

ISABEL SAMILA LIMA CASTRO

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE ANÁLOGO DE GENE DE
RESISTÊNCIA (RGA) ENVOLVIDO NA INTERAÇÃO CAFEIEIRO-*Hemileia*
*vastatrix***

Tese apresentada à Universidade Federal de Viçosa,
como parte das exigências do Programa de Pós-
Graduação em Genética e Melhoramento, para
obtenção do título de *Doctor Scientiae*.

Orientador: Eveline Teixeira Caixeta

Coorientadores: Tiago Antônio de Oliveira Mendes
Laércio Zambolim

**VIÇOSA - MINAS GERAIS
2020**

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

T

Castro, Isabel Samila Lima, 1986-
C346i Identificação e caracterização de análogo de gene de
2020 resistência (RGA) envolvido na interação cafeeiro - *Hemileia
vastatrix* / Isabel Samila Lima Castro. – Viçosa, MG, 2020.
82 f. : il. (algumas color.) ; 29 cm.

Inclui apêndices.

Orientador: Eveline Teixeira Caixeta.

Tese (doutorado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. Bioinformática. 2. Ferrugem-do-cafeeiro. 3. Fungos
fitopatogênicos. 4. Relações hospedeiro-parasito. 5. Resistência.
I. Universidade Federal de Viçosa. Departamento de Fitotecnia.
Programa de Pós-Graduação em Genética e Melhoramento.
II. Título.

CDD 22. ed. 570.285

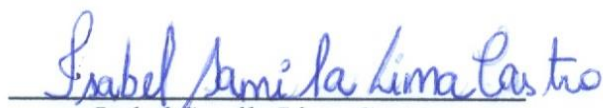
ISABEL SAMILA LIMA CASTRO

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE ANÁLOGO DE GENE DE
RESISTÊNCIA (RGA) ENVOLVIDO NA INTERAÇÃO CAFEIEIRO-*Hemileia
vastatrix***

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

APROVADA: 17 de fevereiro de 2020.

Assentimento:



Isabel Samila Lima Castro

Autora



Eveline Teixeira Caixeta

Orientadora

AGRADECIMENTOS

A Deus, por ter concedido tantas oportunidades e sempre ter me dado força e sabedoria. Obrigada, Senhor!

À Universidade Federal de Viçosa, pela qualidade do ensino.

Ao Instituto de Biotecnologia Aplicada à Agropecuária (Bioagro), pela infraestrutura disponibilizada à realização do trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), à FINEP, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Brasil (CAPES), ao Consórcio pesquisa café (CBP&D Café) e ao Instituto Nacional de Ciência e Tecnologia do Café- INTC Café pelo apoio financeiro.

A minha mãe, Penha, pelo amor, por sempre ter confiado em mim, por ser minha base e meu exemplo de mulher forte.

As minhas irmãs, Paula e Bruna, por todo carinho e por toda torcida pelo meu sucesso.

Ao Daniel, que além de esposo é meu amigo e companheiro, pelo amor e pela cumplicidade em todos os momentos, desde a minha graduação.

A meu filho, Benício, que veio ao mundo durante a realização desse trabalho, eu agradeço por estar aqui e espero que um dia entenda a minha ausência em alguns momentos.

Aos amigos de Viçosa, pela companhia e amizade durante todos esses anos de estudos.

A minha Orientadora, Dra. Eveline Teixeira Caixeta, pela confiança em mim depositada para a realização deste trabalho, pelos ensinamentos, pela paciência e amizade.

Ao meu coorientador Laércio Zambolim, por todos os ensinamentos e pela oportunidade de trabalhar em um excelente laboratório, com toda a infraestrutura necessária para a realização de pesquisas de alta qualidade.

Ao meu coorientador Tiago Antônio de Oliveira Mendes pela valiosa ajuda em todas as etapas desse trabalho e por compartilhar todo seu conhecimento de forma tão generosa.

Ao professor Fabrício de Ávila Rodrigues e a Ananda Pereira Aguilar, pela disponibilidade em participar da banca examinadora deste trabalho.

A Dênia, Danúbia e Edson, pela valiosa ajuda em todas as etapas deste trabalho, e a todos os amigos do Biocafé, pela agradável companhia.

RESUMO

CASTRO, Isabel Samila Lima, D.Sc., Universidade Federal de Viçosa, fevereiro de 2020. **Identificação e caracterização de Análogo de Gene de Resistência (RGA) envolvido na interação cafeeiro-*Hemileia vastatrix***. Orientadora: Eveline Teixeira Caixeta. Coorientadores: Tiago Antônio de Oliveira Mendes e Laércio Zambolim.

A ferrugem do cafeeiro, causada pelo fungo biotrófico *Hemileia vastatrix*, é considerada a principal doença dessa cultura, sendo responsável por perdas na ordem de 30 a 50% da produção, caso medidas de prevenção e controle não sejam adotadas. Apesar da eficiência dos fungicidas, o uso de variedades resistentes é a melhor alternativa para o controle da doença por ser eficiente, econômico e sustentável. No entanto, a obtenção de cultivares com resistência durável tem sido um constante desafio para os melhoristas, devido ao surgimento de novas raças do patógeno capazes de suplantar a resistência. Embora o processo infeccioso de *H. vastatrix* seja bem compreendido, informações acerca da interação com o hospedeiro são escassas. Estudos de genômica funcional tem contribuído para o entendimento dos mecanismos moleculares envolvidos na interação planta-patógeno, assim como para a identificação de genes codificadores de proteínas envolvidas nas vias de resposta de defesa da planta. Sendo assim, o objetivo desse trabalho foi a identificação e caracterização de um Análogo de Gene de Resistência (*Resistance Gene Analog - RGA*) potencialmente envolvido na resposta de defesa do cafeeiro à *H. vastatrix* raça XXXIII. Para isso, um *RGA* candidato foi selecionado a partir de dados do transcriptoma da interação cafeeiro-*H. vastatrix* raça XXXIII, com o potencial envolvimento no mecanismo de defesa da planta. A análise de expressão gênica do *RGA* mostrou diferença significativa entre variedades resistente e suscetível. Foi realizado o rastreamento do *RGA* em uma biblioteca de clones BAC (*Bacterial Artificial Chromosome*) obtida do Híbrido de Timor (HdT) CIFC 832/2. O sequenciamento do clone BAC e a sua análise estrutural revelou um loco gênico altamente similar a um loco do genoma de *Coffea arabica*. Foram identificados três *RGAs* e as análises de filogenia utilizando esses genes confirmaram a similaridade dessa região com os genomas disponíveis de *C. arabica*. Para identificar qual dos três *RGAs* possui importância na

resposta de defesa do cafeeiro à *H. vastatrix*, *primers* específicos para cada *RGA* foram construídos e utilizados para amplificação de DNA de cafeeiros diferenciadores de raças de *H. vastatrix* e para análises de expressão gênica em genótipos resistentes e suscetíveis. Os três *RGAs* foram amplificados em todos os cafeeiros diferenciadores de raças. Assim, a diferença entre os genótipos resistentes e suscetíveis poderia estar no nível de expressão desses genes. Análises de expressão mostraram que o *RGA4* possivelmente é o gene que mantém a função envolvida com a resistência do cafeeiro à *H. vastatrix* raça XXXIII. Além disso, os resultados de expressão corroboram com a ideia de resistência pré-haustorial sugerida para a mesma interação.

Palavras-chave: Bioinformática. Ferrugem do cafeeiro. Fungos fitopatogênicos. Relações hospedeiro-parasito. Resistência.

ABSTRACT

CASTRO, Isabel Samila Lima, D.Sc., Universidade Federal de Viçosa, February, 2020. **Identification and characterization of Resistance Gene Analog (RGA) involved in the coffee-*Hemileia vastatrix* interaction.** Adviser: Eveline Teixeira Caixeta. Co-advisers: Tiago Antônio de Oliveira Mendes and Laércio Zambolim.

Coffee rust, caused by biotrophic fungus *Hemileia vastatrix*, is considered the main disease of this crop, being responsible for losses in the order of 30 to 50% of production, if preventive and control measures are not adopted. Despite the effectiveness of fungicides, the use of resistant varieties is the best alternative for disease control because it is efficient, economical and sustainable. However, obtaining cultivars with durable resistance has been a constant challenge for breeders due to the emergence of new pathogen races capable of supplanting resistance. Although the *H. vastatrix* infectious process is well understood, information about host interaction is scarce. Functional genomics studies have contributed to the understanding of the molecular mechanisms involved in plant-pathogen interaction, as well as to the identification of protein coding genes involved in plant defense response pathways. Thus, the objective of this work was the identification and characterization of a Resistance Gene Analog (RGA) potentially involved in the defense response of the coffee to *H. vastatrix* race XXXIII. For this, a candidate RGA was selected from data from the transcriptome interaction coffee-*H. vastatrix* race XXXIII, with the potential involvement in the plant's defense mechanism. The analysis of RGA gene expression showed a significant difference between resistant and susceptible varieties. The RGA was screened in a library of BAC clones (Bacterial Artificial Chromosome) obtained from the Híbrido de Timor (HdT) CIFC 832/2. The sequencing of the BAC clone and its structural analysis revealed a gene locus highly similar to a locus in the *Coffea arabica* genome. Three RGAs were identified and phylogeny analyzes using these genes confirmed the similarity of this region with available *C. arabica* genomes. In order to identify which of the three RGAs has importance in the defense response of coffee to *H. vastatrix*, specific primers for each RGA were built and used to amplify the DNA of differentiating clones of the *H. vastatrix* races and for analysis of gene expression resistant and susceptible genotypes. The three RGAs were amplified in all the differentiating coffee clones. The result indicates that the difference between resistant and susceptible genotypes could be in the level of expression

of these genes. Expression analyzes suggested that *RGA4* is possibly the gene that maintains the function involved with coffee resistance to *H. vastatrix* race XXXIII. In addition, the expression results corroborate the idea of pre-haustorial resistance suggested for the same interaction.

Keywords: Bioinformatics. Coffee rust. Phytopathogenic fungi. Host-parasite relationships. Resistance.

SUMÁRIO

1. INTRODUÇÃO GERAL.....	9
2. REVISÃO BIBLIOGRÁFICA.....	11
2.1 Importância do cafeeiro.....	11
2.2 Ferrugem do cafeeiro - <i>Hemileia vastatrix</i>	11
2.3 Sistema de defesa da planta.....	15
2.4 <i>Resistance Gene Analogs (RGAs)</i>	17
3. REFERÊNCIAS.....	19
CAPÍTULO 1: Análise de Genômica Integrada Revela um Loco Rico em Análogos de Genes de Resistência (RGAs) com Potencial Envolvimento na Resistência do Cafeeiro à Ferrugem (<i>Hemilia vastatrix</i>)	24
RESUMO.....	25
1. INTRODUÇÃO.....	25
2. MATERIAL E MÉTODOS.....	29
3. RESULTADOS.....	35
4. DISCUSSÃO.....	41
5. REFERÊNCIAS.....	45
CAPÍTULO 2: Análise de Expressão Gênica Sugere um RGA Potencialmente Envolvido na Resposta de Resistência do Cafeeiro à Ferrugem	53
RESUMO.....	54
1. INTRODUÇÃO.....	54
2. MATERIAL E MÉTODOS.....	57
3. RESULTADOS.....	63
4. DISCUSSÃO.....	67
5. REFERÊNCIAS.....	70
CONCLUSÕES GERIAS.....	74
APÊNDICES.....	75

1. INTRODUÇÃO GERAL

O café é um dos principais produtos de exportação do Brasil e possui importante papel no agronegócio do país. Segundo dados atuais do Ministério da agricultura, pecuária e abastecimento (MAPA), o Brasil é o maior produtor e exportador mundial e o segundo maior consumidor. O café é fonte imprescindível de receita para centenas de municípios, além de ser o principal gerador de postos de trabalho na agropecuária nacional (MAPA, 2019).

A ferrugem do cafeeiro, causada pelo fungo biotrófico *Hemileia vastatrix*, é considerada a principal doença dessa cultura. A doença comprometer a produtividade de lavouras de variedades suscetíveis (Zambolim, 2016). *H. vastatrix* pertence a classe *Basidiomicete*, parasita obrigatório, podendo ocorrer em várias espécies do gênero *Coffea*, e apenas nesse gênero, não se conhece outra planta hospedeira. O progresso da infecção provoca a redução da área foliar e, gradativamente, ocorre a desfolha das plantas. Os esporos de *H. vastatrix* só germinam na presença de água e a temperatura é fator limitante, sendo que para germinação e penetração na folha a temperatura ótima fica entre 22 e 23 °C (Carvalho et al., 2002).

Apesar da eficiência dos fungicidas, o uso de variedades resistentes é a melhor alternativa para o controle da doença. No entanto, a obtenção de cultivares com resistência durável tem sido um constante desafio para os melhoristas, uma vez que, ao longo do tempo, novas raças do patógeno surgem com a capacidade de suplantação da resistência (Várzea & Marques, 2005). No Brasil, já foram relatadas 15 raças de *H. vastatrix*, sendo a raça II, que apresenta o gene de virulência *v5*, a mais frequente (Chaves e Pereira 1980; Cardoso et al. 1986, 1988; Zambolim et al. 2005; Cabral et al. 2009; Capucho et al. 2012). No entanto, entre as raças encontradas no Brasil, a XXXIII se destaca por ser complexa, apresentando maior número de genes de virulência (*v5,7* ou *v5,7,9*). Além disso, essa raça foi capaz de suplantar a resistência de cultivares lançadas no mercado como resistentes à ferrugem (Várzea & Marques, 2005; Capucho et al., 2012), o que tem causado preocupação para os produtores. Dessa forma, essa raça é de grande interesse para os estudos de interação com o cafeeiro.

Estudos realizados no Centro de Investigações das Ferrugens do Cafeeiro (CIFC), em Portugal, têm mostrado que a teoria gene-a-gene proposta por Flor (1956) também se

aplica à interação *Hemileia-Coffea* (Noronha et al., 1967). Segundo os autores, a interação planta-patógeno é determinada pelo genótipo do hospedeiro e pelo genótipo do patógeno, em uma interação gene a gene. Dessa forma, o fenótipo de resistência só é observado quando o produto do gene de resistência dominante (gene *R*) da planta hospedeira interage com seu respectivo produto do gene de avirulência (gene *Avr*) do genótipo do patógeno. Até o momento, foram relatados nove genes que conferem resistência à *H. vastatrix*. A resistência de cafeeiros às raças fisiológicas de *H. vastatrix* é condicionada por fatores dominantes, aparentemente independentes. Estes genes receberam a sigla S_H , e são designados por S_{H1} , S_{H2} , S_{H3} , S_{H4} , S_{H5} , S_{H6} , S_{H7} , S_{H8} e S_{H9} (Noronha-Wagner and Bettencurt, 1967). Assim, pode-se inferir a existência de pelo menos nove genes de virulência (*v*) em *H. vastatrix*. Atualmente admite-se haver mais fatores genéticos que ainda não foram identificados (Silva, 2017; Barka, 2017).

Ao longo do processo evolutivo, os fitopatógenos desenvolveram diferentes estratégias para infectar a planta, como o transporte de proteínas efetoras (entre elas as proteínas *Avr*) para o citoplasma do hospedeiro. Da mesma forma, as plantas desenvolveram mecanismos de defesa para restringir a proliferação do patógeno em seu tecido, como o reconhecimento desses efetores pelas proteínas *R* (Jones et al., 2006; Dodds et al., 2010). Como se fosse uma “corrida armamentista”, devido a mutação e a alta capacidade adaptativa, os patógenos adquirem a habilidade de suplantar os mecanismos de defesa da planta, por meio do desenvolvimento de novas proteínas efetoras. Em contrapartida, as plantas desenvolvem novas proteínas *R* mediante ao reconhecimento desses novos efetores (Stahl & Bishop, 2000; Jones & Dangl, 2006). E durante essa “batalha”, a ativação dos mecanismos de defesa da planta, envolve uma cascata de eventos e sinais moleculares que se inicia com o reconhecimento do agente agressor e culmina com a ativação das barreiras físicas e químicas envolvidas no processo.

Os *Resistance Gene Analogs (RGAs)* são uma grande classe de genes *R* potenciais que possuem domínios conservados e características estruturais. Como tal, os *RGAs* podem ser identificados a partir de genomas sequenciados usando abordagens de bioinformática (Sekhwal et al., 2015). Nos últimos anos, devido aos avanços das tecnologias de sequenciamento um grande número de genomas foram sequenciados e montados. Apesar desse grande recurso, apenas um número relativamente pequeno de genes *R* foram

clonados e totalmente caracterizados, fornecendo não apenas informações sobre sua estrutura, função e evolução, mas também gerando recursos genéticos úteis para criar novos componentes resistentes em cultivares (Sekhwal et al., 2015). Em comparação, vários de *RGAs* foram identificados em muitos genomas vegetais (Sekhwal et al., 2015).

A identificação e caracterização de genes envolvidos nas vias de defesa da planta, podem fornecer informações importantes sobre o funcionamento desse mecanismo, assim como serem utilizadas como biomarcadores. Esses biomarcadores, se bem caracterizados, poderão auxiliar na triagem rápida de cultivares resistentes para os programas de melhoramento, além de fornecerem subsídios para o desenvolvimento de cultivares com resistência durável.

2. REVISÃO DE LITERATURA

2.1 Importância do cafeeiro

O café é uma das *commodities* tropicais mais importantes e fornece benefícios econômicos em cada etapa da cadeia produtiva que liga produtores a consumidores. A indústria do café contribui significativamente para a economia de países exportadores e importadores da cultura. Desde 1990, o setor cafeeiro global expandiu-se expressivamente com um aumento da produção de mais de 65%. O principal fator de crescimento foi o aumento do consumo nas economias emergentes e nos países produtores de café (OIC, 2019).

O Brasil é o maior produtor e exportador de café e segundo maior consumidor no mundo. Esse produto se encontra entre os dez principais setores exportadores no país, ocupando a 5ª posição (MAPA, 2019). A safra em 2019, ano de bienalidade negativa, foram obtidas aproximadamente 49,31 milhões de sacas beneficiadas em que o rendimento médio ficou em 27,20 scs/ha. Para o café arábica, a produção foi de 34,30 milhões de sacas e para o café conilon foram cerca de 15,01 milhões de sacas colhidas. (CONAB, 2019).

2.2 Ferrugem do cafeeiro - *Hemileia vastatrix*

A ferrugem do cafeeiro é causada pelo fungo biotrófico *Hemileia vastatrix* Berk. & Br, pertencente ao filo *Basidiomycota* e a ordem *Pucciniales*. Esse fungo foi descrito pela primeira vez por Berkeley em 1969, no Ceilão, onde hoje fica o Sri Lanka. No Brasil, *H. vastatrix* foi identificado em *Coffea arabica* em 1970, no estado da Bahia, e hoje está presente em praticamente todas as regiões produtoras de café do país. As perdas na produção causadas pela doença podem chegar a ordem de 30 a 50%, se não forem adotadas medidas preventivas e de controle da doença (Zambolim, 2016).

Devido aos recentes surtos da doença, chamados coletivamente de *The Big Rust* (a grande ferrugem) pelo biólogo Peter Baker, em países da América Latina, *H. vastatrix* recuperou notoriedade tornando-se uma grande preocupação para os países produtores (Baker, 2014; McCook and Vandermeer, 2015). Severas perdas ocorreram na Colômbia no ano de 2008, levando a queda de 31% da produção. Em 2013/2014 houve perdas significativas na produção da América Central e México e em 2014/2015 também foram relatadas no Peru e Equador (Avelino et al. 2015; McCook and Vandermeer, 2015). De acordo com dados recentes da Organização Internacional do Café, alguns países ainda se encontram em fase de recuperação dos surtos de ferrugem (OIC, 2018).

Um conjunto de fatores podem estar associados a “*The Big Rust*”, entre eles as mudanças climáticas. Uma redução na faixa de temperatura diurna, que encurtou o período de latência da doença, pode ajudar a explicar como o surto produziu perdas primárias altas. A variação da temperatura também poderia estar relacionado aos surtos registrados em altitudes mais elevadas (acima de 1000 m) onde o impacto da ferrugem até então era limitado (Boudrot et al., 2016; Roza et al., 2012). O surgimento de uma nova raça de ferrugem mais agressiva também foi especulado, porém não houve confirmação (Avelino et al., 2015).

H. vastatrix possui ciclo de vida complexo e ainda não bem compreendido, o que tem dificultado os estudos de evolução. Ele é dividido em duas fases distintas: uma fase dicariótica e outra monocariótica. Na primeira fase, ocorre a formação de urediniosporos dicarióticos representando a fase assexuada do ciclo que, em condições ambientais favoráveis, são responsáveis pela infecção no cafeeiro. Ocasionalmente, em condições de chuva moderada e baixas temperaturas, são formados os teliósporos. Os teliósporos podem germinar produzindo basidiósporos monocarióticos, que não infectam o cafeeiro

e representam a fase sexuada do ciclo, não sendo conhecida nenhuma função para esta estrutura ou hospedeiro. Dessa forma, o ciclo de vida de *H. vastatrix* é considerado macrocíclico e autóico (Nutman and Roberts, 1963; Fernandes et al., 2009; Silva et al., 2006). A dispersão do fungo pode ocorrer de várias formas: pela ação do vento, pelas gotas de chuva, pelo homem, por insetos e outros animais que entram em contato com plantas infectadas, sendo a água o mais eficiente agente de disseminação local (Moraes, 1983).

O processo de infecção de *H. vastatrix* no cafeeiro envolve os seguintes eventos: adesão dos uredinióporos na superfície abaxial foliar, germinação dos uredinióporos, formação do apressório sobre o estômato e colonização intra e intercelular (Talhinhas et al., 2017). A etapa de adesão é muito importante para o sucesso da colonização sendo fundamental para a correta detecção dos sinais topográficos envolvidos nas respostas tigmotrópicas e também para diferenciação do apressório (Braun and Howard, 1994). A germinação dos uredinióporos requer água e uma faixa de temperatura ótima estimada em 21-25 °C na ausência de luz (Nutman and Roberts, 1963; Rayner, 1961; Montoya and Chaves, 1974; Capucho et al., 2013). Após a formação do apressório, o fungo penetra no hospedeiro por meio dos estômatos e forma a hifa de penetração, dando origem a célula mãe do haustório (CMH), a partir da qual, os haustórios primários são produzidos nas células subsidiária e adjacente dos estômatos. Quando a hifa de penetração chega ao mesófilo, ocorre o desenvolvimento de uma vesícula típica em forma de âncora na câmara subestomática. De cada ponta dos ramos da âncora, novas CMH são formadas, a partir das quais os haustórios secundários são produzidos nas células do mesófilo. O haustório desempenha importante papel na infecção sendo responsável pela absorção de nutrientes a partir do hospedeiro, além da indução de mudanças estruturais na célula hospedeira (Ramiro et al., 2009).

A doença é caracterizada pelo aparecimento de manchas pequenas (1 a 3 mm) na face abaxial foliar, de coloração amarelo-pálidas, que evoluem atingindo até 2 cm de diâmetro, quando então apresentam aspecto pulverulento com a produção de uredinióporos de coloração amarelo alaranjada. Na face adaxial das folhas, pode ser observada manchas cloróticas amareladas que correspondem aos limites da pústula na face inferior que posteriormente necrosam. A necrose causa a queda precoce de folhas e a seca de ramos e, conseqüentemente, reduz a produção no ano seguinte. A seca progressiva dos ramos reduz a vida útil da lavoura, tornando-a antieconômica. Em regiões

onde predomina alta umidade relativa e nos cafezais sombreados, podem ser observados sinais do fungo hiperparasita *Lecanicillium lecanii*, colonizando urédias de *H. vastatrix* (Zambolim et al., 1997; Zambolim, 2016).

O controle da doença é feito principalmente por meio da aplicação de fungicidas protetores e sistêmicos (Zambolim, 2016). Porém, o uso de variedades de cafeeiro resistentes é o mais indicado pois, além de efetivo, é econômico e sustentável. A resistência à ferrugem do cafeeiro é conferida pelos genes *SH1*, *SH2*, *SH3*, *SH4*, *SH5*, *SH6*, *SH7*, *SH8* e *SH9*, que podem ser suplantados pelos genes de virulência do patógeno *v1*, *v2*, *v3*, *v4*, *v5*, *v6*, *v7*, *v8* e *v9*, respectivamente (Bettencourt e Rodrigues, 1988). Os genes *SH1*, *SH2*, *SH4* e *SH5* foram identificados em *C. arabica* (Bettencourt e Noronha-Wagner, 1971), os genes *SH6* a *SH9* foram identificados em *C. canephora* (Rodrigues - Junior et al., 1975; Bettencourt e Rodrigues, 1988) e o gene *SH3* em *C. liberica* (Prakash et al., 2004).

Foram identificadas mais de 50 raças de *H. vastatrix* a partir de amostras coletadas em diferentes países produtores de café, sendo 15 dessas encontradas no Brasil (Zambolim, 2005; Várzea e Marques, 2005; Zambolim, 2016). A raça II, a mais prevalente no Brasil e no mundo, possui o gene de virulência *v5*. Entretanto, a raça XXXIII possui 2 ou 3 genes de virulência (*v5,7* ou *v5,7,9*), sendo capaz de suplantar a resistência de cultivares que foram liberadas como resistentes à ferrugem (Várzea & Marques, 2005; Capucho et al., 2012) causando preocupação para os produtores. Como essa raça possui além do gene *v5*, outros importantes genes de virulência que tem suplantado a resistência, esse patótipo é de grande interesse para os estudos de interação com o cafeeiro. A variabilidade genética em *H. vastatrix* tem sido determinada pelo padrão de infecção observado em uma série de 24 clones de cafeeiros diferenciadores de raças (Rodrigues & Bettencourt, 1975; Zambolim, 2005).

A grande variabilidade genética do patógeno, o contínuo surgimento de novas raças, assim como a existência de raças complexas, demonstram o alto potencial evolutivo de *H. vastatrix* e representam um grande desafio para os pesquisadores na busca por resistência durável (Várzea e Marques, 2005; Zambolim et al., 2005; Cabral et al., 2009).

2.3 Sistema de defesa da planta

Uma vez que os patógenos superaram as barreiras mecânicas contra infecção, as plantas desenvolveram um complexo sistema capaz de reconhecer moléculas invasoras, realizar a transdução de sinais e responder por meio da expressão de genes de defesa. Os patógenos, por sua vez, tentam interferir nessas vias de respostas das plantas (Andersen et al., 2018).

O sistema imune inato das plantas é dividido em duas linhas de defesa principais. A primeira, denominada imunidade desencadeada por PAMPs/MAMPs (PTI - *PAMP/MAMPs triggered immunity*), baseia-se no reconhecimento de PAMPs/MAMPs (padrões moleculares associados ao patógeno/micróbio) por proteínas receptoras (PRRs - *Pattern recognition receptors*) localizadas na membrana plasmática do hospedeiro (Jones & Dangl, 2006). Os PAMPs/MAMPs geralmente possuem função estrutural, por exemplo a quitina, principal componente da parede celular de fungos. Apesar da eficiência da PTI, ao longo da evolução, os patógenos conseguiram suprimir essa linha de defesa por meio do transporte de proteínas efetoras para o citoplasma do hospedeiro, contribuindo, portanto, para a virulência do patógeno, e resultando na suscetibilidade desencadeada por efetores (ETS - *Effector Triggered Susceptibility*). Segundo Hogenhout et al. (2009), proteínas efetoras são “moléculas secretadas por organismos associados às plantas que modificam a função e estrutura da célula hospedeira”. Uma vez no interior da célula hospedeira, os efetores podem alterar as vias de seu metabolismo e suas vias de defesa.

