



**TATIANA SANTOS COSTA**

**ANÁLISE DO PERFIL TRANSCRIPTÔMICO E  
PROTEÔMICO DE RAÍZES DE DIFERENTES  
CLONES DE *Coffea canephora* EM CONDIÇÕES  
DE DÉFICIT HÍDRICO**

**LAVRAS - MG**

**2014**

**TATIANA SANTOS COSTA**

**ANÁLISE DO PERFIL TRANSCRIPTÔMICO E PROTEÔMICO DE  
RAÍZES DE DIFERENTES CLONES DE *Coffea canephora* EM  
CONDIÇÕES DE DÉFICIT HÍDRICO**

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

Orientador

Dr. Alan Carvalho Andrade

Coorientador

Dr. Pierre Marraccini

**LAVRAS - MG**

**2014**

**Ficha Catalográfica Elaborada pela Coordenadoria de Produtos e  
Serviços da Biblioteca Universitária da UFLA**

Costa, Tatiana Santos.

Análise do perfil transcriptômico e proteômico de raízes de diferentes clones de *Coffea canephora* em condições de déficit hídrico / Tatiana Santos Costa. – Lavras : UFLA, 2014.

235 p. : il.

Tese (doutorado) – Universidade Federal de Lavras, 2014.

Orientador: Alan Carvalho Andrade.

Bibliografia.

1. Cafeeiro. 2. Raiz. 3. Tolerância à seca. 4. Expressão gênica. 5. Espectrometria. I. Universidade Federal de Lavras. II. Título.

CDD – 633.7358

TATIANA SANTOS COSTA

**ANÁLISE DO PERFIL TRANSCRIPTÔMICO E PROTEÔMICO DE  
RAÍZES DE DIFERENTES CLONES DE *Coffea canephora* EM  
CONDIÇÕES DE DÉFICIT HÍDRICO**

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

APROVADA em 30 de abril de 2014.

PhD. Antonio Chalfun Júnior	UFLA
Dr. Hugo Bruno Correa Molinari	EMBRAPA
Dr. Luciano Vilela Paiva	UFLA
Dra. Polyana Kelly Martins	EMBRAPA

Dr. Alan Carvalho Andrade  
Orientador

Dr. Pierre Marraccini  
Coorientador

LAVRAS – MG

2014

*A Deus e a todos que me fizeram acreditar nesse sonho e que trabalharam muito para que eu pudesse realizá-lo.*

**DEDICO**

## AGRADECIMENTOS

Primeiramente agradeço a Deus, pelo dom da vida e pelos impulsos de fé e força durante todo o percurso, sempre iluminando os meus caminhos.

A toda minha família, agradeço pela torcida, compreensão e carinho e por não medirem esforços para a concretização deste meu trabalho. À minha mãe, minha avó e aos meus irmãos, pela paciência, compreensão e amor incondicional. Aos meus tios e primos, em especial a minha bonequinha Maria Luiza, meu muito obrigada a vocês por sempre me passarem mensagens de carinho e força, com vocês aprendi a acreditar nos meus sonhos e batalhar para alcançá-los.

Aos meus grandes amigos, que me mostraram que a distância não separa e não limita uma verdadeira amizade, em especial Ana Paula, Cassia, Eliana e Very e aos amigos da UFS: Franklin, Ivan, Juliano, Márcia e Katy, pelas mensagens de carinho e força. Vocês são anjos que escolhi para ter sempre ao meu lado durante toda a vida!! Muito obrigada!

Aos amigos e pesquisadores da Embrapa Tabuleiro Coteiros: Ana Ledo, Ana Veruska e Leandro Diniz, pelo incentivo para meu ingresso neste programa.

À professora Silmara Pantaleão, pelo incentivo e apoio sempre prestados.

À Universidade Federal de Lavras e aos professores e colegas do Departamento de Biotecnologia Vegetal, em especial ao professor Luciano, coordenador do curso, pela oportunidade para a realização do curso de doutorado.

Aos membros da banca examinadora, Dr<sup>a</sup> Polyana Kelly Martins, Dr. Antonio Chalfun Júnior, Dr. Hugo Molinari e Dr. Luciano Vilela Paiva por aceitarem o convite para fazer parte da banca e contribuírem construtivamente com o meu trabalho.

Aos amigos da UFLA em especial Barretinha, Kátia, Patrick, Larissa, Mayara, Marcelo e Samuel, pela amizade e conversas agradáveis.

As novas amigadas, Sandra, Michelle e Fernanda que compartilharam os anseios de um futuro brilhante, lutando contra todos os empecilhos e desafios do mundo da pesquisa. Jean, Rafa e Valquíria pela companhia agradável, sempre tentando suprir a saudade causada pela distância dos entes queridos. Mas que acima de tudo me proporcionaram muita alegria e experiência de vida.

Ao meu orientador Dr. Alan Carvalho Andrade, pela confiança, genialidade, compreensão, ensinamentos, apoio, amizade e alegria, contribuindo de forma significativa para minha capacitação profissional.

Ao Pierre Marracinni, pela coorientação, simpatia, paciência, pelos ensinamentos e sugestões e principalmente pela humildade sempre demonstrada.

A toda equipe do LGM: Edriana, Érica, Fernanda, Jean, Karol, Michelle, Naty, Rafa, Sinara e Vanessa, por todo o carinho. Muito obrigada por todo apoio nos experimentos e nos momentos difíceis, vocês são extraordinários, sem vocês tudo teria sido muito mais difícil!

Aos funcionários da Embrapa Cenargen, pela prestatividade, atenção e carinho que sempre recebi. Em especial aos amigos que a EMBRAPA me deu de presente, Ada, Bárbara, Éder, Felipe, José Cardoso, Lílian, e ao colega Jorge, pelos ensinamentos, questionamentos, conversas, auxílio e diversão durante todo esse período!

Ao Dr. Carlos Bloch Júnior pela disponibilidade do espaço laboratorial e ao Dr. André Murad, pela disponibilidade e ensinamentos.

Ao Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café e a Embrapa Recursos Genéticos e Biotecnologia, pela infraestrutura disponibilizada, ao Conselho Nacional de Desenvolvimento Científico – CNPq e ao INCT-Café pelo apoio financeiro.

A todos que, direta ou indiretamente, contribuíram para o êxito deste trabalho. Obrigada!

*“De tudo, ficaram três coisas: a certeza de que ele estava sempre começando, a certeza de que era preciso continuar e a certeza de que seria interrompido antes de terminar. Fazer da interrupção um caminho novo. Fazer da queda um passo de dança, do medo uma escada, do sono uma ponte, da procura um encontro.”*

*(Fernando Sabino)*

## RESUMO GERAL

O café é uma das principais commodities agrícolas do mundo, sendo o Brasil o maior produtor. A seca é um problema frequente no país, uma vez que é considerado um fator limitante que possibilita oscilações na produtividade do cafeeiro. Em resposta ao déficit hídrico as plantas ativam mecanismos que auxiliam suportar períodos de seca, promovendo alterações na fotossíntese, respiração, translocação, absorção de íons, o metabolismo de nutrientes e hormônios. O presente trabalho tem por objetivo avaliar a expressão diferencial de genes e proteínas em raízes de clones de *C. canephora* var. Conilon, sendo o clone 22 (sensível à seca) e os clones 14, 73 e 120 (tolerantes à seca), cultivados em condições controladas em condições irrigada e sob estresse hídrico. Para cada clone e regime hídrico foi extraído o RNA total e avaliado o perfil transcriptômico e proteômico das raízes por meio do sequenciamento 454 e da espectrometria LC-MS, respectivamente. Os resultados apresentados possibilitaram uma melhor compreensão de alguns mecanismos de resposta em raízes em clones *C. canephora* quanto a tolerância ao déficit hídrico, por meio da diferença de expressão gênica e dos níveis de proteínas entre as condições avaliadas (I e NI). Foi possível verificar que, entre os clones tolerantes há diferentes mecanismos de respostas ao estresse. O pirossequenciamento 454 permitiu gerar mais de 4 milhões de *reads* a partir de 08 bibliotecas de raízes de *C. canephora* provenientes dos clones nas duas condições, o que possibilitou a ampliação dos dados de sequências para este tecido. Vários genes candidatos foram identificados e avaliados quanto à resposta da planta ao déficit hídrico, como em relação a biossíntese do ABA e aos estresses osmóticos e oxidativos. A técnica LC-MS permitiu a identificação de 598 proteínas, o que representa cerca de 1% do proteoma conhecido de *C. canephora*. Para algumas proteínas foi possível identificar diferentes formas alélicas. De forma geral, a resposta dos clones de *C. canephora* ao déficit hídrico foi evidenciada neste trabalho. Na maioria dos casos há correspondência entre os resultados das técnicas “Omicas” utilizadas. Entretanto, ainda é necessário o aprimoramento das técnicas para identificação proteica, como por meio da utilização de sequências genótipo-específicas, considerada uma alternativa para futuras análises integradas, possibilitando a compreensão dos mecanismos de regulação e dos pontos de controle do fluxo da informação gênica. Em complemento aos trabalhos anteriores, no qual foram utilizadas folhas, também foi possível identificar vários genes de tolerância à seca em raízes de cafeeiro.

Palavras-chave: Cafeeiro. Raiz. Tolerância à seca. Expressão gênica. Espectrometria.

## GENERAL ABSTRACT

Coffee is one of the main agricultural commodities in the world, with Brazil as the largest producer. Drought is a frequent issue in the country, since it is considered a limiting factor that allows oscillations in coffee productivity. In response to the water deficit, the plants activate mechanisms that aid in supporting drought periods, promoting changes in photosynthesis, respiration, translocation, ion absorption and in nutrient and hormone metabolism. The present work has the objective of evaluating the differential expression of genes and proteins in clones of *C. canephora* var. Conilon roots, with clone 22 (sensitive to drought) and clones 14, 73 and 120 (tolerant to drought), cultivated under controlled conditions and water stress. For each clone and water regime, total RNA was extracted and the transcriptomic and proteomic profile of the roots were evaluated by means of the 454 and LC-MS sequencing, respectively. The results presented allowed the better understanding of some of the response mechanisms in *C. canephora* clone roots regarding the tolerance to water deficit by means of the difference in gene expression and of the levels of proteins between the conditions evaluated (I and NI). It was possible to verify that, among the tolerant clones, there are different stress response mechanisms. The 454 pyrosequencing allowed us to generate more than four million reads from eight libraries of *C. canephora* roots derived from the clones in both conditions, which allowed the extension of the data for this tissue. Many candidate genes were identified and evaluated regarding the response of the plant to water deficit, such as in regard to the ABA biosynthesis and the osmotic and oxidative stresses. The LC-MS technique allowed the identification of 598 proteins, which represents around 1% of the proteome known for *C. canephora*. For some of the proteins, it was possible to identify different allele forms. In general, the response of the *C. canephora* clones to water deficit was evidenced in this work. In most cases, there is correspondence between the results of the used "Omics" techniques. However, the improvement of the techniques for protein identification is still necessary, such as by means of the use of genotype-specific sequences, considered an alternative for future integrated analyses, allowing the understanding of regulation mechanisms and of the control points of genic information flow. In complement to previous works, in which leaves were used, it was also possible to identify many drought tolerance genes in coffee roots.

Keywords: Coffee plant. Root. Drought tolerance. Gene expression. Spectrometry.

## LISTA DE FIGURAS

### CAPÍTULO 1

- Figura 1 Regiões produtoras de café. Os dez países maiores produtores de café estão destacados em amarelo .....34
- Figura 2 Média da evolução da produtividade do café beneficiado no Brasil. Produção em milhões de sacas/ano nos últimos 15 anos. Segundo estimativas, a produção de 2014 é considerada de bienalidade positiva .....35
- Figura 3 Grupos de distribuição da diversidade de *Coffea canephora* .....38
- Figura 4 Sistemas radiculares de quatro clones de *C. canephora* cultivados com restrição hídrica .....41
- Figura 5 “Ômicas” - cascata de eventos para a expressão de um gene 49

### CAPÍTULO 2

- Figura 1 Evolução da perda de água nas folhas dos clones de *C. canephora*. Dados baseados no potencial hídrico de antemanhã ( $\Psi_{am}$ ) .....83
- Figura 2 Gráfico de dispersão dos genes no Cromossomo 0, com base nos resultados de pirosequenciamento 454 .....97
- Figura 3 Gráfico de dispersão dos genes no Cromossomo 1, com base nos resultados de pirosequenciamento 454 .....99
- Figura 4 Gráfico de dispersão dos genes no Cromossomo 2, com base nos resultados de pirosequenciamento 454 ..... 100

Figura 5	Gráfico de dispersão dos genes no Cromossomo 3, com base nos resultados de pirosequenciamento 454.....	101
Figura 6	Gráfico de dispersão dos genes no Cromossomo 4, com base nos resultados de pirosequenciamento 454.....	103
Figura 7	Gráfico de dispersão dos genes no Cromossomo 5, com base nos resultados de pirosequenciamento 454.....	104
Figura 8	Gráfico de dispersão dos genes no Cromossomo 6, com base nos resultados de pirosequenciamento 454.....	106
Figura 9	Gráfico de dispersão dos genes no Cromossomo 7, com base nos resultados de pirosequenciamento 454.....	107
Figura 10	Gráfico de dispersão dos genes no Cromossomo 8, com base nos resultados de pirosequenciamento 454.....	108
Figura 11	Gráficos de dispersão dos genes no Cromossomo 9, com base nos resultados de pirosequenciamento 454.....	110
Figura 12	Gráficos de dispersão dos genes no Cromossomo 10, com base nos resultados de pirosequenciamento 454.....	111
Figura 13	Gráficos de dispersão dos genes no Cromossomo 11, com base nos resultados de pirosequenciamento 454.....	112
Figura 14	Genes com maiores valores de Q entre todos os clones tolerantes (14, 73 e 120). O valor Q foi obtido pelo quociente do valor de FI para cada clone em relação ao valor de FI para o clone sensível (22) .....	120
Figura 15	Gráficos ilustrativos da expressão de genes candidatos que apresentaram valores de FI mais significativos para o clone 14 .....	122

Figura 16	Gráficos ilustrativos da expressão dos genes candidatos que apresentaram valores de FI mais significativos para o clone 73 .....	124
Figura 17	Gráficos ilustrativos da expressão dos genes candidatos que apresentaram valores de FI mais significativos para o clone 120 .....	125
Figura 19	Gráficos ilustrativos da expressão dos genes candidatos que apresentaram valores de FI mais significativos para o clone 22 .....	127
Figura 20	Perfis de expressão do gene <i>CcNCED3</i> (GSCOCT00021488001).....	130
Figura 21	Perfis de expressão do gene <i>CcPYL7.2</i> (GSCOCT00027751001).....	132
Figura 22	Perfis de expressão do gene <i>CcPP2C-1</i> (GSCOCT00021820001).....	132
Figura 23	Perfis de expressão do <i>CcAREB</i> (GSCOCT00034982001).....	133
Figura 24	Perfis de expressão do <i>CcNAC</i> (A e B) e <i>CcNAC4</i> (C) na qPCR e <i>in silico</i> , em raízes dos clones de <i>C. canephora</i> 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado).....	134
Figura 25	Perfis de expressão do gene <i>CcLEA</i> codificando uma chaperona LEA (GSCOCT00022121001) .....	136

Figura 26	Perfis de expressão da Dehidrina na qPCR e <i>in silico</i> , em raízes dos clones de <i>C. canephora</i> 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado).....	136
Figura 27	Perfis de expressão do gene <i>CcABA-8-hidrolase</i> (GSCOCT00024390001).....	138
Figura 28	Perfis de expressão do gene <i>CcPRX1</i> que codifica a Peroxidase (GSCOCT00016870001) .....	140
Figura 29	Perfis de expressão <i>CcGST</i> - Glutathiona-S-Transferase (GSCOCT00015916001).....	141
Figura 30	Perfis de expressão do transcrito <i>CcMJE</i> (GSCOCT00039836001) que codifica para a proteína Metil Jasmonato esterase.....	143
Figura 31	Perfis de expressão do gene <i>CcNADP-S6PDH</i> (GSCOCT00011142001) que codifica a NADP dependente de D-Sorbitol-6-fosfatase desidrogenase .....	144
Figura 32	Perfis de expressão do gene <i>CcGols3</i> codificando a Galactinol Sintase <i>CcGols3</i> (GSCOCT00042862001).....	145
Figura 33	Perfis de expressão gênica para o transcrito <i>CcSRP-ABA</i> (GSCOCT00033650001) que codifica a proteína de amadurecimento induzida por ABA.....	147
Figura 34	Perfis de expressão gênica para o transcrito <i>CcNRT1.4</i> (GSCOCT00031111001) que codifica a proteína de transporte de nitrato.....	148

Figura 35	Perfis de expressão do transcrito <i>CcPhLOG</i> (GSCOCT00024672001) que codifica para a proteína Citocinina ribosídeo-5'-monofosfatase fosforibohidrolase.	149
Figura 36	Perfis de expressão do transcrito <i>CcMALD3</i> (GSCOCT00040458001) que codifica para a proteína Nltp .....	150
Figura 37	Perfis de expressão do transcrito <i>CcLTP2</i> (GSCOCT00040232001) que codifica para um tipo de proteína Nltp.....	150
Figura 38	Perfis de expressão de transcritos que codificam proteínas fosfatases <i>CcPP3B</i> .....	152
Figura 39	Perfis de expressão do transcrito <i>CcCLXIP</i> (GSCOCT00013168001) que codifica para a uma Quitinase inibidora de Xilanase.....	153
Figura 40	Perfis de expressão do transcrito <i>CcCPI3</i> (GSCOCT00017904001) que codifica para a uma proteína inibidora de cisteinase .....	154
Figura 41	Perfis de expressão do transcrito <i>CcBSPA</i> (GSCOCT00007152001) que codifica para a uma proteína de armazenamento .....	155
Figura 42	Perfis de expressão de transcritos que codificam <i>No Hits</i> (GSCOCT00005517001, GSCOCT00015847001, GSCOCT00030679001 e GSCOCT00042816001) para as análises de qPCR e <i>in silico</i> em raízes dos clones de <i>C. canephora</i> 14, 73 e 120 (tolerantes à seca) e 22 (sensível	

à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado) .....	157
---	-----

### CAPÍTULO 3

Figura 1	Proporção de proteínas obtidas no clone 14 identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA) .....	180
Figura 2	Proporção de proteínas obtidas no clone 73 identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA) .....	181
Figura 3	Proporção de proteínas obtidas no clone 120 identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA) .....	182
Figura 4	Proporção de proteínas obtidas no clone 22 identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA) .....	182
Figura 5	Perfil diferencial de quantificação de quitinases, identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA) .....	185
Figura 6	Perfil diferencial de quantificação de fosfatases <i>CcPP3B</i> identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA) .....	187
Figura 7	Perfil diferencial de quantificação da proteína de armazenamento 07152001 identificada por LC-MS (NanoAcquity, UPLC-Waters, USA). Os valores de quantificação foram padronizados com base na quantificação da Ubiquitina.....	188

- Figura 8 Perfil diferencial de quantificação da proteína de armazenamento 07291001 identificada por LC-MS (NanoAcquity, UPLC-Waters, USA). Os valores de quantificação foram padronizados com base na quantificação da Ubiquitina..... 188
- Figura 9 Perfil diferencial de quantificação da catecol oxidase 05696001 – *CcCo-1* identificada por LC-MS (NanoAcquity, UPLC-Waters, USA). Os valores de quantificação foram padronizados com os valores de quantificação da Ubiquitina..... 189
- Figura 10 Perfil diferencial de quantificação da Gliceraldeído-3-fosfato desidrogenase 38191001 – *CcGAPDH* identificada por LC-MS (NanoAcquity, UPLC-Waters, USA). Os valores de quantificação foram padronizados com os valores de quantificação da Ubiquitina..... 189
- Figura 11 Perfil diferencial de quantificação da proteína hipotética 31009001 identificada por LC-MS (NanoAcquity, UPLC-Waters, USA). Os valores de quantificação foram padronizados com os valores de quantificação da Ubiquitina..... 190
- Figura 12 Quantificação de proteínas ribossomais 40S identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA). Resultados de quantificação relativa, com os dados normalizados pela Ubiquitina, resultantes das análises realizadas com o “ProteinLynx” (versão 5.2.2), nas amostras de raízes de *C. canephora* com a Base de dados

	do Genoma de <i>C. canephora</i> (Cenargen, Brasília-DF), como referência. Todas as proteínas foram obtidas pela extração com o protocolo de éter de petróleo.....	191
Figura 13	Perfis de quantificação proteica.....	195
Figura 14	Perfil de quantificação da proteína de armazenamento (07152001). Dados obtidos pela identificação por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando as sequências do genoma de referência do banco de dados do Genoma de <i>C. canephora</i> (Cenargen, Brasília-DF), (em preto) e os resultados da análise restrita, com somente as sequências das 15 proteínas e com as formas alélicas de cada clone (em branco).....	196
Figura 15	Perfil de quantificação da D - sorbitol - 6 - fosfatase desidrogenase (11142001). Dados obtidos a partir da identificação por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando o banco de dados com sequências das 15 proteínas nos clones (14, 22, 73 e 120) e suas formas alélicas .....	197
Figura 16	Perfis de quantificação proteica da quitinase .....	198
Figura 17	Perfis de quantificação proteica da fosfatase ácida .....	199
Figura 18	Perfil de quantificação proteica da ascorbato peroxidase 28091001. Dados obtidos a partir da identificação por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando o banco de dados com sequências das 15 proteínas nos clones (14, 22, 73 e 120) e suas formas alélicas.....	199

- Figura 19 Perfil de quantificação proteica da desidrina 37166001. Dados obtidos a partir da identificação por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando o banco de dados com sequências das 15 proteínas nos clones 14, 22, 73 e 120 e suas formas alélicas.....200
- Figura 20 Perfil de quantificação proteica da Gliceraldeído-3-fosfato desidrogenase (GAPDH) 38191001. Dados obtidos a partir da identificação por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando o banco de dados com sequências das 15 proteínas nos clones (14, 22, 73 e 120) e suas formas alélicas .....200
- Figura 21 Fracionamento do extrato proteico de raiz de *Coffea canephora* (14I-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B .....205
- Figura 22 Fracionamento do extrato proteico de raiz de *Coffea canephora* (14NI-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B .....205
- Figura 23 Fracionamento do extrato proteico de raiz de *Coffea canephora* (22I-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B .....206

Figura 24	Fracionamento do extrato proteico de raiz de <i>Coffea canephora</i> (22NI-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B. ....	206
Figura 25	Fracionamento do extrato proteico de raiz de <i>Coffea canephora</i> (73I-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B. ....	207
Figura 26	Fracionamento do extrato proteico de raiz de <i>Coffea canephora</i> (73NI-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B. ....	207
Figura 27	Fracionamento do extrato proteico de raiz de <i>Coffea canephora</i> (120I-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B. ....	208
Figura 28	Fracionamento do extrato proteico de raiz de <i>Coffea canephora</i> (120 NI-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B. ....	208

Figura 29	Fracionamento do extrato proteico de raiz de <i>Coffea canephora</i> (14I-SA) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B .....	209
Figura 30	Fracionamento do extrato proteico de raiz de <i>Coffea canephora</i> (14NI-SA) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B. ....	209
Figura 31	Fracionamento do extrato proteico de raiz de <i>Coffea canephora</i> (22I-SA) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B. ....	210
Figura 32	Fracionamento do extrato proteico de raiz de <i>Coffea canephora</i> (22NI-SA) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B. ....	210
Figura 33	Fracionamento do extrato proteico de raiz de <i>Coffea canephora</i> (73I-SA) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B. ....	211

Figura 34	Fracionamento do extrato proteico de raiz de <i>Coffea canephora</i> (73NI-SA) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B. ....	211
-----------	---	-----

#### CAPÍTULO 4

Figura 1	Análise de expressão <i>in silico</i> (A) e quantificação proteica (B) da quitinase codificada pelo gene GSCOCT00013166001 – <i>CcCLXIP</i> . Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e os da análise <i>in silico</i> por RPKM .....	218
Figura 2	Análise de expressão <i>in silico</i> (A), quantificação proteica (B) e expressão gênica por qPCR (C) da quitinase codificada pelo gene GSCOCT00013168001 – <i>CcCLXIP</i> . Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e da análise <i>in silico</i> por RPKM .....	219
Figura 3	Análise de expressão <i>in silico</i> (A), quantificação proteica (B) e expressão gênica por qPCR (C) da fosfatase ácida do tipo 3B, codificada pelo gene GSCOCT00022022001 – <i>CcPP3B</i> . Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e da análise <i>in silico</i> por RPKM.....	221

- Figura 4 Análise de expressão *in silico* (A) e quantificação proteica (B) da fosfatase ácida do tipo 3B, codificada pelo gene GSCOCT00022023001 – *CcPP3B*. Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e os da análise *in silico* por RPKM.....222
- Figura 5 Análise de expressão *in silico* (A), quantificação de proteínas (B) e expressão gênica por qPCR (C) da fosfatase ácida do tipo 3B codificada pelo gene GSCOCT00021916001 – *CcPP3B*. Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e os da expressão *in silico* por RPKM.....224
- Figura 6 Análise de expressão *in silico* (A) de quantificação proteica (B) e de expressão gênica por qPCR (C) de uma proteína de armazenamento, codificada pelo gene GSCOCT00007152001 –*CcBSPA*. Os valores de quantificação foram padronizados pela quantificação da Ubiquitina e os de expressão *in silico* por RPKM.....225
- Figura 7 Análise de expressão *in silico* (A) e quantificação (B) da proteína Kirola, codificada pelo gene GSCOCT00007291001 – *CcKirola*. Os valores de quantificação foram padronizados pela quantificação da Ubiquitina e da análise *in silico* por RPKM.....226
- Figura 8 Análise de expressão *in silico* (A), quantificação proteica (B) e expressão gênica por qPCR (C) da catecol oxidase, codificada pelo gene GSCOCT00005696001 – *CcCo-1*. Os valores de quantificação foram padronizados com a

quantificação da Ubiquitina e os valores de expressão *in silico* por RPKM.....227

Figura 9 Análise de expressão *in silico* (A) e quantificação (B) da Gliceraldeído-3-fosfato desidrogenase, codificada pelo gene GSCOCT00038191001001 – *CcGAPDH*. Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e os valores de expressão *in silico* foram por RPKM .....228

## LISTA DE TABELAS

### CAPÍTULO 1

Tabela 1	Comparativo de produção de café beneficiado no Brasil, destacando as regiões de maior produção. (S) mil sacas beneficiadas.....	36
----------	---	----

### CAPÍTULO 2

Tabela 1	Genes candidatos e respectivos <i>primers</i> , selecionados com base no Fator de Indução (FI), utilizados nos experimentos de qPCR. <b>GB</b> Número do acesso de proteínas registrados no GenBank. Os <i>primers</i> foram sintetizados utilizando o <i>software</i> Primer Express (Applied Biosystems). <b>pb</b> pares de bases. ....	87
Tabela 2	Bibliotecas geradas e número total de <i>reads</i> produzidos por meio do sequenciamento 454 a partir de raízes de <i>C. canephora</i> em condições Irrigado (I) e Não Irrigado (NI).....	95
Tabela 3	Valores de expressão para os Genes Candidatos (GCs) de tolerância ao déficit hídrico em raízes de clones de <i>C. canephora</i> . Dados obtidos com base nos valores de expressão resultantes no pirosequenciamento-454. <b>Chr</b> (cromossomo). <b>Expressão basal</b> (valor de expressão dos genes para os clones na condição I). <b>Fator de indução FI</b> (I/NI) e <b>Q</b> (FI (clones tolerantes) / FI (clone 22)). Em negrito estão os valores que foram utilizados como critério para a seleção do determinado gene. Os dados foram avaliados utilizando somente os clones tolerantes como referência.....	113
Tabela 4	Genes com maiores valores de Q entre todos os clones tolerantes (14, 73 e 120). O valor Q foi obtido pelo quociente	

	do valor de FI para cada clone em relação ao valor de FI para o clone sensível (22).....	119
Tabela 5	Genes candidatos que obtiveram valor superior no clone 14 quanto ao fator de indução (FI), em resposta ao estresse hídrico, com base nos resultados do sequenciamento obtidos pelo pirosequenciamento 454 .....	121
Tabela 6	Genes candidatos que obtiveram valor superior no clone 73 quanto ao fator de indução (FI), em resposta ao estresse hídrico, com base nos resultados do sequenciamento obtidos pelo pirosequenciamento 454 .....	123
Tabela 7	Genes candidatos que obtiveram valor superior no clone 120 quanto ao fator de indução (FI), em resposta ao estresse hídrico, com base nos resultados do sequenciamento obtidos pelo pirosequenciamento 454 .....	125
Tabela 8	Genes candidatos que obtiveram valor superior no clone 22 quanto ao fator de indução (FI), em resposta ao estresse hídrico, com base nos resultados do sequenciamento obtidos pelo pirosequenciamento 454 .....	126

### **CAPÍTULO 3**

Tabela 1	Quantitativo de proteínas identificadas utilizando a análise LC-MS.....	180
Tabela 2	Dez proteínas com maiores valores de quantificação, identificadas por LC-MS, utilizando o protocolo de solução ácida .....	183
Tabela 3	Dez proteínas com maiores valores de quantificação, identificadas por LC-MS, utilizando o protocolo de éter de petróleo.....	184

Tabela 4	Proteínas reanalisadas por LC-MS, utilizando o banco de dados do Genoma de <i>C. canephora</i> (Cenargen, Brasília-DF).....	193
----------	---	-----

#### **CAPÍTULO 4**

Tabela 1	Proteínas que apresentaram maiores valores de quantificação, com base nos resultados de LC-MS, utilizando o protocolo de éter de petróleo.....	216
----------	--	-----

## SUMÁRIO

CAPÍTULO 1 Introdução Geral.....	30
1 INTRODUÇÃO.....	30
2 CARACTERÍSTICAS GERAIS DO CAFEIEIRO.....	33
2.1 Importância econômica do café.....	34
2.2 <i>Coffea canephora</i> .....	37
2.3 Raízes de cafeeiro.....	39
2.4 Biotecnologia do cafeeiro.....	42
3 ESTRESSE HÍDRICO E PRODUTIVIDADE CAFEIEIRA.....	44
3.1 Resposta do cafeeiro ao estresse hídrico.....	45
4 ANÁLISES “ÔMICAS”.....	49
4.1 PCR Quantitativa em Tempo Real - qPCR.....	52
4.2 Análises <i>in silico</i> , NGS (Nova Geração do Sequenciamento) e bioinformática.....	53
5 PROTEÔMICA.....	56
5.1 Separação de proteínas por cromatografia.....	57
5.2 Identificação de proteínas por espectrometria de massa.....	58
5.3 Análise <i>in silico</i> .....	60
REFERÊNCIAS.....	62
CAPÍTULO 2 Caracterização de genes diferencialmente expressos em raízes de clones de <i>Coffea canephora</i> sensíveis e tolerantes ao estresse hídrico.....	78
1 INTRODUÇÃO.....	80
2 MATERIAL E MÉTODOS.....	82
2.1 Material Vegetal.....	82
2.2 Extração de RNA.....	84
2.3 Identificação <i>in silico</i> de genes candidatos.....	85
2.4 Ensaios de qPCR.....	92
3 RESULTADOS E DISCUSSÃO.....	95
3.1 Análises <i>in silico</i> .....	95
3.2 qPCR para os genes candidatos.....	127
4 CONCLUSÕES.....	158
REFERÊNCIAS.....	160
CAPÍTULO 3 Caracterização de proteínas diferencialmente expressas em raízes de clones de <i>Coffea canephora</i> submetidos à deficiência hídrica.....	170
1 INTRODUÇÃO.....	172
2 MATERIAL E MÉTODOS.....	174
2.1 Material vegetal.....	174
2.2 Fracionamento das proteínas totais.....	176

2.3	Espectrometria de massa.....	177
2.4	Identificação das proteínas <i>in silico</i> .....	178
3	RESULTADOS E DISCUSSÃO .....	179
4	CONCLUSÕES .....	202
	REFERÊNCIAS .....	203
	ANEXOS .....	206
	<b>CAPÍTULO 4 Análise integrada dos perfis transcriptômico e proteômico em raízes de clones de cafeeiro (<i>Coffea canephora</i>) sensíveis e tolerantes ao estresse hídrico .....</b>	<b>212</b>
1	INTRODUÇÃO .....	214
2	MATERIAL E MÉTODOS.....	215
3	RESULTADOS E DISCUSSÃO .....	216
4	CONCLUSÕES .....	231
	REFERÊNCIAS .....	232
	CONCLUSÃO GERAL .....	235

## CAPÍTULO 1 Introdução Geral

### 1 INTRODUÇÃO

O cafeeiro é a *commodity* agrícola mais valiosa do planeta, sendo cultivado em mais de 60 países (INTERNATIONAL COFFEE ORGANIZATION -ICO, 2014). As duas espécies *Coffea arabica* L. (Arábica) e *Coffea canephora* Pierre. (Robusta) predominam na produção mundial representando aproximadamente 65% e 35% do mercado cafeeiro, respectivamente (ICO, 2014; DAVIS et al., 2011).

No Brasil, a cafeicultura é um dos mais importantes complexos agroindustriais. O café é produzido em mais de 1.000 municípios, o que faz do país o maior produtor e exportador do grão em nível mundial, sendo responsável por aproximadamente 35% do valor global das exportações, seguidos por Vietnã e Indonésia (ICO, 2014).

Segundo o Ministério da Agricultura, Pecuária e Abastecimento (BRASIL, 2013), há uma previsão de aumento do percentual mundial de consumo de café em 28,6% e da produção dos grãos em 10,9% para os próximos 10 anos, o que explica a importância desta *commodity* tanto para os pequenos como para os grandes produtores (DAMATTA; RAMALHO, 2006).

Como consequência do aquecimento global, as intempéries climáticas que ocorrem no país estão mais frequentes, provocando mudanças com um aumento previsto em radiação solar, temperatura, umidade, dentre outros, comprometendo o desenvolvimento e a produtividade das plantas (ASSAD et al., 2004).

Dentre os fatores abióticos que afetam a agricultura brasileira, a seca é um dos principais agentes redutores, comprometendo 80% da produção (DAMATTA; RAMALHO, 2006). Neste sentido, a seleção de cultivares

tolerantes à seca é de grande importância, principalmente diante da expansão da cafeicultura brasileira para áreas sujeitas ao déficit hídrico (DAMATTA et al., 2010).

Como resposta ao estresse, as plantas desencadeiam múltiplas vias de sinalização como respostas de defesa ou de adaptação ao ambiente, estas induzem a percepção e transdução do sinal, expressão de genes e alterações metabólicas (AGARWAL et al., 2006). Em decorrência da necessidade em entender como estes mecanismos são regidos, técnicas biotecnológicas vêm sendo aplicadas. No cafeeiro, estas técnicas vêm sendo aplicadas por diversos grupos, com a finalidade de desenvolver cultivares com alta variabilidade genética, maior resistência, características agronômicas superiores, altamente produtivas e adaptadas às regiões de cultivo (DAMATTA; RAMALHO, 2006; MARRACCINI et al., 2011, 2012; RIBAS; PEREIRA; VIEIRA, 2006).

Diversos trabalhos buscam definir e selecionar genes candidatos para análise da expressão diferencial em clones de café sob deficiência hídrica (FREIRE et al., 2013; MARRACCINI et al., 2011, 2012; VIEIRA et al., 2013; VINECKY; SILVA; ANDRADE, 2012). Como resposta ao estresse hídrico, estudos comprovam que genótipos de *C. canephora* Conilon, possuem um sistema radicular para absorção de água mais eficiente (CRUZ et al., 2009; PINHEIRO et al., 2005), o que faz desta espécie alvo de pesquisas relacionadas às respostas do cafeeiro para a tolerância à seca.

Com o avanço dos projetos para sequenciamento do genoma do café (DE KOCHKO et al., 2010), nos últimos anos há um aumento considerável das pesquisas relacionadas à análise de expressão de perfis gênicos, proteicos e metabólicos de cafeeiros quando submetidos a condições de deficiência hídrica. As abordagens das “ômicas”: genômica, transcriptômica, proteômica e metabolômica possibilitam um melhor entendimento da fisiologia da planta. Trabalhos com essa finalidade têm sido publicados para respostas de cafeeiros

ao déficit hídrico (DAMATTA; RAMALHO, 2006; MARRACCINI et al., 2011, 2012; PINHEIRO et al., 2005; VIEIRA et al., 2013), tanto ao nível transcricional como a outros níveis de regulação: (i) estabilidade dos RNAm, (ii) rearranjos de DNA genômico, dentre outros (LOPES et al., 2013). Entretanto, trabalhos que visam entender os genes envolvidos na resposta de raízes de cafeeiro à seca ainda são poucos.

Sendo assim, o presente trabalho tem por finalidade integrar a análise do perfil de genes e proteínas diferencialmente expressos em raízes de clones de *C. canephora* quando submetidos ao déficit hídrico.

## 2 CARACTERÍSTICAS GERAIS DO CAFEEIRO

Pertencente à família *Rubiaceae*, subfamília *Ixoreideae*, o cafeeiro é subdividido em dois gêneros, o *Coffea L.* e o *Coffea Hook f.*, que compreendem mais de 100 espécies as quais diferem na morfologia, tamanho e adaptação ecológica (DAVIS et al., 2011). Nativo do continente africano, o gênero *Coffea L.*, pertencente ao grupo das dicotiledôneas, é caracterizado por possuir folhas persistentes, porte arbustivo, caule lenhoso e flores hermafroditas (FAZUOLI et al., 1986).

*C. arabica* é uma espécie autógama alotetraploide ( $2n = 4X = 44$  cromossomos) que possui um genoma de tamanho aproximado em  $2,62 \times 10^9$  pb, ou seja, 650MB na forma haplótica (CLARINDO; CARVALHO, 2011) e baixa diversidade genética, como consequência de autopolinização de suas flores (HATANAKA et al., 1999). Originou-se nos planaltos da Etiópia Central a aproximadamente um milhão de anos como resultado de um cruzamento entre *C. canephora* e *C. eugenoides* (LASHERMES et al., 1999).

A espécie *C. canephora* é alógama e diploide ( $2n = 2X = 22$  cromossomos), seu genoma tem aproximadamente  $1,41 \times 10^9$  pb, ou seja, em torno de 700 MB na forma haplótica (CLARINDO; CARVALHO, 2011). É nativa das terras baixas florestais da bacia do Rio Congo, onde se estende até o Lago Vitória, em Uganda (MONTAGNON; CUBRY; LEROY, 2012). *C. canephora* apresenta alta variabilidade genética e ótima adaptação às mais variadas condições climáticas (BERTRAND et al., 2003).

Essas duas espécies representam aproximadamente 99% da produção mundial (PARTELLI et al., 2009). Em relação à qualidade da bebida, o *C. arabica* é considerado de melhor qualidade, por possuir baixo teor de cafeína tem menor amargor, sendo a bebida proveniente deste a mais apreciada no mercado mundial, já os grãos de *C. canephora* são mais utilizados pelas

indústrias de café solúvel e na composição dos “*blends*” (HATANAKA et al., 1999; LEROY et al., 2006).

## 2.1 Importância econômica do café

Depois do petróleo, o café é a segunda maior fonte de renda para grande parte dos países em desenvolvimento. A cultura cafeeira está limitada à zona intertropical (Figura 1) e o Brasil é o maior produtor, seguindo por Vietnã, Indonésia e Colômbia, os quais correspondem a mais de 50% da produção mundial (ICO, 2014).

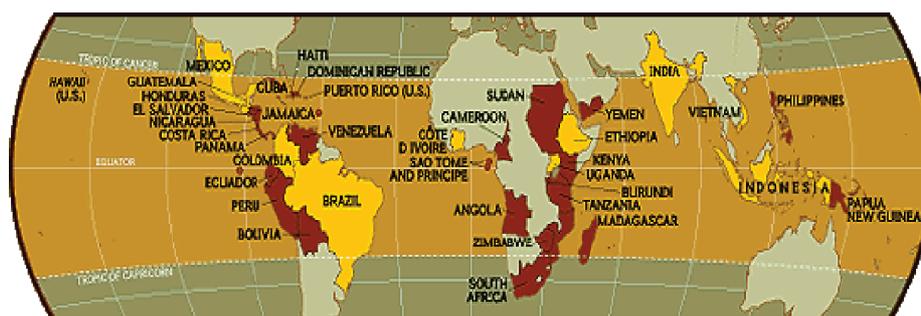


Figura 1 Regiões produtoras de café. Os dez países maiores produtores de café estão destacados em amarelo

Fonte: Major... (2014).

A produção cafeeira do Brasil se destaca na história econômica e social do país desde o período colonial. Hoje, o café é produzido em 15 estados, sendo os maiores produtores do grão Minas Gerais, Espírito Santo, São Paulo, Bahia, Paraná, Rondônia e Rio de Janeiro, que correspondem a 98% da produção brasileira (COMPANHIA NACIONAL DE ABASTECIMENTO - CONAB, 2014; MORAIS; MELO, 2011).

Segundo estimativas da CONAB (2014), neste ano, a safra de grãos de café poderá atingir 44,57 milhões de sacas de grãos beneficiados para as duas espécies (*C. arabica* e *C. canephora*), o que representa uma redução de aproximadamente 9% da produção em relação ao ano de 2013, conforme ilustrado na Figura 2.

Para o Café Arábica, a redução poderá atingir aproximadamente 15% da produção, o que pode ser explicado pela forte estiagem verificada nos primeiros meses de 2014 e à inversão da bienalidade em algumas regiões produtoras. Já para café robusta, estima-se um crescimento de aproximadamente 13% na produção, este aumento na produtividade decorre principalmente do crescimento da área em produção (CONAB, 2014). Em detrimento destes dados, as exportações de café para 2020/2023 estão projetadas em 20 milhões de sacas, com a expansão da taxa anual mundial em 2,7% e nacional em 4,8% (BRASIL, 2013).

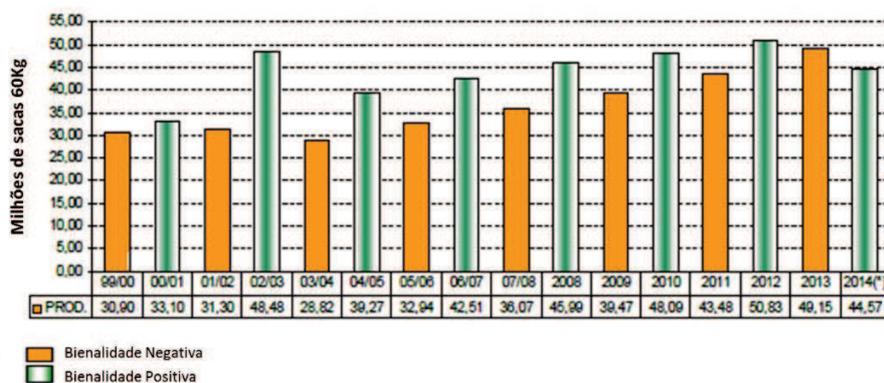


Figura 2 Média da evolução da produtividade do café beneficiado no Brasil. Produção em milhões de sacas/ano nos últimos 15 anos. Segundo estimativas, a produção de 2014 é considerada de bienalidade positiva

Fonte: CONAB (2014).

Tabela 1 Comparativo de produção de café beneficiado no Brasil, destacando as regiões de maior produção. (S) mil sacas beneficiadas

Regiões de maior produção de café Brasil/2013			Regiões de maior produção de café Brasil/2014			Variação (%)
Café Arábica (Região)	Maior Produtor (Estado)	Produção (s)	Café Arábica (Região)	Maior Produtor (Estado)	Produção (s)	
Sudeste	MG	27.380	Sudeste	MG	22.687	17,10**
Sul	PR	1.650	Sul	PR	545	66,97**
Nordeste	BA	1.079	Nordeste	BA	1.217	12,77*
Centro-Oeste	GO	265	Centro-Oeste	GO	261	1,39**
Café Canephora (Região)	Maior Produtor (Estado)	Produção (s)	Café Canephora (Região)	Maior Produtor (Estado)	Produção (s)	Variação (%)
Sudeste	ES	8.211	Sudeste	ES	9.350	13,87*
Norte	RO	1.357	Norte	RO	1.625	19,75*
Nordeste	BA	723	Nordeste	BA	769	6,37*
Centro-Oeste	MT	169	Centro-Oeste	MT	167	1,29**
<b>Total</b>		<b>49.151</b>	<b>Total</b>		<b>44.566</b>	<b>9,33**</b>

Fonte: CONAB (2014).

\*Variação positiva da produção. \*\*Variação negativa da produção.

Conforme ilustrado na Tabela 1, para o ano de 2014, a produção estimada total de *C. arábica* é de 32,2 milhões sacas, sendo o maior produtor o estado de Minas Gerais com 22,7 milhões de sacas. A produção total do café robusta é estimada em 12,3 milhões sacas, sendo o estado do Espírito Santo o maior produtor com a produção de 9,35 milhões de sacas (CONAB, 2014).

Embora forneça uma bebida de qualidade inferior à bebida fornecida pelos grãos de *C. arábica*, os grãos de *C. canephora* vêm apresentando um crescimento constante no Brasil (CONAB, 2014; GUIMARÃES, 2007). O seu preço reduzido no mercado mundial, atrai as indústrias de café solúvel, que os utilizam em misturas com o *C. arábica*, diminuindo o custo do produto. Os

estados que mais se destacam quanto à produção são Espírito Santo, Rondônia e Bahia (Tabela 1), os quais conferem ao país a mais elevada posição em exportação mundial, destacando-se principalmente quanto à exportação de café solúvel (SILVA; LEITE, 2000).

## **2.2 *Coffea canephora***

O *C. canephora* é típico de florestas equatoriais e tropicais, cujas temperaturas médias estão entre 24 e 26°C, as chuvas são abundantes, superiores a 2.000 mm e distribuídas ao longo de 9 a 10 meses, em locais onde o nível do mar atinge em torno de 800m (COSTE, 1989; DAMATTA, 2003). Apresenta ampla diversidade genética, comprovada com estudos de diversidade entre suas populações, os quais possibilitaram reunir esta espécie em dois grupos distintos denominados Congolês e Guineano, estabelecidos de acordo com sua origem geográfica (DUSSERT et al., 1999; MONTAGNON; CUBRY; LEROY, 2012).

O grupo Congolês apresenta vários subgrupos dentre os quais existem dois de maior importância, o subgrupo SG1 (tolerante à seca), típico de uma região de clima árido, localizado numa área mais litorânea do continente, e o subgrupo SG2 (sensível à seca), característico de uma região de clima tropical com índices pluviométricos elevados e bem distribuídos ao longo do ano, o que lhe confere susceptibilidade à seca (Figura 3) (DUSSERT et al., 1999; FAZUOLI, 2007).

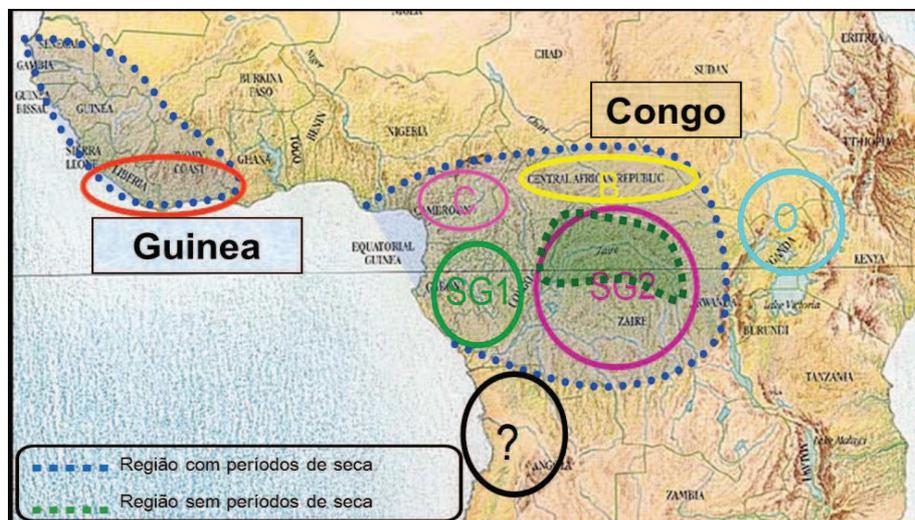


Figura 3 Grupos de distribuição da diversidade de *Coffea canephora*

Legenda: As linhas pontilhadas delimitam regiões com distribuição pluviométrica distintas. O Subgrupo SG1 originou-se em uma região de clima árido, enquanto o SG2 é típico de uma região de clima tropical.

Fonte: Montagnon et al. (2000), adaptado.

Dentro do subgrupo SG1, encontram-se as plantas conhecidas no Brasil como Conilon, considerada eficiente na resposta à tolerância à seca, por proporcionar maior controle de transpiração (LAMBOT et al., 2009). Também apresenta alta tolerância ao fungo *Hymeleia vastatrix*, causador da ferrugem nos cafeeiros (CARVALHO et al., 2001). Desse modo, as plantas de Conilon são alvos de pesquisas que objetivam a introdução destas em regiões de baixa altitude, que possuem temperaturas médias anuais elevadas (de 22°C a 26°C) e, principalmente, em regiões com um período seco mais prolongado (FAZUOLI et al., 1986).

### 2.3 Raízes de cafeeiro

As raízes são órgãos importantes na regulação do desenvolvimento das plantas. Exercem a função de absorção e condução de água e nutrientes minerais, permitem a sustentação mecânica da planta, além de produzir substâncias orgânicas complexas, como fito-hormônios (DAMATTA et al., 1997). O sistema radicular das plantas apresenta características diferenciadas de acordo com a espécie, variedade, idade, época do ano, clima, densidade do plantio, ataque de pragas, doenças, textura e estrutura do solo (PARTELLI et al., 2006).

