
Ccan009g016830	Ceug006g030840	100%	100%
Ccan002g038870	N.R.	N.R.	N.R.
Ccan004g007630	Ceug004g007680	100%	99%
Ccan006g018680	N.R.	N.R.	N.R.
Ccan007g008340	Ceug007g008480	100%	96%
Ccan002g047080	Ceug002g052130	100%	97%
Ccan002g017530	Ceug002g019190	100%	99%
Ccan006g015490	Ceug006g016250	100%	96%
Ccan002g023900	N.R.	N.R.	N.R.

*Algumas proteínas apresentaram similaridade significativa com mais de uma sequência no conjunto comparado (somente foi considerado o resultado com maior significância). Fonte: Do autor

Tabela 2 - Comparação das sequências proteicas geradas, a partir dos 10 genes com maiores valores de *Log2foldchange* (5 up e 5 down-regulated), realizadas com diferentes genoma de referência (*Coffea arabica*, *Coffea canephora* e *Coffea eugenioides*) nos estudos de RNA-seq de dados obtidos de plantas submetidas ao estresse de calor. As lacunas preenchidas com N.R significam que nenhuma similaridade proteica foi encontrada entre o conjunto de proteínas comparado.(continua)

Resultados obtidos na análise do tratamento do estresse térmico de calor

<i>Coffea arabica x Coffea canephora</i>				
Sequências inqueridas	Sequências alinhamentos significativos	com	Porcentagem da abrangência da sequência inquerida	*Porcentagem da identidade
Cara016g024390	Ccan008g027560		100%	99%
Cara012g023820	Ccan006g024380		93%	98%
Cara010g023290	N.R.		N.R.	N.R.
Cara004g010590	N.R.		N.R.	N.R.
Cara003g045870	N.R.		N.R.	N.R.
Cara004g028770	Ccan002g028410		100%	98%

Tabela 2 - Comparação das sequências proteicas geradas, a partir dos 10 genes com maiores valores de *Log2foldchange* (5 up e 5 down-regulated), realizadas com diferentes genoma de referência (*Coffea arabica*, *Coffea canephora* e *Coffea eugenioides*) nos estudos de RNA-seq de dados obtidos de plantulas submetidas ao estresse de calor. As lacunas preenchidas com N.R significam que nenhuma similaridade proteica foi encontrada entre o conjunto de proteínas comparado.(conclusão)

Cara015g018780	Ccan008g022320	98%	95%
Cara016g019120	Ccan008g022320	100%	85%
Cara012g006660	Ccan006g006890	100%	98%
Cara020g012240	Ccan010g012200	87%	95%
<i>Coffea arabica x Coffea eugenioides</i>			
Cara016g024390	Ceug008g031390	100%	100%
Cara012g023820	Ceug006g026690	93%	98%
Cara010g023290	N.R.	N.R.	N.R.
Cara004g010590	N.R.	N.R.	N.R.
Cara003g045870	N.R.	N.R.	N.R.
Cara004g028770	Ceug002g031250	100%	100%
Cara015g018780	Ceug008g025630	84%	94%
Cara016g019120	Ceug008g025630	91%	83%
Cara012g006660	Ceug006g007050	100%	97%
Cara020g012240	Ceug010g013490	97%	93%
<i>Coffea canephora x Coffea eugenioides</i>			
Ccan010g016230	Ceug010g017220	100%	97%
Ccan006g024380	Ceug006g026690	99%	98%
Ccan008g027560	Ceug008g031390	100%	98%
Ccan002g024410	Ceug002g027160	99%	96%
Ccan005g014260	N.R.	N.R.	N.R.
Ccan008g022320	Ceug008g025630	90%	98%
Ccan002g028410	Ceug002g031250	100%	98%
Ccan010g012200	Ceug010g013490	100%	97%
Ccan006g006890	Ceug006g007050	95%	98%
Ccan002g011900	Ceug006g007050	73%	73%

*Algumas proteínas apresentaram similaridade significativa com mais de uma sequência no conjunto comparado (somente foi considerado o resultado com maior significância). Fonte: Do autor

Os dados apresentados na Tabela 3, indicam convergência nos resultados de expressão diferencial aqui utilizados no estudo, porém, das 10 proteínas indenficadas com os maiores valores de *Log2foldchange*, na comparação Frio vs controle, observou-se 80% de concordância nas análises comparativas entre os resultados obtidos com ref: *C.arabica* vs. *C.canephora*, 40% nas análises comparativas entre os resultados obtidos com ref: *C.arabica* vs. *C.eugenioides* e 40% nas análises comparativas entre os resultados obtidos com ref: *C.canephora* vs. *C.eugenioides*.

Tabela 3 - Comparação das sequências proteicas geradas, a partir dos 10 genes com maiores valores de *Log2foldchange* (5 up e 5 down-regulated), realizadas com diferentes genoma de referência (*Coffea arabica*, *Coffea canephora* e *Coffea eugenioides*) nos estudos de RNA-seq de dados obtidos de plantulas submetidas ao estresse térmico de frio. As lacunas preenchidas com N.R significam que nenhuma similaridade proteica foi encontrada entre o conjunto de proteínas comparado.Resultados obtidos na análise do tratamento do estresse térmico de frio.