Em resposta, uma segunda linha de defesa foi desenvolvida pelas plantas, a imunidade desencadeada por efetores (ETI - *Effector Triggered Immunity*), que utiliza receptores intracelulares com domínios NB (*nucleotide-binding*) e LRR (*leucine-rich repeat*) para o reconhecimento das proteínas efetoras produzidas pelo patógeno. Tais receptores são codificados por genes de resistência (*R*), sendo assim, denominados proteínas R. O reconhecimento dos efetores do patógeno pelas proteínas R do hospedeiro pode ocorrer de forma direta ou indireta.

O reconhecimento direto, também denominado como “modelo ligante-receptor” (Stergiopoulos & de Wit, 2009), se dá por meio da associação física direta entre a proteína R e o efector. Os efetores reconhecidos pelas proteínas R são normalmente denominados proteínas de avirulência (*Avr*) (Jones & Dangl, 2006). Quando isso ocorre, é

desencadeada a resposta de hipersensibilidade e fenótipo de resistência. A reação de hipersensibilidade (HR) é uma forma especializada de morte celular ativada por vias de sinalização que conduzem a morte do tecido no sítio de infecção e, conseqüentemente, inibição da proliferação do patógeno (Heath, 2000).

O reconhecimento indireto, também conhecido como “modelo-guarda”, ocorre por meio de proteínas acessórias que são alvos do efector no hospedeiro. Neste segundo modelo, as proteínas R monitoram a presença do efector indiretamente por meio da detecção de mudanças em outras proteínas da planta (Dodds et al., 2006). Sendo assim, quando as proteínas R e Avr estão presentes na planta e no patógeno, respectivamente, ocorre o reconhecimento (de forma direta ou indireta) originando uma interação incompatível, não desenvolvendo a doença. Contudo, quando não há o reconhecimento, o patógeno é dito virulento e a interação compatível, originando a doença (Flor, 1971).

Durante o processo coevolutivo, o patógeno adquiriu a capacidade de suprimir a ETI por meio de mutação e do desenvolvimento de novas moléculas efectoras. Em contra partida, as plantas também evoluíram desenvolvendo novas proteínas R mediante ao reconhecimento desses novos efetores (De Wit, 2007; Jones & Dangl, 2006).

Resumidamente, a “corrida armamentista” existente entre a planta e o patógeno (Stahl & Bishop, 2000) pode ser representada pelo “modelo zig-zag” descrito por Jones & Dangl (2006) (Figura 1). Na fase 1, as proteínas PRRs (*PAMP-recognition receptor*) no hospedeiro reconhecem os PAMPs/MAMPs no patógeno desencadeando a PTI. Na fase 2, os patógenos bem sucedidos translocam proteínas efectoras para o citoplasma do hospedeiro, o que interfere na PTI e resulta na ETS. Na fase 3, as proteínas efectoras são reconhecidas pelas proteínas R do hospedeiro ativando a ETI. Na fase 4, os patógenos são selecionados por adquirirem novas proteínas efectoras suprimindo a ETI. Por fim, a seleção favorece novos genes R do hospedeiro que podem reconhecer os novos efetores do patógeno, resultando, novamente, em ETI.

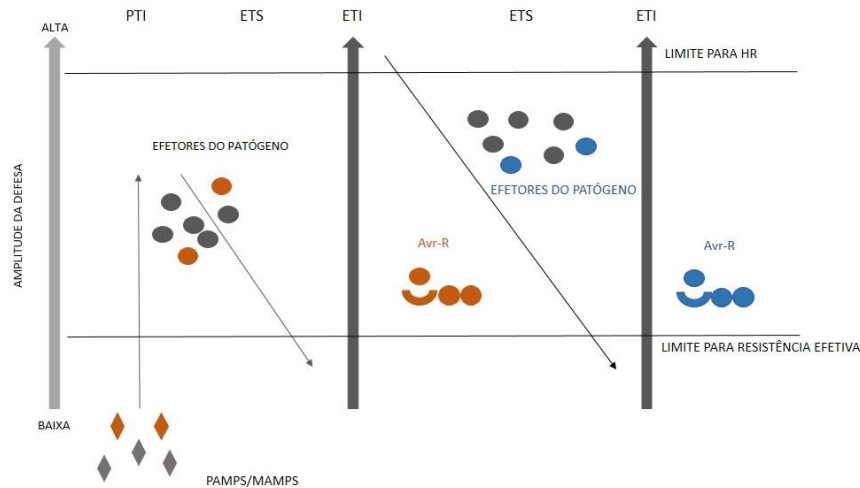


Figura 1: Modelo zig-zag proposto para ilustrar o sistema imune da planta (adaptado de Jones & Dangl, 2006).

Um completo entendimento da base molecular da resistência de plantas a doenças, poderá permitir a aplicação dessas descobertas no desenvolvimento de plantas com novas combinações de vias de resistência que sejam duráveis e capazes de reconhecer um amplo espectro de raças do patógeno. Além disso, moléculas importantes identificadas nessas vias, poderão ser utilizadas como marcadores biológicos para genótipos resistentes a doenças.

2.4 Resistance Gene Analogs (RGAs)

As plantas desenvolveram mecanismos eficazes para reconhecer e responder às infecções causadas pelos patógenos. Os *Resistance Gene Analogs (RGAs)* são genes que possuem, assim como os genes de resistência (genes *R*), domínios e motivos conservados que desempenham papéis específicos no patógeno (Sekhwal et al., 2015). Sendo assim, os *RGAs* são considerados genes *R* putativos.

As proteínas codificadas pelos genes *R* se dividem em 8 diferentes classes, sendo duas principais. A classe I que inclui proteínas citoplasmáticas conhecidas como TIR-NBS-LRR (TNL), as quais contêm três domínios principais, um domínio *Toll/Interleukin-1-like receptor (TIR)* no N-terminal, um *Nucleotide-Binding Site (NBS)* e uma série de *Leucine Rich Repeats (LRR)* no C-terminal. A classe II são as *Coiled Coil-NBS-LRR (CNL)* ou non-TIR-NBS-LRR que também são proteínas citoplasmáticas e semelhantes

à classe I, apenas diferindo na extremidade N-terminal, possuindo um domínio do tipo *Coiled Coil* ou CC. As duas classes têm sido relacionadas com a interação com proteínas efetoras de patógenos e HR, importantes fatores de resistência efetiva a doenças em plantas (Gururani et al., 2012).

Em síntese, a família de *RGAs* mais conhecida é a NBS-LRR (*Nucleotide Binding Site Leucine Rich Repeat*) sendo que as duas classes dessa família se diferem pelo seus domínios TIR e no-TIR na região N-terminal. O domínio LRR, encontrado na região C-terminal, é responsável pela interação proteína-proteína, é altamente irregular e variável. Entre os domínios NBS e LRR existe uma região chamada domínio ARC, assim denominada devido à sua ocorrência na APAF-1, proteína R e CED-4 (Van der Biezen et al., 1998). Este domínio ARC pode ser dividido em subdomínios ARC1 e ARC2. O domínio ARC, juntamente com o domínio NBS, forma uma região para ligação de nucleotídeos (Martin et al., 2003). As duas classes de proteínas R tem como alvo específico proteínas efetoras do patógeno desencadeando a ETI.

Vários estudos revelaram que os *RGAs* estão distribuídos irregularmente nos genomas vegetais e que muitos residem em *clusters*. A distribuição agrupada de *RGAs* fornece um reservatório de variação genética para impulsionar a evolução de novas especificidades do gene *R* (Michelmore et al., 1998; Young et al., 2000; Zhou et al., 2007). A maioria dos *RGAs* identificados estão agrupadas em *clusters*. Por exemplo, em arroz 50% dos genes *NBS* foram identificados agrupados (Yang et al., 2006), em batata, 73% dos genes *NBS-LRR* foram agrupados em 63 clusteres (Jupe et al., 2012). Muitos super-clusters foram identificados para genes *NBS-LRR*, incluindo um super-cluster com 11 genes *NBS* no cromossomo 4 de *Brachypodium* e um com 82 genes codificados por *NBS* em *Medicago* (Tan et al., 2012; Ameline-Torregrosa et al., 2008). Além disso, alguns genes *R* aparecem em clusteres com estruturas complexas formando uma família multigênica, tais como os genes *Cf4/9* em tomate e *L* em linho (Michelmore et al., 1998; Parniske et al., 1997; Dodds et al., 2006). Genes dentro de um único cluster pode determinar a resistência à diferentes patógenos (Michelmore et al., 1998). Pseudogenes também são distribuídos e agrupados em locais cromossômicos específicos da mesma maneira que os genes *NBS-LRR* funcionais. Os pseudogenes evoluem muito mais rapidamente do que genes funcionais e representam um reservatório para a evolução de novas especificidades (Lozano et al., 2012; Ameline-Torregrosa et al., 2008; Ota et al., 1994).

Dessa forma, vários genes *R* clonados até agora apresentam domínios comuns que podem auxiliar na identificação de locos de *RGAs* em *clusters* potencialmente envolvidos na resistência da planta à doenças. Esses genes, quando identificados e caracterizados poderão fornecer informações importantes sobre o mecanismo de resistência da planta e serem utilizados como biomarcadores de resistência.

3. REFERÊNCIAS

AMELINE-TORREGROSA, C.; WANG, B.B.; O'BLENESS, M.S.; DESHPANDE, S.; ZHU, H.; ROE, B.; YOUNG, N.D.; CANNON, S.B. Identification and characterization of nucleotide-binding site-leucine-rich repeat genes in the model plant *Medicago truncatula*. *Plant Physiol.* 146: 5-21. 2008.

AVELINO, J.; CRISTANCHO, M.; GEORGIU, S.; IMBACH, P.; AGUILAR, L.; BORNEMANN, G.; LÄDERACH, P.; ANZUETO, F.; HRUSKA, A.J.; MORALES, C. The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions. *Food Sec* 7: 303-321. 2015.

BAKER, P. The “Big Rust”: An update on the coffee leaf rust situation. *Coffee Cocoa Int.* 40:37-39. 2014.

BRAUN, E.J. AND HOWARD, R.J. Adhesion of fungal spores and germlings to host plant surfaces. In: *The Protistan Cell Surface* (Wetherbee, R., Pickett-Heaps, J.D. and Anderson, R.A., eds), pp. 202–212. Vienna: Springer. 1994.

BOUDROT, A.; PICO, J.; MERLE, I.; GRANADOS, E.; VILCHEZ, S.; TIXIER, P.; et al. Shade effects on the dispersal of airborne *Hemileia vastatrix* uredospores. *Phytopathology* 106:572-580. 2016.

CAPUCHO, A.S.; ZAMBOLIM, L.; CABRAL, P.G.C.; MACIEL-ZAMBOLIM E.; CAIXETA, E.T. Climate favorability to leaf rust in Conilon coffee. *Austral Plant Pathol* 24:511-514. 2013.

CAPUCHO, A.S.; MACIEL-ZAMBOLIM, E.; FREITAS, R.L.; HADDAD, F.; CAIXETA, E.T.; ZAMBOLIM, L. Identification of race XXXIII of *Hemileia vastatrix*

on *Coffea arabica* Catimor derivatives in Brazil. Austral Plant Dis Notes 7:189-191. 2012.

CARDOSO, R.M.L.; ZAMBOLIM, L.; CHAVES, G.M. Novas raças de *Hemileia vastatrix* Berk. et Br. identificadas em cafeeiros de genótipos complexos. Fitopatol Bras 11:318. 1986.

CARDOSO, R.M.L.; ZAMBOLIM, L.; CHAVES, G.M. Ocorrência no Brasil da raça XVI de *Hemileia vastatrix* Berk. et Br. coletada de germoplasma de *Coffea arabica* no Estado de Minas Gerais. Fitopatol Bras 5:393-394. 1988.

DODDS, P.N.; LAWRENCE, G.J.; CATANZARITI, A.M.; TEH, T.; WANG, C.I.; AYLIFFE, M.A. & ELLIS, J.G. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. Proceedings of the National Academy of Sciences 103: 8888-8893. 2006.

FERNANDES, R.C.; EVANS, H.C.; BARRETO, R.W. Confirmation of the occurrence of teliospores of *Hemileia vastatrix* in Brazil with observations on their mode of germination. Tropical Plant Pathology 34:108-113. 2009.

GURURANI, M.A., et al. Plant disease resistance genes: current status and future directions. Physiological and Molecular Plant Pathology 78:51-56. 2012.

JONES, J.D.G. & DANGL, J.L. The plant immune system. Nature 444: 323–329. 2006.

JUPE, F.; PRITCHARD, L.; ETHERINGTON, G.J.; MACKENZIE, K.; COCK, P.J.; WRIGHT, F.; SHARMA, S.K.; BOLSER, D.; BRYAN, G.J.; JONES, J.D. *et al.* Identification and localisation of the NB-LRR gene family within the potato genome. BMC Genom. 13:75-88. 2012,

LOZANO, R.; PONCE, O.; RAMIREZ, M.; MOSTAJO, N.; ORJEDA, G. Genome-wide identification and mapping of NBS-encoding resistance genes in *Solanum tuberosum* group phureja. PLoS ONE 7, e34775. 2012

MARTIN, G.B.; BOGDANOVA, A.J.; SESSA, G. Understanding the functions of plant disease resistance proteins. Annu. Rev. Plant Biol. 54:23-61. 2003.

- MICHELMORE, R.W.; MEYERS, B.C. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* 8:1113–1130. 1998.
- MCCOOK, S.; VANDERMEER, J. The Big Rust and the Red Queen: Long-Term Perspectives on Coffee Rust Research. *Phytopathology* 105: 1164-1173. 2015.
- MICHELMORE, R.W.; MEYERS, B.C. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* 8:1113–1130. 1998.
- MONTOYA, R.H.; CHAVES, G.M. Influência da temperatura e da luz na germinação, infectividade e período de geração de *Hemileia vastatrix* Berk. & Br. *Experientiae* 18:239-266. 1974.
- MORAES, S.A. A ferrugem do cafeeiro: importância, condições predisponentes, evolução e situação no Brasil. Instituto Agronômico de Campinas, Campinas, 50p. 1983.
- NORONHA-WAGNER, M. AND BETTENCURT, A.J. Genetic study of resistance of *Coffea* sp. To leaf rust. I. Identification and behaviour of four factors conditioning disease reaction in *Coffea arabica* to twelve physiologic races *Hemileia vastatrix*. *Can. J. Bot.* 45: 20121-2031. 1967
- NUTMAN, FJ; ROBERTS, FM. Studies on the biology of *Hemileia vastatrix* Berk. & Br. *Trans British Mycological Soc* 46:27-48. 1963.
- OTA, T.; NEI, M. Divergent evolution and evolution by the birth-and-death process in the immunoglobulin *VH* gene family. *Mol. Biol. Evol.* 11:469-482. 1994.
- PARNISKE, M.; HAMMOND-KOSACK, K.E.; GOLSTEIN, C.; THOMAS, C.M.; JONES, D.A.; HARRISON, K.; WULFF, B.B.; JONES, J.D. Novel disease resistance specificities result from sequence Exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato. *Cell* 91:821-832. 1997.
- RAMIRO, D.A.; ESCOUBE, J.; PETITOT, A.S.; NICOLE, M.; MALUF, M.P.; FERNANDEZ, D. Biphasic haustorial differentiation of coffee rust (*Hemileia vastatrix* race II) associated with defence responses in resistant and susceptible coffee cultivars. *Plant Pathology* 58:944-955. 2009.

- RAYNER, R.W. Germination and penetration studies on coffee rust (*Hemileia vastatrix* B. & Br.). *Annals of Applied Biology* 49: 497-505. 1961.
- ROZO, Y.; ESCOBAR, C.; GAITÁN, Á. & CRISTANCHO, M. Aggressiveness and genetic diversity of *Hemileia vastatrix* during an epidemic in Colombia. *Journal of Phytopathology* 160: 732-740. 2012.
- SILVA, M.C.; VÁRZEA, V.; GUIMARÃES, L.G.; AZINHEIRA, H.G.; FERNANDEZ, D.; PETITOT, A.; BERTRAND, B.; LASHERMES, P.; NICOLE, M. Coffee resistance to the main diseases: leaf rust and coffee berry disease. *Brazilian Journal of Plant Physiology* 18:119-147. 2006.
- STAHL, E.A. & BISHOP, J.G. Plant-pathogen arms races at the molecular level. *Current Opinion in Plant Biology* 3:299-304. 2000.
- TAN, S.; WU, S. Genome wide analysis of nucleotide-binding site disease resistance genes in *Brachypodium distachyon*. *Comp. Funct. Genom.* 418208-418220. 2012.
- VAN DER BIEZEN, E.A.; JONES, J.D. Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochem. Sci.* 23:454-456. 1998.
- YANG, S.; FENG, Z.; ZHANG, X.; JIANG, K.; JIN, X.; HANG, Y.; CHEN, J.Q.; TIAN, D. Genome-wide investigation on the genetic variations of rice disease resistance genes. *Plant Mol. Biol.* 62:181-193. 2006.
- YOUNG, N.D. The genetic architecture of resistance. *Curr. Opin. Plant Biol.* 3:285–290. 2000.
- ZAMBOLIM, L. Current status and management of coffee leaf rust in Brazil. *Tropical Plant Pathology* 41: 1-8. doi 10.1007/s40858-016-0065-9. 2016.
- ZAMBOLIM, L., ZAMBOLIM, E.M.; VALE, F.X.R.; PEREIRA, A.A.; SAKYAMA, N.S.; CAIXETA, E.T. Physiological races of *Hemileia vastatrix* Berk. Et Br. In Brazil-Physiological variability, current situation and future prospect. In: Zambolim, L., M. ZE, VÁRSEA, M. P. V. eds. Durable resistance to coffee leaf rust. Viçosa-MG: Universidade Federal de Viçosa-UFV. p. 75-98. 2005.

ZAMBOLIM, L.; CHAVES, G.M. Efeito de baixas temperaturas e do binômio temperatura-umidade relativa sobre a viabilidade dos uredósporos de *Hemileia vastatrix* Berk. et Br. e *Uromyces phaseolityca* Arth. *Experientiae*. 17: 151-184. 1974.

ZHOU, B.; DOLAN, M.; SAKAI, H.; WANG, G.L. The genomic dynamics and evolutionary mechanism of the *Pi2/9* locus in rice. *Mol. Plant Microbe Interact.* 20: 63-71. 2007

CAPÍTULO 1

ANÁLISE GENÔMICA INTEGRADA REVELA UM LOCO RICO EM *RESISTANCE GENES ANALOGS (RGAs)* COM POTENCIAL ENVOLVIMENTO NA RESISTÊNCIA DO CAFEEIRO À FERRUGEM (*Hemileia vastatrix*)

ANÁLISE GENÔMICA INTEGRADA REVELA UM LOCO RICO EM *RESISTANCE GENES ANALOGS (RGAs)* COM POTENCIAL ENVOLVIMENTO NA RESISTÊNCIA DO CAFEEIRO A FERRUGEM (*Hemileia vastatrix*)

RESUMO

O café é uma das *commodity* agrícolas mais importante no mundo. A ferrugem do cafeeiro, causada pelo fungo *Hemileia vastatrix*, é a principal doença dessa cultura e o

controle químico aumenta os custos de produção e a contaminação do meio ambiente. Dessa forma, a utilização de variedades resistentes é o mais indicado por ser econômico, eficaz e sustentável. Nesse trabalho, selecionou-se um *RGA* (*Resistance Gene Analog*), a partir de dados do transcriptoma da interação cafeeiro-*H. vastatrix*, com potencial envolvimento no mecanismo de defesa da planta. A análise de expressão gênica do *RGA* mostrou o aumento da expressão desse gene no cafeeiro resistente às 24 horas após a inoculação (hai), sugerindo o seu envolvimento na resposta de defesa do hospedeiro. Foi realizado o rastreamento do *RGA* em uma biblioteca de clones BAC (*Bacterial Artificial Chromosome*) obtida do Híbrido de Timor (HdT) CIFIC 832/2. O sequenciamento do clone BAC e a sua análise estrutural revelou um loco gênico altamente similar a um loco do genoma de *C. arabica* e com nenhuma similaridade com *C. canephora*. Foram identificados três *RGAs* e as análises de filogenia utilizando esses genes, confirmaram a similaridade dessa região com genomas disponíveis de *C. arabica*. Contudo, análise de filogenia utilizando genes que codificam proteínas do metabolismo primário presente no clone e em espécies de *Coffea* indicaram evolução rápida e específica desse loco em HdT CIFIC 832/2. A identificação e caracterização de genes envolvidos nas vias de defesa da planta podem fornecer informações importantes sobre o funcionamento desse mecanismo, assim como serem utilizados para a identificação e seleção de variedades resistentes nos programas de melhoramento do cafeeiro.

1. INTRODUÇÃO

O café é uma das *commodity* agrícola mais importante no mundo, possuindo papel de destaque na economia de vários países. *Coffea arabica* e *Coffea canephora* são as espécies mais importantes economicamente (Davis et al., 2011). *C. arabica*, um alotetraploide natural ($2n = 4x = 44$) híbrido entre as duas espécies diploides *C. canephora* e *C. eugenioides* (Lashermes et al., 1999), representa 70% da produção mundial (Davis et al., 2011).

A ferrugem do cafeeiro, causada pelo fungo biotrófico *Hemileia vastatrix*, é considerada a principal doença dessa cultura. A doença pode atingir com gravidade toda a lavoura sendo responsável por perdas da ordem de 35 a 50% da produção se não forem implantadas medidas de controle eficiente. O seu controle exige gastos e,

consequentemente, aumenta os custos de produção (Zambolim, 2016). *H. vastatrix* pertence a classe *Basidiomicete*, parasita obrigatório, podendo ocorrer em várias espécies do gênero *Coffea*, e apenas nesse gênero, não se conhece outra planta hospedeira. O aumento da infecção provoca a redução da área foliar e, gradativamente, ocorre a desfolha das plantas.

Após a germinação dos uredinósporos, o tubo germinativo se diferencia em apressório que penetra no hospedeiro por meio dos estômatos e forma a hifa de penetração, dando origem a célula mãe do haustório (CMH), a partir da qual, os haustórios primários são produzidos nas células subsidiárias e adjacentes dos estômatos. Quando a hifa de penetração chega ao mesófilo, ocorre o desenvolvimento de uma vesícula típica em forma de âncora na câmara subestomática. De cada ponta dos ramos da âncora, novas CMH são formadas, a partir das quais os haustórios secundários são produzidos nas células do mesófilo (Talhinhas et al., 2017). O haustório desempenha importante papel na infecção, sendo responsável pela absorção de nutrientes a partir do hospedeiro, além da indução de mudanças estruturais na célula hospedeira (Ramiro et al., 2009). Esse modo de interação entre o patógeno e o hospedeiro, em geral, envolve sofisticados mecanismos de supressão do sistema imunológico e a indução de genes específicos na planta para o estabelecimento da colonização pelo fungo (Schulze-Lefert e Panstruga, 2003; Voegelé e Mendgen, 2003; Guimarães et al., 2015).

O sistema imune das plantas é dividido em duas linhas de defesa principais. A primeira linha consiste no reconhecimento de padrões moleculares associados ao patógeno (PAMPs) por proteínas denominadas receptores de reconhecimento padrão (PRRs) presentes no hospedeiro (Jones e Dangl, 2006; Büttner, 2016; Oliveira e Valente, 2015). Essa linha de defesa ativa a imunidade desencadeada por PAMP (PTI). Patógenos adaptados são capazes de superar a PTI por meio da secreção de pequenas proteínas chamadas de proteínas efetoras ou efetores. Em contrapartida, as plantas são capazes de reconhecer as proteínas efetoras, de forma direta ou indireta, por meio de proteínas de resistência (R) que faz ativar uma segunda linha de defesa, a imunidade desencadeada por efetores (ETI), sendo essa mais eficiente do que a PTI (De Wit et al., 2009). Cada proteína codificada pelo gene *R* reconhece um efector específico codificado pelo gene *Avr* (de avirulência) (Flor, 1942). Entretanto, alguns patógenos diversificaram seus efetores e

adquiriram a capacidade de suprimir a ETI e colonizar o hospedeiro com sucesso (Jones e Dangl, 2006; Bigeard et al., 2015).

A maioria dos genes de resistência (genes *R*) clonados e sequenciados até o momento contêm um sítio de ligação a nucleotídeo (NBS) e um domínio de repetição rica em leucina (LRR), onde são encontrados motivos altamente conservados (Rommens & Kishore, 2000). As proteínas NBS-LRR são geralmente compostas por três domínios principais: a) um domínio N-terminal variável de aproximadamente 200 aminoácidos; b) um domínio NBS de 300 aminoácidos e c) um conjunto tandem mais variável de aproximadamente 10 a 40 curtos motivos LRR (repetição rica em leucina) (Cannon et al., 2002). Acredita-se que o domínio NBS participa da transdução de sinal, enquanto que o domínio LRR está envolvido no reconhecimento do patógeno (Young, 2000). Os *RGAs* (*Resistance Genes Analogs*), sendo genes *R* candidatos, possuem domínios conservados, principalmente NBS, ligados a resistência. Dessa forma, a identificação e caracterização de *RGAs* são fundamentais para o desenvolvimento de estratégias para os programas de melhoramento que visam a obtenção de variedades resistentes.

A identificação dos *RGAs* tem sido favorecido pelo desenvolvimento das tecnologias de sequenciamento de alto rendimento (HTS - *High Throughput Sequencing*). Essas tecnologias tem resultado na obtenção de enorme quantidade de dados de sequenciamento de DNA. Esses dados, associados a ferramentas computacionais, permitem o acesso a informação sobre genomas de várias espécies por meio da montagem de suas sequências parciais ou completas, e posterior caracterização estrutural e funcional. Métodos comparativos têm sido a base de estudos que fazem inferências sobre função e evolução em vários níveis de organização biológica. A disponibilidade de sequências genômicas completas permitiu estender o método comparativo para abranger a evolução da estrutura e função do genoma (Caicedo & Purugganam, 2005). Recentemente, os genomas das espécies *C. canephora* (Denoeud et al., 2014), *C. arabica* e *C. eugenioides* (*National Center for Biotechnology Information - NCBI*, 2018) foram disponibilizados, permitindo a identificação e caracterização de genes de interesse, incluindo *RGAs*, e análises de genômica comparativa do gênero *Coffea*.

