

**KÁTIA NOGUEIRA PESTANA**

**MAPA GENÉTICO INTEGRADO E SELEÇÃO GENÔMICA AMPLA  
VISANDO RESISTÊNCIA DE *Coffea arabica* L. À FERRUGEM**

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*.

VIÇOSA  
MINAS GERAIS – BRASIL  
2013

Ficha catalográfica preparada pela Seção de Catalogação e  
Classificação da Biblioteca Central da UFV

T Pestana, Kátia Nogueira, 1984-  
P476m Mapa genético integrado e seleção genômica ampla  
2013 visando resistência de *Coffea arabica* L. à ferrugem / Kátia  
Nogueira Pestana. – Viçosa, MG, 2013.  
xviii, 109 f. : il. ; 29 cm.

Orientador: Ney Sussumu Sakiyama.  
Tese (doutorado) - Universidade Federal de Viçosa.  
Inclui bibliografia

1. Café – Melhoramento genético. 2. Genoma. 3. *Coffea arabica*. 4. Ferrugem-do-cafeeiro. I. Universidade Federal de Viçosa. Departamento de Genética. Programa de Pós-Graduação em Genética e Melhoramento. II. Título.

CDD 22. ed. 633.733

KÁTIA NOGUEIRA PESTANA

MAPA GENÉTICO INTEGRADO E SELEÇÃO GENÔMICA AMPLA VISANDO  
RESISTÊNCIA DE *Coffea arabica* L. À FERRUGEM

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: 4 de dezembro de 2013.

---

Cosme Damião Cruz  
(Coorientador)

---

Eveline Teixeira Caixeta  
(Coorientadora)

---

Eunize Maciel Zambolim

---

Antonio Carlos Baião de Oliveira

---

Ney Sussumu Sakiyama  
(Orientador)

*“Deus nos fez perfeitos e não escolhe os capacitados, capacita os escolhidos.”*

*Albert Einstein*

A Deus.

Aos meus pais Antônio e Marinalva.

Às minhas irmãs Karla, Renata, Camila e Rafaela.

Ao meu noivo Ubiraci.

Aos meus amigos e familiares.

Aos profissionais que contribuíram para a minha formação profissional.

Dedico

## AGRADECIMENTOS

A Deus, pelo dom da vida e por sempre iluminar meu caminho, dando-me saúde e força para enfrentar os desafios e tornar possível este sonho.

À Universidade Federal de Viçosa e ao Programa de Pós-Graduação em Genética e Melhoramento, pela oportunidade de realização do Curso.

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

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Capes, pela concessão da bolsa.

Aos meus pais Antônio e Marinalva, pelo amor, pela educação, confiança, pelo incentivo e por sempre acreditarem em mim.

Às minhas irmãs Karla, Renata, Camila e Rafaela, pelo incentivo, carinho, pela compreensão e amizade.

Ao meu noivo Ubiraci, pelo carinho, pela atenção, pelo incentivo e pela paciência durante esse tempo.

Ao Professor Ney Sussumu Sakiyama, pela orientação, pelos valiosos ensinamentos, pelo incentivo, pelos aconselhamentos e pela confiança depositada em mim.

À Professora Eveline Teixeira Caixeta, pela orientação, amizade, confiança e incentivo, pelos sábios ensinamentos responsáveis pelo meu crescimento pessoal e profissional, pelos aconselhamentos e por estar sempre disponível e atenciosa para me atender, mesmo com várias outras obrigações.

À Dra. Eunize Maciel Zambolim, pela amizade, confiança depositada em mim, pelo incentivo, pelos ensinamentos e por sempre estar disposta a me ajudar, tanto no lado pessoal quanto no lado profissional.

Ao Professor Cosme, pela orientação, paciência, por sempre estar disponível a me ajudar, pelas valiosas sugestões e ensinamentos e pelo exemplo de profissional a ser seguido.

Ao Dr. Antônio Carlos Baião de Oliveira, pelas valiosas sugestões e pelos aconselhamentos para concretização deste trabalho.

Ao Professor Moysés, pelas sugestões, pelos aconselhamentos e pelas contribuições nas análises estatísticas.

Ao Professor Laércio Zambolim, pela amizade, pelos aconselhamentos e pelas sugestões.

Aos professores do Programa de Pós-Graduação em Genética e Melhoramento, em especial àqueles que contribuíram para a minha formação acadêmica.

Aos meus amigos do Laboratório Biocafé (Dênia, Dani, Rej, Pat, Karla, Alexandre, Brenda, Emilly, Tiago, Samila, Letícia 1, Letícia 2, Gregóry, Luis Felipe, Samuel, Tes, Flávio, Fernando, Bruna, Rose, Juan, Cinthia, Rodrigo, Geleta e Bikila), pela amizade, pelo incentivo, pela descontração nas horas vagas e pelo apoio para a realização deste trabalho.

À Dênia e a Dani, pela amizade e pelo apoio na realização desta pesquisa.

Aos pesquisadores da EMBRAPA Mandioca e Fruticultura Tropical, em especial a Dra. Cláudia, Dr. Edson, Dr. Sebastião e Dr. Onildo pela amizade, pelos aconselhamentos e incentivos que contribuíram para meu crescimento profissional.

Ao Caio, pela amizade e pela força para a realização deste trabalho

.

Ao Rafael Tomaz, pela amizade e pelo apoio na execução deste trabalho.

Aos Senhores Mário, Geraldo e Gerson, pelo apoio nos trabalhos de campo.

Aos meus amigos que aqui conquistei, em especial a Renatinha, Ana Paula, Telma, Thalita, Manu e Chris, pela amizade e pelo companheirismo.

A todos que, de alguma forma, contribuíram para o meu crescimento profissional e para a concretização deste trabalho.

Muito obrigada!

## **BIOGRAFIA**

KÁTIA NOGUEIRA PESTANA, filha de Antônio Nunes Pestana e Marinalva Nogueira Pestana, nasceu em Cruz das Almas, BA, no dia 21 de novembro de 1984.

Em 2003, ingressou no Curso de Engenharia Agrônômica da Universidade Federal da Bahia (UFBA), graduando-se pela Universidade Federal do Recôncavo da Bahia (UFRB) em 2008.

No período de julho de 2004 a julho de 2008, foi bolsista de Iniciação Científica na Embrapa Mandioca e Fruticultura Tropical, desenvolvendo atividades de pesquisa na área de Melhoramento Genético, com a utilização de marcadores moleculares RAPD e SSR.

Em agosto de 2008, iniciou o Programa de Pós-Graduação, em nível de Mestrado, em Fitotecnia da Universidade Federal de Viçosa (UFV), concluindo-o em julho de 2010.

Em agosto de 2010, ingressou no Programa de Pós-Graduação, em nível de Doutorado, em Genética e Melhoramento da UFV, submetendo-se à defesa da Tese em 4 de dezembro de 2013.



2.7. Influência dos marcadores moleculares identificados na severidade da doença.....	18
3. RESULTADOS .....	19
3.1. Avaliação da doença.....	19
3.2. Mapa de ligação.....	21
3.3. Identificação de QTLs .....	23
4. DISCUSSÃO .....	27
4.1. Avaliação da ferrugem .....	27
4.2. Mapa de ligação.....	27
4.3. Identificação de QTLs .....	30
5. CONCLUSÃO .....	34
6. REFERÊNCIAS .....	35
CAPÍTULO 2 MAPA DE LIGAÇÃO INTEGRADO PARA <i>Coffea arabica</i> .....	43
RESUMO .....	44
1. INTRODUÇÃO .....	46
2. MATERIAL E MÉTODOS .....	50
2.1. Material vegetal das populações de mapeamento para o mapa integrado.....	50
2.2. Genotipagem das populações .....	50
2.2.1. Extração de DNA.....	50
2.2.2. Marcadores SSR .....	51
2.2.3. Marcadores AFLP .....	54
2.2.4. Marcadores RAPD.....	54
2.2.5. Marcador <i>primer</i> -específico .....	54
2.3. Construção de mapas parciais .....	55
2.4. Integração de mapas .....	55
2.5. Análise fenotípica para identificação de QTLs .....	56
3. RESULTADOS .....	57
3.1. Genotipagem das populações .....	57
3.2. Construção de mapas parciais .....	57
3.3. Integração de mapas .....	58
3.4. Identificação de QTLs .....	59
4. DISCUSSÃO .....	68
4.1. Genotipagem das populações .....	68
4.2. Construção de mapas parciais .....	68