Coffea arabica x *Coffea canephora*

Sequências inqueridas	Sequências alinhamentos com significativos	Porcentagem da abrangência da sequência inquerida	*Porcentagem da identidade
Cara019g019470	Ccan010g021060	99%	98%
Cara019g019460	Ccan010g021060	99%	89%
Cara015g021070	N.R.	N.R.	N.R.
Cara004g011830	Ccan002g011660	100%	98%
Cara005g020810	N.R.	N.R.	N.R.
Cara019g010200	Ccan010g010260	100%	96%
Cara007g007360	Ccan004g007450	99%	99%
Cara020g010250	Ccan010g010260	100%	100%
Cara009g021510	Ccan005g024050	100%	99%
Cara014g005840	Ccan007g005840	100%	96%

Coffea arabica x *Coffea eugenioides*

Cara019g019470	N.R.	N.R.	N.R.
Cara019g019460	N.R.	N.R.	N.R.
Cara015g021070	N.R.	N.R.	N.R.
Cara004g011830	Ceug002g012340	100%	99%
Cara005g020810	N.R.	N.R.	N.R.
Cara019g010200	Ceug010g011160	100%	100%
Cara007g007360	Ceug004g007480	99%	98%
Cara020g010250	Ceug010g011160	100%	96%
Cara009g021510	N.R.	N.R.	N.R.
Cara014g005840	N.R.	N.R.	N.R.

Coffea canephora x *Coffea eugenioides*

Ccan002g011660	Ceug002g012340	100%	97%
Ccan002g024830	Ceug002g027580	100%	99%
Ccan010g021060	N.R.	N.R.	N.R.
Ccan004g016910	N.R.	N.R.	N.R.
Ccan010g015020	N.R.	N.R.	N.R.
Ccan010g010260	Ceug010g011160	100%	96%
Ccan004g007450	Ceug004g007480	100%	97%
Cara020g010250	N.R.	N.R.	N.R.
Cara009g021510	N.R.	N.R.	N.R.
Cara014g005840	N.R.	N.R.	N.R.

*Algumas proteínas apresentaram similaridade significativa com mais de uma sequência no conjunto comparado (somente foi considerado o resultado com maior significância). Fonte:Do autor

Os dados apresentados na Tabela 4, indicam convergência nos resultados de expressão diferencial aqui utilizados no estudo, sendo que das 10 proteínas indicadas com os maiores valores de *Log2foldchange*, na comparação Luminosidade vs controle, observou-se 60% de concordância nas análises comparativas entre os resultados obtidos com ref: *C.arabica* vs. *C.canephora*, 60% nas análises comparativas entre os resultados obtidos com ref: *C.arabica* vs. *C.eugenioides* e 70% nas análises comparativas entre os resultados obtidos com ref: *C.canephora* vs. *C.eugenioides*.

Tabela 4 - Comparação das sequências proteicas geradas, a partir dos 10 genes com maiores valores de *Log2foldchange* (5 up e 5 down-regulated), realizadas com diferentes genoma de referência (*Coffea arabica*, *Coffea canephora* e *Coffea eugenioides*) nos estudos de RNA-seq de dados obtidos de plantas submetidas ao estresse de luminosidade. As lacunas preenchidas com N.R significam que nenhuma similaridade proteica foi encontrada entre o conjunto de

proteínas comparado.

Resultados obtidos na análise do tratamento do estresse luminoso

Coffea arabica x Coffea canephora

Sequências inqueridas	Sequências alinhamentos significativos com	Porcentagem abrangência da sequência inquerida	*Porcentagem da identidade
Cara010g012390	Ccan005g014130	96%	95%
Cara015g024260	N.R.	N.R.	N.R.
Cara022g022370	N.R.	N.R.	N.R.
Cara012g010380	N.R.	N.R.	N.R.
Cara004g050170	N.R.	N.R.	N.R.
Cara019g008790	Ccan002g023900	72%	26%
Cara003g008210	Ccan002g008110	72%	99%
Cara020g008930	Ccan002g023900	72%	26%
Cara003g023910	Ccan002g023900	80%	98%
Cara001g017020	Ccan001g022880	88%	96%

Coffea arabica x Coffea eugenioides

Cara010g012390	Ceug005g015230	96%	100%
Cara015g024260	N.R.	N.R.	N.R.
Cara022g022370	Ceug011g029420	98%	100%
Cara012g010380	N.R.	N.R.	N.R.
Cara004g050170	N.R.	N.R.	N.R.
Cara019g008790	Ceug010g009470	97%	100%
Cara003g008210	Ceug010g009470	14%	57%
Cara020g008930	Ceug010g009470	97%	26%
Cara003g023910	Ceug002g026620	80%	99%
Cara001g017020	N.R.	N.R.	N.R.

Coffea canephora x Coffea eugenioides

Ccan005g016790	N.R.	N.R.	N.R.
Ccan005g019360	Ceug009g010270	100%	94%
Ccan005g014260	Ceug005g015230	100%	93%
Ccan005g014130	Ceug005g015230	100%	95%
Ccan004g007630	N.R.	N.R.	N.R.
Ccan002g023900	Ceug002g026620	100%	99%
Ccan001g027260	Ceug001g032930	99%	97%
Ccan001g022880	N.R.	N.R.	N.R.
Ccan002g008110	Ceug010g009470	17%	57%
Ccan001g025820	Ceug001g031360	98%	98%

*Algumas proteínas apresentaram similaridade significativa com mais de uma sequência no conjunto comparado (somente foi considerado o resultado com maior significância). Fonte: Do autor

Os dados apresentados na Tabela 5, indicam convergência nos resultados de expressão diferencial aqui utilizados no estudo, sendo que das 10 proteínas indicadas com os maiores valores de *Log2foldchange*, na comparação restrição hídrica vs controle, observou-se 60% de concordância nas análises comparativas entre os resultados obtidos com ref: *C.arabica* vs. *C.canephora*, 40% nas análises comparativas entre os resultados obtidos com ref: *C.arabica* vs. *C.eugenioides* e 60% nas análises comparativas entre os resultados obtidos com ref: *C.canephora* vs. *C.eugenioides*.