Outros estudos em genômica funcional e interação hospedeiro-patógeno tem contribuído para o conhecimento dos genes de resistência a doenças em cafeeiro. Nesse

contexto, o cafeeiro Híbrido de Timor (HdT) tem se destacado por ser fonte de genes de resistência para várias doenças, especialmente para a ferrugem do cafeeiro. O HdT é um híbrido natural interespecífico de *C. arabica* e *C. canephora* com fenótipo de arábica sendo encontrado pela primeira vez em uma plantação de cultivar Typica na Ilha de Timor em 1917 (Bettencourt, 1973). Os derivados deste híbrido têm sido alvo de estudos moleculares com a finalidade de caracterizar os acessos de HdT com relação a diversidade genética, estrutura populacional e a presença de genes de resistência (Setotaw et al., 2013; Sousa et al, 2017; Silva et al, 2018). Cação et al. (2013), com o objetivo de disponibilizar uma ferramenta útil para a identificação de genes de interesse e para estudos de genômica comparativa, construíram uma biblioteca de cromossomo artificial bacteriano (BAC) do HdT CIFC 832/2, genótipo parental de vários cultivares modernas do cafeeiro. Essa biblioteca BAC consiste em 56.832 clones com um tamanho médio de inserção de 118 kb, que representa uma cobertura de cinco a seis vezes o genoma dihaplóide do HdT. Em outro estudo, foi obtido o transcriptoma da interação entre cafeeiros resistentes (foi utilizado o HdT CIFC 832/1) e suscetíveis e *H. vastatrix* raça XXXIII durante o processo infeccioso do fungo (Florez et al., 2017). Os resultados indicaram uma resposta de defesa pré-haustorial do hospedeiro, complementar a resistência pós-haustorial que é geralmente sugerida para esse tipo de interação.

Com base nos estudos prévios de genômica e transcriptômica do cafeeiro, o objetivo desse trabalho foi identificar e caracterizar um *RGA* com potencial envolvimento no mecanismo de defesa do cafeeiro à *H. vastatrix*. O *RGA* identificado a partir dos dados de transcriptoma da interação cafeeiro- *H. vastatrix* foi submetido a estudo de expressão gênica para validar seu possível envolvimento no processo de resposta de defesa da planta durante a infecção do patógeno. Esse gene foi caracterizado por meio do seu rastreamento em uma biblioteca de clones BAC do HdT CIFC 832/2, construída por Cação et al. (2013), sequenciamento do clone selecionado e análises de bioinformática.

2. MATERIAL E MÉTODOS

2.1. Mineração do *RGA*

Em trabalho prévio, Florez et al. (2017) obtiveram o transcriptoma da interação entre cafeeiros resistentes/suscetíveis e *H. vastatrix* raça XXXIII ao longo do tempo de infecção do hospedeiro pelo patógeno. Utilizando esse banco de dados, primeiramente

foram selecionados os genes expressos somente nas bibliotecas de interação incompatível, ou seja, interação entre planta resistente e *H. vastatrix*. Em seguida, foram filtrados os genes que apresentaram anotação relacionada a resistência da planta, chegando até a identificação do RGA. Esse foi considerado um potencial gene envolvido no mecanismo de resposta de defesa da planta e sua sequência foi recuperada a partir dos arquivos referentes a montagem por referência do transcriptoma da interação.

2.2. Material vegetal e inoculação das plantas

Para a avaliação da expressão gênica, os esporos de *H. vastatrix* raça XXXIII foram obtidos por meio da multiplicação do isolado HV-02 caracterizado biologicamente por Capucho et al. (2012). A multiplicação foi realizada de acordo com a metodologia proposta por Zambolim & Chaves (1974), em mudas de cafeeiro Caturra (CIFC 19/01). Para o experimento foram utilizadas as variedades de cafeeiro resistente HdT CIFC 832/1 e a suscetível Caturra CIFC 19/01. O mesmo foi conduzido em câmara de crescimento em condições de luminosidade (foto período de 12h) e temperatura (22°C) controladas. Foi utilizado o delineamento inteiramente casualizado (DIC) com três réplicas biológicas para cada tempo de coleta após a inoculação de urediniósporos de *H. vastatrix* raça XXXIII (0, 12 e 24 horas após a inoculação). A inoculação foi realizada da mesma forma como para a obtenção dos urediniósporos (Zambolim & Chaves, 1974). As folhas coletadas em cada tempo foram imediatamente congeladas em nitrogênio líquido e, em seguida, armazenadas em ultrafreezer -80 °C até a utilização.

2.3. Extração de RNA

As folhas, previamente congeladas, foram maceradas em nitrogênio líquido com ajuda de cadinho e pistilo autoclavados. Durante a maceração foi adicionada uma alíquota de Polivinilpirrolidona (PVPP). Aproximadamente 100 mg de folhas maceradas foram destinadas a extração de RNA utilizando-se o Concert® (Invitrogen) como extrator. O material vegetal foi aliquoteado em tubos de microcentrífuga de 2,0 mL de capacidade. Em seguida, foi adicionado 600 µL de tampão de extração Concert®. A amostra foi homogeneizada por 2 min utilizando vórtex e mantida a temperatura ambiente por 5 min. Depois da incubação, o material foi centrifugado por 15 min (14000 rpm a 4 °C). O sobrenadante foi coletado e transferido para novo tubo contendo 100 µL de NaCl 5M gelado. Em seguida, foi adicionado 300 µL de clorofórmio gelado ao tubo e este

homogeneizado em vórtex por 20 s. O material foi centrifugado por 10 min (14000 rpm a 4 °C), a fase superior foi coletada e transferida para novo tubo. Esse passo foi repetido até que o sobrenadante se mostrou totalmente limpo. A ele foi adicionado um volume igual ao coletado (400 / 450 µL) de isopropanol gelado. A mistura foi homogeneizada por inversão e incubada de 15 a 20 min a -20 °C, seguida de uma nova centrifugação por 15 min (14000 rpm a 4 °C). O sobrenadante foi descartado e ao pellet foi adicionado 1 mL de etanol 80% gelado. Em seguida foi realizada nova centrifugação por 2 min (14000 rpm a 4 °C). O sobrenadante foi descartado e o pellet seco à temperatura ambiente. O pellet foi ressuspensionado em 20 µL de água tratada com dietil pirocarbonato (DEPC) e armazenado a -80 °C até o momento do uso.

O RNA total foi quantificado com o auxílio do Qubit[®] 2.0 fluorometer (*Life Technologies*), seguindo as recomendações do fabricante, e sua integridade foi avaliada por eletroforese em gel de agarose (1,5 %) corado com brometo de etídio (Wilson e Walker, 2000).

2.4. Análise de expressão gênica

A partir da sequência do *RGA* selecionado foram desenhados *primers* que amplificam, utilizando o programa Oligo Explorer 1.2 (Tabela 1). Os *primers* foram usados para a reação de PCR quantitativo em tempo real.

Tabela 1: Lista de *primers* utilizados nas análises de RT-qPCR. **TM:** Temperatura de anelamento. * Genes endógenos utilizados na análise.

Gene	Sequência do <i>primer</i> (5' → 3')	TM (°C)	Amplicom (bp)	Publicação
<i>RGA</i>	F-AGTCCCTGCCACATCAGCTC	56.0	106	Este estudo
	R-GCCGCAAAGTTTCCGAACCA	56.0		
<i>GAPDH</i> *	F-AGGCTGTTGGGAAAGTTCTTC	63.4	70	Cruz et al. 2009
	R-ACTGTTGGAACCTCGGAATGC	64.0		
<i>UBQ</i> *	F-CAGACCAGCAGAGGCTGATT	64.6	100	Cruz et al. 2009
	R-AGAACCAAGTGAAGGGTGGGA	63.5		

A síntese de cDNA foi realizada com 1 µg de RNA total, pré-tratado com 1 µL de DNase por 30 min (50 U/µL, Amplification Grade DNase I, Invitrogen TM). A síntese

da primeira fita de cDNA foi feita utilizando o kit ImProm-II™ *Reverse Transcription System Protocol RT-PCR* (Promega), segundo as orientações do fabricante. O cDNA foi armazenado a -20 °C até a utilização.

Para a realização da PCR quantitativa em tempo real, em aparelho 7500 Real Time PCR Systems (*Applied Biosystems*), foi utilizado o sistema de detecção de fluorescência SYBR Green I (*Applied Biosystems, California, USA*). Para cada reação foi utilizado 1 µL (25 ng) da diluição da reação de síntese de cDNA de fita simples, 1 µL (2 µM) de cada iniciador, 4 µL de SYBR® Green PCR Master Mix (*Applied Biosystems*) e 3 µL de água estéril, ficando com volume final de 10 µL. As condições térmicas da reação foram: 95 °C por 10 min para a desnaturação inicial, seguido de 40 ciclos de 95 °C por 15 s e 60 °C por 60 s e finalizando com 95 °C por 15 s. O nível de expressão dos genes foi calculado utilizando os valores médios de Ct resultante de três réplicas biológicas e três réplicas técnicas por meio de curva padrão relativa. As análises estatísticas foram realizadas usando o *GraphPad Prism* (versão 5; GraphPad Software Inc.; La Jolla, CA, EUA). As médias foram comparadas usando o Teste de Tukey ($p < 0,05$). Para a normalização dos dados foram utilizados 2 genes constitutivos: GAPDH e UBQ (Cruz et al., 2009) (Tabela 1).

2.5. Screening de clones BAC

O RGA selecionado foi rastreado a partir de uma biblioteca de 56.832 clones BAC oriunda de um clone de Híbrido de Timor CIFC 832/2, resistente a ferrugem, construída por Cação et al. (2013). Os clones foram replicados em placas de titulação de 384 poços utilizando replicador autoclavado em uma cabine de fluxo laminar. As culturas antigas foram replicadas para uma nova placa de titulação de 384 poços com 75 µL de meio LB fresco acrescido de 12,5 µg/ml de cloranfenicol em cada poço. A multiplicação da cultura foi realizada por meio da incubação das placas em agitador com temperatura de 37 °C durante 18h a velocidade de 120 rpm.

Na identificação do clone contendo o gene de interesse, utilizou-se o método de decomposição de *pool* de clones BAC. Procedeu-se a identificação de meia placa contendo o clone positivo, seguido pela identificação da coluna da meia placa identificada e, por fim, a seleção do clone positivo, analisando a coluna da etapa anterior. Para o *screening* das BAC, amostras de *pools* de DNA foram amplificadas por meio de PCR

utilizando o *primer* específico do *RGA* em estudo (Tabela 1). A reação de PCR foi otimizada contendo 50 ng de DNA, 0.1 μ M de cada *primer*, 0.15 mM de dNTP (Promega), 1 mM de MgCl₂, 1 U de *Taq* polimerase (Invitrogen) e 1x de tampão de reação de PCR, para um volume final de 20 μ L de reação. A amplificação foi realizada em termociclador (Veriti, *Applied Biosystems*) programado com desnaturação inicial de 94 °C por 5 min, seguido por 35 ciclos de amplificação de desnaturação a 94 °C por 30 s, anelamento a 60 °C por 30 s e extensão a 72 °C por 1 min. Foi realizada extensão final de 72 °C por 10 min. O produto da amplificação foi visualizado em gel de agarose 1,5%.

O DNA da BAC identificada foi extraído utilizando o *Kit Wizard® SV Plus Minipreps DNA Purification System* (Promega), seguindo o protocolo do fabricante. O DNA foi quantificado com o auxílio do Qubit® 2.0 fluorometer (*Life Technologies*).

2.6. Sequenciamento e avaliação da qualidade

O clone contendo o fragmento correspondente do gene de interesse, foi sequenciado por meio da plataforma Illumina MiSeq, utilizando *MiSeq Reagent Kit v2*, seguindo as recomendações do fabricante. Essa etapa de sequenciamento do clone foi repetida com o objetivo de aumentar a cobertura do sequenciamento. A qualidade das *reads* foi avaliada utilizando o *software* FastQC versão 0.11.5 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). O *software* trimmomatic versão 0.36 foi utilizado, com valor de Phred=30, para a seleção das *reads* de boa qualidade e eliminação dos adaptadores. Além disso, esse *software* gera arquivos separados com as *reads* órfãs (*Unpair*). As *reads* com qualidade e que possuem seu par (*pair*), dos dois sequenciamentos realizados, foram então unidas e submetidas a montagem (Figura 1).

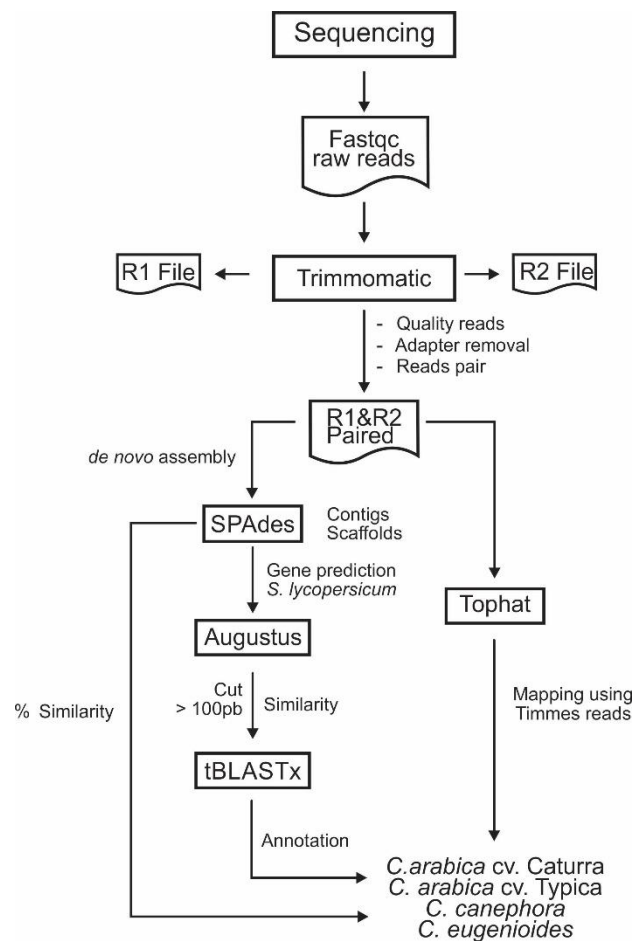


Figura 1: Pipeline do sequenciamento, qualidade, filtragem das *reads* de montagem, predição gênica e análises de comparação entre genomas. R1 *read* forward, R2 *read* reverse.

2.7. Montagem, anotação e mapeamento das *reads*

A montagem da sequência do clone foi realizada utilizando a estratégia *de novo* por meio do *software* SPAdes 3.9.0 (<http://cab.spbu.ru/software/spades/>). A predição de genes foi feita por meio do *software* AUGUSTUS versão 3.3.1 (<http://augustus.gobics.de/>). *Solanum lycopersicum* foi utilizado como genoma de referência por ser a espécie mais próxima filogeneticamente disponível nesse *software* até a presente data. A anotação gênica foi feita com base nos genomas das espécies *C. arabica* cv. Caturra (Assembly GCA_003713225.1 Cara_1.0), *C. eugenioides* (Assembly GCA_003713205.1 Ceug_1.0) e *C. canephora* (<http://www.coffee-genome.org/coffeacanephora>). Foi utilizada a ferramenta BLAST *all*, por linha de comando na linguagem de programação *python*, na busca por similaridade de sequências

com parâmetro *e-value* < 10⁻⁵. Foram filtrados somente os melhores *hits* entre todos os genomas. As sequências anotadas foram submetidas ao *software* Pfam (<https://pfam.xfam.org/>) para a busca por domínios conservados. A análise estrutural do loco de resistência identificado foi feita por meio de comparação com as sequências dos diferentes genomas e da utilização da ferramenta *Genoma Data Viewer* disponível no NCBI (<https://www.ncbi.nlm.nih.gov/genome/gdv/>).

Para o mapeamento foram utilizados dois genomas disponíveis de *C. arabica* (Caturra e Typica), um de *C. canephora* e um de *C. eugenioides*. Arquivos índice dos genomas foram obtidos por meio do *software* Bowtie2-Build versão 2.2.8. Em seguida as *reads* foram alinhadas utilizando o *software* Bowtie2 versão 2.2.5 e Tophat versão 2.1.1 (Trapnell et al., 2009). Além do mapeamento utilizando as *reads* trimadas, foi realizada também análise de comparação de sequências, com *scripts* em *python*, utilizando a ferramenta do BLAST tBLASTx, entre os *scaffolds* do genoma montado do clone e os quatro genomas disponíveis de *Coffea* (Figura 1).

2.8. Análise filogenética

A fim de identificar o *RGA* de interesse e entender a provável origem do loco que contém os *RGAs*, os genes identificados nessa região foram submetidos a análise de filogenia. Foi incluída nessa análise a sequência da qual foram obtidos os *primers* utilizados para o rastreamento do *RGA* nos clones BAC. Além disso, cinco genes relacionados ao metabolismo primário da planta presentes em outras regiões da sequência do clone, também foram submetidos a análise filogenética. Os genes do metabolismo primário selecionados foram: NODE34 (transketolase chloroplast-like); NODE82 (fumarate hydratase mitochondrial); NODE98 (Protein TRIGALACTOSYLDIACYLGLYCEROL chloroplast); NODE138 (Aspartate aminotransferase cytoplasmic) e NODE162 (RNA polymerase beta' subunit). Sequências similares aos genes selecionados foram obtidas dos quatro genomas das espécies do gênero *Coffea* por meio da ferramenta tBLASTx, considerando *e-value* < 10⁻⁵. Foram selecionados os melhores *hits* de cada genoma. Arquivos *multifasta* com todas as sequências foram construídos e utilizados para as análises de filogenia por meio do *software* MEGA7 versão 7.0.26 (Kumar et al., 2016). As sequências foram alinhadas utilizando o ClustalW com parâmetros padrões. As árvores filogenéticas foram obtidas

utilizando o método *maximum likelihood* com os seguintes parâmetros: deleção completa dos *gaps*, modelo Tamura-Nei (Tamura et al., 2007) e 1000 réplicas de *bootstrap*.

3. RESULTADOS

Com base no banco de dados de transcriptoma da interação cafeeiro-*H. vastatrix* raça XXXIII (Florez et al., 2017), foram identificados os genes expressos exclusivamente nas bibliotecas de interação incompatível. Em seguida, foram selecionados os genes que apresentaram anotação relacionada com a resistência da planta, até a identificação do *RGA*. Esse *RGA* foi considerado um potencial gene envolvido no mecanismo de resposta de defesa da planta e sua sequência foi recuperada a partir dos arquivos referentes a montagem por referência do transcriptoma da interação.

O *RGA* selecionado foi submetido a análise de expressão gênica temporal por meio da técnica de RT-qPCR. Foram avaliados cafeeiros resistente (HdT CIFC 832/1) e suscetível (Caturra CIFC 19/01), observando sua expressão às 0, 12 e 24 horas após a inoculação (hai). Os resultados mostraram diferença significativa de expressão entre os genótipos resistente e suscetível nos tempos de 12 e 24 hai, além de expressão maior e mais rápida no genótipo resistente (Figura 2).

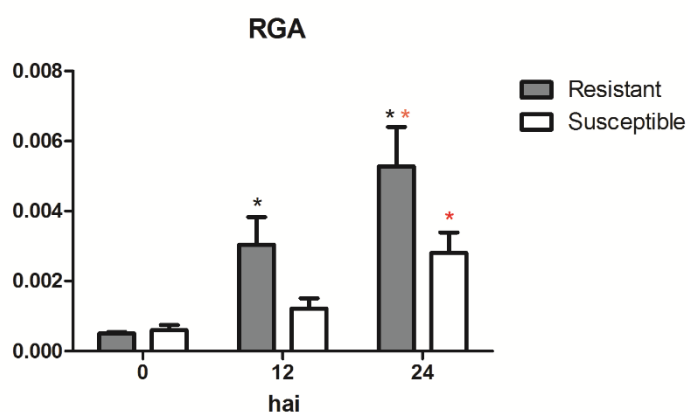


Figura 2: Validação do *RGA* por qPCR. O eixo y representa o nível de expressão em relação aos genes endógenos em três momentos, 0, 12 e 24 horas após a inoculação (hai) por *Hemileia vastatrix*. Barras de erros = SEM, n = 3 réplicas biológicas independentes. Asterisco (*) indica diferença significativa do nível de expressão na mesma hai entre as interações resistente e suscetível. Asterisco (*) indica diferença significativa do nível de expressão entre as hai dentro da mesma interação.

Com o objetivo de obter a sequência total do gene e caracterizá-lo, o *RGA* foi clonado de uma biblioteca BAC do cafeeiro HdT CIFC 832/2. Usando a estratégia de decomposição de *pools* de clones BAC e amplificação do gene em estudo por PCR, foi obtido um clone localizado na placa 101, poço 2P (clone 101-2P), contendo o fragmento do gene de interesse. Esse clone foi sequenciado e um total de 31019 *reads* (R1-*forward* e R2-*reverse*), com tamanho 35-151 pares de bases, foram obtidos. Um segundo sequenciamento foi realizado, resultando em um total de 52605 *reads* (R1 e R2) com o mesmo tamanho de *reads*. Os arquivos de *reads* de boa qualidade, dos dois sequenciamentos, que possuem seu par (*pair*), foram então unidos e posteriormente utilizados para a montagem do clone 101-2P. Um total, 81352 *reads pair* de boa qualidade foram obtidas a partir dos dois sequenciamentos (Tabela 2).

Tabela 2: Parâmetros estatísticos para a montagem do clone 101-2P utilizando o *software* SPAdes.

Características do Genoma	Números
<i>Reads</i> de alta qualidade	81352
Tamanho total	221871pb*
Números de <i>scaffolds</i>	2597
Maior <i>scaffold</i>	56551pb
N50	1615
L50	11
Tamanho total (≥ 1000 pb)	121403pb
Tamanho total (≥ 5000 pb)	94573pb
Tamanho total (≥ 10000 pb)	85323pb
Tamanho total (≥ 25000 pb)	56551pb
Tamanho total (≥ 50000 pb)	56551pb
CG (%)	43,53

pb*: pares de bases

A montagem realizada com o *software* SPAdes gerou 2600 *contigs* e 2597 *scaffolds* com valor de N50=1615. O tamanho total da montagem foi de 221871pb sendo o maior *scaffold* com 56551 pb. A partir dos *scaffolds* foram preditos 1645 ORFs (*Open*

Reading Frame) por meio do *software* AUGUSTUS. Foram filtrados somente as ORFs que apresentaram tamanho acima de 100 pares de bases (pb), totalizando 1636 ORFs. Esses possíveis genes foram então anotados utilizando a ferramenta tBLASTx e os genomas de *C. arabica*, *C. canephora* e *C. eugenioides*. Um total de 297 genes foram anotados (Material Suplementar 1). Vinte genes (~7%) foram anotados como proteínas não caracterizadas, enquanto que a grande maioria foi anotada como proteína que participa do metabolismo primário da planta. Três genes (*scaffold1g1*, *scaffold1g3* e *scaffold3g5*) foram anotados como proteínas RGAs, confirmando que o clone 101-2P sequenciado provavelmente contém o gene de interesse. As sequências desses três genes foram submetidas ao *software* Pfam para a análise de domínios conservados em proteínas. O resultado indicou que as proteínas codificadas pelos genes possuem os domínios Rx-N e NB-ARC, ambos encontrados em proteínas envolvidas na resistência de plantas à patógenos (Rommens & Kishore, 2000).

Com o objetivo de identificar o *RGA* de interesse e entender a provável origem do loco que contém os *RGAs*, foi realizada análise filogenética utilizando as sequências dos genes presentes no loco e a sequência que deu origem aos *primers* utilizados para o rastreamento na BAC. O resultado mostrou um agrupamento contendo os três *RGAs*, *scaffold1g1*, *scaffold1g3* e *scaffold3g5*, e a sequência de construção do *primer* separada. Todos os *RGAs* analisados apresentaram relação filogenética com os genomas de *C. arabica* cv. Caturra e *C. arabica* cv. Typica (Figura 3).

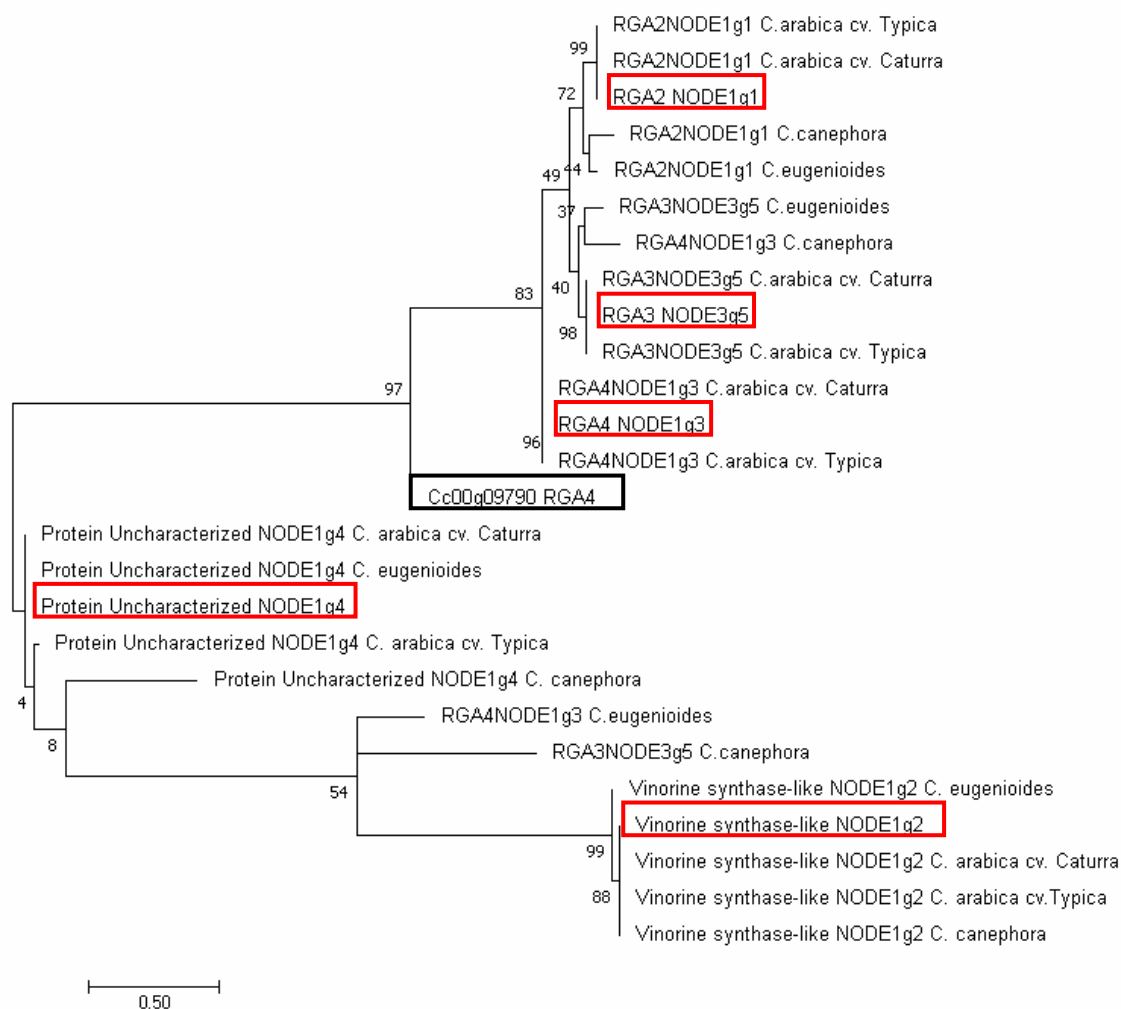


Figura 3: Análise filogenética dos genes presentes no loco *RGA* do clone 101-2P pelo Método da Máxima Verossimilhança. Foi incluída na análise a sequência do transcriptoma que deu origem aos *primers* utilizados no rastreamento do clone na BAC (Cc00g09790 RGA4, destacado em preto). O valor de *bootstrap* é mostrado ao lado das ramificações. A árvore foi projetada em uma escala com comprimentos de ramificações medidas no número de substituições por sítio. Genes destacados em vermelho são os genes identificados no clone sequenciado. Os demais, representam os genes similares nos quatro genomas de referência.

