

EXPRESSÃO ALELO ESPECÍFICA EM *Coffea canephora* ASSOCIADOS À RESISTÊNCIA A NEMATOIDES¹

Jessica Coelho Valeriano²; Edriana Araújo de Lima³; Sinara Oliveira de Aquino⁴; Fernanda de Araújo Carneiro⁵; Wendel William de Jesus Carneiro⁶; Nathália Gomes Mattos⁷; Regina Maria Dechechi Gomes Carneiro⁸; Alan Carvalho Andrade⁹

¹Trabalho financiado pelo Consórcio Pesquisa Café e INCT-Café (CNPq/FAPEMIG).

²Doutoranda em Biotecnologia Vegetal, UFLA, Lavras-MG, Brasil, jessicacoelho_bio@hotmail.com

³Doutora em Fitopatologia, Universidade de Brasília - UNB, DF, edrianaal@gmail.com

⁴Bolsista Pós-Doutorado PNP/CAPE, PhD, UFLA, Lavras-MG, Brasil, saquinobiotec@gmail.com

⁵Doutora em Biotecnologia Vegetal, UFLA, Lavras-MG, Brasil, fearca14@gmail.com

⁶Bolsista Consórcio Pesquisa Café, wendelwill11@gmail.com

⁷Doutoranda em Biotecnologia Vegetal, UFLA, Lavras-MG, Brasil, nagm7@hotmail.com

⁸Pesquisadora, PhD, Embrapa Recursos Genéticos e Biotecnologia, Brasília-DF, regina.carneiro@embrapa.br

⁹Pesquisador, PhD, Embrapa Café/INOVACAFÉ, Lavras - MG, alan.andrade@embrapa.br

RESUMO: O café representa uma das principais commodities do mundo, sendo o Brasil considerado o maior produtor mundial de *C. arabica* e o segundo maior produtor de *C. canephora*. Entretanto, o parasitismo por nematoides limita o cultivo do cafeeiro, sendo considerado um dos principais fatores responsáveis por perdas na produtividade das lavouras cafeeiras. Dentre os nematoides, o gênero *Meloidogyne* spp. destaca-se como um dos principais, já que possui uma ampla disseminação causando inúmeros prejuízos. A utilização de técnicas de biologia molecular e análise genômica permitem a identificação e caracterização de genes candidatos à resistência aos nematoides no café e oferecem uma alternativa para o controle de doenças, pois esses genes podem ser utilizados como marcadores em programas de melhoramento por seleção assistida. O objetivo do presente trabalho foi identificar a frequência e os modos de variação alélica regulatória através de dados de RNAseq de clones de *C. canephora* previamente caracterizados como resistente e suscetível a infecção por nematoides. Análises comparativas do perfil de expressão gênica de estudos de RNAseq de RNAs extraídos de raízes do clone 14 de *C. canephora* (resistente) e do clone 22 (suscetível) infectadas e não infectadas por *M. paranaensis* foram realizadas. A partir dos resultados das análises de RNAseq, genes com domínio NBS-LRR foram identificados como possíveis candidatos. A expressão alélica específica foi calculada por meio do ASEReadCounter (GATK). Foi possível identificar a partir das análises de expressão alélica específica que 74 genes candidatos apresentam especificidade alélica significativa, estes genes já foram descritos em trabalhos anteriores como relacionados à resposta a estresses bióticos.

PALAVRAS-CHAVE: *Coffea canephora*, *Meloidogyne*, RNAseq.