Tabela 5 - Comparação das sequências proteicas geradas, a partir dos 10 genes com maiores

valores de *Log2foldchange* (5 up e 5 down-regulated), realizadas com diferentes genoma de referência (*Coffea arabica*, *Coffea canephora* e *Coffea eugenioides*) nos estudos de RNA-seq de dados obtidos de plantas submetidas ao estresse de restrição hídrica. As lacunas preenchidas com N.R significam que nenhuma similaridade proteica foi encontrada entre o conjunto de proteínas comparado.

Resultados obtidos na análise do tratamento de estresse à restrição hídrica
Coffea arabica x *Coffea canephora*

Comparação de proteínas do estresse de restrição hídrica

Sequências inqueridas	Sequências alinhamentos com significativos	Porcentagem da abrangência da sequência inquerida	*Porcentagem da identidade
Cara015g015450	N.R.	N.R.	N.R.
Cara003g021750	N.R.	N.R.	N.R.
Cara008g014730	Ccan004g015300	84%	97%
Cara013g016870	N.R.	N.R.	N.R.
Cara012g019470	N.R.	N.R.	N.R.
Cara012g016970	Ccan006g017360	81%	95%
Cara014g022740	Ccan007g022760	100%	97%
Cara011g016910	Ccan006g017360	81%	100%
Cara013g020990	Ccan007g022760	100%	100%
Cara016g022060	Ccan008g025340	87%	97%

Coffea arabica x *Coffea eugenioides*

Cara015g015450	N.R.	N.R.	N.R.
Cara003g021750	N.R.	N.R.	N.R.
Cara008g014730	Ceug004g016270	84%	100%
Cara013g016870	N.R.	N.R.	N.R.
Cara012g019470	N.R.	N.R.	N.R.
Cara012g016970	Ceug006g018650	83%	93%
Cara014g022740	N.R.	N.R.	N.R.
Cara011g016910	Ceug006g018650	83%	90%
Cara013g020990	N.R.	N.R.	N.R.
Cara016g022060	Ceug008g028840	78%	99%

Coffea canephora x *Coffea eugenioides*

Ccan007g011530	Ceug007g012190	98%	99%
Ccan004g015300	Ceug004g016270	N.R.	N.R.
Ccan011g024430	N.R.	N.R.	N.R.
Ccan011g020190	Ceug011g024760	100%	99%
Ccan010g012310	N.R.	N.R.	N.R.
Ccan006g017360	Ceug006g018650	83%	93%
Ccan010g012310	N.R.	N.R.	N.R.
Ccan002g023850	Ceug002g026570	100%	96%
Ccan002g008250	N.R.	N.R.	N.R.
Ccan008g025340	Ceug008g028840	69%	97%

*Algumas proteínas apresentaram similaridade significativa com mais de uma sequência no conjunto comparado (somente foi considerado o resultado com maior significância). Fonte: Do autor

6 CONCLUSÃO

Por meio das análises realizadas neste trabalho, foi possível verificar que houveram discrepância nos resultados de expressão diferencial de RNA-seq em função do genoma de referência utilizado no estudo. Entretanto, foi também observado alta convergência dos

resultados em algumas das análises realizadas (tratamentos), independente do genoma utilizado.

Observou-se, de forma geral, uma tendência de pior predição gênica nas anotações do genoma de *Coffea eugenioides*, o que pode ter proporcionado os menores valores percentuais de concordância nas análises. De modo geral, com a estratégia utilizada, ficou demonstrado a possibilidade de se identificar com sucesso os homeólogos de *C.canephora* (*CaCc*) e *C.eugenioides* (*CaCe*) presentes no genoma de *C. arabica*.

REFERÊNCIAS

- AGARWAL, M. et al. A R2R3 type MYB transcription factor is involved in the cold regulation of CBF genes and in acquired freezing tolerance. **Journal of Biological Chemistry**, Rockville, v. 281, n. 49, p. 37636–37645, 2006.
- ALVES, G. C. S. **Characterization of a candidate gene for drought tolerance in *Coffea*: the *CcDREB1D* gene, in contrasting genotypes of *Coffea canephora* and related species. 2015. 191 p.** PhD Thesis - Montpellier SupAgro, France / Federal University of Lavras, Lavras, Brazil, 2015.
- ALVES, Rita C. et al. State of the art in coffee processing by-products. In: **Handbook of coffee processing by-products**. Academic Press, 2017. p. 1-26.
- ANDERS, S; HUBER, W. Differential expression analysis for sequence count data. **Genome Biology**, London, v. 11, n. 10, p. R106, 2010.
- ARVE, L. E. et al. **Stomatal Responses to Drought Stress and Air Humidity**. In: SHANKER, A; VENKATESWARLU, B. *Abiotic Stress in Plants - Mechanisms and Adaptations*. 1.ed., 2011.
- ASSELBERGH, B; De VIEESSCHAUWER, D; HOFTE, M. Global switches and fine-tuning—ABA modulates plant pathogen defense. **Molecular Plant-Microbe Interactions**, Minnesota, v. 21, n. 6, p. 709–719, 2008b.
- ATKINSON, N. J; URWIN, P. E. The interaction of plant biotic and abiotic stresses: from genes to the field. **Journal of Experimental Botany**, Oxford, v. 63, n. 10, p. 3523–3544, 2012.
- BATES, G. W. et al. A comparative study of the *Arabidopsis thaliana* guard-cell transcriptome and its modulation by sucrose. **PLoS ONE**, Cambridge, v. 7, n. 11, p. e49641, 2012.
- BENTLEY, D. R. et al. Accurate whole human genome sequencing using reversible terminator chemistry. **Nature**, London, v. 456, n. 7218, p. 53–9, 2008.
- BARTHOLOMEW, Dolores M.; BARTLEY, Glenn E.; SCOLNIK, Pablo A. Abscisic acid control of *rbcS* and *cab* transcription in tomato leaves. **Plant Physiology**, v. 96, n. 1, p. 291-296, 1991.
- CAMPOS, P. S. et al. Electrolyte leakage and lipid degradation account for cold sensitivity in leaves of *Coffea* sp. plants. **Journal of Plant Physiology**, Amsterdam, v. 160, n. 3, p. 283-292, 2003.
- CHANG, Yun C.; WALLING, Linda L. Abscisic acid negatively regulates expression of chlorophyll a/b binding protein genes during soybean embryogeny. **Plant Physiology**, v. 97, n. 3, p. 1260-1264, 1991.