4.3. Integração de mapas .....	70
4.4. Identificação de QTLs .....	72
5. CONCLUSÃO .....	73
6. REFERÊNCIAS .....	74
CAPITULO 3 SELEÇÃO GENÔMICA AMPLA EM <i>Coffea arabica</i> .....	80
RESUMO .....	81
1. INTRODUÇÃO .....	83
2. MATERIAL E MÉTODOS .....	86
2.1. Material vegetal .....	86
2.2. Extração de DNA .....	86
2.3. Genotipagem da população .....	86
2.3.1. Marcadores SSR .....	87
2.3.2. Marcadores AFLP.....	87
2.3.3. Marcadores RAPD .....	88
2.3.4. Marcador primer-específico.....	88
2.4. Codificação dos marcadores.....	89
2.5. Seleção genômica ampla ( <i>Genome Wide Selection-GWS</i> ) – Lasso bayesiano (BLASSO).....	89
2.6. Análise fenotípica e mapeamento de QTLs associados a <i>H. vastarix</i> .....	90
3. RESULTADOS E DISCUSSÃO .....	92
3.1. Genotipagem dos marcadores.....	92
3.2. Capacidade preditiva da seleção genômica ampla .....	93
4. CONCLUSÃO .....	104
5. REFERÊNCIAS .....	105

## LISTA DE TABELAS

CAPÍTULO 1	MAPEAMENTO DE GENES E QTLs PARA A RESISTÊNCIA DO CAFEEIRO À FERRUGEM .....	6
Tabela 1 -	Proporção esperada e observada para dois e três genes responsáveis pela resistência, nos progenitores F1, população F2, retrocruzamento resistente (RCr) e retrocruzamento suscetível (RCs) .....	21
Tabela 2 -	Resumo do mapa de ligação para a população F2 de <i>C. arabica</i> .....	22
Tabela 3 -	Marcadores moleculares associados a QTLs para resistência do cafeeiro à ferrugem, identificados por intervalo simples por meio da análise de regressão ( $LOD > 3$ ) .....	25
CAPÍTULO 2	MAPA DE LIGAÇÃO INTEGRADO PARA <i>Coffea arabica</i> .....	43
Tabela 1 -	Descrição dos primers SSR utilizados para genotipagem das populações 1 e 2 de mapeamento. ....	52
Tabela 2 -	Resumo dos mapas parciais de ligação e mapa integrado. ....	60
Tabela 3 -	Resumo dos grupos de ligação dos mapas parciais que foram utilizados para integração e apenas grupo de ligação integrados .....	61
CAPÍTULO 3	SELEÇÃO GENÔMICA AMPLA EM <i>Coffea arabica</i> .....	80
Tabela 1 -	Capacidade preditiva da seleção genômica ampla utilizando os mesmos indivíduos para população de treinamento e validação para estimação dos valores genômicos dos indivíduos da população de seleção, para três isolados de <i>H. vastatrix</i> em <i>C. arabica</i> . ....	93

Tabela 2 - Capacidade preditiva da seleção genômica ampla com indivíduos selecionados de forma aleatória, para constituir a população de treinamento e validação para estimação dos valores genéticos genômicos dos indivíduos da população de seleção, para três isolados de <i>H. vastatrix</i> em <i>C. arabica</i> . .....	95
Tabela 3 - Capacidade preditiva da seleção genômica ampla com indivíduos selecionados de forma aleatória, para constituir a população de treinamento e validação para estimação dos valores genético-genômicos dos indivíduos da população de seleção, para três isolados de <i>H. vastatrix</i> em <i>C. arabica</i> . Selecionando marcadores de maior efeito. ....	96
Tabela 4 - Estimativa do valor genômico do indivíduo para a raça I de <i>H. vastatrix</i> , considerando-se apenas os marcadores de maiores efeitos. Dividindo a população F2 em dois grupos, amostrados aleatoriamente, sendo o primeiro grupo formado por 200 indivíduos para população de treinamento e o segundo por 45 indivíduos para população de validação. Ind = indivíduo da população F2; F = fenotípico; e G = genotípico. ....	98
Tabela 5 - Estimativa do valor genômico do indivíduo para a raça II de <i>H. vastatrix</i> , considerando-se apenas os marcadores de maiores efeitos. Dividindo a população F2 em dois grupos, amostrados aleatoriamente, sendo o primeiro grupo formado por 200 indivíduos para população de treinamento e o segundo por 45 indivíduos para população de validação. Ind = indivíduo da população F2; F = fenotípico; e G = genotípico. ....	99
Tabela 6 - Estimativa do valor genômico do indivíduo para o patótipo 001 de <i>H. vastatrix</i> , considerando-se apenas os marcadores de maiores efeitos. Dividindo a população F2 em dois grupos, amostrados aleatoriamente, sendo o primeiro grupo formado por 200 indivíduos para população de treinamento e o segundo grupo por 45 indivíduos para população de validação. Ind = indivíduo da população F2; F = fenotípico; e G = genotípico.....	100

Tabela 7 - Marcadores moleculares identificados associados às raças I e II e ao patótipo 001 de *H. vastatrix* em mapa genético de ligação e pelo método de seleção genômica ampla/BLASSO, considerando-se todos os indivíduos para população de treinamento e validação..... 102

## LISTA DE FIGURAS

CAPÍTULO 1	MAPEAMENTO DE GENES E QTLS PARA A RESISTÊNCIA DO CAFEEIRO À FERRUGEM .....	6
Figura 1 -	Frequência de cafeeiros F2– seguiu-se a escala de Tamayo et al. (1995). Indivíduos resistentes (1 a 3), suscetíveis (4 a 6). 1- Ausência de sintomas; 2- lesões cloróticas pequenas; 3- lesões cloróticas grandes sem esporulação; 4- lesões cloróticas grandes com poucos uredósporos, ocupando menos de 25% da área da lesão; 5- lesões com esporulação ocupando de 25 a 50% da área; e 6- lesões com esporulação ocupando mais de 50% da área com uredósporos.....	20
Figura 2 -	Mapa de ligação de <i>C. arabica</i> contendo 111 marcadores. À direita da barra estão os nomes dos marcadores moleculares e à esquerda, as distâncias em cM. As barras ao lado do nome dos marcadores indicam a posição dos QTLs para as raças I e II e o patótipo 001. ....	23
CAPÍTULO 2	MAPA DE LIGAÇÃO INTEGRADO PARA <i>Coffea arabica</i> .....	43
Figura 1 -	Grupo de ligação (GL) 1 do mapa integrado, proveniente da integração dos GLs 1 e 8 do mapa 1 (população 1) e GL 1 do mapa 2 (população 2). Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais.....	62
Figura 2 -	Grupo de ligação (GL) 2 do mapa integrado, proveniente da integração do GL 2 do mapa 1 (população 1) e GL 4 do mapa 2 (população 2). Linhas tracejadas indicam a posição dos	

	marcadores âncoras nos mapas parciais. A barra vermelha ao lado do GL 2 indica a posição do QTL 1 identificado associado à resistência do cafeeiro à raça II de <i>H. vastatrix</i> . ....	63
Figura 3 - Grupo de ligação (GLs) 3 do mapa integrado, proveniente da integração do GL 3 do mapa 1 (população 1) e GL 3 do mapa 2 (população 2). Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais.....		64
Figura 4 - Grupo de ligação (GL) 4 do mapa integrado, proveniente do GL 6 do mapa 1 (população 1) e GL 2 do mapa 2 (população 2). Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais.....		65
Figura 5 - Grupo de ligação (GL) 5 do mapa integrado, resultante da integração do GL 7 do mapa 1 e GL 6 do mapa 2. Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais.....		65
Figura 6 - Grupo de ligação (GL) 6 do mapa integrado, proveniente da integração do GL 11 do mapa 1 e do GL 8 do mapa 2. Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais. ....		66
Figura 7 - Grupo de ligação (GL) 7 do mapa integrado, resultante da integração do GL 12 do mapa 1 e GL 10 do mapa 2. Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais.....		66
Figura 8 - Demais grupos de ligação dos mapas parciais não resultantes do processo de integração GL 8 a 16 correspondem a GL não integrados. A barra vermelha ao lado dos GLs 2 e 11 indica a posição dos QTLs 1 e 2 identificados para a raça II de <i>H. vastatrix</i> , respectivamente. ....		67

## RESUMO

PESTANA, Kátia Nogueira, D.Sc., Universidade Federal de Viçosa, dezembro de 2013. **Mapa genético integrado e seleção genômica ampla visando resistência de *Coffea arabica* L. à ferrugem.** Orientadores: Ney Sussumu Sakiyama. Coorientadores: Eveline Teixeira Caixeta e Cosme Damião Cruz.