ALLELE SPECIFIC EXPRESSION IN *Coffea canephora* ASSOCIATED WITH NEMATOIDES RESISTANCE

ABSTRACT: Coffee is one of the most important agricultural commodity in the world, Brazil is the first producer of *C. arabica* and the second largest producer of *C. canephora*. However, parasitism by nematodes limits coffee cultivation and it has been considered one of the main factors responsible for losses in coffee crop yields. Among the nematodes, the genus *Meloidogyne* spp. stands out as one of the main, since it has widespread dissemination causing numerous damages. The use of molecular biology and genomic analysis techniques allow the identification and characterization of nematode resistance candidate genes in coffee and offer an alternative for disease control because these genes can be used as markers in assisted selection breeding programs. The goal of the present study was to identify the frequency and the modes of variation of allelic regulation through RNAseq data from *C. canephora* clones previously characterized as resistant and susceptible to nematode infection. Comparative analyzes of the gene expression profile from RNAseq data of the RNAs extracted from roots of clone 14 (resistant) and clone 22 (susceptible) inoculated and not inoculated by *M. paranaensis* were performed. Data analyzes showed that genes with NBS-LRR domain were identified as possible candidates. Specific allelic expression was calculated using ASEReadCounter (GATK). It was possible to identify from the analysis of specific allelic expression that 74 candidate genes have significant allelic specificity. These genes have been described in previous studies as related to the response to abiotic and biotic stresses.

KEYWORDS: *Coffea canephora*, *Meloidogyne*, RNAseq.

INTRODUÇÃO

O café é considerado a segunda commodity mais negociada no mundo, depois do petróleo, com um grande impacto econômico em muitos países, especialmente na América do Sul (CAMPOS; PANIS; CARPENTIER, 2017). No entanto, alguns fatores que podem comprometer o rendimento desta cultura, principalmente aqueles relacionados aos aspectos fitossanitários, com destaque ao ataque de nematoides (OLIVEIRA e ROSA, 2018). Estima-se que a ocorrência destes microrganismos possa reduzir até 15% da produção mundial e 20% da produção brasileira (ITO et al., 2008).

Dentre os nematoides fitoparasitas que afetam o cafeeiro no Brasil, o *M. paranaensis* se destaca das demais espécies pela sua agressividade e forte dano ao sistema radicular do cafeeiro, com elevado grau de depauperamento das plantas. As raízes parasitadas por *M. paranaensis* apresentam descascamento e rachaduras, com alguns pontos de engrossamento que se mostram lesões do tipo “cancro” e descorticação (CASTRO et al., 2008). O desenvolvimento de cultivares resistentes a nematoides oferece uma alternativa de controle eficiente e ambientalmente segura. As plantas de *C. canephora* destacam-se por apresentar uma alta variabilidade genética, rusticidade e tolerância a várias doenças. Visando a exploração dessas características a espécie passou a ser alvo em programas de melhoramento genético, utilizadas como porta-enxertos e em cruzamentos interespecíficos com *C. arabica* (GARAVITO et al. 2016).

Os marcadores moleculares identificados pela genômica permitem a seleção de genótipos sem a necessidade de cruzamentos teste. Dos quais destacam-se os SNPs, úteis no desenvolvimento de mapas de ligação gênica e na genotipagem, ao revelar sítios polimórficos e suas localidades no genoma (OLIVER et al., 2011). Além disso, abordagens transcriptômicas têm sido empregadas na caracterização de mudanças na expressão global de genes em resposta a patógenos (SANA et al., 2010). A expressão gênica é um processo complexo que é regulado por variações genéticas e epigenéticas em resposta a estímulos desenvolvimentais e ambientais e que dá origem a fenótipos dinâmicos. Diferenças hereditárias na expressão gênica, denominada expressão específica de alelo (ASE), entre indivíduos são causadas por polimorfismos que afetam o nível de expressão de um ou ambos os alelos em um organismo diploide. Uma abordagem poderosa para a identificação da variação reguladora da ação é a medição da expressão específica do alelo (ASE) (ZHUO, LAMONT e ABASHT, 2017; SHAO et al., 2019). A diferença de expressão causada pelo ASE pode levar a variações fenotípicas dependendo da função dos genes (SHAO et al., 2019).

A ASE está se tornando uma ferramenta cada vez mais importante, pois quantifica a variação de expressão entre os dois haplótipos de um indivíduo diploide, distinguido por locais heterozigotos (CASTEL et al., 2015).

O objetivo do presente trabalho foi identificar a frequência e os modos de variação regulatória alélica através de dados de RNAseq de clones de *C. canephora* previamente caracterizados como resistente e suscetível a infecção por nematoides.