O clone BAC em estudo corresponde a uma região do genoma do HdT, um híbrido natural interespecífico entre *C. arabica* e *C. canephora*. Dessa forma, com o objetivo de verificar se essa região possui genes de *C. arabica* ou se é resultante da introgressão de *C. canephora*, foi realizada busca por similaridade dos *scaffolds* do clone 101-2P montado contra os quatro genomas de *Coffea* disponíveis. Nos genomas utilizados como referência estão incluídos dois de *C. arabica* (variedades Caturra e Typica) e um de *C. canephora* (duplohaploide acesso DH200-94). Foi analisado, também, um genoma de *C. eugenioides*, uma vez que *C. arabica* corresponde a um alotetraploide originado de *C.*

canephora e *C. eugenioides*. Observou-se que o clone BAC 101-2P apresenta maior porcentagem de similaridade de sequência com o *C. arabica* cv. Caturra (45,72 %), seguido do *C. arabica* cv. Typica (43,4 %) e *C. eugenioides* (10,9 %). Não foram encontradas sequências similares com o genoma de *C. canephora* (0 %) (Figura 4a).

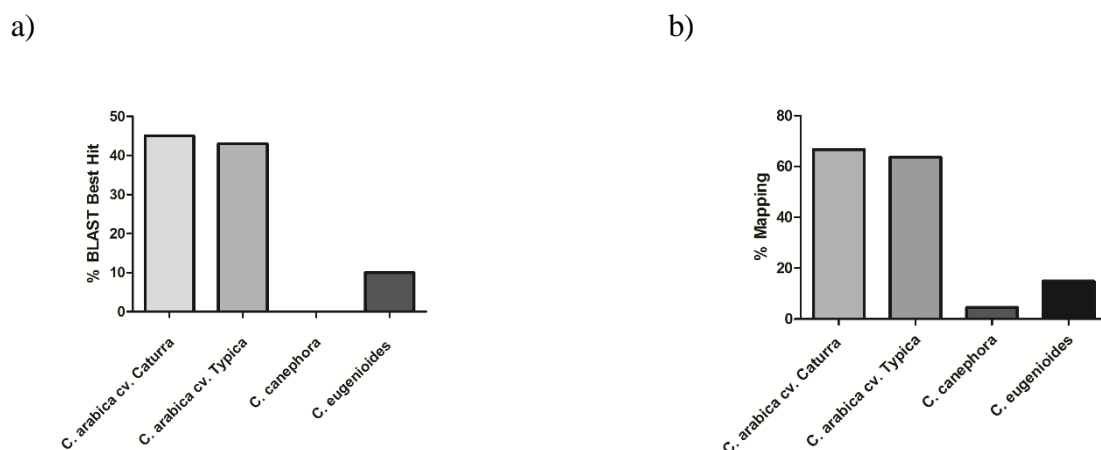


Figura 4: a) Porcentagem de BLAST best hits observada entre os *scaffolds* do genoma montado do clone 101-2P e os quatro genomas disponíveis das espécies do gênero *Coffea*. b) Porcentagem de mapeamento observada entre as *reads* trimadas e os quatro genomas.

Tendo em vista que a sequência montada do clone BAC 101-2P pode conter artefatos de montagem e, portanto, influenciar os resultados de busca por similaridade, foi realizado o mapeamento das *reads* trimadas utilizando os quatro genomas disponíveis de *Coffea* como referência (Figura 4b). O genoma em que houve o maior número de *reads* alinhadas foi o de *C. arabica* cv. Caturra (66,7%), seguido por *C. arabica* cv. Typica (63,7%). Apenas 14,9% das *reads* foram alinhadas no genoma de *C. eugenioides*, enquanto que no genoma de *C. canephora* apenas 4,6%.

A fim de estudar a estrutura da região contendo os *RGAs*, o *scaffold1* e o *scaffold3* da montagem foram submetidos a análise de similaridade com sequências disponíveis do gênero *Coffea* presentes no NCBI. Essa análise mostrou que essas regiões contendo *RGAs* são altamente conservadas no genoma de *C. arabica*, apresentando o melhor *hit* no cromossomo 1e (cromossomo vindo de *C. eugenioides*) dessa espécie. Nessa região foram identificados cinco genes com as mesmas anotações dos genes encontrados no *scaffold1* e no *scaffold3* (Figura 5). Entretanto, na região de maior similaridade com a espécie *C. eugenioides* não foram identificados genes, enquanto que na espécie *C. canephora* não foram observadas sequências similares.

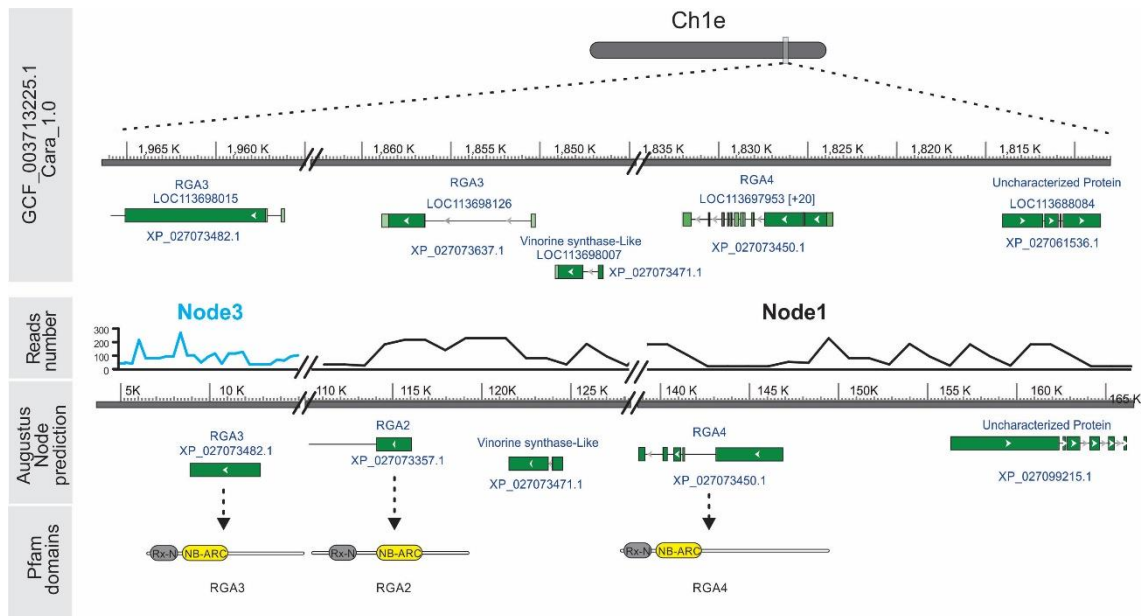


Figura 5: Representação do loco de resistência identificado no clone 101-2P e do loco similar presente no genoma de *C. arabica* Chr1e. Blocos verdes representam os éxons, setas representam a orientação da fita. O ID do loco e da proteína, e a anotação estão identificados ao lado de cada gene. GCF_003713225.1 Cara_1.0 (código de identificação do genoma disponível no NCBI).

Análises filogenéticas dos genes codificadores de proteínas do metabolismo primário presentes no clone 101-2P e nos quatro genomas de *Coffea* disponíveis, mostraram os genes do clone em *clusters* separados do restante dos genomas com valores significativos de *bootstrap* (Figura 6).

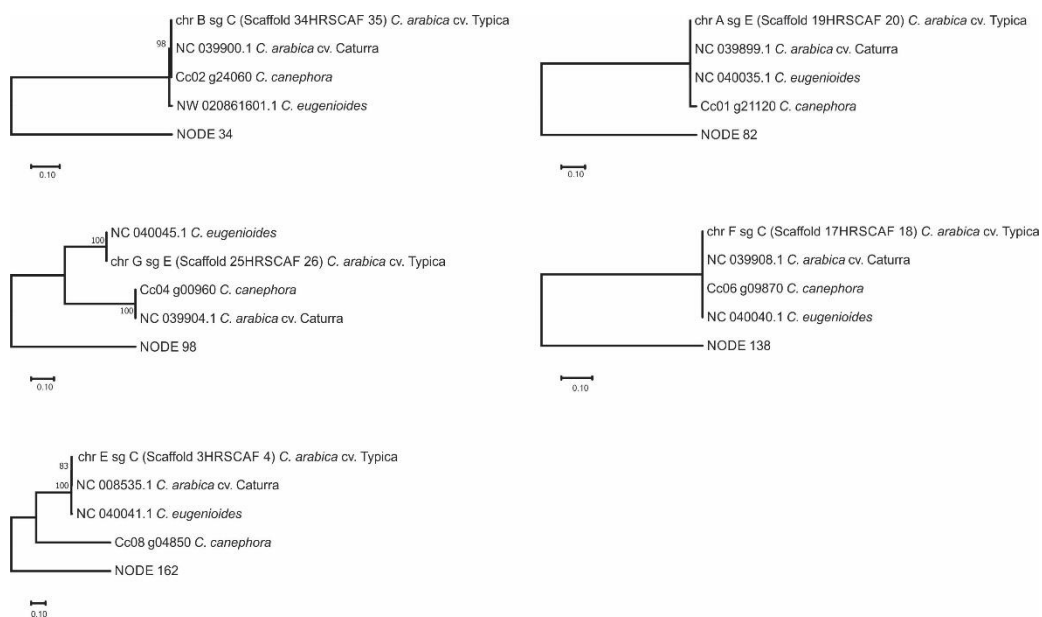


Figura 6: Análise filogenética dos genes selecionados do metabolismo primários presentes no clone 101-2P pelo Método da Máxima Verossimilhança. O valor de *bootstrap* é mostrado ao lado das ramificações. A árvore foi projetada em uma escala com comprimentos de ramificações medidas no número de substituições por site. Códigos de identificação de sequência usados para construir as árvores estão disponíveis nos materiais suplementares 2 (SM2). NODE (representação dos genes selecionados)

4. DISCUSSÃO

Os *RGAs* (*Resistance Gene Analogs*) são genes que codificam proteínas que possuem domínios conservados presentes em proteínas de resistência caracterizadas. Dessa forma, a caracterização de *RGAs* é extremamente importante para o desenvolvimento de estratégias de melhoramento de plantas visando à resistência a patógenos. O *RGA* selecionado a partir de dados de transcriptômica (Florez et al., 2017) foi identificado somente nas bibliotecas de interação incompatível. A fim de validar esse dado, primeiramente foi realizada a análise de expressão temporal desse gene. Os resultados mostraram expressão significativa diferenciada entre o genótipo resistente e suscetível em 12 e 24 hai, ocorrendo aumento da expressão desse gene no genótipo resistente. O padrão de expressão do gene se mostrou crescente durante as horas após a infecção nos dois genótipos, porém com uma maior e mais rápida expressão no genótipo resistente na fase inicial (12-24 hai) da infecção. Esse padrão de expressão reforça a ideia de que a resistência a patógenos em plantas pode depender da taxa e extensão da síntese de uma ou mais proteínas induzidas no hospedeiro pelo patógeno (Diniz et al. 2012). Diniz et al. (2012) sugeriram que a resposta rápida de resistência limita a formação de

haustórios e isso pode ser a base para o sucesso da resistência da planta à infecção pelo patógeno.

Confirmando a sua importância durante a interação planta-patógeno, o clone BAC 101-2P contendo o *RGA* foi sequenciado. Um total de 81352 *reads pair* com boa qualidade foram utilizadas para a montagem. Apesar dos genomas de referência de *C. canephora* e *C. arabica* estarem disponíveis, optou-se pela estratégia de montagem *de novo*, pois o loco sequenciado pertence a um híbrido interespecífico, o Híbrido de Timor. Dessa forma, mudanças estruturais nesse loco poderiam não ser identificadas por meio da montagem utilizando um genoma de referência. Com a montagem *de novo* do clone, obteve-se 2597 *scaffolds*, sendo que os seis maiores *scaffolds* correspondem a 47% do clone.

Na anotação gênica do clone 101-2P, três *RGAs* foram identificados nos *scaffold* 1 e 3, confirmando que a montagem foi eficiente, apesar de relativa fragmentação observada. No *scaffold1* foram preditos dois *RGAs* (anotados como *RGA2* e *RGA4*) e no *scaffold3* foi predito um (*RGA3*). Todos esses genes codificam proteínas que possuem domínios conservados relacionados a resistência, NB-ARC e Rx-N, segundo o *software* Pfam. A maioria das proteínas de resistência nas plantas possuem sítio de ligação de nucleotídeos e repetição rica em leucina (NLR, também conhecidos como superfamília NB-LRR) (Lee and Yeom, 2015). Tais proteínas possuem domínios NB-ARC na região central e LRR na região C-terminal. Já foi demonstrado que genes codificadores de NB sem LRRs também pode funcionar na imunidade da planta (Nandety et al., 2013).

Com o objetivo de identificar qual o *RGA* de interesse e a provável origem do loco, análise de filogenia foi realizada. O resultado indicou relação filogenética entre os *RGAs* identificados no clone, ficando esses genes em *clusters* separados do restante dos genes com anotação diferente. Porém, o *RGA* utilizado para desenho dos *primers* apresentou relação filogenética com todos os *RGAs* presentes no clone. Esse resultado sugere que qualquer um dos *RGAs* preditos no clone 101-2P pode ser o gene de interesse, podendo ser até mesmo locos duplicados. A duplicação de genes é um dos processos evolutivos mais importantes que gera diversidade genética e novidade funcional e, portanto, é fundamental na adaptação e na especiação (Kondrashov, 2012; Magadum et al., 2013; Panchy et al., 2016). A análise de filogenia também mostrou que todos os

RGAs presentes no clone possuem relação filogenética com os genes de *C. arabica* cv. Caturra e *C. arabica* cv. Typica. Esse resultado sugere que provavelmente esse loco tenha origem do genoma de *C. arabica*.

Foi realizada análise de similaridade entre a sequência montada do clone 101-2P e os quatro genomas disponíveis de *Coffea*. Os resultados indicaram porcentagem de similaridade maior com o genoma de *C. arabica* cv. Caturra. As *reads* trimadas do clone foram também mapeadas nos quatro genomas de *Coffea*. Essa análise confirmou uma porcentagem maior de mapeamento no genoma de *C. arabica* cv. Caturra. Observou-se baixa porcentagem de similaridade com o genoma de *C. canephora* utilizando as *reads* trimadas, enquanto que na análise utilizando o genoma montado do clone BAC não foram observadas sequências similares. Apesar do genoma clonado ser o do HdT CIFC 832/2, um híbrido natural entre *C. arabica* e *C. canephora*, trabalhos anteriores mostraram que esses híbridos interespecíficos apresentam uma alta relação genética natural com *C. arabica* (Lashermes et al, 1993, 1996, 2000).

A ausência de similaridade da sequência do clone BAC com a espécie *C. canephora* pode também ser resultante da distância genética existente entre o HdT (DNA da BAC) e o genoma de referência de *C. canephora* utilizado na análise. Diferente de *C. arabica*, *C. canephora* apresenta alta variabilidade genética, sendo constituída por dois grupos varietais distintos. O Grupo Guineano compreende os genótipos do oeste africano e é fenotipicamente caracterizado como suscetível à ferrugem. O segundo grupo, conhecido como Grupo Congolês, é dividido em cinco subgrupos, SG1, SG2, B, C e Ugandence. Os genótipos do SG1 apresentam características semelhantes ao Grupo Guineano, sobretudo a suscetibilidade à ferrugem. Por outro lado, genótipos SG2, B, C e Ugandence são altamente resistentes à ferrugem (Cubry et al., 2008; Musoli et al., 2009). Os subgrupos SG1 são representados por variedades conhecidas localmente como "café Conilon", enquanto SG2, B, C e Ugandence são conhecidas como "café Robusta". O genoma sequenciado de *C. canephora* é um duplo haploide proveniente do clone IF200. Esse clone apresenta características intermediárias entre os grupos genéticos Guineano e Congolês (Lashermes et al., 1993). Já o HdT, é um híbrido entre as espécies *C. arabica* e *C. canephora*, porém acredita-se que o *C. canephora* pertença ao grupo varietal robusta. Como os genótipos utilizados nas análises possuem origens diferentes, esse pode ser o

motivo pelo qual foi observada ausência de similaridade entre o loco sequenciado que pertence ao HdT CIFC 832/2 e o genoma de *C. canephora*.

Por meio de análise de similaridade entre os *scaffolds1* e *3* (em que foram encontrados os *RGA*s) e os três genomas disponíveis no NCBI de *C. arabica*, *C. eugenioides* e *C. canephora*, foi identificada uma região altamente similar com 97 % de cobertura, 99,94 % de identidade e *e-value* = 0.0, no genoma de *C. arabica* (cromossomo 1e). Nessa região, também estão preditos cinco genes. A análise estrutural desse loco revelou alta sintenia e baixa divergência entre as sequências. Estudo da evolução do loco de resistência *SH3* do genoma das espécies *C. arabica*, *C. eugenioides* e *C. canephora*, por meio do sequenciamento de clones BACs e análises comparativas também revelou uma sintenia genética quase perfeita e baixa divergência de sequências. Em comparação com os resultados de análises semelhantes em outros gêneros (*Aegilops/Triticum* e *Oryza*), os genomas de *Coffea* parecem ser muito menos divergentes, o que é consistente com a radiação relativamente recente do gênero *Coffea* (Cenci et al., 2012).

A evolução de sequências codificadoras de região NBS em cafeeiros envolvem um acúmulo gradual de mutações, o que reflete uma evolução lenta (Noir et al., 2001) e que pode explicar a alta similaridade existente entre os *RGA*s presentes no clone com a região no genoma de *C. arabica*. Entretanto, não houve similaridade entre as sequências desse loco e o genoma de *C. canephora*, e a região similar no genoma de *C. eugenioides* não apresentou genes (dados não mostrados). O resultado sugere que esse loco *RGA* pode não estar presente em *C. canephora*. No entanto, essa falta de similaridade pode também ser explicada pela grande variabilidade genética dessa espécie (Dussert et al., 1999; Prakash et al., 2005; Gomez et al., 2009; Musoli et al., 2009), além do fato de que o genoma de *C. canephora* disponível e a sequência montada do clone 101-2P possuem origens diferentes como citado anteriormente. Vários estudos mostraram que dentro de um *cluster* de genes de resistência à doenças são encontrados genes que apresentam altos níveis de similaridade a genes *R*, mas não codificam um produto funcional (Ori et al., 1997; Song et al., 1997). Outros estudos revelaram que *clusters* de genes *R* podem conter sequências relacionadas em função mas não em sequência (Dixon et al., 1996; Milligan et al., 1998). Esses estudos são coerentes com o resultado de similaridade obtido com o genoma de *C. eugenioides*, onde foi observada alta similaridade com uma região que não apresenta genes.

Foi realizada análise de filogenia utilizando sequências de nucleotídeos de genes codificadores de proteínas de metabolismo primário presentes no genoma do clone 101-2P e nos quatro genomas disponíveis. Por serem genes conservados e geralmente de cópia simples, a análise poderia mostrar a história evolutiva dos genes presentes no clone inteiro, não somente no loco dos *RGAs*, e mais próximo de qual espécie esses genes estariam. Contudo, todos os genes utilizados na análise ficaram em *clusters* isolados, não apresentando relação filogenética com o restante dos genomas, o que também pode indicar uma evolução rápida e específica desse loco. Essa característica pode ter sido desenvolvida após o evento de hibridação, que deu origem ao HdT, por meio de fatores evolutivos como a recombinação entre regiões homólogas de *C. arabica* e *C. canephora*, pais de HdT.

Os resultados obtidos nesse trabalho foram úteis para o entendimento da origem do loco *RGA* presente no cafeeiro HdT 832/2, além de fornecerem informações importantes sobre a estrutura do loco de *RGAs*. Observou-se que o loco *RGA* em estudo apresenta três diferentes *RGAs* possivelmente provenientes de *C. arabica*. Estudos funcionais desses genes deverão ser realizados. A união dessas informações poderá auxiliar o desenvolvimento de estratégias que visam o controle da doença por meio da identificação de variedades resistentes. Os três *RGAs* identificados constituem alvos importantes para os programas de melhoramento do cafeeiro que visam a resistência do cafeeiro à *H. vastatrix*.

5. REFERÊNCIAS

Bettencourt AJ, Rodrigues Jr CJ (1988) Principles and practice of coffee breeding for resistance to rust and other diseases. In: Clarke RJ, Macrae R (eds) London and New York: Coffee Agronomy, 199-234, 4v. Elsevier Applied Scienci.

Bettencourt AJ, Noronha-Wagner M (1971) Genetic Factors Conditioning Resistance of *Coffea arabica* L. To *Hemileia vastatrix* Berk. & Br. Agron Lusit 31:285-292

Bigeard J, Colcombet J, Hirt H (2015) Signaling Mechanisms in Pattern-Triggered Immunity (PTI). Molecular Plant 8: 521-539

- Bettencourt AJ (1973) Considerações gerais sobre o “Híbrido de Timor.” Circ. No 23 20.
- Büttner D (2016) Behind the lines-actions of bacterial type III effector proteins in plant cells. FEMS Microbiol Rev 40:894-937
- Caicedo AL, Purugganan MD (2005) Comparative Plant Genomics. Frontiers and Prospects. Plant Physiology 138: 545–547
- Cação SMB, Silva NV, Domingues DS et al. (2013) Construction and Characterization of a BAC Library From the *Coffea arabica* Genotype Timor Hybrid CIFIC 832/2. Genetica. doi: 10.1007/S10709-013-9720-Y
- Cannon SB, Zhu H, Baumgarten AM, Spangler R, May G, Cook DR, Young ND (2002) Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. Journal of Molecular Evolution, New York 54: 548-62
- Capucho AS, Zambolim EM, Freitas RL, Haddad F, Caixeta ET, Zambolim L (2012) Identification of race XXXIII of *Hemileia vastatrix* on *Coffea arabica* Catimor derivatives in Brazil. Austral Plant Dis Notes 7:189-191.
- Cenci A, Combes MC, Lashermes P (2012) Genome evolution in diploid and tetraploid coffee species as revealed by comparative analysis of orthologous genome segments. Plant Molecular Biology 78: 135.
- Cubry P, Musoli P, Legnate’ H, Pot D, De Bellis F, Poncet V (2008) Diversity in coffee assessed with SSR markers: structure of the genus *Coffea* and perspectives for breeding. Genome 51: 50–63
- Cruz F, Kalaoun S, Nobile P et al (2009) Evaluation of coffee reference genes for relative expression studies by quantitative real-time RT-PCR. Mol Breed 23:607-616. doi: 10.1007/S11032-009-9259-X
- Davis AP, Tosh J, Ruch N, Fay MF (2011) Growing coffee: *Psilanthus* (*Rubiaceae*) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. Bot J Linn Soc 167:357–377. doi:10.1111/j.1095-8339.2011.01177.x.

De Wit PJGM, Mehrabi R, Van Den Burg HA, Stergiopoulos I (2009) Fungal effector proteins: past, present and future. *Mol Plant Pathol* 10:735–747. doi:10.1111/j.1364-3703.2009.00591.x.

Denoeud F, Paulet LC, Dereeper A et al. (2014) The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science* 345: 1181-1184 doi: 10.1126/science.1255274

Diniz I, Talhinhos P, Azinheira HG et al. (2012) Cellular and molecular analyses of coffee resistance to *Hemileia vastatrix* and nonhost resistance to *Uromyces vignae* in the resistance-donor genotype HDT832/2. *European Journal Plant Pathology* 133: 141-157.

Dixon MS, Jones DA, Heddie JS, Thomas CM, Harrison H, Jones JDJ (1996) The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell* 84: 451-459.

Fernandez D, Tisserant E, Talhinhos P, Azinheira H, Vieira A et al. (2012) 454-pyrosequencing of *Coffea arabica* leaves infected by the rust fungus *Hemileia vastatrix* reveals in plant expressed pathogen-secreted proteins and plant functions in a late compatible plant-rust interaction. *Mol Plant Pathol* 13:17-37.doi:10.1111/j.1364-3703.2011.00723.x

Flor H (1942) Inheritance of pathogenicity in *Melampsora lini*. *Phytopathology* 32: 653-669.

Florez JC, Mofatto LS, Lopes RLF, Ferreira SS, Zambolim EM, Carazzolle MF, Zambolim L, Caixeta ET (2017) High throughput transcriptome analysis of coffee reveals prehaustorial resistance in response to *Hemileia vastatrix* infection. *Plant Mol Biol* 95: 607-623 DOI 10.1007/s11103-017-0676-7

Gomez C, Dussert S, Hamon P, Hamon S, de Kochko A, Poncet V (2009) Current genetic differentiation of *Coffea canephora* Pierre ex A. Froehn in the Guineo-Congolian African zone: cumulative impact of ancient climatic changes and recent human activities. *BMC Evol Biol* 9:167.

Guimarães LG, Tenente R, Pinheiro C, Chaves I, Silva MC, Cardoso FNH et al. (2015) Proteomic analysis of apoplastic fluid of *Coffea arabica* leaves highlights novel biomarkers for resistance against *Hemileia vastatrix*. *Frontiers In Plant Science* 6:478.

Guzzo SD (2004) Aspectos bioquímicos e moleculares da resistência sistêmica adquirida em cafeeiro contra *Hemileia vastatrix*. Dissertação, Universidade de São Paulo.