A obtenção de cultivares de cafeeiros resistentes e produtivos, assim como de informações detalhadas sobre o genoma dessa espécie, demanda uso de ferramentas de análise genética refinada. Por isso, a identificação de marcadores moleculares associados a genes de resistência à ferrugem em mapa genético de ligação tem sido estratégia proposta para incorporar a Seleção Assistida por Marcadores no melhoramento de *Coffea arabica*. Neste trabalho, foi caracterizada a herança da resistência do Híbrido de Timor UFV 443-03 às raças I e II e ao patótipo 001 de *Hemileia vastatrix*, assim como localizados em mapa genético de ligação parcial do cafeeiro os genes/QTLs identificados. O mapa genético construído foi integrado em outro mapa parcial pré-existente visando obter um mapa saturado, de referência para a espécie. Nesse mapa, composto por 191 marcadores, cobrindo 1.421,32 cM do genoma, foram identificados dois QTLs associados à raça II de *H. vastatrix*. Esses marcadores moleculares associados aos QTLs identificados podem ser utilizados para seleção de cafeeiros resistentes à ferrugem. Além dessa estratégia de melhoramento molecular, foi também proposta a seleção genômica com o objetivo de prever a performance fenotípica dos indivíduos e identificar marcadores moleculares associados à característica de resistência à ferrugem. Neste estudo foram utilizados 245 indivíduos F<sub>2</sub>, resultantes do cruzamento da cv suscetível Catuaí

Amarelo IAC 64 (UFV 2148-57) e da fonte de resistência Híbrido de Timor UFV 443-03, genotipados com 137 marcadores moleculares (74 AFLP, 58 SSR, 4 RAPD e 1 *primer*-específico). As estimativas dos parâmetros genéticos foram obtidas via BLASSO. A capacidade preditiva (CP) foi medida utilizando todos os indivíduos da população como população de treinamento e validação. Além disso, a população foi dividida aleatoriamente em 200 indivíduos para treinamento e 45 indivíduos para validação. No primeiro caso, a maior CP preditiva foi para o patótipo 001 (0,78) e, no segundo caso, para a raça I (0,49). Quando se utilizaram apenas os marcadores de maiores efeitos para cada isolado, a maior CP na população de validação foi para o patótipo 001 (0,49). Ainda foi possível observar que as herdabilidades estimadas por meio dos marcadores moleculares para a resistência aos três isolados foram inferiores às obtidas a partir de dados fenotípicos. Isso demonstra maior predição dos valores genômicos dos indivíduos na seleção genômica em comparação com a fenotípica. A maioria dos marcadores identificados associados aos QTLs no mapa genético de ligação foi também identificada pela GWS, confirmando a acurácia da metodologia utilizada. Além disso, a GWS permitiu predizer, com sucesso, os valores genômicos dos indivíduos.

## ABSTRACT

PESTANA, Kátia Nogueira, D. Sc., Universidade Federal de Viçosa, December, 2013. **Genetic integrated map and genome wide selection for resistance of the rust *Coffea arabica* L.** Adviser: Ney Sussumu Sakiyama. Co-Advisers: Eveline Teixeira Caixeta and Cosme Damião Cruz.

The obtaining of coffee cultivars resistant and productive, as well as detailed information about the genome of this species demand use of refined genetic analysis tools. Therefore, the identification of molecular markers linked to rust resistance genes in the genetic linkage map has been proposed a strategy to incorporate marker-assisted selection in the improvement of *Coffea arabica*. In this study, was characterized the resistance inheritance of the Híbrido de Timor UFV 443-03 to races I, II and pathotype 001 of *Hemileia vastatrix* as well as located in the partial genetic linkage map of coffee identified genes/QTLs. The integrated genetic map was constructed into other pre-existing partial map in order to obtain a saturated map reference for the species. In this map, composed of 191 markers, covering 1421.32 cM genome, QTL associated with race II were identified in *H. vastatrix*. These molecular markers associated with QTLs identified, could be used for selection of rust-resistant coffee. In addition to this molecular breeding strategy was also proposed to genomic selection in order to predict the phenotypic performance of individuals and to identify molecular markers associated with the trait of resistance to rust. In this study 245 F<sub>2</sub> individuals, resulting from the crossing of the susceptible cv. Catuaí Amarelo IAC 64 (UFV 2148-57) and the resistance source of Híbrido de Timor UFV 443-03, genotyped with 137 molecular markers (74 AFLP, 58 SSR, 4 RAPD and 1 specific primer). Estimates of genetic parameters were obtained by

BLASSO. The predictive capacity (CP) was measured using all individuals in the population as population training and validation. Moreover, the population was randomly divided into 200 individuals for training and 45 individuals for validation. In the first case, the CP was most predictive for the pathotype 001 (0.78), and in the second case for the race I (0.49). When using only the markers for each single greatest effect, the greater the CP validation population for the pathotype was 001 (0.49). Although it was observed that the estimated heritabilities through molecular markers for resistance to three isolates were lower than those obtained from phenotypic data. This demonstrates a higher prediction genomic values individuals in the genomic selection of compared to the phenotype. Most of the identified markers associated with QTLs in the genetic linkage map was also identified by the GWS, confirming the accuracy of the methodology used. Additionally, the GWS possible to predict successfully genomic values of individuals.

## 1. INTRODUÇÃO GERAL

Os avanços em genômica permitem sequenciar e desvendar a informação genética contida no genoma de vários organismos. Entre elas, pode-se destacar: a obtenção de marcadores moleculares cobrindo todo o genoma ligados a características agronômicas desejáveis, a construção de mapas genéticos cobrindo todos os cromossomos de um organismo, o desenvolvimento e a automação de métodos de sequenciamento de DNA genômico e de cDNA em larga escala, a transformação genética e o aperfeiçoamento de algoritmos para análise de grandes quantidades de dados (bioinformática) (BORÉM et al., 2013).

Diante desse novo cenário, as pesquisas moleculares têm-se dividido em diferentes áreas, entre elas a detecção de marcadores moleculares associados a QTLs (*Quantitative Trait Loci*) em mapa genético de ligação e o uso dos marcadores nos programas de seleção assistida por marcadores (SAM) e seleção genômica ampla (Genome Wide Selection – GWS), cada uma com particularidades importantes.

O uso de marcadores, isto é, de polimorfismo de sequência de DNA, como referência ao longo dos cromossomos, tem sido a tecnologia básica para construção de mapas genéticos de várias espécies (BOREM; CAIXETA, 2009). Um mapa genético ilustra a ordem dos genes ou marcadores genéticos em um cromossomo, assim como a distância relativa entre eles (LANDER; BOTSTEIN, 1989). Devido aos avanços obtidos em genômica e automatização, mapas genéticos de alta densidade estão disponíveis para diversas espécies.