MATERIAL E MÉTODOS

Para o presente estudo, foram utilizados dados de sequenciamento de RNAs extraídos de raízes dos clones 14 (resistente ao *M. paranaensis*) e 22 (suscetível ao *M. paranaensis*) desenvolvidos pelo Instituto Capixaba de Pesquisa e Extensão Rural (INCAPER). Os RNAs foram extraídos das raízes coletadas em diferentes tempos após a inoculação do nematoide *M. paranaensis*: 4, 8, 12, 20, 32 e 45 dias após a infecção (DAI). Os RNAs foram sequenciados por meio da técnica Illumina (HiSeq 2000) e foram construídas 24 bibliotecas *paired-end* (6 coletas x 2 clones x inoculada e não inoculada). A qualidade das sequências geradas foi verificada pelo software FastQC. Como nos casos das amostras inoculadas, também existia RNA de *M. paranaensis*, também foi possível se separar e analisar os dados de RNAseq de nematoide nos diferentes tempos e tratamentos. Para essa separação foram utilizados os programas FastQScreen e Xenome (CONWAY et al., 2012). O mapeamento foi executado pelo algoritmo STAR usando o genoma de um duplo haploide de *C. canephora* como referência (DENOEU et al., 2014). A identificação da expressão diferencial dos genes foi realizada por meio do pacote estatístico Deseq2, utilizando a contagem dos *reads* brutos obtidos no mapeamento de cada biblioteca. A chamada de SNPs foi realizada utilizando a ferramenta *HaplotypeCaller* do GATK v3.6 (*Genotyper in the Genome Analyses Toolkit*). Uma análise preliminar da expressão alélica específica foi realizada por meio da ferramenta do GATK, *ASEReadCounter*, usando marcadores bialélicos heterozigotos selecionados a partir dos dados obtidos pelo *HaplotypeCaller*.

RESULTADOS E DISCUSSÃO

Baseado nas análises do programa FASTQC, observou-se que todas as bibliotecas analisadas apresentaram ótimos índices de qualidade. O programa permitiu extrair relatórios dos conjuntos de sequências a partir dos dados brutos, gerando gráficos da distribuição das sequências levando em consideração a qualidade média total por base, a qualidade média de todos os *reads* e o conteúdo G + C dos dados.

Após o mapeamento dos *reads* no genoma de referência de *C. canephora* com o software STAR foram gerados, somando-se todas as bibliotecas, um número total de 1.000.287.108 *reads* e uma média geral de 41.678.629,5 *reads*, sendo estes distribuídos de maneira similar por biblioteca, indicando uma boa preparação das mesmas.

Os estudos de RNAseq realizados demonstraram que, as amostras que tiveram maior variação na expressão gênica foram os tratamentos do clone 14 (8dai-X8RNA4) e clone 22 (20 dai-X20RNA3), conforme apresentado na Figura 1.