Jones, JDG & Dangl, JL (2006) The Plant Immune System. *Nature* 444: 323-29.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33: 1870-1874

Kondrashov FA (2012) Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proc Biol Sci* 279: 5048-5057. doi: 10.1098/rspb.2012.1108

Lashermes P, Andrzejewski S, Bertrand B, Combes MC, Dussert S, Graziose G, Trouslot P, Anthony F (2000) Molecular analysis of introgressive breeding in coffee (*Coffea arabica* L.). *Theor Appl Genet* 100: 139-146. doi: 10.1007/s001220050019

Lashermes P, Combes MC, Robert J, Trouslot P, D'Hont A, Anthony F, Charrier A (1999) Molecular characterization and origin of the *Coffea arabica* L. genome. *Mol Gen Genet* 261:259-266. doi: 10.1007/s004380050965

Lashermes P, Cros J, Marmey P, Charrier A (1993) Use of random amplified DNA marker to analyse genetic variability and relationships of *Coffea* species. *Gen Crop Evolution* 40: 91-99. doi: 10.1007/BF00052639

Lee HA, and Yeom SI (2015) Plant NB-LRR proteins: tightly regulated sensors in a complex manner. *Brief Funct Genomics* 14: 233-242. doi: 10.1093/bfgp/elv012

Le Roux C, Huet G, Jauneau A, Camborde L, Tremousaygue D, Kraut A, et al. (2015) A receptor pair with an integrated decoy converts pathogen disabling of transcription factors to immunity. *Cell* 161: 1074-1088. doi: 10.1016/j.cell.2015.04.025

Magadum S, Banerjee U, Murugan P, Gangapur D, and Ravikesavan R (2013) Gene duplication as a major force in evolution. *J. Genet.* 92: 155-161. doi: 10.1007/s12041-013-0212-8

- Milligan SB, Bodeau J, Yaghoobi J, Haloshian I, Zabel P, Uil- liamson VM (1998) The root knot nematode resistance gene *Wi* from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. *Plant Cell* 10: 1307-1319
- Musoli P, Cubry P, Aluka P, Billot C, Dufour M, De Bellis T et al. (2009) Genetic differentiation of wild and cultivated populations: diversity of *Coffea canephora* Pierre in Uganda. *Genome* 52: 634-46.
- Nandety RS, Caplan JL, Cavanaugh K, Perroud B, Wroblewski T, Michelmore RW et al. (2013) The role of TIR-NBS and TIR-X proteins in plant basal defense responses. *Plant Physiol.* 162: 1459-1472. doi: 10.1104/pp.113.219162
- Noir S, Combes MC, Anthony F, Lashermes P (2001) Origin, diversity and evolution of NBS-type disease-resistance gene homologues in coffee trees (*Coffea* L.). *Mol Genet Genomics* 265: 654-662. doi: 10.1007/s004380100459
- Oliveira-Garcia E, Valent B (2015) How eukaryotic filamentous pathogens evade plant recognition. *Curr Opin Microbiol* 26: 92-101
- Ori N, Eshed Y, Paran I, Presting G, Aviv D, Tanksley S, Zamir D, Fluhr R (1997) The I2Cl family from the wilt disease resistance locus I2 belong to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. *Plant Cell* 9: 521-532.
- Panchy N, Lehti-Shiu M, and Shiu SH (2016) Evolution of gene duplication in plants. *Plant Physiol.* 171: 2294-2316. doi: 10.1104/pp.16.00523
- Prakash NS, Marques DV, Várzea VMP, Silva MC, Combes MC, Lashermes P (2004) Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into *Coffea arabica* L. *Theorl Appl Genet* 109:1311-1317
- Ramiro DA, Escoube J, Petitot AS, Nicole M, Maluf MP, Fernandez D (2009) Biphasic haustorial differentiation of Coffee Rust (*Hemileia vastatrix* Race II) associated with defence responses in resistant and susceptible coffee cultivars. *Plant Pathology* 58: 944-955.
- Rodrigues CJ, Bettencourt AJ, Rijo L (1975) Races of the pathogen and resistance to coffee rust. *Annu Rev Phytopathol* 13:49-70. doi:10.1146/annurev.py.13.090175.000405

Rommens CM, Kishore GM (2000) Exploiting the full potential of disease-resistance genes for agricultural use. *Current Opinion Biotechnology*, London, 11: 120-25.

Sarris PF, Duxbury Z, Huh SU, Ma Y, Segonzac C, Sklenar J et al. (2015) A plant immune receptor detects pathogen effectors that target WRKY transcription factors. *Cell* 161: 1089-1100. doi: 10.1016/j.cell.2015.04.024

Schulze-Lefert P, Panstruga R (2003) Establishment of biotrophy by parasitic fungi and reprogramming of host cells for disease resistance. *Annu Rev Phytopathol* 41:641–667. doi:10.1146/annurev. phyto.41.061002.083300

Setotaw TA, Caixeta ET, Pereira AA et al. (2013) Coefficient of parentage in *Coffea arabica* L. cultivars grown in Brazil. *Crop Science* 53:1237-1247. doi: 10.2135/cropsci2012.09.0541

Sekhwil MK, Li P, Lam I, Wang X, Cloutier S, You FM (2015) Disease Resistance Gene Analogs (RGAs) in Plants. *Int. J. Mol. Sci.* 16: 19248-19290. doi:10.3390/ijms160819248

Silva RA, Zambolim L, Castro ISL, Rodrigues HS, Cruz CD, Caixeta ET (2018) The Híbrido de Timor germplasm: identification of molecular diversity and resistance sources to coffee berry disease and leaf rust. *Euphytica* 214:153 <https://doi.org/10.1007/s10681-018-2231-2>.

Silva MC, Nicole M, Guerra-Guimarães L, Rodrigues CJ (2002) Hypersensitive cell death and post-haustorial defence responses arrest the orange rust (*Hemileia vastatrix*) growth in resistant coffee leaves. *Physiol Mol Plant Pathol* 60: 169-183. doi:10.1006/pmpp.2002.0389

Song UY, Pi LY, Uang GL, Gardner J, Holsten T, Ronald PC (1997) Evolution of the rice Xa21 disease resistance gene family. *Plant Cell* 9: 1279-1287.

Talhinhas P, Batista D, Diniz I, Vieira A, Silva DN, Loureiro A et al. (2017) The coffee leaf rust pathogen *Hemileia vastatrix*: one and a half centuries around the tropics. *Mol Plant Pathol*. 18: 1039-1051. <https://doi.org/10.1111/mpp.12512>

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol.* 10: 512-526.

Trapnell C, Pachter L, Salzberg SL (2009) Tophat: Discovering splice junctions with Rna-Seq. *bioinformatics* 25:1105-1111. doi:10.1093/Bioinformatics/Btp120.

Várzea VMP, Marques DV (2005) Population Variability of *Hemileia vastatrix* vs. coffee durable resistance. In: Zambolim L, Zambolim EM, Várzea VMP (Ed.). *Durable Resistance to Coffee Leaf Rust*. Viçosa: UFV. 53-74.

Voegelé RT, Mendgen K (2003) Rust haustoria: nutrient uptake and beyond. *New Phytol* 159: 93-100. doi:10.1046/j.1469-8137.2003.00761.x

Wilson K & Walker K (2000) *Principles and Techniques of Practical Biochemistry*. 5^a Ed. Cambridge: Cambridge University Press, 784 P.

Young ND (2000) The genetic architecture of resistance. *Current Opinion Plant Biology*, Reino Unido 3: 285-290.

Zambolim L (2016) Current Status and management of coffee leaf rust in Brazil. *Tropical Plant Pathology* 41: 1-8. doi 10.1007/S40858-016-0065-9

Zambolim L, Zambolim EM, Vale FXR, Pereira AA, Sakyama NS, Caixeta ET (2005) Physiological races of *Hemileia vastatrix* Berk. et Br. in Brazil-physiological variability, current situation and future prospect. In: Zambolim L, M. Ze, Várzea MPV (eds) *Durable Resistance to Coffee Leaf Rust*. Viçosa-MG: Universidade Federal de Viçosa-UFV pp 75-98.

Zambolim L, Chaves GM (1974) Efeito de baixas temperaturas e do binômio temperatura-umidade relativa sobre a viabilidade dos uredósporos de *Hemileia vastatrix* Berk. *Et Br. e Uromyces phaseolityca* Arth. *Experientiae* 17: 151-184.

Material Suplementar 1: Anotação das ORFs, preditas pelo *software* AUGUTUS, presentes no clone 101-2P. (Apêndice 1)

Material Suplementar 2: Códigos de identificação (ID) de sequências do NCBI usadas para as análises filogenéticas. (Apêndice 2)

CAPÍTULO 2: ANÁLISE DE EXPRESSÃO GÊNICA SUGERE UM RGA POTENCIALMENTE ENVOLVIDO NA RESPOSTA DE RESISTÊNCIA DO CAFEIEIRO À FERRUGEM

**ANÁLISE DE EXPRESSÃO GÊNICA SUGERE UM RGA POTENCIALMENTE
ENVOLVIDO NA RESPOSTA DE RESISTÊNCIA DO CAFEIEIRO A FERRUGEM**

RESUMO

O cafeeiro é uma importante *commoditie* agrícola. A principal doença que afeta a cultura é a ferrugem, causada pelo fungo biotrófico *Hemileia vastatrix*. O método de controle da doença mais recomendado é o desenvolvimento de variedade resistentes. Estudos

mostraram que a resistência do cafeeiro à *H. vastatrix* é governada pela interação gene-gene. Os *Resistance Gene Analogs (RGAs)* são uma classe de genes que possuem domínios conservados e características estruturais presentes em genes *R*. Dessa forma, os *RGAs* são considerados genes *R* putativos e possuem potencial envolvimento no mecanismo de defesa da planta durante a interação com o patógeno. O objetivo desse trabalho foi identificar se os três *RGAs*, identificados no capítulo 1, possuem importância na resposta de defesa do cafeeiro à *H. vastatrix* e investigar se os mesmos poderão ser utilizados como biomarcadores de resistência a ferrugem. *Primers* específicos para cada *RGA* foram construídos e utilizados para amplificação de DNA de cafeeiros diferenciadores de raças de *H. vastatrix* e para análises de expressão gênica em genótipos resistentes e suscetíveis. Os três *RGAs* foram amplificados em todos os clones de cafés diferenciadores de raças. Esses dados sugerem que a diferença entre genótipos resistentes e suscetíveis poderia estar no nível de expressão dos *RGAs*. As análises de expressão sugeriram que o *RGA4* possivelmente é o gene que mantém a função envolvida com a resistência do cafeeiro a *H. vastatrix* raça XXXII. Além disso, os resultados de expressão corroboram com a ideia de resistência pré-haustorial sugerida para a interação cafeeiro-*H. vastatrix*.

1. INTRODUÇÃO

As duas espécies de café de maior importância econômica são *Coffea arabica* (café arábica) e *Coffea canephora* (café canephora), representando, respectivamente, 75 e 25% da produção nacional (MAPA, 2019). *C. arabica* é uma espécie tetraploide, com 90% de autofecundação, enquanto as demais espécies são diploides, autoincompatíveis e de fecundação cruzada (Charrier et al., 1985). Um dos principais desafios para os produtores de café é o controle da ferrugem do cafeeiro, doença responsável por danos severos à cultura (Zambolim, 2016).

A ferrugem do cafeeiro, causada pelo fungo *Hemileia vastatrix* Berkeley & Broome, é considerada a principal doença da cultura. O fungo está amplamente distribuído nas regiões produtoras de café, causando prejuízos econômicos. Os sintomas da ferrugem do café são observados na face inferior das folhas, onde aparecem pequenas manchas amarelo-pálidas, as quais evoluem para lesões maiores de aspecto pulverulento

(urediniósporos) de coloração amarelo-alaranjada. Na face superior, observam-se manchas cloróticas amareladas que correspondem aos limites da pústula na face inferior, a qual necrosa e causa a queda da folha infectada (Zambolim et al., 1999).

Apesar da eficiência dos fungicidas no controle da ferrugem, o emprego de cultivares resistentes é considerado o método mais econômico e eficiente, além de não causar impactos ambientais. O grande desafio para os melhoristas é o surgimento de novas raças do patógeno capazes de suplantar a resistência de cultivares resistentes desenvolvidas. As cultivares Oeiras e Icatú Vermelho, por exemplo, tiveram sua resistência suplantada cerca de 9 e 15 anos, respectivamente, após seu lançamento (Capucho *et al.*, 2009). A baixa durabilidade da resistência de cultivares é devido à grande variabilidade e evolução de novas raças virulentas do patógeno, onde as condições climáticas são favoráveis à doença. No Brasil, desde 1971, já foram relatadas 15 destas raças (Zambolim, 2016). Tradicionalmente, a variabilidade fisiológica em *H. vastatrix* tem sido determinada pelo padrão de infecção observado em uma série de 24 clones de cafeeiros diferenciadores de raças (Várzea & Marques, 2005). Raças distintas são identificadas por meio da observação da virulência de isolados nestes clones diferenciadores que possuem diferentes combinações de genes de resistência.

Estudos mostraram que a interação entre o cafeeiro e *H. vastatrix* é governada pela teoria gene-a-gene sendo a resistência do cafeeiro condicionada por, pelo menos, nove genes dominantes isolados ou em conjunto. Nessa interação, esses genes são nomeados pela sigla S_H (Rodrigues Jr et al., 1975; Bettencourt & Rodrigues Jr., 1988). Os genes de resistência S_{H1} , S_{H2} , S_{H4} e S_{H5} foram encontrados em *C. arabica*, os genes S_{H6} , S_{H7} , S_{H8} e S_{H9} em *C. canephora* e o gene S_{H3} em *C. liberica*. Os genes S_{H6} , S_{H7} , S_{H8} e S_{H9} também foram encontrados em derivados do Híbrido de Timor, um híbrido natural proveniente do cruzamento entre *C. arabica* e *C. canephora*, o que indica que esses genes são originados do café Robusta (*C. canephora*) (Rodrigues et al., 1975; Bettencourt et al., 1988). Dessa forma, de acordo com a teoria gene-a-gene, pode-se inferir que existe pelo menos nove genes de avirulência (*Avr*) em *H. vastatrix*.

As proteínas codificadas pelos genes de resistência, também chamados de genes *R*, são capazes de reconhecer, de forma direta ou indireta, proteínas efetoras secretadas pelo fungo e codificadas pelos genes de avirulência (*Avr*), ativando, assim, os

mecanismos de resistência da planta, resultando em uma interação incompatível. Quando a planta não é capaz de reconhecer as proteínas efetoras, o patógeno consegue estabelecer a colonização do hospedeiro resultando em uma interação compatível (Ellis et al., 2009). Como já foi observado em diferentes patossistemas, mutações nos genes *Avr* permitem que os patógenos suplantem a resistência conferida por um determinado gene *R* (Stergiopoulos & de Wit, 2009). Assim, enquanto os patógenos evoluem adquirindo novos efetores para manipular os processos de defesa das plantas, estas também evoluem obtendo novas proteínas R devido a pressão de seleção, o que sugere uma constante e indefinida “corrida armamentista” na interação planta-patógeno (Coll et al., 2011). Dessa forma, a identificação de genes expressos pela planta durante a infecção pelo patógeno pode fornecer informações importantes a respeito dos mecanismos que levam a resistência ou à doença no hospedeiro.

Os genes de resistência isolados até o momento podem ser classificados em três classes diferentes com base na presença ou ausência de seus domínios principais: i) genes tipo NBS-LRR (sítio de ligação a nucleotídeo-repetições ricas em leucina); ii) genes R contendo domínio LRR e sem domínio NBS e iii) Outros genes *R* (Oth-R) que codificam proteínas sem LRR e/ou NBS. A maioria dos genes *R* isolados de várias espécies pertence à classe NBS-LRR e está presente principalmente em clusters nos genomas (Reddy et al., 2018). A classe NBS-LRR ainda pode ser dividida em TIR-NBS-LRR (Toll/interleukin-1-receptor homology region-NBS-LRR) e no-TIR-NBS-LRR, com base da região N-terminal do domínio NBS. O domínio NBS (sítio de ligação a nucleotídeo) atua como uma chave molecular para controlar o status de ativação da proteína. O domínio LRR (repetição rica em leucina) na extremidade C-terminal é necessário para o reconhecimento do efector *Avr*. Sendo assim, os genes *R* de várias espécies de cultivares compartilham domínios conservados que podem ser usados para rastrear genomas de plantas em busca de genes *R* putativos (Reddy et al., 2018).

Os *Resistance Gene Analogs (RGA)* são uma classe de genes que possuem domínios conservados e características estruturais presentes em genes *R*. Dessa forma, são considerados genes *R* putativos e possuem potencial envolvimento no mecanismo de defesa da planta durante a interação com o patógeno. Os *RGAs* podem ser identificados a partir de genomas sequenciados utilizando abordagens de bioinformática. Com o avanço das tecnologias de sequenciamento de nova geração vários genomas estão disponíveis

atualmente o que tem possibilitado a identificação de muitos *RGAs*. No entanto, apesar desse recurso, poucos genes *R* putativos foram clonados e totalmente caracterizados (Sekhwal et al., 2015).

Em trabalho prévio (Capítulo 1) foi identificado um *cluster* de *RGAs* possivelmente envolvido no processo de resposta de resistência do cafeeiro. Esse *cluster* foi caracterizado por meio de comparações genômicas. Essas análises sugeriram que os *RGAs* identificados possuem sequências similares, possivelmente originadas de eventos de duplicação, impossibilitando a identificação do *RGA* de interesse. A duplicação de genes é um dos processos evolutivos mais importantes que gera diversidade genética e novidade funcional e, portanto, é fundamental na adaptação e na especiação (Kondrashov, 2012; Magadum et al., 2013; Panchy et al., 2016). Essa observação é comum em interações específicas de espécies, nas quais uma cópia do gene geralmente mantém a função básica do gene progenitor. Dessa maneira, outras cópias estão disponíveis para evoluir (Hughes, 2005).

Dessa forma, o objetivo desse trabalho foi identificar o *RGA*, entre os três *RGAs* candidatos presente em *cluster*, que possui importância na resposta de defesa do cafeeiro Híbrido de Timor CIFC838/2 à raça XXXIII de *H. vastatrix* e investigar se o mesmo poderia ser utilizado como biomarcador de resistência à ferrugem.

2. MATERIAL E MÉTODOS

2.1. Análises comparativas de sequências e desenho de *primers* específicos

Em trabalho prévio (Capítulo 1), um loco rico em *RGAs* foi identificado e caracterizado. A fim de identificar o *RGA* que poderá ser usado como biomarcador de resistência, as sequências dos mesmos (*RGA2*, *RGA3* e *RGA4*) foram alinhadas e comparadas utilizando o *software* Clustal Omega (Sievers et al., 2011). Regiões específicas de cada *RGA* foram identificadas e utilizadas para o desenho de *primers* utilizando o *software* Primer3 versão 0.4.0 (Untergasser et al., 2012) (Tabela 1). Os *primers* obtidos foram avaliados quanto a sua especificidade por meio da ferramenta Primer-BLAST. Os *primers* que foram capazes de amplificar somente o *RGA* correspondente foram selecionados para serem utilizados nas análises de expressão

gênica. Os mesmos foram utilizados para as análises de reação em cadeia da polimerase (PCR) e PCR em tempo real (RT-qPCR).

Tabela 1: *Primers* selecionados para cada *RGA* identificado. (F) sequência do *primer* forward; (R) sequência do *primer* reverse. *Genes endógenos utilizados nas análises de RT-qPCR.

Gene	Sequência do primer (5' → 3')	Amplicom (pb)	Eficiência (%)	Publicação
<i>RGA2</i>	F: GAGGGTCTCATTGTGGAAGAG R: CTTCTAATTCATTGTCTGGGGA	151	107,58	Este estudo
<i>RGA3</i>	F: GGTCACCTGGGATTGATCAACC R: GATCTTCCTACAACGTTTCGGA	101	94,17	Este estudo
<i>RGA4</i>	F: ATCGTGAAGGAGAGTCAAACCTG R: CCATCAATCTTTGGTTTTGGCG	214	95,48	Este estudo
GAPDH*	F: AGGCTGTTGGGAAAGTTCTTC R: ACTGTTGGAACCTCGGAATGC	70	101,12	Cruz et al. 2009
UBQ*	F: CAGACCAGCAGAGGCTGATT R: AGAACCAAGTGAAGGGTGA	100	95,49	Cruz et al. 2009

2.2.Extração de DNA genômico

Folhas jovens e completamente expandidas foram coletadas das plantas de cafeeiro utilizadas como diferenciadoras de raças de *H. vastatrix*. O DNA dessas folhas foi extraído conforme a metodologia descrita por Diniz et al. (2005). Após a extração, o mesmo foi quantificado em espectrofotômetro Nanodrop 2000 e a sua qualidade avaliada por eletroforese em gel de agarose (1,5 %) corado com brometo de etídio (10 mg/mL). Os DNAs foram diluídos para a concentração final de 25 ng/μl e armazenados a - 20°C até o uso.

2.3.Identificação dos *RGAs* em clones diferenciadores de raças

Os *primers* específicos para cada *RGA* identificado foram utilizados para amplificação das regiões correspondentes por meio da reação em cadeia da polimerase (PCR). Para tal foram utilizados os seguintes reagentes: 2,0 μL do DNA na concentração de 25 ng/μL, 2,0 μM de cada primer, 0,2 mM de cada dNTP (Promega), 2,0 mM de MgCl₂, 1,0 U Taq DNA polimerase (Invitrogen) e tampão de reação de PCR 1X, para um volume final de 20 μL. O termociclador utilizado foi o Veriti (Applied Biosystems) programado para 94°C por 1 min e 35 ciclos de 94 °C 30 s, 65 °C 30 s e 72 °C 1 min. O produto da PCR obtido foi visualizado em gel de agarose 1,5%. A reação foi realizada em 22 cafeeiros diferenciadores (Tabela 2).

Tabela 2: Clones de cafeeiros diferenciadores de raças de *Hemileia vastatrix* e os respectivos fenótipos de interação com a raça XXXIII.

CAFEIROS	FENÓTIPOS
CIFC 849/1	Suscetível
CIFC 7963	Suscetível
CIFC 128/2	Resistente
CIFC 635/3	Resistente
CIFC 1343/269	Resistente
CIFC 32/1	Resistente
CIFC 33/1	Resistente
CIFC 644/18	Resistente
CIFC 419/20	Resistente
CIFC 420/2	Resistente
CIFC 420/10	Resistente
CIFC 829/1	Resistente
CIFC 110/5	Resistente
CIFC 1006/10	Resistente
CIFC 153/2	Resistente
CIFC 635/2	Resistente
CIFC 152/3	Resistente
CIFC 17/12	Resistente
CIFC 147/1	Resistente
CIFC 832/1	Resistente
CIFC 134/4	Resistente
CIFC 87/1	Resistente
CIFC 832/2*	Resistente
CIFC 419*	Suscetível
CIFC 19/1*	Suscetível

*Genótipos que não fazem parte da série diferenciadoras de raças de *Hemileia vastatrix*, mas foram incluídos na análise. 832/2 é o genótipo clonado na biblioteca BAC. CT 419 e 19/1 são variedades suscetíveis à maioria das raças de *Hemileia vastatrix*. CIFC (Centro de Investigações das Ferrugens do Cafeeiro).

2.4.Obtenção dos isolados de *H. vastatrix*

Para a análise de PCR em tempo real, foi utilizado o isolado da raça XXXIII (isolado Hv-02) caracterizado biologicamente por Capucho *et al.* (2012) e desde então mantido no Laboratório de Biotecnologia do Cafeeiro (Biocafé) localizado no Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO), na Universidade Federal de Viçosa (UFV), em Viçosa-MG.

A multiplicação do isolado foi realizada de acordo com a metodologia proposta por Zambolim & Chaves (1974), em mudas de cafeeiro Caturra (CIFC 19/1). A inoculação dos urediniosporos foi feita na face abaxial de folhas jovens e completamente expandidas com o auxílio de um pincel. Em seguida, as folhas foram aspergidas com água

destilada, utilizando um borrifador, até a obtenção de um leve molhamento superficial. Posteriormente, as plantas foram cobertas com saco plástico e mantidas no escuro por 48 horas a 22°C. Após este período, as plantas foram mantidas em câmara de crescimento com temperatura de 22°C e fotoperíodo de 12 horas até o aparecimento de urediniósporos. Os mesmos foram coletados e acondicionados em cápsulas de gelatina dentro de um dessecador com solução de ácido sulfúrico (densidade de 1,8 e concentração 32,6%) de modo a manter a umidade relativa do ambiente interno em torno de 50%. O dessecador foi mantido em refrigerador a 4 °C para a preservação da viabilidade dos urediniósporos. Este procedimento foi repetido até que se obteve quantidade suficiente de esporos para a realização do trabalho.

2.5. Teste de viabilidade dos esporos

A viabilidade dos urediniósporos de *H. vastatrix* foi avaliada pelo teste de germinação em meio ágar-água 2%, utilizando o método descrito por Zambolim & Chaves (1974), antes da inoculação nas plantas. Somente os urediniósporos com viabilidade superior a 30% foram considerados adequados para iniciar o trabalho. A metodologia consistiu em obter uma suspensão de urediniósporos a uma concentração de 2mg/ml em água destilada, acrescentando-se *tween* (0,01%). Espalhou-se 300µL dessa suspensão na superfície das placas de *petri*, as quais foram incubadas por 24 horas no escuro a 22°C. Após esse período, procedeu-se a contagem dos urediniósporos com auxílio de microscópio óptico.

2.6. Material vegetal e experimentação para o PCR em tempo real

Foram utilizados três genótipos de cafeeiro resistentes à raça XXXIII de *H. vastatrix* e três suscetíveis (Tabela 2). As plantas inoculadas com o fungo foram mantidas em câmara de crescimento com temperatura e luminosidade controladas, utilizando a mesma metodologia para a obtenção dos esporos, descrita no item 2.2. Foi utilizado o delineamento experimental inteiramente casualizado (DIC), com três réplicas biológicas para cada tempo de coleta após a inoculação em folhas jovens totalmente expandidas. As coletas foram realizadas em 0, 12, 24 e 96 horas após a inoculação (hai), sendo o tempo de 0 hora utilizado como referência para as análises de expressão gênica temporal. Durante as coletas, as amostras (folhas) foram imediatamente congeladas em N₂ líquido e mantidas em ultrafreezer -80 °C até a utilização.

Tabela 2: Variedades de cafeeiro utilizadas para análise de expressão gênica.

Identificação	Descrição	Fenótipo	S_H*
CIFC 832/2	Híbrido de Timor	Resistente	5,6,7,8,9,?
CIFC 110/5	S4 Agaro	Resistente	4,5
CIFC H419/20	Mundo Novo x HW26/13 (Caturra x CIFC 832/1)	Resistente	5,6,9,?
Catuai IAC 44	Cultivar	Suscetível	5
CIFC 7963/117	Catimor (Caturra x CIFC 832/1)	Suscetível	5, 7 ou 5,7,9
CIFC 849/1	Matari	Suscetível	?

* genes de resistência inferidos por meio de inoculação de diferentes raças de *Hemileia vastatrix* e teoria gene-a-gene de Flor (1956).

2.7.Extração de RNA

As folhas, previamente congeladas, foram maceradas em nitrogênio líquido com ajuda de cadinho e pistilo devidamente esterilizados. Durante a maceração foi adicionada uma alíquota de PVPP (Polivinilpolipirrolidona). Aproximadamente 100mg de folhas maceradas foram destinada a extração de RNA utilizando-se o Concert® (Invitrogen) como extrator, seguindo as recomendações do fabricante.

O material vegetal foi alíquotado em tubos de microcentrífuga de 2,0 mL de capacidade. Em seguida, foi adicionado 600 µL de tampão de extração Concert®. A amostra foi homogeneizada por 2 min utilizando um vórtex e mantida a temperatura ambiente por 5 min. Depois da incubação, o material foi centrifugado por 15 min (14000 rpm a 4 °C). O sobrenadante foi coletado e transferido para novo tubo contendo 100 µL de NaCl 5M gelado. Em seguida, foi adicionado 300 µL de clorofórmio gelado ao tubo, e este homogeneizado em vórtex por 20s. O material foi centrifugado por 10 min (14000 rpm a 4 °C), a fase superior foi coletada e transferida para novo tubo. Esse passo foi repetido até que o sobrenadante se mostrou totalmente limpo. A ele foi adicionado um volume igual ao coletado (400/450 µL) de isopropanol gelado. A mistura foi homogeneizada por inversão e incubada por 20 min a -20 °C, seguida de uma nova centrifugação por 15 min (14000 rpm a 4 °C). O sobrenadante foi descartado e ao pellet foi adicionado 1 ml de etanol 80% gelado. Em seguida foi realizada nova centrifugação por 2 min (14000 rpm a 4 °C). O sobrenadante foi descartado e o pellet seco à temperatura ambiente. O pellet foi ressuscitado em 20 µL de água tratada com DEPC (Diethyl pyrocarbonate) e armazenado a -80 °C até o momento do uso.

O RNA total foi quantificado com o auxílio do Qubit RNA BR (*Life Technologies*) e sua integridade foi avaliada por eletroforese em gel de agarose (1,5%) corado com brometo de etídio (Wilson e Walker, 2000).