CHEN, M. et al. Cold-induced modulation and functional analyses of the DRE-binding transcription factor gene, *GmDREB3*, in soybean (*Glycine max*). *Journal of Experimental Botany*, Oxford, v. 60, n. 1, p. 121–135, 2009.

CHEN, W. Q. et al. Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative functions in response to environmental stresses. *Plant cell*, Rockville, v. 14, n. 3, p. 559-574, mar. 2002.

CHINI, A. et al. Drought tolerance established by enhanced expression of the *CC-NBSLRR* gene, *ADRI*, requires salicylic acid, *EDS1* and *ABII*. **Plant Journal**, Oxford, v. 38, n. 5, p. 810-822, June 2004.

CHINNUSAMY, V. et al. ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. **Genes & Development**, Woodbury, v. 17, n. 8, p. 1043–1054, 2003.

CHINNUSAMY, V.; ZHU, J.; ZHU, J. K. Cold stress regulation of gene expression in plants. **TREND in Plant Science**, Cambridge, v. 12, n. 10, p. 444-451, 2007.

CHRISTMANN, A. et al. A hydraulic signal in root-to-shoot signaling of water shortage. **The Plant Journal**, Malden, v. 52, n. 1, p. 167–174, 2007.

CIARMIELLO, L. F. et al. **Plant Genes for Abiotic Stress**. In: *Abiotic Stress in plants - Mechanisms and Adaptations*, Prof. Arun Shanker (Ed.), ISBN: 978-953-307-394-1, 2011.

CLARINDO, W. R.; CARVALHO, C. R. Flow cytometric analysis using SYBR Green I for genome size estimation in coffee. **Acta Histochemica**, Jena, v. 113, n. 2, p. 221-225, Feb. 2011.

CLEMENT, M. et al. Identification of new up-regulated genes under drought stress in soybean nodules. **Gene**, Amsterdam, v. 426, n. 1-2, p.15-22, 2008.

COMBES, M. C. et al. Contribution of subgenomes to the transcriptome and their intertwined regulation in the allopolyploid *Coffea arabica* grown at contrasted temperatures. **The New Phytologist**, Lancaster v. 200, n. 1, p. 251–260, 2013.

COSTA, Tatiana Santos. Análise do perfil transcriptômico e proteômico de raízes de diferentes clones de *Coffea canephora* em condições de déficit hídrico. **Embrapa Recursos Genéticos e Biotecnologia-Tese/dissertação (ALICE)**, 2014.

COTTA, M. G. et al. Lipid transfer proteins in coffee: isolation of *Coffea* orthologs, *Coffea arabica* homeologs, expression during coffee fruit development and promoter analysis in transgenic tobacco plants. **Plant Molecular Biology**, Basel, v. 85, n. 1-2, p. 11–31, 2014.

COTTA, Michelle Guitton. **Molecular mechanisms in the first step of ABA mediated response in *Coffea* ssp.** 2017. Tese de Doutorado. Montpellier SupAgro.

CRUZ, J. R. G. A. **Evaluating differential gene expression using RNA-sequencing data: a case study in host-pathogen interaction upon *Listeria monocytogenes* infection.** 2013.

159 p. Thesis (Master of Science Degree in Biomedical Engineering) – IFT Técnico Lisboa, Lisboa, 2013.

CUSHMAN, J. C; BOHNERT, H. J. Genomic approaches to plant stress tolerance. **Current Opinion in Plant Biology**, Amsterdam, v. 3, n. 2, p. 117-24, 2000.

DAMATTA, F. M; RAMALHO, J. D. C. Impacts of drought and temperature stress on coffee physiology and production: a review. **Brazilian Journal of Plant Physiology**, Campos dos Goytacazes, v. 18, n. 1, p. 55–81, 2006.

Davis AP, Gole TW, Baena S, Moat J. The impact of climate change on indigenous Arabica coffee (*Coffea arabica*): predicting future trends and identifying priorities. **PLoS ONE**. 2012;

DAVIS ET.AL . The impact of climate change on indigenous Arabica coffee (*Coffea arabica*): predicting future trends and identifying priorities. **PLoS ONE**. 2012;

DEGNER, J. F. et al. Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. **Bioinformatics**, Oxford, v. 25, n. 24, p. 3207–12, Dec. 2009.