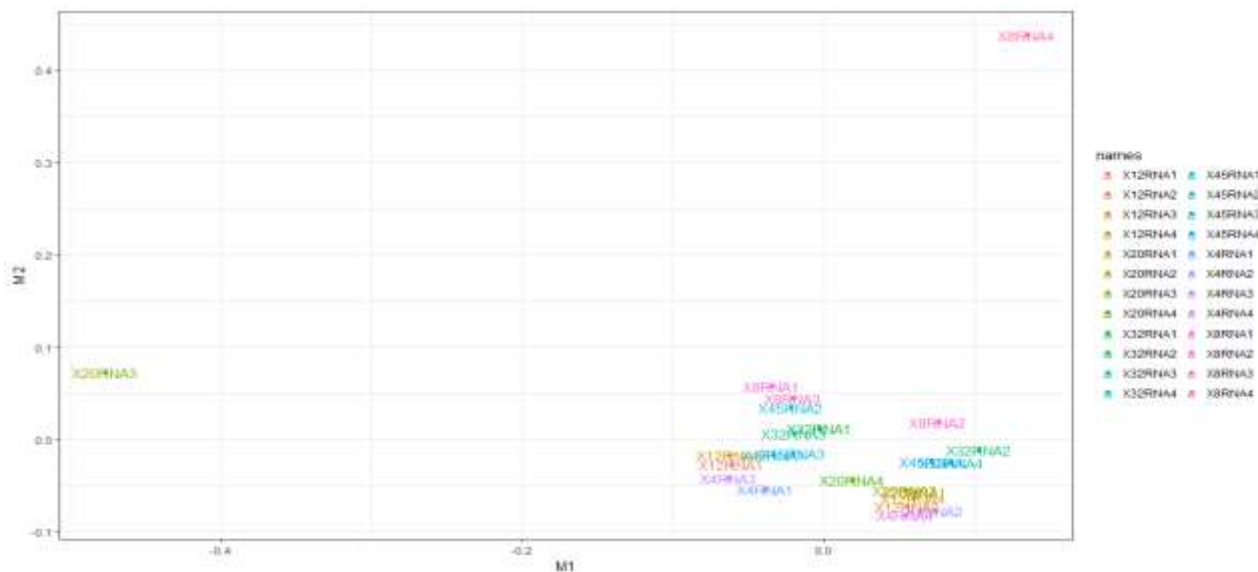


Figura 1. Análise de componentes principais das amostras de RNAseq analisadas.

A partir de uma análise dos padrões de expressão gênica, dos dados analisados, pode se observar grupos de genes respondendo de maneira similar em função dos tratamentos (Figura 2). Nota-se claramente, pelos dados apresentados na figura 2 que os tratamentos com maior alteração nos padrões de expressão gênica foram novamente os tratamentos do clone 14 (8dai-X8RNA4) e clone 22 (20dai-X20RNA3). Nos grupos 1, 4 e 6, observa-se um conjunto de genes que tiveram uma clara redução na expressão gênica. No grupo 1, genes que tiveram redução na expressão especificamente no tratamento do clone 22 (20 dai), no grupo 4 genes que apresentaram significativa redução na expressão gênica no tratamento do clone 14 (8dai) e no grupo 6, um conjunto de genes com redução significativa em ambos os tratamentos (clone 14-8dai e clone 22-20 dai). Por outro lado, chama a atenção, o grupo 3 (Figura 2) com um conjunto de genes com aumento significativo da expressão nos mesmos tratamentos (14-8dai e 22-20dai).

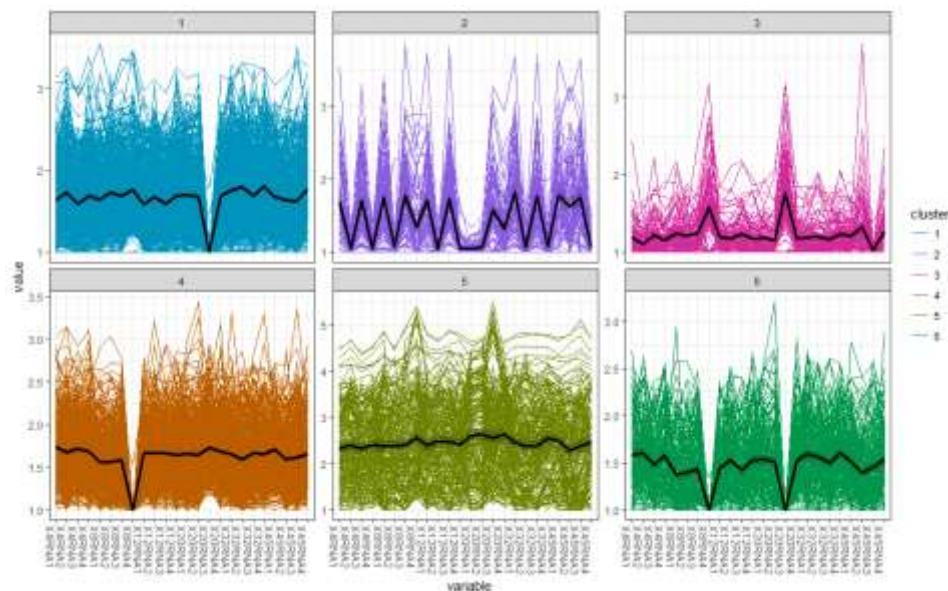


Figura 2. Padrões de alterações na expressão gênica das amostras de RNAseq analisadas. As amostras do clone 22 (1- controle e 3-inoculadas com *M. paranaensis*) e clone 14 (2-controle e 4- inoculadas com *M. paranaensis*). Para ambos os clones, os tempos 4, 8, 12, 20, 32 e 45 com (3 e 4) ou sem (1 e 2) inoculação de *M. paranaensis*.