2.8. Análises de expressão gênica

A síntese de cDNA foi realizada com 1 µg de RNA total, pré-tratado com 1 µL de DNase à 37 °C por 30 min, utilizando RQ1 RNase- Free DNase (Promega), para retirar os possíveis contaminantes de DNA genômico. A síntese da primeira fita de cDNA foi feita utilizando o kit ImProm-II™ *Reverse Transcription System Protocol RT-PCR* (Promega), segundo as orientações do fabricante. O cDNA foi armazenado a -20 °C até a utilização.

Para a realização da técnica de PCR quantitativo em tempo real, em aparelho 7500 Real Time PCR Systems (*Applied Biosystems*), foi utilizado o sistema de detecção de fluorescência SYBR Green I (*Applied Biosystems, California, USA*). Para cada reação foi utilizado 1µl (25ng) da diluição da reação de síntese de cDNA de fita simples, 1µl (2µM) de cada iniciador, 4µl de GoTaq® qPCR Master Mix (*Promega*) e 3µl de água estéril, para um volume final de 10µl. As condições térmicas da reação foram: 95°C por 10 minutos para a desnaturação inicial, seguido de 40 ciclos de 95°C por 15 segundos e 60°C por 60 segundos e finalizando com 95°C por 15 segundos. O nível de expressão dos genes foi calculado utilizando os valores médios de Ct resultante de três réplicas biológicas e três réplicas técnicas. As análises estatísticas foram realizadas usando o *GraphPad Prism* (versão 5; GraphPad Software Inc.; La Jolla, CA, EUA). As médias foram comparadas usando o Teste de Tukey ($p < 0,05$). Para a normalização dos dados foram utilizados 2 genes constitutivos: GAPDH e UBQ, cujas expressões foram encontradas estáveis (Cruz et al., 2009).

3. RESULTADOS

A fim de identificar o RGA envolvido na resposta de resistência do cafeeiro HdT CIFC 838/2 à raça XXXIII de *H. vastatrix*, as sequências dos três RGAs (*RGA2*, *RGA3* e *RGA4*) identificadas no loco foram alinhadas e comparadas. *Primers* específicos para cada sequência foram contruídos (Tabela 1) e utilizados para a amplificação das

regiões correspondentes em 22 clones de cafeeiros diferenciadores de raças de *H. vastatrix* (Tabela 2). A análise indicou que os três *RGAs* estão presentes em todos os clones diferenciadores, não apresentando polimorfismo entre os cafeeiros (Figura 1).

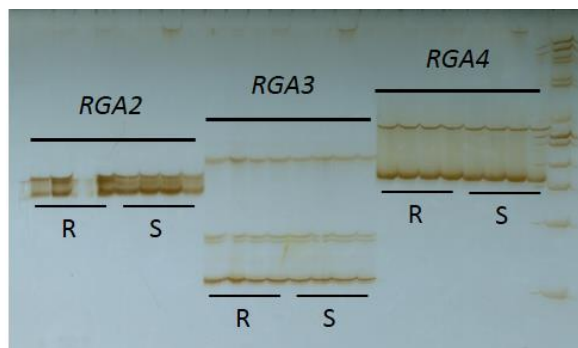


Figura 1: Amplificação dos três *RGAs* candidatos. Gel de poliacrilamida (6%). R corresponde a quatro cafeeiros resistentes: 419/20, 832/1, 832/2 e 1343/269. S corresponde a quatro cafeeiros suscetíveis: CT 419, 19/1, 849/1 e 7963. Esse mesmo padrão de amplificação foi observado nos 25 cafeeiros analisados. Marcador de peso molecular (100 pb) está representado no último poço do lado direito.

Como os três *RGAs* estão presentes em todos os clones diferenciadores, foi realizada a análise de expressão desses genes para verificar se existe diferença que possa justificar a resposta de compatibilidade ou incompatibilidade na interação cafeeiro-*H. vastatrix*. Analisando os genótipos resistentes, observou-se que a expressão do gene *RGA2* não apresentou diferença significativa ao longo das horas após a infecção (hai) para os cafeeiros CIFC 832/2 e CIFC 110/5. No genótipo CIFC H419/20 ocorreu redução significativa da expressão às 12 hai. Em genótipos suscetíveis foi observado aumento significativo da expressão às 12 hai para CIFC 849/1, às 24 hai para CIFC 7963/117 e às 96 hai para Catuaí IAC144 (Figura 2).

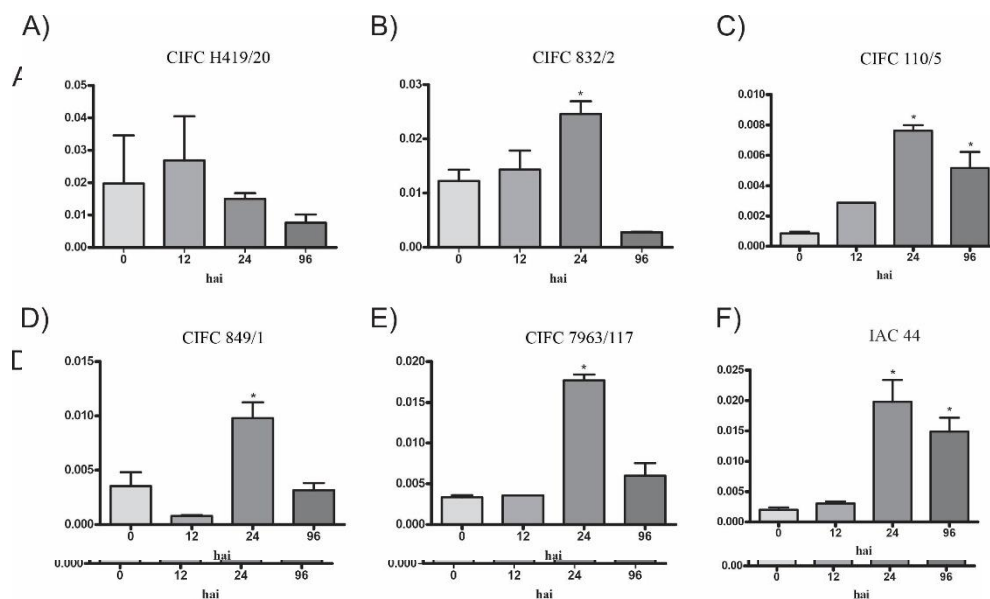


Figura 2: Expressão relativa do *RGA2*. A-C) Genótipos resistentes e D-F) Genótipos suscetíveis à raça XXXIII de *Hemileia vastatrix*. O eixo y representa o nível de expressão em relação aos genes endógenos em quatro momentos, 0, 12, 24 e 96 horas após a inoculação (hai) por *Hemileia vastatrix*. Barras de erros = SEM, n = 3 réplicas biológicas independentes. Asterisco (*) indica diferença significativa do nível de expressão em relação a amostra controle (0 hai).

Para o *RGA3*, com exceção do cafeeiro CIFC H419-20 que não apresentou diferença significativa de expressão ao longo das horas após inoculação, os demais cafeeiros, tanto resistentes quanto suscetíveis, apresentaram aumento de expressão às 24 hai (Figura 3). Portanto, não foi observada diferença de expressão entre as interações compatíveis e incompatíveis.

Figura 3: Expressão relativa do *RGA3*. A-C) Genótipos resistentes e D-F) Genótipos suscetíveis à raça XXXIII de *Hemileia vastatrix*. O eixo y representa o nível de expressão em relação aos genes endógenos em quatro momentos, 0, 12, 24 e 96 horas após a inoculação (hai) por *Hemileia vastatrix*. Barras de erros = SEM, n = 3 réplicas biológicas independentes. Asterisco (*) indica diferença significativa do nível de expressão em relação a amostra controle (0 hai).

Para o gene *RGA4* foi observado o aumento de expressão significativo às 12 hai nos três genótipos resistentes. Nos genótipos suscetíveis não foi observada diferença significativa de expressão no genótipo CIFC 849/1, o nível de expressão diminuiu significativamente no genótipo CIFC 7963/117 às 96 hai e aumentou no genótipo IAC 44 às 96 hai (Figura 4).

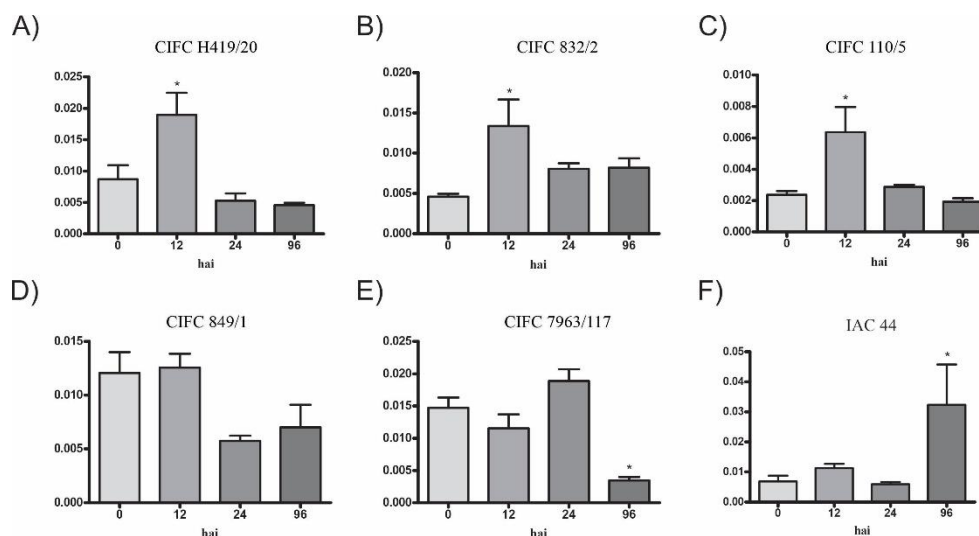


Figura 4: Expressão relativa do *RGA4*. A-C) Genótipos resistentes e D-F) Genótipos suscetíveis à raça XXXIII de *Hemileia vastatrix*. O eixo y representa o nível de expressão em relação aos genes endógenos em quatro momentos, 0, 12, 24 e 96 horas após a inoculação (hai) por *Hemileia vastatrix*. Barras de erros = SEM, n = 3 réplicas biológicas independentes. Asterisco (*) indica diferença significativa do nível de expressão em relação a amostra controle (0 hai).

Para a melhor visualização do perfil de expressão gênica dos *RGAs*, os resultados obtidos foram analisados com a função *Heatmap* do pacote *ComplexHeatmap*, em ambiente R. Essa função possibilita agrupar o conjunto de dados utilizando diferentes

métodos de distância (aqui utilizado a distância de Pearson) e representar o perfil de expressão na forma de *heatmap*. Para o agrupamento hierárquico (HCL) foi utilizado o método completo, que possibilitou agrupar os genótipos em função da similaridade do perfil de expressão de cada gene.

Na análise da expressão do gene *RGA2*, dois cafeeiros resistentes, CIFIC H419/20 e CIFIC 832/2, ficaram agrupados. Os genótipos IAC 44 (Catuaí) e CIFIC 849/1 (Matari), considerados suscetíveis para a maioria das raças de *H. vastatrix*, ficaram em um grupo diferente. Um terceiro grupo foi formado pelo cafeeiro resistente CIFIC 110/5 e o CIFIC 7963/117, que apesar de suscetível à raça XXXIII, possui dois ou três genes de resistência caracterizados (Figura 5A).

Para o *RGA3*, os agrupamentos formados não apresentaram relação com a resistência e suscetibilidade dos genótipos (Figura 5B). Na expressão do gene *RGA4*, foram formados dois grandes grupos, um contendo os três genótipos resistentes e outro contendo os três genótipos suscetíveis à raça XXXIII de *H. vastatrix*. No *cluster* dos cafeeiros resistentes, os genótipos CIFIC H419/20 e CIFIC 832/2 ficaram em um subgrupo e o CIFIC 110/5 em outro. O *cluster* dos cafeeiros suscetíveis ficou dividido em dois subgrupos, um contendo os genótipos CIFIC 7963/117 e IAC 44 (Catuaí) e outro contendo apenas o CIFIC 849/1 (Figura 5C). Por meio da análise de *Heatmap* também foi possível observar que os níveis de expressão do gene *RGA4* foi maior às 12 hai, que corresponde ao estágio inicial da infecção do fungo.

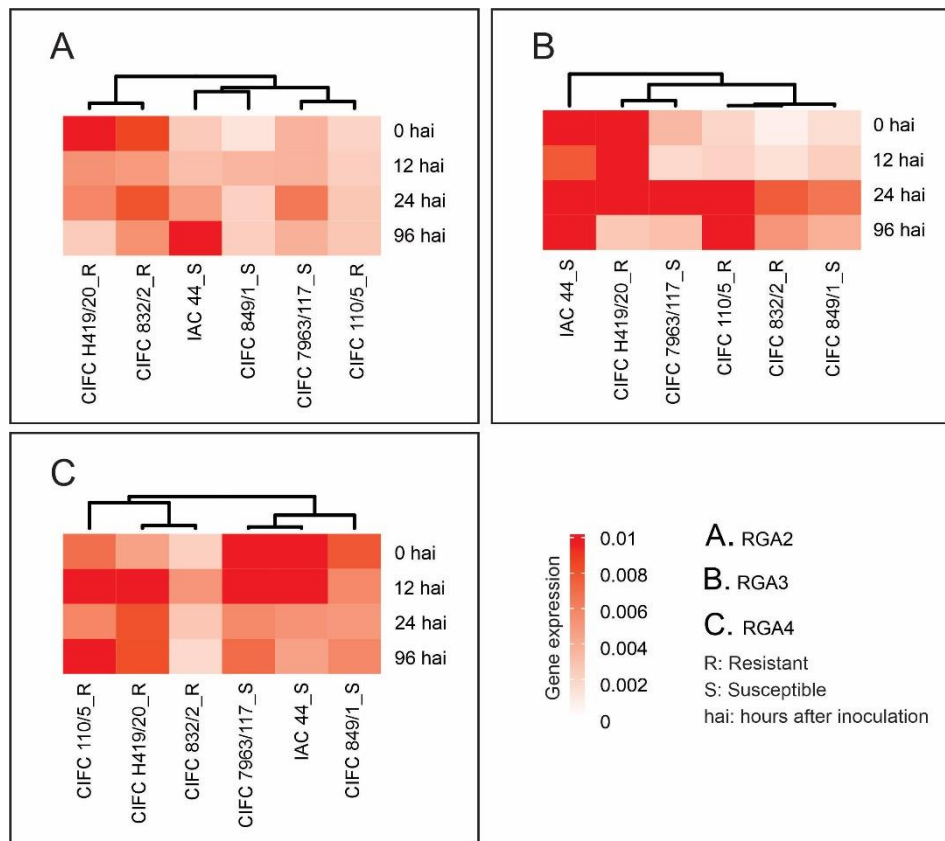


Figura 5: Perfil de expressão dos *RGAs* e clusterização hierárquica de acordo com o perfil de expressão ao logo do tempo de infecção.

4. DISCUSSÃO

Com o objetivo de relacionar os genes S_H , previamente caracterizados nos cafeeiros diferenciadores de raças de *H. vastatrix*, com os *RGAs* clonados, *primers* específicos para cada *RGA* foram construídos e utilizados para amplificação do DNA desses diferenciadores. Os três *RGAs* analisados foram clonados de uma região do genoma do HdT C1FC 832/2. Esse cafeeiro tem sido considerado uma das principais fontes de resistência à *H. vastatrix* nos programas de melhoramento do cafeeiro, não tendo nenhuma raça brasileira suplantado sua resistência até o momento.

Por meio da amplificação dos *primer* desenhados, observou-se que os três *RGAs* (*RGA2*, *RGA3* e *RGA4*) estão presentes em todos os cafeeiros diferenciadores de raça de *H. vastatrix*. Portanto, não foi possível relacionar os genes clonados com os S_H previamente inferidos usando essa metodologia. Análise de expressão diferencial dos genes foi então realizada para a obtenção do perfil de expressão gênica em interações

compatíveis (genótipos suscetíveis) e incompatíveis (genótipos resistentes). Estudos da expressão gênica na interação cafeeiro-*Hemileia* tem permitido a identificação de genes diferencialmente expressos ao longo do processo infeccioso. Em cafeeiros infectados pelo fungo foi verificado a expressão de genes relacionados com os processos de reconhecimento, sinalização, comunicação celular, defesa e metabolismo (Fernandez et al., 2004; Guzzo et al., 2009; Diola et al., 2013).

No presente trabalho, as análises de expressão do *RGA2* na interação compatível indicaram o aumento de sua expressão às 12 hai no genótipo CIFC 849/1, às 24 hai no genótipo CIFC 7963/117 e às 96 hai no genótipo IAC 44. Curiosamente, na interação incompatível não foram observadas diferenças significativas de expressão (para os genótipos CIFC 832/2 e CIFC 110/5) ou houve decréscimo no nível de expressão do gene ao longo das horas de infecção (CIFC H419/20). Além disso, a análise de clusterização hierárquica dos dados de expressão do *RGA2* não permitiu separar os genótipos resistentes e os suscetíveis. Esses resultados sugerem não ser o *RGA2* o gene envolvido na resistência do HdT CIFC 832/2 à raça XXXIII de *H. vastatrix*. Durante a resistência do hospedeiro, em geral, os genes de resistência aumentam o nível de expressão para permitir resposta efetiva de defesa. Em genótipos de cafeeiros resistentes à *H. vastatrix*, o crescimento do fungo é impedido nas primeiras fases do processo infeccioso devido à reação de hipersensibilidade - HR (Silva et al., 2002, 2006). Em cafeeiros suscetíveis à doença, o processo de infecção se completa com a esporulação sobre a superfície externa da folha, via estômatos, em torno de 20 dias após a penetração do fungo (Silva et al., 1999). Dessa forma, o esperado é que a planta não gaste energia com a expressão de genes de resistência ao longo do tempo em genótipos suscetíveis, já que o patógeno superou as barreiras iniciais à infecção.

A análise de expressão do *RGA3* mostrou um perfil semelhante nas duas interações, compatível e incompatível, sendo observado um aumento da expressão do gene às 24 hai. Observações citológicas, realizadas por Lopes (2015), mostraram que no mesmo patossistema, a resistência do cafeeiro HdT (CIFC 832/1) à *H. vastatrix* raça XXXIII é pré-haustorial, podendo ser observada com 17 hai. Portanto, por apresentar aumento significativo de expressão às 24 hai, e não apresentar diferença entre os genótipos resistentes e suscetíveis, provavelmente, o *RGA3* é um gene essencial à planta, porém pouco relacionado com a resposta de resistência efetiva.

O perfil de expressão do *RGA4* se apresentou diferente para os dois tipos de interação. Na interação compatível, não foi observada diferença significativa de expressão (genótipo CIFC 849/1), a expressão diminuiu de forma significativa (genótipo CIFC 7963/117) ou aumentou na fase tardia da infecção (genótipo IAC 44), às 96 hai. Contudo, na interação incompatível, o gene apresentou aumento significativo de expressão na fase inicial da interação, às 12 hai, em todos os genótipos. O estágio inicial da infecção (12 e 24 hai) é considerado a fase pré-haustorial, pois os haustórios não estão completamente formados, enquanto que o estágio tardio da infecção (acima de 24 hai) é considerado a fase pós-haustorial, quando espera-se que as estruturas já estejam completamente formadas (Silva et al., 2006; Freitas et al., 2014; Florez et al., 2017). O aumento da expressão desse gene nas horas iniciais da infecção pode ser responsável pelo sucesso da resistência da planta e corrobora com a resistência pré-haustorial sugerida para essa interação. Outros estudos sugeriram a resistência pré-haustorial na interação cafeeiro-*Hemileia* (Lopes, 2015; Florez et al., 2017).

Foi relatado que em cafeeiros com resistência completa à *H. vastatrix*, a invasão do fungo cessa nos estágios iniciais da infecção onde ocorre a perturbação do conteúdo citoplasmático e das estruturas de infecção do fungo (Diniz et al., 2012). Portanto, a regulação positiva da expressão do *RGA4* pode ser uma consequência direta da infecção pelo fungo e pode ser responsável por uma defesa rápida e eficaz. A resposta pré-haustorial também foi observada em outros patossistemas. Na interação cevada-*Puccinia graminis f. sp. tritici*, o gene de resistência *RPG1* foi expresso alguns minutos após a inoculação com raças de *P. graminis* (Nirmala et al., 2010).

A análise de agrupamento hierárquico corrobora com a indicação de ser o *RGA4* o gene envolvido na resposta de resistência do cafeeiro HdT CIFC 832/2 à raça XXXIII de *H. vastatrix*, pois observou-se separação dos genótipos resistentes e genótipos suscetíveis. Os dados obtidos não permitem afirmar que esse é um gene *R* e está ligado ao reconhecimento específico de efetores do patógeno, mas sugere que ele desempenha função importante no processo de defesa da planta.

5. REFERÊNCIAS

- BETTENCOURT, A.J.; RODRIGUES JR. C.J. Principles and practice of coffee breeding for resistance to rust and other diseases. In: CLARKE, R. J.; MACRAE, R. (Eds). Coffee Agronomy, Elsevier Applied Science Publishers LTD, London and New York, v. 4, p. 199-234. 1988.
- CAPUCHO, A.S.; CAIXETA, E.T.; ZAMBOLIM, E.M. & ZAMBOLIM, L. Herança da resistência do Híbrido de Timor UFV 443-03 à ferrugem do cafeeiro. Pesq. Agrop. Bras., Brasília. 44: 276-282. 2009.
- CAPUCHO, A.S.; MACIEL-ZAMBOLIM, E.; FREITAS R.L.; HADDAD, F.; CAIXETA, E.T.; ZAMBOLIM L. Identification of race XXXIII of *Hemileia vastatrix* on *Coffea arabica* Catimor derivatives in Brazil. Austral Plant Dis Notes 7:189-191. 2012.
- COLL, N.S.; EPPLER P.; DANGL, J.L. Programmed cell death in the plant immune system. Cell Death and Differentiation 18: 1247-1256. 2011.
- CHARRIER, A.; BERTHAUD, J. Botanical classification of *Coffea*. In: CLIFFORD, MN; WILLSON, KC. Coffee: botany, biochemistry and production of beans and beverage. Inc. Westport, Connecticut, USA: The AVI Publishing Company, p. 13-47. 1985.
- CRUZ, F.; KALAOUN, S.; NOBILE, P. et al. Evaluation of coffee reference genes for relative expression studies by quantitative real-time RT-PCR. Mol Breed 23:607-616. 2009. doi: 10.1007/S11032-009-9259-X
- DIOLA, V. et al. A new set of differentially expressed signaling genes is early expressed in coffee leaf rust race II incompatible interaction. Functional & Integrative Genomics 13: 379-389. 2013.
- DINIZ, L.E.C.; SAKIYAMA, N.S.; LASHERMES, P.; CAIXETA, E.T.; OLIVEIRA, A.C.B.; MACIEL-ZAMBOLIM, E.; LOUREIRO, M.E.; PEREIRA, A.A.; ZAMBOLIM, L. Analysis of AFLP markers associated to the Mex-1 resistance locus in Icatu progenies. Crop Breeding and Applied Biotechnology 5: 387-393. 2005.

ELLIS, J.G. et al. Recent progress in discovery and functional analysis of effector proteins of fungal and oomycete plant pathogens. *Current Opinion in Plant Biology* 12: 399-405. 2009

FERNANDEZ, D. et al. Coffee (*Coffea arabica* L.) genes early expressed during infection by the rust fungus (*Hemileia vastatrix*). *Molecular Plant Pathology* 5: 527-536. 2004

FLOREZ, J.C.; MOFATTO, L.S.; LOPES, R.L.F.; FERREIRA, S.S.; ZAMBOLIM, E.M.; CARAZZOLLE, M.F.; ZAMBOLIM, L.; CAIXETA, E.T. High throughput transcriptome analysis of coffee reveals prehaustorial resistance in response to *Hemileia vastatrix* infection. *Plant Mol Biol* 95: 607-623. 2017. doi: 10.1007/s11103-017-0676-7

FREITAS, R.L.; ZAMBOLIM, E.; SILVA, M.C.; CAIXETA, E.T.; LELIS, D.T.; ZAMBOLIM, L.; SAKIYAMA, N. Cytological evaluation of the infection process of *Hemileia vastatrix* (race XXXIII) in resistant and susceptible coffee. In: Proceedings of the 25th international conference on coffee science (ASIC), p. 42-46. Armenia, Colombia. 2014

GUZZO, S.D.; HAKAKAVA, R.; TSAI, S.M. Identification of coffee genes expressed during systemic acquired resistance and incompatible interaction with *Hemileia vastatrix*. *Journal of Phytopathology* 157: 625-638. 2009

HUGHES, A. L. Gene duplication and the origin of novel proteins. *Proceedings of the National Academy of Sciences of the USA* 102: 8791-8792. 2005

KONDRASHOV, F.A. Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proc Biol Sci* 279: 5048-5057. 2012. doi: 10.1098/rspb.2012.1108

LOPES, R.L.F. Análise citológica e perfil de expressão gênica de *Hemileia vastatrix* (raça XXXIII) na interação com o cafeeiro. Tese de Doutorado. UFV. 2015

MAGADUM, S.; BANERJEE, U.; MURUGAN, P.; GANGAPUR, D.; RAVIKESAVAN, R. Gene duplication as a major force in evolution. *J. Genet.* 92: 155-161. 2013. doi: 10.1007/s12041-013-0212-8

NIRMALA, J.; DRADER, T.; CHEN, X.; STEFFENSON, B.; KLEINHOF, A. Stem rust spores elicit rapid RPG1 phosphorylation. *Mol Plant-Microbe Interact* 23: 1635-1642. 2010. doi:10.1094/MPMI-06-10-0136

PANCHY, N.; LEHTI-SHIU, M.; SHIU, S.H. Evolution of gene duplication in plants. *Plant Physiol.* 171: 2294-2316. 2016. doi: 10.1104/pp.16.00523

REDDY, A.C.; NARESH, P.; REDDY, D.C.L. The crucial role of R-genes/RGAs in host-microbial interactions and plant immunity. *Research Journal of Biotechnology* 13: 4. 2018

RODRIGUES JR., C.J.; BETTENCOURT, A. J.; RIJO, L. Races of the pathogen and resistance to coffee rust. *Annual Review of Phytopathology* 13: 49-70. 1975

SEKHWAL, M.K.; LI, P.; LAM, I.; WANG, X.; CLOUTIER, S.; YOU, F.M. Disease Resistance Gene Analogs (RGAs) in Plants. *Int. J. Mol. Sci.* 16: 19248-19290. 2015. doi:10.3390/ijms160819248

SIEVERS, F.; WILM, A.; DINEEN, D.; GIBSON, T.J.; KARPLUS, K.; LI, W.; LOPEZ, R.; MCWILLIAM, H.; REMMERT, M.; SÖDING, J.; THOMPSON, J.D.; HIGGINS, D.G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7: 539. 2011. doi: 10.1038/msb.2011.75

SILVA, M.C.; VÁRZEA, V.; GUIMARÃES, L.G.; AZINHEIRA, H.G.; FERNANDEZ, D.; PETITOT, A.; BERTRAND, B.; LASHERMES, P.; NICOLE, M. “Coffee resistance to the main diseases: leaf rust and coffee berry disease”, *Brazilian Journal of Plant Physiology* 18: 119-147. 2006

SILVA, M.C. et al. Cytochemical aspects of the plant–rust fungus interface during the compatible interaction *Coffea arabica* (cv. Caturra) – *Hemileia vastatrix* (race III). *International Journal of Plant Sciences* 160: 79-91. 1999

SILVA, M.C. et al. Hypersensitive cell death and post-haustorial defence responses arrest the orange rust (*Hemileia vastatrix*) growth in resistant coffee leaves. *Physiological and Molecular Plant Pathology* 60: 169-183. 2002

STERGIOPOULOS, I.; DE WIT, P.J. Fungal effector proteins. Annual Review of Phytopathology 47: 233-263. 2009

UNTERGASSER, A.; CUTCUTACHE, I.; KORESSAAR, T.; YE, J.; FAIRCLOTH, B.C.; REMM, M.; ROZEN, S.G. Primer3 - new capabilities and interfaces. Nucleic Acids Research 40(15):e115. 2012

VÁRZEA, V.M.P.; MARQUES, D.V. Population variability of *Hemileia vastatrix* vs. coffee durable resistance. In: ZAMBOLIM, L.; ZAMBOLIM, E.M.; VÁRZEA, V.M.P. (Ed.). Durable resistance to coffee leaf rust. Viçosa: UFV. p. 53-74. 2005

WILSON K & WALKER K. Principles and techniques of practical biochemistry. 5a ed. Cambridge: Cambridge University Press, p. 784, 2000

ZAMBOLIM, L; CHAVES, G.M. Efeito de baixas temperaturas e do binômio temperatura-umidade relativa sobre a viabilidade dos uredósporos de *Hemileia vastatrix* Berk. et Br. e *Uromyces phaseolityca* Arth. Experientiae 17: 151-184. 1974.