DENOEUDE, F. et al. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. **Science**, New York, v. 345, n. 6201, p. 1181-1185, 2014.

DEREEPER, A. et al. The coffee genome hub: a resource for coffee genomes. **Nucleic Acids Research**, Oxford, v. 43, p. 1028-1035, 2015.

DESCROIX, F; SNOECK; J. **Environmental factors suitable for coffee cultivation**. In: *Coffee: Growing, Processing, Sustainable Production: A Guidebook for Growers, Processors, Traders, and Researchers*. (ed J. N. Wintgens), Wiley-VCH Verlag GmbH, Weinheim, Germany. p. 164–177, 2008.

DESCROIX, F; SNOECK; J. **Environmental factors suitable for coffee cultivation**. In: *Coffee: Growing, Processing, Sustainable Production: A Guidebook for Growers, Processors, Traders, and Researchers*. (ed J. N. Wintgens), Wiley-VCH Verlag GmbH, Weinheim, Germany. p. 164–177, 2008.

DUBOUZET, J. G. et al. Os *DREB* genes in rice, *Oryza sativa* L., encode transcription activators that function in drought, high-salt and cold responsive gene expression. **The Plant Journal**, Oxford, v. 33, n. 4, p. 751-763, Feb. 2003.

FANG, Z. W. et al. Characterization of *FeDREB1* promoter involved in cold- and drought-inducible expression from common buckwheat (*Fagopyrum esculentum*). **Genetics and Molecular Research**, Ribeirão Preto, v. 14, n. 3, p. 7990–8000, 2015.

FAROOQ, M. et al. Plant drought stress: effects, mechanisms and management. **Agronomy for Sustainable Development**, Paris, v. 29, n. 1, p. 185–212, 2009.

FAZUOLI, L. C. **Genética e melhoramento do cafeeiro**. In: RENA, A. B. et al. (Ed.). *Cultura do cafeeiro: fatores que afetam a produtividade*. Piracicaba: Associação Brasileira para Pesquisa da Potassa e do Fósforo, p. 87-113, 1986.

FERNANDEZ, D. et al. Coffee (*Coffea arabica* L.) genes early expressed during infection by

the rust fungus (*Hemileia vastatrix*). **Molecular Plant Pathology**, Lancaster, v. 5, n. 6, p. 527–536, 2004.

FREIRE, Luciana Pereira et al. Análise da expressão do gene manose 6 fosfato redutase em cafeeiros submetidos ao déficit hídrico. 2013.

FUJITA M. et al. Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. **Current Opinion in Plant Biology**, Amsterdam, v. 9, n. 4, p. 436–442, 2006.

GALVEZ-VALDIVIESO, G. et al. The High Light Response in *Arabidopsis* involves ABA Signaling between vascular and Bundle Sheath Cells. **Plant Cell Reports**, Berlin Heidelberg, v. 21, n. 7, p. 2143–2162, July 2009.

GARCION, E. et al. RNA mutagenesis and sporadic prion diseases. **Journal of Theoretical Biology**, Amsterdam, v. 230, n. 2, p. 271–4, Sep. 2004.

GASCÓ, A; NARDINI, A; SALLES, S. Resistance to water flow through leaves of *Coffea arabica* is dominated by extra-vascular tissues. **Functional Plant Biology**, v. 31, p. 1161–1168, 2004.

GAY, Carlos et al. Potential impacts of climate change on agriculture: a case of study of coffee production in Veracruz, Mexico. **Climatic Change**, v. 79, n. 3, p. 259–288, 2006.

GEROMEL, Clara et al. Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. **Journal of Experimental Botany**, v. 57, n. 12, p. 3243–3258, 2006.

GILL, S. S; TUTEJA, N. Reactive Oxygen Species and antioxidant machinery in abiotic stress tolerance in crop plants. **Plant Physiology and Biochemistry**, Paris, v. 48, n. 12, p. 909–930, 2010.

GONGORA, C. E; MOLINA, D; ACUÑA, R. Heterologous gene approaches to introduce resistance to coffee insects and diseases in Colombia. In: PROCEEDINGS OF THE 25TH ASIC INTERNATIONAL CONFERENCE ON COFFEE SCIENCE, 2015, Armenia, Colombia. **Anais...** Armenia: 2015.

GRIGOROVA, B. et al. Combined drought and heat stress in wheat: changes in some heat shock proteins. **Biologia Plantarum**, Basel, v. 55, n. 1, p. 105–111, 2011.

GUTTMAN, M. et al. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. **Nature biotechnology**, New York, v. 28, n. 5, p. 503–10, May 2010.

HASANUZZAMAN, M. et al. Physiological, Biochemical, and Molecular Mechanisms of Heat Stress Tolerance in Plants. **International Journal of Molecular Sciences**, Basel, v. 14, n. 5, p. 9643–9684, 2013.

HEIDARVAND, L; AMIRI, R. M. What happens in plant molecular responses to cold stress? **Acta Physiologiae Plantarum**, Berlin Heidelberg, v. 32, n. 3, p. 419–431, 2010.

HINNIGER, Cécile et al. Isolation and characterization of cDNA encoding three dehydrins expressed during *Coffea canephora* (Robusta) grain development. **Annals of Botany**, v. 97, n. 5, p. 755-765, 2006.

HOGEWEG, P. The roots of bioinformatics in theoretical biology. **PLoS Computational Biology**, Cambridge, v. 7, n. 3, p. e1002021, Mar. 2011.