Uma análise mais detalhada dos genes do grupo 3 (Figura 2), por meio do “GO enrichment” resultou na constatação de que a maior parte desses, estão relacionadas com os processos de defesa das plantas ao ataque de patógenos (Figura 3).

Além disto, vários desses genes são genes de Resistência (R genes) já descritos como importantes determinantes moleculares no reconhecimento e início das repostas de defesa das plantas ao ataque de patógenos. Os genes NBS-LRR representam uma grande família de genes de resistência (genes R) em plantas. Após o reconhecimento específico, os produtos dos genes R iniciam um conjunto de defesas ativas como morte celular localizada ou uma resposta de hipersensibilidade (HR) no local da infecção. O gene *Mi* do tomateiro é um exemplo de gene dessa família que está relacionado com mecanismos de defesa contra algumas espécies de nematoides do gênero *Meloidogyne* (HWANG e WILLIAMSON, 2003).

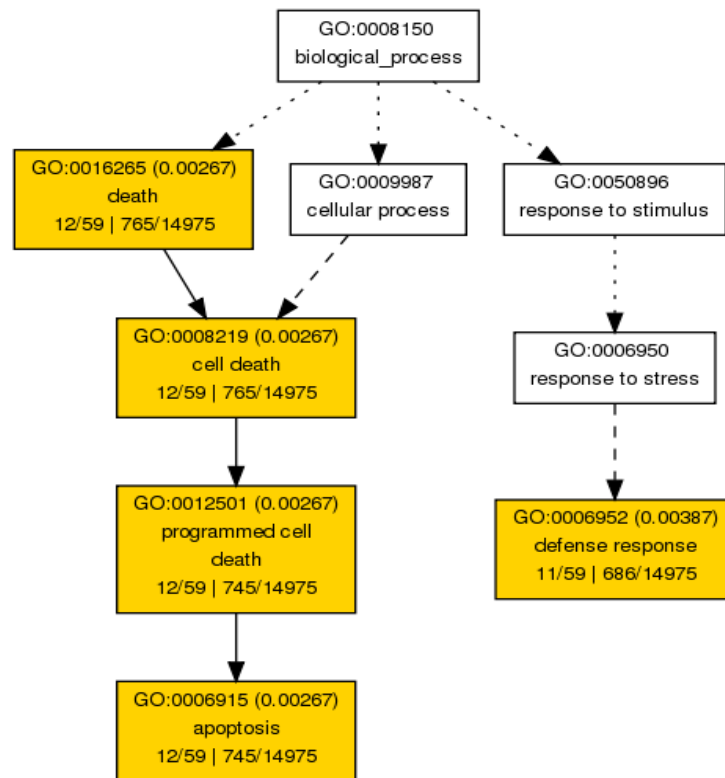


Figura 3. Função molecular e biológica dos genes categorizados no Grupo 3 (Figura 2) de padrão de resposta à inoculação de raízes de café com *M. parananensis*.

O padrão de expressão de alguns desses R-genes identificados neste trabalho, são apresentados na figura 4. Pode-se observar um aumento significativo da expressão desses genes nos tratamentos clone 14 (8dai) e clone 22 (20 dai).