ZAMBOLIM, L. et al. Manejo integrado das doenças do cafeeiro. In: ENCONTRO SOBRE PRODUÇÃO DE CAFÉ COM QUALIDADE, 1999, Viçosa, MG. Anais. Viçosa, MG: UFV, 1999. v. 1, p. 134-215.

ZAMBOLIM L. Current Status and management of coffee leaf rust in Brazil. Tropical Plant Pathology 41: 1-8. 2016. doi 10.1007/S40858-016-0065-9

CONCLUSÕES GERAIS

Neste trabalho, foi identificada uma região no genoma de HdT CIFIC 832/2 com três *RGAs* (*RGA2*, *RGA3* e *RGA4*) com potencial envolvimento na resposta de defesa da planta à *Hemileia vastatrix*. Foram obtidas informações importantes sobre a estrutura desse loco e a sua provável origem da espécie *Coffea arabica*. Análises de expressão gênica permitiram identificar o *RGA4* como o provável gene que mantém a função relacionada a resistência. O perfil de expressão desse gene foi capaz de separar os genótipos resistentes dos genótipos suscetíveis à *H. vastatrix*. Dessa forma, esse gene tem potencial para ser utilizado como biomarcador de resistência do cafeeiro à *H. vastatrix* raça XXXIII.

APÊNDICES

Apêndice 1: Tabela de anotação dos genes preditos no Clone 101-2P. Em amarelo está o loco *RGA* e em laranja estão os genes selecionados para análise filogenética relacionados com o metabolismo primário da planta.

ORF	e-value	Anotação	Espécie
NODE_1_length_56551_cov_82.9898:g1.t1	0.0	disease resistance protein RGA2-like isoform X3	<i>C. arabica</i>
NODE_1_length_56551_cov_82.9898:g2.t1	0.0	vinorine synthase-like	<i>C. arabica</i>
NODE_1_length_56551_cov_82.9898:g3.t1	0.0	putative disease resistance protein RGA4 isoform X1	<i>C. arabica</i>
NODE_1_length_56551_cov_82.9898:g3.t2	0.0	putative disease resistance protein RGA4 isoform X1	<i>C. arabica</i>
NODE_1_length_56551_cov_82.9898:g4.t1	0.0	uncharacterized protein LOC113718514	<i>C. arabica</i>
NODE_3_length_11907_cov_75.166:g5.t1	0.0	putative disease resistance protein RGA3	<i>C. arabica</i>
NODE_3_length_11907_cov_75.166:g5.t2	0.0	putative disease resistance protein RGA3	<i>C. arabica</i>
NODE_6_length_4548_cov_104.836:g12.t1	1.00E-76	uncharacterized protein LOC113708274	<i>C. arabica</i>
NODE_6_length_4548_cov_104.836:g13.t1	4.00E-82	uncharacterized protein LOC113714139	<i>C. arabica</i>
NODE_19_length_947_cov_1.26667:g24.t1	3.00E-41	AAA-type ATPase family protein	<i>C. canephora</i>
NODE_34_length_789_cov_1.11236:g37.t1	2.00E-102	LOW QUALITY PROTEIN: transketolase%2C chloroplastic-like	<i>C. arabica</i>
NODE_36_length_777_cov_1.67429:g39.t1	2.00E-09	uncharacterized protein LOC113768482	<i>C. eugenioides</i>
NODE_41_length_756_cov_0.726068:g44.t1	9.00E-37	Succinate dehydrogenase [ubiquinone] flavoprotein subunit 1%2C mitochondrial	<i>C. canephora</i>
NODE_52_length_695_cov_1.26052:g55.t1	3.00E-10	uncharacterized protein LOC113714019	<i>C. arabica</i>
NODE_55_length_681_cov_1.02318:g57.t1	2.00E-61	6-phosphogluconate dehydrogenase%2C decarboxylating 1	<i>C. arabica</i>
NODE_60_length_668_cov_0.988156:g62.t1	1.00E-11	alpha-glucosidase-like	<i>C. eugenioides</i>
NODE_63_length_658_cov_0.886403:g65.t1	5.00E-31	sulfite reductase 1 [ferredoxin]%2C chloroplastic-like	<i>C. arabica</i>
NODE_65_length_653_cov_1.11632:g67.t1	2.00E-62	Elongation factor Tu%2C mitochondrial	<i>C. canephora</i>
NODE_66_length_652_cov_1.26783:g68.t1	7.00E-43	epimerase family protein SDR39U1 homolog%2C chloroplastic isoform X2	<i>C. eugenioides</i>
NODE_70_length_638_cov_1.47772:g72.t1	8.00E-15	ribonuclease 3-like protein 3 isoform X2	<i>C. eugenioides</i>

NODE_74_length_630_cov_1.10669:g76.t1	6.00E-08	ribulose-phosphate 3-epimerase%2C cytoplasmic isoform-like	<i>C. eugenioides</i>
NODE_75_length_623_cov_1.02381:g77.t1	6.00E-08	DNA mismatch repair protein MSH7	<i>C. arabica</i>
NODE_78_length_614_cov_1.17691:g79.t1	7.00E-22	Glycine--tRNA ligase 2%2C chloroplatic/mitochondrial	<i>C. canephora</i>
NODE_79_length_612_cov_1.36822:g80.t1	1.00E-09	two-component response regulator-like PRR37 isoform X2	<i>C. arabica</i>
NODE_82_length_608_cov_1.16008:g83.t1	2.00E-47	fumarate hydratase 1%2C mitochondrial	<i>C. eugenioides</i>
NODE_83_length_607_cov_1.64528:g84.t1	3.00E-10	ERAD-associated E3 ubiquitin-protein ligase component HRD3A isoform X2	<i>C. arabica</i>
NODE_87_length_599_cov_1.04789:g87.t1	2.00E-72	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	<i>C. eugenioides</i>
NODE_91_length_592_cov_1.29126:g91.t1	2.00E-29	LOW QUALITY PROTEIN: ornithine carbamoyltransferase%2C chloroplatic-like	<i>C. eugenioides</i>
NODE_93_length_591_cov_1.11845:g93.t1	2.00E-59	Acetate--CoA ligase ACS%2C chloroplatic/glyoxysomal	<i>C. canephora</i>
NODE_97_length_588_cov_1.08023:g97.t1	6.00E-42	thylakoid membrane protein TERC%2C chloroplatic isoform X3	<i>C. arabica</i>
NODE_98_length_586_cov_0.866405:g98.t1	2.00E-13	Protein TRIGALACTOSYLDIACYLGLYCEROL 3%2C chloroplatic	<i>C. canephora</i>
NODE_104_length_580_cov_1.01789:g103.t1	9.00E-18	protease Do-like 1%2C chloroplatic	<i>C. arabica</i>
NODE_113_length_564_cov_0.903491:g112.t1	2.00E-14	Putative Beta-glucosidase 24	<i>C. canephora</i>
NODE_114_length_564_cov_1.5462:g113.t1	4.00E-42	Heat shock 70 kDa protein%2C mitochondrial	<i>C. canephora</i>
NODE_118_length_560_cov_1.24638:g117.t1	2.00E-17	RNA polymerase alpha subunit	<i>C. arabica</i>
NODE_119_length_560_cov_1.30228:g118.t1	1.00E-40	Lipoyl synthase%2C chloroplatic	<i>C. canephora</i>
NODE_130_length_544_cov_1.04497:g128.t1	2.00E-21	dicarboxylate transporter 1%2C chloroplatic	<i>C. eugenioides</i>
NODE_133_length_542_cov_1.05161:g131.t1	3.00E-112	uncharacterized protein LOC113758044	<i>C. eugenioides</i>
NODE_135_length_540_cov_1.10799:g133.t1	2.00E-18	RNA pseudouridine synthase 7	<i>C. eugenioides</i>
NODE_138_length_537_cov_0.984783:g136.t1	2.00E-52	Aspartate aminotransferase%2C cytoplasmic	<i>C. canephora</i>
NODE_152_length_521_cov_0.995495:g147.t1	6.00E-17	ABC transporter F family member 3-like	<i>C. arabica</i>
NODE_154_length_520_cov_1:g149.t1	1.00E-19	LOW QUALITY PROTEIN: ABC transporter F family member 4-like	<i>C. arabica</i>
NODE_162_length_516_cov_1.75399:g155.t1	1.00E-56	RNA polymerase beta' subunit	<i>C. arabica</i>
NODE_167_length_512_cov_1.57931:g160.t1	1.00E-08	Acetolactate synthase 2%2C chloroplatic	<i>C. canephora</i>

NODE_168_length_511_cov_1.00922:g161.t1	4.00E-11	uncharacterized protein LOC113751769	<i>C. eugenioides</i>
NODE_177_length_504_cov_1.38642:g169.t1	1.00E-14	mannose-6-phosphate isomerase 1	<i>C. eugenioides</i>
NODE_179_length_502_cov_1.89647:g171.t1	1.00E-13	histidine kinase 5-like	<i>C. arabica</i>
NODE_181_length_501_cov_0.728774:g173.t1	7.00E-12	putative multidrug resistance protein	<i>C. arabica</i>
NODE_194_length_493_cov_0.903846:g184.t1	3.00E-10	Putative N-acylneuraminate-9-phosphatase	<i>C. canephora</i>
NODE_197_length_493_cov_1.0601:g187.t1	8.00E-12	Probable anion transporter 2%2C chloroplastic	<i>C. canephora</i>
NODE_201_length_490_cov_0.830508:g190.t1	3.00E-07	isoamylase 1%2C chloroplastic-like isoform X2	<i>C. arabica</i>
NODE_203_length_490_cov_0.975787:g193.t1	4.00E-10	D-lactate dehydrogenase [cytochrome]%2C mitochondrial	<i>C. arabica</i>
NODE_206_length_489_cov_1.37136:g196.t1	2.00E-33	Thioredoxin reductase 2	<i>C. canephora</i>
NODE_206_length_489_cov_1.37136:g196.t2	1.00E-26	thioredoxin reductase NTRC-like	<i>C. arabica</i>
NODE_211_length_486_cov_0.723716:g201.t1	2.00E-09	4-coumarate--CoA ligase-like 10	<i>C. canephora</i>
NODE_224_length_475_cov_1.26131:g212.t1	2.00E-19	adenine DNA glycosylase-like isoform X2	<i>C. arabica</i>
NODE_228_length_474_cov_0.858942:g214.t1	2.00E-16	uncharacterized protein LOC113772441 isoform X1	<i>C. eugenioides</i>
NODE_231_length_473_cov_0.772727:g217.t1	2.00E-22	probable uridine nucleosidase 2	<i>C. eugenioides</i>
NODE_234_length_471_cov_1.1269:g220.t1	9.00E-14	extradiol ring-cleavage dioxygenase-like isoform X1	<i>C. eugenioides</i>
NODE_239_length_468_cov_1.18159:g225.t1	8.00E-29	LOW QUALITY PROTEIN: glycine--tRNA ligase%2C chloroplastic/mitochondrial 2	<i>C. eugenioides</i>
NODE_240_length_467_cov_0.915385:g226.t1	8.00E-23	GTP-binding protein OBGC%2C chloroplastic	<i>C. arabica</i>
NODE_241_length_466_cov_0.930591:g227.t1	4.00E-35	ribosomal protein L2	<i>C. arabica</i>
NODE_242_length_466_cov_0.822622:g228.t1	3.00E-32	Putative GTP pyrophosphokinase	<i>C. canephora</i>
NODE_243_length_465_cov_1.58763:g229.t1	3.00E-22	leucine aminopeptidase 1-like	<i>C. eugenioides</i>
NODE_245_length_464_cov_0.927649:g231.t1	1.00E-60	LOW QUALITY PROTEIN: uncharacterized protein LOC113783053	<i>C. eugenioides</i>
NODE_247_length_463_cov_1.09067:g233.t1	2.00E-23	50S ribosomal protein L1%2C chloroplastic	<i>C. eugenioides</i>
NODE_252_length_460_cov_1.15144:g238.t1	1.00E-26	glutamate synthase 1 [NADH]%2C chloroplastic isoform X1	<i>C. eugenioides</i>
NODE_254_length_460_cov_1.22454:g239.t1	2.00E-16	Acetolactate synthase 2%2C chloroplastic	<i>C. canephora</i>
NODE_257_length_458_cov_1.08136:g241.t1	3.00E-11	abscisic-aldehyde oxidase-like isoform X2	<i>C. eugenioides</i>
NODE_263_length_455_cov_1.15608:g246.t1	2.00E-84	uncharacterized protein LOC113766439	<i>C. eugenioides</i>

NODE_267_length_453_cov_1.56117:g250.t1	1.00E-49	Glucose-6-phosphate 1-dehydrogenase%2C cytoplasmic isoform	<i>C. canephora</i>
NODE_283_length_448_cov_0.797844:g259.t1	5.00E-39	Valine--tRNA ligase	<i>C. canephora</i>
NODE_292_length_441_cov_1.00275:g266.t1	2.00E-34	gamma carbonic anhydrase 1%2C mitochondrial-like	<i>C. arabica</i>
NODE_294_length_440_cov_1.22039:g267.t1	7.00E-11	LOW QUALITY PROTEIN: ABC transporter B family member 29%2C chloroplastic-like	<i>C. arabica</i>
NODE_300_length_438_cov_0.778393:g272.t1	6.00E-33	thiosulfate/3-mercaptopyruvate sulfurtransferase 1%2C mitochondrial-like	<i>C. arabica</i>
NODE_304_length_437_cov_1.53333:g276.t1	2.00E-51	DNA gyrase subunit A%2C chloroplastic/mitochondrial-like	<i>C. eugenioides</i>
NODE_308_length_436_cov_1.03064:g278.t1	1.00E-08	uncharacterized protein LOC113772609	<i>C. eugenioides</i>
NODE_309_length_435_cov_1.64804:g279.t1	5.00E-09	L-ascorbate peroxidase%2C cytosolic-like	<i>C. arabica</i>
NODE_314_length_434_cov_1.59944:g283.t1	5.00E-07	Glutathione S-transferase T1	<i>C. canephora</i>
NODE_315_length_434_cov_0.613445:g284.t1	3.00E-42	DNA mismatch repair protein MSH3-like	<i>C. arabica</i>
NODE_330_length_427_cov_1.05143:g297.t1	4.00E-24	Putative Transaldolase	<i>C. canephora</i>
NODE_334_length_426_cov_1.05731:g301.t1	4.00E-44	glyceraldehyde-3-phosphate dehydrogenase GAPCP2%2C chloroplastic-like	<i>C. arabica</i>
NODE_337_length_425_cov_1.05747:g304.t1	2.00E-13	probable enoyl-CoA hydratase 2%2C mitochondrial	<i>C. eugenioides</i>
NODE_344_length_424_cov_1.26225:g310.t1	2.00E-09	Putative 2-oxoglutarate/malate translocator%2C chloroplastic	<i>C. canephora</i>
NODE_359_length_418_cov_1.11144:g321.t1	1.00E-40	Serine hydroxymethyltransferase%2C mitochondrial	<i>C. canephora</i>
NODE_363_length_417_cov_1.19706:g324.t1	2.00E-12	probable GTP diphosphokinase RSH2%2C chloroplastic	<i>C. arabica</i>
NODE_365_length_416_cov_1.9351:g326.t1	3.00E-15	glutamate--glyoxylate aminotransferase 2	<i>C. arabica</i>
NODE_374_length_412_cov_0.859701:g334.t1	5.00E-25	glyoxylate/hydroxypyruvate reductase HPR3	<i>C. eugenioides</i>
NODE_381_length_410_cov_0.855856:g340.t1	3.00E-06	dicarboxylate transporter 2.1%2C chloroplastic	<i>C. eugenioides</i>
NODE_392_length_406_cov_0.81459:g347.t1	7.00E-25	Protein TRIGALACTOSYLDIACYLGLYCEROL 3%2C chloroplastic	<i>C. canephora</i>
NODE_409_length_402_cov_0.904615:g362.t1	5.00E-18	Probable mannitol dehydrogenase	<i>C. canephora</i>
NODE_411_length_402_cov_1.25846:g363.t1	1.00E-26	DEAD-box ATP-dependent RNA helicase 53%2C mitochondrial- like isoform X2	<i>C. arabica</i>
NODE_414_length_401_cov_0.941358:g364.t1	5.00E-15	threonine dehydratase 1 biosynthetic%2C chloroplastic isoform X2	<i>C. arabica</i>

NODE_416_length_400_cov_1.14551:g366.t1	4.00E-10	ABC transporter G family member 22	<i>C. canephora</i>
NODE_418_length_399_cov_1.13043:g368.t1	8.00E-13	Delta-1-pyrroline-5-carboxylate synthase	<i>C. canephora</i>
NODE_422_length_398_cov_1.72586:g371.t1	2.00E-09	dicarboxylate transporter 2.1%2C chloroplastic	<i>C. eugenioides</i>
NODE_432_length_396_cov_0.9279:g378.t1	2.00E-21	tropinone reductase homolog At5g06060-like	<i>C. arabica</i>
NODE_432_length_396_cov_0.9279:g378.t2	2.00E-21	tropinone reductase homolog At5g06060-like	<i>C. arabica</i>
NODE_434_length_396_cov_0.683386:g380.t1	2.00E-36	histidinol dehydrogenase%2C chloroplastic-like isoform X6	<i>C. arabica</i>
NODE_435_length_396_cov_1.15674:g381.t1	2.00E-17	endonuclease V-like isoform X2	<i>C. arabica</i>
NODE_452_length_392_cov_1.15556:g391.t1	7.00E-36	carbamoyl-phosphate synthase large chain%2C chloroplastic-like	<i>C. arabica</i>
NODE_458_length_391_cov_0.929936:g396.t1	6.00E-20	ATP synthase gamma chain%2C chloroplastic	<i>C. canephora</i>
NODE_470_length_390_cov_0.846645:g405.t1	3.00E-19	Putative Cystathionine beta-lyase	<i>C. canephora</i>
NODE_489_length_387_cov_1.55161:g422.t1	8.00E-21	uncharacterized protein LOC113703666	<i>C. arabica</i>
NODE_495_length_386_cov_1.18447:g428.t1	2.00E-34	50S ribosomal protein L3-2%2C chloroplastic	<i>C. eugenioides</i>
NODE_501_length_385_cov_0.769481:g433.t1	8.00E-27	3-methyl-2-oxobutanoate hydroxymethyltransferase 1%2C mitochondrial-like isoform X1	<i>C. arabica</i>
NODE_502_length_384_cov_0.957655:g434.t1	1.00E-29	phosphoenolpyruvate carboxykinase (ATP)-like	<i>C. arabica</i>
NODE_510_length_382_cov_0.914754:g439.t1	8.00E-08	ABC transporter B family member 25 isoform X1	<i>C. arabica</i>
NODE_512_length_382_cov_1.25902:g440.t1	2.00E-07	protein TRIGALACTOSYLDIACYLGLYCEROL 3%2C chloroplastic	<i>C. eugenioides</i>
NODE_514_length_381_cov_1.21382:g442.t1	1.00E-07	37 kDa inner envelope membrane protein%2C chloroplastic	<i>C. canephora</i>
NODE_519_length_380_cov_0.851485:g446.t1	2.00E-11	Putative Hypoxanthine-guanine phosphoribosyltransferase	<i>C. canephora</i>
NODE_523_length_380_cov_1.05281:g449.t1	6.00E-17	probable copper-transporting ATPase HMA5 isoform X2	<i>C. arabica</i>
NODE_525_length_380_cov_1.19472:g451.t1	3.00E-22	octanoyltransferase LIP2p%2C chloroplastic-like isoform X3	<i>C. arabica</i>
NODE_528_length_379_cov_1.03974:g453.t1	2.00E-14	Putative Gamma-glutamyltranspeptidase 1	<i>C. canephora</i>
NODE_529_length_379_cov_0.97351:g454.t1	2.00E-14	Acetylornithine aminotransferase%2C mitochondrial	<i>C. canephora</i>
NODE_538_length_377_cov_1.01667:g461.t1	3.00E-18	serine acetyltransferase 5-like	<i>C. arabica</i>
NODE_539_length_377_cov_0.983333:g462.t1	9.00E-18	ABC transporter B family member 13-like isoform X1	<i>C. eugenioides</i>
NODE_545_length_376_cov_1.46823:g467.t1	4.00E-18	DExH-box ATP-dependent RNA helicase DExH1 isoform X3	<i>C. eugenioides</i>
NODE_550_length_375_cov_0.738255:g469.t1	7.00E-14	Zinc-metallopeptidase%2C peroxisomal	<i>C. canephora</i>

NODE_557_length_373_cov_0.902027:g473.t1	2.00E-15	imidazole glycerol phosphate synthase hisHF%2C chloroplastic	<i>C. eugenioides</i>
NODE_564_length_371_cov_1.2551:g476.t1	7.00E-18	alpha-1%2C4 glucan phosphorylase L-2 isozyme%2C chloroplastic/amyloplastic-like	<i>C. arabica</i>
NODE_565_length_371_cov_1.0068:g477.t1	1.00E-07	ABC transporter I family member 17	<i>C. canephora</i>
NODE_567_length_371_cov_1.2449:g479.t1	5.00E-13	dephospho-CoA kinase-like	<i>C. eugenioides</i>
NODE_570_length_371_cov_0.819728:g481.t1	1.00E-07	Probable rhamnose biosynthetic enzyme 1	<i>C. canephora</i>
NODE_572_length_370_cov_1:g483.t1	8.00E-16	probable phosphoribosylformylglycinamide synthase%2C chloroplastic/mitochondrial	<i>C. arabica</i>
NODE_574_length_370_cov_1.00683:g485.t1	8.00E-09	secoisolariciresinol dehydrogenase-like	<i>C. eugenioides</i>
NODE_575_length_370_cov_1.24232:g486.t1	2.00E-08	uncharacterized protein LOC113770668	<i>C. eugenioides</i>
NODE_580_length_369_cov_1.00685:g490.t1	5.00E-25	Heat shock protein 83	<i>C. canephora</i>
NODE_583_length_369_cov_0.928082:g493.t1	2.00E-17	acetyl-CoA carboxylase beta subunit	<i>C. arabica</i>
NODE_584_length_368_cov_1.19244:g494.t1	2.00E-16	acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha%2C chloroplastic-like isoform X1	<i>C. arabica</i>
NODE_585_length_368_cov_0.721649:g495.t1	5.00E-14	50S ribosomal protein L5%2C chloroplastic	<i>C. canephora</i>
NODE_591_length_366_cov_1.27682:g500.t1	2.00E-10	GTP 3'%2C8-cyclase%2C mitochondrial-like isoform X2	<i>C. arabica</i>
NODE_599_length_365_cov_0.770833:g507.t1	3.00E-09	50S ribosomal protein L21%2C chloroplastic	<i>C. eugenioides</i>
NODE_609_length_364_cov_1.02091:g513.t1	2.00E-21	Putative protease Do-like 14	<i>C. canephora</i>
NODE_619_length_363_cov_1.02098:g519.t1	3.00E-16	aspartate carbamoyltransferase 1%2C chloroplastic	<i>C. eugenioides</i>
NODE_619_length_363_cov_1.02098:g519.t2	8.00E-17	aspartate carbamoyltransferase 1%2C chloroplastic	<i>C. eugenioides</i>
NODE_624_length_363_cov_1.48601:g523.t1	3.00E-38	Heat shock 70 kDa protein%2C mitochondrial	<i>C. canephora</i>
NODE_636_length_361_cov_1.03873:g529.t1	3.00E-19	allantoate deiminase 2-like isoform X1	<i>C. arabica</i>
NODE_638_length_361_cov_1.03169:g531.t1	2.00E-32	Fructose-1%2C6-bisphosphatase%2C cytosolic	<i>C. canephora</i>
NODE_642_length_361_cov_1.02465:g535.t1	1.00E-18	Putative tRNA (cytidine(34)-2'-O)-methyltransferase	<i>C. canephora</i>
NODE_643_length_361_cov_0.757042:g536.t1	3.00E-41	delta-aminolevulinic acid dehydratase%2C chloroplastic-like	<i>C. eugenioides</i>
NODE_652_length_359_cov_0.787234:g541.t1	3.00E-12	ABC transporter B family member 25	<i>C. canephora</i>
NODE_656_length_359_cov_1.30142:g545.t1	2.00E-16	LOW QUALITY PROTEIN: transketolase%2C chloroplastic-like	<i>C. arabica</i>
NODE_658_length_358_cov_1.04626:g547.t1	1.00E-14	L-idonate 5-dehydrogenase	<i>C. canephora</i>