HUANG, G. T. et al. Signal transduction during cold, salt, and drought stresses in plants. **Molecular Biology Reports**, Basel, v. 39, n. 2, p. 969–987, 2012.

HUNG, S. H; YU, C. W; LIN, C. H. Hydrogen peroxide functions as a stress signal in plants. **Botanical Bulletin of Academia Sinica**, Taipei, v. 46, n. 1, p. 1-10, 2005.

ITO, Y. et al. Functional analysis of rice DREB1/CBF-type transcription factors involved in cold-responsive gene expression in transgenic rice. **Plant & Cell Physiology**, Kyoto, v. 47, n. 1, p. 141–153, 2006.

JOHNSON, M. P. et al. Photoprotective Energy Dissipation Involves the Reorganization of Photosystem II Light-Harvesting Complexes in the Grana Membranes of Spinach Chloroplasts. **The Plant cell**, Waterbury, v. 23, n. 4, p. 1468-1479, 2011.

KLIMCZAK, L. J. et al. Reconstitution of Arabidopsis casein kinase II from recombinant subunits and phosphorylation of transcription factor GBF1. **The Plant Cell**, Waterbury, v. 7, n. 1, p. 105–115, 1995.

KNIGHT, H; KNIGHT, M. R. Abiotic stress signalling pathways: specificity and cross-talk. **Trends in Plant Science**, Kidlington, v. 6, n. 6, p. 262-267, June 2001

KOTAK, S. et al. Complexity of the heat stress response in plants. **Current Opinion in Plant Biology**, Amsterdam, v. 10, n. 3, p. 310–316, 2007.

KRASENSKY, J; JONAK, C. Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. **Journal of Experimental Botany**, Oxford, v 63, n. 4, p. 1593-1608, 2012.

KUDOYAROVA, G. et al. Involvement of root ABA and hydraulic conductivity in the control of water relations in wheat plants exposed to increased evaporative demand. **Planta**, Berlin, v. 233, n. 1, p. 87–94, 2011.

LALONDE, E. et al. RNA sequencing reveals the role of splicing polymorphisms in regulating human gene expression. **Genome research**, Cold Spring, v. 21, n. 4, p. 545–54, Apr. 2011.

LASHERMES, P. et al. Molecular characterisation and origin of the *Coffea arabica* L. genome. **Molecular and General Genetics**, New York, v. 261, n. 2, p. 259-266, Mar. 1999.

LATA, C; PRASAD, M. Role of DREBs in regulation of abiotic stress responses in Plants. **Journal of Experimental Botany**, Oxford, v. 62, n. 14, p. 4731–4748, 2011.

LEE, J. et al. Analysis of transcription factor HY5 genomic binding sites revealed its hierarchical role in light regulation of development. **The Plant Cell**, Waterbury, v. 9, n. 3, p.

731–749, 2007.

LEPELLEY, Maud et al. Coffee cysteine proteinases and related inhibitors with high expression during grain maturation and germination. **BMC Plant Biology**, v. 12, n. 1, p. 1-16, 2012.

LIM, C. J. et al. Over-expression of the *Arabidopsis* DRE/CRT-binding transcription factor *DREB2C* enhances thermotolerance. **Biochemical and Biophysical Research Communications**, Amsterdam, v. 362, n. 2, p. 431–436, 2007.

Mahajan S, Tuteja N. Cold, salinity and drought stresses: An overview. *Arch Biochem Biophys* 2005; 444:139 – 158.

MAHALINGAM, R; FEDOROFF, N. Stress response, cell death and signaling: the many faces of reactive oxygen species. **Physiologia Plantarum**, Malden, v.119, n. 1, p. 56-68, 2003.

MAHER, C. A. et al. Transcriptome sequencing to detect gene fusions in cancer. **Nature**, London, v. 458, n. 7234, p. 97–101, Mar. 2009.

MARRACCINI, P. et al. Differentially expressed genes and proteins upon drought acclimation in tolerant and sensitive genotypes of *Coffea canephora*. **Journal of Experimental Botany**, Oxford, v. 63, n. 11, p. 4191-4212, Apr. 2012.

MARRACCINI, Pierre. Gene expression in coffee. **Progress in Botany Vol. 82**, p. 43-111, 2020.

MITTLER, R. Oxidative stress, antioxidants and stress tolerance. **Trends in Plant Science**, Kidlington, v. 7, n. 9, p. 405-410, 2002.

MIZOI, J; SHINOZAKI, K; YAMAGUCHI-SHINOZAKI, K. AP2/ERF family transcription factors in plant abiotic stress response. **Biochimica and Biophysica Acta**, Amsterdam, v. 1819, n. 2, p. 86–96, 2012.

MOFATTO, Luciana Souto et al. Identification of candidate genes for drought tolerance in coffee by high-throughput sequencing in the shoot apex of different *Coffea arabica* cultivars. **BMC Plant Biology**, v. 16, n. 1, p. 1-18, 2016.

MØLLER, I. M; SWEETLOVE, L. J. ROS signalling – specificity is required. **Trends in Plant Science**, Kidlington, v. 15, n. 7, p. 370-374, 2010.

MONDEGO, J. M. C. et al. An EST-based analysis identifies new genes and reveals distinctive gene expression features of *Coffea arabica* and *Coffea canephora*. **BMC Plant Biology**, London, v. 11, n. 30, p. 30-34, Feb. 2011.