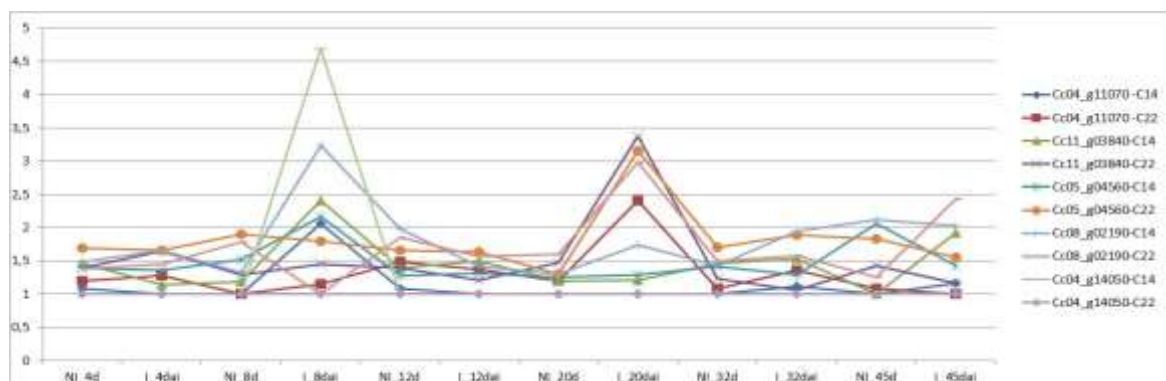


Figura 4. Genes de resistência NBS-LRR apresentando níveis de expressão mais altos aos 8 DAI no clone 14 inoculado e aos 20 DAI no clone 22 inoculado.

Esse resultado corrobora com as análises histológicas feitas por Lima et al. (2015), que observou uma forte reação de hipersensibilidade (RH) no clone 14 inoculado logo no início do período de infecção pelo nematoide. A reação precoce de resistência RH bloqueou o nematoide impedindo sua migração através do tecido cortical e chegada ao cilindro central e consequentemente, o seu desenvolvimento e multiplicação. Durante a migração, os J2 podem ficar retidos

devido à morte das células circundantes e a RH (ANTHONY et al., 2005). Também foram observados diversos genes “no hits” cuja função é desconhecida, mas que apresentaram altos níveis de expressão indicando que essa classe de genes pode estar associada à mecanismos de defesa, como já foi descrito por Vieira et al. (2012) que observaram “no hits” diferencialmente expressos em clones de *C. canephora* tolerantes e sensíveis à seca.