A arquitetura da raiz é importante para determinar a capacidade da planta em explorar os recursos que a circunda (LYNCH, 1995), assim como sua capacidade em responder à dinâmica e disponibilidade de recursos por meio da atividade meristemática, funções de transporte e distribuição dos produtos capturados (PINHEIRO et al., 1995).

Numerosos fatores podem influenciar no desenvolvimento radicular: diferenças genéticas, produção e disponibilidade de carboidratos, variação sazonal, tipo de propagação e formação de mudas, espaçamento do plantio, frequência de podas aéreas e radiculares, ambiente físico e químico do solo, fatores bióticos como micorrizas, dentre outros (BRESSAN; VASCONCELOS, 2002). Desse modo, o conhecimento radicular de uma espécie proporciona melhor entendimento da fisiologia da planta de forma geral (PARTELLI et al., 2006; SOARES et al., 2007).

No cafeeiro, as raízes apresentam-se curtas, grossas, terminando abruptamente, são múltiplas e estendem-se no máximo 45cm abaixo da superfície (RENA; GUIMARÃES, 2000). O aprofundamento radicular é uma característica fenotípica importante e, quando se desenvolve pela indisponibilidade de água e nutrientes no solo, possibilita a identificação de

possíveis variedades resistentes à seca (AMARAL, 2002; BARRETO et al., 2006).

A distribuição radicular espacial permite maximizar o aproveitamento de diversos nutrientes pelas diversas culturas (BARRETO et al., 2006). Em situações de déficit hídrico, o sistema radicular do cafeeiro desenvolve-se mais do que outras partes da planta (THOMAZIELLO et al., 2000). Estes e outros mecanismos fisiológicos são importantes para que a planta tolere melhor aos períodos mais secos (DAMATTA et al., 2003; DAMATTA; RAMALHO, 2006; FREITAS et al., 2000; PINHEIRO et al., 2005).

Clones que possuem elevada produção em condições de seca têm a capacidade de manter os potenciais hídricos adequados, em decorrência da combinação do aprofundamento do sistema radicular e aumento do controle estomático. A manutenção da capacidade de transporte de fotoassimilados da parte aérea para as raízes é um dos fatores que possibilita o aumento da tolerância do estresse hídrico, pois permite maior crescimento radicular e, dessa forma, maior acesso a horizontes mais profundos do solo, de modo a aumentar a absorção de água (SILVA et al., 2010).

Por sua vez, os clones sensíveis ao déficit hídrico possuem controle deficiente da transpiração, resposta limitada das células estomáticas à redução da disponibilidade de água do solo e a baixa capacidade de absorção de água nas raízes, uma vez que possuem sistema radicular pouco desenvolvido e superficial (PINHEIRO et al., 2005).

Estudos de associação da tolerância à seca, enraizamento e controle estomático em folhas de clones de *C. canephora* foram realizados por Pinheiro et al. (2005) e os resultados de morfologia radicular mostraram que as raízes dos clones tolerantes à seca desenvolveram mais profundamente e com melhor distribuição se comparadas aos dos clones sensíveis (Figura 4).

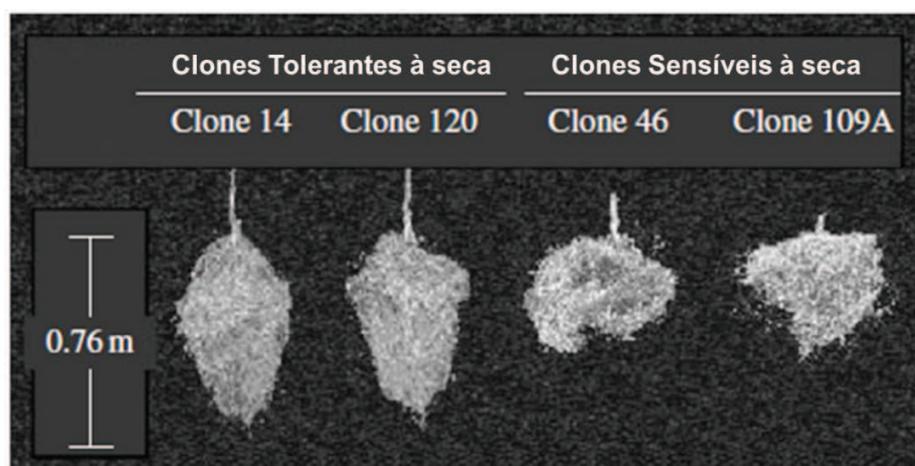


Figura 4 Sistemas radiculares de quatro clones de *C. canephora* cultivados com restrição hídrica

Fonte: Pinheiro et al. (2005) adaptado.

Espécies ou cultivares tolerantes à seca geralmente diferem morfológica e fisiologicamente das sensíveis, pois detêm mecanismos que possibilitam maior capacidade no abastecimento de água, como (i) maximização da absorção pelos sistemas radiculares, (ii) menor perda de água pelo fechamento estomático e (iii) redução da área foliar (KRAMER; BOYER, 1995). O ajuste osmótico ou a mudança na elasticidade da parede celular podem ser alcançados pela melhoria do turgor celular, o que é essencial para manter a atividade fisiológica da planta durante longos períodos de seca (TURNER, 1997).

No cafeeiro, o sistema radicular é importante por absorver nutrientes e água do solo para a planta, destacando-se na assimilação de nitrato, que é responsável pela acumulação de nitrito, aminoácidos e amônia nas raízes proporcionando melhor desenvolvimento da planta (DAMATTA; AMARAL; RENA, 1999), bem como melhor crescimento (principalmente de pequenas raízes) durante as estações secas (CANNELL; HUXLEY, 1969).

## 2.4 Biotecnologia do cafeeiro

Com os avanços das pesquisas em biologia molecular, o melhoramento genético vem sendo aprimorado e a engenharia genética amplamente utilizada com a finalidade de inserir novas características de interesse agrônomo (BORÉM, 1998). O que resulta em um rápido e atraente modo para aprimorar os mecanismos de tolerância das espécies a múltiplos estresses (WANG; VINOGRAD; ALTMAN, 2003).

Reduzir os custos de produtividade e aumentar a produção são fatores que afetam diretamente o crescimento e potencial competitivo do país no que diz respeito ao melhoramento de cafeeiro. Nos últimos 15 anos, com o avanço da biotecnologia, os métodos vegetativos de seleção aprimoraram o desempenho de clones de cafeeiro em condições de seca, tornando a cultura menos dependente de irrigação.

Inúmeros trabalhos estão sendo realizados priorizando entender os mecanismos fisiológicos de tolerância à seca em café (DAMATTA; RAMALHO, 2006; PINHEIRO et al., 2005). Pouco se sabe sobre os mecanismos moleculares envolvidos na resposta ao estresse hídrico nas raízes em clones de *C. canephora*. Segundo Damatta et al. (2003), o melhor rendimento na colheita do clone tolerante, em comparação com um clone sensível, está associado à manutenção da área foliar e do potencial hídrico dos tecidos. O desenvolvimento do sistema radicular também é uma resposta fisiológica da planta para a tolerância à seca (PINHEIRO et al., 2005).

Como resultado do sequenciamento do genoma do café (Projeto Genoma Café), foram geradas mais de 155 mil sequências de genes (*expressed sequence tag* - EST) que estão sendo caracterizadas funcionalmente (MONDEGO et al., 2011; VIEIRA et al., 2006). Porém, as sequências geradas

não fornecem respostas quantitativas nem qualitativas do produto final de um gene, as proteínas.

Análises de transcriptoma mostram que o acúmulo de mRNA não significa necessariamente que os mRNAs dos genes ativamente expressos são traduzidos, além do perfil de transcriptoma, é importante examinar quais mRNAs são traduzidos, degradados ou temporariamente armazenados durante a resposta ao estresse (URANO et al., 2010).

A resposta do transcrito ao estresse é mediada por mecanismos pós-transcricionais de processamento do mRNA, como *splicing* alternativo e silenciamento do mRNA. Mecanismos pós-traducionais como remoção de peptídeos sinais, fosforilação, glicosilação, dentre outros, regulam e ativam as moléculas pré-existentes garantindo uma rápida resposta. Correlacionar estes mecanismos permitem o entendimento a nível temporal e espacial dos padrões adequados de expressão (MAZZUCOTELLI, 2008).

O sequenciamento dos genes de resposta ao estresse hídrico em cafeeiro já está sendo realizado, inclusive técnicas de RNA-Seq em folhas em *C. canephora* e *C. arabica*. Como forma de ampliar as informações geradas pelo sequenciamento, este trabalho tem por finalidade ampliar os dados de RNA-Seq nas análises com raízes.

Para um melhor entendimento dos mecanismos que regem o sistema fisiológico de cafeeiro, dados de proteômica também foram analisados. A identificação das proteínas envolvidas nas respostas ao déficit hídrico nos clones contrastantes permitirá correlacioná-las aos mecanismos de tolerância diferencial.

### 3 ESTRESSE HÍDRICO E PRODUTIVIDADE CAFEIEIRA

Para sua sobrevivência, as plantas se adaptam às variáveis ambientais as quais são submetidas durante todo o seu ciclo vital. Seja em relação a intempéries climáticas, como intensidade de luz, temperatura, disponibilidade de água (estresse abiótico), seja para proteger-se de predadores (estresse biótico). Assim, o desenvolvimento das plantas depende da rápida regulação da expressão dos genes, com a finalidade de adaptar a sua fisiologia aos estresses ao qual são submetidas (FLORIS et al., 2009).

O estresse ambiental apresenta um grande desafio na busca de uma produção sustentável de alimentos, pois reduz os rendimentos potenciais em até 70% das plantas cultivadas (NELSON et al., 2009). O estresse hídrico ocasionado pela seca e severidade da temperatura é o estresse abiótico mais prevalente, que limita o crescimento e a produtividade das plantas. As plantas respondem e se adaptam a essas condições com uma série de alterações bioquímicas e fisiológicas.

A disponibilidade da água é reconhecida como uma restrição primária sobre a produtividade dos ecossistemas florestais (BRÉDA et al., 2006; BRESHEARS et al., 2005). Existem vários estudos sobre o desenvolvimento fisiológico das espécies vegetais ao longo de vários ambientes (MONTAGNON et al., 2000; WU; MATHESON, 2005), mas para árvores de colheita, principalmente para espécies perenes, o aumento das mudanças climáticas vem dificultando as pesquisas em decorrência das frequentes mudanças no desenvolvimento das plantas em respostas a diversos tipos de estresses (GWAZE; BRIDGWATER; WILLIAMS, 2002).

Os estresses osmótico e oxidativo resultam na interrupção da homeostase osmótica e iônica e em danos às proteínas e membranas (DAMATTA et al., 2003). Altas temperaturas por tempo prolongado levam à

morte da planta e, em períodos moderados, afetam a floração, o desenvolvimento do grão e, conseqüentemente, a produção do café (DAMATTA; RAMALHO, 2006). O frio também é uma problemática para a produção de café, provocando diversas alterações fisiológicas, morfológicas e biológicas no cafeeiro (CAMPOS et al., 2003; RAMALHO et al., 2003; SILVA et al., 2004). Grandes variações pluviométricas e de temperatura também comprometem os frutos, modificando sua composição bioquímica e a qualidade final da bebida (SILVA et al., 2005).

Os efeitos primários de estresses abióticos, como seca e frio, causam danos celulares e estresses secundários nas plantas, desencadeando inúmeras reações, dentre as quais estão: alteração da expressão gênica; acúmulo de metabólitos como o fito-hormônio ácido abscísico (ABA) ou compostos osmóticos ativos e a síntese de proteínas específicas (grandes proteínas hidrofílicas, proteínas antioxidantes, chaperonas, etc.) (REDDY; CHAITANYA; VIVEKANANDAN, 2004).

Os danos causados pela seca se desenvolvem lentamente e a intensidade do dano aumenta ao longo tempo. Em longo prazo, a deficiência hídrica possibilita o desenvolvimento de respostas morfológicas e fisiológicas na planta (DAMATTA; RAMALHO, 2006).

### **3.1 Resposta do cafeeiro ao estresse hídrico**

A produtividade do cafeeiro é fortemente influenciada pelo suprimento adequado de água e nutrientes (COELHO et al., 2009) e está sujeita a oscilações regulares decorrentes do ciclo bienal natural, bem como dos efeitos adversos das condições climáticas, principalmente as altas temperaturas (CANNEL et al., 1985; DAMATTA; RAMALHO, 2006), frio e grandes variações pluviométricas (SILVA et al., 2004).

No início do século XXI, o Brasil teve uma acentuada redução da área cafeeira devido a diversos fatores que têm determinado oscilações na produção total do café, dentre estes, fatores abióticos como seca, alta salinidade, mudanças de temperatura e fatores bióticos como pragas e doenças (GONÇALVES, 1999). Devido à mudança climática global, os períodos de seca podem tornar-se mais frequentes, comprometendo a sustentabilidade e a produtividade do cafeeiro (ASSAD et al., 2004; INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE -IPCC, 2014).

Diz-se que uma planta é tolerante à seca quando ela possui a capacidade de desenvolvimento e produção de maneira satisfatória, independente de como a habilidade se desenvolve (DAMATTA et al., 2003). Diversos genes são induzidos quando a planta sofre um tipo de estresse abiótico (BOHNERT et al., 2001; SEKI et al., 2001; ZHU et al., 2001) e estes podem ser agrupados conforme a função dos seus produtos:

Genes que codificam para proteínas funcionais: proteínas de membrana (proteínas de canais de água e transportadores de membrana); as enzimas (prolina e betaína) e transportadores de açúcares (sorbitol e manitol), fundamentais na biossíntese dos compostos osmóticos e enzimas detoxificantes, que auxiliam no equilíbrio metabólico da planta (glutathione-S-transferase, hidrolase, catalase, superóxido-dismutase, ascorbato peroxidase, etc.); e outras proteínas de proteção às macromoléculas (proteínas LEA, osmotina, chaperonas e proteínas de ligação ao mRNA, etc.) são alvos de pesquisa para entendimento dos mecanismos fisiológicos do cafeeiro (SHINOZAKI et al., 2003; SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2000).

Genes que codificam para proteínas regulatórias ou fatores de transcrição: (bZIP, MYC, MYB e DREB, etc.) as proteínas quinases (MAP quinase, CDP quinase, e proteínas quinases receptoras, ribossomais e reguladoras da transcrição) e proteinases (fosfotransferases e fosfolipases C)

envolvidas na regulação da transmissão de sinais e expressão gênica (SHINOZAKI et al., 2003; SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2000). Elementos cruciais de respostas adaptativas dos cafeeiros às condições ambientais adversas, como estresse hídrico provocado pela seca, o ABA (ácido abscísico) e o peróxido de hidrogênio ( $H_2O_2$ ) são produzidos e acumulados nas plantas em resposta ao fechamento dos estômatos (SCHROEDER; KWAK; ALLEN, 2001; ZHANG et al., 2001) e na defesa das plantas contra agentes antioxidantes (ZHANG et al., 2005). Para o cafeeiro, existe uma comprovada relação positiva entre o fechamento estomático e a formação de acúmulo de peróxido de hidrogênio (DEUNER et al., 2011). O fechamento dos estômatos é um dos primeiros mecanismos de resposta à deficiência hídrica (NASCIMENTO et al., 2008).

O ABA reduz o turgor das células-guarda, induzindo o fechamento estomático, limitando a perda de água e a assimilação de carbono pela fotossíntese (MARRACCINI et al., 2011). Como a assimilação de carbono é a etapa inicial da inativação de reações de transferência de elétrons, nas plantas estressadas há uma redução de energia, diminuindo o oxigênio molecular e aumentando o número de moléculas reativas de oxigênio que causam danos foto-oxidativos (PRAXEDES et al., 2006). Em consequência, há aumento na atividade das enzimas antioxidantes, bem como nos teores de ascorbato e hidroascorbato, atuando na detoxificação dos radicais livres formados em função do estresse (DEUNER et al., 2011).

A tolerância à seca é um fenômeno decorrente de diversas características anatômicas, morfológicas e fisiológicas, de natureza constitutiva ou indutiva, que operam cooperativamente e possibilitam a manutenção do crescimento e desenvolvimento das plantas submetidas a condições edafoclimáticas que levam à deficiência hídrica

A fim de evitar perdas adicionais de água, a planta promove o fechamento dos estômatos resultando no bloqueio de dióxido de carbono (CO<sub>2</sub>) e consequente alteração na taxa fotossintética. A redução do fluxo de vapor e da absorção de água e nutrientes pelo sistema radicular resultam na diminuição da produção (KUMAR, 1999; MATIELLO; DANTAS, 1987). A escassez de água no ambiente leva a perda do turgor celular, ao aumento da concentração do citossol e a consequente desidratação do protoplasma (PINHEIRO et al., 2005).

#### 4 ANÁLISES “ÔMICAS”

As respostas moleculares e celulares ao estresse abiótico obedecem a uma cascata de eventos que incluem a percepção do estímulo, transmissão do sinal ao citoplasma e núcleo, expressão gênica e por fim alterações metabólicas que resultam na manifestação da tolerância, finalizada quando este produto atinge o nível fenotípico, conforme ilustrado na Figura 5.

##### A cascata das “OMICAS”

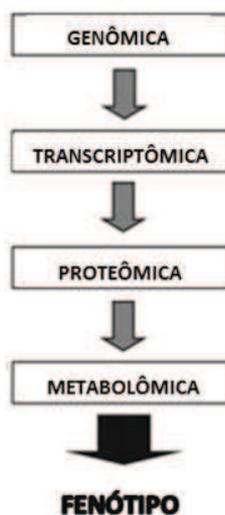


Figura 5 “Ômicas” - cascata de eventos para a expressão de um gene  
Fonte: Dettmer, Aronov e Hammock (2007), adaptado.

O genoma é definido como o conjunto de genes presentes em uma célula. A sequência de nucleotídeos encontradas nestes genes é transcrita em mRNA e, em seguida, traduzidas em proteínas, que desempenham as diversas funções celulares que mantêm os organismos vivos (BAN et al., 1999). O estudo

da sequência que compõe um gene ou uma proteína de interesse é considerado o marco crucial da biologia molecular.

Em adição às informações obtidas pela análise do genoma é importante ter conhecimento sobre quais proteínas estão sendo expressas, quando e em quais níveis esta expressão ocorre, bem como quais as eventuais modificações pós-traducionais. Portanto, em paralelo às disponibilizações dos dados genômicos, foram desenvolvidas outras técnicas para o estudo do genoma funcional, como a análise dos transcritos de mRNA (transcriptoma ou transcriptômica) e a análise do conjunto de proteínas expressas a partir do genoma (análise de proteoma ou proteômica).

Com o advento do sequenciamento genômico do cafeeiro (MONDEGO et al., 2011; VIEIRA et al., 2006), foi possível identificar inúmeros genes de interesse agrônomo, o que possibilitou aos biólogos moleculares a manipulação e estudo destes genes.

Mesmo com o auxílio de técnicas modernas, o sequenciamento completo de espécies com genomas grandes é demorado, oneroso e trabalhoso. O sequenciamento de cDNAs por ser rápido e possuir boa relação de custo benefício tem crescido nas publicações nas bases de dados públicos. É uma técnica bastante utilizada para avaliar o funcionamento das células, bem como dos organismos. Uma maneira corriqueiramente utilizada para estas análises é por meio do estudo de ESTs.

Uma alternativa simples e eficiente para analisar ESTs em condições ambientais e tecido-temporais específicas é a produção de sequências randômicas de transcritos gênicos. Com o objetivo de se estudar o genoma funcional de diversas plantas, diversos marcadores moleculares EST-PCR foram desenvolvidos e são muito utilizados para o estudo de mapeamento comparativo (SARGENT et al., 2007). Para o cafeeiro, mais de 250 mil sequências de ESTs

estão depositadas em domínio público essencialmente de *C. arabica*, *C. canephora* e *C. racemosa* (VIEIRA et al., 2006).

Por iniciativa do Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café (CBP&D-CAFÉ), coordenado pela Embrapa Café em colaboração com a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) e a Embrapa Recursos Genéticos e Biotecnologia (Cenargen), teve início, em fevereiro de 2002, o Projeto Brasileiro do Genoma Café (PBGC), em sua base de dados (CafEST) estão armazenadas mais de 30 mil unigenes. Em raízes, existiam somente 2.071 *reads*, o presente trabalho permitiu a ampliação de dados de raízes de cafeeiro por meio do sequenciamento de cDNA, utilizando a técnica de pirosequenciamento 454, para a construção de bibliotecas do transcriptoma de raízes de *C. canephora* provenientes dos clones 14, 22, 73 e 120 nas condições irrigado (I) e não irrigado (NI).

O grupo de pesquisa CENICAFE, produziu 32.961 seqüências de EST de três tecidos diferentes (folhas, frutos com 31 semanas após a antese e flores) de *C. arabica* (cv. Caturra) montadas em 10.799 unigenes (MONTROYA et al., 2006). Apesar da obtenção dessas seqüências em 2006, nenhum desses grupos de pesquisa as tornou imediatamente públicas.

Na França, o grupo do IRD (Institut de Recherche pour le Développement) produziu 10.420 seqüências de ESTs (agrupadas em 5.534 unigenes potenciais) a partir das bibliotecas de cDNA de frutos e folhas de *C. canephora* (PONCET et al., 2006) que unidas às 47.000 seqüências de EST disponíveis, representam 13.175 unigenes, publicados pela Nestlé e a Universidade de Cornell (LIN et al., 2005), estes perfazem o total de 55.694 seqüências disponíveis e formam uma das principais fontes de pesquisa para a comunidade científica. Em decorrência de uma parceria entre o Projeto Genoma Café Brasileiro e o consórcio feito pela Universidade de Cornell e a Nestlé, foi possível a montagem de ESTs de *C. arabica* e *C. canephora* que revelaram mais

de 32.000 *clusters* de *C. arabica* e 16.665 *clusters* de *C. canephora* (MONDEGO et al., 2011).

Desse modo, as ESTs, representam milhares de genes expressos em diferentes tecidos do cafeeiro, obtidos em vários estágios de desenvolvimento da planta e submetidos a condições de estresse biótico e abiótico. Estas podem ser utilizadas para diversos fins, como a detecção dos perfis de expressão entre as principais espécies comerciais de cafeeiro, no que diz respeito à estrutura genômica; bem como para comparação de dados de expressão das sequências dos *contigs* de cada um dos grupos (*C. arabica* e *C. canephora*), resultando na identificação de genes diferencialmente expressos relacionados à defesa da planta, estresses abióticos e à qualidade da bebida (MONDEGO et al., 2011; VIDAL et al., 2010).

#### 4.1 PCR Quantitativa em Tempo Real - qPCR

Aplicada inicialmente em 1984 por Kary Mullis, a *Polymerase Chain Reaction* (PCR) foi uma das maiores invenções da ciência (MULLIS; FALOONA, 1987). A PCR Quantitativa em Tempo Real (qPCR) é uma técnica da PCR que possibilita medir e distinguir sequências específicas de ácidos nucleicos em uma amostra, por meio da fluorescência gerada por sondas durante a amplificação de sequências de DNA alvo em “tempo real”, proporcionando o monitoramento dos produtos da PCR à medida que eles vão sendo amplificados, em aproximadamente cinco ciclos, o que permite avaliar a expressão dos genes de interesse pela quantidade destes a cada ciclo de amplificação. Mede a formação do produto na fase exponencial, ainda no início da reação, quando a amplificação dos fragmentos é mais eficiente (VALASEK; REPA, 2005). O ponto que detecta o ciclo na qual a reação atinge o limiar da fase exponencial é denominado de *Cycle Threshold* ( $C_T$ ) (NOVAES; PIRES-ALVES, 2004).

Dentre as vantagens da qPCR em relação a PCR convencional destacam-se a maximização da rapidez e da reprodutibilidade e a minimização do risco de contaminações. O processo de qPCR é comumente subdividido em três estágios: o primeiro caracteriza-se pela desnaturação (separação) da dupla fita de DNA, que ocorre em altas temperaturas, por volta de 90°C; o segundo estágio refere-se à ligação dos *primers* (iniciadores) à fita molde, que ocorre por volta de 50°C a 60°C e o terceiro estágio está relacionado à extensão e ao simultâneo anelamento da nova fita, e ocorre em aproximadamente 70°C. Esta temperatura é usualmente referida como Temperatura de Melt (TM) (MACKAY, 2004), e é um dado importante na qPCR, pois indica a temperatura em que 50% dos oligonucleotídeos estão em dupla-fita (VALASEK; REPA, 2005).

Devido à precisão e sensibilidade do tempo real da qPCR, mesmo as mudanças sutis na expressão do gene podem ser detectadas. Assim a qPCR pode ser utilizada para avaliar tanto DNA quanto RNA com grande sensibilidade e precisão (VALASEK; REPA, 2005). Esta técnica é o método mais sensível para quantificar mRNA, especialmente para moldes pouco abundantes (GAUSE; ADAMOVICZ, 1994), o mRNA serve de moldes para a produção de cDNAs e estes servirão de matéria-prima para a qPCR.

Esta técnica vem crescendo rapidamente entre as pesquisas científicas, sendo uma forte ferramenta para a análise da expressão de genes.

#### **4.2 Análises *in silico*, NGS (Nova Geração do Sequenciamento) e bioinformática**

O processo pioneiro do sequenciamento do DNA teve início no século XXI com a metodologia de Sanger, Nicklen e Coulson (1977). Novas tecnologias de Sequenciamento vêm sendo desenvolvidas ao longo do tempo, dentre as plataformas da Nova Geração do Sequenciamento (NGS), o 454

(Roche Applied Science), a Solexa (Illumina Genome Analyzer), a SOLiDTM (Applied Biosystems) e a Helicos BioSciences são as principais (MARDIS, 2008).

Das NGS, a pioneira no sequenciamento foi a plataforma 454, pela qual os fragmentos são digeridos e selecionados por tamanho e então sequenciados, produzindo uma cobertura parcial de todo o genoma. Somente as extremidades dos fragmentos são sequenciadas (DAVEY et al. 2011).

Apesar dos avanços nas tecnologias NGS o sequenciamento de espécies com grandes genomas é lento, trabalhoso e dispendioso. Uma alternativa é o sequenciamento parcial do genoma utilizando ESTs (Etiquetas de Sequências Expressas) para analisar e identificar genes expressos a partir de transcritos gênicos (NAGARAJ et al., 2007). Alguns trabalhos já estão sendo produzidos em relação aos dados de RNA-Seq de raízes de clones de *C. canephora* (COSTA et al., 2013a, 2013b).

Mondego et al. (2005) isolaram e caracterizaram genes envolvidos na resistência ao bicho-mineiro em café, selecionando ESTs expressos nas bibliotecas de cafés resistentes à *Leucoptera coffeella*. Marraccini et al. (2012) identificaram 13 genes candidatos em café diferencialmente expressos em resposta ao estresse hídrico, utilizando análises de bioinformática. Também foram identificados genes de resistência à ferrugem (PRAKASH et al., 2004; TEDESCO et al., 2000), entretanto, nenhum trabalho foi publicado de genes envolvidos na resposta à seca em raízes de cafeeiro.

Desde o advento do GenBank e do Projeto Genoma Humano, a bioinformática se tornou ferramenta essencial na análise e gerenciamento de dados de sequências de DNA e proteína, e tende a continuar se desenvolvendo com o surgimento de novos dados (ZATZ, 2002).

Diversas ferramentas criadas pela bioinformática possibilitam o acesso e análise dos bancos de dados. Para a comparação de sequências de DNA com as

dos bancos de dados genômicos, pode-se utilizar o BLAST – *Basic Local Alignment Search Tool* (ALTSCHUL et al., 1990), o qual tem a finalidade de encontrar, no banco de dados, a presença de uma sequência com certa similaridade à pesquisada.

Várias regiões de DNA podem ser anotadas pelo BLAST o que possibilita detectar a função que um determinado segmento de DNA representa comparando-o com sequências de similaridade significativa depositada no GenBank que tem função experimentalmente conhecida (SANTOS; ORTEGA, 2003). Para o estudo em cafeeiro, existem diversas plataformas para busca de sequências como Coffee Genome Hub, MoccaDB, Sol Genomics Network.

## 5 PROTEÔMICA

O estudo sistemático do proteoma expresso nas células em condições fisiológicas específicas é definido como proteômica, que tem por objetivo identificar e relacionar proteínas, bem como caracterizar modificações pós-traducionais ou interações proteína-proteína (WIJK, 2001). A regulação da expressão gênica pode ocorrer em diferentes estádios e pode ser particularmente importante durante a transcrição. Portanto, os dados proteômicos são bastante utilizados na melhoria das anotações genômicas, pois permitem a validação de ESTs por meio da comparação da proteína expressa com o transcrito do gene correspondente (NAGARAJ et al., 2007).

A espectrometria de massa e as técnicas de separação, identificação e quantificação de proteínas são bastante utilizadas, entretanto a identificação de todas as proteínas é um desafio técnico. Enquanto em organismos unicelulares a cobertura do proteoma pode alcançar mais de 50%, em organismos superiores a cobertura raramente ultrapassa os 10%. Em se tratando de quantificação de proteínas, esses números são ainda significativamente menores, pois a qualidade dos dados, em termos de conteúdo de informação necessária para a quantificação é muito inferior ao número existente de proteínas (BANTSCHEFF et al., 2007).

Uma maneira de identificá-las é separar as proteínas presentes num extrato proteico e, em seguida, caracterizá-las isoladamente pela espectrometria de massa, que combinada com análises computacionais e bases de consulta de dados é uma poderosa ferramenta para identificação, quantificação e caracterização de proteínas (WILDGRUBER et al., 2000).

Dentre os métodos de quantificação de proteínas, a utilização de fluoróforos, corantes, ou radioatividade é bem difundida e apresenta excelente sensibilidade, linearidade e faixa dinâmica, mas passa por dois problemas importantes: requer a separação de proteínas de alta resolução normalmente

fornecidas pelos géis 2-DE, o que limita a sua aplicabilidade em proteínas mais solúveis e abundantes e ainda não revela a identidade da proteína subjacente. Ambos os problemas podem ser superados pelas técnicas modernas LC-MS. No entanto, a espectrometria de massa é inerentemente não quantitativa porque peptídeos proteolíticos exibem uma vasta gama de propriedades físico-químicas, tais como tamanho, carga, hidrofobicidade, entre outras, que levam as grandes diferenças na resposta de espectrometria de massa para uma precisa quantificação (BANTSCHEFF et al., 2007).

A capacidade para determinar a concentração absoluta de uma proteína em uma mistura proteica complexa é crucial para a compreensão de como as proteínas interagem dentro da célula em resposta a uma alteração adaptativa (SILVA et al., 2005).

Muitos estudos proteômicos já foram relatados utilizando a eletroforese bidimensional em gel (2-DE). No entanto, é uma técnica extremamente demorada, pois requer o uso de replicatas de géis, bem como procedimentos de excisão de gel e digestão, estes passos podem ser complicados e lentos. Como forma de minimizar as limitações sofridas pela eletroforese bidimensional em gel de poliacrilamida, a cromatografia líquida tem sido integrada com a espectrometria de massa (LC-MS) e compõe uma nova plataforma para análise do proteoma (MURAD; RECH, 2012). Neste tipo de análise, a identificação e a quantificação das proteínas são obtidas a partir do extrato proteico digerido com a enzima tripsina sem a necessidade de numerosos peptídeos externos de referência ou de métodos de radiomarcção (SILVA et al., 2005).

### **5.1 Separação de proteínas por cromatografia**

Cromatografia líquida acoplada com espectrometria de massa é uma ferramenta poderosa para a análise de teores de proteína e modificações de

expressão (MURAD; RECH, 2012). É uma técnica de separação físico-química dos elementos de uma mistura por meio da distribuição destes componentes em duas fases: a fase estacionária ou a fase fixa que compreende uma coluna ou uma superfície sólida e a fase móvel, que corresponde a um gás, um líquido ou um fluido supercrítico (LANÇAS, 2009).

A técnica a ser utilizada varia de acordo com o enfoque da pesquisa: a cromatografia líquida, a cromatografia em coluna ou planar, a cromatografia por afinidade ou exclusão molecular e a cromatografia gasosa são as mais aplicadas. (LANÇAS; MCHAIR, 1983). Os tipos de cromatografia líquida diferem de acordo com a interação ocorrida entre a fase estacionária e o soluto; a cromatografia por afinidade que pode ser adsorção, partição ou troca iônica e a cromatografia por exclusão nuclear, que pode ser por permeação ou filtração em gel (LANÇAS, 2009).

A cromatografia líquida de fase reversa (RPLC) é caracterizada pela hidrofobicidade das interações entre a parte não polar do soluto e a fase estacionária, ou seja, pela repulsão do soluto pela fase móvel aquosa. Por possuir alta resolução, é uma técnica bastante utilizada na purificação e análise de biomoléculas, como a separação de peptídeos decorrentes de uma digestão enzimática, bem como para separar peptídeos naturais e sintéticos (SCARBOROUGH et al., 1984). Esta técnica tem a capacidade em separar polipeptídios de sequências muito próximas até mesmo os que diferem de apenas um resíduo de aminoácido (LANÇAS, 2009; RIVIER; MCCLINTOCK, 1983).

## **5.2 Identificação de proteínas por espectrometria de massa**

A espectrometria de massa é uma técnica que caracteriza as proteínas por meio da avaliação da massa molecular dos peptídeos (YERGEY et al.,

2002), tendo por base a conversão das moléculas neutras em íons e a determinação precisa da relação  $m/z$  (massa/carga) destes em fase gasosa, com a utilização de campos eletromagnéticos (BANDIL, 2008).

É uma ferramenta crucial na proteômica sendo utilizada na determinação precisa da massa molecular proteica, quantificação de cisteínas, localização de pontes dissulfeto, determinação de modificações pós-traducionais tais como glicosilações e fosforilações, determinação das sequências de aminoácidos e, principalmente, na identificação da estrutura primária das proteínas (LARSEN et al., 2002).

O espectrômetro de massa é composto por uma fonte de ionização, que introduz uma carga e vaporiza a amostra, analisadores de massa, que determinam a relação  $m/z$  do analito e um detector que registra a presença dos analitos (GRAVES; HAYSTEAD, 2002). Há dois tipos básicos de fontes de ionização que são amplamente utilizados para volatilização e ionização de biomoléculas de alto peso molecular: a Ionização por Dessorção à Laser Auxiliada por uma Matriz (*Matrix-assisted Laser Desorption/Ionization* – MALDI) (KARAS; HILLENKAMP, 1998) e a ionização por eletropulverização (*Eletron Spray Ionization* – ESI) (FENN et al., 1989). MALDI-MS é empregada normalmente para analisar misturas simples de peptídeos, enquanto a cromatografia líquida acoplada a ESI-MS é mais eficiente em amostras mais complexas (AEBERSOLD; MANN, 2003).

Os principais analisadores de massa que acompanham os sistemas de ionização, responsáveis pela acurácia dos espectros são o aprisionamento de íons (*Ion Trape* – IT), o tempo de voo (*Time of Flight* – TOF), o quadrupolo (Q) e a transformada de Fourier (*Fourier transform ion ciclotron* – FT) (MANN; HENDRICKSON; PANDEY, 2001). Alguns destes podem ser associados em série (tandem) para unir as vantagens de cada um, como o Q-IT, TOF-TOF, Q-TOF, dentre outros (AEBERSOLD; MANN, 2003).

Uma análise bastante utilizada para a caracterização e identificação proteica é a *Peptide Sequence* ou *Peptide Sequence Tag Query*, que emprega a espectrometria de massa em tandem (MS/MS), possibilitando a fragmentação dos polipeptídios, gerados por digestão enzimática. Neste procedimento são usados espectrômetros que possuem analisadores de sequência para CID (*Collision Induced Dissociation*) e para separação de fragmentos formados. O valor de  $m/z$  analisado é o dos íons-produto da fragmentação, além do íon precursor ou parental. Empregam-se *softwares* especiais que realizam as identificações proteicas pela comparação de uma sequência parcial, aliada à massa total do peptídeo que gerou tal sequência e a diferença de massa entre a sequência parcial com aquelas presentes em bancos de dados (CHAKRAVARTI; CHAKRAVARTI; MOUATSOS, 2002).

### 5.3 Análise *in silico*

A análise *in silico* dos dados obtidos pela espectrometria de massa (MS) é realizada com os peptídeos oriundos da digestão da proteína com uma protease específica e os resultados dos dados de massa são correlacionados com os dados de proteínas armazenados nos bancos de dados. Esta correlação é feita por *softwares* desenvolvidos para busca, comparação e certificação da identificação da proteína com análises estatísticas (MANN; HENDRICKSON; PANDEY, 2001).

Para tanto, as sequências depositadas são fragmentadas teoricamente (*in silico*) pelo mesmo método de digestão experimental e os valores das massas dos peptídeos teóricos são calculados e comparados com os valores experimentais. Essa abordagem de busca de coincidências de massa peptídica foi associada ao termo “identificação de assinaturas peptídicas” (PMF), pois cada proteína digerida por protease, como a tripsina, apresenta um único conjunto de peptídeos

específicos, semelhante à particularidade da impressão digital de cada pessoa (ASHCROFT, 2003; THIEDE; KOEHLER, 2010).

Para as análises de PMF é preciso que os valores de massa sejam obtidos com exatidão. Alguns programas que usam algoritmos sofisticados são utilizados na busca de PMF, como por exemplo, o “Bayesian” do programa ProFound, “MOWSE modificado” e “número de matches” no MSFit e “MOWSE baseado em probabilidade” no Mascot. A identificação da proteína depende da comparação entre as massas obtidas experimentalmente e aquelas obtidas teoricamente (STEEN; MANN, 2004).

## REFERÊNCIAS

AEBERSOLD, R.; MANN, M. Mass spectrometry-based proteomics. **Nature**, London, v. 422, n. 6928, p. 198-207, Mar. 2003.

AGARWAL, P. K. et al. Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. **Plant Cell Reports**, Berlin, v. 25, n. 12, p. 1263-1274, July 2006.

ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, London, v. 215, n. 3, p. 403-410, Oct. 1990.

AMARAL, J. F. T. **Eficiência de produção de raízes, absorção, translocação e utilização de nutrientes em cultivares de café arábica**. 2002. 97 p. Tese (Doutorado em Fitotecnia) - Universidade Federal de Viçosa, Viçosa, MG, 2002.

ASHCROFT, A. E. Protein and peptide identification: the role of mass spectrometry in proteomics. **Natural Product Reports**, London, v. 20, n. 2, p. 202-215, Feb. 2003.

ASSAD, E. D. et al. Climatic changes impact in agroclimatic zoning of coffee in Brazil. **Pesquisa Agropecuaria Brasileira**, Brasília, v. 39, n. 11, p. 1057-1064, nov. 2004.

BAN, N. et al. Placement of protein and RNA structures into a 5 Å: resolution map of the 50S ribosomal subunit. **Nature**, London, v. 400, n. 6747, p. 841-847, Aug. 1999.

BANDIL, G. B. **Análise proteômica comparativa do fruto de café (*Coffea arabica*) em dois estádios iniciais de desenvolvimento**. 2008. 84 p. Dissertação (Mestrado em Agronomia) - Universidade Estadual de Ponta Grossa, Ponta Grossa, 2008.

BANTSCHIEFF, M. et al. Quantitative mass spectrometry in proteomics: a critical review. **Analytical and Bioanalytical Chemistry**, Heidelberg, v. 389, n. 4, p. 1017-1031, Oct. 2007.

BARRETO, C. V. G. et al. Distribuição espacial do sistema radicular do cafeeiro fertirrigado por gotejamento em Campinas. **Bragantia**, Campinas, v. 65, n. 4, p. 641-647, ago. 2006.

BERTRAND, B. et al. Impact of the *Coffea canephora* gene introgression on beverage quality of *C. arabica*. **Theoretical and Applied Genetics**, Berlin, v. 107, n. 3, p. 387-394, Aug. 2003.

BOHNERT, H. J. et al. A genomic approach towards salt stress tolerance. **Plant Physiology and Biochemistry**, Paris, v. 39, n. 3/4, p. 295-311, Mar./Apr. 2001.

BORÉM, A. **Biotechnologia no melhoramento de plantas**: melhoramento de plantas. 2. ed. Viçosa, MG: UFV, 1998. 374 p.

BRASIL. Ministério da Agricultura Pecuária e Abastecimento. **Projeções do agronegócio Brasil 2012/2013 e 2022/2023**. Brasília, 2013. 96p. Disponível em: <<http://www.agricultura.gov.br/>>. Acesso em: 20 abr. 2014.

BRÉDA, N. et al. Temperate forest trees and stands under severe drought: a review of ecophysiological responses, adaptation processes and long-term consequences. **Annals of Forest Science**, Les Ulis, v. 63, n. 6, p. 625-644, Sept. 2006.

BRESHEARS, D. D. et al. Regional vegetation die-off in response to global-change-type drought. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 102, n. 42, p. 15144-15148, Oct. 2005.

BRESSAN, W.; VASCONCELOS, C. A. Alterações morfológicas no sistema radicular do milho induzidas por fungos micorrízicos e fósforo. **Pesquisa Agropecuária Brasileira**, Brasília, v. 37, n. 4, p. 509-517, abr. 2002.

CAMPOS, P. S. et al. Electrolyte leakage and lipid degradation account for cold sensitivity in leaves of *Coffea* sp plants. **Journal of Plant Physiology**, Stuttgart, v. 160, n. 1, p. 283-292, Jan. 2003.

CANNELL, M. G. R. Physiology of the coffee crop. In: \_\_\_\_\_. **Coffee**. Berlin: Springer, 1985. p. 108-134.

CANNELL, M. G. R.; HUXLEY, P. A. Seasonal differences in the pattern of assimilate movement in branches of *Coffea arabica* L. **Annals of Applied Biology**, Warwick, v. 64, p. 345-357, 1969.

CARVALHO, L. M. de et al. Aspectos morfofisiológicos das cultivares de cafeeiro Catuaí-Vermelho e Conilon. **Pesquisa Agropecuária Brasileira**, Brasília, v. 36, n. 3, p. 411-416, mar. 2001.

CHAKRAVARTI, D. N.; CHAKRAVARTI, B.; MOUATSOS, I. Informatic tools for proteome profiling. **Computational Proteomics Supplement**, Highlands, v. 32, n. 1, p. 4-15, Mar. 2002.

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.

COELHO, G. et al. Efeito de épocas de irrigação e de parcelamento de adubação sobre a produtividade do cafeeiro 'Catuaí'. **Ciência e Agrotecnologia**, Lavras, v. 33, n. 1, p. 67-73, jan./fev. 2009.

COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira: café**. Disponível em: <[http://www.conab.gov.br/OlalaCMS/uploads/arquivos/12\\_01\\_10\\_10\\_54\\_22\\_boletim\\_cafe\\_1a\\_estimativa\[1\].pdf](http://www.conab.gov.br/OlalaCMS/uploads/arquivos/12_01_10_10_54_22_boletim_cafe_1a_estimativa[1].pdf)>. Acesso em: 22 abr. 2014.

COSTA, T. S. et al. Análise preliminar para integração do perfil genômico e proteômico de raízes de *Coffea canephora* submetidos a diferentes condições hídricas. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 8., 2013, Salvador. **Anais...** Salvador: UFBA, 2013a. 1 CD-ROM.

COSTA, T. S. et al. Preliminary analysis for integrating transcripts and proteomic profiling of *Coffea canephora* roots subjected to different water conditions. In: WORKSHOP ON BIOTIC AND ABIOTIC STRESS TOLERANCE IN PLANTS, 21., 2013, Ilhéus. **Anais...** Ilhéus: CIBA, 2013b. 1 CD-ROM.

COSTE, R. **Caféiers et cafés**. Paris: Maisonneuve et Larose, 1989. 373 p.

CRUZ, F. et al. Evaluation of coffee reference genes for relative expression studies by quantitative real-time RT-PCR. **Molecular Breeding**, Dordrecht, v. 23, n. 4, p. 607-616, Feb. 2009.

DAMATTA, F. M. Drought as a multidimensional stress affecting photosynthesis in tropical tree crops. In: HEMANTARANJAN, A. (Ed.). **Advances in plant physiology**. Washington: Scientific, 2003. p. 227-265.

DAMATTA, F. M.; AMARAL, J. A.; RENA, A. B. Growth periodicity in trees of *Coffea arabica* L. in relation to nitrogen supply and nitrate reductase activity. **Field Crops Research**, Amsterdam, v. 60, n. 1, p. 223-229, Jan./Mar. 1999.

DAMATTA, F. M. et al. Drought tolerance of two field-grown clones of *Coffea canephora*. **Plant Science**, Shannon, v. 164, n. 1, p. 111-117, Jan. 2003.

DAMATTA, F. M. et al. Impacts of climate changes on crop physiology and food quality. **Food Research International**, Barking, v. 43, n. 7, p. 1814-1823, Aug. 2010.

DAMATTA, F. M. et al. Photosynthetic performance of two coffee species under drought. **Photosynthetica**, Prague, v. 34, n. 2, p. 257-264, Jan. 1997.

DAMATTA, F. M.; RAMALHO, J. D. C. Impacts of drought and temperature stress on coffee physiology and production. **Brazilian Journal of Plant Physiology**, Piracicaba, v. 18, n. 1, p. 55-81, Jan./Mar. 2006.

DAVEY, J. W. et al. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. **Nature**, London, v. 12, n. 7, p. 499-510, Aug. 2011.

DAVIS, A. P. et al. Growing coffee: *Psilantus* (Rubiaceae) subsumed on the basis of molecular and morphological data: implications for the size, morphology, distribution and evolutionary history of *Coffea*. **Botanical Journal of the Linnean Society**, London, v. 167, n. 4, p. 357-377, Dec. 2011.

DE KOCHKO, A. et al. Advances in *Coffea* genomics. In: KADER, J. C.; DELSENY, M. (Ed.). **Advances in botanical research**. New York: Elsevier, 2010. p. 23-53.

DETTMER, K.; ARONOV, P. A.; HAMMOCK, B. D. Mass spectrometry-based metabolomics. **Mass Spectrometry Review**, London, v. 26, n. 1, p. 51-78, June 2007.

DEUNER, S. et al. Stomatal behavior and components of the antioxidative system in coffee plants under water stress. **Scientia Agricola**, Piracicaba, v. 68, n. 1, p. 77-85, jan./fev. 2011.

DUSSERT, D. et al. Le caféier, *Coffea canephora*. In: HAMON, P. (Ed.). **Diversité génétique des plantes tropicales cultivées**. Montpellier: Cirad, 1999. p. 239-258.

FAZUOLI, L. C. Cultivares de café arábica do IAC: um patrimônio da cafeicultura brasileira. **O Agrônomo**, Campinas, v. 59, n. 1, p. 12-15, 2007.

FAZUOLI, L. C. et al. 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 Fosfato, 1986. p. 87-113.

FENN, J. B. et al. Electrospray ionization for mass spectrometry of large biomolecules. **Science**, New York, v. 246, p. 64-71, 1989.

FLORIS, M. et al. Post-transcriptional regulation of gene expression in plants during abiotic stress. **International Journal of Molecular Science**, Basel, v. 10, n. 7, p. 3169-3185, July 2009.

FREIRE, L. P. et al. Análise da expressão do gene manose 6 fosfato redutase em cafeeiros submetidos ao déficit hídrico. **Coffee Science**, Lavras, v. 8, n. 1, p. 17-23, jan./mar. 2013.

FREITAS, R. B. et al. Comportamento fisiológico de dois cultivares de *Coffea arabica* L. submetidos a duas condições de disponibilidade hídrica. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 2., 2000, Poços de Caldas. **Anais...** Brasília: EMBRAPA Café, 2000. p. 917-919.

GAUSE, W. C.; ADAMOVICZ, J. The use of the PCR to quantitate gene expression. **Genome Research**, Cold Spring Harbor, v. 3, n. 6, p. 123-135, June 1994.

GONÇALVES, W. Melhoramento do cafeeiro visando resistência a nematóide. In: SIMPÓSIO DE ATUALIZAÇÃO EM GENÉTICA E MELHORAMENTO DE PLANTAS, 3., 1999, Lavras. **Anais...** Lavras: UFLA, 1999. 1 CD-ROM.

GRAVES, P. R.; HAYSTEAD, T. A. Molecular biologist's guide to proteomics. **Microbiology and Molecular Biology Reviews**, New York, v.66, n. 1, p. 39-63, Mar. 2002.

GUIMARÃES, B. L. S. **Proteômica diferencial em clones de *Coffea canephora* sob condições de déficit hídrico**. 2007. 71 p. Dissertação (Mestrado em Fisiologia Vegetal) - Universidade Federal de Viçosa, Viçosa, MG, 2007.

GWAZE, D. P.; BRIDGWATER, F. E.; WILLIAMS, C. G. Genetic analysis of growth curves for a woody perennial species, *Pinus taeda* L. **Theoretical and Applied Genetics**, Berlin, v. 105, n. 4, p. 526-531, Sept. 2002.

HATANAKA, T. et al. Transgenic plants of coffee *Coffea canephora* from embryogenic callus via *Agrobacterium tumefaciens*-mediated transformation. **Plant Cell Reports**, Berlin, v. 19, n. 2, p. 106-110, Dec. 1999.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE. Disponível em: <<http://www.ipcc.ch/>>. Acesso em: 10 maio 2014.

INTERNATIONAL COFFEE ORGANIZATION. **Trade statistics**. Disponível em: <[http://www.ico.org/trade\\_statistics.asp?section=Statistics](http://www.ico.org/trade_statistics.asp?section=Statistics)>. Acesso em: 22 maio 2014.