A espécie *Coffea arabica*, contudo, ainda não dispõe de um mapa genético saturado. A obtenção de um mapa de ligação saturado para essa espécie será um referencial para localizar os genes de interesse agronômico e elucidar as relações de ligação de tais genes associados a características qualitativas e quantitativas. De fato,

mapas genéticos podem fornecer informações importantes a respeito da organização e localização de genes, assim como estrutura para identificação de regiões cromossômicas que interferem em características de importância econômica (QTL) (CRITTENDEN et al., 1993). Assim, os resultados de ligação entre marcador e QTL e também entre os próprios marcadores são essenciais no contexto da seleção genética, empregando-se informações genômicas. Além disso, a disponibilidade de um mapa genético saturado poderá facilitar enormemente os estudos genéticos e projetos de sequenciamento do genoma da espécie. A integração do mapa genético e físico proporcionará maior precisão no ordenamento dos genes nos cromossomos, facilitando o mapeamento comparativo entre espécies do mesmo gênero. Análise comparativa com base em mapas genéticos, citogenéticos e físicos, bem como informações sobre a sequência do DNA, fornecerá novos conhecimentos sobre a evolução de genomas. O mapeamento de fatores que controlam características de interesse possibilitará a clonagem e caracterização funcional de novos genes, que funcionam no desenvolvimento da planta, na adaptação aos estresses abiótico e biótico, ou na formação de outros caracteres agrônômicos (GEBHARDT et al., 2005).

Especificamente para resistência a doenças, a identificação de marcadores moleculares associados aos genes de resistência pode ser utilizada na seleção precoce de genótipos, reduzindo o tempo necessário para o desenvolvimento de grandes populações e para a estimativa dos parâmetros usados na seleção indireta (CAIXETA et al., 2009), além de servirem como fontes para a clonagem desses genes.

A SAM é especialmente útil quando a avaliação fenotípica apresenta elevado custo e exige ambientes específicos quando é subjetiva ou, ainda, quando exige destruição da planta. Além disso, quando o caráter de interesse só se manifesta em fases avançadas do desenvolvimento da planta, a SAM pode ser realizada nos estágios iniciais, o que, geralmente, não é possível quando são avaliados caracteres morfológicos (DOMINGUES et al., 2006).

Além da SAM, uma técnica que pode acelerar a identificação de indivíduos superiores em programas de melhoramento é a GWS. Por meio da seleção genômica ampla, os efeitos de todos os marcadores genéticos são estimados simultaneamente em uma população de treinamento, e os modelos são desenvolvidos para prever o valor genômico de indivíduos nas gerações futuras (MEUWISSEN et al., 2001). A GWS apresenta alta eficiência seletiva, uma vez que utiliza informação direta do

DNA na seleção. Como consequência, obtêm-se ganhos genéticos com a seleção com perspectiva de baixo custo, em comparação com a seleção tradicional baseada em dados fenotípicos (RESENDE JR. et al., 2012).

Diante disso, os objetivos deste estudo foram: i) estudar a resistência do Híbrido de Timor UFV 443-3 às raças I e II e ao patótipo 001 de *H. vastatrix* e confirmar o número e posição dos genes identificados em mapa genético de ligação, com o intuito de serem usados na SAM; ii) construir um mapa integrado saturado para *C. arabica* e localizar os genes qualitativos ou locos controladores de característica quantitativa (QTLs) associados à resistência do Híbrido de Timor UFV 443-03 à ferrugem no mapa; iii) verificar a capacidade preditiva da seleção genômica ampla em prever os valores genéticos genômicos de indivíduos superiores para resistência à ferrugem do cafeeiro sem a necessidade de fenotipá-los; e iv) comparar os marcadores de maior efeito identificados na GWS com os associados a QTLs em mapa genético de ligação.

## 2. REFERÊNCIAS

BOREM, A.; CAIXETA, E. T. **Marcadores Moleculares. 2a. ed. Visconde do Rio Branco: Suprema Grafica e Editora** 2009. 532.

BORÉM, A.; DIOLA, V.; FRITSCHÉ-NETO, R. O Melhoramento de Plantas e os Avanços Biotecnológicos. In: BORÉM, A. e FRITSCHÉ-NETO, R. (Ed.). **Biotechnologia Aplicada ao Melhoramento de Plantas**. Viçosa-MG, 2013. p.9-30.

CAIXETA, E. T.; OLIVEIRA, A. C. B.; BRITO, G. G. Tipos de marcadores moleculares. In: BORÉM, A. e CAIXETA, E. T. (Ed.). **Marcadores moleculares**. Viçosa, MG, 2009. p.11-93.

CRITTENDEN, L. B.; PROVENCHER, L.; SANTANGELO, L.; LEVIN, I.; ABPLANALP, H.; BRILES, R. W.; BRILES, W. E.; DODGSON, J. B. Characterization of a Red Jungle Fowl by White Leghorn backcross reference population for molecular mapping of the chicken genome. **Poultry Science**, v. 72, n. 2, p. 334-348, 1993. ISSN 0032-5791.

DOMINGUES, D. S.; FARRO, A. P. C.; COSCRATO, V. E.; DE MELLO, E. J.; ODA, S.; MARINO, C. L. Identificação de marcador RAPD e SCAR relacionados ao caractere florescimento precoce em *Eucalyptus grandis*. **Ciência Florestal**, v. 16, n. 3, 2006. ISSN 1980-5098.

GEBHARDT, C.; SCHMIDT, R.; SCHNEIDER, K. Plant genome analysis: the state of the art. **International review of cytology**, v. 247, p. 223-284, 2005. ISSN 0074-7696.

LANDER, E. S.; BOTSTEIN, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, v. 121, n. 1, p. 185-199, 1989. ISSN 0016-6731.

MEUWISSEN, T. H.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, p. 1819-1829, 2001.

RESENDE JR, M. F. R.; MUÑOZ, P.; RESENDE, M. D. V.; GARRICK, D. J.; FERNANDO, R. L.; DAVIS, J. M.; JOKELA, E. J.; MARTIN, T. A.; PETER, G. F.; KIRST, M. Accuracy of genomic selection methods in a standard data set of loblolly pine (*Pinus taeda* L.). **Genetics**, v. 190, n. 4, p. 1503-1510, 2012. ISSN 0016-6731.

## **CAPÍTULO 1**

### **MAPEAMENTO DE GENES E QTLS PARA A RESISTÊNCIA DO CAFEIRO À FERRUGEM**

## RESUMO

A identificação de marcadores moleculares associados a genes de resistência à ferrugem (*Hemileia vastatrix*) é estratégia essencial para a incorporação da Seleção Assistida por Marcadores no melhoramento de *Coffea arabica* que visa resistência a essa doença. Neste trabalho, a herança da resistência do Híbrido de Timor UFV 443-03 às raças I e II e ao patótipo 001 de *H. vastatrix* foi caracterizada, e os genes/QTLs identificados foram localizados em mapa genético do cafeeiro. A herança foi estudada por meio da inoculação das plantas das populações F<sub>2</sub>, retrocruzamento suscetível e retrocruzamento resistente, além do híbrido F<sub>1</sub> e dos parentais das populações (Híbrido de Timor UFV 443-03 e o Catuaí Amarelo IAC 64). Para os três isolados analisados entre várias hipóteses formuladas, foram encontradas duas hipóteses aceitáveis P (>0,05): 1) a resistência do Híbrido de Timor é governada por dois genes dominantes independentes; e 2) a resistência é conferida por três genes, dois dominantes e um recessivo. A confirmação da presença ou ausência do gene recessivo foi elucidada em mapa genético de ligação através da identificação de QTLs. O mapa genético de ligação foi construído com plantas da população F<sub>2</sub>, genotipadas com marcadores AFLP, SSR, RAPD e *primer*-específico, formando 12 grupos de ligação (GL). Foram identificados, no mapa, QTLs para os três isolados do fungo em estudo. Para a raça II, identificaram-se dois QTLs nos GLs 2 e 10, enquanto para as raças I e o patótipo 001 foram identificados quatro QTLs nos GLs 2, 3 e 10. O gene recessivo não foi identificado na análise genômica, o que possibilita inferir que dois genes dominantes e independentes conferem resistência aos isolados de *H. vastatrix* estudados. Para a raça II, os dois QTLs identificados confirmam a hipótese de dois genes dominantes envolvidos na resistência do Híbrido de Timor a

esse isolado. No entanto, os quatro QTLs identificados para a raça I e ao patótipo 001 sugerem que a resistência do Híbrido de Timor a esses dois isolados pode ser controlada por quatro genes dominantes, o que de fato pode ser justificado pelo maior número de genes de virulência presentes nesses isolados. Esses marcadores identificados, ligados aos QTLs para resistência à ferrugem, podem ser úteis para seleção assistida por marcadores moleculares, em programas de melhoramento do cafeeiro que visam resistência durável a diferentes raças de *H. vastatrix*.