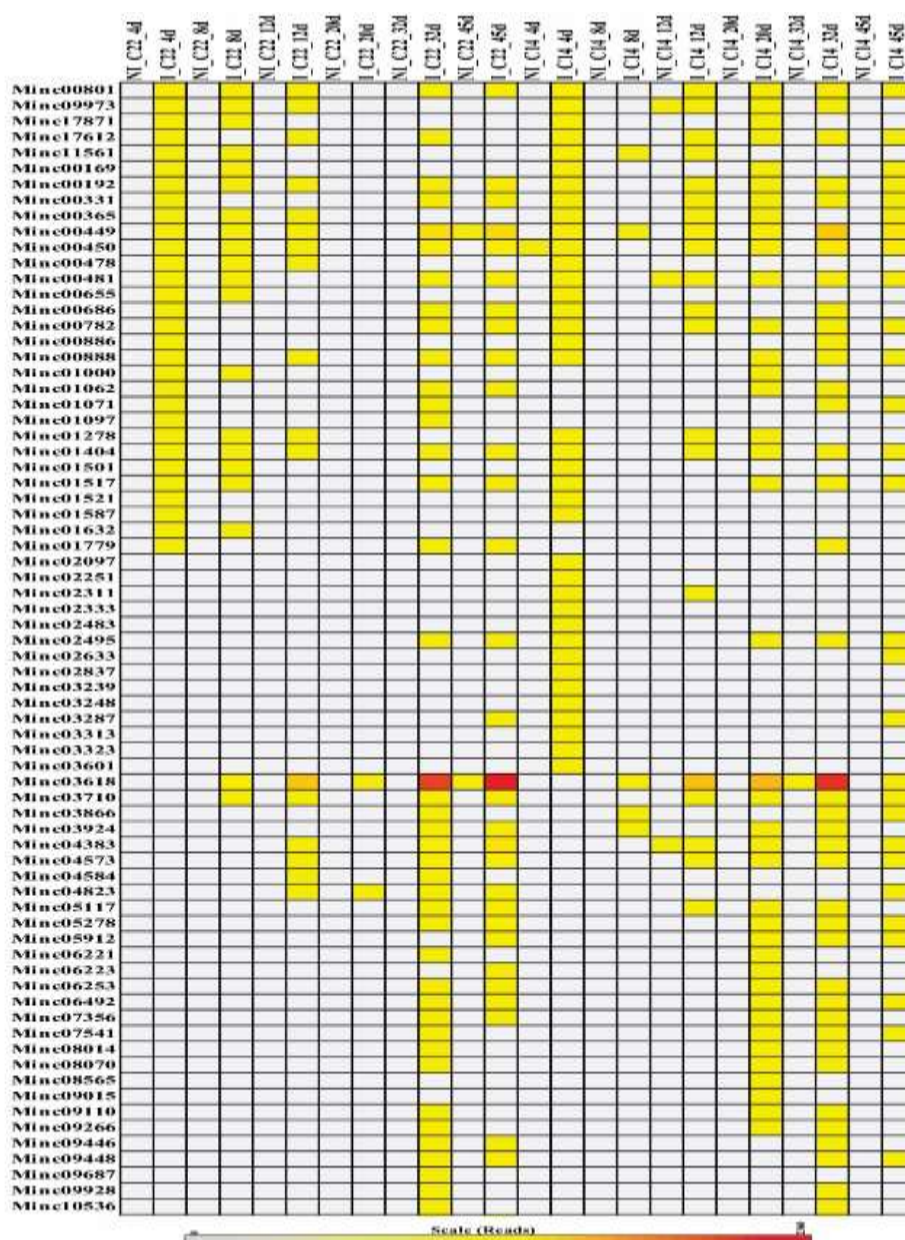


Figura 5. Heat-map de análises *in silico* com potenciais efetores de *M. paranaensis* nos diferentes tratamentos.

Por outro lado, também foi possível acompanhar a expressão gênica de nematoides nos tratamentos com inoculação nos dois clones testados (14 vs 22). Pode-se observar que vários potenciais efetores de nematoides também puderam ser identificados com trabalho realizado (Figura 5). Os resultados de RT-qPCR desses genes estão em andamento. Os resultados obtidos demonstram que o clone 14 pode ser importante fonte de genes relacionados com a resistência à *Meloidogyne* spp. e que esses genes devem ser estudados mais profundamente no futuro a fim de obter maior conhecimento sobre os mecanismos moleculares envolvidos no processo de resistência aos nematoides das galhas e os possíveis pontos de intersecção entre os estresses bióticos e abióticos. Com o objetivo de complementar os estudos foi realizada uma análise preliminar de alelo especificidade que podem agregar informações a respeito de polimorfismos que afetam o nível de expressão de um ou ambos os alelos, modificando as respostas fenotípicas. Os marcadores SNPs foram identificados a partir dos dados de RNAseq das bibliotecas do clone 14 com 8 DAI e controle, por meio do *pipeline* GATK. Identificou-se um total de 309.652 SNPs, destes, 19.462 são bialélicos, heterozigotos e de alta qualidade com uma cobertura mínima de dez vezes. Foi calculado a expressão alélica específica por meio do *ASEReadCounter* (GATK). Para eliminar qualquer erro de mapeamento dos *reads*, o que poderia levar a erros de

inferência alélica na abundância de transcritos, removeu-se todos os *reads* potencialmente duplicados por PCR através da ferramenta *Picard Tools*. Obtivemos contagem de *reads* específicos para alelos de aproximadamente 14 mil regiões transcritas. Dentre estas, 74 genes candidatos já descritos como relacionados a resposta à estresses abióticos e bióticos, foram identificados com especificidade alélica significativa. Entretanto, as análises com todas as 24 bibliotecas ainda estão em andamento. Os resultados, com a análise ASE, poderão fornecer informações úteis sobre como os padrões da variação alélica dos genes podem estar correlacionados com as respostas ao estresse biótico em *C. canephora*. A estratégia de análise ASE se apresenta como uma abordagem para identificar SNPs e genes candidatos que exibem resposta alélica diferencial à infecção ao nematoide.