NODE_664_length_358_cov_1.02847:g551.t1	5.00E-40	Succinate-semialdehyde dehydrogenase%2C mitochondrial	<i>C. canephora</i>
NODE_685_length_354_cov_1.18773:g564.t1	4.00E-16	uncharacterized protein LOC113778856	<i>C. eugenioides</i>
NODE_699_length_352_cov_0.792727:g573.t1	1.00E-09	Putative Epimerase family protein slr1223	<i>C. canephora</i>
NODE_722_length_350_cov_0.787546:g586.t1	1.00E-11	1%2C4-alpha-glucan-branching enzyme 2-2%2C chloroplastic/amyloplastic-like isoform X2	<i>C. eugenioides</i>
NODE_725_length_349_cov_1.15809:g589.t1	4.00E-10	RNA polymerase sigma factor sigF%2C chloroplastic	<i>C. eugenioides</i>
NODE_729_length_349_cov_1.08088:g591.t1	1.00E-08	uncharacterized protein LOC113783762	<i>C. eugenioides</i>
NODE_733_length_348_cov_1.0738:g593.t1	3.00E-14	probable copper-transporting ATPase HMA5	<i>C. eugenioides</i>
NODE_735_length_348_cov_1.07749:g594.t1	2.00E-21	Putative GTPase Der	<i>C. canephora</i>
NODE_753_length_346_cov_1.91078:g609.t1	2.00E-18	50S ribosomal protein L3%2C chloroplastic-like isoform X1	<i>C. arabica</i>
NODE_772_length_343_cov_1.09774:g626.t1	3.00E-09	ATP-dependent DNA helicase SRS2-like protein At4g25120 isoform X4	<i>C. arabica</i>
NODE_776_length_343_cov_1.10526:g628.t1	1.00E-27	aconitate hydratase%2C cytoplasmic	<i>C. arabica</i>
NODE_786_length_341_cov_0.700758:g637.t1	2.00E-24	LOW QUALITY PROTEIN: uncharacterized protein LOC113771516	<i>C. eugenioides</i>
NODE_787_length_341_cov_0.704545:g638.t1	3.00E-26	threonine dehydratase 1 biosynthetic%2C chloroplastic isoform X1	<i>C. eugenioides</i>
NODE_793_length_339_cov_1.12595:g642.t1	2.00E-14	ferredoxin-dependent glutamate synthase%2C chloroplastic-like	<i>C. arabica</i>
NODE_796_length_339_cov_1.12214:g644.t1	3.00E-37	lon protease homolog 2%2C peroxisomal isoform X2	<i>C. arabica</i>
NODE_805_length_338_cov_1.39464:g649.t1	8.00E-18	uncharacterized protein LOC113741456 isoform X3	<i>C. arabica</i>
NODE_818_length_337_cov_0.730769:g659.t1	7.00E-25	3-ketoacyl-CoA thiolase 2%2C peroxisomal	<i>C. arabica</i>
NODE_831_length_336_cov_1:g665.t1	5.00E-10	gamma carbonic anhydrase 1%2C mitochondrial-like	<i>C. arabica</i>
NODE_842_length_334_cov_0.856031:g673.t1	1.00E-19	Probable phosphoribosylformylglycinamide synthase%2C chloroplastic/mitochondrial	<i>C. canephora</i>
NODE_844_length_334_cov_1.14008:g675.t1	5.00E-08	Glutamyl-tRNA reductase 1%2C chloroplastic	<i>C. canephora</i>
NODE_845_length_334_cov_0.856031:g676.t1	5.00E-52	phosphomethylpyrimidine synthase%2C chloroplastic isoform X2	<i>C. eugenioides</i>
NODE_846_length_334_cov_1.14786:g677.t1	2.00E-12	Methylmalonate-semialdehyde dehydrogenase [acylating]%2C mitochondrial	<i>C. canephora</i>
NODE_847_length_334_cov_0.692607:g678.t1	6.00E-11	2-methylacyl-CoA dehydrogenase%2C mitochondrial-like	<i>C. arabica</i>

NODE_848_length_333_cov_0.867188:g679.t1	4.00E-23	bifunctional protein FOLD 1%2C mitochondrial-like	<i>C. arabica</i>
NODE_849_length_333_cov_0.726562:g680.t1	3.00E-08	xaa-Pro dipeptidase	<i>C. eugenioides</i>
NODE_850_length_333_cov_1.4375:g681.t1	3.00E-17	Ammonium transporter 3 member 1	<i>C. canephora</i>
NODE_869_length_330_cov_0.87747:g694.t1	1.00E-12	tRNA-dihydrouridine(20) synthase [NAD(P)]-like	<i>C. canephora</i>
NODE_894_length_327_cov_0.852:g711.t1	4.00E-15	ABC transporter F family member 5	<i>C. canephora</i>
NODE_900_length_326_cov_1.10442:g717.t1	6.00E-28	50S ribosomal protein L13%2C chloroplastic	<i>C. canephora</i>
NODE_903_length_325_cov_1.17742:g720.t1	1.00E-09	histidine kinase 2-like isoform X2	<i>C. arabica</i>
NODE_926_length_322_cov_0.926531:g734.t1	1.00E-36	Protein translocase subunit SECA2%2C chloroplastic	<i>C. canephora</i>
NODE_933_length_322_cov_1.13469:g740.t1	6.00E-16	ABC transporter I family member 17	<i>C. canephora</i>
NODE_992_length_314_cov_0.932489:g772.t1	4.00E-12	Ribonucleoside-diphosphate reductase large subunit	<i>C. canephora</i>
NODE_994_length_314_cov_0.932489:g774.t1	1.00E-10	Bifunctional purine biosynthesis protein PurH	<i>C. canephora</i>
NODE_995_length_314_cov_1.8692:g775.t1	4.00E-08	ABC transporter B family member 1	<i>C. canephora</i>
NODE_1017_length_312_cov_0.931915:g793.t1	5.00E-20	Soluble inorganic pyrophosphatase	<i>C. canephora</i>
NODE_1022_length_312_cov_2.49787:g795.t1	4.00E-12	ABC transporter A family member 2-like	<i>C. eugenioides</i>
NODE_1027_length_311_cov_0.940171:g798.t1	2.00E-47	phosphomethylpyrimidine synthase%2C chloroplastic isoform X3	<i>C. eugenioides</i>
NODE_1060_length_306_cov_1.27074:g820.t1	1.00E-19	S-formylglutathione hydrolase isoform X2	<i>C. eugenioides</i>
NODE_1068_length_305_cov_1.19737:g827.t1	2.00E-21	Probable ornithine aminotransferase	<i>C. canephora</i>
NODE_1105_length_301_cov_1.29464:g849.t1	1.00E-12	protein TRIGALACTOSYLDIACYLGLYCEROL 3%2C chloroplastic	<i>C. eugenioides</i>
NODE_1135_length_299_cov_1.14414:g867.t1	2.00E-36	Phosphoenolpyruvate carboxykinase [ATP]	<i>C. canephora</i>
NODE_1137_length_298_cov_0.972851:g869.t1	3.00E-11	pyridoxal kinase-like isoform X2	<i>C. arabica</i>
NODE_1163_length_296_cov_1.43836:g880.t1	9.00E-08	probable polyol transporter 4 isoform X2	<i>C. arabica</i>
NODE_1170_length_296_cov_0.694064:g887.t1	4.00E-07	uncharacterized protein LOC113722469	<i>C. arabica</i>
NODE_1179_length_295_cov_1.53211:g894.t1	3.00E-18	Putative Methionine aminotransferase	<i>C. canephora</i>
NODE_1193_length_293_cov_1.01852:g902.t1	5.00E-12	uncharacterized protein LOC113742933 isoform X2	<i>C. arabica</i>
NODE_1202_length_293_cov_1.2963:g905.t1	2.00E-09	folate synthesis bifunctional protein%2C mitochondrial-like isoform X2	<i>C. arabica</i>
NODE_1218_length_292_cov_1.69767:g917.t1	6.00E-27	histidinol dehydrogenase%2C chloroplastic-like isoform X4	<i>C. arabica</i>

NODE_1222_length_291_cov_1.37383:g920.t1	2.00E-14	8-amino-7-oxononanoate synthase-like isoform X1	<i>C. eugenioides</i>
NODE_1225_length_291_cov_1.02804:g921.t1	8.00E-14	70 kDa peptidyl-prolyl isomerase-like isoform X1	<i>C. arabica</i>
NODE_1233_length_290_cov_1.03286:g927.t1	3.00E-16	50S ribosomal protein L17	<i>C. canephora</i>
NODE_1239_length_290_cov_1.06103:g929.t1	2.00E-07	putative elongation factor TypA-like SVR3%2C chloroplastic isoform X1	<i>C. arabica</i>
NODE_1242_length_289_cov_0.90566:g931.t1	8.00E-12	long chain base biosynthesis protein 1-like isoform X1	<i>C. arabica</i>
NODE_1255_length_288_cov_0.630332:g936.t1	4.00E-19	3-ketoacyl-CoA thiolase 2%2C peroxisomal	<i>C. canephora</i>
NODE_1262_length_287_cov_1.04286:g940.t1	5.00E-14	ATP-dependent DNA helicase homolog RECG%2C chloroplastic isoform X2	<i>C. eugenioides</i>
NODE_1265_length_287_cov_1.2:g943.t1	2.00E-06	DUF21 domain-containing protein At1g55930%2C chloroplastic-like isoform X3	<i>C. arabica</i>
NODE_1268_length_287_cov_1.69048:g945.t1	3.00E-38	xylose isomerase-like	<i>C. arabica</i>
NODE_1269_length_287_cov_1.40476:g946.t1	2.00E-18	dihydropyrimidine dehydrogenase (NADP(+))%2C chloroplastic-like	<i>C. arabica</i>
NODE_1270_length_287_cov_0.890476:g947.t1	8.00E-23	threonylcarbamoyladenosine tRNA methylthiotransferase isoform X2	<i>C. eugenioides</i>
NODE_1293_length_285_cov_0.740385:g960.t1	2.00E-24	citrate synthase%2C glyoxysomal	<i>C. arabica</i>
NODE_1340_length_281_cov_1.44608:g988.t1	3.00E-07	RNA polymerase sigma factor sigD%2C chloroplastic-like	<i>C. arabica</i>
NODE_1347_length_280_cov_0.729064:g990.t1	9.00E-15	glutamine--tRNA ligase-like	<i>C. arabica</i>
NODE_1349_length_280_cov_1.20197:g991.t1	5.00E-21	superoxide dismutase [Mn]%2C mitochondrial	<i>C. eugenioides</i>
NODE_1361_length_280_cov_1.13793:g1000.t1	3.00E-24	isoleucine--tRNA ligase%2C chloroplastic/mitochondrial-like	<i>C. arabica</i>
NODE_1373_length_279_cov_0.727723:g1007.t1	4.00E-23	Adenylate kinase B	<i>C. canephora</i>
NODE_1385_length_278_cov_0.776119:g1013.t1	5.00E-25	ATP-dependent RNA helicase-like protein DB10	<i>C. eugenioides</i>
NODE_1393_length_278_cov_0.731343:g1016.t1	4.00E-25	NADPH:quinone oxidoreductase-like	<i>C. eugenioides</i>
NODE_1403_length_277_cov_0.725:g1023.t1	1.00E-60	Cysteine-rich receptor-like protein kinase 10	<i>C. canephora</i>
NODE_1406_length_277_cov_0.725:g1026.t1	4.00E-08	Putative ATP-dependent DNA helicase pcrA	<i>C. canephora</i>
NODE_1418_length_277_cov_1.47:g1032.t1	1.00E-17	anthranilate synthase alpha subunit 2%2C chloroplastic	<i>C. eugenioides</i>
NODE_1419_length_277_cov_0.81:g1033.t1	2.00E-29	chaperone protein ClpB4%2C mitochondrial isoform X1	<i>C. arabica</i>

NODE_1435_length_276_cov_0.738693:g1043.t1	3.00E-26	asparagine synthetase [glutamine-hydrolyzing] 2	<i>C. arabica</i>
NODE_1440_length_276_cov_1.48744:g1046.t1	1.00E-21	DNA gyrase subunit A%2C chloroplastic/mitochondrial	<i>C. eugenioides</i>
NODE_1447_length_276_cov_0.738693:g1050.t1	4.00E-48	carbamoyl-phosphate synthase large chain%2C chloroplastic	<i>C. eugenioides</i>
NODE_1451_length_276_cov_1.36181:g1052.t1	5.00E-30	Isoleucine--tRNA ligase	<i>C. canephora</i>
NODE_1459_length_275_cov_0.737374:g1057.t1	2.00E-18	uncharacterized protein LOC113763603	<i>C. eugenioides</i>
NODE_1463_length_275_cov_1.4596:g1059.t1	2.00E-13	alpha-glucan phosphorylase%2C H isozyme	<i>C. arabica</i>
NODE_1481_length_274_cov_0.751269:g1068.t1	2.00E-13	phosphoserine aminotransferase 2%2C chloroplastic-like	<i>C. eugenioides</i>
NODE_1490_length_273_cov_1.11224:g1071.t1	2.00E-17	Putative GTPase HflX	<i>C. canephora</i>
NODE_1509_length_273_cov_0.714286:g1084.t1	2.00E-08	NADH dehydrogenase [ubiquinone] flavoprotein 1%2C mitochondrial	<i>C. eugenioides</i>
NODE_1531_length_272_cov_0.723077:g1095.t1	5.00E-18	Ribosomal protein L4/L1 family	<i>C. canephora</i>
NODE_1538_length_271_cov_0.762887:g1098.t1	1.00E-21	chaperonin CPN60-like 2%2C mitochondrial	<i>C. eugenioides</i>
NODE_1539_length_271_cov_0.757732:g1099.t1	5.00E-26	ATPase WRNIP1	<i>C. eugenioides</i>
NODE_1554_length_270_cov_1.5285:g1109.t1	3.00E-11	fe-S cluster assembly factor HCF101%2C chloroplastic-like	<i>C. arabica</i>
NODE_1584_length_269_cov_0.765625:g1127.t1	2.00E-22	uncharacterized protein LOC113749318	<i>C. eugenioides</i>
NODE_1590_length_269_cov_1.14583:g1131.t1	2.00E-08	Putative Amino-acid acetyltransferase	<i>C. canephora</i>
NODE_1599_length_268_cov_0.774869:g1136.t1	7.00E-13	uncharacterized protein LOC113770237	<i>C. eugenioides</i>
NODE_1601_length_268_cov_0.769634:g1138.t1	2.00E-15	leucine--tRNA ligase%2C chloroplastic/mitochondrial	<i>C. eugenioides</i>
NODE_1610_length_268_cov_0.764398:g1142.t1	2.00E-11	ABC transporter G family member 10	<i>C. eugenioides</i>
NODE_1614_length_268_cov_0.764398:g1143.t1	1.00E-08	NADH dehydrogenase [ubiquinone] flavoprotein 2%2C mitochondrial	<i>C. eugenioides</i>
NODE_1615_length_268_cov_0.769634:g1144.t1	1.00E-42	nucleolar and coiled-body phosphoprotein 1-like	<i>C. eugenioides</i>
NODE_1622_length_268_cov_1.05236:g1147.t1	6.00E-34	Putative NADP-specific glutamate dehydrogenase	<i>C. canephora</i>
NODE_1639_length_267_cov_1.14737:g1154.t1	9.00E-26	pyruvate kinase%2C cytosolic isozyme-like	<i>C. eugenioides</i>
NODE_1646_length_266_cov_0.772487:g1159.t1	2.00E-09	protein lojap-related%2C mitochondrial	<i>C. eugenioides</i>
NODE_1680_length_265_cov_0.781915:g1178.t1	4.00E-18	Alpha-glucan phosphorylase%2C H isozyme	<i>C. canephora</i>
NODE_1700_length_263_cov_0.784946:g1187.t1	4.00E-18	uncharacterized protein LOC113768461 isoform X2	<i>C. eugenioides</i>

NODE_1709_length_263_cov_1.17742:g1192.t1	5.00E-18	xylose isomerase-like	<i>C. arabica</i>
NODE_1721_length_262_cov_0.789189:g1198.t1	4.00E-12	ABC transporter I family member 10	<i>C. eugenioides</i>
NODE_1722_length_262_cov_1.16757:g1199.t1	3.00E-09	succinate dehydrogenase [ubiquinone] flavoprotein subunit%2C mitochondrial	<i>C. eugenioides</i>
NODE_1731_length_262_cov_0.794595:g1204.t1	2.00E-34	phosphoribosylamine--glycine ligase	<i>C. eugenioides</i>
NODE_1738_length_261_cov_1.20109:g1208.t1	1.00E-22	Probable 1-deoxy-D-xylulose-5-phosphate synthase 2%2C chloroplastic	<i>C. canephora</i>
NODE_1742_length_261_cov_0.793478:g1209.t1	4.00E-15	Geranylgeranyl pyrophosphate synthase%2C chloroplastic	<i>C. canephora</i>
NODE_1743_length_261_cov_1.04891:g1210.t1	1.00E-18	50S ribosomal protein L5%2C chloroplastic	<i>C. eugenioides</i>
NODE_1784_length_260_cov_1.03279:g1229.t1	7.00E-19	transketolase%2C chloroplastic	<i>C. arabica</i>
NODE_1794_length_259_cov_0.813187:g1235.t1	3.00E-35	Putative Lysyl-tRNA synthetase%2C class II	<i>C. canephora</i>
NODE_1798_length_259_cov_1.23626:g1238.t1	5.00E-10	uncharacterized protein LOC113776159	<i>C. eugenioides</i>
NODE_1802_length_259_cov_1.2033:g1240.t1	1.00E-15	putative deoxyribonuclease TATDN1	<i>C. eugenioides</i>
NODE_1825_length_258_cov_0.78453:g1252.t1	2.00E-07	succinate dehydrogenase [ubiquinone] flavoprotein subunit%2C mitochondrial	<i>C. eugenioides</i>
NODE_1832_length_257_cov_1.21667:g1256.t1	2.00E-27	DNA repair protein recA homolog 3%2C mitochondrial	<i>C. eugenioides</i>
NODE_1856_length_256_cov_1.65922:g1267.t1	1.00E-13	thioredoxin reductase NTRC	<i>C. eugenioides</i>
NODE_1863_length_256_cov_0.899441:g1272.t1	6.00E-15	dihydroorotate dehydrogenase (quinone)%2C mitochondrial	<i>C. eugenioides</i>
NODE_1899_length_254_cov_0.824859:g1290.t1	9.00E-11	3-oxoacyl-[acyl-carrier-protein] synthase%2C mitochondrial isoform X1	<i>C. arabica</i>
NODE_1905_length_254_cov_1.24859:g1292.t1	6.00E-29	putative elongation factor TypA-like SVR3%2C chloroplastic isoform X1	<i>C. arabica</i>
NODE_1917_length_254_cov_1.0791:g1296.t1	3.00E-08	RNA polymerase sigma factor sigF%2C chloroplastic isoform X1	<i>C. arabica</i>
NODE_1922_length_253_cov_1.59659:g1299.t1	2.00E-19	Putative Oligopeptidase A	<i>C. canephora</i>
NODE_1937_length_253_cov_0.835227:g1306.t1	7.00E-11	NADPH-dependent thioredoxin reductase 3	<i>C. canephora</i>
NODE_1975_length_251_cov_0.833333:g1324.t1	4.00E-07	uncharacterized protein LOC113697941 isoform X1	<i>C. arabica</i>
NODE_1989_length_251_cov_0.988506:g1331.t1	6.00E-13	DNA gyrase subunit B%2C chloroplastic/mitochondrial-like	<i>C. arabica</i>
NODE_2007_length_250_cov_1.7052:g1338.t1	3.00E-40	Elongation factor Tu%2C chloroplastic	<i>C. canephora</i>

NODE_2011_length_250_cov_0.849711:g1340.t1	7.00E-10	4-hydroxy-tetrahydrodipicolinate synthase%2C chloroplastic-like	<i>C. eugenioides</i>
NODE_2034_length_249_cov_1.27326:g1353.t1	6.00E-11	3-phosphoshikimate 1-carboxyvinyltransferase 2	<i>C. eugenioides</i>
NODE_2042_length_249_cov_0.854651:g1358.t1	2.00E-17	Putative Hypoxanthine-guanine phosphoribosyltransferase	<i>C. canephora</i>
NODE_2064_length_248_cov_1.2924:g1368.t1	1.00E-08	omega-amidase%2C chloroplastic-like isoform X1	<i>C. arabica</i>
NODE_2067_length_248_cov_0.853801:g1371.t1	1.00E-25	glycine cleavage system H protein 2%2C mitochondrial	<i>C. eugenioides</i>
NODE_2088_length_247_cov_0.858824:g1382.t1	5.00E-10	glycogen phosphorylase 1-like isoform X2	<i>C. eugenioides</i>
NODE_2093_length_247_cov_1.29412:g1383.t1	3.00E-14	uncharacterized protein LOC113739699 isoform X2	<i>C. arabica</i>
NODE_2114_length_246_cov_0.869822:g1393.t1	9.00E-09	Putative copper-transporting ATPase HMA5	<i>C. canephora</i>
NODE_2119_length_246_cov_1.68639:g1396.t1	8.00E-08	Putative uncharacterized protein	<i>C. canephora</i>
NODE_2131_length_245_cov_0.880952:g1404.t1	2.00E-18	Homocysteine S-methyltransferase 2	<i>C. canephora</i>
NODE_2146_length_245_cov_0.869048:g1411.t1	1.00E-07	nitrogen regulatory protein P-II homolog	<i>C. arabica</i>
NODE_2172_length_244_cov_1.32335:g1422.t1	9.00E-19	Lipoyl synthase%2C chloroplastic	<i>C. canephora</i>
NODE_2181_length_243_cov_0.891566:g1426.t1	6.00E-12	Putative Ribokinase	<i>C. canephora</i>
NODE_2184_length_243_cov_0.873494:g1427.t1	3.00E-09	ATP-dependent DNA helicase SRS2-like protein At4g25120 isoform X1	<i>C. arabica</i>
NODE_2186_length_243_cov_0.891566:g1428.t1	3.00E-22	aspartate aminotransferase%2C cytoplasmic-like	<i>C. arabica</i>
NODE_2192_length_243_cov_1.78313:g1431.t1	4.00E-27	adenyllyl-sulfate kinase 3 isoform X2	<i>C. arabica</i>
NODE_2207_length_242_cov_0.890909:g1436.t1	3.00E-31	Glyceraldehyde-3-phosphate dehydrogenase%2C cytosolic	<i>C. canephora</i>
NODE_2208_length_242_cov_0.884848:g1437.t1	3.00E-28	alcohol dehydrogenase class-3	<i>C. eugenioides</i>
NODE_2220_length_242_cov_1.21212:g1439.t1	4.00E-09	Putative aconitate hydratase%2C cytoplasmic	<i>C. canephora</i>
NODE_2221_length_242_cov_1.33939:g1440.t1	9.00E-13	ethanolamine-phosphate cytidyltransferase	<i>C. eugenioides</i>
NODE_2257_length_240_cov_1.25767:g1455.t1	3.00E-30	biotin synthase%2C mitochondrial-like	<i>C. eugenioides</i>
NODE_2270_length_240_cov_0.90184:g1463.t1	4.00E-09	zinc protease PQQL-like isoform X1	<i>C. arabica</i>
NODE_2281_length_239_cov_1.34568:g1467.t1	8.00E-21	Putative tRNA-dihydrouridine synthase A	<i>C. canephora</i>
NODE_2314_length_238_cov_0.913043:g1479.t1	4.00E-13	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase%2C chloroplastic	<i>C. canephora</i>
NODE_2324_length_238_cov_0.658385:g1483.t1	2.00E-11	alpha-1%2C4 glucan phosphorylase L-1 isozyme%2C chloroplastic/amyloplastic isoform X2	<i>C. arabica</i>

NODE_2336_length_238_cov_0.888199:g1490.t1	6.00E-09	probable low-specificity L-threonine aldolase 1 isoform X1	<i>C. arabica</i>
NODE_2407_length_235_cov_1.39241:g1520.t1	2.00E-08	Acetolactate synthase 2%2C chloroplastic	<i>C. canephora</i>
NODE_2410_length_235_cov_1.86076:g1522.t1	4.00E-09	Elongation factor Ts	<i>C. canephora</i>
NODE_2437_length_234_cov_1.05732:g1533.t1	1.00E-17	glycogen phosphorylase 1-like isoform X3	<i>C. eugenioides</i>
NODE_2456_length_233_cov_1.36538:g1543.t1	4.00E-08	ABC transporter G family member 8	<i>C. canephora</i>
NODE_2471_length_232_cov_0.948387:g1548.t1	3.00E-19	glycerol kinase-like	<i>C. arabica</i>
NODE_2575_length_229_cov_0.973684:g1598.t1	1.00E-08	aldehyde dehydrogenase family 2 member C4-like	<i>C. eugenioides</i>
NODE_2576_length_229_cov_1.66447:g1599.t1	2.00E-21	monofunctional riboflavin biosynthesis protein RIBA 3%2C chloroplastic	<i>C. eugenioides</i>

Apêndice 2: Tabelas com os códigos de identificação (ID) de sequências do NCBI usadas nas análises de filogenia.

Sequence ID	Genome
NODE 34	Clone 101-2P
NW_020861601.1	<i>C. eugenoides</i>
Cc02_g24060	<i>C. canephora</i>
chr_B_sg_C_(Scaffold_34;HRSCAF=35)	<i>C. arabica</i> cv. Typica
NC_039900.1	<i>C. arabica</i> cv. Caturra

Sequence ID	Genome
NODE 138	Clone 101-2P
NC_040040.1	<i>C. eugenoides</i>
Cc06_g09870	<i>C. canephora</i>
chr_F_sg_C_(Scaffold_17;HRSCAF=18)	<i>C. arabica</i> cv. Typica
NC_039908.1	<i>C. arabica</i> cv. Caturra

Sequence ID	Genome
NODE 82	Clone 101-2P
NC_040035.1	<i>C. eugenoides</i>
Cc01_g21120	<i>C. canephora</i>
chr_A_sg_E_(Scaffold_19;HRSCAF=20)	<i>C. arabica</i> cv. Typica
NC_039899.1	<i>C. arabica</i> cv. Caturra

Sequence ID	Genome
NODE 162	Clone 101-2P
NC_040041.1	<i>C. eugenoides</i>
Cc08_g04850	<i>C. canephora</i>
chr_E_sg_C_(Scaffold_3;HRSCAF=4)	<i>C. arabica</i> cv. Typica
NC_008535.1	<i>C. arabica</i> cv. Caturra

Sequence ID	Genome
NODE 98	Clone 101-2P
NC_040045.1	<i>C. eugenoides</i>
Cc04_g00960	<i>C. canephora</i>
chr_G_sg_E_(Scaffold_25;HRSCAF=26)	<i>C. arabica</i> cv. Typica
NC_039904.1	<i>C. arabica</i> cv. Caturra

Sequence ID	Genome
RGA2	clone 101-2P
NC_039899.1	<i>C. arabica</i> cv. Caturra
chr_A_sg_E_(Scaffold_19;HRSCAF=20)	<i>C. arabica</i> cv. Typica
Cc00_g32920	<i>C. canephora</i>
NC_040035.1	<i>C. eugenioides</i>

Sequence ID	Genome
Vinorine synthase-Like NODE1g2	Clone 101-2P
NC_039899.1	<i>C. arabica</i> cv. Caturra
chr_A_sg_E_(Scaffold_19;HRSCAF=20)	<i>C. arabica</i> cv. Typica
Cc01_g01700	<i>C. canephora</i>
NC_040035.1	<i>C. eugenioides</i>

Sequence ID	Genome
RGA3	clone 101-2P
NC_039899.1	<i>C. arabica</i> cv. Caturra
chr_A_sg_E_(Scaffold_19;HRSCAF=20)	<i>C. arabica</i> cv. Typica
Cc00_g32920	<i>C. canephora</i>
NC_040045.1	<i>C. eugenioides</i>

Sequence ID	Genome
Protein Uncharacterized NODE1g4	Clone 101-2P
NC_039899.1	<i>C. arabica</i> cv. Caturra
chr_B_sg_E_(Scaffold_35;HRSCAF=36)	<i>C. arabica</i> cv. Typica
Cc03_g08290	<i>C. canephora</i>
NC_040045.1	<i>C. eugenioides</i>

Sequence ID	Genome
RGA4	clone 101-2P
NC_039899.1	<i>C. arabica</i> cv. Caturra
chr_A_sg_E_(Scaffold_19;HRSCAF=20)	<i>C. arabica</i> cv. Typica
Cc00_g30480	<i>C. canephora</i>
NC_040035.1	<i>C. eugenioides</i>