KARAS, M.; HILLENKAMP, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. **Analytical Chemistry**, Washington, v. 60, p. 2299-2301, 1988.

KRAMER, P. J.; BOYER, J. S. **Water relations of plants and soils**. New York: Academic, 1995. 495 p.

KUMAR, L. S. DNA markers in plant improvement: an overview. **Biotechnology Advances**, New York, v. 17, n. 2/3, p. 143-182, Sept. 1999.

LAMBOT, C. et al. Evaluation of Conilons for genetic diversity, cup quality and biochemical composition. In: INTERNATIONAL CONFERENCE ON COFFEE SCIENCE, 22., 2008, Campinas. **Proceedings...** Campinas: ASIC, 2009. p. 1537-1544.

LANÇAS, F. M. **Cromatografia líquida moderna: HPLC/CLAE**. Campinas: Átomo, 2009. 382 p.

LANÇAS, F. M.; MCNAIR, H. Cromatografia em fase gasosa: I., teoria elementar. **Química Nova**, São Paulo, v. 6, n. 1, p. 6-13, 1983.

LARSEN, M. R. et al. Graphite powder as an alternative or supplement to reversed-phase material for desalting and concentration of peptide mixtures prior to matrix-assisted laser desorption/ionization-mass spectrometry. **Proteomics**, Weinheim, v. 2, n. 9, p. 1277-1287, Sept. 2002.

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.

LEROY, T. et al. Genetics of coffee quality. **Brazilian Journal of Plant Physiology**, Piracicaba, v. 18, n. 1, p. 229-242, Jan./Mar. 2006.

LIN, C. et al. Coffee and tomato share common gene repertoires as revealed by deep sequencing of seed and cherry transcripts. **Theoretical and Applied Genetics**, Berlin, v. 112, n. 1, p. 114-130, Dec. 2005.

LOPES, F. R. et al. Transcriptional activity, chromosomal distribution and expression effects of transposable elements in coffee genomes. **PloS One**, San Francisco, v. 8, n. 11, Nov. 2013. Disponível em: <<http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0078931>>. Acesso em: 10 mar. 2014.

LYNCH, L. Root architecture and plant productivity. **Plant Physiology**, Bethesda, v. 109, n. 1, p. 7-13, Sept. 1995.

MACKAY, I. M. Real-time PCR in the microbiology laboratory. **Clinical and Microbiology and Infection**, Paris, v. 10, n. 3, p. 190-212, Mar. 2004.

MAJOR coffee producers. Disponível em: <<http://www.nationalgeographic.com/coffee/map.html#1>>. Acesso em: 10 maio 2014.

MANN, M.; HENDRICKSON, R. C.; PANDEY, A. Analysis of proteins and proteomes by mass spectrometry. **Annual Review of Biochemistry**, Palo Alto, v. 70, n. 1, p. 437-473, July 2001.

MARDIS, E. R. Next-generation DNA sequencing methods. **Annual Review of Genomics and Human Genetics**, Palo Alto, v. 9, p. 387-402, 2008.

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

MARRACCINI, P. et al. *RBCS1* expression in coffee: *Coffea* orthologs, *Coffea arabica* homeologs, and expression variability between genotypes and under drought stress. **BMC Plant Biology**, London, v. 11, n. 85, p. 1-23, May 2011.

MATIELLO, J. B.; DANTAS, S. F. Desenvolvimento do cafeeiro e do sistema radicular, com e sem irrigação, em Brejões, PE. In: CONGRESSO BRASILEIRO DE PESQUISAS CAFEEIRAS, 14., 1987, Campinas. **Resumos...** Campinas: Instituto Brasileiro do Café, 1987. p. 165.

MAZZUCOTELLI, A. Abiotic stress response in plants: when post-transcriptional and post-translational regulations control transcription. **Plant Science**, Shannon, v. 174, n. 4, p. 420-431, Apr. 2008.

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. 1-22, Feb. 2011.

MONDEGO, J. M. C. et al. Isolation and characterization of *Coffea* genes induced during coffee leaf miner (*Leucoptera coffeella*) infestation. **Plant Science**, Shannon, v. 169, n. 2, p. 351-360, Aug. 2005.

MONTAGNON, C.; CUBRY, P.; LEROY, T. Amélioration génétique du caféier *Coffea canephora* Pierre: connaissances acquises, stratégies et perspectives. **Cahiers Agricultures**, Madrid, v. 21, n. 2/3, p. 143-153, mar./jun. 2012.

MONTAGNON, C. et al. Genotype-location interactions for *Coffea canephora* yield in the Ivory Coast. **Agronomie**, Paris, v. 20, n. 1, p. 101-109, Jan. 2000.

MONTOYA, G. et al. Sequence analysis from leaves, flowers and fruits of *Coffea arabica* var Caturra. In: INTERNATIONAL CONFERENCE ON COFFEE SCIENCE, 21., 2006, Montpellier. **Proceedings...** Montpellier: ASIC, 2006.1 CD-ROM.

MORAIS, T. P. de; MELO, B. de. Biotecnologia aplicada ao melhoramento genético do cafeeiro. **Ciência Rural**, Santa Maria, v. 41, n. 5, p. 756-760, May 2011.

MULLIS, K. B.; FALOONA, F. A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. **Methods in Enzymology**, New York, v. 155, p. 335-350, 1987.

MURAD, A. M.; RECH, E. NanoUPLC-MSE proteomic data assessment of soybean seeds using the Uniprot database. **BMC Biotechnology**, London, v. 12, n. 82, p. 1-17, Nov. 2012.

NAGARAJ, S. H. et al. EST Explorer: an expressed sequence tag (EST) assembly and annotation platform. **Nucleic Acids Research**, Oxford, v. 35, n. 1, p. 143-147, July 2007.

NASCIMENTO, M. N. et al. Biochemical alterations of plants and bud morphology of coffee tree associated to events on flowering in response to meteorological elements. **Ciência Rural**, Santa Maria, v. 38, n. 5, p. 1300-1307, ago. 2008.

NELSON, G. C. et al. **Climate change: impact on agriculture and costs of adaptation**. Washington: International Food Policy Research Institute, 2009. 19 p.

NOVAES, C. M.; PIRES-ALVES, M. PCR em tempo real: uma inovação tecnológica da reação em Cadeia Polimerase - PCR. **Biotecnologia Ciência e Desenvolvimento**, Brasília, v. 33, n. 1, p. 10-13, jul./dez. 2004.

PARTELLI, F. L. et al. Cold impact on photosynthetic related parameters in coffee genotypes. **Pesquisa Agropecuária Brasileira**, Brasília, v. 44, n. 11, p. 1404-1415, nov. 2009.

PARTELLI, F. L. et al. Produção e desenvolvimento radicular de plantas de café “Conilon” propagadas por sementes e por estacas. **Pesquisa Agropecuária Brasileira**, Brasília, v. 41, n. 6, p. 949-954, jun. 2006.

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, July 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 and Genomics**, Berlin, v. 276, n. 5, p. 436-449, Aug. 2006.

PRAKASH, N. S. et al. Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into *C. arabica* L. **Theoretical and Applied Genetics**, Berlin, v. 109, n. 6, p. 1311-1317, July 2004.

PRAXEDES, S. C. et al. Effects of long-term soil drought on photosynthesis and carbohydrate metabolism in mature robusta coffee (*Coffea canephora* Pierre var. kouillou) leaves. **Environmental and Experimental Botany**, Elmsford, v. 56, n. 3, p. 263-273, July 2006.

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

REDDY, R. A.; CHAITANYA, K. V.; VIVEKANANDAN, M. Drought-induced responses of photosynthesis and antioxidant metabolism in higher plants. **Journal of Plant Physiology**, Stuttgart, v. 161, n. 11, p. 1189-1202, Nov. 2004.

RENA, A. B.; GUIMARÃES, P. T. G. **Sistema radicular do cafeeiro: estrutura, distribuição, atividade e fatores que o influenciam**. Belo Horizonte: EPAMIG, 2000. 80p.

RIBAS, A. F.; PEREIRA, L. F. P.; VIEIRA, L. G. Genetic transformation of coffee. **Brazilian Journal of Plant Physiology**, Piracicaba, v. 18, n. 1, p. 83-94, Jan./Mar. 2006.

RIVIER, J.; MCCLINTOCK, R. Reversed-phase high performance liquid chromatography of insulins from different species. **Journal of Chromatography**, Amsterdam, v. 268, p. 112-119, 1983.

SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 74, p. 5463-5467, 1977.

SANTOS, F. R.; ORTEGA, J. M. Bioinformática. In: BORÉM, A.; DEL GIÚDICE, M.; SEDYIAMA, T. (Ed.). **Melhoramento genômico**. Viçosa, MG: UFV, 2003. v.1, p. 165-185.

SARGENT, D. J. et al. The development and mapping of functional markers in *Fragaria* and their transferability and potential for mapping in other genera. **Theoretical and Applied Genetics**, Berlin, v. 114, n. 2, p. 373-384, Jan. 2007.

SCARBOROUGH, R. M. et al. Isolation and primary structure of two peptides with cardio acceleratory and hyperglycemic activity from the corpora cardiac of periplaneta Americana. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 81, p. 5575-5579, 1984.

SCHROEDER, J. I.; KWAK, J. M.; ALLEN, G. J. Guard cell abscisic acid signalling and engineering drought hardiness in plants. **Nature**, London, v. 410, p. 327-330, Mar. 2001.

SEKI, M. et al. Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. **Plant Cell**, Rockville, v. 13, n. 1, p. 61-72, Jan. 2001.

SHINOZAKI, K. et al. Regulatory network of gene expression in the drought and cold stress responses. **Current Opinion in Plant Biology**, Oxford, v. 6, n. 5, p. 410-417, Oct. 2003.

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**, Oxford, v. 3, n. 3, p. 217-223, June 2000.

SILVA, E. A. da et al. The influence of water management and environmental conditions on the chemical composition and beverage quality of coffee beans. **Brazilian Journal Plant Physiology**, Piracicaba, v. 17, n. 2, p. 229-238, Apr./June 2005.

SILVA, E. A. da et al. Seasonal changes in vegetative growth and photosynthesis of Arabica coffee trees. **Field Crops Research**, Amsterdam, v. 89, n. 2/3, p. 349-357, Oct. 2004.

SILVA, J. C. et al. Absolute quantification of proteins by LCMS<sup>E</sup>: a virtue of parallel MS acquisition. **Molecular & Cellular Proteomics**, Rockville, v. 5, n. 1, p. 144-156, Jan. 2005.

SILVA, O. M.; LEITE, C. A. M. Competitividade e custos do café no Brasil e no exterior. In: ZAMBOLIM, L. (Ed.). **Café: produtividade, qualidade e sustentabilidade**. Viçosa, MG: UFV, 2000. p. 27-48.

SILVA, V. A. et al. Resposta fisiológica de clone de café Conilon sensível à deficiência hídrica enxertado em porta-enxerto tolerante. **Pesquisa Agropecuária Brasileira**, Brasília, v. 45, n. 5, p. 457-464, maio 2010.

SOARES, B. M. et al. Produção e distribuição de raízes em cafeeiros irrigados por gotejamento. **Engenharia na Agricultura**, Viçosa, MG, v. 15, n. 3, p. 130-140, abr./jun. 2007.

STEEN, H.; MANN, M. The ABC's (and XYZ's) of peptide sequencing. **Nature Reviews Molecular Cell Biology**, London, v. 5, n. 9, p. 699-711, Sept. 2004.

TEDESCO, N. S. et al. Marcadores OPF-151280 e OPC-091120 ligados ao gene de resistência à ferrugem do cafeeiro (*Hemileia Vastatrix* Berk. Et Br). In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 1., 2000, Poços de Caldas. **Anais...** Poços de Caldas: SBC, 2000.p. 169-171.

THIEDE, B.; KOEHLER, C. **Mass spectrometry-based quantitative proteomics using isobaric peptide termini labeling**. Oslo: Universit Oslo, 2010. 47p.

THOMAZIELLO, R. A. et al. **Café arábica: cultura e técnicas de produção**. Campinas: IAC, 2000. 82 p. (Boletim Técnico, 187).

TURNER, N. C. Further progress in crop water relations. **Advances in Agronomy**, San Diego, v. 58, n. 1, p. 298-338, Apr. 1997.

URANO, K. et al. ‘Omics’ analyses of regulatory networks in plant abiotic stress responses. **Current Opinion in Plant Biology**, Oxford, v. 13, n. 2, p. 132-138, Apr. 2010.

VALASEK, M. A.; REPA, J. J. The power of real-time PCR. **Advances in Physiology Education**, Bethesda, v. 29, n. 3, p. 151-159, Sept. 2005.

VIDAL, R. et al. A high-throughput data mining of single nucleotide polymorphisms in *Coffea* species expressed sequence tags suggests differential homeologous gene expression in the allotetraploid *Coffea arabica*. **Plant Physiology**, Bethesda, v. 154, n. 3, p. 1053-1066, Nov. 2010.

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

VIEIRA, N. G. et al. Different molecular mechanisms account for drought tolerance in *Coffea canephora* var. Conilon. **Tropical Plant Biology**, New York, v. 6, n. 4, p. 181-190, Dec. 2013.

VINECKY, F.; SILVA, F. R. da; ANDRADE, A. C. Análise *in silico* das bibliotecas de cDNA SH2 e SH3 para a identificação de genes responsivos à seca em cafeeiro. **Coffee Science**, Lavras, v. 7, n. 1, p. 1-19, jan./abr. 2012.

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, Nov. 2003.

WIJK, K. J. van. Challenges and prospects of plant proteomics. **Plant Physiology**, Bethesda, v. 126, n. 2, p. 501-508, June 2001.

WILDGRUBER, R. et al. Towards higher resolution: two-dimensional electrophoresis of *Saccharomyces cerevisiae* proteins using overlapping narrow immobilized pH gradients. **Electrophoresis**, Weinheim, v. 21, n. 13, p. 2610-2616, July 2000.

WU, H. X.; MATHESON, A. C. Genotype by environment interactions in an Australia-wide radiata pine diallel mating experiment: implications for regionalised breeding. **Forest Science**, Lawrence, v. 51, n. 13, p. 29-40, Feb. 2005.

YERGEY, A. L. et al. De novo sequencing of peptides using MALDI/TOF-TOF. **Journal of the American Society for Mass Spectrometry**, New York, v. 13, n. 7, p. 784-791, July 2002.

ZATZ, M. M. Bioinformatics training in the USA. **Briefings in Bioinformatics**, London, v. 3, n. 4, p. 353-360, Dec. 2002.

ZHANG, X. et al. Biosolids impact on tall fescue drought resistance. **Journal of Residuals Science & Technology**, Denton, v. 2, n. 3, p. 173-180, 2005.

ZHANG, X. et al. Hydrogen peroxide is involved in abscisic acid-induced stomatal closure in *Vicia faba*. **Plant Physiology**, Bethesda, v. 126, n. 4, p. 1438-1448, Aug. 2001.

ZHU, T. et al. Toward elucidating the global expression patterns of developing *Arabidopsis*: parallel analysis of 8300 genes by a high-density oligonucleotide probe array. **Plant Physiology and Biochemistry**, Dordrecht, v. 39, n. 3, p. 221-242, Mar. 2001.

## **CAPÍTULO 2 Caracterização de genes diferencialmente expressos em raízes de clones de *Coffea canephora* sensíveis e tolerantes ao estresse hídrico**

### **RESUMO**

Genes que codificam proteínas funcionais são alvos de pesquisa para o entendimento dos mecanismos que regem a fisiologia do cafeeiro e, como tal, a tolerância à seca. Em raízes, estudos sobre os genes relacionados com o desenvolvimento radicular da planta sob seca vem sendo realizados e são de grande importância para o melhoramento genético do cafeeiro. O presente trabalho tem como objetivo caracterizar o perfil de expressão gênica em raízes de clones de *C. canephora* variedade Conilon sensíveis e tolerantes, que são diferencialmente expressos quanto ao déficit hídrico quando submetidos a condições controladas e sob estresse hídrico, além de selecionar Genes Candidatos (GC) que indiquem respostas do cafeeiro a este evento e validá-los por meio da PCR Quantitativa (qPCR). Para cada clone e regime hídrico, o RNA total foi extraído e utilizado tanto para obtenção do perfil do transcriptoma das raízes, pelo sequenciamento 454, que possibilitou análises *in silico* da expressão por *Northern* eletrônico, bem como para a validação por qPCR dos genes selecionados entre os clones nas condições I vs. NI. Foi possível verificar que, entre os clones tolerantes há diferentes mecanismos de respostas ao estresse. A técnica de pirosequenciamento 454 permitiu gerar mais de quatro milhões de *reads* a partir de oito bibliotecas de raízes de *C. canephora* provenientes dos clones nas duas condições, o que possibilitou a ampliação dos dados para este tecido. Vários genes candidatos foram identificados e avaliados quanto à resposta da planta ao déficit hídrico, bem como em relação à biossíntese do ABA e aos estresses osmóticos e oxidativos. Com base nos resultados foi possível obter uma melhor compreensão de alguns mecanismos de resposta em raízes em clones *C. canephora* quanto a tolerância ao déficit hídrico, além de verificar que existe resposta diferencial para os mecanismos de resposta ao estresse entre os clones tolerantes avaliados.

Palavras-chave: Cafeeiro. Raiz. Tolerância à seca. Expressão gênica. Pirosequenciamento.

**Characterization of genes differentially expressed in *Coffea canephora* clone roots sensitive and tolerant to water stress**

**ABSTRACT**

Genes that codify functional proteins are research targets for understanding the mechanisms that guide coffee plant physiology and, as such, the tolerance to drought. For roots, studies on genes related to the radicular development of the plant under drought have been performed and are of great importance for the genetic improvement of coffee. The present work had the objective of characterizing the profile of gene expression in sensitive and tolerant *C. canephora* clone roots of the variety Conilon, which are differentially expressed regarding water deficit when submitted to controlled conditions and under water stress, in addition to selecting Candidate Genes (CG) that indicate responses of the coffee plant to this event and validate them by means of Quantitative PCR (qPCR). For each clone and water regime, the total RNA was extracted and used for obtaining the transcriptome profile of the roots, by means of the 454 sequencing, which allowed the *in silico* analysis of the expression by electronic *Northern*, as well as to validate the genes selected between the clones in conditions I vs. NI, by means of qPCR. It was possible to verify that, between the tolerant clones, there are different response mechanisms for stress. The 454-pyrosequencing technique allowed the generation of more than four million reads from eight *C. canephora* root libraries originated from the clones on both conditions, which allowed the amplification of the data for this tissue. Many candidate genes were identified and evaluated regarding the plant's response to water stress, as well as concerning the biosynthesis of ABA and the osmotic and oxidative stresses. Based on the results, it was possible to obtain a better understanding of a few response mechanisms for *C. canephora* clone roots regarding the tolerance to water deficit, in addition to verifying that there is differential response for the response mechanisms to stress between the tolerant clones evaluated.

Keywords: Coffee plant. Root. Drought tolerance. Gene expression. Pyrosequencing.

## 1 INTRODUÇÃO

O estresse ambiental é um grande desafio para a produção de alimentos, uma vez que reduz os rendimentos em até 70% (AGARWAL et al., 2006). Em virtude do aquecimento global, estresses abióticos como seca e chuvas abundantes são cada vez mais frequentes no Brasil. Dentre os fatores climáticos que afetam o país, a seca é o mais preocupante. Altas temperaturas prolongadas levam à morte das plantas e, em períodos moderados, afetam a floração, o desenvolvimento do grão, e, conseqüentemente, a produção do café (MARRACCINI et al., 2011).

Genes que codificam proteínas funcionais bem como proteínas regulatórias ou fatores de transcrição estão envolvidos na regulação da transmissão de sinais e expressão gênica e, portanto, são alvos de pesquisa para entendimento dos mecanismos que regem a fisiologia do cafeeiro e, como tal, a tolerância à seca (AGARWAL et al., 2006).

Diversidade genética e fenotípica para a tolerância à seca, já foi observada em plantas de *C. canephora* do grupo Conilon, sendo que alguns genótipos foram considerados mais eficientes na tolerância à seca, uma vez que possuem maior controle de transpiração foliar, maior aprofundamento e distribuição radicular e melhor adaptação a climas quentes (DAMATTA; RAMALHO, 2006; LEROY et al., 2006; PINHEIRO et al., 2005).

Diversos fatores influenciamno desenvolvimento das raízes: diferenças genéticas, produção e disponibilidade de carboidratos, variação sazonal, ambiente físico e químico do solo, dentre outros (BRESSAN; VASCONCELOS, 2002). Portanto, o conhecimento radicular de uma espécie proporciona um melhor entendimento da fisiologia da planta como um todo (PARTELLI et al., 2006; SOARES et al., 2007).

Sequências de genes encontrados em café já foram identificadas para as espécies *C. arábica*, *C. canephora* e *C. racemosa* (MONTTOYA et al., 2006; PONCET et al., 2006; VIEIRA et al., 2006). E novas sequências são constantemente depositadas e utilizadas em pesquisas para a identificação de perfis de expressão que fornecem informações a respeito das respostas fisiológicas das plantas em situações de estresses bióticos e abióticos, bem como, relacionadas às qualidades de bebida.

A produção de variedades de café tolerantes à seca é uma das prioridades de muitos institutos de pesquisas e a utilização de técnicas biotecnológicas como PCR quantitativa em Tempo Real (qPCR), Espectrometria de Massa (MS), gel bidimensional de proteínas, ferramentas de bioinformática dentre outros são o foco das pesquisas. Diversos trabalhos têm evidenciado os avanços na área de biotecnologia vegetal e as técnicas apresentam-se rápidas, eficientes e úteis para aprimorar o entendimento dos mecanismos de tolerância do cafeeiro aos diversos estresses (DAMATTA; RAMALHO, 2006; DUNWELL, 2000; MARRACCINI et al., 2012; RIBAS; PEREIRA; VIEIRA, 2006; WANG; VINOUCUR; ALTMAN, 2003).

Desta forma, o objetivo do presente trabalho foi caracterizar o perfil de expressão gênica em raízes de clones de *C. canephora* sensíveis e tolerantes, diferencialmente expressos quanto ao déficit hídrico nas condições de irrigação (I) e não irrigação (NI), além de selecionar Genes Candidatos (GC) que indiquem resposta do cafeeiro a este evento e validá-los por meio da PCR Quantitativa (qPCR).

## 2 MATERIAL E MÉTODOS

### 2.1 Material Vegetal

Para as análises foram utilizadas raízes de quatro clones de *Coffea canephora* var Conilon (14, 22, 73 e 120). Estes clones foram anteriormente mantidos em estacas no Instituto de Pesquisa e Assistência Rural (Incaper, Vitória, Espírito Santo, Brasil), cada estaca tinha aproximadamente 12 meses e 40 cm de altura (FERRÃO et al., 2000; FONSECA et al., 2004). Estas estacas foram transportadas para a Universidade de Viçosa (UFV, Minas Gerais, Brasil), onde foram plantadas em vasos de 12 litros e aclimatadas em casa de vegetação em condições controladas de temperatura (25°C), umidade relativa (70%) e fluxo fotossintético (PPF de 900  $\mu\text{mol}^{-2}\text{s}^{-1}$ ).

Em seguida, foram submetidas a ensaios de estresse e amostragens de tecidos. Para tal, após seis meses de aclimação, as plantas passaram por dois tipos de tratamento: em condições de irrigação (I) e sem irrigação (NI).

Diversos parâmetros foram utilizados para avaliar o nível de estresse hídrico dos clones analisados, conforme apresentado por Marraccini et al. (2011):  $\Psi_{\text{am}}$  (Potencial hídrico de antemanhã na folha), RDPWP (taxa de diminuição do  $\Psi_{\text{am}}$ ), A (taxa de assimilação de CO<sub>2</sub>), gs (condutância estomatal), Ci / Ca (taxa de concentração de CO<sub>2</sub> no ambiente interno), Fv / Fm (eficiência fotoquímica máxima do FSII),  $\Phi\text{PSII}$  (rendimento quântico do transporte de elétrons), qp (coeficiente de extinção fotoquímica), qN (coeficiente de extinção não fotoquímica de Stern Volmer), PE (fração de energia não utilizada na fotoquímica nem dissipada termicamente).

Para a condição NI, a irrigação das plantas foi suspensa até o momento em que o potencial hídrico de antemanhã ( $\Psi_{\text{am}}$ ) atingiu o valor de -3,0 Mpa. O clone 22, considerado sensível ao déficit hídrico, obteve o valor de  $\Psi_{\text{am}}$  após seis

dias sem irrigação, os clones 14 e 120 após 12 dias, e o clone 73 em 15 dias (Figura 1). A taxa de diminuição do potencial de água na folha antes do amanhecer-RDPWP diferenciou as respostas entre os clones 14 e 120, uma vez que houve atraso na diminuição de RDPWP para o clone 120 em comparação ao clone 14 (MARRACCINI et al., 2011). Para cada clone e regime hídrico, seis plantas (repetições biológicas) foram analisadas.

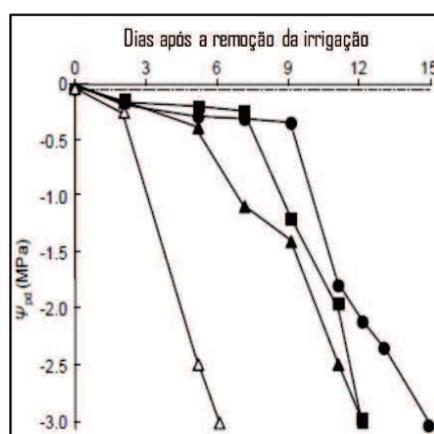


Figura 1 Evolução da perda de água nas folhas dos clones de *C. canephora*. Dados baseados no potencial hídrico de antemanhã ( $\Psi_{am}$ )

Legenda:  $\Delta$  clone 22,  $\blacktriangle$  clone 14,  $\bullet$  clone 73 e  $\blacksquare$  clone 120

Fonte: Marraccini et al. (2011), adaptado.

Ao final do experimento, 5 gramas de raízes das plantas foram coletadas para as análises moleculares, nas mesmas condições fisiológicas já descritas por Marraccini et al. (2011) e Vieira et al. (2013) para folhas. Em seguida, foram imediatamente congeladas em nitrogênio líquido e armazenadas em freezer -80°C. Para a extração de RNA Total e proteínas, inicialmente foi preparado um *pool* das seis repetições biológicas destas raízes para cada clone e em cada condição.

## 2.2 Extração de RNA

A extração do RNA total foi efetuada utilizando o protocolo sugerido pelo fabricante TRIzol® Reagent (Life technologies), com adaptações. Para evitar a oxidação do material, uma etapa preliminar foi realizada com a adição de 300 µL de tampão CTAB (CTAB 2%, Tris-HCL pH 8,0 1M, EDTA 0,5M pH 8,0, NaCl 1,4M, PVP 1%) e 06 µL de β-mercaptoetanol 2%.

Em seguida, acrescentou-se 306 µL de clorofórmio gelado e centrifugou-se durante 10 minutos a 4°C, velocidade de 15.000 x g. Após a centrifugação, foi recolhido o sobrenadante e este foi transferido para novos tubos, onde foi adicionado 800 µL de TRIzol® por repetição. As amostras permaneceram em repouso por cinco minutos à temperatura ambiente, em seguida, foram adicionados 500 µL de clorofórmio (por amostra) e a solução foi misturada por cerca de 30 segundos e centrifugada (15min, 4°C, 15.000 x g), como descrito anteriormente. A fase aquosa foi transferida para um novo tubo e 400 µL de isopropanol foi adicionado. Os tubos permaneceram em repouso por cerca de 10 minutos à temperatura ambiente. Posteriormente, as amostras foram centrifugadas (15min, 4°C, 15.000 x g). O sobrenadante foi removido e lavado com 600 µL de etanol 70% e, em seguida, novamente centrifugado (5min, 4°C, 9,500 x g). O etanol foi removido e o material foi seco à 37°C para que o *pellet* fosse ressuscitado em 25µL de água miliQ.

Para avaliar a qualidade e concentração dos RNAs, as amostras foram quantificadas em NanoVue Plus (GE Healthcare, EUA) e por meio de eletroforese, em gel de agarose 0,8% e corrida em d.d.p de 80v. Após a análise de qualidade do RNA, procedeu-se o tratamento enzimático por meio do Kit RQ1 DNaseI-RNase “free” (Promega) para a retirada dos resquícios de DNA genômico.

### 2.3 Identificação *in silico* de genes candidatos

O sequenciamento do cDNA dos clones de *C. canephora*, nas condições I e NI, foi realizado pelo método de pirosequenciamento-454, plataforma Roche. Para a análise *in silico*, a expressão dos genes foi avaliada por *Northern* eletrônico utilizando o pacote DNASTar, *software* Q-seq, versão 7.0 (Lasergene), tendo por base a quantificação dos *reads* mapeados na sequência de referência de cada gene, cujos valores foram normalizados por RPKM (Reads Per Kilobases Per Million), portanto, a expressão reflete a quantidade de *reads* do transcrito normalizado pelo comprimento do RNA e pelo número total de *reads* na amostra (MORTAZAVI et al., 2008).

$$\text{RPKM} = \frac{10^6 C}{\text{NL}/10^3}$$

Onde C é o número de reads que unicamente alinharam no gene A;

N é o número de reads únicos mapeados em todos os genes (número de reads mapeados);

L é o número de bases do gene A.

Foram selecionados 100 genes como candidatos à tolerância à seca em cafeeiros, de acordo com o fator de indução (FI), calculado pelo quociente da expressão do gene, em cada clone, na condição NI e pela expressão do mesmo gene no respectivo clone na condição I. Alguns genes foram escolhidos por estarem ligados na transcrição de proteínas de alto valor quantitativo em raízes (com base no experimento de proteômica) e, alguns genes porque já foram relatados em trabalhos anteriores relacionados com a resposta do cafeeiro à tolerância à seca.

A partir da seleção dos genes candidatos, foram sintetizados pares de *primers*, que possibilitaram a validação da expressão gênica por meio de qPCR. Todos os *primers* sintetizados foram testados quanto à eficiência e, somente aqueles que apresentaram resultados mais consistentes quanto à amplificação e, conseqüentemente, à expressão do gene, foram utilizados nas análises de qPCR. Foram utilizados 31 pares de *primers* para a análise de expressão quantitativa de seus respectivos genes, conforme ilustrado na Tabela 1.

Para a normalização e standardização da expressão dos genes, foi utilizado o gene que expressa a Ubiquitina-*CcUBQ10* como calibrador interno de referência, por ser considerada uma proteína constitutiva, na qual a expressão é constante independente da condição em que planta se encontra (BARSALOBRES-CAVALLARI et al., 2009). A amostra 14I foi utilizada como calibrador interno, apresentando sempre quantificação relativa 1.

Tabela 1 Genes candidatos e respectivos *primers*, selecionados com base no Fator de Indução (FI), utilizados nos experimentos de qPCR. **GB** Número do acesso de proteínas registrados no GenBank. Os *primers* foram sintetizados utilizando o *software* Primer Express (Applied Biosystems). **pb** pares de bases.

Config GB	Nome	Sequência do primer	Tamanho do Amplicon (pb)
<b>Proteínas regulatórias</b>			
21488	<i>CcNCED3</i> (9-cis-epoxycarotenoid dioxygenase 3)	F 5'-GCCTGGGAAGAGCCTGAAAC-3'	90
ABA43901.1		R 5'-CCCCTCGTCACATTCATTGAA-3'	
PYL 7	<i>CcPYL7.2</i>	F 5'-GAGAAAGCACATTC TTGGGATCAA-3'	80
DV704550		R 5'-GGATGCACGGTAAGGATGGA-3'	
21820	<i>CcPP2C-1</i> (protein phosphatase 2C ABI2 homologue)	F 5'-ATGGCTTGTGGGATGTCATGA-3'	81
NP-001234686.1		R 5'-CGTTCTTCTTGTGCCAAAGCA-3'	
34982	<i>CcAREB</i> (ABA responsive element-binding protein)	F 5'-TGCCTTATGCCACTCCCATT-3'	78
AFA37978.1		R 5'-GTCCAACAATCCCACCCCTAA-3'	
06253	<i>CcNAC</i> (NAC domain)	F 5'-CCATGGGAGCTCAATGAGAAAG-3'	79
XP-002264894.2		R 5'-TAGCTGCGCCCTTGGATCACTT-3'	
21185	<i>CcNAC4</i> (NAC domain protein NAC4)	F 5'-AGGTCATCACCCGAAAGGAC-3'	80
ACI15345.1		R 5'-GGTGCCTTTAGGTGCTTTGC-3'	
39597	<i>CcNAC</i> (PREDICTED: NAC domain)	F 5'-TTTCTCCCCTGCTTCCTTGATT-3'	80
XP-002283811.1		R 5'-CCATATGCGGCATGATCTGA-3'	
21916	<i>CcPP3B</i> (putative acid phosphatase 2B homologue HAD)	F 5'-GCAGCAACTTCTCATGCAACA-3'	74
AFX66999.1		R 5'-GCATGGGAGGTTTGGTCAATC-3'	

“Tabela 1, continua”

Contig GB	Nome	Sequência do primer	Tamanho do Amplicon (pb)
<b>Proteínas regulatórias</b>			
22022	<i>CcPP3B</i> (putative acid phosphatase 2B homologue HAD)	F 5'-ATTTGTTCTGACGGCGGAAAGT -3' R 5'-CCGCCCTCGACTATAGCTTGGTA -3'	71
<b>Proteínas funcionais</b>			
24390	<i>CcABA</i> 8-hidroxiase (ABA 8'-hidroxylase CYP707A1)	F 5'-CTGCGATTTGCTCGGATCAT -3' R 5'-CACGGCAGCAAAGATAACA -3'	95
33650	<i>CcSRP-ABA</i> (abscisic acid stress ripening protein)	F 5'-AGTGGGTGCCGGTGGATT -3' R 5'-TTCCCTCAGCCCTTTTCTCTTCTT -3'	80
40232	<i>CcMLTP2</i> (putative non-specific lipid-transfers - subfamily 2)	F 5'-GTTGCAGCCTGGTTGGTGAT -3' R 5'-ACGGACTCAGCTCTGCTATGCT -3'	91
22121	<i>CcLEA</i> (late embryogenesis abundant protein)	F 5'-ATGACGCTGCCCATGAGAAG -3' R 5'-GTGGGTTCTGAAA TCTTGTGTC -3'	78
37166	<i>CcDHIa</i> (dehydrin DH1a)	F 5'-AAGGCGGAGGAGGAAAGAA -3' R 5'-GTCCAGGTTGAGCCTCCTTGT 3'	80

“Tabela 1, continuação”

Contig GB	Nome	Sequência do primer	Tamanho do Amplicon (pb)
14237 ABC55672.1	<i>CcDH2a</i> (dehydrin DH2a)	F 5'-CTCTGGCTCGTCGGAGGAT -3' R 5'-CTCCCGCCGTTAGTTTCTC 3'	90
24547 ABC68275.1	<i>CcDH3</i> (dehydrinDH3)	F 5'-GCCATCTGCGGAGTTGA -3' R 5'-TTGAGCTGCTGGAGCTGCTA -3'	121
31111 AFR11354.1	<i>CcNRT1.4</i> (nitrate transporter)	F 5'-CACGGGCCCTCTTCTTGACAA -3' R 5'-TTGTGCCCCCTTCACTATG -3'	80
24672 BAK52670.1	<i>CcPhLOG</i> (cytokinin riboside 5'-monophosphate fosforibohidrolase)	F 5'-TGCCCAAACTGCTCATGAACT -3' R 5'-GGAAGCAACCCCGTTATGCT -3'	70
40458 Q9M5X7.1	<i>CcMALD3</i> (NLTP-MALDO RecName: Full=Non-specific lipid transfers)	F 5'-ATAACATGGGGCCAAAGTATCG -3' R 5'-CCTAATCCCACCGCAACAATA -3'	96
07152 XP_002262699.1	<i>CcBSPA</i> (bark storage protein)	F 5'-TGGAACTGGACTGGTGAATGC -3' R 5'-CCCCAGTGCCAAAAGTGAT -3'	90
<b>Estresse oxidativo</b>			
15916 AAAY84147.1	<i>CcGST</i> (Glutathione S-transferase)	F 5'-CTTTACCTGGCGGATCTCAA -3' R 5'-ACACTTTCCTGGCGGTCGTA -3'	71
16870 ACU82387.2	<i>CcPRX1</i> (peroxidase 1)	F 5'-ACAGTACCCCTCGGCAGAAAGA -3' R 5'-GGAGGAGGGATTGGCTGTT -3'	71

“Tabela 1, continuação”

Contig GB	Nome	Sequência do primer	Tamanho do Amplicon (pb)
<b>Ajuste osmótico</b>			
11142 XP-002269232.1	<i>CcNADP-S6PDH</i> (PREDICTED: NADP-dependent D-sorbitol-6-fostafase desidrogenase)	F 5'-CGTTCTCGAGGCTTGCAAAG-3' R 5'-ATGCCTTGTGGTACTGGAAAAT-3'	91
42862 ADM92589.1	<i>CcGoIS3</i> (Galactinol synthase)	F 5'-CCCTTTGGTGGTGCAGTTT-3' R 5'-AGGCTCGATCTCCCGGACTATA-3'	91
<b>Estresse biótico</b>			
13168 ADZ48381	<i>CcCLXIP</i> (chitinase-like xylanase inhibitor protein)	F 5'-CGTTGCTCCTCGCTTAGTT-3' R 5'-GCATCATCGCGTGAAGAAAG-3'	105
39836 ABY59789.1	<i>CcMJE</i> (methyl jasmonate esterase)	F 5'-AGGCCCATAGCACTGTTTGTCT-3' R 5'-ACGACGACAGCTCCGTAATT-3'	81
17904 AEQ54768	<i>CcCPI3</i> (cysteine proteinase inhibitor CPI-3)	F 5'-AACCACGCTGCAAGAATG-3' R 5'-CAGGGTGGTGAGTAGGAGGAGAT-3'	58
<b>No Hits</b>			
05517	<i>CcNohits</i> (NO HITS)	F 5'-AAAAATTCACCACGGCAAGCT-3' R 5'-TTGCTCCCTCACATTTCCA-3'	72
15847	<i>CcNohits</i> (NO HITS)	F 5'-GGAGGCCACAAGGATGATCAT-3' R 5'-TCACCACTATGGCCTTCCTTCTT-3'	89

“Tabela 1, conclusão”

Contig GB	Nome	Sequência do primer	Tamanho do Amplicon (pb)
<i>No Hits</i>			
30679	CcNohits (NO HITS)	F 5'- ATGGCGGAAGTTAGCATTGC -3'	92
		R 5'- GAGGGAGACTCCTGTGCTGATG -3'	
42816	CcNohits (NO HITS)	F 5'- GCCTGAAAGAGGTCCAACCA -3'	81
		R 5'- ATTCTCCGAACCTTGAAAACAGAA -3'	
NUBI	CcUBQ (Ubiquitin)	F 5'- AAGACAGCTTCAACAGAGTACAGCAT -3'	120
		R 5'- GGCAGGACCTTGGCTGACTATA -3'	

## 2.4 Ensaios de qPCR

**Síntese de cDNA:** A síntese da primeira fita de cDNA foi realizada pela manipulação de 1 µg de RNA total com o kit ImProm-II-Sistema de Transcrição Reversa com oligo dT<sub>15</sub> de acordo com as recomendações do fabricante (Promega). A reação de PCR para sintetizar cDNA foi realizada no termociclador PTC-100 (MJ Research), usando o kit GoTaq DNA polimerase, seguindo as instruções do fabricante (Promega) sob as seguintes condições: desnaturação inicial 94°C por dois minutos, seguida de 40 ciclos de 94°C por 30 segundos, 55°C por 30 segundos e 73°C por três minutos, e a fase de extensão final de 72°C por seis minutos.

Foram selecionados genes potencialmente envolvidos na resposta da planta ao estresse hídrico para a análise por qPCR, com a finalidade de integrar os resultados de expressão com os resultados do perfil proteico. Além disso, o material vegetal utilizado nos ensaios de qPCR foi o mesmo usado para o sequenciamento. Dentre os genes candidatos, foram selecionados ainda aqueles genes com características fisiológicas de interesse, em relação ao metabolismo do cafeeiro quando submetido ao estresse hídrico (Tabela 1). As sequências dos *primers* (*Foward e Reverse*) foram obtidas utilizando o Software *Primer Express*. Todos os *primers* foram desenhados em regiões com 40 a 60% dos nucleotídeos citosina (C) e guanina (G) em sua sequência e temperatura de desnaturação (T<sub>m</sub>) entre 57°C e 63°C e os *amplicons* gerados apresentaram tamanho entre 60 e 126 pares de base (Tabela 1).

**Testes de qPCR:** Após a síntese dos cDNAs, utilizando o protocolo recomendado para o Fast 7.500 Real-Time PCR Systems (Applied Biosystems), os cDNAs preparados foram diluídos nas condições 1/20, 1/25 e 1/40 e testados por qPCR, utilizando pares de *primers* GCs projetados pelo *software* Primer

Express (Applied Biosystems). Preliminarmente os *primers* foram testados por sua eficiência e especificidade com o mix de cDNA das raízes de todos os clones nas duas condições I e NI.

Para a qPCR, 1µl de cDNA diluído e 0.2µM (concentração final) de cada *primer* são utilizados, em um volume final de 10µL com o fluorocromo SYBR® green (SYBRGreen qPCR MIX-UDG/ROX, Introvigen). O mix da reação é incubado por dois minutos a 50°C (tratamento da uracil DNA-Glycosylase), em seguida, por cinco minutos a 95°C para inativação da uracil DNA glycosylase, seguidos de 40 ciclos de amplificação de três segundos a 95°C, 30 segundos a 60°C. Os dados são analisados utilizando o *software* SDS 2.1 (Applied Biosystems) para determinar o valor do CT (*Cycle Threshold*). A etapa de pré-incubação, para a ativação da enzima UDGase, tem como função evitar a contaminação do mix, uma vez que em vez de dTTP tem dUTP.

A especificidade dos produtos da PCR de cada *primer* é verificada analisando o tempo médio de dissociação dos produtos amplificados. A eficiência da PCR (E) é calculada utilizando os dados de fluorescência absoluta capturados durante a fase exponencial da amplificação de cada reação pela equação  $(1+E)^{1/4}(-1/\text{slope})$  (RAMAKERS et al., 2003). Os valores da eficiência são obtidos pela contagem de todos os cálculos subsequentes.

O coeficiente de expressão gênica é normalizado pelos níveis de expressão da *CcUBQ10* como uma referência constitutiva (BARSALOBRES-CAVALLARI et al., 2009). Na quantificação relativa, há uma comparação entre as concentrações relativas, cujo gene que se utiliza para comparar, o gene de referência, tem que ter a expressão sempre constante, mesmo quando o indivíduo a ser analisado sofre algum tipo de estresse. A expressão relativa é quantificada pela aplicação da fórmula  $(1+E)^{-\Delta\Delta Ct}$ , em que  $\Delta Ct = Ct$  do gene alvo -  $Ct$  do gene endógeno e  $\Delta\Delta Ct = \Delta Ct$  da amostra -  $\Delta Ct$  do calibrador. O calibrador é uma das amostras analisadas, em geral seleciona-se como calibrador

uma amostra “controle”, ou seja, que não sofreu um tipo de estresse. Em todas as reações de qPCR também serão utilizadas amostras consideradas controles negativos (NTC), uma vez que colocamos água em vez de DNA junto com o mix da reação para certificação de que o mix está livre de DNA contaminante.

### 3 RESULTADOS E DISCUSSÃO

#### 3.1 Análises *in silico*

Bibliotecas do transcriptoma de raízes de *C. canephora* foram geradas a partir dos dados de sequenciamento de cDNA, utilizando a técnica de pirosequenciamento 454. Foram gerados 4.397.628 *reads* a partir de oito bibliotecas de raízes de *C. canephora* provenientes dos clones 14, 22, 73 e 120 nas condições irrigado (I) e não irrigado (NI), conforme descrito na Tabela 2.

Esta iniciativa é a primeira a produzir dados significativos de raízes de cafeeiro, tendo em vista que na base de dados do Projeto Genoma Café (VIERA et al., 2006) existiam somente 2.071 *reads* de raízes.

Tabela 2 Bibliotecas geradas e número total de *reads* produzidos por meio do sequenciamento 454 a partir de raízes de *C. canephora* em condições Irrigado (I) e Não Irrigado (NI)

Nome da biblioteca	Descrição	Número de <i>reads</i>
Cc-14I	Raiz de <i>C. canephora</i> irrigado	559.079
Cc-14NI	Raiz de <i>C. canephora</i> não-irrigado	603.473
Cc-22I	Raiz de <i>C. canephora</i> irrigado	560.932
Cc-22NI	Raiz de <i>C. canephora</i> não-irrigado	630.779
Cc-73I	Raiz de <i>C. canephora</i> irrigado	499.501
Cc-73NI	Raiz de <i>C. canephora</i> não-irrigado	539.102
Cc-120I	Raiz de <i>C. canephora</i> irrigado	450.456
Cc-120NI	Raiz de <i>C. canephora</i> não-irrigado	554.306

Os dados de sequências produzidos foram então utilizados para a realização de análises de expressão *in silico* do perfil do transcriptoma das raízes nas condições I *versus* NI. Para tanto, as 25.574 sequências de cDNA resultantes do Projeto Genoma Estrutural de *Coffea canephora* foram utilizadas como referência nas análises, utilizando-se o *software* Q-Seq do pacote DNASTar

(Lasergene). Os dados foram normalizados por RPKM e os resultados quantitativos de expressão obtidos.

Para cada gene foi obtido um dado de expressão por clone e respectiva condição (I/NI). Além disto, um fator de indução (FI) para cada gene também foi calculado, a partir do quociente da expressão gênica de cada clone na condição de déficit hídrico, relativo à condição controle (NI/I). Além disso, com a finalidade de entender quão induzida é a expressão dos genes nos clones tolerantes em relação ao clone sensível, o valor Q de expressão também foi calculado e indica o quociente entre o valor de FI dos clones tolerantes (14, 73 e 120) em relação ao FI do clone sensível (22).

Visando melhor representação dos perfis diferenciais de expressão obtidos, principalmente no que tange aos clones tolerantes, os dados de cada clone na condição controle (I) e os fatores de indução calculados foram plotados por cromossomo, conforme figuras abaixo. Além dos 11 cromossomos que existem em *C. canephora*, um cromossomo extra foi gerado, denominado cromossomo 0, formado a partir de *reads* que não foram mapeados em nenhum dos 11 cromossomos.

Conforme os dados de expressão basal (condição controle-I) dos genes alocados no Cromossomo 0, apresentados na Figura 2A, pode-se observar que existem diferenças de expressão para alguns genes entre os clones de *C. canephora* analisados. Por exemplo, o gene GSCOCT00007152001 (Putative Bark storage protein A - BSPA) apresenta uma expressão basal bem mais elevada no clone 14, quando comparado aos clones tolerantes (Figura 2A). Outro gene que apresenta expressão basal diferenciada é o gene GSCOCT00000746001 (Unknown Gene), o qual apresenta uma expressão basal mais alta no clone 73, comparativamente aos demais clones analisados (Figura 2A).

Para os dados de expressão diferencial entre os clones analisados, também há diferenças quanto aos valores de FI para alguns genes, conforme apresentado na Figura 2B. O clone 120 destaca-se com mais elevado valor de FI para alguns genes, como o gene GSCOCT00011142001 (NADP-dependent D-sorbitol-6-phosphate dehydrogenase - S6PDH). Por outro lado, em alguns genes o clone que se destaca quanto ao FI é o clone 14, como pode-se observar para o caso do gene GSCOCT00012078001 (Acidic endochitinase - CHIB1) (Figura 2B).

Analisando a indução da expressão dos genes entre os clones tolerantes em relação ao clone sensível para a condição NI (Fator Q), conforme apresentado na Figura 2C, observa-se que o gene que apresentou expressão mais induzida é o GSCOCT00012078001, cuja indução foi significativamente maior no clone 14.

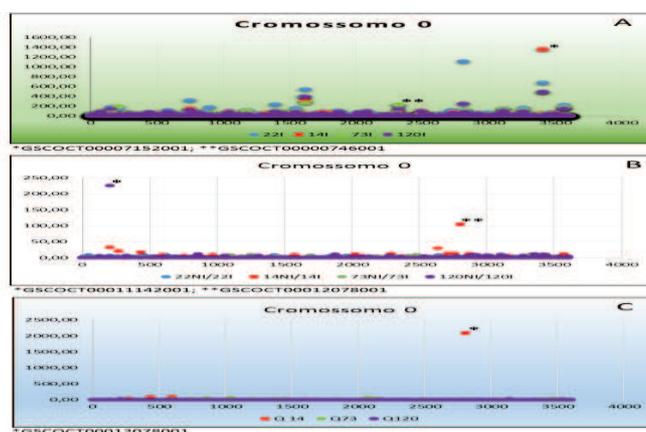


Figura 2 Gráfico de dispersão dos genes no Cromossomo 0, com base nos resultados de pirosequenciamento 454

Legenda: **A.** Valor de expressão dos clones na condição controle (I), dados normalizados por RPKM. **B.** Valor de FI (I/NI) para todos os clones. **C.** Valores de Q (FI clones tolerantes/FI clone 22).