CONCLUSÕES

1. Genes de resistência com domínio rico em leucina (LRR) tiveram expressão diferencial detectada em ensaio *in silico*, bem como em análises de RT-qPCR, apresentando-se como genes candidatos a resistência ao nematoide *M. paranaensis* no clone 14.
2. análise preliminar de expressão alélica diferencial no clone 14, no período de 8 dias após a inoculação e sem inocular, permitiu a identificação de 74 genes relacionados a respostas a estresses bióticos, com expressão alélica específica significativa.
3. Análises de expressão alélica diferencial em todos os tempos de inoculação e no controle dos clones 14 (resistente) e 22 (susceptível) estão em andamento, estas análises objetivam a seleção e o desenvolvimento de marcadores moleculares para identificação de alelos com variação regulatória, com aplicação potencial na seleção assistida.

AGRADECIMENTOS

Ao Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural (INCAPER) pelo provimento dos clones de *C. canephora* utilizados neste estudo. A UFLA por ceder a estrutura para o desenvolvimento dos projetos e a CAPES pelo apoio financeiro.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANTHONY, François et al. Hypersensitive- like reaction conferred by the *Mex- 1* resistance gene against *Meloidogyne exigua* in coffee. **Plant Pathology**, v. 54, n. 4, p. 476-482, 2005.
- CASTEL, STEPHANE E. et al. Tools and best practices for data processing in allelic expression analysis. **Genome biology**, v. 16, n. 1, p. 195, 2015.
- CASTRO, J. M. C. et al. Levantamento de fitonematóides em cafezais do sul de Minas Gerais. **Nematologia Brasileira**, Piracicaba, v. 32, p. 56-64, 2008.
- CAMPOS, N. A.; PANIS, B.; CARPENTIER, S. C. Somatic embryogenesis in coffee: The evolution of biotechnology and the integration of omics technologies offer great opportunities. **Frontiers in Plant Science**, v. 8, n. August, p. 1–12, 2017.
- CONWAY, T. et al. Xenome—a tool for classifying reads from xenograft samples. **Bioinformatics**, v. 28, n. 12, p. i172-i178, 2012.
- DENOEU, France et al. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. **science**, v. 345, n. 6201, p. 1181-1184, 2014.
- GARAVITO, A. et al. Identification by the DArTseq method of the genetic origin of the *Coffea canephora* cultivated in Vietnam and Mexico. **BMC Plant Biology**, v. 16, n. 1, p. 242, 2016.
- Hwang C. F.; Williamson V. M. Leucine-rich repeat-mediated intramolecular interactions in nematode recognition and cell death signaling by the tomato resistance protein Mi. **The Plant Journal** 34:585-593, 2003.
- ITO, D. S. et al. Progenies de café com resistência aos nematoides *Meloidogyne paranaensis* e raça 2 de *Meloidogyne incognita*. **Coffee Science**, Lavras, v. 3, p. 156-163, 2008.
- LIMA, Edriana A. et al. The multi-resistant reaction of drought-tolerant coffee ‘Conilon clone 14’ to *Meloidogyne* spp. and late hypersensitive-like response in *Coffea canephora*. **Phytopathology**, v. 105, n. 6, p. 805-814, 2015.
- SANA, THEODORE R. et al. Metabolomic and transcriptomic analysis of the rice response to the bacterial blight pathogen *Xanthomonas oryzae* pv. *oryzae*. **Metabolomics**, v. 6, n. 3, p. 451-465, 2010.
- OLIVER, R. E. et al. Model SNP development for complex genomes based on hexaploid oat using high-throughput 454 sequencing technology. **BMC Genomics**, v. 12, n. 1, p. 77, 2011.
- OLIVEIRA, C. M. G.; ROSA, J. M. O. **Nematoides parasitos do cafeeiro**. São Paulo: Instituto Biológico. N.32. p. 1-28, 2018.
- SHAO, LIN et al. Patterns of genome-wide allele-specific expression in hybrid rice and the implications on the genetic basis of heterosis. **Proceedings of the National Academy of Sciences**, v. 116, n. 12, p. 5653-5658, 2019.
- ZHUO, ZHU et al. RNA-Seq analyses identify frequent allele specific expression and no evidence of genomic imprinting in specific embryonic tissues of chicken. **Scientific reports**, v. 7, n. 1, p. 11944, 2017.