## 1. INTRODUÇÃO

Existem duas espécies de importância econômica nos cafeeiros *Coffea arabica* L. e *Coffea canephora* Pierre ex A. Froehner. *C. arabica* é a única espécie autofértil e alotetraploide ( $2n = 4x = 44$ ) do gênero *Coffea*. As plantas oriundas dessa espécie são bastante aparentadas devido à sua biologia reprodutiva (autógama), fatores evolutivos e pequeno número de acessos introduzidos no Brasil (SETOTAW et al., 2013). Em contraste, a espécie *C. canephora* é diploide ( $2n = 2x = 22$ ), alógama e apresenta elevada variabilidade genética (MONDEGO et al., 2011). *C. arabica* é a espécie mais comercializada mundialmente devido à sua alta produtividade e superioridade na qualidade da bebida, em comparação com outras espécies do gênero (MENDONÇA et al., 2005; CHALFOUN et al., 2013). No entanto, apresenta suscetibilidade à ferrugem, principal doença do cafeeiro causada pelo fungo biotrófico *Hemileia vastatrix* Berk. et Br. Esse fungo ataca, principalmente, as folhas, podendo causar redução na produção e vida útil da lavoura em até 50%. A forma de controle ideal dessa doença é a obtenção de variedades resistentes, por ser um método menos oneroso e causar menor impacto ao meio ambiente (ZAMBOLIM et al., 1999). No entanto, a obtenção de cultivares resistentes tem sido constante desafio para os melhoristas, uma vez que, ao longo do tempo, as plantas podem tornar-se suscetíveis a novas raças do patógeno (VÁRZEA; MARQUES, 2005; CABRAL et al., 2009; MAIA et al., 2013).

Dessa forma, diferentes programas de melhoramento têm concentrado esforços na incorporação de resistência às variedades de *C. arabica*. A resistência do cafeeiro à *H. vastatrix* tem sido atribuída a nove genes dominantes ( $S_H1$  a  $S_H9$ ), simples ou em associação. Os genes  $S_H1$ ,  $S_H2$ ,  $S_H4$  e  $S_H5$  foram caracterizados em

cafeeiros da espécie *C. arabica* e S<sub>H</sub>3 em derivados de seleções de *C. liberica* (WAGNER; BETTENCOURT, 1965; NORONHA-WAGNER; BETTENCOURT, 1967; BETTENCOURT; CARVALHO, 1968; BETTENCOURT; NORONHA-WAGNER, 1971; BETTENCOURT; RODRIGUES JR., 1988). *C. canephora* é considerada mais resistente a doenças e pragas, e nesta espécie estão presentes os genes S<sub>H</sub>6, S<sub>H</sub>7, S<sub>H</sub>8 e S<sub>H</sub>9 (BETTENCOURT et al., 1980).

Como os genes identificados em *C. arabica* foram suplantados pelas raças de *H. vastatrix*, a transferência de genes desejáveis de *C. canephora* para *C. arabica* tem sido uma das estratégias de sucesso adotadas por melhoristas (OWUOR; VAN DER VOSSSEN, 1981; BERTHAUD; CHARRIER, 1988). Assim, uma das principais fontes de resistência utilizadas nos Programas de Melhoramento de *C. arabica* visando obtenção de variedades resistentes a *H. vastatrix* tem sido o Híbrido de Timor (HT). Esse híbrido é resultante do cruzamento natural entre *C. arabica* e *C. canephora* e apresenta resistência a diversas raças de *H. vastatrix*. Por ser tetraploide, cruza-se facilmente com as cultivares de *C. arabica*, favorecendo, assim, a transferência de sua resistência (CARVALHO et al., 1989). Os genes S<sub>H</sub>6, S<sub>H</sub>7, S<sub>H</sub>8 e S<sub>H</sub>9 foram identificados em derivados do Híbrido de Timor (BETTENCOURT et al., 1980; BETTENCOURT et al., 1992) e em outros híbridos tetraploides interespecíficos, como o Icatu (RODRIGUES JR. et al., 1975; BETTENCOURT; RODRIGUES JR., 1988). A existência de outros genes de resistência nesses materiais genéticos tem sido confirmada pela suplantação da resistência de alguns Híbridos de Timor por novas raças de *H. vastatrix* (RODRIGUES JR. et al., 2000). Além disso, alguns autores têm demonstrado a existência de outros genes de efeitos menores conferindo resistência incompleta a derivados do Híbrido de Timor (ESKES; Da COSTA, 1983; ESKES et al., 1990).

Apenas genes de efeito maior têm sido, no entanto, manipulados em programas de melhoramento visando à obtenção de cafeeiros resistentes. Embora a base genética da resistência duradoura em plantas não seja ainda bem compreendida (JOHNSON, 1981), frequentemente ela tem sido atribuída a genes de efeito menor. Dessa forma, tem sido consenso a estratégia de combinar genes de efeitos maiores e de efeitos menores para obter resistência durável.

Acumular diferentes genes de resistência em um mesmo genótipo é tarefa complexa e laboriosa de ser realizada pelos métodos tradicionais de melhoramento. Uma ferramenta eficiente e que pode auxiliar nessa estratégia seria o uso de

marcadores moleculares. Esses marcadores possibilitam inferir de forma rápida e eficiente a presença dos genes de resistência nas progêneses sem a necessidade de avaliar o fenótipo da característica, por meio da seleção assistida por marcadores moleculares (SAM) (CAIXETA et al., 2009). O uso de marcadores ligados a genes de interesse é de grande importância na seleção de genótipos, principalmente quando o programa de melhoramento tem como objetivo introduzir dois ou mais genes, quando o fenótipo é de determinação complexa, ou quando o processo de avaliação requer destruição da planta. Para melhoramento visando resistência a doenças, a SAM é particularmente importante, pois permite a seleção na ausência do patógeno ou da raça do patógeno (CAIXETA; PESTANA, 2013).

Para incorporar a SAM no melhoramento do cafeeiro, primeiramente é necessário identificar marcadores moleculares ligados aos genes de efeitos maiores e menores associados à resistência. Isso pode ser realizado detectando e mapeando *Quantitative Trait Loci* (QTLs) para a resistência ao patógeno em mapas genéticos. A identificação de locos ligados a características quantitativas é estratégia importante para conhecer melhor a genética de uma espécie. Por meio de marcadores moleculares, podem-se mapear QTLs e conhecer o número, a localização no genoma, o efeito dos genes de resistência a doenças e a interação entre eles (TAKAGI et al., 2013). Esses conhecimentos poderiam facilitar a busca e manipulação da resistência a doenças.

A construção de mapas genéticos saturados em *C. arabica* tem sido dificultada devido ao baixo nível de polimorfismo no germoplasma e complicações advindas da poliploidia. Além disso, poucos são os trabalhos que mapearam marcadores moleculares ligados aos genes de resistência à ferrugem (PRAKASH et al., 2004; MAHÉ et al., 2008; BRITO et al., 2010; DIOLA et al., 2011). Em nenhum dos trabalhos foram identificados genes de efeito menor no patossistema *Coffea-Hemileia vastatrix*. O trabalho de Prakash et al. (2004) foi realizado em uma população F<sub>2</sub> derivada de *C. liberica*, em que foram mapeadas 21 marcas AFLP ligadas ao gene de resistência S<sub>H3</sub>. Dando sequência a esse estudo, Mahé et al. (2008), trabalhando com marcadores SSR, AFLP e SCAR, identificaram três marcadores cossegando perfeitamente com o gene S<sub>H3</sub>. Brito et al. (2010) mapearam duas marcas AFLP ligadas a um dos genes de resistência à raça II de *H. vastatrix* presente no Híbrido de Timor UFV 427-15. Posteriormente, Diola et al. (2011), trabalhando com esse mesmo gene, apresentaram uma região de alta

densidade com seis marcadores SCAR delimitando uma região cromossômica de 9,45 cM e flanqueando o gene de resistência em 0,7 e 0,9 cM.