Os genes alocados no Cromossomo 1 também apresentam valores de expressão diferencial para alguns genes entre os clones estudados. Em relação aos dados de expressão basal (condição controle-I) apresentados na Figura 3A, pode-se observar que alguns genes se destacam quanto aos valores de expressão, o gene GSCOCT00028209001 (Protein Heat Shock) apresenta expressão basal mais elevada no clone 120 e o gene GSCOCT00028440001 (Asparagine synthetase 1) apresenta expressão basal mais elevada no clone 73 (Figura 3A). Em relação aos dados de expressão diferencial (FI), apresentados na Figura 3B, também há diferenças quanto aos valores para alguns genes, o clone 120 destaca-se com maior valor de FI em alguns genes, como o gene GSCOCT00016031001 (Eukaryotic aspartyl protease family protein BG) e o gene GSCOCT00016033001 (Predicted Basic 75 Globulin) (Figura 3B). Por outro lado, o clone 14 também se destaca em alguns genes, como pode-se observar no caso do gene GSCOCT00015920001 (Putative uncharacterized protein NFYB5) (Figura 3B).

Analisando-se o diferencial de expressão dos genes entre os clones tolerantes em relação ao clone sensível para a condição NI (Fator Q), conforme apresentado na Figura 3C, observa-se que o clone 73 destaca-se quanto à expressão de alguns genes, como é o caso do gene GSCOCT00024137001 (xyloglucan endotransglucosylase/hydrolase protein 32 XTH32). Por outro lado, o gene GSCOCT00016554001 (Predicted: uncharacterized protein) evidencia maior expressão no clone 120 (Figura 3C).

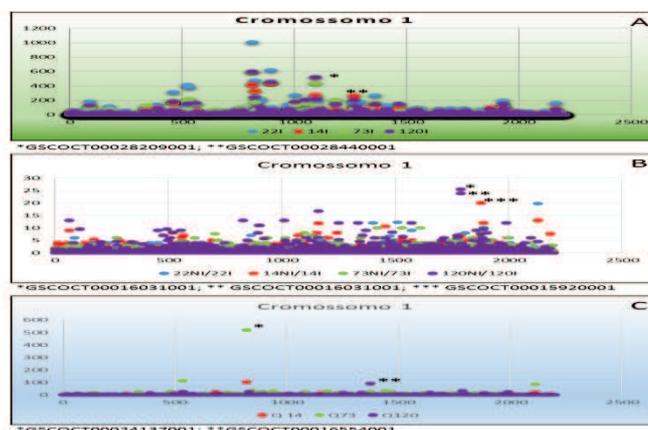


Figura 3 Gráfico de dispersão dos genes no Cromossomo 1, com base nos resultados de pirosequenciamento 454

Legenda: **A.** Valor de expressão dos clones na condição controle (I), dados normalizados por RPKM. **B.** Valor de FI (I/NI) para todos os clones. **C.** Valores de Q (FI clones tolerantes/FI clone 22).

Para o Cromossomo 2, observando os dados de expressão basal representados pela Figura 4A, nota-se que um dos genes que se destaca quanto à expressão na condição controle (I) é o GSCOCT00014291001 (Unknown gene) cujo valor é mais elevado no clone 14 em relação aos outros clones para o mesmo gene (Figura 4A). Outro gene de destaque é o GSCOCT00029121001 (Auxin-repressed 12.5 kDa protein), cujo valor de expressão basal é elevado em todos os clones tolerantes em relação aos outros genes (Figura 4A).

Com relação aos dados de expressão diferencial, tendo como base os valores de FI, conforme apresentado na Figura 4B, para grande parte dos genes, os valores mais elevados de expressão diferencial é observado no clone 120. Como é o caso dos genes GSCOCT00029971001 (Putative unknown protein) e GSCOCT00042863001 (Galactinol synthase), este último também apresenta elevado valor de FI no clone 14.

Analisando os dados apresentados na Figura 4C que representam a indução de expressão de um gene nos clones tolerantes em relação ao clone

sensível para a condição NI (Fator Q), observa-se que o clone 73 destaca-se com maior valor de Q para alguns genes, como é o caso do gene GSCOCT00029096001 (ROX1 homologue). Para o Clone 14, também existem alguns genes com alto valor de Q, como o gene GSCOCT00039176001 (Predicted Leucoanthocyanidin dioxygenase – ANS).

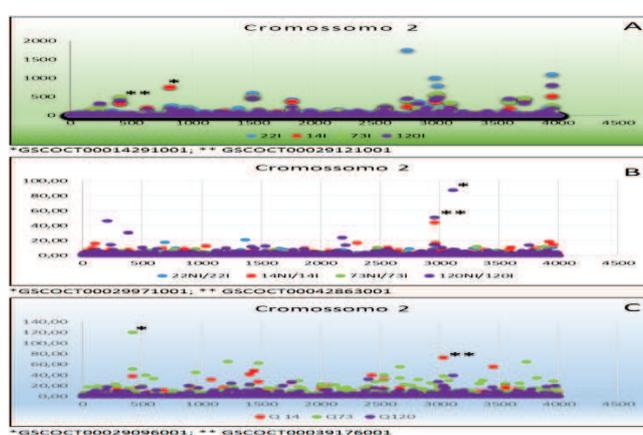


Figura 4 Gráfico de dispersão dos genes no Cromossomo 2, com base nos resultados de pirosequenciamento 454

Legenda: **A.** Valor de expressão dos clones na condição controle (I), dados normalizados por RPKM. **B.** Valor de FI (I/NI) para todos os clones. **C.** Valores de Q (FI clones tolerantes/FI clone 22).

Para o Cromossomo 3, a expressão basal dos genes é apresentada na Figura 5A. O gene GSCOCT00042727001 (CYC02 protein) foi o que apresentou maior valor de expressão basal na condição controle (I), o gene GSCOCT00031814001 (Plant basic secretory protein - BSP) também apresentou alto valor de expressão basal no clone 73, mas nos outros clones tolerantes à expressão também foi elevada (Figura 5A). Em relação aos valores de expressão diferencial, alguns genes apresentam alto valor de FI no clone 120, como é observado no gene GSCOCT00025904001 (1-deoxy-D-xylulose-5-phosphate synthase, chloroplastic - TKT2) que apresenta o mais elevado valor de FI para os

genes alocados neste cromossomo (Figura 5B). Outro gene que também se destaca quanto à expressão é o GSCOCT00026853001 (NADP-dependent alkenal double bond reductase P1) no qual a expressão diferencial é alta também no clone 14 (Figura 5B).

Para os valores de indução observados nos clones tolerantes em relação ao clone sensível, como apresentado na Figura 5C, pode-se observar que para uma parte dos genes alocados neste cromossomo os mais elevados valores de indução são apresentados pelo clone 73. O gene GSCOCT00042703001 (xyloglucan endotransglucosylase/hydrolase protein 16 - XTH16) é o que apresenta mais elevado valor de indução (Figura 5C). O gene GSCOCT00033258001 (HXXXD-type acyl-transferase family protein - HHT1) apresenta maior indução no clone 14 em relação aos outros clones (Figura 5C).

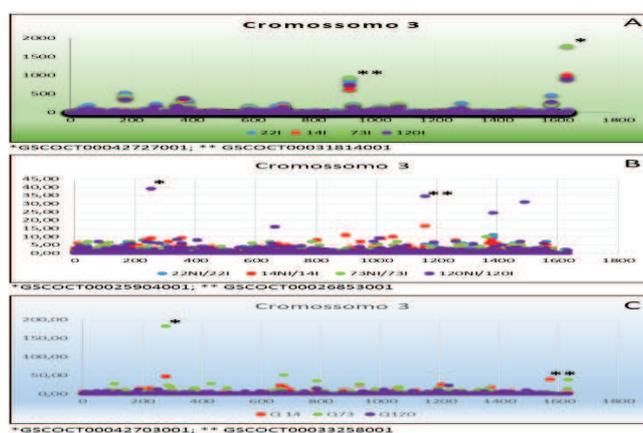


Figura 5 Gráfico de dispersão dos genes no Cromossomo 3, com base nos resultados de pirosequenciamento 454

Legenda: **A.** Valor de expressão dos clones na condição controle (I), dados normalizados por RPKM. **B.** Valor de FI (I/NI) para todos os clones. **C.** Valores de Q (FI clones tolerantes/FI clone 22).

Conforme os dados de expressão basal (condição controle-I) dos genes alocados no Cromossomo 4, apresentados na Figura 6A, pode-se observar que

há diferenças de expressão para alguns genes entre os clones estudados. Por exemplo, o gene GSCOCT00022022001 (Putative Acid phosphatase 1 - APS1) apresenta uma expressão basal bem mais elevada no clone 14, apesar de nos outros clones tolerantes também apresentar alta expressão (Figura 6A) e o gene GSCOCT00018903001 (Chaperone protein ClpB1 - CLPB1) que apresenta expressão basal mais elevada no clone 120.

Para os dados de expressão diferencial entre os clones analisados, também há diferenças quanto aos valores de FI para alguns genes, conforme apresentado na Figura 6B. O clone 14 destaca-se com mais elevado valor de FI em alguns genes, como o gene GSCOCT00021907001 (Unknow gene). Por outro lado, em alguns genes o clone que se destaca quanto ao FI é o clone 120, como se pode observar para o gene GSCOCT00022121001 (Late embryogenesis abundant protein Dc3) (Figura 6B).

Analisando-se a indução da expressão dos genes entre os clones tolerantes em relação ao clone sensível para a condição NI (Fator Q), conforme apresentado na Figura 6C, observa-se que o clone 73 destaca-se quanto aos mais elevados valores de Q, o gene com maior indução na expressão na condição NI é o GSCOCT00022066001 (Gibberellin 3-beta-dioxygenase 4 - 20ox1). Por outro lado, o gene GSCOCT00037380001 (Predicted protein) apresenta maior indução no clone 120.

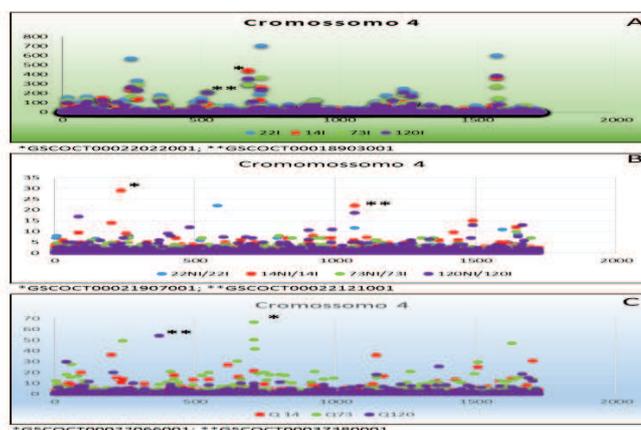


Figura 6 Gráfico de dispersão dos genes no Cromossomo 4, com base nos resultados de pirosequenciamento 454

Legenda: **A.** Valor de expressão dos clones na condição controle (I), dados normalizados por RPKM. **B.** Valor de FI (I/NI) para todos os clones. **C.** Valores de Q (FI clones tolerantes/FI clone 22).

Para o Cromossomo 5, observando-se a expressão basal apresentada na Figura 7A, o gene GSCOCT00013166001 (Acidic endochitinase) apresentou maior valor de expressão basal na condição controle (I) no clone 14 em relação aos outros clones. O gene GSCOCT00038706001 (Heat shock cognate protein 80 - HSC80) também apresentou alta expressão basal só que no clone 120 (Figura 7A). Em relação aos valores de expressão diferencial, alguns genes apresentam alto valor de FI no clone 14, como é observado para o gene GSCOCT00009865001 (Beta-amylase - BMY1), que apresenta o mais elevado valor de FI para os genes alocados neste cromossomo e o gene GSCOCT00021185001 (NAC domain-containing protein 72 - NAC072) (Figura 7B).

Para os valores de indução observados nos clones tolerantes em relação ao clone sensível, como apresentado na Figura 7C, pode-se observar que para uma parte dos genes alocados neste cromossomo os mais elevados valores de indução estão no clone 73. O gene GSCOCT00035120001 (Xyloglucan

endotransglucosylase/hydrolase protein 9 - XTH9) é o que apresenta o mais elevado valor de indução. Por outro lado, o gene GSCOCT00017064001 (Acidic endochitinase - CHIT3) apresenta maior indução no clone 14 em relação aos outros clones.

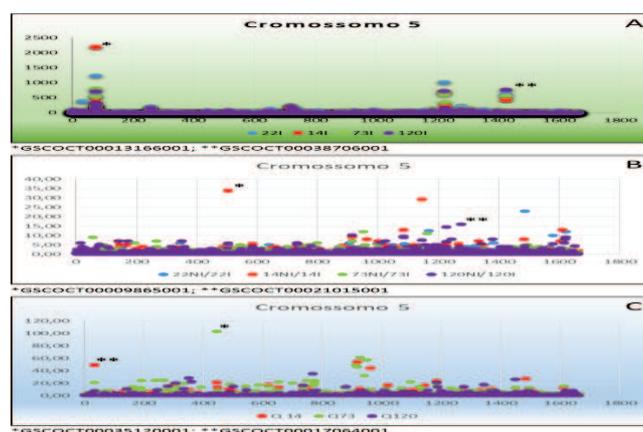


Figura 7 Gráfico de dispersão dos genes no Cromossomo 5, com base nos resultados de pirosequenciamento 454

Legenda: **A.** Valor de expressão dos clones na condição controle (I), dados normalizados por RPKM. **B.** Valor de FI (I/NI) para todos os clones. **C.** Valores de Q (FI clones tolerantes/FI clone 22).

Conforme os dados de expressão basal (condição controle-I) dos genes alocados no Cromossomo 6, apresentados na Figura 8A, pode-se observar que há diferenças de expressão para alguns genes entre os clones estudados, sendo que os valores mais elevados de expressão basal são apresentados no clone 22. Por exemplo, o gene GSCOCT00028895001 (unknown gene) apresenta uma expressão basal bem mais elevada no clone 22 em relação aos clones tolerantes (Figura 8A). Outro gene que apresenta valor de expressão basal mais elevado em relação aos genes alocados no Cromossomo 6 é o gene GSCOCT00041507001 (Sucrose synthase), que também apresenta altos valores de expressão basal nos clones tolerantes (Figura 8A). Observando-se somente o gene que se destaca

quanto ao valor de expressão basal dentre os clones tolerantes, nota-se que o gene GSCOCT00004140001 (Probable galactinol--sucrose galactosyltransferase 2 - RFS2) apresenta mais elevado valor de expressão basal no clone 120 (Figura 8A).

Para os dados de expressão diferencial entre os clones analisados, também há diferenças quanto aos valores de FI para alguns genes, conforme apresentado na Figura 8B. O clone 14 destaca-se com mais elevado valor de FI em alguns genes, como o gene GSCOCT00031111001 (Nitrate transporter 1.4 - NRT1.4) e o gene GSCOCT00041611001 (Threonine synthase) (Figura 8B).

Analisando a indução da expressão dos genes entre os clones tolerantes em relação ao clone sensível para a condição NI (Fator Q), conforme apresentado na Figura 8C, observa-se que o clone 73 destaca-se quanto aos mais elevados valores de Q e o gene com maior indução na expressão na condição NI é o GSCOCT00031383001 (U3 small nucleolar ribonucleoprotein protein IMP4 - Imp4) com mais elevado valor de Q no clone 120 (Figura 7C). Por outro lado, o gene GSCOCT00015869001 (Inactive purple acid phosphatase 27 - PAP27) apresenta maior indução no clone 73 (Figura 8C).

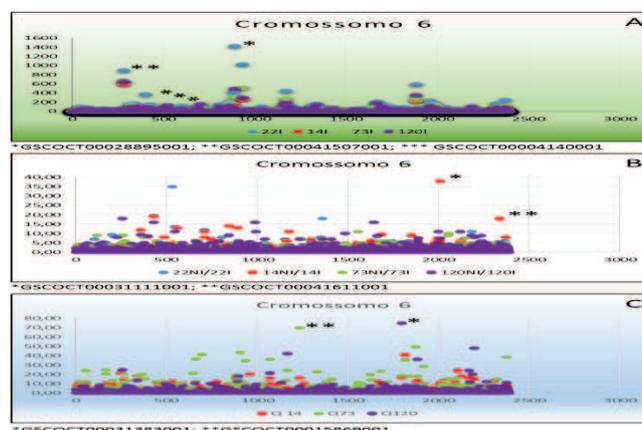


Figura 8 Gráfico de dispersão dos genes no Cromossomo 6, com base nos resultados de pirosequenciamento 454

Legenda: **A.** Valor de expressão dos clones na condição controle (I), dados normalizados por RPKM. **B.** Valor de FI (I/NI) para todos os clones. **C.** Valores de Q (FI clones tolerantes/FI clone 22).

Conforme os dados de expressão basal (condição controle-I) dos genes alocados no Cromossomo 7, apresentados na Figura 9A, pode-se observar que há diferenças de expressão para alguns genes entre os clones estudados. Por exemplo, o gene GSCOCT00039723001 (Sodium-coupled neutral amino acid transporter 3 - SLC38A3) e o gene GSCOCT00036596001 (Homeobox-leucine zipper protein ATHB-16 - ATHB-6) apresentam maior expressão basal no clone 14 (Figura 9A).

Para os dados de expressão diferencial entre os clones analisados, também há diferenças quanto aos valores de FI para alguns genes, conforme apresentado na Figura 9B. O clone 120 destaca-se com mais elevado valor de FI em alguns genes, como o gene GSCOCT00037166001 (Dehydrin DH1a) e o gene GSCOCT00039509001 (ATPase 4, plasma membrane-type - AHA4) (Figura 9B).

Analisando a indução da expressão dos genes entre os clones tolerantes em relação ao clone sensível para a condição NI (Fator Q), conforme

apresentado na Figura 9C, observa-se que o clone 73 destaca-se quanto aos mais elevados valores de Q, o gene com maior indução na expressão na condição NI é o GSCOCT00040013001 (Peroxidase 73 - PER73) (Figura 9C). Por outro lado, o gene GSCOCT00037166001 (Dehydrin DH1a), apresenta maior indução no clone 120 (Figura 9C).

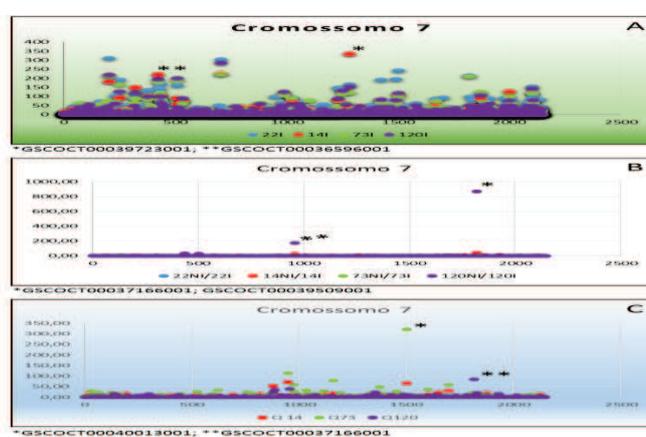


Figura 9 Gráfico de dispersão dos genes no Cromossomo 7, com base nos resultados de pirosequenciamento 454

Legenda: **A.** Valor de expressão dos clones na condição controle (I), dados normalizados por RPKM. **B.** Valor de FI (I/NI) para todos os clones. **C.** Valores de Q (FI clones tolerantes/FI clone 22).

Para o Cromossomo 8, observando-se a expressão basal, apresentada na Figura 10A. O gene GSCOCT00035443001 (Heat shock protein 83- HSP83A) e o gene GSCOCT00030391001 (Heat shock protein STI - STI), apresentam alto valor de expressão basal na condição controle (I) no clone 120 (Figura 10A).

Em relação aos valores de expressão diferencial, alguns genes apresentam alto valor de FI, como é observado para o gene GSCOCT00035651001 (Homeobox Leucine-zipper - ATHB-7) que apresenta o mais elevado valor de FI para os genes alocados neste cromossomo, sendo o maior valor apresentado no clone 73. Outro gene que também destaca-se quanto

à expressão é o GSCOCT00027142001 (Expansin-like B1 - EXLB1) (Figura 10B).

Para os valores de indução observados nos clones tolerantes em relação ao clone sensível, como apresentado na Figura 10C, pode-se observar que para uma parte dos genes alocados neste cromossomo os mais elevados valores de indução estão no clone 73. Como mostrado para o gene GSCOCT00026947001 (Amino acid permease 8 - AAP8) que apresenta mais elevado valor de indução e para o gene GSCOCT00030430001 (14 kDa proline-rich protein DC2.15 - At4g00165).

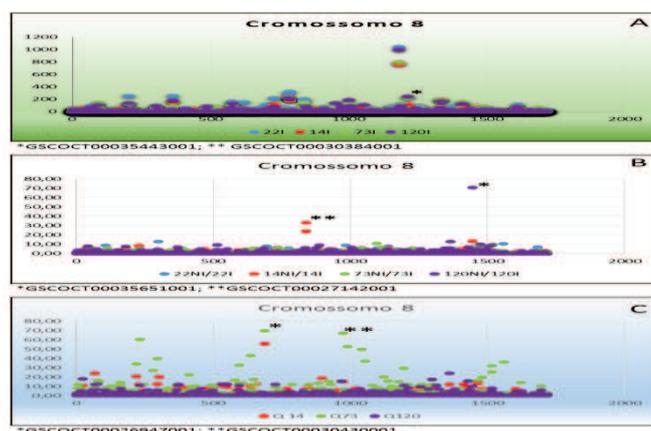


Figura 10 Gráfico de dispersão dos genes no Cromossomo 8, com base nos resultados de pirosequenciamento 454

Legenda: **A.** Valor de expressão dos clones na condição controle (I), dados normalizados por RPKM. **B.** Valor de FI (I/NI) para todos os clones. **C.** Valores de Q (FI clones tolerantes/FI clone 22).

Conforme os dados de expressão basal (condição controle-I) dos genes alocados no Cromossomo 9, apresentados na Figura 11A, pode-se observar que há diferenças de expressão para alguns genes entre os clones estudados. O gene GSCOCT00017678001 (Inorganic phosphate transporter 1-11 - PHT1-11) apresenta o mais elevado valor de expressão basal no clone 120 (Figura 11A).

Para os dados de expressão diferencial entre os clones analisados, também há diferenças quanto aos valores de FI para alguns genes, conforme apresentado na Figura 11B. O clone 120 destaca-se com mais elevado valor de FI para alguns genes, como é o caso do gene GSCOCT00041005001 (Thioredoxin-like 1-1, chloroplastic - At1g08570) e do gene GSCOCT00017783001 (Soluble starch synthase 3, chloroplastic/amyloplastic - SS3) (Figura 11B).

Analisando-se a indução da expressão dos genes entre os clones tolerantes em relação ao clone sensível para a condição NI (Fator Q), conforme apresentado na Figura 11C, observa-se que o clone 14 é o que apresenta maior valor de indução para o gene GSCOCT00017380001 (Apyrase - RROP1). Por outro lado, o gene GSCOCT00005099001 (protein At3g02645) apresenta maior indução no clone 73 (Figura 11C).

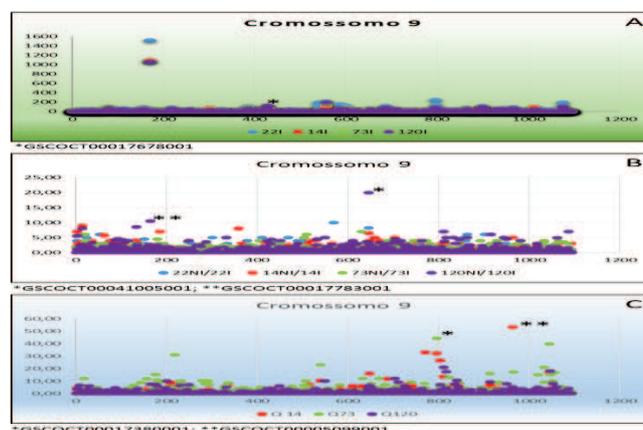


Figura 11 Gráficos de dispersão dos genes no Cromossomo 9, com base nos resultados de pirosequenciamento 454

Legenda: **A.** Valor de expressão dos clones na condição controle (I), dados normalizados por RPKM. **B.** Valor de FI (I/NI) para todos os clones. **C.** Valores de Q (FI clones tolerantes/FI clone 22).

Para o Cromossomo 10, observando-se a expressão basal, apresentada na Figura 12A, nota-se que o gene GSCOCT00024515001 (Catalase isozyme 3 - CAT3) apresenta o mais elevado valor de expressão basal na condição controle (I), valor este apresentado no clone 120 (Figura 12A). O gene GSCOCT00024721001 (Cysteine proteinase 15A) também apresenta alto valor de expressão basal, mas no clone 73 (Figura 12A). Em relação aos valores de expressão diferencial, alguns genes apresentam alto valor de FI, sendo o clone 120 o clone com maiores valores de expressão diferencial, como é observado para os genes GSCOCT00024587001 (nucleoredoxin 2) e GSCOCT00034647001 (Purine permease 3 - PUP3) (Figura 12A). Para os valores de indução observados nos clones tolerantes em relação ao clone sensível, como apresentado na Figura 12C, pode-se observar que para os genes alocados neste cromossomo os mais elevados valores de indução são apresentados para o clone 73. O gene GSCOCT00034917001 (Subtilisin-like

protease - ARA12) é o que apresenta mais elevado valor de indução (Figura 12C).

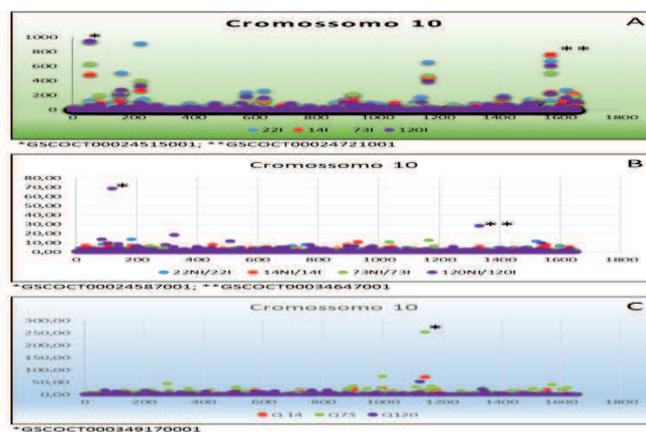


Figura 12 Gráficos de dispersão dos genes no Cromossomo 10, com base nos resultados de pirosequenciamento 454

Legenda: **A.** Valor de expressão dos clones na condição controle (I), dados normalizados por RPKM. **B.** Valor de FI (I/NI) para todos os clones. **C.** Valores de Q (FI clones tolerantes/FI clone 22).

Conforme os dados de expressão basal (condição controle-I) dos genes alocados no Cromossomo 11, apresentados na Figura 13A, pode-se observar que há diferenças de expressão para alguns genes entre os clones estudados. O maior valor de expressão basal é apresentado para o clone 22, entretanto dentre os clones tolerantes, o clone 73 destaca-se com o maior valor basal de expressão para o gene GSCOCT00019481001 (Acidic endochitinase - SE2) (Figura 13A).

Para os dados de expressão diferencial entre os clones analisados, também há diferenças quanto aos valores de FI para alguns genes, conforme apresentado na Figura 13B. O clone 120 destaca-se com mais elevado valor de FI para o gene GSCOCT00020463001 (Unknown gene). Por outro lado, o clone 14 também destaca-se com a expressão do gene GSCOCT00032601001 (nitrite transporter At1g68570) (Figura 13B).

Analisando-se a indução da expressão dos genes entre os clones tolerantes em relação ao clone sensível para a condição NI (Fator Q), conforme apresentado na Figura 13C, observa-se que o clone 73 é o que apresenta maior valor de indução para o gene GSCOCT00038067001 (Lipase). Por outro lado, o gene GSCOCT00019567001 (Pleiotropic drug resistance protein 1 - PDR1) apresenta maior indução no clone 14 (Figura 13C).

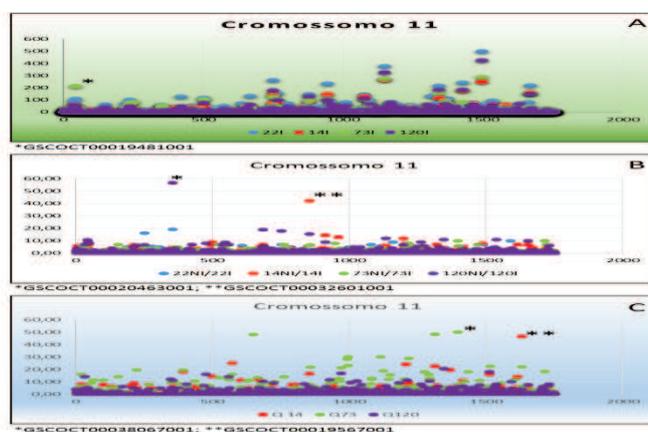


Figura 13 Gráficos de dispersão dos genes no Cromossomo 11, com base nos resultados de pirosequenciamento 454

Legenda: **A.** Valor de expressão dos clones na condição controle (I), dados normalizados por RPKM. **B.** Valor de FI (I/NI) para todos os clones. **C.** Valores de Q (FI clones tolerantes/FI clone 22).

Por meio da análise *in silico* realizados neste trabalho, foram identificados e selecionados genes com expressão diferencial em raízes dos clones estudados (Tolerantes vs sensível). Genes estes que foram selecionados a partir dos valores diferenciais em cada cromossomo, os valores foram: expressão basal (Controle – I), Fator de Indução (FI) e valor de indução em relação ao clone sensível (Q), conforme apresentado na Tabela 3.

Tabela 3 Valores de expressão para os Genes Candidatos (GCs) de tolerância ao déficit hídrico em raízes de clones de *C. canephora*. Dados obtidos com base nos valores de expressão resultantes no pirosequenciamento-454. **Chr** (cromossomo). **Expressão basal** (valor de expressão dos genes para os clones na condição I). **Fator de indução FI** (I/NI) e **Q** (FI (clones tolerantes) / FI (clone 22)). Em negrito estão os valores que foram utilizados como critério para a seleção do determinado gene. Os dados foram avaliados utilizando somente os clones tolerantes como referência.

Chr	Gene	Descrição	Expressão Basal			FI			Q		
			14	73	120	14	73	120	14	73	120
0	GSCOC00007152001	Putative Bark storage protein A~BSPA~ modules	1.361,00	490,00	490,00	0,41	1,30	0,75	0,22	0,69	0,40
0	GSCOC00000746001	Unknown Gene	129,00	<b>230,00</b>	143,00	0,19	0,75	0,10	2,24	9,07	1,26
0	GSCOC00011142001	D-sorbitol-6-phosphate dehydrogenase - S6PDH	36,00	192,00	4,00	32,31	1,22	<b>226,25</b>	4,29	0,16	30,01
0	GSCOC00006614001	60S ribosomal protein L10-2~ RPL10B	1,00	10,00	0,00	<b>30,00</b>	1,30	0,00	6,49	0,28	0,00
0	GSCOC00012078001	Acidic endochitinase - CHIB1	1,00	0,00	245,00	105,00	0,00	0,13	<b>2.086,00</b>	0,00	0,68
1	GSCOC00028209001	Protein Heat Shock	273,00	440,00	<b>521,00</b>	0,41	1,11	0,81	1,05	2,82	2,06
1	GSCOC00028440001	Asparagine synthetase 1	<b>261,00</b>	190,00	195,00	0,46	1,15	0,69	0,57	1,43	0,85
1	GSCOC00016031001	Eukaryotic aspartyl protease family protein BG	5,00	6,00	4,00	34,00	11,00	<b>102,00</b>	8,74	2,36	32,79
1	GSCOC00016033001	Predicted Basic 75Globulin	16,00	51,00	11,00	5,25	1,20	<b>23,91</b>	0,73	0,17	3,31
1	GSCOC00024137001	Xyloglucan endotransglucosylase/hydrolase protein 32 XTH32	332,00	172,00	242,00	0,21	1,10	0,00	101,37	<b>520,85</b>	0,00

“Tabela 3, continua”

Chr	Gene	Descrição	Expressão Basal				FI			Q		
			14	73	120	19,00	14	73	120	14	73	120
2	GSCOC00014291001	Unknown gene	765,00	123,00	19,00	0,04	0,30	2,26	0,01	0,10	0,75	
2	GSCOC00029121001	Auxin-repressed protein-like	330,00	500,00	402,00	1,55	1,14	2,23	0,42	0,31	0,70	
2	GSCOC00029971001	Unknown protein	12,00	37,00	1,00	3,00	0,95	88,00	1,35	0,43	39,60	
2	GSCOC00042863001	Galactinol synthase	6,00	65,00	3,00	44,50	1,14	51,00	4,95	0,13	5,68	
2	GSCOC00029096001	ROX1 homologue	38,00	45,00	54,00	0,45	1,40	0,00	38,47	120,40	0,00	
3	GSCOC00042727001	CYC02 protein	1.008,00	1.791,00	900,00	0,56	1,20	1,78	0,32	0,69	1,03	
3	GSCOC00031814001	Plant basic secretory protein - BSP	631,00	936,00	748,00	0,70	1,27	0,78	1,27	2,31	1,42	
3	GSCOC00025904001	1-deoxy-D-xylulose-5-phosphate synthase, chloroplastic - TKT2	3,00	15,00	1,00	9,00	1,33	39,00	2,84	0,42	12,32	
3	GSCOC00026853001	NADP-dependent alkenal double bond reductase P1	4,00	24,00	2,00	16,50	0,96	34,50	2,98	0,17	6,23	
3	GSCOC00042703001	Xyloglucan endotransglucosylase/hydrolase protein 16 - XTH16	143,00	129,00	160,00	0,40	1,56	0,01	46,64	182,30	0,73	
4	GSCOC00022022001	Putative Acid phosphatase 1 - APS1	442,00	293,00	356,00	0,45	1,35	0,23	2,20	6,65	1,13	
4	GSCOC00018903001	Chaperone protein ClpB1 - CLPB1	81,00	226,00	213,00	1,15	1,20	0,82	0,92	0,96	0,57	
4	GSCOC00021907001	Unknown gene.	1,00	13,00	2,00	29,00	1,15	4,50	10,55	0,42	1,64	
4	GSCOC00022121001	Late embryogenesis abundant protein Dc3	1,00	13,00	4,00	22,00	2,38	18,75	1,87	0,20	1,60	

“Tabela 3, continuação”

Chr	Gene	Descrição	Expressão Basal			FI			Q		
			14	73	120	14	73	120	14	73	120
4	GSCOC00022066001	Gibberellin 3-beta-dioxygenase 4 - 20ox1.	125,00	39,00	53,00	0,33	1,03	0,15	21,32	66,67	9,81
5	GSCOC00013166001	Acidic endochitinase	<b>2.196,00</b>	813,00	716,00	2,02	8,11	4,16	0,45	0,24	0,92
5	GSCOC00038706001	Heat shock cognate protein 80 - HSC80	149,00	563,00	<b>770,0</b>	0,62	1,27	0,84	0,86	1,76	1,15
5	GSCOC00009865001	Beta-amylase - BMY1	4,00	2,00	5,00	<b>34,00</b>	0,00	5,80	6,69	0,0	1,14
5	GSCOC00021185001	NAC domain-containing protein 72 - NAC072	3,00	21,00	4,00	<b>29,33</b>	1,86	7,50	16,39	1,74	4,19
5	GSCOC00035120001	Xyloglucan endotransglucosylase/hydrolase protein 9 - XTH9	41,00	41,00	64,00	0,32	1,56	0,02	29,93	<b>103,02</b>	1,03
6	GSCOC00004140001	Probable galactinol--sucrose galactosyltransferase 2 - RFS2	75,00	130,00	<b>155,00</b>	2,13	1,28	1,21	1,43	0,86	0,81
6	GSCOC0003111001	Nitrate transporter 1.4 - NRT1.4	1,00	2,00	2,00	<b>38,00</b>	1,00	7,50	10,17	0,27	2,01
6	GSCOC00041611001	Threonine synthase 1, chloroplastic - TS1	1,00	3,00	0,00	<b>18,00</b>	1,33	0,00	5,54	0,41	0,0
6	GSCOC00031383001	U3 small nucleolar ribonucleoprotein protein IMP4	5,00	3,00	3,00	1,20	1,33	5,33	16,80	18,67	<b>74,67</b>

“Tabela 3, continuação”

Chr	Gene	Descrição	Expressão Basal			FI			Q		
			14	73	120	14	73	120	14	73	120
7	GSCOC00036596001	Putative Homeobox-leucine zipper protein ATHB-16-ATH-6	221,00	180,00	196,00	0,43	1,09	0,57	0,61	1,55	0,81
7	GSCOC00037436001	Actin-7 - ACT7	226,00	223,00	<b>285,00</b>	0,76	1,23	0,86	1,17	1,90	1,28
7	GSCOC00037166001	Dehydrin DH1a	28,00	210,00	2,00	43,07	1,05	<b>870,5</b>	4,22	0,10	85,23
7	GSCOC00039509001	ATPase 4, plasma membrane-type - AH/A4	7,00	47,00	1,00	23,14	1,49	<b>175,00</b>	5,16	0,33	39,03
7	GSCOC00040013001	Peroxidase 73 - PER73	116,0	96,00	119,00	0,27	1,32	0,01	65,68	<b>321,47</b>	2,04
8	GSCOC00035443001	Heat shock protein 83 - HSP83A	110,00	224,00	<b>241,00</b>	1,37	1,21	1,30	0,48	0,43	0,46
8	GSCOC00030391001	Heat shock protein ST1 - ST1	72,00	120,00	<b>189,00</b>	0,76	1,50	1,08	0,86	1,69	1,22
8	GSCOC00035651001	(Homeobox Leucine-zipper - ATHB-7)	8,00	33,00	2,00	13,13	1,24	<b>71,00</b>	1,84	0,17	9,95
8	GSCOC00027142001	Expansin-like B1 - EXLB1	10,00	72,00	0,00	<b>33,00</b>	1,21	0,00	8,05	0,29	0,0
8	GSCOC00026947001	Amino acid permease 8 - AAP8	34,00	49,00	35,00	0,74	0,92	0,00	55,88	<b>69,80</b>	0,0
9	GSCOC00017678001	Inorganic phosphate transporter 1-11 - PHT1-11	<b>89,00</b>	9,00	119,00	1,15	2,33	0,01	0,00	0,00	0,00
9	GSCOC00041005001	Thioredoxin-like 1-1, chloroplastic - Atlg08570	13,00	42,0	7,00	6,47	1,07	<b>20,00</b>	0,79	0,13	2,44
9	GSCOC00017783001	Starch synthase 3	8,00	15,00	2,00	1,25	1,00	<b>10,50</b>	0,89	0,71	7,50

“Tabela 3, conclusão”

Chr	Gene	Descrição	Expressão Basal			FI			Q		
			14	73	120	14	73	120	14	73	120
9	GSCOC00017380001	Apyrase - RROP1	39,00	54,00	78,00	1,05	1,44	0,05	32,20	44,24	1,57
9	GSCOC00005099001	Protein A13g02645	1,00	4,00	2,00	5,00	0,00	0,50	58,33	0,00	5,33
10	GSCOC00024515001	Catalase isozyme 3 - CAT3	489,00	638,00	950,00	0,96	1,18	0,67	1,75	2,14	1,22
10	GSCOC00024721001	Cysteine proteinase 15A - RD19A	273,00	402,00	336,00	1,94	1,03	1,98	2,88	1,53	2,94
10	GSCOC00024587001	Nucleoredoxin 2	10,00	19,00	1,00	2,80	1,53	69,00	0,46	0,25	11,39
10	GSCOC00034647001	Purine permease 3 - PUP3	9,00	12,00	2,00	5,89	2,25	28,50	4,76	1,83	23,04
10	GSCOC00034917001	Subtilisin-like protease - ARA12	62,00	86,00	104,00	0,40	1,43	0,01	71,37	215,13	1,70
11	GSCOC00019481001	Putative Acidic endochitinase - SE2	51,00	213,00	52,00	5,63	1,15	10,63	0,87	0,18	1,64
11	GSCOC00020463001	Unknown gene	20,0	52,00	7,00	4,75	1,44	56,57	0,25	0,08	2,95
11	GSCOC00032601001	Nitrite transporter At1g68570	1,00	12,00	2,00	42,00	1,33	15,50	16,52	0,52	6,10
11	GSCOC00038067001	Lipase.	5,00	1,00	3,00	0,60	10,0	1,00	3,00	50,00	5,00
11	GSCOC00019567001	Pleiotropic drug resistance protein 1 - PDR1	2,00	7,00	7,00	1,50	0,71	0,14	46,50	22,40	4,43

Genes candidatos para tolerância à seca, em folhas, já foram descritos pelo grupo em trabalhos anteriores (FREIRE, 2010; MARRACCINI et al., 2011, 2012; VIEIRA et al., 2013; VINECKY; SILVA; ANDRADE, 2012). Utilizando folhas Marraccini et al. (2012) e Vieira et al. (2013) já analisaram alguns genes candidatos para os mesmos clones e nas mesmas condições abordadas neste trabalho. Marraccini et al. (2012) utilizaram inclusive abordagens semelhantes a esta, com as quais compararam dados de qPCR com dados de espectrometria de massa MS/MS associada ao gel 2-D. Entretanto, todos os trabalhos anteriores foram focados em folhas e neste trabalho, abordamos estudos em raízes.

Para os valores de Q foram selecionados os 15 genes que apresentaram valores mais significativos, ou seja, os genes com maior indução na condição de déficit hídrico entre os clones tolerantes (Figura 14). É observado que para a maioria dos genes o clone 73 é o que apresenta maior valor, com exceção dos genes GSCCT00037166001 e GSCOCT00016554001 cujos valores de Q foram maiores no clone 120, já em GSCOCT00005624001, GSCOCT00005883001 e GSCOCT00012078001 os valores de Q foram maiores no clone 14. O maior valor apresentado de Q é observado em GSCOCT00012078001, 2,086,88 para o clone 14 e 2,68 para o clone 120. No clone 73, não houve expressão para este gene, conforme apresentado na Tabela 4.

Tabela 4 Genes com maiores valores de Q entre todos os clones tolerantes (14, 73 e 120). O valor Q foi obtido pelo quociente do valor de FI para cada clone em relação ao valor de FI para o clone sensível (22).

Contig	Descrição (proteína)	Q14	Q73	Q120
GSCOCT00012078001	Acidic endochitinase - CHIB1 Probable xyloglucan	2086,9	0,0	2,7
GSCOCT00024137001	endotransglucosylase/hydrolase Protein 32 XTH32	101,4	520,8	0,0
GSCOCT00040013001	Peroxidase 73 - PER73	65,7	321,5	2,0
GSCOCT00034917001	Putative Subtilisin-like protease - ARA12 Probable xyloglucan	71,4	253,2	1,7
GSCOCT00042703001	endotransglucosylase/hydrolase protein 16~ - XTH16	46,6	182,3	0,7
GSCOCT00029096001	ROX1 homologue Bifunctional	38,5	120,4	0,0
GSCOCT00033883001	monodehydroascorbate reductase and carbonic anhydrase nectarin-3 - NEC3	72,0	114,1	8,5
GSCOCT00009386001	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family unknown- gene Xyloglucan	26,0	113,0	23,6
GSCOCT00035120001	endotransglucosylase/hydrolase protein 9 - XTH9 - complete	20,9	103,0	1,0
GSCOCT00005883001	Putative Subtilisin-like protease SDD1 - SDD1	96,9	0,0	10,2
GSCOCT00016554001	Unknown protein	5,6	27,5	90,0
GSCOCT00005624001	8-hydroxyquercetin 8-O- methyltransferase - OMT2	88,1	4,5	1,0
GSCOCT00019030001	Secologanin synthase CYP72A1	20,3	85,5	4,2
GSCOCT00037166001	Dehydrin DH1a - XERO1	4,2	0,1	85,2
GSCOCT00007209001	Putative Subtilisin-like protease - ARA12	17,1	78,3	6,1

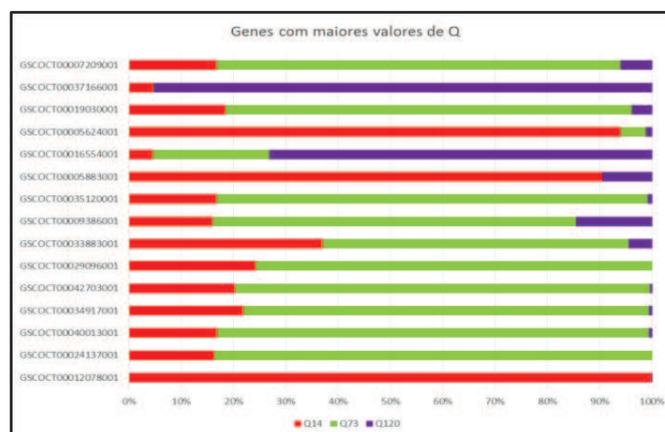


Figura 14 Genes com maiores valores de Q entre todos os clones tolerantes (14, 73 e 120). O valor Q foi obtido pelo quociente do valor de FI para cada clone em relação ao valor de FI para o clone sensível (22)

No Clone 14, os genes candidatos que mais se destacaram quanto ao Fator de Indução (FI), quando comparados aos outros clones (73, 120 e 22) está apresentado na Tabela 5. Dentro desse grupo, foram encontrados genes que expressam proteínas de considerável importância na fisiologia da raiz do cafeeiro como a galactinol sintase - *CcGols3* (GSCOCT00042863001) e a proteína transportadora de nitrato *CcNRT1.4* (GSCOCT00031111001), dentre outras. Também foram encontrados genes que expressam *No Hits* (GSCOCT00042816001 e GSCOCT00004150001).

O gene que codifica a proteína transportadora de nitrato *CcNRT1.4* (GSCOCT00031111001) apresentou mais alto valor de FI para o clone 14, a expressão foi consideravelmente maior na condição NI (Figura 15). A maior expressão deste gene foi observada no clone 22 na condição I, no Clone 73, para as duas condições, houve baixa expressão; e no clone 120 na condição I a expressão foi nula (Figura 15A). O nitrogênio é um dos compostos químicos mais requeridos pelo cafeeiro, sendo adquirido principalmente por meio da absorção do solo pelas raízes na forma de nitrato (MEDA et al., 2009).

O gene que codifica a proteína Galactinol Sintase (E C. 2.4.1.123) (GSCOCT00042863001) também apresentou uma elevada indução nos clones na condição de estresse (NI). O valor da expressão para o clone 73 nas duas condições foi igual (Figura 15C). Em clones de *C. arabica*, a expressão da Galactinol sintase só foi observada em folhas, a isoforma 3 teve sua expressão avaliada por *Northern* eletrônico e apresentou alta expressão nas condições de seca (SANTOS et al., 2011). Em situações de estresse por deficiência hídrica, é acumulada pela planta e atua como osmoprotetora em plantas tolerantes (TAJI et al., 2002).

Tabela 5 Genes candidatos que obtiveram valor superior no clone 14 quanto ao fator de indução (FI), em resposta ao estresse hídrico, com base nos resultados do sequenciamento obtidos pelo pirosequenciamento 454

<b>Contigs</b>	<b>Descrição</b>	<b>FI</b>
<b>GSCOCT00031111001</b>	Nitrate transporter	77,62
<b>GSCOCT00021907001</b>	Unknown	46,82
<b>GSCOCT00042863001</b>	Galactinol synthase	46,50
<b>GSCOCT00042816001</b>	No hits	31,28
<b>GSCOCT00004150001</b>	No hits	12,80

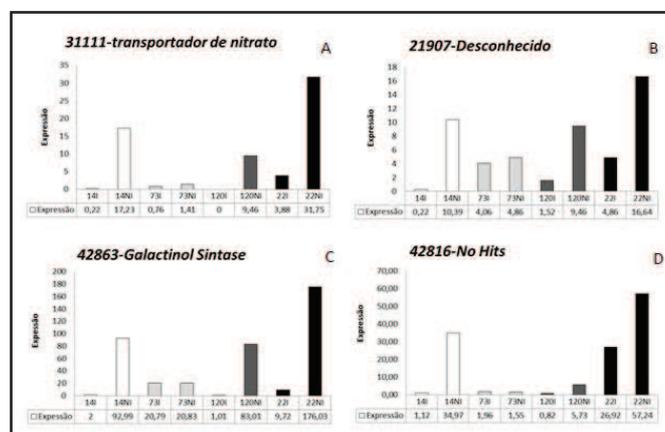


Figura 15 Gráficos ilustrativos da expressão de genes candidatos que apresentaram valores de FI mais significativos para o clone 14

Legenda: **A.** gene que codifica a proteína transportadora de nitrato (*Contig 31111*); **B.** gene que codifica a proteína cuja função é desconhecida (*Contig21907*); **C.** gene que codifica a proteína Galactinol Sintase (*Contig 42863*); **D.** gene que codifica *No Hits* (*Contig 42816*). O fator de indução foi calculado pelo quociente do valor da expressão no clone não irrigado sobre o clone irrigado. Os resultados da expressão foram obtidos pela análise *in silico* dos hits encontrados nos clones analisados por meio da ferramenta BlastP com sequências gênicas do CAFest.

Para a avaliação do FI no Clone 73, a Tabela 6 ilustra os quatro principais genes que codificam proteínas que apresentam valores diferenciais quando comparados com os outros clones (14, 120 e 22). O gene que induz a expressão do GAGA (GSCOCT00020529001) foi o que apresentou maior valor do FI neste clone e tem valor de expressão induzido nos clones na condição NI, o que pode estar relacionado à resposta da planta ao estresse sofrido (Figura 16A). O GAGA é um fator de transcrição que se liga aos promotores de histonas, protegendo as proteínas da degradação por choque térmico (BERGER; DUBREUCQ, 2012; PARSELL; LINDQUIST, 1993).