MONTAGNON, C. et al. Amélioration génétique du caféier *Coffea canephora* Pierre: connaissances acquises, stratégies et perspectives. **Cahiers Agricultures**, Elmsford, v. 21, n. 2-3, p. 143-153, 2012.

MORTAZAVI, A. et al. Mapping and quantifying mammalian transcriptomes by RNA-Seq. **Nature Methods**, New York, v. 5, n. 7, p. 621-628, 2008.

NAKANO, T. et al. Genome-wide analysis of the *ERF* gene family in *Arabidopsis* and rice. **Plant Physiology**, Rockville, v. 140, n. 2, p. 411–432, 2006.

NAMBARA, E; MARION-POLL, A. Abscisic acid biosynthesis and catabolism. **Annual Review of Plant Biology**, Palo Alto, v. 56, p. 165–185, 2005.

NIYOGI, K. K. Photoprotection revisited: genetic and molecular approaches. **Annual Review of Plant Biology**, Palo Alto, v. 50, p. 333–359, 1999.

NOVILLO, F; MEDINA, J; SALINAS, J. *Arabidopsis* CBF1 and CBF3 have a different function than CBF2 in cold acclimation and define different gene classes in the CBF regulon. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 104, n. 52, p. 21002–21007, 2007.

OVALLE-RIVERA, Oriana et al. Projected shifts in *Coffea arabica* suitability among major global producing regions due to climate change. **PloS one**, v. 10, n. 4, p. e0124155, 2015.

PAN, Q. et al. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. **Nature genetics**, New York, v. 40, n. 12, p. 1413–5, Dec. 2008.

PEI, B. et al. The GENCODE pseudogene resource. **Genome Biology**, London, v. 13, p. R51, 2012.

PINHEIRO, H. A. et al. Drought tolerance is associated with rooting depth and stomatal control of water use in clones of *Coffea canephora*. **Annals of Botany**, Oxford, v. 96, n. 1, p. 101-108, 2005.

PONCET, V. et al. SSR mining in coffee tree EST databases: potential use of EST–SSRs as markers for the *Coffea* genus. **Molecular Genetics Genomics**, Basel, v. 276, n. 5, p. 436–449, 2006.

POP, M; SALZBERG, S. L. Bioinformatics challenges of new sequencing technology. **Trends in genetics**, Cambridge, v. 24, n. 3, p. 142–9, Mar. 2008.

QUAIL, P. H. Phytochrome photosensory signalling networks. **Nature Reviews Molecular Cell Biology**, London, v. 3, p. 85–93, 2002b.

QUEIROZ, C. G. S. et al. Chilling-induced changes in membrane fluidity and antioxidant enzyme activities in *Coffea arabica* L. roots. **Biologia Plantarum**, Basel, v. 41, n. 3, p. 403-413, 1998.

RAMALHO, J. C. et al. Cold acclimation ability of photosynthesis among species of the tropical *Coffea* genus. **Plant Biology**, v. 5, n. 6, p. 631-641, 2003.

RAMALHO, J. C. et al. Nitrogen dependent changes in antioxidant systems and in fatty acid composition of chloroplast membranes from *Coffea arabica* L. plants submitted to high irradiance. **Plant Science**, Amsterdam, v. 135, n. 2, p. 115-124, 1998.

ROBINSON, M. D; MCCARTHY, D. J; SMYTH, G. K. edgeR: A Bioconductor package for

differential expression analysis of digital gene expression data. **Bioinformatics**, Oxford, England, v. 26, n. 1, p. 139-40, 2010.

ROSSEL, J. B; WILSON, I. W; POGSON, B. J. Global Changes in Gene Expression in Response to High Light in Arabidopsis. **Plant Physiology**, Rockville, v. 130, n. 3, p. 1109–1120, 2002.

SAIBO, N. J. M; LOURENCO, T; OLIVEIRA, M. M. Transcription factors and regulation of photosynthetic and related metabolism under environmental stresses. **Annals of Botany**, Oxford, v. 103, n. 4, p. 609–623, 2009.

SAKUMA, Y. et al. DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis DREBs*, transcription factors involved in dehydration- and cold-inducible gene expression.

SANTOS, Adriana Brombini dos; MAZZAFERA, Paulo. Aquaporins and the control of the water status in coffee plants. **Theoretical and Experimental Plant Physiology**, v. 25, n. 2, p. 79-93, 2013.

SANTOS, Adriana Brombini; MAZZAFERA, Paulo. Dehydrins are highly expressed in water-stressed plants of two coffee species. **Tropical plant biology**, v. 5, n. 3, p. 218-232, 2012.

SEKI, M. et al. Molecular responses to drought, salinity and frost: common and different paths for plant protection. **Current Opinion of Biotechnology**, London, v. 14, n. 2, p. 194-199, Apr. 2003.

SEO, M. et al. The Arabidopsis aldehyde oxidase 3 (*AAO3*) gene product catalyzes the final step in abscisic acid biosynthesis in leaves. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 97, n. 23, p. 12908-12913, 2000.

SHINOZAKI, K; YAMAGUCHI-SHINOZAKI, K. Gene expression and signal transduction in water-stress response. **Plant Physiology**, Rockville, v. 115, n. 2, p. 327±334, 1997.

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

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

SIMKIN, Andrew J. et al. An investigation of carotenoid biosynthesis in *Coffea canephora* and *Coffea arabica*. **Journal of plant physiology**, v. 165, n. 10, p. 1087-1106, 2008.