Visando caracterizar outros genes de resistência à ferrugem presentes no Híbrido de Timor, os objetivos deste estudo foram estudar a herança da resistência do Híbrido de Timor UFV 443-3 às raças I e II e ao patótipo 001 de *H. vastatrix* e confirmar o número e posição dos genes identificados no estudo de herança em mapa genético de ligação, a fim de serem usados na Seleção Assistida por Marcadores Moleculares.

## 2. MATERIAL E MÉTODOS

### 2.1. Material vegetal

Para o estudo da herança da resistência do cafeeiro à ferrugem, foram utilizadas populações derivadas do cruzamento entre o Híbrido de Timor UFV 443-03 (progenitor resistente) e o Catuaí Amarelo IAC 64 (UFV 2148-57) (progenitor suscetível). A planta  $F_1$ , denominada H-511-1, foi autofecundada em condições controladas para se obter a população segregante  $F_2$  (247 indivíduos). Além disso, foi feito o retrocruzamento com o Híbrido de Timor UFV 443-03 e com o Catuaí Amarelo UFV 2148-57, para obtenção dos retrocruzamentos resistente (RCr) (87 indivíduos) e suscetível (RCs) (115 indivíduos), respectivamente. Para o mapeamento genético, foi utilizada apenas a população  $F_2$ .

### 2.2. Análise fenotípica

Os uredósporos das raças I e II e do patótipo 001 de *H. vastatrix* foram multiplicados e inoculados de acordo com a metodologia descrita por Capucho et al. (2009). A inoculação foi efetuada em um total de 449 plantas, sendo 247 correspondentes à população  $F_2$ , 115 RCs e 87 RCr, bem como os progenitores e a planta  $F_1$ , usando-se a estratégia de discos foliares descrita por Eskes (1982).

Foram realizadas três repetições das plantas, que corresponderam a inoculações em épocas diferentes, das populações  $F_2$  e RCs; e duas repetições da RCr. A avaliação da resistência/suscetibilidade foi efetuada de acordo com Capucho et al. (2009), seguindo a escala de Tamayo et al. (1995), que se baseia na ausência ou presença de uredósporos. De acordo com essa escala, são consideradas resistentes as

notas 1 a 3; a ausência de sintomas, nota 1; as lesões cloróticas pequenas, nota 2; e as lesões cloróticas grandes sem esporulação, nota 3. A suscetibilidade (notas 4 a 6) é atribuída a lesões cloróticas grandes com poucos uredósporos, ocupando menos de 25% da área foliar (nota 4), lesões com esporulação ocupando de 25 a 50% da área (nota 5) e lesões com esporulação ocupando mais de 50% da área com uredósporos (nota 6).

### **2.3. Estudo de herança**

Com o intuito de verificar a segregação fenotípica para resistência/suscetibilidade das plantas, foi utilizado o teste qui-quadrado ( $\chi^2$ ), com o auxílio do *Software Genes* (CRUZ, 2013). Para a população F<sub>2</sub> foram testadas as proporções fenotípicas esperadas para um gene (3:1 e 1:3), dois genes (15:1, 13:3, 7:9, 9:7, 3:13 e 1:15) e três genes (63:1, 61:3, 55:9, 37:27, 27:37, 9:55, 3:61, 1:63, 57:7, 51:13, 49:15, 43:21, 25:39 e 19:45), significativos a 5% de probabilidade. Nas populações de retrocruzamento foram testadas as proporções para um gene (1:1) e para dois e três genes (3:1).

### **2.4. Genotipagem com marcadores moleculares**

#### **2.4.1. Extração DNA**

A extração de DNA dos progenitores (Híbrido de Timor UFV 443-03 e Catuaí Amarelo UFV 2148-57), da planta F<sub>1</sub> H 511-1 e das 247 plantas da população F<sub>2</sub> foi realizada utilizando a metodologia descrita por Diniz et al. (2005). A qualidade do DNA foi avaliada em gel de agarose (1%) e sua quantificação, em espectrofotômetro NanoDrop 2000 Spectrophotometer Thermo Scientific. As amostras de DNA foram padronizadas em 25 ng/ $\mu$ L e armazenadas a -20 °C.

### 2.4.2. Marcadores SSR

Inicialmente, foram efetuados testes de polimorfismo dos *primers* SSR utilizando o DNA dos progenitores e de 10 indivíduos da população F<sub>2</sub>. Foram testados 373 pares de *primers* SSR de cafeeiro (COMBES et al., 2000; ROVELLI et al., 2000; BARUAH et al., 2003; COULIBALY et al., 2003; MONCADA; MCCOUCH, 2004; PONCET et al., 2004; BHAT et al., 2005; HENDRE et al., 2008; MISSIO et al., 2009b; MISSIO et al., 2011). As reações de PCR foram realizadas de acordo com as metodologias descritas por Missio et al. (2009a). A detecção do polimorfismo dos marcadores foi realizada em gel de poliacrilamida desnaturante 6% e corada com nitrato de prata, conforme protocolo descrito por Brito et al. (2010), e em sequenciador automático *ABI 3130 xl Genetic Analyzer (Applied Biosystems)*. Para análise no sequenciador, um dos *primers* do par que amplifica cada loco microssatélite, o *forward*, foi marcado com um fluoróforo específico (6-FAM, azul, ou NED, amarelo, ou HEX, verde). No entanto, as reações de PCR para cada *primer* foram efetuadas separadamente e a reação de três fluoróforos específicos, misturada para ser analisada, caracterizando um “falso multiplex”. Cada reação de PCR foi diluída na proporção de 1:10, e 1 µL dessa diluição foi transferido para uma placa específica do sequenciador contendo 6,8 µL de formamida HI-DI Applied Biosystems e 0,2 µL do Rox-500 (padrão de peso molecular marcado com fluoróforo RED). A reação foi desnaturada em termociclador a 95 °C, por 3 min. Em seguida, as amostras foram eletroinjetadas no sequenciador, utilizando capilares de sílica de 50 cm de comprimento e 50 µm de diâmetro e preenchidas com polímero POP 7, conforme recomendação do fabricante. Os fragmentos foram analisados por meio do *software GeneMapper V4.1 Applied Biosystems*.

### 2.4.3. Marcadores AFLP

A técnica de AFLP foi realizada conforme descrito por Brito et al. (2010), com algumas modificações. Os DNAs dos progenitores (Híbrido de Timor UFV 443-03 e Catuaí Amarelo UFV 2148-57) e dos indivíduos F<sub>2</sub> resultantes desse cruzamento foram digeridos com as enzimas de restrição *EcoRI* e *MseI* por 8 h, a 37 °C. Os fragmentos gerados foram ligados aos adaptadores específicos, utilizando-se a

enzima *T4 DNA* ligase por 14 h, a 4 °C. O material digerido e ligado foi diluído 1:10 em TE pH 8,0 e armazenado a -20 °C. A amplificação pré-seletiva foi realizada com *primers* complementares aos adaptadores, contendo um nucleotídeo adicional (N) na extremidade 3'. O produto pré-amplificado foi diluído 1:40 em TE pH 8,0 e armazenado a -20 °C. Para a reação de amplificação, foram utilizadas 17 combinações de *primers* seletivos resultantes de *EcoRI* e *MseI* (E-CAG/M-TGC; E-CAT/M-AGT; E-CCT/M-ATA; E-CCC/M-AGC; E-CCC/MAGA; E-CGA/M-ACA; E-CGG/M-ACG; E-CGT/M-TCT; E-CGA/M-TCC; E-CGC/M-ATA; E-CGT/M-ATC; E-CTG/M-AAT; E-CTG/M-TAC; E-CTT/M-TGC; E-CTC/M-AGT; E-CTC/M-AAG; e E-CTT/MTCG). Aos produtos amplificados foram adicionados 8 µL de formamida (formamida 98%, EDTA pH 8,0, 10 mM, *xylene cyanol* 1 mg/mL e *bromophenol blue* 1 mg/mL). O polimorfismo dos marcadores foi observado em gel de poliacrilamida desnaturante (6%) e corado com nitrato de prata.