O gene que expressa a proteína PHD – proteína de inibição de crescimento 2 (GSCOCT00026541001), apresentou maior expressão no clone

73 e nos clones 14, 22, e 120 não apresentou indução (Figura 16B). Proteínas PHD são proteínas que podem associar-se com histonas HAT (acetiltransferases) ou HDAC (desacetilases) formando um complexo cujas funções são reparar moléculas de DNA ou realizar apoptose de células (ALBERTS et al., 2002).

O gene que codifica a CDK – Ciclinas Dependentes de Quinase (GSCOCT00031590001) apresenta indução nos clones 73 e 120 (Figura 16C), o valor da expressão é maior na condição NI. No clone sensível, a expressão foi maior na condição controle (irrigado).

A Figura 16D ilustra o gene que codifica a proteína rica em leucina (GSCOCT00014070001) e este se destaca por não apresentar expressão no clone 22 (clone sensível à seca). A maior expressão foi verificada no clone 14 na condição I. No clone 73, o valor de FI foi de 7,1 e no clone 120 somente na condição I a expressão foi detectada. Os fatores transcricionais da família HD-Zip são exclusivamente encontrados no reino vegetal e a expressão destes é regulada por fatores ambientais como seca, altas temperaturas e estresse osmótico (RÉ et al., 2010).

Tabela 6 Genes candidatos que obtiveram valor superior no clone 73 quanto ao fator de indução (FI), em resposta ao estresse hídrico, com base nos resultados do sequenciamento obtidos pelo pirosequenciamento 454

<b>Contigs</b>	<b>Descrição</b>	<b>FI</b>
<b>GSCOCT00020529001</b>	Ativador Transcricional de Ligação ao GAGA	12,74
<b>GSCOCT00026541001</b>	Prevista: Proteína "finger" PHD ING2	10,35
<b>GSCOCT00031590001</b>	Suposta CDK	7,96
<b>GSCOCT00014070001</b>	Proteína rica em leucina	7,14

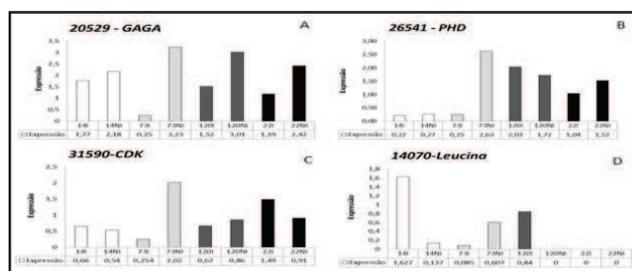


Figura 16 Gráficos ilustrativos da expressão dos genes candidatos que apresentaram valores de FI mais significativos para o clone 73

Legenda: **A.** gene que codifica o ativador transcricional de ligação ao GAGA (*Contig 20529*); **B.** gene que codifica a Proteína “finger” PHD (*Contig 26541*); **C.** gene que codifica a proteína Suposta CDK (*Contig 31590*); **D.** gene que codifica a proteína rica em Leucina (*Contig 14070*). O fator de indução foi calculado pelo quociente do valor da expressão no clone não irrigado sobre o clone irrigado. Os resultados da expressão foram obtidos pela análise *in silico* dos *hits* encontrados nos clones por meio da ferramenta BlastP com sequências gênicas do CAFEst.

No Clone 120, os quatro principais genes que codificam proteínas que apresentam valores diferenciados quando comparados com os outros clones (14, 73 e 22) estão apresentados na Tabela 7. O gene que induz a expressão da Dehidrina (GSCOCT00037166001) apresentou maior valor de FI para os clones na condição NI, principalmente no clone 22 e no clone 120 (Figura 17A). Esta proteína tem expressão induzida em tecidos vegetativos quando a planta sofre algum tipo de estresse que cause desidratação da célula (seca, salinidade, frio ou calor) protegendo-a da desidratação. A expressão desta aumenta a produção de ABA, o que pode aumentar a tolerância da planta ao estresse (HANIN et al., 2011).

Os genes que expressam *No Hits* (GSCOCT00024650001 e GSCOCT00041957001) (Figuras 17B e 17D) apresentaram valores de expressão nulo no clone sensível à seca (clone 22) e expressão elevada no clone 120, na condição de estresse (NI). No clone 14, não há indução do gene GSCOCT00024560001 entre as duas condições (Figura 17B). O gene que

codifica a F-Box (GSCOCT00013922001) apresentou indução na condição NI em todos os clones (Figura 17C). Esta proteína tem importante função na degradação das células pelo proteossoma, também no silenciamento gênico do RNAi (ALBERTS et al., 2002).

Tabela 7 Genes candidatos que obtiveram valor superior no clone 120 quanto ao fator de indução (FI), em resposta ao estresse hídrico, com base nos resultados do sequenciamento obtidos pelo pirosequenciamento 454

Contigs	Descrição	FI
GSCOCT00037166001	dehydrin DH1a	446,72
GSCOCT00024650001	No hits	17,82
GSCOCT00013922001	f-box family protein	10,18
GSCOCT00041957001	No hits	7,63

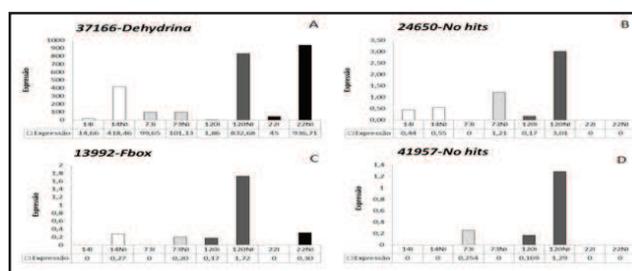


Figura 17 Gráficos ilustrativos da expressão dos genes candidatos que apresentaram valores de FI mais significativos para o clone 120

Legenda: **A.** gene que codifica a Dehidrina (*Contig* 37166); **B.** gene que codifica *No hits* (*Contig* 24650) **C.** gene que codifica a proteína F-Box (*Contig* 13992). **D.** gene que codifica *No hits* (*Contig* 41957). O fator de indução foi calculado pelo quociente do valor da expressão no clone não irrigado sobre o clone irrigado. Os resultados da expressão foram obtidos pela análise *in silico* dos *hits* encontrados nos clones por meio da ferramenta BlastP com sequências gênicas do CAFEst.

No Clone 22, sensível à seca, os genes candidatos que obtiveram maior FI apresentaram valores entre 60 e 34. A Tabela 8 ilustra os cinco principais

genes que codificam proteínas com valores significativos quando comparados com os clones tolerantes. O gene que codifica para uma proteína prevista (GSCOCT00018761001) foi o que apresentou maior valor do FI e tem expressão induzida nos clones na condição NI, principalmente no clone 22, o que pode estar relacionado à resposta da planta ao estresse sofrido (Figura 18A).

O gene que expressa a Citocinina-Ribosídeo-5'-monofosfatase (GSCOCT00024672001) possui uma alta expressão 120 e 22 nas condições de seca (NI) e no clone 73 não houve indução (Figura 18B). A citocinina é um hormônio de crescimento produzido nas raízes, sendo transportado pelo xilema para o restante das plantas na forma de ribosídeos (PINO-NUNES et al., 2009).

O gene que codifica a proteína BRCA (GSCOCT00042965001) apresenta indução para todos os clones na condição NI. Somente no clone 14 não há valor significativo de FI. O maior valor é apresentado no clone 22 (Figura 18C). Junto com a ubiquitina, esta proteína atua prevenindo a proliferação descontrolada das células (ATIPAIRIN; RATANAPHAN, 2011).

O gene que codifica o No Hits (GSCOCT00015847001) não apresenta indução nos clones 14 e 73 e no clone 120, o FI não é significativo. Somente o clone 22 apresenta FI com valor significativo (Figura 18D).

Tabela 8 Genes candidatos que obtiveram valor superior no clone 22 quanto ao fator de indução (FI), em resposta ao estresse hídrico, com base nos resultados do sequenciamento obtidos pelo pirosequenciamento 454

<b>Contigs</b>	<b>Descrição</b>	<b>FI</b>
<b>GSCOCT00018761001</b>	Predicted protein	60,49
<b>GSCOCT00024672001</b>	Cytokinin riboside 5'-monophosphate	52,43
<b>GSCOCT00042965001</b>	BRCA1	44,36
<b>GSCOCT00015847001</b>	No hits	36,29
<b>GSCOCT00026138001</b>	Heat shock factor protein HSF30	34,28

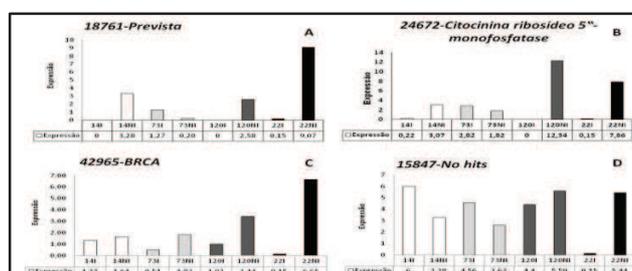


Figura 18 Gráficos ilustrativos da expressão dos genes candidatos que apresentaram valores de FI mais significativos para o clone 22

Legenda: **A.** gene que expressa uma proteína prevista (*Contig 18761*); **B.** gene que expressa a proteína Citocinina ribosídeo 5'-monofosfatase (*Contig 24672*). **C.** gene que expressa a proteína BRCA (*Contig 42965*). **D.** gene que expressa *No hits* (*Contig 15847*). O fator de indução foi calculado pelo quociente do valor da expressão no clone não irrigado sobre o clone irrigado. Os resultados da expressão foram obtidos pela análise *in silico* dos *hits* encontrados nos clones por meio da ferramenta BlastP com sequências gênicas do CAFest.

### 3.2 qPCR para os genes candidatos

Com base nos resultados obtidos, foi possível analisar a expressão de alguns genes responsivos à deficiência hídrica em raízes do cafeeiro. Houve resposta diferencial tanto entre clones, como entre o mesmo clone nas duas condições analisadas. Para a maioria dos resultados, a expressão dos genes mostrou-se induzida com o déficit hídrico. Em comparação com os resultados de análise *in silico*, a maior parte dos transcritos analisados mostrou correspondência com os resultados de qPCR.

O ABA é um hormônio de grande importância no desenvolvimento da planta, bem como em resposta desta a estresses abióticos (SOON et al., 2012), sendo sintetizado em raízes e folhas (THOMPSON et al., 2007). Sua ação é controlada pelo exato equilíbrio entre a sua biossíntese e o catabolismo.

Nas raízes de cafeeiro, foram analisados os elementos cruciais de sua rota metabólica: O NCED, responsável pela biossíntese deste hormônio; Sinalizadores de ABA, como o PYL e a Fosfatase 2C crucial para ativação de

quinases, principalmente a SRNK2 ativadora da rota do ABA; os fatores de transcrição AREB e NAC, que atuam nas vias dependente e independente do ABA, respectivamente e os genes que regulam o nível de ABA, possibilitando o equilíbrio exato deste no cafeeiro.

Também foram analisados genes que atuam no mecanismo responsivo do cafeeiro: Dehidrinas, LEA, Galactinol Sintase, genes que induzem a síntese de proteínas que protegem a planta contra estresse oxidativos (Peroxidase, Glutathione-S-Transferase, AKR), que são responsáveis pelo desenvolvimento da planta (Transportadores de nitrato, citocininas, MALD3, quinases (MAPK2)); elementos responsivos a estresses bióticos, mas que estão interligados à resposta a estresses abióticos, promovendo um *cross-talk* entre os eventos responsivos da planta à deficiência hídrica (quitinase, jasmonato e cisteína proteínases).

Dentro das análises também foi possível observar genes com funções desconhecidas em plantas (*No Hits*) que apresentaram expressão diferencial entre os clones analisados, bem como em relação às duas condições.

## **BIOSSÍNTESE DO ÁCIDO ABISCÍSICO**

### **Expressão do transcrito que codifica para a enzima dioxigenase 9-cis-epoxicarotenoide - NCED**

Conforme mostramos resultados para a NCED3 (GSCOCT00021488001), a expressão desta foi maior na condição de estresse (NI) para todos os clones analisados (Figura 19A). Entre os clones tolerantes, o clone 14NI foi o que apresentou maior indução deste gene em resposta à seca, com valor quatro vezes maior que os outros clones na mesma condição e cinco vezes maior que o mesmo clone na condição de irrigação (I). No clone sensível (22), a NCED3 apresentou baixa expressão na condição NI em comparação com o clone 14.

Por ser a enzima-chave na biossíntese do ABA e o acúmulo deste hormônio provocar a expressão de muitos genes responsivos ao estresse, aumentando a tolerância da planta (LUCHI et al., 2001), é justificável a expressão e o acúmulo de NCED nas raízes analisadas ser maior para os clones na condição NI, uma vez que na literatura já foi relatado que o ABA pode ser acumulado de 10 a 30 vezes mais nas plantas submetidas a estresse hídrico em comparação a plantas não estressadas (LEUNG et al., 2012), também mostrado no trabalho publicado por Simkin et al. (2008), no qual foi observado que o nível de transcrição da NCED3 foi rapidamente aumentado em folhas de *C. canephora* sob condições de deficiência hídrica.

Em comparação com os resultados obtidos *in silico* (Figura 19B), é possível ressaltar correspondência nos resultados com qPCR. Entretanto, nas análises *in silico* foi observado uma maior expressão deste gene no clone sensível. Esta discrepância em relação ao quantitativo de expressão para o clone 22 pode estar relacionado a diversos fatores, entre eles o processo de normalização.

Resultados que corroboram com os encontrados por Khodambashi, Shiran e Graraghanipour (2013), nos quais a avaliação deste gene em tecidos de grãos de feijão (*Phaseolus vulgaris*), sob condições de estresse por deficiência hídrica, possibilitou observar maior expressão destes nos genótipos tolerantes que nos genótipos sensíveis.

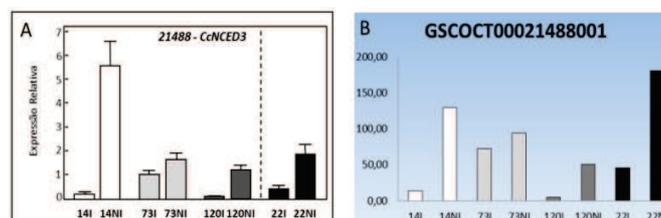


Figura 19 Perfis de expressão do gene *CcNCED3* (GSCOCT00021488001)

Legenda: Por qPCR (A) e quantificação *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

#### **Expressão dos transcritos que codifica para o receptor de ABA PYL (PYL7.2) e para a fosfatase 2C – PP2C-1**

Para a análise de expressão do gene que codifica para o receptor de ABA PYL 7, é observado que houve indução da expressão na condição NI para os clones tolerantes 14 e 120 (Figura 20A), sendo que a maior indução foi observada no clone 14 evidenciando que, para este clone a expressão deste gene é de grande importância para ativação dos mecanismos de resposta dependentes da sinalização do ácido abscísico. Nos outros clones, 73 e 22 os valores de expressão relativa foram similares.

Para a expressão da fosfatase 2C tipo 1 (GSCOCT00021820001), em todos os clones, a expressão foi induzida na condição de estresse (NI) (Figura 21A). Entre os clones tolerantes, o clone 14NI foi o que apresentou maior expressão, com valor quatro vezes maior que os outros clones na mesma condição e seis vezes maior que o mesmo clone na condição de irrigação (I). No clone sensível (22) a PP2C-1 apresentou baixa expressão.

Comparando com os resultados anteriormente demonstrados por Marraccini et al. (2012) e Vieira et al. (2013) para análises de expressão de uma

fosfatase 2C para estes clones em folhas, pode-se observar diferenças nos resultados, uma vez que para este gene houve uma redução da expressão durante a seca.

Os resultados mostram que apesar de o ABA ser sintetizado nas raízes, a sinalização nestes tecidos é mais tardia, uma vez que primeiramente as folhas necessitam receber o sinal, onde ocorre o fechamento estomático, e, conseqüentemente, a ativação dos receptores PYL/PYR/Rcars e inativação das fosfatases, para que ocorra um aumento da condutância de água das raízes para as outras partes das plantas.

Em comparação com os resultados obtidos *in silico* (Figura 21B) é possível observar correspondência para a maioria dos resultados. A maior quantidade desta foi observada no clone 120 na condição NI. O clone 73 comportou-se de maneira contrária aos outros clones, apresentando maior expressão na condição controle (I). No clone sensível, é notável uma grande quantidade desta fosfatase também na condição NI, enfatizando que por possuir grande concentração desta na planta, não há necessidade de ser expressa mais fosfatase, a expressão desta, para esta análise foi maior no clone sensível.

Para a sinalização de ABA nas células, estudos recentes indicam a interação direta entre receptores intracelulares de ABA (PYR/PYL/RCARS), proteínas fosfatases do tipo 2C (ANTONI et al., 2012; CUTLER et al., 2010) e proteínas relacionadas com quinases SnRK2s (HAUSER; WAADT; SCHROEDER, 2011). A ligação do ABA aos receptores intracelulares forma um complexo que ao ligar-se à fosfatase 2C inativa-a, ativando quinases como a SnRK2 que atuam fosforilando os fatores de transcrição de genes responsivos ao ABA (LEUNG et al., 2012). Os fatores de transcrição ABF/AREB reconhecem sequências do promotor ABRE, induzindo a expressão de componentes regulatórios para a abertura estomática (FUJII et al., 2009), garantindo tolerância da planta ao estresse hídrico.

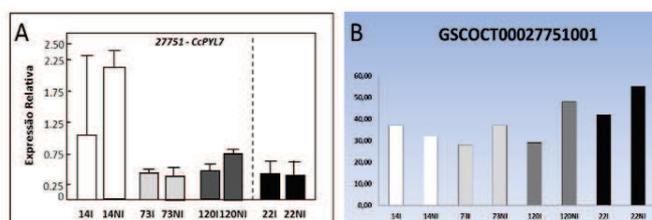


Figura 20 Perfis de expressão do gene *CcPYL7.2* (GSCOCT00027751001)

Legenda: Por qPCR (A) e *in silico* (B) em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

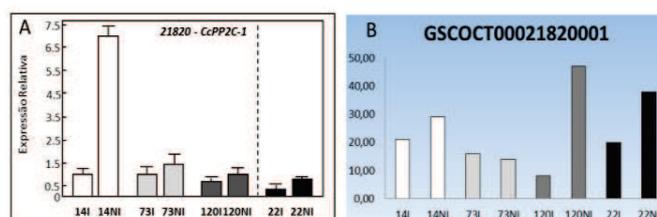


Figura 21 Perfis de expressão do gene *CcPP2C-1* (GSCOCT00021820001)

Legenda: Por qPCR (A) e *in silico* (B) em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

### Expressão do transcrito que codifica para o AREB

Para o transcrito que codifica o Fator de Transcrição responsivo ao ABA – AREB (GSCOCT00034982001), foi possível observar maior expressão deste na condição de estresse (NI) apresentando alta indução em todos os clones avaliados (Figura 22A). Entre os clones tolerantes, a maior expressão deste foi observada no clone 120NI, sendo que na condição I a expressão deste transcrito

para este clone foi nula. Também foi apresentada alta indução da expressão do AREB para o clone 14 na condição de estresse (NI), com valor de expressão cinco vezes maior que este clone na condição de irrigação (I). No clone sensível (22) a expressão do AREB também foi maior para os clones na condição NI. Estes resultados corroboram com os adquiridos por Marraccini et al. (2012) e Vieira et al. (2013) para análises em folhas nos mesmos clones e nas mesmas condições.

Nos resultados da análise *in silico*, os maiores valores de expressão foram observados no clone 73 na condição I e no clone 22 na condição NI, conforme Figura 22B. No clone 73, o valor de expressão foi ligeiramente maior para a condição controle (I), contradizendo com os resultados de qPCR (Figura 22B).

Fatores de transcrição AREB/ABFs estão envolvidos na resposta da planta por meio do mecanismo ABA-responsivo e, como tal, os fatores têm sua transcrição elevada como resposta a mecanismos de tolerância à seca (SHINOZAKY; YAMAGUSHI-SHINOZAKI, 2007).

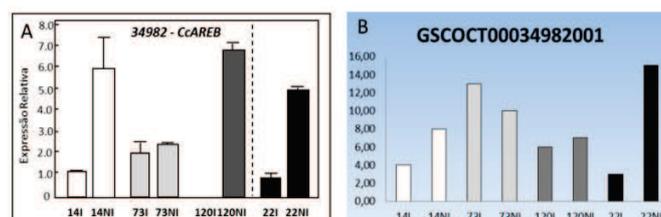


Figura 22 Perfis de expressão do *CcAREB* (GSCOCT00034982001)

Legenda: Por qPCR (A) e quantificação *in silico* (B) em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

### Expressão dos transcritos que codificam para os Fatores de Transcrição NAC

Para genes que codificam Fatores de Transcrição NAC, foram avaliados em três *contigs* (GSCOCT00006253001, GSCOCT00021185001 e GSCOCT00039597001). Houve expressão diferencial e indução da expressão na condição NI em os todos os clones (Figuras 23A, 23B, 23C). Para os NACs, foi observada maior expressão no clone 120, na condição de estresse (NI) (Figuras 23A e 23B). Para o NAC4, a maior expressão foi observada no clone 14, também na condição de seca (Figura 23C). No clone 22, todos os NACs apresentaram elevada indução de expressão na condição de estresse.

Os fatores de transcrição NAC têm função mediadora na formação de raízes laterais e sinalização de auxina, bem como resposta a estresse biótico e abiótico (XIA et al., 2010). Podem atuar na via de sinalização ao estresse por déficit hídrico, controlando a expressão do gene RD29A (na via dependente de ABA) e na codificação de LEAs (via independente de ABA) (TRAN et al., 2004).

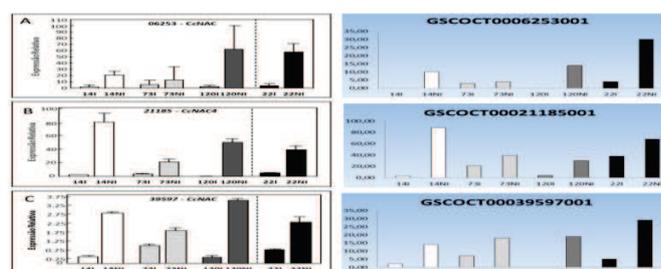


Figura 23 Perfis de expressão do *CcNAC* (A e B) e *CcNAC4* (C) na qPCR e *in silico*, em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado)

Legenda: **A** (GSCOCT00006253001) **B** (GSCOCT00021185001) e **C** (GSCOCT00039597001). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

### **Expressão dos transcritos que codificam para LEA**

Para o transcrito que codifica a LEA (GSCOCT00022121001), com base nas análises de qPCR, foi possível observar maior expressão desta na condição de estresse (NI) em todos os clones estudados (Figura 24A). Entre os clones tolerantes, o clone 14NI foi o que apresentou maior expressão para esta chaperona, com valor vinte vezes maior que os outros clones na mesma condição e 45 vezes maior que o mesmo clone na condição de irrigação (I). Com base nas análises *in silico*, os resultados mostram que houve alta concentração de LEA nos clones 22 e 120 na condição NI (Figura 24B).

Proteínas LEA são de grande abundância nas células vegetais em condições de estresse, protegendo-as da desidratação, evitando a cristalização de proteínas e outras moléculas durante o período de seca, capturando íons e restabelecendo as pontes de hidrogênio que formam a membrana celular (SEKI et al., 2001; SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2007), consequentemente tendem a estar em maior nível em plantas sensíveis à seca (CHAVES; OLIVEIRA, 2004).

Dentre as proteínas LEA, o grupo das Dehidrinas é difundidamente analisado para estudos de respostas da planta à seca e ao frio. Neste trabalho, foram analisados três tipos de Dehidrinas. (DH1a, DH2a e DH3) (GSCOCT00037166001, GSCOCT00014237001, GSCOCT00024547001). Nas duas análises, análise *in silico* e qPCR, foi observada uma alta indução dos genes que codificam as Dehidrinas DH1a e DH3 para o clone 22 na condição de seca (NI) (Figura 25A e 25C). Para a DH2a, a expressão foi maior no clone 73 para qPCR, apesar de o maior acúmulo desta ter sido observado no clone 120 (Figura 25).

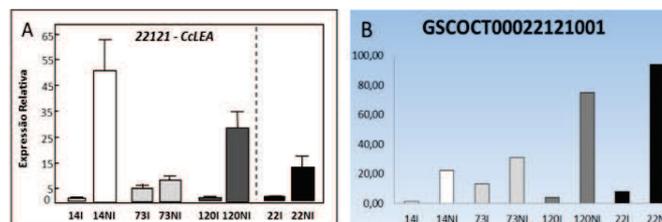


Figura 24 Perfis de expressão do gene *CcLEA* codificando uma chaperona LEA (GSCOCT00022121001)

Legenda: Na qPCR (A) e *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

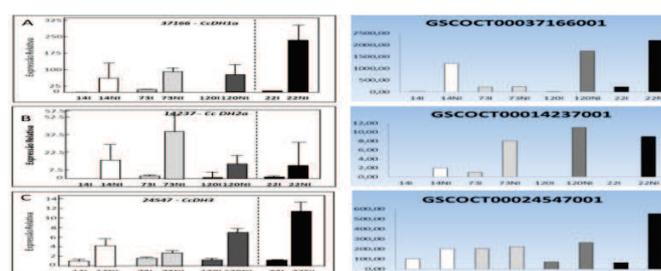


Figura 25 Perfis de expressão da Dehidrina na qPCR e *in silico*, em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado)

Legenda: A. *DH1a* (GSCOCT00037166001), B. *DH2a* (GSCOCT00014237001), C. *DH3* (GSCOCT00024547001). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

### **Análise de expressão para o transcrito que codifica para a enzima ABA-8-hidrolase**

Com base nos resultados para o gene que codifica a ABA-8-hidrolase (GSCOCT00024390001), é possível observar que estes correspondem com os resultados esperados, uma vez que o transcrito apresenta expressão baixa ou nula nos clones em condição de estresse (Figura 26). Para os resultados de qPCR, entre os clones tolerantes o que apresentou valor de expressão mais significativo para este gene foi o clone 120 na condição de irrigação (I). Não houve expressão neste clone na condição NI (Figura 26A).

Comparando os dados de qPCR e *Northern* Eletrônico, foi possível observar nos dados apresentados que, para o clone 73, houve valor significativo de expressão na análise de *Northern*. Mesmo na condição de estresse, a expressão foi maior que para o restante dos clones analisados (Figura 26B).

A enzima ABA-8-hidrolase, catalisada pelo citocromo P450, atua no controle do nível de ABA em plantas (KUSHIRO et al., 2004). Tem função de clivar a molécula de ABA por hidrólise, inativando-a em situações de estresse na planta por desidratação (KUSHIRO et al., 2004; UMEZAWA et al., 2006). Em condições de estresse de seca, a expressão deste gene é diminuída, sendo aumentada quando a planta é hidratada. Sendo assim, o nível de ABA na planta é aumentado, possibilitando a regulação da síntese de proteínas e metabólitos que atuam em resposta ao estresse por deficiência de água.

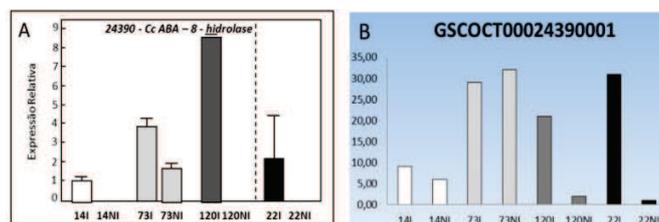


Figura 26 Perfis de expressão do gene *CcABA-8-hidrolase* (GSCOCT00024390001)

Legenda: Por qPCR (A) e *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

## ESTRESSE OXIDATIVO

Dentre os estresses sofridos pelas plantas, a seca é o mais frequente e o potencialmente mais comprometedor para a produtividade, conseqüentemente, pode provocar o aumento de Espécies Reativas de Oxigênio (EROs), altamente tóxicas, que podem levar à degradação de proteínas e até à morte da planta (GILL et al., 2013; SOARES; MACHADO, 2007). Para minimizar este problema, as células vegetais possuem a maquinaria fundamental para combater o estresse oxidativo, estimulando a biossíntese de componentes antioxidantes, além de aumentar a atividade destas pela ação coordenada de agentes antioxidantes de natureza enzimática e não enzimática (MITTLER, 2002).

As enzimas antioxidantes mais conhecidas são: a superóxido dismutase (SOD), responsável pela conversão do radical superóxido em peróxido de hidrogênio; ascorbato peroxidase (APX), que converte o peróxido de hidrogênio em ascorbato; catalase (CAT), que transforma peróxido de hidrogênio em água; glutatona redutase (GR), que controla os níveis de glutatona reduzida e glutatona S-transferase (GST), que pode atuar no combate ao peróxido de

hidrogênio, peroxidases (PRX) e outros tipos de redutases que são ativadas em resposta ao estresse oxidativo induzido pelo déficit hídrico (GILL et al., 2013; MILLER et al., 2010).

As glutatonas GST e GR são os dois principais componentes da via do ascorbato, cujas ações estão interligadas. A GR é um antioxidante enzimático dependente de NADPH que atua principalmente mantendo os níveis de GST nas células (GILL et al., 2013).

Outro elemento de grande importância na atuação contra o estresse oxidativo é o ácido jasmônico, um importante hormônio de defesa vegetal, que também atua na produção peróxido de hidrogênio, uma espécie reativa de oxigênio que, no início do estresse, pode atuar como molécula sinalizadora de defesa em plantas (SOARES; MACHADO, 2007). O ácido jasmônico induz a expressão de diversos genes relacionados à defesa contra estresses, agindo, portanto, num complexo de redes com *cross-talk* para vários outros tipos de vias de sinalização fito-hormonal.

#### **Análises de expressão do transcrito que codifica para a Peroxidase**

As peroxidases são classes de enzimas presentes em tecidos de animais, plantas e micro-organismos, catalisando as reações de óxido redução entre peróxido de hidrogênio e uma variedade de substratos (HIRAGA et al., 2001). As peroxiredoxinas (PRXs) são uma família ubíqua de peroxidases à base de tiol com função enzimática antioxidante (WOOD et al., 2003). Com base nos resultados para o transcrito que codifica a peroxidase PRX1 (GSCOCT00016870001), foi possível observar alta expressão desta no clone 14 na condição de estresse (NI) (Figuras 27A e 27B), podendo elucidar que, para este mecanismo de defesa, o clone 14 é o mais eficiente, em comparação aos outros clones analisados. Os resultados obtidos pelas análises *in silico* e de qPCR (Figura 27) evidenciam o mesmo padrão de expressão entre os clones

analisados, indução da expressão do gene *PRX1* na condição NI. Tais resultados corroboram com estudos anteriores realizados em plantas. Em *Xerophyta viscosa*, a expressão do gene *I-Cys Prx* é induzida em resposta a estresses abióticos, tais como: tratamentos com ABA e NaCl, desidratação, altas temperaturas e intensidade luminosa (MOWLA et al., 2002). No trabalho de Kim et al. (2009), observou-se que a proteína 2-Cys Prx1 além de desempenhar papéis antioxidativos durante o desenvolvimento da planta e fotossíntese, atuando como peroxidase, possui função de chaperonas moleculares.

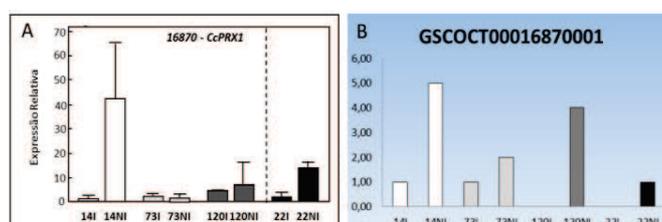


Figura 27 Perfis de expressão do gene *CcPRX1* que codifica a Peroxidase (GSCOCT00016870001)

Legenda: Na qPCR (A) e na análise *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

### **Análise de expressão do transcrito que codifica para Glutathione-S-Transferase**

Tanto para a análise *in silico* como pela validação *in vivo*, os resultados da expressão da Glutathione-S-Transferase (GSCOCT00015916001) mostram que ocorre aumento da expressão do gene nas condições de déficit hídrico, sendo a maior expressão apresentada no Clone 14 (Figura 28). Entretanto, houve significativa expressão em todos os clones tolerantes, evidenciando que a expressão deste gene pode ser uma resposta comum aos clones analisados.

A Glutathione-S-Transferase (GST) é uma proteína multifuncional codificada por uma extensa família gênica identificada em plantas, animais, fungos e bactérias (SHEEHAN et al., 2001). As GSTs têm sido identificadas como diferencialmente reguladas por diversos tipos de estresses, incluindo desidratação (BIANCHI; ROUX; VARTANIANA, 2002) e estresse oxidativo (SMIRNOFF, 1998).

Bianchi, Roux e Vartaniana (2002) identificaram o acúmulo de uma GST em plantas de *A. thaliana*, sujeitas à seca progressiva, denominado *GST8*. Após seis semanas, estas plantas apresentaram aumento nos níveis de *GST8* e evidenciaram que essa resposta não é transiente e que o respectivo gene é responsivo à seca de maneira independente de ABA. A indução da expressão do gene *CcGST*, em *C. canephora*, para todos os clones tolerantes após o tratamento de déficit hídrico pode evidenciar que este é um gene importante para a resposta de tolerância, uma vez que quanto maior o estresse oxidativo sofrido pela planta, menor são os níveis de glutathione (SMIRNOFF, 1998), como é claramente observado no clone 22, sensível ao déficit hídrico.

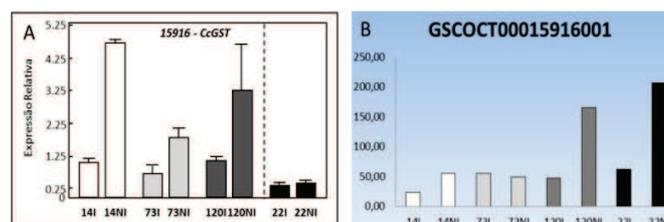


Figura 28 Perfis de expressão *CcGST* - Glutathione-S-Transferase (GSCOCT00015916001)

Legenda: Na qPCR (A) e na análise *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

### **Expressão do transcrito que codifica para o Metil Jasmonato esterase – MJE**

Para a expressão do gene *CcMJE* que codifica a MJE (GSCOCT00039836001), o clone 14 destaca-se com alta indução deste na condição NI (Figura 29A), resultado que é esperado, uma vez que a Metil Jasmonato Esterase está relacionada com a biossíntese do hormônio Ácido Jasmônico, responsável por diversas vias de sinalização e combate a estresses bióticos e abióticos.

Além de sua atuação na defesa da planta contra estresses oxidativos, sabe-se que o Ácido Jasmônico também tem importante atuação na resposta da planta contra estresses abióticos, além de estar diretamente relacionado com a expressão de receptores funcionais de ABA, responsáveis pela inibição de fosfatases do tipo 2C, proteínas conhecidas por ser a chave da regulação negativa do ABA (LACKMAN et al., 2011). Portanto, os dados aqui analisados corroboram com os resultados de expressão da fosfatase 2C acima relatados. A alta expressão de Jasmonato é requerida pelas plantas para a síntese de receptores funcionais de ABA, que inibem a função da proteína PP2C, altamente expressa nas raízes de *C. canephora*.

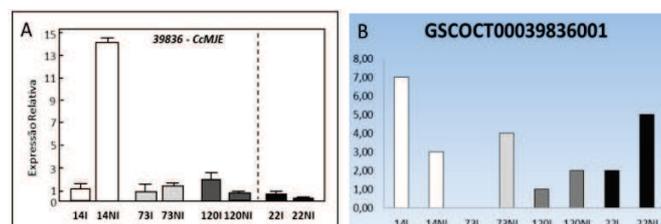


Figura 29 Perfis de expressão do transcrito *CcMJE* (GSCOCT00039836001) que codifica para a proteína Metil Jasmonato esterase

Legenda: Na análise de qPCR (A) e *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

## ESTRESSE OSMÓTICO

Em resposta ao estresse osmótico, as plantas sintetizam e acumulam osmoprotetores, como os polipídeos que são formas reduzidas de açúcares aldoses ou cetoses (MENG et al., 2008). O sorbitol, a forma reduzida de glicose-6-fosfato, é um dos principais produtos fotossintéticos de muitas famílias de plantas, também é o principal carboidrato que muitas espécies de plantas armazenam principalmente quando se deseja adquirir tolerância ao estresse abiótico por meio do ajuste osmótico (LI; LI, 2007).

### **Análise de expressão do transcrito que codifica para D-sorbitol-6-fosfato desidrogenase**

Os resultados de expressão da D-sorbitol-6-phosphate desidrogenase (GSCOCT00011142001), como elucidado nos resultados abaixo, estão apresentados na Figura 30. Quanto à expressão por qPCR, é observada uma maior expressão desta no clone tolerante 14NI, o que enfatiza a indução da expressão para as condições de estresse (NI) e nos clones que toleram melhor a deficiência hídrica (Figura 30A). Para os resultados de expressão *in silico*, a

maior expressão foi apresentada na condição NI para todos os clones avaliados, sendo maior no clone 14 (Figura 30B). A enzima D-sorbitol-6-fostato desidrogenase atua na resposta da planta contra estresses oxidativos (ZHANG et al., 2005) e osmóticos (MENG et al., 2008).

Li e Li (2007) observaram que em folhas e raízes de maçã e pêssgo sob estresse abiótico, o aumento de S6PDH e a diminuição na atividade SDH são sinais de regulação dos níveis de sorbitol em plantas sob estresse, consequentemente induziram o desenvolvimento destes tecidos. Nos resultados deste trabalho, em raízes de cafeeiro também observou-se maior expressão de S6PDH nos clones tolerantes na condição de estresse NI, o que corrobora com os dados de Li e Li (2007), podendo correlacionar este aumento da expressão com maior acúmulo de sorbitol e o melhor desenvolvimento das raízes em condições de estresse. É observado que nesta via de sinalização, o clone 14 é o que melhor responde no que tange ao acúmulo de sorbitol em *C. canephora*.

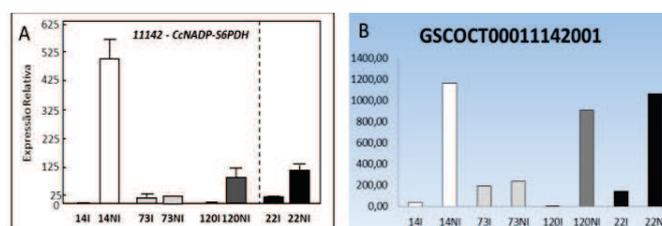


Figura 30 Perfis de expressão do gene *CcNADP-S6PDH* (GSCOCT00011142001) que codifica a NADP dependente de D-Sorbitol-6-fosfatase desidrogenase

Legenda: Por qPCR (A) e na análise *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

### Análise de expressão do transcrito que codifica para a Galactinol Sintase

Nas análises de qPCR, foi possível observar uma alta expressão do gene *CcGolS3* que codifica a Galactinol Sintase (GSCOCT00042862001) no clone 14 na condição de estresse (NI) (Figura 31A), com valor quatro vezes maior que os outros clones na mesma condição. Com base nos resultados de análise *in silico*, pode-se observar que houve alta concentração da GolS3 nos clones tolerantes 14 e 120 e no clone sensível (22) na condição de deficiência hídrica (Figura 31B).

A GolS3 é uma isoforma da Galactinol Sintase altamente expressa em condições de seca moderada e grave, atuando como osmoprotetoras em plantas tolerantes ao estresse hídrico (TAJI et al., 2002), como mostrado neste trabalho, em que houve uma alta expressão desta enzima no clone tolerante 14 na condição de déficit hídrico. Estudos mostram o acúmulo destas em folhas de *C. arabica* (SANTOS et al., 2011), mas nenhum trabalho relatou a expressão da Galactinol Sintase em raízes de clones de *C. canephora*.

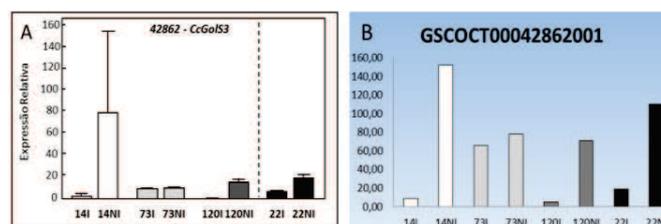


Figura 31 Perfis de expressão do gene *CcGolS3* codificando a Galactinol Sintase *CcGolS3* (GSCOCT00042862001)

Legenda: Na qPCR (A) e na análise *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

## **RESPOSTAS FISIOLÓGICAS DE RAÍZES DO CAFEEIRO PARA O ESTRESSE POR DEFICIÊNCIA HÍDRICA**

Em resposta à deficiência de água, as plantas desempenham respostas fisiológicas que auxiliam na adaptação contra o estresse abiótico permitindo-lhes sobreviver. A seca diminui a absorção de nutrientes pela raiz, bem como a inibição de enzimas implicadas no metabolismo do nitrogênio. A deficiência de nitrogênio induz a expressão de diversos transportadores da família NRT (MILLER et al., 2007). Estes transportadores regulam o nível de nitrato nas plantas, que agem na regulação de inúmeros genes (VIDAL; GUTIERREZ, 2008), como por exemplo, a NCED, enzima atuante na biossíntese do ABA (BOYD et al., 2009).

A biossíntese de hormônios também é afetada na planta para promover o desenvolvimento destas em resposta às adaptações. A citocinina é produzida principalmente nas raízes, de onde são transportadas para o caule pelo xilema na forma de ribosídeos, sendo um hormônio crucial na transmissão de informações sobre a disponibilidade de nutrientes na planta (PINO-NUNES, 2009).

### **Expressão do transcrito que codifica para a proteína de amadurecimento induzida por ABA**

Dentre os genes que conferem respostas fisiológicas para o desenvolvimento das raízes do cafeeiro, o transcrito que codifica uma proteína responsável pelo amadurecimento dos frutos (SRP) (GSCOCT00033650001), apresentou alto valor de indução em todos os clones para a condição NI (Figura 32A). Na validação por qPCR, a maior expressão foi observada no clone 14, entretanto, comparando com os resultados de análise *in silico* observa-se maior expressão nos clones 120 e 22 (Figura 32B).

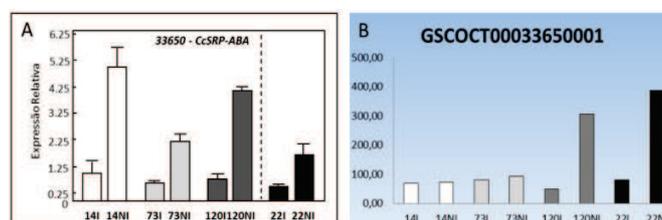


Figura 32 Perfis de expressão gênica para o transcrito *CcSRP-ABA* (GSCOCT00033650001) que codifica a proteína de amadurecimento induzida por ABA

Legenda: Nas análises de qPCR (A) e *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

#### Expressão do transcrito que codifica para a proteína de transporte de nitrato – NRT1.4

A expressão do gene *CcNRT1.4* (GSCOCT00031111001) codificado para o transportador de nitrato, permitiu observar que houve uma alta indução de expressão deste para os clones na condição de seca (Figura 33), tanto na análise *in silico* (Figura 33B) como nos dados de qPCR (Figura 33A), sendo que o maior valor de indução deste gene ocorreu no clone 22. O que reforça o importante papel fisiológico apresentado pelo nitrato para o ajuste osmótico e como molécula sinalizadora (MEDA et al., 2009), como observado nos resultados abaixo, tendo por base, a maior indução no clone sensível ao estresse hídrico.

Sánchez-Rodríguez, Romero e Ruiz (2013) avaliaram as respostas fisiológicas de tomateiros quanto à absorção de nitrogênio em plantas sob deficiência hídrica e concluíram que a partir do estudo de enxertia entre cultivares de tomate (*Lycopersicon esculentum*) tolerante (variedade Luciana) e sensível (variedade Josefina) para estresse por deficiência hídrica, houve uma

maior assimilação de nitrogênio nas plantas sensíveis, corroborando com uma melhor absorção deste pelas raízes e fotorrespiração pelas folhas.

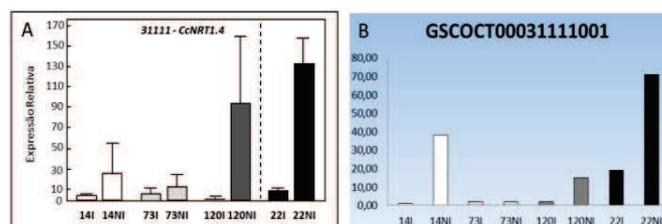


Figura 33 Perfis de expressão gênica para o transcrito *CcNRT1.4* (GSCOCT00031111001) que codifica a proteína de transporte de nitrato

Legenda: Na análise de qPCR (A) e *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

#### **Expressão do transcrito que codifica para a Citocinina ribosídeo-5'-monofosfatase fosforibohidrolase – PhLOG**

Conforme resultados, a Citocinina na forma de ribosídeo-5-monofosfatase (GSCOCT00024672001) apresentou maior valor de expressão no clone tolerante 120NI (não irrigado) e no clone 22NI (sensível). Houve indução da citocinina na condição de estresse para todos os clones (Figura 34). A citocinina é um hormônio presente em todos os tecidos vegetais, sendo sintetizado nas raízes, onde sua concentração é abundante, bem como em ápices caulinares e sementes imaturas (PINO-NUNES, 2009).

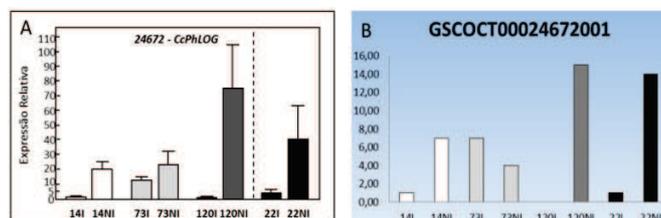


Figura 34 Perfis de expressão do transcrito *CcPhLOG* (GSCOCT00024672001) que codifica para a proteína Citocinina ribosídeo-5'-monofosfatase fosforibohidrolase

Legenda: Na análise de qPCR (A) e *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

#### Expressão dos transcritos que codificam para a nsLTPs

As nsLTP atuam basicamente na transferência de fosfolipídios nas células, bem como em várias funções biológicas na planta, atuando na formação de cera e cutina nas paredes celulares, no mecanismo de defesa contra fitopatógenos, bem como na adaptação das plantas a diferentes condições ambientais, tais como estresse por seca (JANG et al., 2004; TAPIA et al., 2013).

Como mostra os resultados para estes dois genes, em condições de estresse por déficit hídrico, a expressão é induzida, como forma de resposta à seca, uma vez que em células vegetais, a falta de água na planta reduz o turgor celular, como forma de proteção, as células acumulam compostos que garantem maior estabilidade à estrutura da parede celular (SHARP et al., 2004).

Para os resultados de expressão dos genes que codificam nsLTPs (proteínas não específicas de transferência de lipídeos), *CcMALD3* (GSCOCT00040458001), apresentado na Figura 35, e *CcLTP2* (GSCOCT00040232001), apresentado na Figura 36, é observado que os resultados para os dois genes se correspondem. Em todos os clones é

apresentada indução destes genes para a condição NI nas duas análises (qPCR e expressão *in silico*). Analisando os resultados de expressão relativa por qPCR, observa-se que o clone 14 na condição NI apresentou maior valor de expressão para os dois genes (Figura 35A e Figura 36A). Já os resultados de expressão *in silico* mostram que a maior expressão destes genes é apresentada no clone 22 (Figura 35B e Figura 36B).

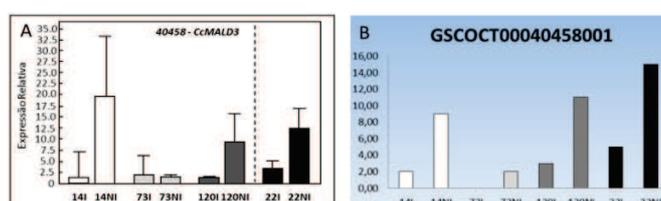


Figura 35 Perfis de expressão do transcrito *CcMALD3* (GSCOCT00040458001) que codifica para a proteína nLTP

Legenda: Na análise de qPCR (A) e *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

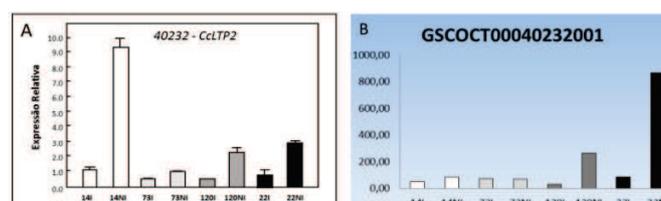


Figura 36 Perfis de expressão do transcrito *CcLTP2* (GSCOCT00040232001) que codifica para um tipo de proteína Nltp

Legenda: Na análise de qPCR (A) e *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e o clone 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

### **Expressão dos transcritos que codificam para Fosfatases ácidas**

Para os dados de expressão de duas isoformas da fosfatase ácida *CcPP3B* (GSCOCT00021916001 e GSCOCT00022022001), observa-se que o perfil de expressão do gene GSCOCT00021916001 corrobora com os valores de expressão apresentados na análise *in silico* para todos os clones (Figura 37A), sendo que os maiores valores são apresentados na condição I para os clones 14, 22 e 120 e na condição NI para o clone 73. O maior valor de expressão relativa (qPCR) é apresentado no clone 14 e de expressão *in silico* é no clone 22 (Figura 37A).

Os perfis de expressão do gene GSCOCT00022022001 estão apresentados na Figura 37B. Neste caso, somente no clone 14 o perfil de expressão por qPCR não corresponde com o perfil de expressão *in silico*. Para os resultados de qPCR o maior valor de expressão deste gene (GSCOCT00022022001) é apresentado no clone 14 na condição NI, para os dados de expressão *in silico* o maior valor também é apresentado no clone 14, mas na condição I (Figura 37B).