SIMKIN, Andrew J. et al. Carotenoid profiling and the expression of carotenoid biosynthetic genes in developing coffee grain. **Plant Physiology and Biochemistry**, v. 48, n. 6, p. 434-442, 2010.

SIRICHANDRA, C. et al. The guard cell as a single-cell model towards understanding drought tolerance and abscisic acid action. **Journal of Experimental Botany**, Oxford, v. 60, n. 5, p. 1439–1463, 2009.

STOCKINGER, E. J; GILMOUR, S. J; THOMASHOW, M. F. *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat DRE, a *cis*-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 94, p. 1035–1040, 1997.

TAIZ, L; ZEIGER, E. **Fisiologia Vegetal**. Porto Alegre: Artmed; 2004.

THIOUNE, El-Hadji et al. A humidity shock leads to rapid, temperature dependent changes in coffee leaf physiology and gene expression. **Tree Physiology**, v. 37, n. 3, p. 367-379, 2017.

THIOUNE, El-Hadji et al. Temperature impacts the response of *Coffea canephora* to decreasing soil water availability. **Tropical Plant Biology**, v. 13, n. 3, p. 236-250, 2020.

THOMASHOW, M. F. Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. **Annual Review of Plant Physiology and Plant Molecular Biology**, Palo Alto, v. 50, p. 571-599, June, 1999.

TORRES, Luana Ferreira. Expression of the CcDREB1D promoter in *Coffea arabica*: functional genomics and transcriptome analysis. 2017.

TORRES, T. T. et al. Gene expression profiling by massively parallel sequencing. **Genome Research**, Nova York, v. 18, n. 1, p. 172–7, 2008.

TRAAS, J; VERNOUX, T. The shoot apical meristem: the dynamics of a stable structure. **Philosophical Transactions of the Royal Society of London**, London, v. 357, n. 1422, p. 737–47, 2002.

ULM, R; BAUMANN, A; ORAVECZ, A. et al. Genome-wide analysis of gene expression reveals function of the bZIP transcription factor HY5 in the UV-B response of *Arabidopsis*. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 101, n. 5, p. 1397-1402, 2004.

VAAHTERA, L; BROSCHE, M. More than the sum of its parts. How to achieve a specific transcriptional response to abiotic stress. **Plant Science**, Amsterdam, v. 180, n. 3, p. 421–430, 2011.

VIEIRA, L. G. E. et al. Brazilian coffee genome project: an EST-based genomic resource. **Brazilian Journal of Plant Physiology**, Campos dos Goytacazes, v. 18, n. 1, p. 95-108, 2006.

VINOCUR, B; ALTMAN, A. Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. **Current Opinion in Biotechnology**, Amsterdam, v. 16, n. 2, p. 123-132, 2005.

VOIGT, Christian et al. In-depth analysis of the distinctive effects of norflurazon implies that tetrapyrrole biosynthesis, organellar gene expression and ABA cooperate in the GUN-type of plastid signalling. **Physiologia Plantarum**, v. 138, n. 4, p. 503-519, 2010.

VRANOVÁ, E; INZÉ, D; VAN BREUSEGEM, F. Signal transduction during oxidative stress. **Journal of Experimental Botany**, Oxford, v. 53, n. 372, p.1227-1236, 2002.

WAN, X. R.; LI, L. Regulation of ABA level and water-stress tolerance of *Arabidopsis* by ectopic expression of a peanut 9-cis-epoxycarotenoid dioxygenase gene. **Biochemical and Biophysical Research Communications**, Basel, v.347, n. 4, p. 1030-1038, 2006.

WANG, W; VINOCUR, B; ALTMAN, A. Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. **Planta**, Berlin, v. 218, n. 1, p. 1-14, 2003.

XING, Y. et al. MKK5 Regulates High Light-Induced Gene Expression of Cu/Zn Superoxide Dismutase 1 and 2 in *Arabidopsis*. **Plant Cell Physiology**, Kyoto, v. 54, n. 7, p. 1217–1227, 2013.

XIONG, L; SCHUMAKER, K; ZHU, J-K. Cell Signaling during Cold, Drought, and Salt Stress. **The Plant Cell**, Waterbury, v. 14, n. 1, p. S165-S183, 2002.

YAMAGUCHI-SHINOZAKI, K; SHINOZAKI, K. A novel cis-Acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low temperature, or high-salt stress. **The Plant Cell**, Waterbury, v. 6, n. 2, p. 251-264, 1994.

YAMAGUCHI-SHINOZAKI, K; SHINOZAKI, K. Characterization of the expression of a desiccation-responsive rd29 gene of *Arabidopsis thaliana* and analysis of its promoter in transgenic plants. **Molecular and General Genetics**, Basel, v. 236, n. 2, p. 331–340, 1993.

YAMAGUCHI-SHINOZAKI, K; SHINOZAKI, K. Organization of *cis*-acting regulatory elements in osmotic- and cold-stress-responsive promoters. **Trends in Plant Science**, Kidlington, v. 10, n. 2, p. 88–94, 2005.

ZANDKARIMI, H. et al. Analyzing the Expression Profile of *AREB/ABF* and *DREB/CBF* Genes under Drought and Salinity Stresses in Grape (*Vitis vinifera* L.). **PLoS ONE**, Cambridge, v. 10, n. 7, p. e0134288, 2015.

ZARKA, D. G. et al. Cold induction of *Arabidopsis CBF* genes involves multiple (inducer of CBF expression) promoter elements and a cold-regulatory circuit that is desensitized by low temperature. **Plant Physiology**, Campinas, v. 133, n. 2, p. 910–918, 2003.