#### **2.4.4. Marcadores RAPD**

Conforme a metodologia descrita por Ferrão et al. (2013), foram utilizados quatro *primers* RAPD da Operon Technologies (OPA 001, OPA 014, OPA 017 e OPAB 009). Cada reação de amplificação dos *primers* RAPD consistiu de um volume total de 25 µl, contendo os seguintes componentes: 25 ng de DNA genômico, 1 U de *Taq DNA* polimerase, 0,1 mM de cada dNTP, 0,2 µM de *primer*, 50 mM de KCl, 10 mM de Tris HCl pH 8,3 e 2 mM de MgCl<sub>2</sub> e o volume final completado com água milli-Q estéril.

#### **2.4.5. Marcador primer-específico**

O *primer* específico CARF 005 foi utilizado de acordo com a metodologia descrita por Alvarenga et al. (2011). Esse *primer* amplifica uma região *Disease Resistance Gene Analog* (RGA) de 400 pb de um potencial gene de resistência a doenças. O CARF 005 foi desenvolvido a partir de sequência minerada do Projeto Brasileiro do Genoma Café e corresponde a uma porção do gene de resistência com

anotação “gi|24459841| emb|CAC82597.1| disease resistance-like protein” de *C. arabica*.

## 2.5. Construção do mapa de ligação

O mapa foi construído com base nos dados de genotipagem da população F<sub>2</sub>, utilizando-se um total de 17 combinações de *primers* AFLP, 58 *primers* SSR, quatro RAPD e um *primer* específico, o CARF 005.

Os dados dos marcadores foram codificados e analisados no *Software* GQMOL (CRUZ, 2008). A representação gráfica do mapa de ligação foi efetuada no *software* MapChart (VOORRIPS, 2002). A segregação em cada marcador foi avaliada por meio do teste qui-quadrado ( $\chi^2$ ) ( $p = 0,05$ ), usando-se a proteção de Bonferroni. A hipótese de segregação mendeliana foi de 3:1 para cada um dos marcadores. Os marcadores que apresentaram distorção de segregação foram considerados na construção do mapa, uma vez que não alteraram a posição inicial dos marcadores no mapa de ligação construído sem marcadores distorcidos.

Os grupos de ligação foram formados e ordenados, utilizando-se LOD score mínimo de 3,0 e máxima recombinação de 30%. As frequências de recombinação estimadas foram convertidas em distância genética (centiMorgans – cM).

## 2.6. Mapeamento de QTLs associados à resistência do cafeeiro à ferrugem

Os genes, previamente identificados no estudo de herança e envolvidos na resistência do cafeeiro, foram tratados como QTLs (*Quantitative Trait Loci*), com os objetivos de localizá-los no mapa genético e identificar marcadores a eles ligados. O objetivo foi identificar genes de efeitos maiores e menores envolvidos na resistência. Para isso, foram tomados os dados fenotípicos de cada planta da população F<sub>2</sub>, considerando-se a nota máxima das três repetições de inoculação. A identificação de QTL associados à resistência do cafeeiro à *H. vastatrix* foi realizada com o auxílio do *Software* GQMOL (CRUZ, 2008). Utilizou-se a metodologia de intervalo simples, por meio da análise de regressão. Os possíveis QTLs foram verificados segundo a metodologia descrita por Lander e Botstein (1989), em que o coeficiente de

determinação do QTL corresponde ao pico de maior significância estatística do QTL, sendo adotado um valor de LOD score superior a 3,0. A posição do possível QTL foi declarada quando os valores de LOD excederam os valores críticos de corte em cada grupo de ligação.

## **2.7. Influência dos marcadores moleculares identificados na severidade da doença**

Os marcadores moleculares identificados associados aos QTLs foram utilizados para verificar o seu efeito na severidade da doença. Dessa forma, a média da severidade da doença dos indivíduos contendo o marcador ligado ao QTL foi comparada com a média de doença dos indivíduos que não continham o marcador. A comparação foi realizada pelo teste F a 5% de probabilidade, pelo método de marca simples, usando-se o *Software* GQMOL (CRUZ, 2008).

A severidade média da doença em cada planta foi estimada pela nota máxima de três repetições, utilizando-se a escala de Tamayo et al. (1995) para as raças I e II e para o patótipo 001 de *H. vastatrix*.

### 3. RESULTADOS

#### 3.1. Avaliação da doença

O Híbrido de Timor UFV 443-03, a planta F<sub>1</sub> e as plantas do RCr foram resistentes às raças I e II e ao patótipo 001 de *H. vastatrix*, em todas as inoculações, enquanto o Catuaí Amarelo UFV 2148-57 foi suscetível. Esse resultado sugere que a resistência desse Híbrido de Timor é dominante para todos os três isolados do patógeno.

Nas populações F<sub>2</sub> e RCs, observou-se segregação na resistência dos três isolados de *H. vastatrix*. Na inoculação das plantas F<sub>2</sub> com a raça I do patógeno foram observadas 226 plantas resistentes (93%) e 17 suscetíveis (7%), e na população de RCs verificaram-se 84 plantas resistentes (73,7%) e 30 suscetíveis (26,3%). Na raça II, das plantas avaliadas na população F<sub>2</sub>, 235 foram resistentes (95,52%) e 11 suscetíveis (4,48%); e no RCs, 84 resistentes (73,04%) e 30 suscetíveis (26,96%). No patótipo 001, verificaram-se 231 plantas resistentes (93,90%) e 15 suscetíveis (6,10%) na população F<sub>2</sub> e 85 resistentes (73,91%) e 30 suscetíveis (26,09%) no RCs. A frequência de indivíduos F<sub>2</sub> em cada nota da escala de Tamayo et al. (1995) (notas variando de 1 a 3 (resistentes) e de 4 a 6 (suscetíveis)) pode ser verificada na Figura 1.

Com base nos dados fenotípicos obtidos na população F<sub>2</sub>, foram testadas 22 hipóteses de segregação esperada para um, dois e três genes, considerando-se o nível de significância de 5%. Nas três raças, duas proporções foram significativas: uma delas esperada para dois genes dominantes e independentes (15:1) e outra para três genes independentes, sendo dois dominantes e um recessivo (61:3) (Tabela 1). Na população de retrocruzamento suscetível, a proporção de segregação R:S significativa encontrada foi de 3:1 (Tabela 1). Essa segregação é esperada tanto para

dois genes dominantes quanto para três genes (dois dominantes e um recessivo), confirmando o resultado encontrado na população  $F_2$ .

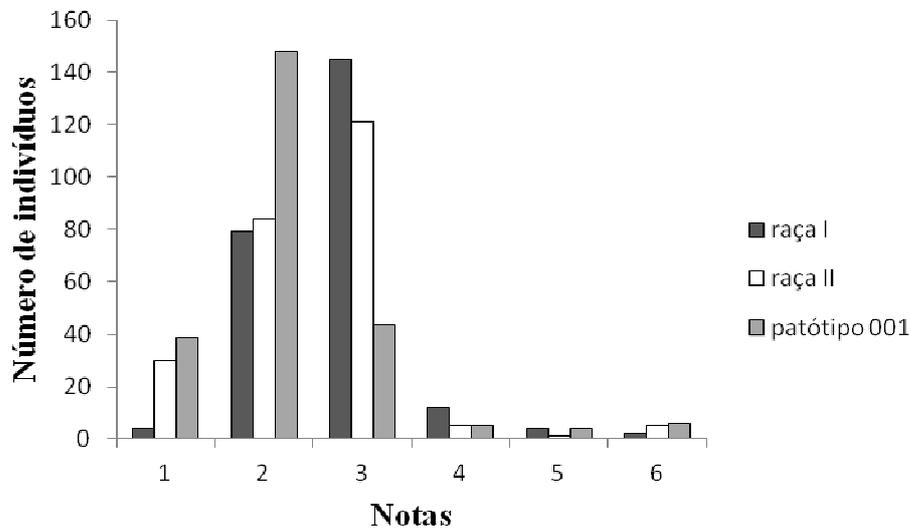


Figura 1 - Frequência de cafeeiros  $F_2$ — seguiu-se a escala de Tamayo et al. (1995). Indivíduos resistentes (1 a 3), suscetíveis (4 a 6). 1- Ausência de sintomas; 2- lesões cloróticas pequenas; 3- lesões cloróticas grandes sem esporulação; 4- lesões cloróticas grandes com poucos uredósporos, ocupando menos de 25% da área da lesão; 5- lesões com esporulação ocupando de 25 a 50% da área; e 6- lesões com esporulação ocupando mais de 50% da área com uredósporos.