A fosfatase ácida é uma enzima utilizada na caracterização de plantas, uma vez que sua atividade é induzida na deficiência de fósforo, como pode ser mostrado nos trabalhos de Li et al. (2011) e Malone et al. (2007). Esta enzima captura o fósforo inorgânico presente em solos ácidos, transportando-o para o interior de células radiculares através de transportadores de fosfomembranares. Portanto, em meio com deficiência de fósforo, a expressão desta enzima é induzida com a finalidade de tentar suprir a falta deste nutriente na planta, possibilitando assim o desenvolvimento da planta mesmo em condições de redução dos níveis de fósforo, como ocorre em solos com déficit hídrico (LI et al., 2011).

Curiosamente, os resultados expostos na Figura 37 mostram que para o gene GSCOCT00021916001 somente o clone 73 apresentou maior expressão

desta enzima quando submetido ao déficit hídrico, condição NI (Figura 37A). Para o gene GSCOCT0002202201, a expressão desta enzima na condição NI foi apresentada somente no clone 14 (Figura 37B).

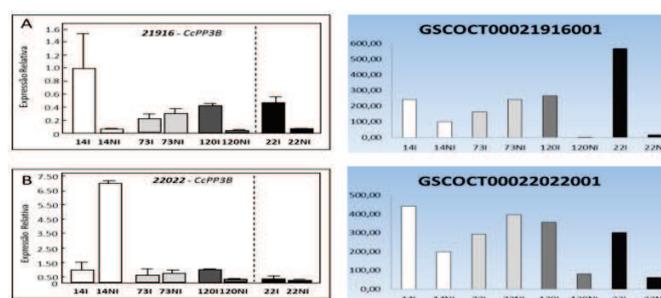


Figura 37 Perfis de expressão de transcritos que codificam proteínas fosfatases *CcPP3B*

Legenda: Para as análises de qPCR e *in silico* em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). **A.** Perfis de expressão da fosfatase GSCOCT00021916001. **B.** Perfis de expressão da fosfatase GSCOCT00022022001. Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

### Expressão do transcrito que codifica para um tipo de Quitinase inibidora de Xilanase – CLXIP

A expressão do gene *CcCLXIP* (GSCOCT00013168001) foi alta somente para o clone 14, na condição NI, que apresentou uma alta indução (Figura 38A). Para os resultados de análise *in silico*, a maior expressão é apresentada em todos os clones na condição NI, destacando os clones 14, 120 e 22, sendo maior no clone sensível 22 (Figura 38B).

Em folhas de café, o gene *CaCLXIP* é altamente transcrito quando as folhas estão infectadas com *Xylella fastidiosa* (VASCONCELOS et al., 2011). Em arroz, a indução ocorre pela infecção por *Aureobasidium pullulans* ou *Trichoviride* (WU et al., 2013). A *CaCLXIP* também pode ser induzida por patógenos e

por estresse abiótico, como injúria ou tratamento com metil Jasmonato (IGAWA et al., 2005), tendo como principal finalidade o suporte mecânico da célula vegetal, sendo sinalizada pelo acúmulo de Jasmonato (WEI et al., 2009), o que corrobora com os dados já apresentados para o Jasmonato, nos quais este apresentou alta expressão no clone 14 na condição NI e acúmulo em todos os clones.

Apesar de não conter quitina em seu organismo, a maioria das plantas sintetizam quitinases, dasquais destacam-se as quitinases de classe III de cafeeiro. A *CcCLXIP* é uma enzima que inibe em até 50% o efeito de xilanases na parede celular vegetal (VASCONCELOS et al., 2011), cuja expressão pode ser inibida tanto na proteção contra patógenos como por estresses abióticos, estando diretamente envolvida na tolerância das plantas à seca (WU et al., 2013; YU; GRIFFITH, 1999).

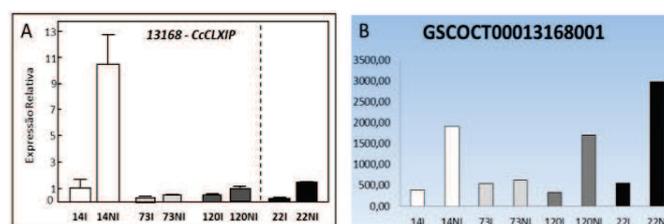


Figura 38 Perfis de expressão do transcrito *CcCLXIP* (GSCOCT00013168001) que codifica para a uma Quitinase inibidora de Xilanase

Legenda: Nas análises de qPCR (A) e *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

### Expressão dos transcritos que codificam a CPI3 – Proteína inibidora de cisteína

A expressão do gene que codifica uma proteína inibidora de cisteinase é apresentada na Figura 39. Para os dados de expressão relativa por qPCR, observa-se que a expressão do gene *CcCPI3* apresenta-se induzida na condição NI para todos os clones analisados, sendo que a maior expressão é apresentada no clone 14 na condição NI.

Esta proteína atua nas células vegetais desempenhando diversas funções: pode atuar na morte celular programada, no desenvolvimento dos tecidos, bem como na sinalização da planta em resposta a patógenos (LEPELLEY et al., 2012). Portanto, o fato de a maior expressão deste gene ser apresentada no clone 14 pode estar relacionado à resposta da planta para tolerância à seca, na defesa da planta a patógenos, uma vez que este clone é considerado resistente a patógenos, como por exemplo, ao nematoide de galhas (NG) *Meloidogyne* spp. (LIMA et al., 2013).

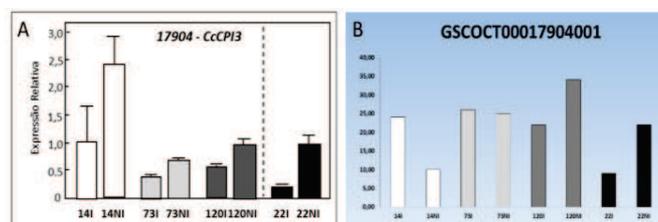


Figura 39 Perfis de expressão do transcrito *CcCPI3* (GSCOCT00017904001) que codifica para a uma proteína inibidora de cisteinase

Legenda: Nas análises de qPCR (A) e *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

### Expressão dos transcritos que codificam a BSPA – Proteína de armazenamento

A expressão do gene que codifica um tipo de proteína de armazenamento BSPA é apresentada na Figura 40. É possível observar que a expressão do gene *CcBSPA* é induzida na condição NI somente nos clones 14 e 22. Essas proteínas estão diretamente ligadas ao mecanismo de armazenamento de nutrientes nas plantas, direcionando recursos e nutrientes necessários para o desenvolvimento, e adaptação ao estresse abiótico (ATKINSON; BRENNAM; JONES, 2013; PETTENGILL; PETTENGILL; COLEMAN, 2013). Como mostrado na Figura 40, dentre os clones tolerantes, o clone 14 foi o que apresentou melhor resposta quanto à expressão deste gene. Supostamente a expressão deste gene é um tipo de resposta rápida da planta, uma vez que o clone 14 é o que apresentou sinal de estresse por seca.

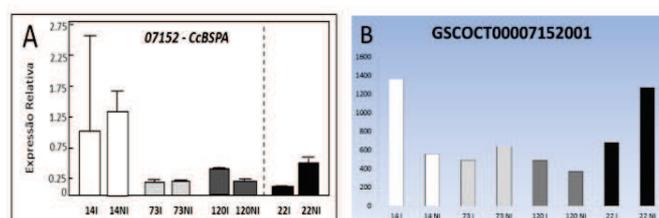


Figura 40 Perfis de expressão do transcrito *CcBSPA* (GSCOCT00007152001) que codifica para a uma proteína de armazenamento

Legenda: Nas análises de qPCR (A) e *in silico* (B), em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado). Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

### Expressão dos transcritos que codificam *No Hits*

Os resultados de expressão dos genes que codificam *No Hits* (GSCOCT00005517001, GSCOCT00015847001, GSCOCT00030679001 e

GSCOCT00042816001) são apresentados na Figura 41. Os dados de qPCR mostram que os quatro genes analisados apresentam indução de expressão na condição NI para todos os clones. Para os genes GSCOCT00005517001, GSCOCT00015847001 e GSCOCT00042816001, a maior expressão é apresentada no clone 22 (Figura 41A, Figura 41B e Figura 41D, respectivamente).

Os dados de expressão *in silico* mostram que, para o gene GSCOCT00015847001, a maior expressão é apresentada no clone 73 na condição I (Figura 41B), para os genes GSCOCT00005517001, GSCOCT00030679001 e GSCOCT00042816001, a maior expressão é apresentada no clone 22 na condição NI (Figuras 41A, 41C e 41D).

Ainda são poucos os relatos na literatura sobre os “genes órfãos”, que receberam esta nomenclatura por possuírem funções desconhecidas para os mecanismos fisiológicos. Este trabalho selecionou alguns desses genes, cujos valores de expressão apresentados destacaram-se na lista dos genes selecionados em raízes de cafeeiro. Apesar da indução destes genes na condição NI em todos os clones, a expressão destes genes é significativa para todos os clones, o que pode estar diretamente relacionado a um mecanismo natural da planta, que é induzida a uma maior expressão em situação de estresse hídrico para os clones de *C. canephora* estudados.

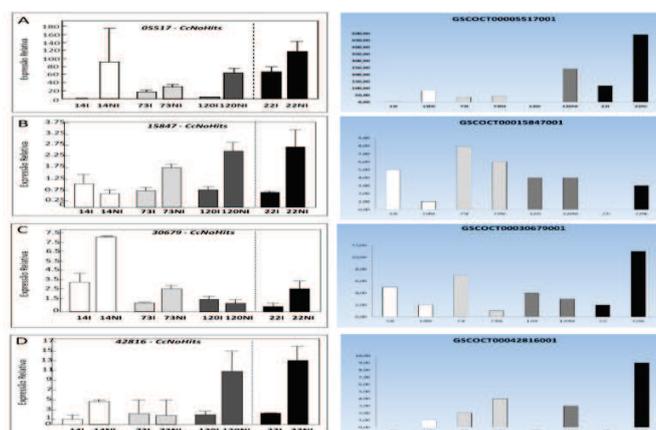


Figura 41 Perfis de expressão de transcritos que codificam *No Hits* (GSCOCT0005517001, GSCOCT00015847001, GSCOCT00030679001 e GSCOCT00042816001) para as análises de qPCR e *in silico* em raízes dos clones de *C. canephora* 14, 73 e 120 (tolerantes à seca) e 22 (sensível à seca) submetidos ao estresse hídrico (NI: não irrigado) e sem estresse (I: Irrigado)

Legenda: **A** perfil de expressão do *No Hit* GSCOCT00055177001. **B** perfis de expressão do *No Hit* GSCOCT00015847001. **C** perfis de expressão do *No Hit* GSCOCT00030679001. **D** perfis de expressão do *No Hit* GSCOCT00042816001. Para as reações de qPCR, a expressão relativa (ER) é avaliada (em unidades arbitrárias) em comparação à expressão do gene de referência (expressão constitutiva) *CcUBQ10*. O calibrador utilizado para normalização dos dados da qPCR foi o 14I (ER = 1).

#### 4 CONCLUSÕES

A técnica de pirosequenciamento 454 permitiu gerar 4.397.628 *reads* a partir de oito bibliotecas de raízes de *C. canephora* provenientes dos clones 14, 22, 73 e 120 nas condições irrigado (I) e não irrigado (NI). O que possibilitou a ampliação dos dados de sequências gênicas de raízes de *C. canephora*, tendo em vista que na base de dados do Projeto Genoma Café existiam somente 2.071 *reads*.

As análises realizadas permitiram a identificação de vários genes candidatos com expressão diferencial entre os clones analisados.

Entre os genes selecionados com maior valor de indução nos clones tolerantes (Q), grande parte se destaca quanto à expressão no clone 73. Exceto para os genes GSCCT00037166001 (Dehydrin DH1a) e GSCOCT00016554001 (Unknown protein), cujos valores de Q foram maiores no clone 120, GSCOCT00005624001 (8-hydroxyquercetin 8-O-methyltransferase), GSCOCT00005883001 (Putative Subtilisin-like protease SDD1) e GSCOCT00012078001 (Acidic endochitinase), cujos valores de Q foram maiores no clone 14.

O gene que mais se destaca quanto à expressão nos clones tolerantes é o gene GSCOCT00012078001 (Acidic endochitinase).

Dos genes analisados na biossíntese do ABA, o clone 14 apresentou melhor resposta na expressão dos genes *CcNCED3*, *CcPP2C-1* e *CcNAC*. Por outro lado, o clone 120 apresentou melhor resposta nos genes *CcAREB*, *CcNAC*, *CcABA-8-hidrolase*.

Quando se trata dos genes de resposta do cafeeiro para os mecanismos de proteção das células contra a desidratação, o clone 14 apresentou melhor resposta quanto à expressão nos genes *CcLEA* e *CcGols3* e o clone 73, para o gene *CcDH2a*.

O clone 14 apresentou melhor resposta quanto à expressão dos genes *Cc-S6PDH* e *CcGolS3* relacionados ao ajuste osmótico e dos genes *CcPRX1* e *CcGST*, relacionados ao estresse oxidativo.

Foram analisados também genes de resposta à deficiência de água que auxiliam na adaptação contra o estresse hídrico. O clone 14 apresentou melhor resposta quanto à expressão nos genes *CcMJE*, *CcSRP-ABA*, *CcMALD3*, *CcLTP2*, *CcPP3B*, *CcCLXIP*, *CcCPI3*; o clone 120, no gene *CcPhLOG* e o clone 22, no gene *CcNRT1.4*.

Já em relação à expressão dos *No Hits*, o clone 120 apresentou maior expressão para o *No Hits* relacionado ao *contig* GSCOCT00030679001.

## REFERÊNCIAS

- AGARWAL, P. K. et al. Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. **Plant Cell Reports**, Berlin, v. 25, n. 12, p. 1263-1274, July 2006.
- ALBERTS, B. et al. Control of gene expression. In: . **Molecular biology of the cell**. 4<sup>th</sup> ed. New York: Garland Science, 2002. p. 1531-1533.
- ANTONI, R. et al. Selective inhibition of clade a phosphatases type 2C by PYR/PYL/RCAR abscisic acid receptors. **Plant Physiology**, Bethesda, v. 158, n. 2, p. 970-980, Feb. 2012.
- ATIPAIRIN, A.; RATANAPHAN, A. *In Vitro* enhanced sensitivity to cisplatin in D67Y BRCA1 RING domain protein. **Breast Cancer: Basic and Clinical Research**, The Hague, v. 5, n. 25, p. 201-208, Sept. 2011.
- ATKINSON, C. J.; BRENNAM, R. M.; JONES, H. G. Declining chilling and its impact on temperature perennial crops. **Environmental and Experimental Botany**, Elmsford, v. 91, n. 1, p. 48-62, July 2013.
- BARSALOBRES-CAVALLARI, C. F. et al. Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. **BMC Molecular Biology**, London, v. 10, n. 1, p. 1-11, Jan. 2009.
- BERGER, N.; DUBREUCQ, B. Evolution goes GAGA: GAGA binding proteins across kingdoms. **Biochimica et Biophysica Acta**, Alberta, n. 1819, p. 863-868, 2012.
- BIANCHI, M. W.; ROUX, C.; VARTANIANA, N. Drought regulation of GST8, encoding the *Arabidopsis* homologue of ParC/Nt107 glutathione transferase/peroxidase. **Physiologia Plantarum**, Copenhagen, v. 116, n. 1, p. 96-105, Sept. 2002.

BOYD, J. et al. Sesquiterpene-like inhibitors of a 9-cis-epoxycarotenoid dioxygenase regulating abscisic acid biosynthesis in higher plants. **Bioorganic & Medicinal Chemistry**, New York, v.17, n. 7, p. 2902-2912, Apr. 2009.

BRESSAN, W.; VASCONCELOS, C. A. Alterações morfológicas no sistema radicular do milho induzidas por fungos micorrízicos e fósforo. **Pesquisa Agropecuária Brasileira**, Brasília, v. 37, n. 4, p. 509-517, abr. 2002.

CHAVES, M. M.; OLIVEIRA, M. M. Mechanism underlying plant resilience to water deficits: prospects for water-saving agriculture. **Journal of Experimental Botany**, Oxford, v. 55, n. 407, p. 2365-2384, Oct. 2004.

CUTLER, S. R. et al. Abscisic acid: emergence of a core signaling network. **Annual Review of Plant Biology**, Palo Alto, v. 61, p. 651-679, 2010.

DAMATTA, F. M.; RAMALHO, J. D. C. Impacts of drought and temperature stress on coffee physiology and production. **Brazilian Journal of Plant Physiology**, Piracicaba, v. 18, n. 1, p. 55-81, Jan./Mar. 2006.

DUNWELL, J. M. Transgenic approaches to crop improvement. **Journal of Experimental Botany**, Oxford, v. 51, n. 1, p. 487-496, Feb. 2000.

FERRÃO, R. G. et al. EMCAPA 8141-Robustão Capixaba, variedade clonal de café conilon tolerante à seca, desenvolvida para o estado do Espírito Santo. **Ceres**, Viçosa, MG, v. 47, n. 273, p. 555-560, set./out. 2000.

FONSECA, A. F. A. et al. Conilon Vitória-Incaper 8142: improved *Coffea canephora* var. Kouillou clone cultivar for the State Espírito Santo. **Crop Breeding and Applied Biotechnology**, Londrina, v. 4, n. 4, p. 1-3, Dec. 2004.

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

FUJII, H. et al. *In vitro* reconstitution of an abscisic acid signaling pathway. **Nature**, London, v. 462, n. 7723, p.660-664, Dec. 2009.

GILL, S. S. et al. Glutathione and glutathione reductase: a boon in disguise for plant abiotic stress defense operations. **Plant Physiology and Biochemistry**, New Delhi, v. 70, n. 1, p. 204-212, Sept. 2013.

HANIN, M. et al. Plant dehydrins and stress tolerance Versatile proteins for complex mechanisms. **Plant Signaling & Behavior**, Bonn, v. 6, n. 10, p. 1503-1509, 2011.

HAUSER, F.; WAADT, R.; SCHROEDER, J. I. Evolution of abscisic acid synthesis and signaling mechanisms. **Current Biology**, London, v. 21, n. 9, p. 346-355, May 2011.

HIRAGA, S. et al. A large family of class III plant peroxidases. **Plant Cell Physiology**, Oxford, v. 42, n. 5, p. 462-468, Feb. 2001.

IGAWA, T. et al. A wheat xylanase inhibitor gene, *Xip-1*, but not *Taxi-1*, is significantly induced by biotic and abiotic signals that trigger plant defense. **Bioscience, Biotechnology, and Biochemistry**, Tokyo, v. 69, n. 5, p. 1058-1063, May 2005.

JANG, C. S. et al. Expression and promoter analysis of the *TaLTP1* gene induced by drought and salt stress in wheat (*Triticum aestivum* L.). **Plant Science**, Shannon, v. 167, n. 5, p. 995-1001, Nov. 2004.

KHODAMBASHI, M.; SHIRAN, B.; GRARAGHANIPOUR, N. Differential expression of *CA7* and *NCED* Genes in common bean genotypes under drought stress. **Journal of Agricultural Science and Technology**, London, v. 15, p. 1491-1499, Dec. 2013. Supplement.

KIM, S. T. et al. Oligomerization and chaperone activity of a plant 2-Cys peroxiredoxin in response to oxidative stress. **Plant Science**, Shannon, v. 177, n. 3, p. 227-232, Sept. 2009.

KUSHIRO, T. et al. The *Arabidopsis* cytochrome P450 CYP707A encodes ABA 80-hydroxylases: key enzymes in ABA catabolism. **European Molecular Biology Organization**, Heidelberg, v. 23, n. 7, p. 1647-1656, Mar. 2004.

LACKMAN, P. et al. Jasmonate signaling involves the abscisic acid receptor PYL4 to regulate metabolic reprogramming in *Arabidopsis* and tobacco. **Plant Biology**, Stuttgart, v. 108, n. 14, p. 5891-5896, May 2011.

LEPELLEY, M. et al. Coffee cysteine proteinases and related inhibitors with high expression during grain maturation and germination. **BMC Plant Biology**, London, v. 12, n. 31, p. 1-16, Mar. 2012.

LEROY, T. et al. Genetics of coffee quality. **Brazilian Journal of Plant Physiology**, Piracicaba, v. 18, n. 1, p. 229-242, Jan./Mar. 2006.

LEUNG, J. et al. Mécanisme moléculaire d'action de l'acide abscissique en réponse à la sécheresse chez les végétaux: the ABC of abscisic acid action in plant drought stress responses. **Biologie Aujourd'hui**, Les Ulis, v. 206, n. 5, p. 301-312, Feb. 2012.

LI, T. H.; LI, S. H. Enzymatic regulation of sorbitol metabolism in micropropagated apple plants in response to water stress. **European Journal of Horticultural Science**, Stuttgart, v. 72, n. 1, p. 12-19, Feb. 2007.

LI, Y. et al. Stimulation of root acid phosphatase by phosphorus deficiency is regulated by ethylene in *Medicago falcata*. **Environmental and Experimental Botany**, Elmsford, v. 71, n. 1, p. 114-120, Apr. 2011.

LIMA, E. A. de et al. Resistência múltipla a resposta de hipersensibilidade do cafeeiro "Conilon 14" a *Meloidogyne* spp. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 8., 2013, Salvador. **Anais...** Salvador: UFBA, 2013. 1 CD-ROM.

LUCHI, S. et al. Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in *Arabidopsis*. **The Plant Journal**, Oxford, v. 27, n. 4, p. 325-333, Mar. 2001.

MALONE, G. et al. Expressão diferencial de isoenzimas durante o processo de germinação de sementes de arroz em grandes profundidades de semeadura. **Revista Brasileira de Sementes**, Brasília, v. 29, n. 1, p. 61-67, abr. 2007.

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

MARRACCINI, P. et al. *RBCS1* expression in coffee: *Coffea* orthologs, *Coffea arabica* homeologs, and expression variability between genotypes and under drought stress. **BMC Plant Biology**, London, v. 11, n. 85, p. 1-23, May 2011.

MEDA, A. R. et al. Análise *in silico* de genes envolvidos no transporte de nitrato, amônio e uréia em cafeeiro. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 6., 2009, Vitória. **Resumos Expandidos...** Vitória: EMBRAPA Café, 2009. p. 1-7.

MENG, Y. et al. The contribution of abscisic acid to sorbitol accumulation in drought-stressed *Malus hupehensis*. **Journal of Food, Agriculture & Environment**, Helsinki, v. 6, n. 2, p. 319-326, Apr. 2008.

MILLER, G. et al. Nitrate transport and signalling. **Journal of Experimental Botany**, Oxford, v. 58, n. 9, p. 2297-306, May 2007.

MILLER, G. et al. Reactive oxygen species homeostasis and signalling during drought and salinity stresses. **Plant Cell & Environment**, Oxford, v. 33, n. 4, p. 453-467, Apr. 2010.

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

MONTOYA, G. et al. Sequence analysis from leaves, flowers and fruits of *Coffea arabica* var Caturra. In: INTERNATIONAL CONFERENCE ON COFFEE SCIENCE, 21., 2006, Montpellier. **Proceedings...** Montpellier: ASIC, 2006. 1 CD-ROM.

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

MOWLA, S. B. et al. A novel stress-inducible antioxidant enzyme identified from the resurrection plant *Xerophyta viscosa* Baker. **Planta**, Berlin, v. 215, n. 5, p. 716-726, Sept. 2002.

PARSELL, D. A.; LINDQUIST, S. The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. **Annual Review of Genetics**, Palo Alto, v. 27, p. 437-496, 1993.

PARTELLI, F. L. et al. Produção e desenvolvimento radicular de plantas de café “Conilon” propagadas por sementes e por estacas. **Pesquisa Agropecuária Brasileira**, Brasília, v. 41, n. 6, p. 949-954, jun. 2006.

PETTENGILL, E. A.; PETTENGILL, J. B.; COLEMAN, G. Elucidating the evolutionary patterns of nucleoside phosphorylase paralogs (vegetative storage proteins) in *Populus* and the plant kingdom. **BMC Plant Biology**, London, v. 13, n. 118, p. 1-35, Aug. 2013.

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, July 2005.

PINO-NUNES, L. E. **Controle do desenvolvimento vegetal pela interação auxina-citocinina: uma nova abordagem baseada no estudo de mutantes de tomateiro (*Solanum Lycopersicum* cv Micro-Tom)**. 2009. 140 p. Tese (Doutorado em Ciências) - Universidade de São Paulo, São Paulo, 2009.

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

RAMAKERS, C. et al. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. **Neuroscience Letters**, Limerich, v. 339, n. 1, p. 62-66, Mar. 2003.

RÉ, D. A. et al. *Nicotiana attenuata NaHD20* plays a role in leaf ABA accumulation during water stress, benzylacetone emission from flowers, and the timing of bolting and flower transitions. **Journal of Experimental Botany Advance**, Oxford, v. 16, n. 1, p. 155-166, Aug. 2010.

RIBAS, A. F.; PEREIRA, L. F. P.; VIEIRA, L. G. Genetic transformation of coffee. **Brazilian Journal of Plant Physiology**, Piracicaba, v. 18, n. 1, p. 83-94, Jan./Mar. 2006.

SÁNCHEZ-RODRÍGUEZ, E.; ROMERO, L.; RUIZ, J. M. Role of grafting in resistance to water stress in tomato plants: ammonia production and assimilation. **Journal of Plant Growth Regulation**, Dordrecht, v. 32, n. 4, p. 831-842, Dec. 2013.

SANTOS, T. B. et al. Expression of three galactinol synthase isoforms in *Coffea arabica* L. and accumulation of raffinose and stachyose in response to abiotic stresses. **Plant Physiology and Biochemistry**, New Delhi, v. 49, n. 4, p. 441-448, Apr. 2011.

SEKI, M. et al. Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. **Plant Cell**, Rockville, v. 13, n. 1, p. 61-72, Jan. 2001.

SHARP, R. E. et al. Root growth maintenance during water deficits: physiology to functional genomics. **Journal of Experimental Botany**, Oxford, v. 55, n. 407, p. 2343-2351, Sept. 2004.

SHEEHAN, D. et al. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of ancient enzyme superfamily. **Biochemical Journal**, London, v. 360, n. 1, p. 1-16, Nov. 2001.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Gene networks involved in drought stress response and tolerance. **Journal of Experimental Botany**, Oxford, v. 58, n. 2, p. 221-227, Oct. 2007.

SIMKIN, A. J. et al. An investigation of carotenoid biosynthesis in *Coffea canephora* and *Coffea arabica*. **Journal of Plant Physiology**, Stuttgart, v. 165, n. 10, p. 1087-1106, July 2008.

SMIRNOFF, N. Plant resistance to environmental stress. **Current Opinion in Biotechnology**, London, v. 9, n. 2, p. 214-219, Apr. 1998.

SOARES, A. M. dos; MACHADO, O. L. T. Defesa de plantas: sinalização química e espécies reativas de oxigênio. **Revista Trópica - Ciências Agrárias e Biológicas**, Chapadinha, v. 1, n. 1, p. 9-19, 2007.

SOARES, B. M. et al. Produção e distribuição de raízes em cafeeiros irrigados por gotejamento. **Engenharia na Agricultura**, Viçosa, MG, v. 15, n. 3, p. 130-140, abr./jun. 2007.

SOON, F. F. et al. Molecular mimicry regulates ABA signaling by SnRK2 kinases and PP2C phosphatases. **Science**, New York, v. 335, n. 6064, p. 85-88, Jan. 2012.

TAJI, T. et al. Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. **The Plant Journal**, Oxford, v. 29, n. 4, p. 417-426, Feb. 2002.

TAPIA, G. et al. Study of nsLTPs in *Lotus japonicus* genome reveal a specific epidermal cell member (LjLTP10) regulated by drought stress in aerial organs with a putative role in cutin formation. **Plant Molecular Biology**, Dordrecht, v. 82, n. 4/5, p. 485-501, July 2013.

THOMPSON, A. J. et al. Regulation and manipulation of ABA biosynthesis in roots. **Plant, Cell and Environment**, Oxford, v. 30, n. 1, p. 67-78, Jan. 2007.

TRAN, L. P. et al. Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought-responsive *cis*-element in the early responsive to dehydration stress 1 Promoter. **The Plant Cell**, Rockville, v. 16, n. 2, p. 2481-2498, Sept. 2004.

UMEZAWA, T. et al. Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. **Current Opinion in Biotechnology**, London, v. 17, n. 2, p. 113-122, Apr. 2006.

VASCONCELOS, E. A. R. et al. A new chitinase-like xylanase inhibitor protein (XIP) from coffee (*Coffea arabica*) affects Soybean Asian rust (*Phakopsora pachyrhizi*) spore germination. **BMC Biotechnology**, London, v. 11, n. 14, p. 1-8, Feb. 2011.

VIDAL; GUTIÉRREZ. A systems view of nitrogen nutrient and metabolite responses in Arabidopsis. **Current Opinion in Plant Biology**, Oxford, v. 11, n. 5, p. 521-529, Oct. 2008.

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

VIEIRA, N. G. et al. Different molecular mechanisms account for drought tolerance in *Coffea canephora* var. Conilon. **Tropical Plant Biology**, New York, v. 6, n. 4, p. 181-190, Dec. 2013.

VINECKY, F.; SILVA, F. R. da; ANDRADE, A. C. Análise *in silico* das bibliotecas de cDNA SH2 e SH3 para a identificação de genes responsivos à seca em cafeeiro. **Coffee Science**, Lavras, v. 7, n. 1, p. 1-19, jan./abr. 2012.

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, Nov. 2003.

WEI, H. et al. Natural paradigms of plant cell wall degradation. **Theoretical and Applied Genetics**, Berlin, v. 20, n. 3, p. 330-338, June 2009.

WOOD, Z. A. et al. Structure, mechanism and regulation of peroxiredoxins. **Trends in Biochemical Sciences**, London, v. 28, n. 1, p. 32-40, Jan. 2003.

WU, J. et al. Characterization of a newly identified rice chitinase-like protein (OsCLP) homologous to xylanase inhibitor. **BMC Biotechnology**, London, v. 51, n. 13, p.29-40, Feb. 2013.

XIA, N. et al. *TaNAC8*, a novel NAC transcription factor gene in wheat, responds to stripe rust pathogen infection and abiotic stresses. **Physiological and Molecular Plant Pathology**, London, v. 74, n. 5/6, p. 394-402, Sept. 2010.

YU, X. M.; GRIFFITH, M. Antifreeze proteins in winter rye leaves form oligomeric complexes. **Plant Physiology**, Bethesda, v. 119, n. 4, p. 1361-1369, Apr. 1999.

ZHANG, X. et al. Biosolids impact on tall fescue drought resistance. **Journal of Residuals Science & Technology**, Denton, v. 2, n. 3, p. 173-180, Dec. 2005.

### **CAPÍTULO 3 Caracterização de proteínas diferencialmente expressas em raízes de clones de *Coffea canephora* submetidos à deficiência hídrica**

#### **RESUMO**

A diminuição da produtividade cafeeira dos últimos anos, causada principalmente pelas secas severas, têm preocupado os cafeicultores. O que têm os levado a buscarem novas estratégias para o aumento da produção, como a seleção de genótipos que toleram melhor aos períodos de seca. Para a identificação destes genótipos, estudos de caracterização do perfil proteômico vêm sendo utilizados aliados aos estudos genômicos para possibilitar melhor entendimento dos mecanismos de resposta à seca nas plantas. Avanços nas análises proteômicas permitiram o desenvolvimento da cromatografia líquida acoplada à espectrometria de massa – LC-MS, uma poderosa ferramenta para as análises de teores de proteínas e modificações de expressão. Trabalhos que buscam avaliar os perfis de expressão proteica em *Coffea* sp. por meio desta técnica têm sido descritos. O objetivo do presente trabalho foi caracterizar o perfil proteômico em raízes de clones de *C. canephora*, sensíveis e tolerantes, diferencialmente expressos quanto ao déficit hídrico nas condições de irrigação (I) e não irrigação (NI) por espectrometria de massa. Foram utilizadas raízes de clones de *C. canephora* var Conilon tolerantes (14, 73 e 120) e sensível (22) ao estresse hídrico nas condições anteriormente citadas. As proteínas totais foram obtidas por dois protocolos distintos: o de Solução Ácida, precipitada por sulfato de amônio e o protocolo Sussulini et al. (2007), precipitado com acetona. Após a obtenção do extrato proteico, as amostras foram submetidas a uma etapa de digestão enzimática com tripsina, fracionadas e analisadas no LC-MS, utilizando a cromatografia líquida de fase reversa acoplada à espectrometria de massa. Os resultados de identificação dos íons tripsinizados foram obtidos com o auxílio do software “Protein Lynx” utilizando o Banco de dados de *C. canephora*. A técnica LC-MS permitiu a identificação de 598 proteínas, o que representa cerca de 1% do proteoma conhecido de *C. canephora*. Para algumas proteínas, foi possível identificar diferentes formas alélicas.

Palavras-chave: Cafeeiro. Raiz. Tolerância à seca. Proteômica. Espectrometria.

**Characterization of proteins differentially expressed in *Coffea canephora* clone roots submitted to water deficiency**

**ABSTRACT**

The decrease in coffee production in recent years, caused mainly by severe drought, has worried coffee producers. This has led to seeking new strategies to increase to population, such as the selection of genotypes that better tolerate the drought periods. To identify these genotypes, studies on proteomic profile characterization allied to genomic studies have been used to allow the better understanding of the response mechanisms to drought in plants. Advances in proteomic analyses allowed the development of liquid chromatography coupled to mass spectrometry - LC-MS, a powerful tool to analyze protein contents and expression modification. Works seeking to evaluate protein expression profiles in *Coffea* sp. by means of this technique have been described. The objective of the present work was to characterize the proteomic profile in sensitive and tolerant *C. canephora* clone roots, differentially expressed regarding water deficit in irrigated (I) and non-irrigated (NI) conditions, by means of mass spectrometry. *C. canephora* clone roots of the Comilon variety, tolerant (14, 73 and 120) and sensitive (22) to water deficit, were used in the previously cited conditions. The total proteins were obtained by two distinct protocols: Acid Solution, precipitated by ammonium sulfate and the Sussulini et al. (2007) protocol, precipitated with acetone. After obtaining the protein extract, the samples were submitted to an enzyme digestion stage with trypsin, fractionated and analyzed on the LC-MS, using reverse phase liquid chromatography coupled with mass spectrometry. The results of the identification of the trypsinized ions were obtained with the aid of the ProteinLynx software using the *C. canephora* databank. The LC-MS technique allowed the identification of 598 proteins, which represents around 1% of the proteome known for *C. canephora*. For some of the proteins, it was possible to identify different allele forms.

Keywords: Coffee plant. Root. Drought tolerance. Proteomic. Spectrometry.

## 1 INTRODUÇÃO

O estudo proteômico de uma célula ou tecido específico possibilita identificar proteínas, bem como caracterizar modificações pós-traducionais ou interações proteína-proteína (WIJK, 2001), sendo estes dados cruciais para a melhoria das anotações genômicas, uma vez que possibilitam a validação de ESTs (*Expressed Sequence Tags*) por meio da comparação entre proteína expressa e o transcrito do gene correspondente (VIEIRA et al., 2006).

Uma maneira de identificação proteica de uma célula ou tecido é separar as proteínas presentes num extrato proteico e, em seguida, caracterizá-las isoladamente por espectrometria de massa, que combinada com análise computacional de imagens e bases de consulta se torna uma poderosa ferramenta para identificação, quantificação e caracterização de proteínas (WILDGRUBER et al., 2000).

Para minimizar as limitações sofridas pelo gel bidimensional 2-DE, foi desenvolvida uma nova técnica de análise proteômica, a cromatografia líquida acoplada à espectrometria de massa - LC-MS (MURAD; RECH, 2012). Esta alia a técnica de separação de proteínas, com base nas características físico-químicas, com a técnica de identificação por meio da avaliação da massa molecular dos peptídeos (LANÇAS, 2009; YERGEY et al., 2002).

Trabalhos que buscam avaliar os perfis de expressão proteica em *Coffea* sp. já descritos em folhas e raízes nos clones 109 e 120 (GUIMARÃES, 2007) e em folhas nos clones 14, 22, 73 e 120 (MARRACCINI et al., 2012) de *C. canephora* sob déficit hídrico, assim como em frutos de *C. arábica* (HEIMBECK, 2011).

Desta forma, o objetivo do presente trabalho foi caracterizar o perfil proteômico em raízes de clones de *C. canephora*, sensíveis e tolerantes, diferencialmente expressos quanto ao déficit hídrico nas condições de irrigação

(I) e não irrigação (NI) por espectrometria de massa (LC-MS – *Liquid Chromatography-Mass-Spectrometry*).

## 2 MATERIAL E MÉTODOS

### 2.1 Material vegetal

Para análises de proteômica, foram utilizadas raízes de clones de *C. canephora* var Conilon tolerantes (14, 73 e 120) e sensível (22) ao estresse hídrico. Estes clones foram analisados nas condições de irrigação (I) e sem irrigação (NI). As proteínas totais foram obtidas por dois protocolos distintos, com o objetivo de alcançar o maior número possível de proteínas de raízes de cafeeiro.

Os protocolos utilizados neste trabalho foram o de Solução Ácida, precipitada por sulfato de amônio e o protocolo Sussulini et al. (2007), precipitado com acetona. Após a obtenção do extrato proteico, as amostras foram submetidas a uma etapa de digestão enzimática com tripsina e, sequencialmente, fracionadas e analisadas no LC-MS.

#### **Protocolo de Solução Ácida, com precipitação em sulfato de amônio – adaptado**

Foi utilizado 2g de material vegetal, pulverizado em nitrogênio líquido. Adicionou-se 50 mL de tampão de extração ácido (5 $\mu$ L de HCl, 1 $\mu$ L de NaCl e 44  $\mu$ L de água deionizada). A solução permaneceu sob agitação em câmara fria (4°C) por 12 horas. O material foi filtrado e centrifugado (20min, 4°C, 10.000 x g). O *pellet* foi descartado e o sobrenadante teve seu pH ajustado para 7,0 com NaOH. Em seguida, a solução permaneceu em repouso, no gelo, por uma hora. A solução foi centrifugada nas condições descritas anteriormente e o *pellet* foi descartado novamente. Foi adicionado sulfato de amônio 60% e a solução permaneceu na câmara fria por 12 horas, sob agitação, para que houvesse a

precipitação das proteínas. Em seguida, a solução foi centrifugada nas condições acima descritas, o *pellet* foi ressuspenso em 10 mL de água deionizada e dialisado por dois dias. A amostra foi liofilizada e ressuspenso em 50 mM de bicarbonato de amônio a uma concentração de proteína de  $1 \mu\text{g} \cdot \mu\text{L}^{-1}$ .

Em seguida a amostra foi digerida com tripsina, e, posteriormente, preparada para a análise de espectrometria (LC-MS) seguindo o protocolo sugerido pelo fabricante (NanoAcquity, UPLC-Waters, USA). Para iniciar a digestão, foi utilizado  $50 \mu\text{L}$  de extrato proteico na concentração de  $1 \mu\text{g} / \mu\text{L}$ , no qual foi adicionado  $10 \mu\text{L}$  de bicarbonato de amônio ( $\text{NH}_4\text{HCO}_3$ ) na concentração de 50mM e  $25 \mu\text{L}$  de solução RAPIGEST SF 2%. Em seguida, a solução foi aquecida a  $80^\circ\text{C}$  durante 15 minutos. Logo após, foram adicionados  $2,5 \mu\text{L}$  de ditioneitol (DTT) na concentração de 100mM e a solução foi aquecida novamente a  $60^\circ\text{C}$  durante 30 minutos. Após a amostra ser resfriada à temperatura ambiente,  $2,5 \mu\text{L}$  de iodoacetamida (IAA) na concentração de 300mM foram adicionados e a solução permaneceu por 30 minutos no escuro.  $10 \mu\text{L}$  da solução de tripsina (promega) e bicarbonato de amônio 50mM foram adicionados. Para facilitar a digestão enzimática, a solução foi submetida à temperatura de  $37^\circ\text{C}$  por 16 horas. Em seguida, foi adicionado  $10 \mu\text{L}$  de ácido trifluoroacético (TFA) na concentração de 5% e a solução foi a  $37^\circ\text{C}$  por 90 minutos. A solução foi centrifugada (30min,  $4^\circ\text{C}$ ,  $10.000 \times g$ ) e o sobrenadante foi transferido para outro tubo onde foi adicionado  $10 \mu\text{L}$  do calibrante interno (de álcool desidrogenase (ADH) na concentração de  $1 \text{pmol} / \mu\text{L}$ . Finalizando esta etapa,  $85 \mu\text{L}$  da solução contendo acetonitrila (ACN) 3% e ácido fórmico (FA) 0,1% foram adicionados. E  $25 \text{fmol} / \mu\text{L}$  de ADH foi inserido, para ser utilizado como calibrador interno para a amostra. A concentração final do ADH foi de  $25 \text{fmol} / \mu\text{L}$ .

### **Protocolo Sussulini et al. (2007) (éter de petróleo) – adaptado**

A extração de proteínas totais a partir de raízes de café iniciou-se com a pulverização de 100mg de cada amostra refrigerada em nitrogênio líquido. As amostras foram pesadas e colocadas em tubos de centrifuga de 2ml com tampa. 1ml de éter de petróleo foi adicionado, e as amostras foram suavemente agitadas durante 15 minutos. O sobrenadante foi descartado e este passo foi repetido duas vezes. O éter de petróleo foi retirado totalmente e 1 mL de solução de extração contendo Tris-HCl 20 mM, pH 8,3, 1,5 mM de KCl, 10 mM de DTT, 1 mM, PMSF e 0,1% de SDS v/v foi adicionado. As amostras foram lentamente agitadas à temperatura ambiente durante 10min e centrifugadas (5min, 4°C, 10.000 x g). O sobrenadante foi então transferido para um novo tubo. Para cada 200 µL de amostra, 800 µL de acetona fria foi adicionado. As amostras foram incubadas à temperatura de -20°C durante 1h, com agitação em vórtex a cada 15 minutos. As amostras foram então centrifugadas (10min, 4°C, 15.700 x g). O sobrenadante foi descartado, e o pellet, depois de seco, foi cuidadosamente dissolvido em 500µL de bicarbonato de amônio 50 mM e quantificado utilizando o Nano Vue Plus (GE Healthcare, EUA). A amostra foi diluída com 50 mM de bicarbonato de amônio a uma concentração proteica em 1 µg.µL<sup>-1</sup>.

Em seguida, a amostra foi digerida com tripsina, e posteriormente, preparada para a análise de espectrometria (LC-MS), seguindo o protocolo sugerido pelo fabricante (NanoAcquity, UPLC-Waters, USA) conforme descrição acima.

### **2.2 Fracionamento das proteínas totais**

As frações referentes aos clones nas duas condições foram individualmente separadas por cromatografia líquida de fase reversa acoplada à

espectrometria de massa (LC-MS), usando uma coluna analítica capilar modelo Acquity UPLC BEH C18 (Waters, Milford, USA) com partículas de 1,7  $\mu\text{m}$ . (NanoAcquity, UPLC-Waters, USA) com capilaridade de 17 $\mu\text{m}$ . A fase móvel A constituiu-se de água + ácido fórmico 0,1% e a fase móvel B, de acetonitrila + ácido fórmico 0,1%. Para tais análises, foi empregado o gradiente de eluição: 2min – 3% B; 2-175min, 3-50% B; 175-180min – 90-95% B; 180-190min – 95% B e 190-195min 3% B. A etapa de lavagem da amostra com o solvente A (99%) a 7 $\mu\text{L}/\text{min}$  ocorreu por dois minutos. O fluxo da corrida foi de 0,6  $\mu\text{L}/\text{min}$ . Para cada análise, foram efetuadas três replicatas técnicas.

### 2.3 Espectrometria de massa

As análises foram conduzidas em um espectrômetro de massas de alta resolução, modelo Synapt G1 (Waters, Manchester, Reino Unido) operando em modo W. O mesmo é composto por um arranjo híbrido de dois analisadores de massa, um quadrupolo (Q) e um tempo de voo (TOF-*Time-of-flight*), respectivamente. A fonte de *electrospray* foi operada no modo positivo, com fluxo de gás ( $\text{N}_2$ ) constante a 0,25  $\text{L}\cdot\text{min}^{-1}$  e temperatura ambiente de nebulização entre 25 e 28°C. As tensões utilizadas no capilar e cone de amostragem foram de 3.0 kV 40V, respectivamente. A faixa de  $m/z$  analisada foi de 50-1990 Da. As amostras foram introduzidas diretamente na fonte ESI por meio do cromatógrafo líquido UPLC NanoAcquity (Waters, Milford, USA) a um fluxo constante de 0.6 $\mu\text{L}/\text{min}$ . Neste equipamento, os experimentos MS foram realizados por meio da seleção do sinal da intensidade do íon de interesse operando no modo  $\text{MS}^E$ . O íon de interesse é dissociado na câmara de colisão (q2), após colidir com o gás argônio (*Collision Induced Dissociation* – CID). Os íons produzidos neste processo foram, em seguida, separados no analisador de massa TOF. Neste método de análise, o espectrômetro de massa é

constantemente calibrado com um fluxo constante de 25 $\mu$ L/min de uma solução de Fibrinopeptídeo B humano [Glu<sup>1</sup>] a 100 fmol/ $\mu$ L (Sigma-Aldrich, F3261, No. CAS 103213-49-6) em 50:50 metanol/água e 0,1% ácido fórmico (BANTSCHIEFF et al., 2007).

#### **2.4 Identificação das proteínas *in silico***

Os dados brutos obtidos, resultantes dos experimentos de LC-MS, foram processados utilizando o programa ProteinLynx Global Server (versão 3.0) da Water/Micromass, o qual já possui um aplicativo para identificação e quantificação das proteínas.

Os resultados de quantificação foram normalizados com a Ubiquitina. A normalização de dados é uma etapa fundamental para controlar artefatos e erros que podem ocorrer na preparação das amostras, nas etapas do experimento e na análise dos dados (GONZÁLEZ-AGÜERO et al., 2013). Diante disso, várias estratégias são aplicadas para a normalização dessas variações, como o uso de genes de controle interno, chamados também de *housekeeping* genes, que podem manter o nível de expressão independente da célula, tecido ou tratamento experimental (VANDESOMPELE et al., 2002). Dentre esses genes, um comumente utilizado como gene de controle interno para o café é a ubiquitina, como no trabalho com *C. arabica* de Barsalobres-Cavallari et al. (2009).

### 3 RESULTADOS E DISCUSSÃO

Por meio das análises acima citadas, foram obtidos espectros com picos de ótima intensidade (ver anexo), os quais representam a massa molecular dos íons obtidos. Com a comparação entre os resultados dos produtos da fragmentação trípica das amostras analisadas e os resultados da fragmentação teórica das proteínas do banco de dados utilizado, foi possível identificar por similaridade entre a massa dos íons, 598 proteínas diferenciais em raízes de clones de *C. canephora*. Destas houve um diferencial de proteínas obtidas para cada clone nas duas condições estudadas, como ilustrado na Tabela 1.

Os extratos proteicos obtidos pelo protocolo de éter de petróleo, resultaram na identificação de um maior quantitativo de proteínas no clone 120 na condição de seca (NI) e a menor quantidade de proteínas foi observada no clone 22 na condição de seca (NI). O quantitativo de proteínas para a condição NI foi maior para todos os clones, com exceção do 120NI. Para o protocolo de solução ácida, o maior número de proteínas foi também observado no clone 120, mas na condição de irrigação (I) e o menor quantitativo ocorreu para o mesmo clone, na condição NI. No total de proteínas encontradas, para os dois protocolos, o clone 120 na condição irrigada obteve o maior número de proteínas, enquanto o 22NI apresentou o menor número (Tabela 1).

Tabela 1 Quantitativo de proteínas identificadas utilizando a análise LC-MS

Clone	Éter de petróleo (N° de PTN)			Solução ácida (SA) (N° de PTN)			Total
	Exclusivas (E)	Comuns (C)	Total	Exclusivas (E)	Comuns (C)	Total	
<b>14I</b>	81	76	157	14	33	47	204
<b>14NI</b>	76	76	152	19	33	52	204
<b>22I</b>	116	91	207	91	57	148	355
<b>22NI</b>	40	91	131	17	57	74	205
<b>73I</b>	98	119	217	39	55	94	311
<b>73NI</b>	37	119	156	39	55	94	250
<b>120I</b>	53	89	142	241	29	270	411
<b>120NI</b>	138	89	227	9	29	38	266

ET - protocolo de éter de petróleo, SA- protocolo de solução ácida. N° de PTN – Número de proteínas identificadas. E – proteínas exclusivas do clone. C – Proteínas comuns ao clone.

No clone 14, a maior quantidade de proteínas foi identificada nos resultados do extrato proteico obtido pelo protocolo de éter de petróleo (Figuras 1A e 1B). O quantitativo de proteínas nas condições I e NI é similar, apresentando uma alta proporção de proteínas comuns (Figuras 1C).

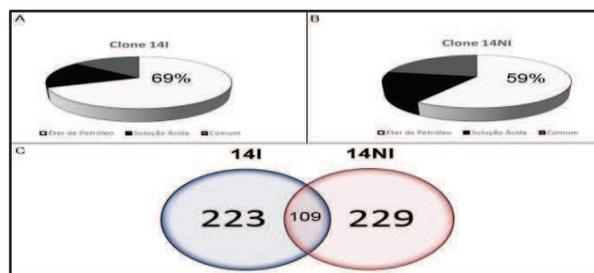


Figura 1 Proporção de proteínas obtidas no clone 14 identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA)

Legenda: **A.** Diferencial de proteínas em cada protocolo na condição I. **B.** Diferencial de proteínas em cada protocolo na condição NI. **C.** Total de proteínas identificadas para o clone 14, mostrando o quantitativo de proteínas comuns ao clone, independente da condição.

No clone 73, a amostra de extrato proteico obtido pelo protocolo de éter de petróleo identificou a maior proporção de proteínas para as duas condições (Figuras 2A e 2B), para este clone o quantitativo de proteínas é também parecido, para as duas condições (Figura 2C).

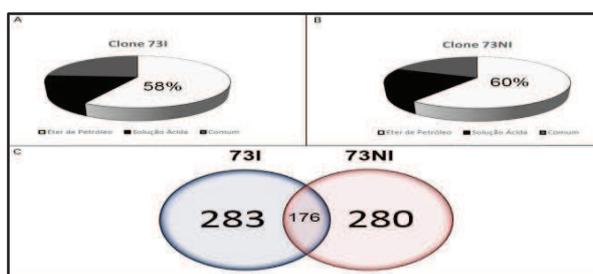


Figura 2 Proporção de proteínas obtidas no clone 73 identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA)

Legenda: **A.** Diferencial de proteínas em cada protocolo na condição I. **B.** Diferencial de proteínas em cada protocolo na condição NI. **C.** Gráfico ilustrativo do total de proteínas identificadas para o clone 73, mostrando o quantitativo de proteínas comuns ao clone, independente da condição.

Para o Clone 120, a proporção de proteínas variou de acordo com a condição. Para a condição I, o maior número de proteínas identificadas foi observado nos resultados obtidos pelo protocolo de solução ácida (Figura 3A), enquanto para a condição NI, a proporção de proteínas identificadas foi maior com os resultados do protocolo de éter de petróleo (Figura 3B). No total, o maior número de proteínas identificadas foi observado na condição I (Figura 3C).



Figura 3 Proporção de proteínas obtidas no clone 120 identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA)

Legenda: **A.** Diferencial de proteínas em cada protocolo na condição I. **B.** Diferencial de proteínas em cada protocolo na condição NI. **C.** Gráfico ilustrativo do total de proteínas identificadas para o clone 120, mostrando o quantitativo de proteínas comuns ao clone, independente da condição.

Para o clone 22, sensível ao estresse hídrico, a proporção de proteínas identificadas foi maior no protocolo de éter de petróleo para as duas condições (Figuras 4A e 4B). No total, o valor de proteínas identificadas para o clone na condição controle (I) foi maior (Figura 4A).

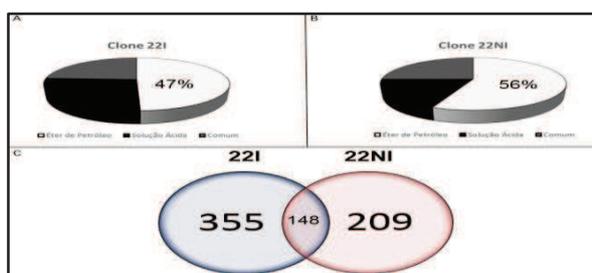


Figura 4 Proporção de proteínas obtidas no clone 22 identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA)

Legenda: **A.** Diferencial de proteínas em cada protocolo na condição I. **B.** Diferencial de proteínas em cada protocolo na condição NI. **C.** Gráfico ilustrativo do total de proteínas identificadas para o clone 22, mostrando o quantitativo de proteínas comuns ao clone, independente da condição.

Os dados obtidos de quantificação para as proteínas foram calculados a partir da normalização da quantificação da ubiquitina. Dentre as dez proteínas com maiores valores, algumas delas são encontradas nos dois protocolos, sendo que as mais significativas são as proteínas de armazenamento (*Contig* GSCOCT00007152001), quitinases (*Contigs* GSCOCT00013166001 e GSCOCT00013168001) e fosfatases ácidas (*Contigs* GSCOCT00022022001 e GSCOCT00022023001) (Tabelas 2 e 3).

Para o protocolo de Solução Ácida, o extrato proteico obtido para cada clone possibilitou a identificação de um tipo de proteína de armazenamento, quatro tipos de quitinases, dois tipos de fosfatases, dois tipos de apirases e uma miraculina (Tabela 2).

Tabela 2 Dez proteínas com maiores valores de quantificação, identificadas por LC-MS, utilizando o protocolo de solução ácida

<i>Contig</i>	Proteína	14I	14NI	73I	73NI	120I	120NI	22I	22NI
<b>07152001</b>	Proteína de Armazenamento	3.571	2.881	8.365	6.231	851	445	994	5.099
<b>13166001</b>	Quitinase	1.939	712	5.638	424	667	298	635	5.019
<b>13168001</b>	Quitinase	605	596	3.234	2.805	246	284	483	1.837
<b>22022001</b>	Fosfatase Ácida	1.667	731	2.392	199	682	129	882	434
<b>13173001</b>	Quitinase	251	192	2.195	1.958	120	352	596	823
<b>22023001</b>	Fosfatase Ácida	1.687	2.121	1.590	660	651	179	816	881
<b>17380001</b>	Apirase	233	200	454	833	241	153	334	401
<b>13169001</b>	Quitinase	337	213	548	303	106	108	155	314
<b>17376001</b>	Apirase	195	214	173	144	58	0	187	398
<b>07300001</b>	Miraculina	302	268	0	0	49	235	0	115

I - condição de irrigação. NI - condição de déficit hídrico. *Contigs* sem o prefixo GSCOCT000.

Para o protocolo de Éter de Petróleo, o extrato proteico obtido para cada clone possibilitou a identificação de um tipo de proteína de armazenamento, dois tipos de quitinases, três tipos de fosfatases, uma proteína predita, uma kirola, uma catecol oxidase e um tipo de GAPDH (Tabela 3).

Tabela 3 Dez proteínas com maiores valores de quantificação, identificadas por LC-MS, utilizando o protocolo de éter de petróleo

Contig	Proteína	14I	14NI	73I	73NI	120I	120NI	22I	22NI
13166001	Quitinase	301	1.337	1.212	846	104	895	1.900	3.151
22022001	Fosfatase Ácida	2.918	2.186	875	614	278	547	317	710
07152001	Proteína de Armazenamento	879	1.920	848	686	29	636	2.526	1.966
22023001	Fosfatase Ácida	1.545	1.198	989	534	321	641	641	388
13168001	Quitinase	59	296	291	178	11	149	286	1.013
07291001	Kirola	841	250	140	191	47	265	324	132
31009001	Proteína Predita	530	462	215	128	240	50	213	393
21916001	Fosfatase Ácida	520	414	222	330	103	337	528	342
05696001	Catecol Oxidase	345	295	190	233	67	241	166	185
38191001	GAPDH	336	152	146	185	176	197	206	136

I - condição de irrigação. NI - condição de déficit hídrico. Contigs sem o prefixo GSCOCT000.

Conforme ilustrado nas tabelas acima (Tabela 2 e Tabela 3), pode-se observar que o Protocolo de Éter de Petróleo mostrou-se mais eficiente no que diz respeito à identificação de proteínas. Uma vez que o protocolo de Solução ácida não permitiu a identificação de proteínas em alguns clones. Como é o caso dos clones 73I, 73NI, 22I para a proteína GSCOCT00007300001 (Miraculina) e do clone 120NI para a proteína GSCOCT00017376001 (Apirase).

Com base nos resultados de quantificação proteica, foi possível agrupar as proteínas de acordo com sua funcionalidade, conforme figuras abaixo.

Dentre as proteínas que foram identificadas com maiores valores de quantificação em raízes de clones de *C. canephora*, dois tipos de quitinases foram identificados: 13166001 e 13168001, estes representam quitinases indutoras de xilanases, identificadas no GenBank pelo código ADZ48381.1.

Para a quitinase 13166001, é observado que a maior quantidade desta proteína é apresentada no clone 22 na condição NI. Entre os clones tolerantes, esta proteína foi identificada em maior quantidade no clone 14 na condição NI. No clone 73, a quantificação desta proteína foi maior na condição I e no clone

120 na condição NI (Figura 5A). Para os dados de quantificação da quitinase 13168001 (Figura 5B), observa-se que a maior quantidade desta também é apresentada no clone 22 na condição NI. Entre os clones tolerantes, esta proteína foi identificada em maior quantidade nos clones 14 na condição NI e no 73 na condição I. O menor valor de quantificação desta proteína foi apresentado pelo clone 120 para as duas condições (Figura 5B).

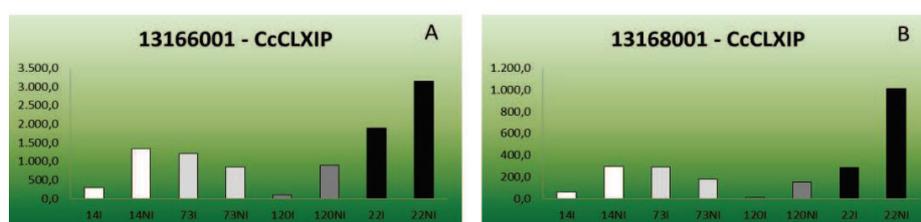


Figura 5 Perfil diferencial de quantificação de quitinases, identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA)

Legenda: **A.** Proteína 13166001. **B.** Proteína 13168001. Os valores de quantificação foram padronizados com base na quantificação da Ubiquitina.

Outro grupo de proteínas identificadas com alto valor de quantificação em raízes de clones de *C. canephora*, é o grupo das fosfatases ácidas 21916001, 22022001 e 22023001, identificadas no GenBank pelo código AFX66999.1.

Quanto ao perfil proteômico para a fosfatase ácida 21916001, é observado que o maior valor de quantificação foi apresentado para o clone 73 na condição I, conforme apresentado pela Figura 6A, valores estes que se assemelham aos apresentados pelo clone 14 para a mesma proteína (Figura 6A). Para os clones 22 e 120, os valores de quantificação foram mais elevados na condição NI (Figura 6A).

Para a fosfatase ácida 22022001, o perfil de quantificação proteica mostra que o maior valor de quantificação foi também apresentado no clone 14 na condição I (Figura 6B). Para os outros clones tolerantes, os perfis também

foram similares, sendo a quantificação da proteína maior na condição NI para o clone 73 e na condição I para o clone 120 (Figura 6B). Em relação ao clone 22, os perfis são totalmente inversos, uma vez que a maior quantidade desta proteína neste clone é apresentada na condição NI.

O perfil proteômico da fosfatase ácida 22023001, conforme apresentado pela Figura 6C, mostra que o maior valor de quantificação foi apresentado para o clone 14 na condição I (Figura 6C), para os outros clones tolerantes, 73 e 120, o maior valor de quantificação também foi apresentado na condição I (Figura 6C). Já para o clone 22, o perfil de quantificação desta proteína foi diferenciado, em relação aos outros clones, sendo maior na condição NI (Figura 6C).

Comparando os dados para as três fosfatases 21916001, 22022001 e 22013001, observa-se que, para os clones, os perfis de quantificação se assemelham para as três fosfatases, como é o caso dos perfis para os clones 14 e 22. Para o clone 73, os perfis se assemelham entre as proteínas 21916001 e 22023001, já para o clone 120, os perfis se assemelham para as proteínas 22022001 e 22023001.

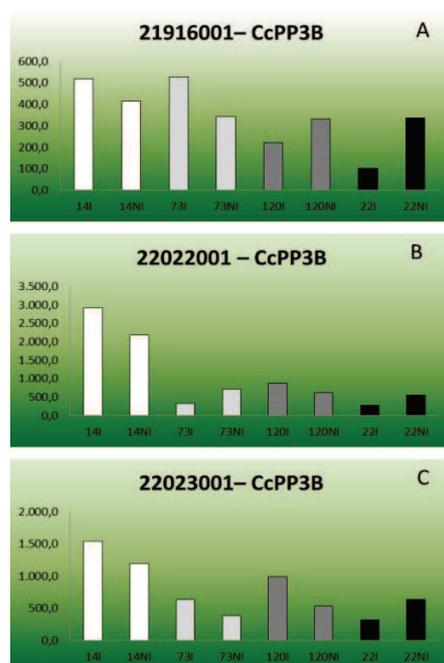


Figura 6 Perfil diferencial de quantificação de fosfatases *CcPP3B* identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA)

Legenda: **A.** Proteína 21916001. **B.** Proteína 22022001 e **C.** Proteína 22023001. Os valores de quantificação foram padronizados com base na quantificação da Ubiquitina.

O valor de quantificação da proteína de armazenamento 07152001 é apresentado na Figura 7. Observa-se que a maior quantidade desta proteína foi apresentada para o clone 22 na condição I. Entre os clones tolerantes, esta proteína foi identificada em maior quantidade no clone 14 na condição NI e no clone 73 na condição I, no clone 120 a maior quantificação foi apresentada na condição NI (Figura 7).

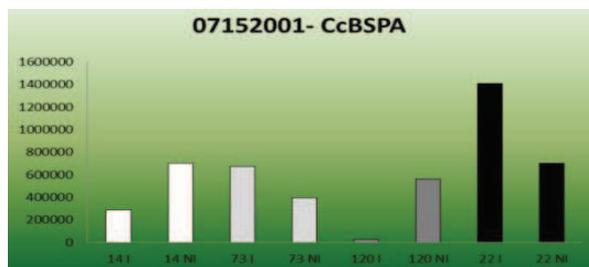


Figura 7 Perfil diferencial de quantificação da proteína de armazenamento 07152001 identificada por LC-MS (NanoAcquity, UPLC-Waters, USA). Os valores de quantificação foram padronizados com base na quantificação da Ubiquitina

Observando os dados de quantificação proteica para 07291001 (Kirola), apresentados na Figura 8, observa-se que a maior quantidade desta é apresentada no clone 14 na condição I. No clone 72, a maior quantificação é apresentada na condição I (Figura 8). Nos clones 120 e 22 a maior quantificação foi maior na condição NI (Figura 8).

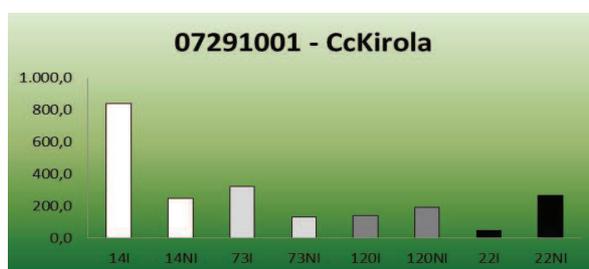


Figura 8 Perfil diferencial de quantificação da proteína de armazenamento 07291001 identificada por LC-MS (NanoAcquity, UPLC-Waters, USA). Os valores de quantificação foram padronizados com base na quantificação da Ubiquitina

Em relação ao perfil quantitativo da proteína catecol oxidase 05696001, apresentado na Figura 9, observa-se que a maior quantidade desta proteína é

apresentada no clone 14 na condição I. Nos outros clones, a quantificação é maior na condição NI (Figura 9).

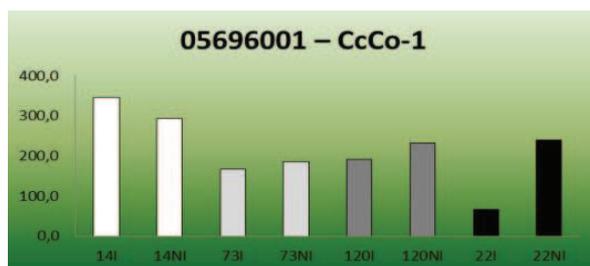


Figura 9 Perfil diferencial de quantificação da catecol oxidase 05696001 – *CcCo-1* identificada por LC-MS (NanoAcquity, UPLC-Waters, USA). Os valores de quantificação foram padronizados com os valores de quantificação da Ubiquitina

Para a proteína 38191001 (Gliceraldeído-3-fosfato-desidrogenase), nota-se que a maior quantidade desta proteína foi apresentada para o clone 14 na condição I (Figura 10). No clone 73, a maior quantificação foi apresentada também na condição I (Figura 10). Já para os clones 22 e 120 houve um pequeno aumento da quantificação desta proteína na condição NI (Figura 10).

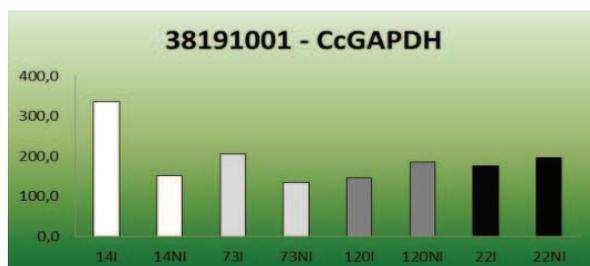


Figura 10 Perfil diferencial de quantificação da Gliceraldeído-3-fosfato desidrogenase 38191001 – *CcGAPDH* identificada por LC-MS (NanoAcquity, UPLC-Waters, USA). Os valores de quantificação foram padronizados com os valores de quantificação da Ubiquitina

Para a proteína 31009001 (Proteína predita), nota-se que a maior quantidade desta proteína é apresentada para o clone 14 na condição I e nos clones 22 e 120, também foi observado maior quantificação desta proteína na condição I (Figura 11). No clone 73, a maior quantificação foi apresentada na condição NI.

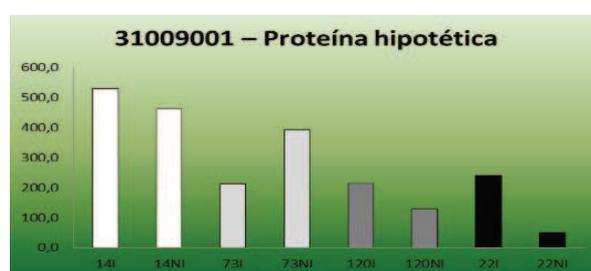


Figura 11 Perfil diferencial de quantificação da proteína hipotética 31009001 identificada por LC-MS (NanoAcquity, UPLC-Waters, USA). Os valores de quantificação foram padronizados com os valores de quantificação da Ubiquitina

Para a proteína ribossomal, foram identificados dois tipos de proteínas: *CcRPS4* e *CcRPS16* (Figura 12). A proteína *CcRPS4* foi identificada nos clones 14, 73 e 120 na condição de irrigação, e no clone 22, na condição de seca, sendo que a maior quantificação foi observada para o clone 14 (Figura 12A). A proteína *CcRPS16* foi identificada em todos os clones para as duas condições. Para os clones 14, 73 e 22, a maior quantificação foi observada para a condição irrigada, sendo que a maior quantidade foi observada no clone 14 (Figura 12B).

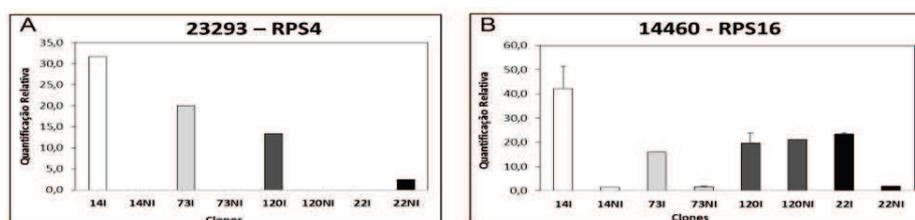


Figura 12 Quantificação de proteínas ribossomais 40S identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA). Resultados de quantificação relativa, com os dados normalizados pela Ubiquitina, resultantes das análises realizadas com o “ProteinLynx” (versão 5.2.2), nas amostras de raízes de *C. canephora* com a Base de dados do Genoma de *C. canephora* (Cenargen, Brasília-DF), como referência. Todas as proteínas foram obtidas pela extração com o protocolo de éter de petróleo

Legenda: **A.** Proteína ribossomal 40S 23293001. **B.** Proteína ribossomal 40S 14460001.

Proteínas ribossomais são peças-chave na síntese de proteínas, sendo portanto cruciais no desenvolvimento celular e, como tal, devem estar presentes em grande quantidade nos organismos para que exerçam sua função. Em decorrência desta característica, esse grupo de proteínas foi utilizado como grupo controle para avaliar a eficácia da análise utilizada. Como ilustrado no resultado acima, é notável que não foi possível a identificação destas proteínas em todos os clones nas duas condições, o que deixa uma dúvida quanto à eficiência da análise.

Uma alternativa para a não identificação destas proteínas pode estar relacionada com a etapa de fragmentação destas, uma vez que estas podem ter polimorfismos em suas sequências nos resíduos de lisina e arginina, o que modifica a fragmentação triptica, podendo alterar o padrão de fragmentação teórico e a identificação da proteína.

Como forma de averiguar esta hipótese, foi verificado a presença de SNPs na sequência do gene que codifica estas proteínas nos diferentes clones. Estes SNPs foram anotados e as formas polimórficas destas proteínas para cada

clone foram confrontadas com os fragmentos trípticos teóricos a fim de verificar se há diferenças nos resultados. Freire (2010) descreveu a ocorrência de polimorfismos em genes envolvidos na resposta ao estresse hídrico em *C. arabica*, também evidenciou diferenças no padrão de expressão de alguns genes integrantes da via de resposta à desidratação, quando expostos a diferentes níveis de estresse hídrico.

Algumas proteínas se destacaram quanto à quantificação proteica e foram selecionadas para a verificação destas diferenças (Tabela 4). Em cada clone analisado, as sequências proteicas foram anotadas com seus respectivos SNPs, possibilitando a construção de um banco de dados específicos com as sequências das 15 proteínas analisadas.

Os fragmentos trípticos resultantes da espectrometria foram comparados com os fragmentos de digestão *in silico* de dois bancos de dados distintos: (i) o banco de dados com as sequências da proteína para cada clone e (ii) o banco de dados com a sequência da proteína do genoma de referência.

O valor de quantificação proteica apresentado em cada clone foi mais elevado que na análise anterior, na maioria das proteínas. Além disso, em muitas proteínas foram identificadas as duas formas alélicas, conforme é observado nos resultados apresentados abaixo.

Tabela 4 Proteínas reanalisadas por LC-MS, utilizando o banco de dados do Genoma de *C. canephora* (Cenargen, Brasília-DF)

<b>Código da proteína (Contig)</b>	<b>Proteína</b>
GSCOCT00005696001	Catechol Oxidase
GSCOCT00007152001	Protein storage
GSCOCT00007291001	Kirola
GSCOCT00007300001	Miraculin
GSCOCT00011142001	D-sorbitol-6-phosphate dehydrogenase
GSCOCT00013168001	Chitinase
GSCOCT00013169001	Chitinase
GSCOCT00013173001	Chitinase
GSCOCT00014460001	40S ribosomal protein
GSCOCT00021916001	Acid phosphatase
GSCOCT00022022001	Acid phosphatase
GSCOCT00023293001	40S ribosomal protein
GSCOCT00028091001	Ascorbate peroxidase
GSCOCT00037166001	Dehydrin Dh1a
GSCOCT00038191001	GAPDH

Os resultados da espectrometria utilizando os dois bancos foram comparados e a quantificação destas proteínas, para cada clone e em cada condição, analisadas. Em algumas proteínas não foram identificados SNPs (Figura 13), para estas, foi realizada somente uma comparação entre a quantificação das proteínas utilizando o banco do genoma de referência desta análise (análise nova) e a quantificação anterior na qual se utilizou todas as sequências do banco de dados do genoma de referência (análise antiga).

Os dados de quantificação da catecol oxidase (05696001) estão apresentados na Figura 13A, na qual é observado que os valores de quantificação entre as análises apresentaram valores distintos. Para o clone 14, o valor de quantificação aumentou com a utilização do banco novo nas duas condições; para o clone 22, os valores foram semelhantes; para o clone 73 e 120 o valor de quantificação aumentou somente para a condição I na nova análise.

Os resultados da proteína Kirola (07291001) estão apresentados na Figura 13B, na qual é observado que os valores de quantificação entre as análises também apresentaram valores distintos. Comparando os resultados, na

análise nova, para o clone 14, o valor de quantificação foi menor nas duas condições; para o clone 22, na condição I a quantificação diminuiu; na condição NI permaneceu semelhante nas duas análises; para o clone 73 e 120 na condição I a quantificação foi maior e na condição NI, menor.

Para a proteína Miraculina (07300001), observou-se que não há aumento de quantificação desta proteína nos clones (14NI e 120NI) em que esta foi identificada, quando se compara os resultados

Nos resultados da proteína ribossomal (14460001), é observado que o valor de quantificação aumentou somente nos clones 22 e 73 na condição NI. Curiosamente no clone 14 na condição NI, a proteína não foi identificada (Figura 13D).

Os dados de quantificação da fosfatase ácida (22022001) estão apresentados na Figura 13E, na qual é observado que o valor de quantificação diminuiu somente nos clones 22, 73 e 120 na condição NI. Curiosamente no clone 14 na condição NI, a proteína não foi identificada.

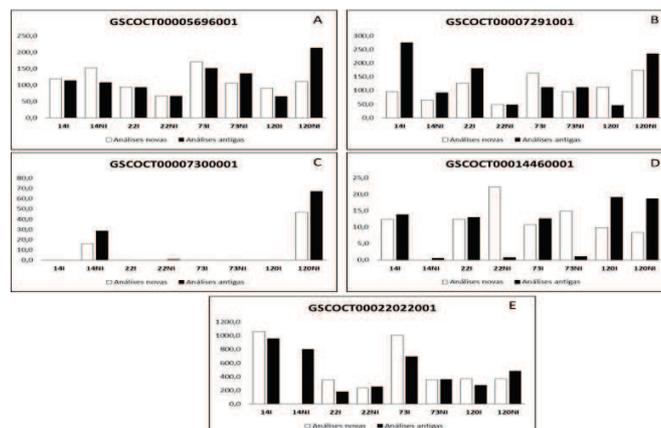


Figura 13 Perfis de quantificação proteica

Legenda: **A.** Catecol oxidase (05696001). **B.** Kirola (07291001). **C.** Miraculina (07300001). **D.** Proteína ribossomal (14460001). **E.** Fosfatase ácida (22022001), identificadas por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando as sequências do genoma de referência do banco de dados do Genoma de *C. canephora* (Cenargen, Brasília-DF), (em preto) e os resultados da análise restrita, com somente as sequências das 15 proteínas e com as formas alélicas de cada clone (em branco).

Baseado nas análises descritas acima, pode-se afirmar que ao utilizar as sequências correspondentes a cada clone como banco de dados para identificação das proteínas houve diferenças significativas entre os valores de quantificação.

Para a proteína de armazenamento (07152001), apesar de existir duas formas alélicas, somente uma foi identificada por esta análise. Os dados de quantificação estão apresentados na Figura 14. O valor de quantificação diminuiu somente nos clones 14 na condição I e 73 e 120 na condição NI. Nos clones 14 e 22 na condição NI, os valores de quantificação foram semelhantes. (Figura 14).

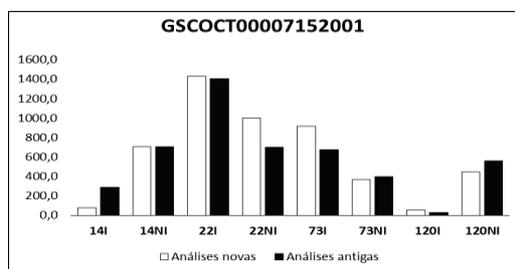


Figura 14 Perfil de quantificação da proteína de armazenamento (07152001). Dados obtidos pela identificação por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando as sequências do genoma de referência do banco de dados do Genoma de *C. canephora* (Cenargen, Brasília-DF), (em preto) e os resultados da análise restrita, com somente as sequências das 15 proteínas e com as formas alélicas de cada clone (em branco)

Os dados abaixo descrevem os resultados de espectrometria nos quais foi possível identificar as duas formas alélicas das proteínas, quando se utilizou o banco de dados de sequências das proteínas específicas para cada clone (15 proteínas). Considerou-se a forma HT1 aquela cuja sequência é igual a sequência do genoma de referência e HT2 a forma polimórfica desta proteína, ou seja aquela com modificação na sua sequência (SNP).

As duas formas alélicas da proteína D-sorbitol-6-fosfatase desidrogenase (11142001) foram identificadas. Curiosamente, somente no clone 22, na condição NI, a forma alélica HT2 foi encontrada (Figura 15).

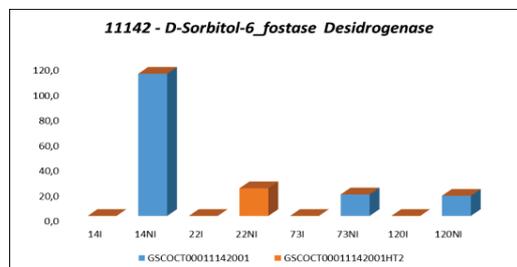


Figura 15 Perfil de quantificação da D - sorbitol - 6 - fosfatase desidrogenase (11142001). Dados obtidos a partir da identificação por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando o banco de dados com sequências das 15 proteínas nos clones (14, 22, 73 e 120) e suas formas alélicas

Para as três quitinases (13168001, 13169001 e 13173001) analisadas, foram identificadas suas duas formas alélicas. Na quitinase 13168001, as duas formas alélicas foram identificadas nos clones 14 e 120 na condição NI, nos demais clones foi identificada somente a forma HT1 (Figura 16A). A quitinase 13169001 foi identificada somente nos clones 22 e 120 na condição NI, sendo que no clone 22 somente foi encontrada a forma alélica HT1 e no 120 HT2 (Figura 16B). Na quitinase 13173001, a forma alélica HT1 foi identificada nos clones 73 e 120 na condição NI; a forma alélica HT2 foi identificada nos clones 14 na condição NI e 22 na condição I, e nos clones 22 na condição NI e 73 na condição I, as duas formas alélicas (Figura 16C).

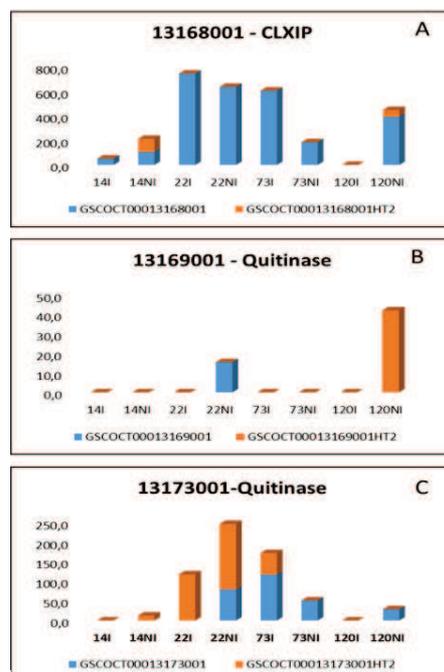


Figura 16 Perfis de quantificação proteica da quitinase

Legenda: **A.** 13168001. **B.** 13169001. **C.** 13173001. Dados obtidos a partir da identificação por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando o banco de dados com as sequências das 15 proteínas nos clones (14, 22, 73 e 120) e suas formas alélicas.

Para as duas fosfatases ácidas (21916001 e 22023001) analisadas, foram identificadas suas duas formas alélicas. Na fosfatase ácida 21916001, foi identificada a forma HT1 nos clones 22 e 120 nas duas condições, nos demais clones (14 e 73) foram identificadas as duas formas alélicas (Figura 17A). Na fosfatase ácida 22023001, somente no clone 120 nas duas condições foi identificada a forma alélica HT1 e nos demais clones foram identificadas as duas formas alélicas (Figura 17B).

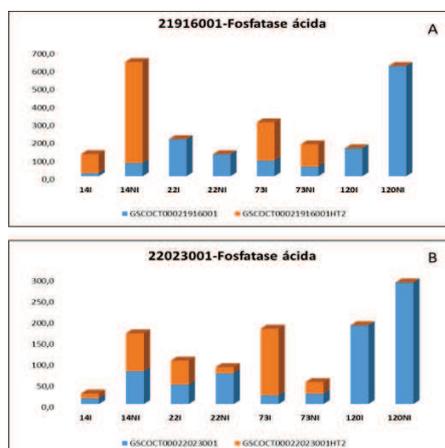


Figura 17 Perfis de quantificação proteica da fosfatase ácida

Legenda: **A.** 21916001. **B.** 22023001. Dados obtidos a partir da identificação por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando o banco de dados com seqüências das 15 proteínas nos clones (14, 22, 73 e 120) e suas formas alélicas.

Para a ascorbato peroxidase (28091001), foram identificadas suas duas formas alélicas. A forma HT1 nos clones 14 nas duas condições, 22 na condição I e 73 na condição NI; a forma HT2 nos clones 22 na condição NI, 73 na condição I e 120 na condição NI, somente no clone 120 na condição I as duas formas HT1 e HT2 foram identificadas (Figura 18).

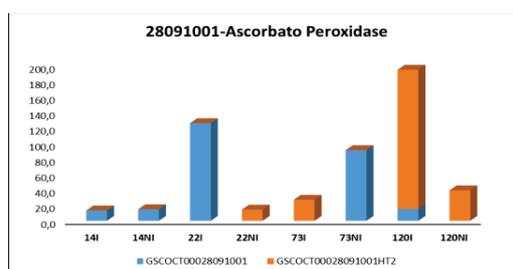


Figura 18 Perfil de quantificação proteica da ascorbato peroxidase 28091001. Dados obtidos a partir da identificação por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando o banco de dados com seqüências das 15 proteínas nos clones (14, 22, 73 e 120) e suas formas alélicas

Para a desidrina (37166001), foram identificadas suas duas formas alélicas: a forma HT1 nos clones 14 e 73 na condição I, 22 e 120 na condição NI e a forma HT2 somente no clone 120 na condição I e no clone 14, na condição NI, as duas formas HT1 e HT2 foram identificadas (Figura 19).

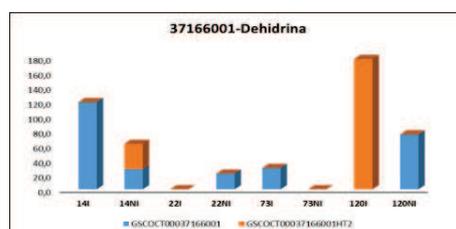


Figura 19 Perfil de quantificação proteica da desidrina 37166001. Dados obtidos a partir da identificação por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando o banco de dados com sequências das 15 proteínas nos clones 14, 22, 73 e 120 e suas formas alélicas

Para a Gliceraldeído-3-fosfato desidrogenase (38191001), foram identificadas somente uma forma alélica (HT1). Esta proteína não foi identificada nos clones 14 e 120 na condição I (Figura 20).

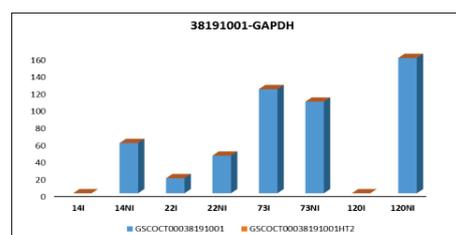


Figura 20 Perfil de quantificação proteica da Gliceraldeído-3-fosfato desidrogenase (GAPDH) 38191001. Dados obtidos a partir da identificação por LC-MS (NanoAcquity, UPLC-Waters, USA), utilizando o banco de dados com sequências das 15 proteínas nos clones (14, 22, 73 e 120) e suas formas alélicas

Aprimorando e refinando o banco de dados, por meio da inserção das sequências com SNPs, foi possível identificar valores quantitativos superiores aos da análise anterior. Possivelmente, os valores de quantificação identificados nas proteínas analisadas para a forma HT2 representam este aumento na quantificação para cada proteína, ou seja, o aprimoramento da técnica possibilitou a quantificação de uma forma alélica destas proteínas que antes era desconsiderada.

#### 4 CONCLUSÕES

Foram identificadas 598 proteínas em raízes de clones de *C. canephora*, o que representa cerca de 1% do proteoma conhecido do cafeeiro, indicando uma certa limitação da técnica.

Ao comparar os resultados nos dois protocolos conclui-se que o protocolo de Éter de Petróleo foi mais eficiente na identificação de proteínas.

Ao comparar os resultados de quantificação utilizando-se dois bancos de dados (completo *versus* restrito) nas análises com o ProteinLinx, observou-se diferenças entre os resultados.

Da mesma forma, ao comparar os resultados das análises de quantificação utilizando-se os bancos de dados com proteínas específicas (SNPs) para cada clone analisado observou-se um aumento significativo na identificação destas proteínas.

Com os resultados obtidos também foi possível a identificação de diferentes formas alélicas.

Entretanto, podemos concluir com os resultados obtidos que as análises proteômicas realizadas, conforme descrito neste trabalho, ainda necessitam de aprimoramento.

## REFERÊNCIAS

BANTSCHIEFF, M. et al. Quantitative mass spectrometry in proteomics: a critical review. **Analytical and Bioanalytical Chemistry**, Heidelberg, v. 389, n. 4, p. 1017-1031, Oct. 2007.

BARSALOBRES-CAVALLARI, C. F. et al. Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. **BMC Molecular Biology**, London, v. 10, n. 1, p. 1-11, Jan. 2009.

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

GONZÁLEZ-AGÜERO, M. et al. Identification of two putative reference genes from grapevine suitable for gene expression analysis in berry and related tissues derived from RNA-Seq data. **BMC Genomics**, London, v. 14, n. 878, p. 1-12, Dec. 2013.

GUIMARÃES, B. L. S. **Proteômica diferencial em clones de *Coffea canephora* sob condições de déficit hídrico**. 2007. 71 p. Dissertação (Mestrado em Fisiologia Vegetal) - Universidade Federal de Viçosa, Viçosa, MG, 2007.

HEIMBECK, I. G. R. **Desenvolvimento de metodologias para a identificação de proteínas e metabólitos durante o desenvolvimento de frutos de café (*Coffea arabica* L.)**. 2011. 110 p. Dissertação (Mestrado em Biotecnologia Vegetal) - Universidade Federal de Lavras, Lavras, 2011.

LANÇAS, F. M. **Cromatografia líquida moderna: HPLC/CLAE**. Campinas: Átomo, 2009. 382 p.

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

MURAD, A. M.; RECH, E. NanoUPLC-MSE proteomic data assessment of soybean seeds using the Uniprot database. **BMC Biotechnology**, London, v. 12, n. 82, p. 2-17, Nov. 2012.

SUSSULINI, A. et al. Evaluation of soybean seed protein extraction focusing on metalloprotein analysis. **Microchimica Acta**, Dordrecht, v. 158, n. 1/2, p. 173-180, Nov. 2007.

VANDESOMPELE, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. **Genome Biology**, London, v. 3, n. 7, p. 1-17, June 2002.

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

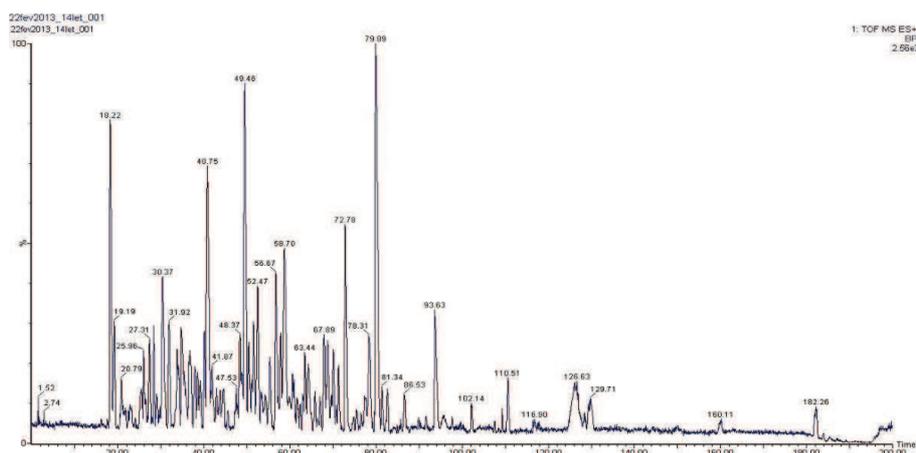
WIJK, K. J. van. Challenges and prospects of plant proteomics. **Plant Physiology**, Bethesda, v. 126, n. 2, p. 501-508, June 2001.

WILDGRUBER, R. et al. Towards higher resolution: two-dimensional electrophoresis of *Saccharomyces cerevisiae* proteins using overlapping narrow immobilized pH gradients. **Electrophoresis**, Weinheim, v. 21, n. 13, p. 2610-2616, July 2000.

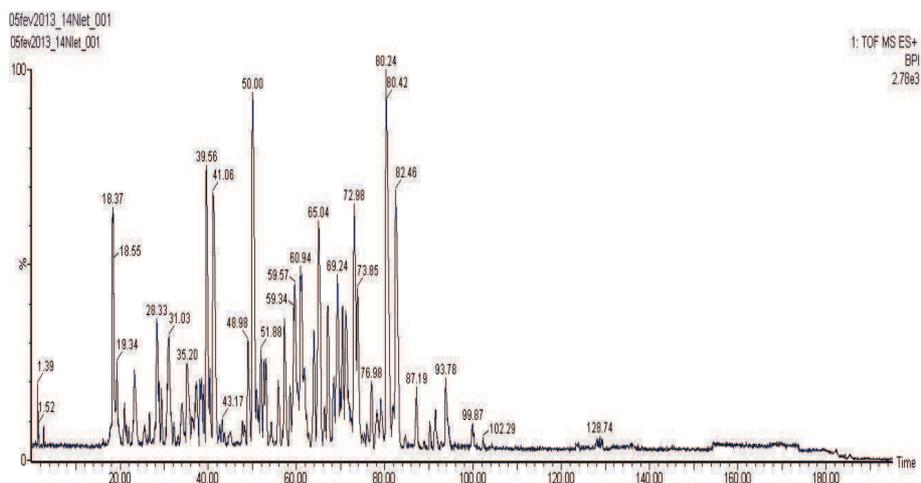
YERGEY, A. L. et al. De novo sequencing of peptides using MALDI/TOF-TOF. **Journal of the American Society for Mass Spectrometry**, New York, v. 13, n. 7, p. 784-791, July 2002.

## ANEXOS

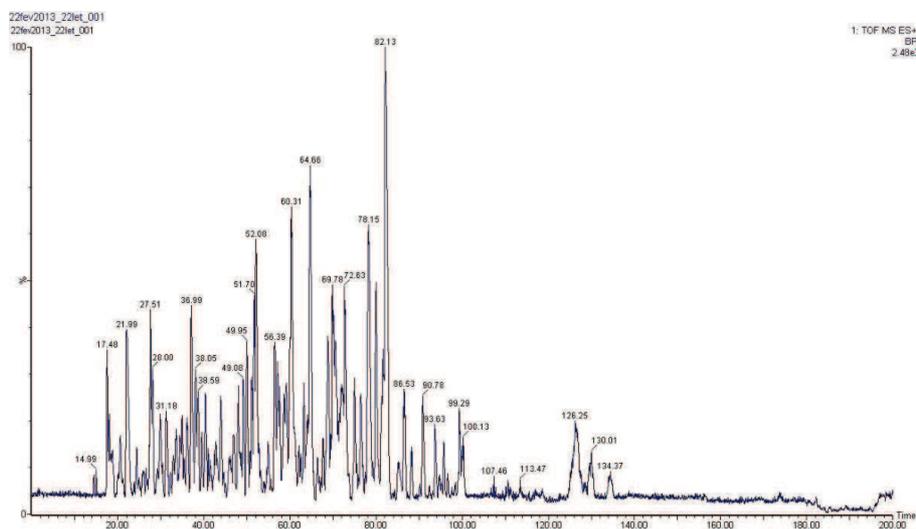
## ANEXO A – Figuras



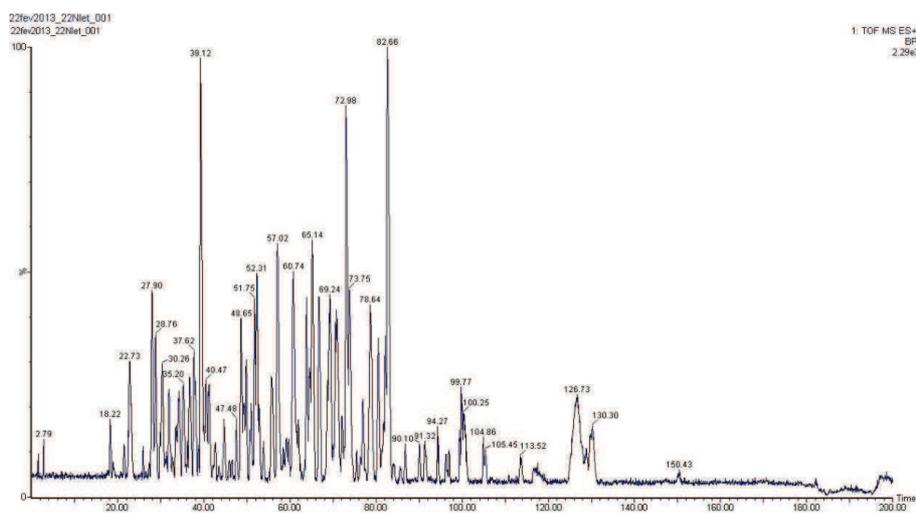
**Figura 21** Fracionamento do extrato proteico de raiz de *Coffea canephora* (14I-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B



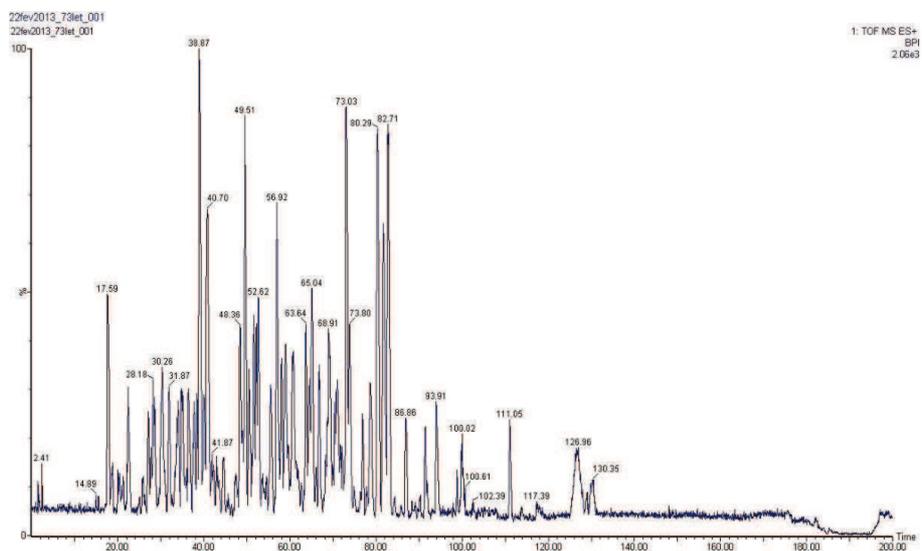
**Figura 22** Fracionamento do extrato proteico de raiz de *Coffea canephora* (14NI-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B



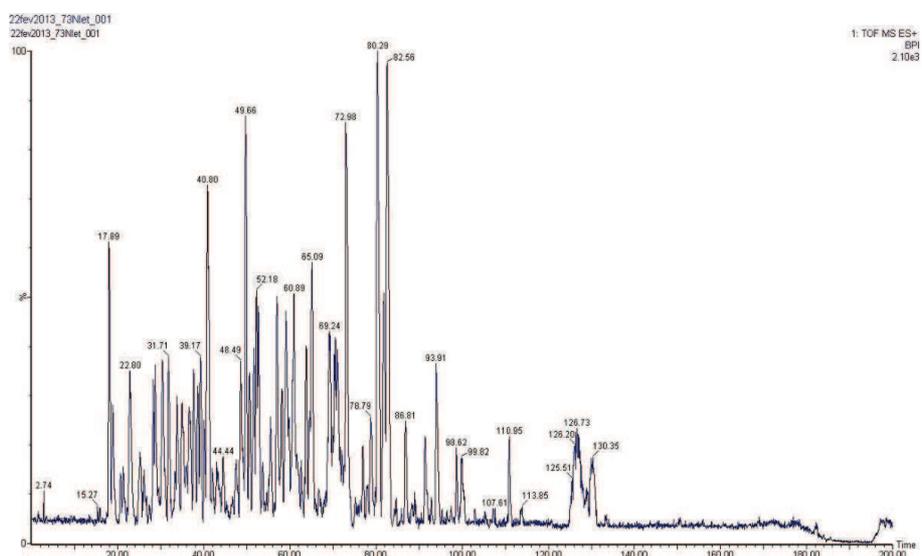
**Figura 23** Fracionamento do extrato proteico de raiz de *Coffea canephora* (22I-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B



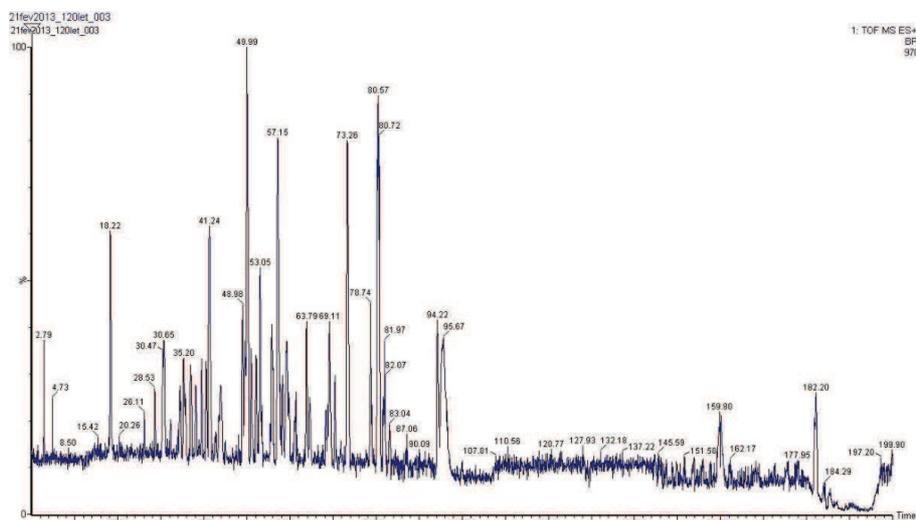
**Figura 24** Fracionamento do extrato proteico de raiz de *Coffea canephora* (22NI-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B.



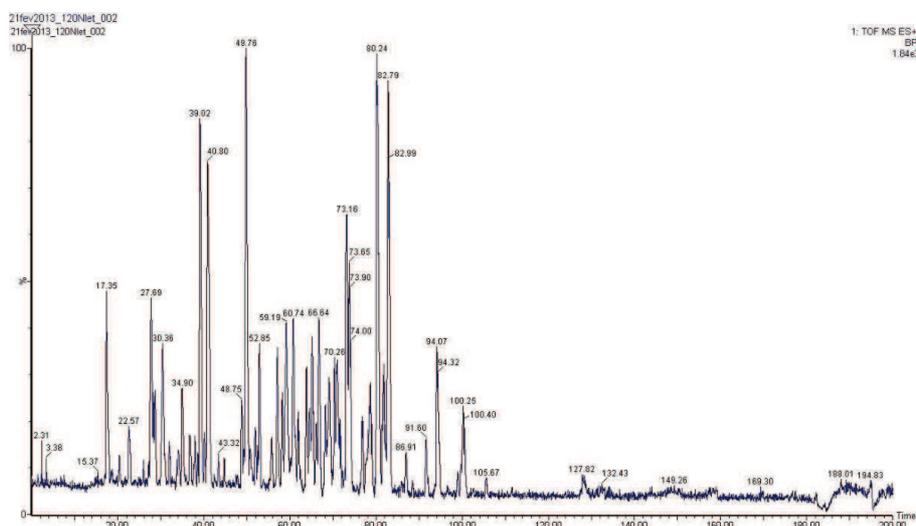
**Figura 25** Fracionamento do extrato proteico de raiz de *Coffea canephora* (73I-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B.



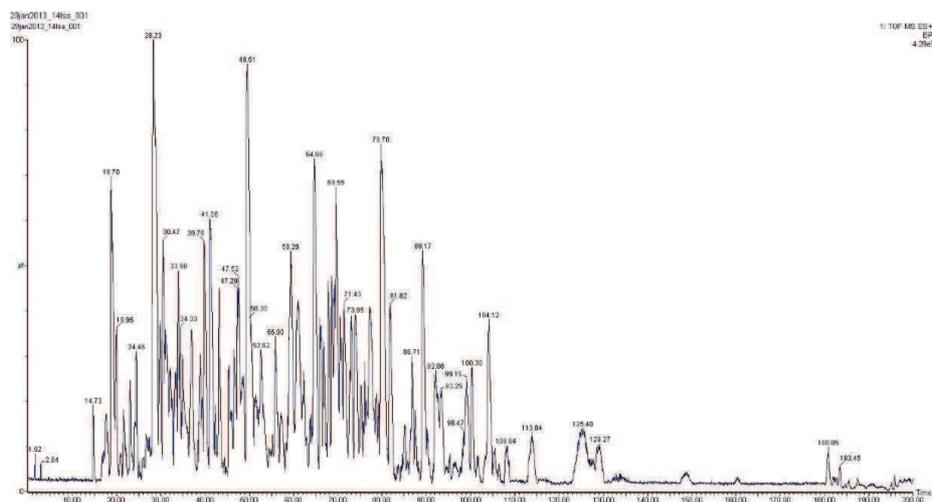
**Figura 26** Fracionamento do extrato proteico de raiz de *Coffea canephora* (73NI-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B.



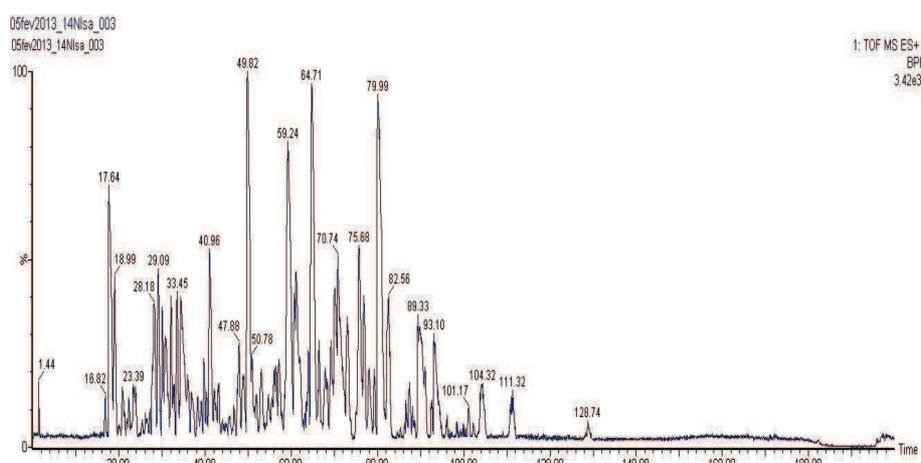
**Figura 27** Fracionamento do extrato proteico de raiz de *Coffea canephora* (120I-ET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B



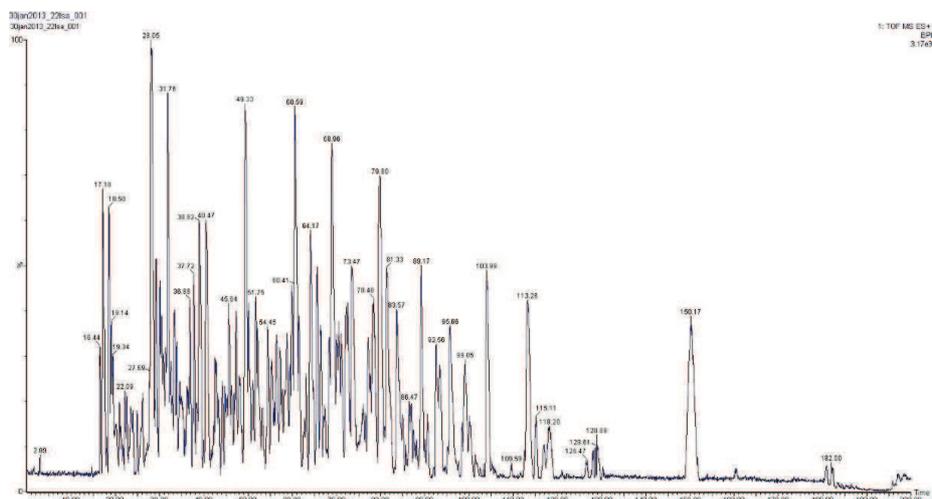
**Figura 28** Fracionamento do extrato proteico de raiz de *Coffea canephora* (120 NIET) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B.



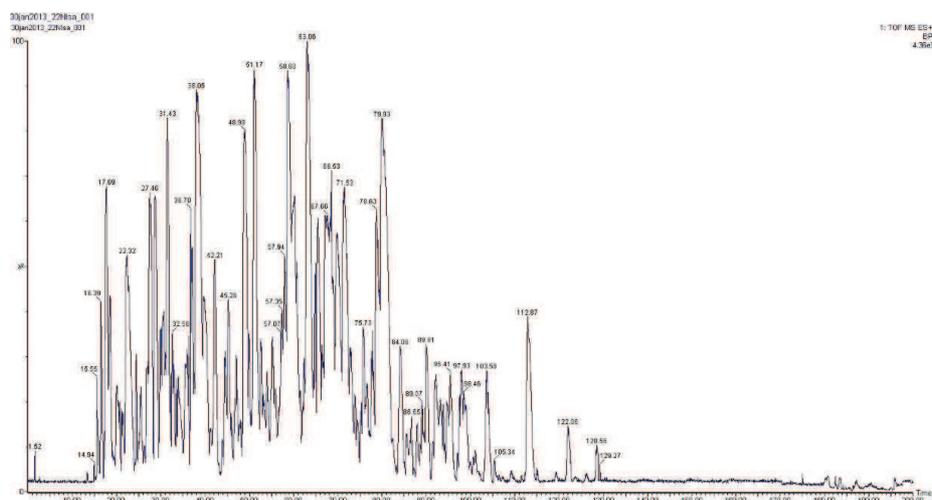
**Figura 29** Fracionamento do extrato proteico de raiz de *Coffea canephora* (14I-SA) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B



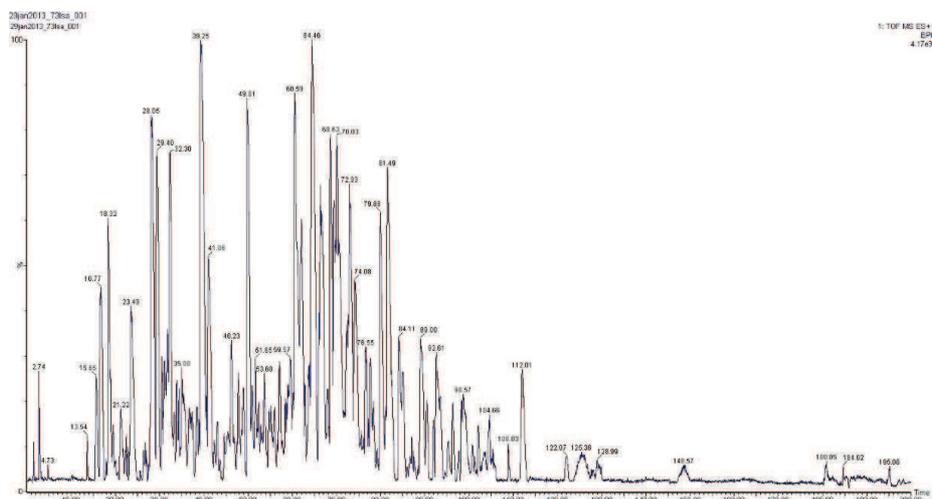
**Figura 30** Fracionamento do extrato proteico de raiz de *Coffea canephora* (14NI-SA) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B.



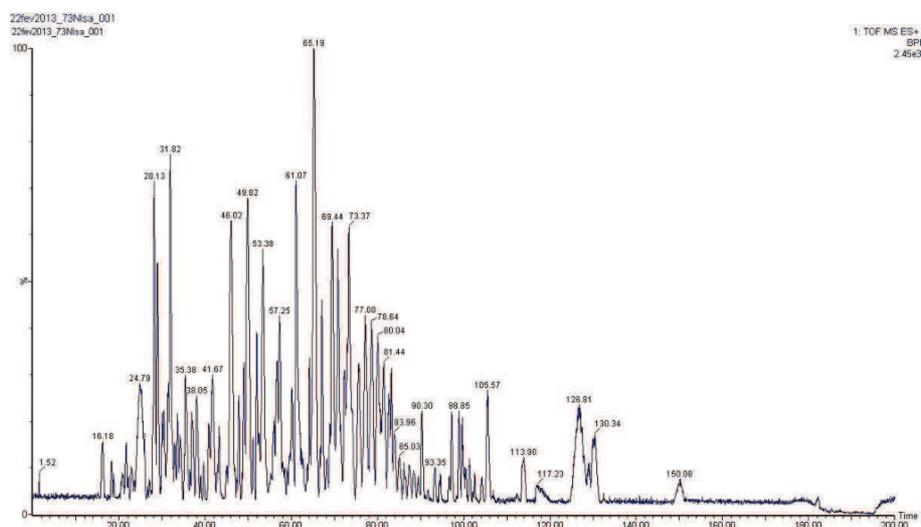
**Figura 31** Fracionamento do extrato proteico de raiz de *Coffea canephora* (22I-SA) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B.



**Figura 32** Fracionamento do extrato proteico de raiz de *Coffea canephora* (22NI-SA) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B.



**Figura 33** Fracionamento do extrato proteico de raiz de *Coffea canephora* (73I-SA) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B.



**Figura 34** Fracionamento do extrato proteico de raiz de *Coffea canephora* (73NI-SA) por cromatografia de alta performance, em modo de fase reversa, utilizando uma coluna capilar (NanoAcquity, UPLC-Waters, USA) em um gradiente de 3% a 85%B.

**CAPÍTULO 4 Análise integrada dos perfis transcriptômico e proteômico em raízes de clones de cafeeiro (*Coffea canephora*) sensíveis e tolerantes ao estresse hídrico**

**RESUMO**

Para obtenção de uma visão geral das respostas da planta às condições ambientais, estudos de associação entre as técnicas “omicas” são amplamente utilizados. Até chegar à resposta ao estresse, uma cascata de eventos ocorre na planta que resultam na manifestação da resposta, possibilitando que a planta sobreviva e se recupere do estresse sofrido. Diante do suposto, o objetivo do presente trabalho foi integrar os resultados dos perfis transcriptômico e proteômico em raízes de clones de *C. canephora*, sensíveis e tolerantes ao déficit hídrico nas condições de irrigação (I) e não irrigação (NI). Para análises de proteômica, 09 genes e suas correspondentes proteínas foram avaliados, obtidos a partir do mesmo material vegetal. Os resultados de transcriptômica foram obtidos seguindo a mesma metodologia do capítulo 2, a normalização dos dados de expressão foram obtidos por *Northern* eletrônico utilizando o *software* Q-Seq (Lasergene), cujos valores foram normalizados pela abordagem RPKM, já os dados de expressão relativa obtidos por qPCR foram normalizados e estandardizados com base na expressão da Ubiquitina-*CcUBQ10*, que é utilizado como calibrador interno de referência. As proteínas totais foram obtidas pelo protocolo de Sussulini et al. (2007), precipitado com acetona. As amostras foram submetidas a uma etapa de digestão enzimática com tripsina e, posteriormente, fracionadas e analisadas no LC-MS, utilizando a cromatografia líquida de fase reversa acoplada à espectrometria de massa. Os resultados de identificação dos íons tripsinizados foram obtidos com o auxílio do *software* “ProteinLynx” utilizando o Banco de dados de *C. canephora*. Apesar de alguns resultados não corresponderem, o aprimoramento das técnicas de identificação proteica, por meio da utilização de sequências genótipo-específicas, pode ser uma alternativa para futuras análises integradas, uma vez que permite a compreensão do mecanismo de regulação e dos pontos de controle do fluxo da informação gênica.

Palavras-chave: Cafeeiro. Raiz. Tolerância à seca. “Omicas”.

**Integrated analysis of the transcriptomic and proteomic profiles in coffee  
(*Coffea canephora*) clone roots sensitive and tolerant to water stress**

**ABSTRACT**

To obtain a general view of the responses of plants to environmental conditions, association studies between omics techniques are widely used. While reaching the response to stress, a cascade of events occurs in the plant, which results in the manifestation of the response, allowing the plant to survive and recover from the suffered stress. With this, the objective of the present work was to integrate the results of transcriptomic and proteomic profiles in *C. canephora* clone roots, sensitive and tolerant to water deficit under irrigated (I) and non-irrigated (NI) conditions. For the proteomic analysis, 09 genes and their correspondent proteins, obtained from the same plant material, were evaluated. The results of the transcriptome were obtained following the same methodology as in chapter 2. The normalization of the expression data were obtained by electronic Northern using the Q-Seq software (Lasergene), of which values were normalized by the RPKM approach, while the relative expression data obtained by qPCR were normalized and standardized based on the Ubiquitin-*CcUBQ10* expression, which is used as internal reference calibrator. The total proteins were obtained by the Sussulini et al. (2007) protocol, precipitated with acetone. The samples were submitted to an enzyme digestion stage with trypsin and, posteriorly, fractioned and analyzed on the LC-MS, using reverse phase liquid chromatography coupled to mass spectrometry. The results of identifying the trypsinized ions were obtained with the aid of the ProteinLynx software using the *C. canephora* databank. Despite a few results not corresponding, the improvement of the protein identification techniques, by means of the use of genotype-specific sequences, might be an alternative for future integrated analyses, since it allows the understanding of the regulation mechanism and of the control points for genic information flow.

Keywords: Coffee plant. Root. Drought tolerance. Omics.

## 1 INTRODUÇÃO

As técnicas ômicas, que incluem a genômica (DNA), transcriptômica (RNA), proteômica (proteína) e metabolômica (metabólitos) são amplamente desenvolvidas e utilizadas para obter uma visão geral da resposta das plantas às condições ambientais, tais como respostas aos diversos tipos de estresses, por meio da associação entre os dados de expressão gênica diferencial, abundância de proteínas e acúmulo de metabólitos, que alteram o perfil das “ômicas” (KOMATSU; HIRAGA; YANAGAWA, 2012; KOMATSU; SHIRASAKA; SAKATA, 2013).

Na natureza, as plantas são normalmente submetidas a uma combinação de múltiplos estresses abióticos, por exemplo, a seca no verão é muitas vezes acompanhada de estresse de altas temperaturas (LIU et al., 2003). Em condições de campo, as repostas a estes estresses podem ser modificadas de forma sinérgica ou antagônica pela sobreposição de outros estresses (DEBNATH; PANDEY; BISEN, 2011).

A capacidade das plantas em lidar com o estresse ambiental depende de mudanças rápidas nos níveis proteicos, que podem ser aumentados ou diminuídos como resultado de alterações na expressão do gene (HAKEEN et al., 2012). Entretanto, até chegar à resposta ao estresse, uma cascata de eventos ocorre na planta que incluem a percepção do estímulo, transmissão do sinal ao citoplasma e núcleo, expressão gênica e por fim alterações metabólicas que resultam na manifestação da resposta, possibilitando que a planta sobreviva e se recupere do estresse sofrido (SALAVATI et al., 2012).

Diante do suposto, o objetivo do presente trabalho foi integrar os resultados dos perfis transcriptômico e proteômico em raízes de clones de *C. canephora*, sensíveis e tolerantes ao déficit hídrico nas condições de irrigação (I) e não irrigação (NI).

## 2 MATERIAL E MÉTODOS

Utilizando a metodologia descrita no Capítulo 2, para obtenção dos resultados de análises de expressão, e a metodologia descrita no Capítulo 3, para obtenção do perfil proteico, foi possível abordar os resultados com base na integração dos perfis transcriptômico e proteômico.

Para cada clone e regime hídrico, o RNA total foi extraído e utilizado tanto para obtenção do perfil do transcriptoma das raízes, pelo sequenciamento 454, que possibilitou análises *in silico* da expressão por *Northern* eletrônico, bem como para a validação por qPCR dos genes selecionados entre os clones nas condições I vs. NI.

Para os dados de transcriptômica, os resultados de expressão foram obtidos por *Northern* eletrônico utilizando o *software* Q-Seq (Lasergene), cujos valores foram normalizados pela abordagem RPKM (Reads Per Kilobases Per Million) (MORTAZAVI et al., 2008). Para os dados de qPCR, a normalização e standardização da expressão dos genes foi efetuada com o gene que expressa a Ubiquitina-*CcUBQ10*, que é utilizado como calibrador interno de referência, por ser considerada uma proteína constitutiva, na qual a expressão é constante independente da condição em que a planta se encontra (BARSALOBRES-CAVALLARI et al., 2009).

Para os dados de proteômica, foram utilizados os resultados obtidos pelo protocolo de éter de petróleo, cuja eficiência foi comprovada pelos resultados obtidos no Capítulo 2 (Tabela 2 e Tabela 3), uma vez que o protocolo de Solução ácida não permitiu a identificação de proteínas em alguns clones.

### 3 RESULTADOS E DISCUSSÃO

Genes candidatos para tolerância à seca já foram descritos em trabalhos anteriores (FREIRE, 2010; MARRACCINI et al., 2011, 2012; VIEIRA et al., 2013). Entretanto, no nosso conhecimento, nenhum trabalho molecular em raízes de cafeeiro, foi descrito. Com o objetivo de complementar os resultados acima citados, foram integrados dados de análises de transcriptômica e proteômica, com a possibilidade de encontrar possíveis correlações entre a expressão do gene e a formação do seu produto final, as proteínas, em raízes de clones *C. canephora* (14, 73 e 120, tolerantes e 22, sensível).

Dentre as proteínas identificadas por espectrometria de massa (LC-MS) foram selecionadas aquelas com maiores valores de quantificação, conforme Tabela 1. Dentro desse grupo de proteínas, destacam-se fosfatases e quitinases, que apresentaram os maiores valores entre as proteínas selecionadas.

Tabela 1 Proteínas que apresentaram maiores valores de quantificação, com base nos resultados de LC-MS, utilizando o protocolo de éter de petróleo

Contig	Proteína	14I	14NI	73I	73NI	120I	120NI	22I	22NI
13166001	Quitinase	301	1.337	1.212	846	104	895	1.900	3.151
22022001	Fosfatase Ácida	2.918	2.186	875	614	278	547	317	710
07152001	Proteína de Armazenamento	879	1.920	848	686	29	636	2.526	1.966
22023001	Fosfatase Ácida	1.545	1.198	989	534	321	641	641	388
13168001	Quitinase	59	296	291	178	11	149	286	1.013
07291001	Kirola	841	250	140	191	47	265	324	132
21916001	Fosfatase Ácida	520	414	222	330	103	337	528	342
05696001	Catecol Oxidase	345	295	190	233	67	241	166	185
38191001	GAPDH	336	152	146	185	176	197	206	136

I - condição de irrigação. NI - condição de déficit hídrico. Contigs sem o prefixo GSCOCT000

Dos genes candidatos utilizados nos experimentos de qPCR, somente cinco codificam para a expressão de proteínas encontrados nos resultados LC-MS. Quando se compara as duas análises (transcriptômica e proteômica) para as raízes de *C. canephora*, pode-se observar considerável correspondência para a maioria dos dados (Figuras 2, 3, 5, 6 e 8), para os outros resultados, foram comparados os dados de expressão *in silico* e de quantificação proteica.

### **Resultados de “Ômicas” correspondentes**

No presente estudo, foi feita uma tentativa preliminar de integrar dois níveis de “ômicas” e o impacto que o déficit hídrico pode causar no desenvolvimento dos clones de cafeeiro, por meio das análises de expressão de alguns genes e proteínas envolvidas na resposta da planta ao estresse. Aqui estão apresentados os dados em que pelo menos duas técnicas e dois clones, entre os analisados, tiveram correspondência entre os dados de “ômicas”.

Dentre as proteínas identificadas em raízes de clones de *C. canephora*, destacam-se dois tipos de quitinases indutoras de xilanases: 13166001 e 13168001, identificadas no GenBank pelo código ADZ48381.1.

Para a quitinase 13166001, é observado que os perfis de expressão do gene que a codifica e de quantificação proteica são correspondentes para a maioria dos clones analisados, conforme representado na Figura 1. O maior valor de expressão para o gene *CcCLXIP*, que codifica a quitinase 13166001, foi apresentado pelo clone 22 na condição NI, entretanto, entre os clones tolerantes, o clone 14 foi o que apresentou maior expressão, sendo que na condição NI a expressão foi induzida em 5x mais que na condição I para o mesmo clone (Figura 1A). O clone 120 também apresentou indução deste gene na condição NI (Figura 1A). Para o clone 73, os valores de expressão deste gene foram relativamente iguais nas duas condições (Figura 1A).

Em relação aos dados de quantificação proteica (Figura 1B), os resultados corroboram com os de expressão gênica, uma vez que a maior quantidade desta proteína também é apresentada no clone 22 na condição NI. Para os outros clones, os perfis também são similares (Figura 1B).

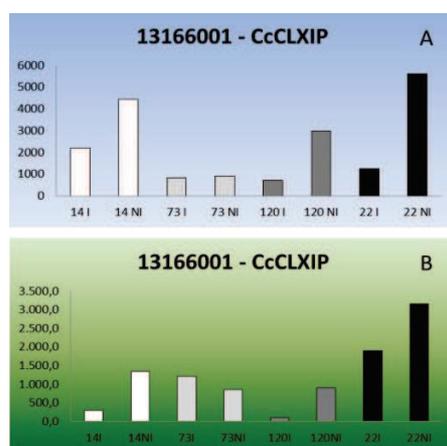


Figura 1 Análise de expressão *in silico* (A) e quantificação proteica (B) da quitinase codificada pelo gene GSCOCT00013166001 – *CcCLXIP*. Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e os da análise *in silico* por RPKM

Observando o perfil de expressão do gene que codifica para a quitinase 13168001 e o perfil de quantificação desta proteína, nota-se que para a maioria dos clones analisados os perfis são correspondentes, como representado na Figura 2. O maior valor de expressão para o gene GSCOCT00013168001, que codifica a quitinase 13168001, foi apresentado pelo clone 22 na condição NI, entretanto, entre os clones tolerantes, os clones 14 e 120 apresentaram valor de expressão similares, sendo induzidos na condição NI, já o clone 73 apresentou valores de expressão, para este gene, relativamente iguais nas duas condições (Figura 2A).

Para os dados de quantificação da proteína quitinase 13168001 (Figura 2B), observa-se que a maior quantidade desta também é apresentada no clone 22 na condição NI. Entre os clones tolerantes, esta proteína foi identificada em maior quantidade nos clones 14 na condição NI e 73 na condição I. O menor valor de quantificação desta proteína foi apresentado pelo clone 120 para as duas condições (Figura 2B).

Para os dados de expressão por qPCR, conforme representado na Figura 2C, é observado que o maior valor de expressão do gene GSCOCT00013168001 é apresentado no clone 14 na condição NI, corroborando com os dados de expressão *in silico* (Figura 2A) e de quantificação proteica (Figura 2C), no que diz respeito à indução de expressão e quantificação, respectivamente.

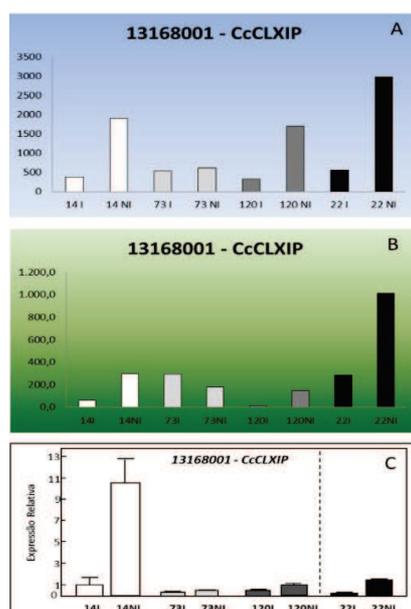


Figura 2 Análise de expressão *in silico* (A), quantificação proteica (B) e expressão gênica por qPCR (C) da quitinase codificada pelo gene GSCOCT00013168001 – *CcCLXIP*. Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e da análise *in silico* por RPKM

Outro grupo de proteínas identificadas em raízes de clones de *C. canephora*, com alta expressão, foi o das fosfatases ácidas 22022001, 22023001 e 21916001, identificadas no GenBank pelo código AFX66999.1. Para a fosfatase ácida 22022001, é observado que o perfil de expressão do gene GSCOCT00022022001, que codifica para esta proteína e o perfil de quantificação proteica são correspondentes para alguns clones entre as três análises (*in silico*, proteômica e qPCR), como apresentado na Figura 3.

Para o perfil de expressão *in silico*, apresentado na Figura 3A, é observado que o maior valor de expressão do gene GSCOCT00022022001, que codifica a quitinase 22022001, foi apresentado para o clone 14 na condição I. Com exceção do clone 73, em todos os outros clones a maior expressão deste gene foi apresentada na condição I (Figura 3A).

Já observando o perfil proteômico, conforme apresentado na Figura 3B, pode-se notar que o maior valor de quantificação também foi apresentado no clone 14 na condição I (Figura 3B), corroborando com os dados de análise de expressão *in silico* (Figura 3A). Para os outros clones tolerantes, os perfis também são similares, a quantificação da proteína é maior na condição NI para o clone 73 e na condição I para o clone 120 (Figura 3B). Em relação ao clone 22, os perfis são totalmente inversos, uma vez que a maior quantidade desta proteína neste clone foi apresentada na condição NI.

Quanto à análise da expressão do gene GSCOCT00022022001 por qPCR (Figura 3C), é observado que houve alta expressão somente no clone 14, sendo esta altamente induzida na condição NI, dados estes totalmente contrários aos da análise *in silico* (Figura 3A), bem como aos de quantificação proteica (Figura 3B).

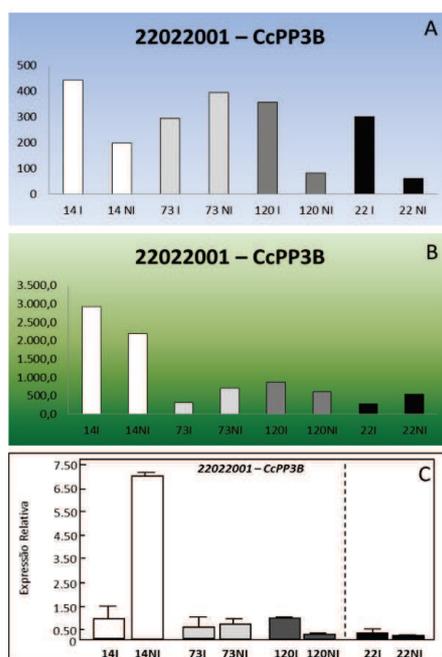


Figura 3 Análise de expressão *in silico* (A), quantificação proteica (B) e expressão gênica por qPCR (C) da fosfatase ácida do tipo 3B, codificada pelo gene GSCOCT00022022001 – *CcPP3B*. Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e da análise *in silico* por RPKM

Para o perfil de expressão *in silico*, apresentado na Figura 4A, é observado que o maior valor de expressão do gene GSCOCT00022023001, que codifica para a quitinase 22023001, foi apresentado no clone 73 na condição I. Em todos os outros clones, a maior expressão deste gene foi apresentada na condição I, sendo que o menor valor de expressão foi apresentado pelo clone 22 (Figura 4A).

Já observando o perfil proteômico, conforme apresentado na Figura 4B, pode-se notar os perfis de expressão do gene e quantificação proteica são semelhantes entre os clones tolerantes (Figura 4B). No clone 22, a maior quantificação desta proteína é apresentada na condição NI. (Figura 4B).

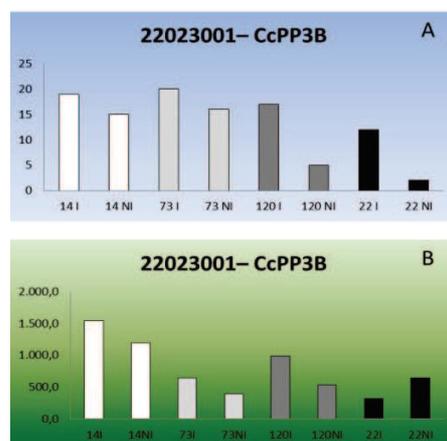


Figura 4 Análise de expressão *in silico* (A) e quantificação proteica (B) da fosfatase ácida do tipo 3B, codificada pelo gene GSCOCT00022023001 – *CcPP3B*. Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e os da análise *in silico* por RPKM

### Resultados de “Ômicas” Divergentes

Na literatura, é muito frequente a não correspondência dos dados de “ômicas”, como já apresentado em muitos trabalhos que utilizam esta abordagem.

Diversos são os fatores que justificam a limitada correspondência entre análises transcriptômicas e de proteômica: como alterações no acúmulo de transcritos e proteínas, que ocorrem de forma independente (AMIOUR et al., 2012), *turnover* diferencial entre proteínas e transcritos além das deficiências inerentes nas técnicas bioinformáticas de análises.

Para a fosfatase ácida 21916001, é observado que o perfil de expressão do gene GSCOCT00021916001, que codifica para esta proteína e o perfil de quantificação desta correspondem apenas no clone 14, para as três análises (*in silico*, proteômica e qPCR), como apresentado na Figura 5.

Para o perfil de expressão *in silico*, apresentado na Figura 5A, é observado que o maior valor de expressão do gene GSCOCT00021916001, que codifica a fosfatase 2191601, foi apresentado para o clone 14 na condição I (Figura 5A). Nos clones tolerantes 73 e 120, a maior expressão deste gene foi apresentada na condição NI (Figura 5A). Para o clone 22 (sensível), o perfil de expressão apresentado assemelha-se ao do clone 14, com indução da expressão neste clone na condição I (Figura 5A).

Já observando o perfil proteômico, conforme apresentado na Figura 5B, pode-se notar que o maior valor de quantificação foi apresentado no clone 73 na condição I (Figura 5B). Valores estes que se assemelham aos apresentados pelo clone 14 para a mesma proteína (Figura 5B). Para os clones 22 e 120, os valores de quantificação foram mais elevados na condição NI (Figura 5B).

Quanto à análise da expressão do gene GSCOCT00021916001 por qPCR (Figura 5C), é observado que o perfil de expressão deste gene para os clones tolerantes 14, 73 e para o clone sensível 22, corroboram com os valores de expressão apresentados na análise *in silico* (Figura 5A), com maior valor de expressão na condição I (Figura 5C). Para o clone 120, o perfil de expressão por qPCR foi diferente aos das outras duas análises (Figura 5C).

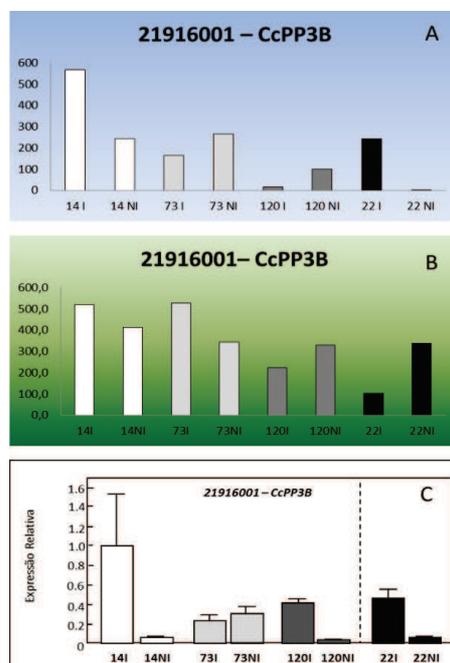


Figura 5 Análise de expressão *in silico* (A), quantificação de proteínas (B) e expressão gênica por qPCR (C) da fosfatase ácida do tipo 3B codificada pelo gene GSCOCT00021916001 - *CcPP3B*. Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e os da expressão *in silico* por RPKM

O maior valor de expressão para o gene GSCOCT00007152001, que codifica a proteína de armazenamento 07152001, foi apresentado pelo clone 14 na condição I (Figura 6A). Para os clones tolerantes 73 e 120, os valores de expressão são similares (Figura 6A). Para o clone 22, a expressão foi induzida na condição NI (Figura 6A).

Em relação aos dados de proteômica (Figura 6B), observa-se que a maior quantidade desta proteína é apresentada no clone 22 na condição I. Entre os clones tolerantes, esta proteína foi identificada em maior quantidade no clone 14 na condição NI e no clone 73 na condição I, no clone 120a maior quantificação foi apresentada na condição NI (Figura 6B).

Para os dados de expressão por qPCR, conforme representado na Figura 6C, é observado que o maior valor de expressão do gene GSCOCT00007152001 é apresentado no clone 14 na condição NI.

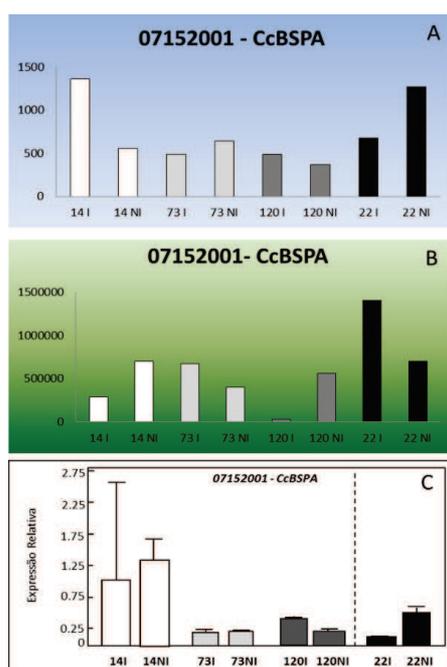


Figura 6 Análise de expressão *in silico* (A) de quantificação proteica (B) e de expressão gênica por qPCR (C) de uma proteína de armazenamento, codificada pelo gene GSCOCT00007152001 - *CcBSPA*. Os valores de quantificação foram padronizados pela quantificação da Ubiquitina e os de expressão *in silico* por RPKM

Observando os dados de expressão (*in silico*) para o gene GSCOCT00007291001 e de quantificação proteica para a respectiva proteína (Kirola), nota-se que apenas para o clone 14 os perfis não se correlacionam, como representado na Figura 7.

Para os valores de expressão *in silico* do gene GSCOCT00007291001, o maior valor é apresentado pelo clone 14 na condição I (Figura 7A). Para o clone

73 e para os clones 22 e 120, o maior valor de expressão é na condição NI (Figura 7A).

Em relação aos dados de proteômica (Figura 7B), observa-se que a maior quantidade da proteína 07291001 – *CcKirola* é apresentada no clone 14 na condição I. No clone 72, a maior quantificação é apresentada na condição I (Figuras 7B). Nos clones 120 e 22, a maior quantificação foi na condição NI (Figura 7B).

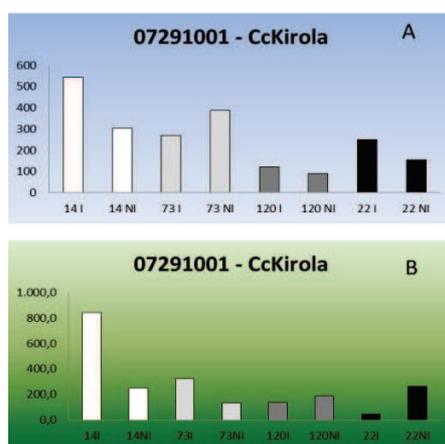


Figura 7 Análise de expressão *in silico* (A) e quantificação (B) da proteína Kirola, codificada pelo gene GSCOCT00007291001 – *CcKirola*. Os valores de quantificação foram padronizados pela quantificação da Ubiquitina e da análise *in silico* por RPKM

Observando os dados de expressão (*in silico* e qPCR) para o gene GSCOCT00005696001 e de quantificação proteica para a respectiva proteína (Catecol oxidase 1), nota-se que, para a maioria dos clones, os perfis não se correlacionam, conforme representado na Figura 8.

Para os valores de expressão *in silico* do gene GSCOCT00005696001, o maior valor de expressão é apresentado pelo clone 14 na condição NI (Figura 8A). Para os clones tolerantes 73 e 120, os valores de expressão também são

induzidos na condição NI (Figura 8A). Para o clone 22, a expressão foi maior na condição I (Figura 8A).

Em relação aos dados de proteômica (Figura 8B), observa-se que a maior quantidade desta proteína é apresentada no clone 14 na condição I. Nos outros clones, a quantificação é maior na condição NI (Figura 8B). Para os dados de expressão por qPCR, conforme representado na Figura 8C, é observado que o maior valor de expressão do gene GSCOCT00005696001 é apresentado no clone 14 na condição NI (Figura 8C).

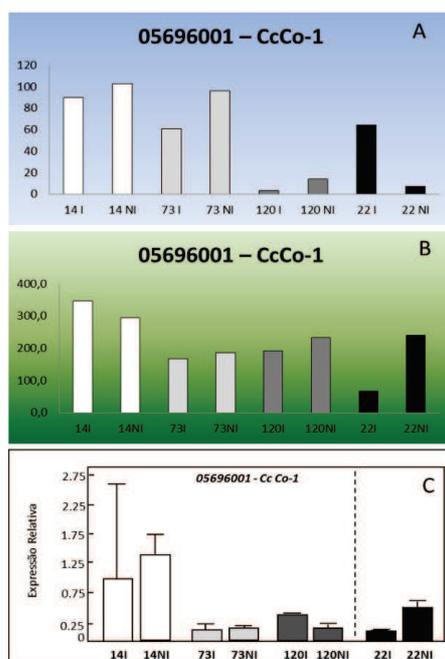


Figura 8 Análise de expressão *in silico* (A), quantificação proteica (B) e expressão gênica por qPCR (C) da catecol oxidase, codificada pelo gene GSCOCT00005696001 – *CcCo-1*. Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e os valores de expressão *in silico* por RPKM

Observando os dados expressão *in silico* para o gene GSCOCT00038191001 e de quantificação proteica para a respectiva proteína 38191001 (Gliceraldeído-3-fosfato-desidrogenase), nota-se que para a maioria dos clones os perfis não se correlacionam, conforme é representado na Figura 9.

Para os valores de expressão *in silico* do gene GSCOCT00039191001, o maior valor de expressão é apresentado pelo clone 14 na condição I (Figura 9A). Para os clones tolerantes 22 e 120, os valores de expressão também são maiores na condição I. Para o clone 73, a expressão foi maior na condição NI (Figura 9A).

Em relação aos dados de proteômica (Figura 9B), observa-se que a maior quantidade desta proteína é apresentada no clone 14 na condição I, correspondendo com os dados de transcriptômica (Figura 9A). No clone 73, a maior quantificação foi apresentada também na condição I (Figura 9B).

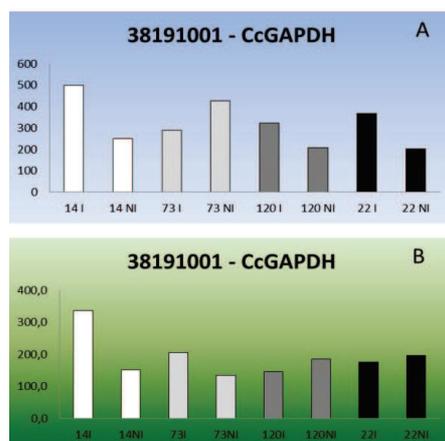


Figura 9 Análise de expressão *in silico* (A) e quantificação (B) da Gliceraldeído-3-fosfato desidrogenase, codificada pelo gene GSCOCT00038191001001 – *CcGAPDH*. Os valores de quantificação foram padronizados com a quantificação da Ubiquitina e os valores de expressão *in silico* foram por RPKM

A proteína predita 3100901 não apresentou valor de expressão *in silico*, portanto ela não será apresentada aqui neste capítulo, uma vez que não haverá possibilidade de comparar seus resultados para as análises de proteômica com outro tipo de análise.

Comparando os resultados apresentados neste trabalho, oriundos das diferentes técnicas utilizadas, padrões de resposta diferentes foram obtidos: (i) correlação entre os perfis de transcriptômica e proteômica e (ii) ausência de correlação. Para o segundo caso, a não correspondência entre os níveis de transcritos e proteínas pode evidenciar mecanismo de regulação pós-transcricional e pós-traducional.

Os processos pós-transcricionais afetam a disponibilidade de mRNA, ou seja, a quantidade de mRNAs disponíveis para a tradução pode ser alterada em diferentes etapas do processo de maturação de RNA, transporte para o início da tradução e degradação por RNAi (MAZZUCOTELLI et al., 2008). Em relação ao *splicing* alternativo, por exemplo, dependendo da espécie de planta considerada, de 20-60% dos genes podem ser afetados por esse mecanismo (NER-GAON et al., 2007). No entanto, tal processo não ocorre de maneira aleatória, no mRNA de todos os genes, mas parece ocorrer preferencialmente em mRNAs de certas classes de genes relacionados à transdução de sinal ou codificando enzimas, receptores e fatores transcricionais (NER-GAON; FLUHR, 2006).

Em condições de estresse, a regulação da tradução dos mRNAs é necessária para conservar energia, uma vez que, a síntese de proteínas é um processo com intenso gasto energético (BRANCO-PRICE et al., 2005, 2008). Em milho, muitos dos mRNAs que acumularam em condições de privação de oxigênio não foram traduzidos em proteínas devido à tradução seletiva de mRNAs (FENNOY; BAILEY-SERRES, 1995).

Além disso, em nível pós-traducional, a cinética de *turnover* de proteínas são processos regulatórios celulares críticos que permitem às plantas responderem rapidamente às diferentes condições ambientais, modificando os níveis de proteínas-chave (YANG et al., 2010).

A não correspondência entre os perfis transcriptômicos e proteômicos é esperada e já foi apresentada em trabalhos anteriores em cafeeiro. No trabalho de Marraccini et al. (2011), foi observado que para os genes que codificam a subunidade menor da ribulose 1,5-bis-fosfato carboxilase (Rubisco) em folhas dos clones 14 e 22 de *C. canephora*, os dados de quantificação da proteína, e de expressão do gene em qPCR não se correlacionam. Em trabalho posterior, Marraccini et al. (2012) observaram que para os genes candidatos *CcPP2C*, que codifica fosfatase tipo 2, e os genes *CcPSBP*, *CcPSBQ*, *CcPSBO* que codificam subunidades do fotossistema III também não houve correspondência entre os perfis.

#### 4 CONCLUSÕES

A capacidade dos clones de *C. canephora* de responder ao déficit hídrico foi evidenciada neste trabalho por meio da diferença de expressão gênica, bem como, nos níveis de proteínas entre as condições avaliadas (I e NI).

Em linhas gerais foi possível observar, que, na maioria dos casos, há correspondência entre os resultados das técnicas “ômicas” utilizadas neste trabalho.

Entretanto, apesar de alguns resultados não corresponderem, o aprimoramento das técnicas de identificação proteica, por meio da utilização de sequências genótipo-específicas, pode ser uma alternativa para futuras análises integradas, uma vez que permite a compreensão do mecanismo de regulação e dos pontos de controle do fluxo da informação gênica.

## REFERÊNCIAS

AMIOUR, N. et al. The use of metabolomics integrated with transcriptomic and proteomic studies for identifying key steps involved in the control of nitrogen metabolism in crops such as maize. **Journal of Experimental Botany**, Oxford, v. 63, n. 14, p. 5017-5033, Sept. 2012.

BARSALOBRES-CAVALLARI, C. F. et al. Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. **BMC Molecular Biology**, London, v. 10, n. 1, p. 1-11, Jan. 2009.

BRANCO-PRICE, C. et al. Genome-wide analysis of transcript abundance and translation in arabidopsis seedlings subjected to oxygen deprivation. **Annals of Botany**, London, v. 96, n. 4, p. 647-660, Aug. 2005.

BRANCO-PRICE, C. et al. Selective mRNA translation coordinates energetic and metabolic adjustments to cellular oxygen deprivation and reoxygenation in *Arabidopsis thaliana*. **The Plant Journal**, Oxford, v. 56, n. 5, p. 743-755, Dec. 2008.

DEBNATH, M.; PANDEY, M.; BISEN, P. S. An Omics approach to understand the plant abiotic stress. **OMICS A Journal of Integrative Biology**, Tsukuba, v. 15, n. 11, p. 739-762, Nov. 2011.

FENNOY, S. L.; BAILEY-SERRES, J. Post-transcriptional regulation of gene expression in oxygen-deprived roots of maize. **The Plant Journal**, Oxford, v. 7, n. 2, p. 287-295, Feb. 1995.

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

HAKEEM, K. et al. Relevance of proteomic investigations in plant abiotic stress physiology. **Proteomics and Plant Abiotic Stress**, Weinheim, v. 16, n. 11, p. 621-635, Nov. 2012.

KOMATSU, S.; HIRAGA, S.; YANAGAWA, Y. Proteomics techniques for the development of flood tolerant crops. **Journal of Proteome Research**, Washington, v. 11, n. 1, p. 68-78, Oct. 2012.

KOMATSU, S.; SHIRASAKA, N.; SAKATA, K. 'Omics' techniques for identifying flooding-response mechanisms in soybean. **Journal of Proteomics**, New York, v. 93, n. 1, p. 169-178, Nov. 2013.

LIU, T. et al. Identification of proteins regulated by ABA in response to combined drought and heat stress in maize roots. **Acta Physiologiae Plantarum**, Copenhagen, v. 35, n. 2, p. 501-513, Feb. 2003.

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

MARRACCINI, P. et al. *RBCS1* expression in coffee: *Coffea* orthologs, *Coffea arabica* homeologs, and expression variability between genotypes and under drought stress. **BMC Plant Biology**, London, v. 11, n. 85, p. 1-23, May 2011.

MAZZUCOTELLI, A. Abiotic stress response in plants: when post-transcriptional and post-translational regulations control transcription. **Plant Science**, Shannon, v. 174, n. 4, p. 420-431, Apr. 2008.

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

NER-GAON, H. et al. Comparative cross-species alternative splicing in plants. **Plant Physiology**, Bethesda, v. 144, n. 3, p. 1632-1641, July 2007.

NER-GAON, H.; FLUHR, R. Whole-genome microarray in Arabidopsis facilitates global analysis of retained introns. **DNA Research**, Oxford, v. 13, n. 3, p. 111-121, June 2006.

SALAVATI, A. et al. Analysis of proteomic changes in roots of soybean seedlings during recovery after flooding. **Journal of Proteomics**, New York, v. 75, n. 3, p. 878-893, Jan. 2012.

VIEIRA, N. G. et al. Different molecular mechanisms account for drought tolerance in *Coffea canephora* var. Conilon. **Tropical Plant Biology**, New York, v. 6, n. 4, p. 181-190, Dec. 2013.

YANG, X. et al. Measuring the turnover rates of Arabidopsis proteins using deuterium oxide: an auxin signaling case study. **The Plant Journal**, Oxford, v. 63, n. 4, p. 680-695, Aug. 2010.

## CONCLUSÃO GERAL

Os resultados apresentados no presente trabalho possibilitaram uma melhor compreensão de alguns mecanismos de resposta em raízes em clones *C. canephora* quanto a tolerância ao déficit hídrico.

Além disso foi possível verificar que os mecanismos de resposta ao estresse entre os clones tolerantes são diferentes.

Da mesma forma que nos trabalhos anteriores utilizando folhas, também é possível identificar vários genes de tolerância à seca em raízes de cafeeiro, complementando os trabalhos anteriores.