Visando ampliar o conhecimento dos genes envolvidos na resistência do Híbrido de Timor à ferrugem, bem como identificar marcadores moleculares ligados aos genes caracterizados, foi construído um mapa genético de ligação contendo os fatores de resistência. Dessa forma, os genes identificados foram tratados como QTL.

Para construção do mapa de ligação, foram analisados 373 pares de *primers* SSR nos progenitores da população em estudo, sendo 58 polimórficos (15,55%) e 271 monomórficos (72,65%) e 44 não amplificaram (11,80%). A ausência de amplificação pode ser explicada pela falta de homologia de alguns *primers* construídos para *C. canephora*. Os 58 marcadores SSR polimórficos foram analisados na população  $F_2$ . Observou-se que 21 desses marcadores se comportaram como diploides e os outros 37, como poliploides.

Na genotipagem com marcadores AFLP, mais de um loco polimórfico foi verificado em cada combinação de *primer*. Assim, das 17 combinações AFLP

utilizadas, foram obtidas 74 bandas polimórficas, com média de 4,35 bandas polimórficas/*primer*. Para o RAPD, do total de 46 marcadores analisados, apenas quatro foram polimórficos, sendo observada uma banda polimórfica/*primer*.

Tabela 1 - Proporção esperada e observada para dois e três genes responsáveis pela resistência, nos progenitores F<sub>1</sub>, população F<sub>2</sub>, retrocruzamento resistente (RCr) e retrocruzamento suscetível (RCs)

	População	Proporção Esperada		Número observado		Número esperado		$\chi^2$	Probabilidade %
		R	S	R	S	R	S		
Raça I	UFV 443-03	1	0	1	0	1	0	-	-
	UFV 2148-57	0	1	0	1	0	1	-	-
	F <sub>1</sub>	1	0	1	0	1	0	-	-
	RCr <sup>1</sup>	1	0	87	0	87	0	-	-
	F <sub>2</sub>	15 <sup>a</sup>	1	226	17	227,8	15,2	0,23	63,10
	F <sub>2</sub>	61 <sup>b</sup>	3	226	17	231,6	11,4	2,90	8,87
	RCs <sup>2</sup>	3 <sup>ab</sup>	1	84	30	85,5	28,5	0,10	74,56
Raça II	UFV 443-03	1	0	1	0	1	0	-	-
	UFV 2148-57	0	1	0	1	0	1	-	-
	F <sub>1</sub>	1	0	1	0	1	0	-	-
	RCr <sup>1</sup>	1	0	87	0	87	0	-	-
	F <sub>2</sub>	15 <sup>a</sup>	1	235	11	230,6	15,4	1,33	24,92
	F <sub>2</sub>	61 <sup>b</sup>	3	235	11	234,5	11,5	0,02	87,27
	RCs <sup>2</sup>	3 <sup>ab</sup>	1	80	35	86,25	28,75	1,81	17,83
Patóti- po 001	UFV 443-03	1	0	1	0	1	0	-	-
	UFV 2148-57	0	1	0	1	0	1	-	-
	F <sub>1</sub>	1	0	1	0	1	0	-	-
	RCr <sup>1</sup>	1	0	87	0	87	0	-	-
	F <sub>2</sub>	15 <sup>a</sup>	1	231	15	230,6	15,4	0,01	92,13
	F <sub>2</sub>	61 <sup>b</sup>	3	231	15	234,5	11,5	1,09	29,54
	RCs <sup>2</sup>	3 <sup>ab</sup>	1	85	30	86,25	28,75	0,07	78,78

<sup>1</sup> Retrocruzamento resistente; <sup>2</sup> Retrocruzamento suscetível; a proporção esperada para dois genes (15:1, na geração F<sub>2</sub>); <sup>b</sup> proporção esperada para três genes (61:3, na geração F<sub>2</sub>); <sup>ab</sup> proporção esperada para dois e três genes, no retrocruzamento suscetível; R- plantas resistentes; S- suscetíveis; e  $\chi^2$ - qui-quadrado.

### 3.2. Mapa de ligação

Com o intuito de localizar todos os marcadores dominantes (AFLP e RAPD), codominantes (SSR), diploides e poliploides e o *primer*-específico em um mesmo mapa, bem como identificar os QTLs, todos os marcadores foram codificados de forma binária. Assim, para os SSR poliploides cada banda foi tratada como um loco

e para os SSR diploides, apenas o alelo acoplado ao progenitor resistente (Híbrido de Timor UFV 443-03) foi considerado. Dessa forma, um total de 137 marcadores polimórficos (74 AFLP, 58 SSR, 4 RAPD e 1 *primer*-específico (CARF 005)) foram analisados em toda a população F<sub>2</sub>. Desses, 21 apresentaram distorção na razão de segregação (15,32%), sendo 12 AFLP, 8 SSR e o *primer*-específico (Tabela 2). Os marcadores com distorção de segregação foram considerados para construção do mapa de ligação, uma vez que a ordem dos marcadores não foi alterada quando estes foram adicionados ao mapa. Os marcadores com distorção de segregação foram identificados com dois asteriscos em sua nomenclatura.

O mapa ficou, portanto, com 12 grupos de ligação (GL), contendo 111 marcadores e cobrindo uma distância de 976,78 cM do genoma (Figura 2 e Tabela 2). Vinte e seis marcadores (12 AFLP, 13 de SSR e o *primer*-específico) não se ligaram a nenhum dos GL formados. A distância entre dois marcadores adjacentes variou de 0 (isto é, dois marcadores completamente ligados) a 29,39 cM. Os tamanhos individuais dos GL variaram de 18,39 cM (GL 12) a 234,61 cM (GL 1). A maior parte dos intervalos (84,85%) não excedeu a 20 cM. Dentro de cada grupo, a distância média entre marcadores variou de 5,29 cM (GL 5) a 20,06 cM (GL 10). E para todos os grupos de ligação, a distância média foi de 9,86 cM.

Tabela 2 - Resumo do mapa de ligação para a população F<sub>2</sub> de *C. arabica*

GL	Comp. (cM)	Nº	AFLP	SSR	RAPD	RGA	Dist. média (cM)	Dist. min. e max. (cM)
1	234,60	33	25	7	1	-	11,17	0,0-25,39
2	137,10	13	11	1	1	-	13,71	0,0-22,69
3	108,70	14	8	6	-	-	15,52	0,0-23,10
4	10,60	3	3	-	-	-	10,60	0,0-10,59
5	160,40	14	3	10	1	-	13,36	0,0-25,39
6	24,60	4	2	2	-	-	24,60	0,0-24,60
7	25,90	4	-	4	-	-	12,95	0,0-13,80
8	42,40	7	3	4	-	-	10,60	0,0-15,50
9	51,10	6	-	6	-	-	10,22	0,80-25,60
10	140,50	8	7	1	-	-	20,07	8,59-29,39
11	22,00	3	-	2	1	-	11,00	9,90-12,09
12	18,40	2	-	2	-	-	18,40	18,39
Geral	976,30	110	62	45	4	-	9,86	0,00-29,39
NL		26	12	13	-	1		
Distorc.		21	12	8	-	1		

**GL**- grupo de ligação; **Comp.**- comprimento em cM; **cM**- unidade em centiMorgan; **Nº**- número de marcadores; **Dist. média (cM)**- distância média entre marcadores; **Dist. min. e max.**- distância mínima e máxima entre marcadores; **NL**- marcadores não ligados; e **Distorc.**- marcadores com segregação distorcida.

### 3.3. Identificação de QTLs

A análise conjunta dos dados genotípicos e fenotípicos permitiu a identificação de QTLs para os três isolados analisados. Apenas marcadores moleculares associados aos QTLs com valor de LOD superior a 3,0 foram considerados (Figura 2). Os intervalos de maior associação com os QTLs identificados estão indicados pela barra cheia na Figura 2. Essa barra indica várias posições do QTL, iniciando e finalizando com menor e maior LOD superior a 3,0, respectivamente. A linha fina indica a distância de marcadores adjacentes associados ao QTL com valor de LOD superior a 3,0.

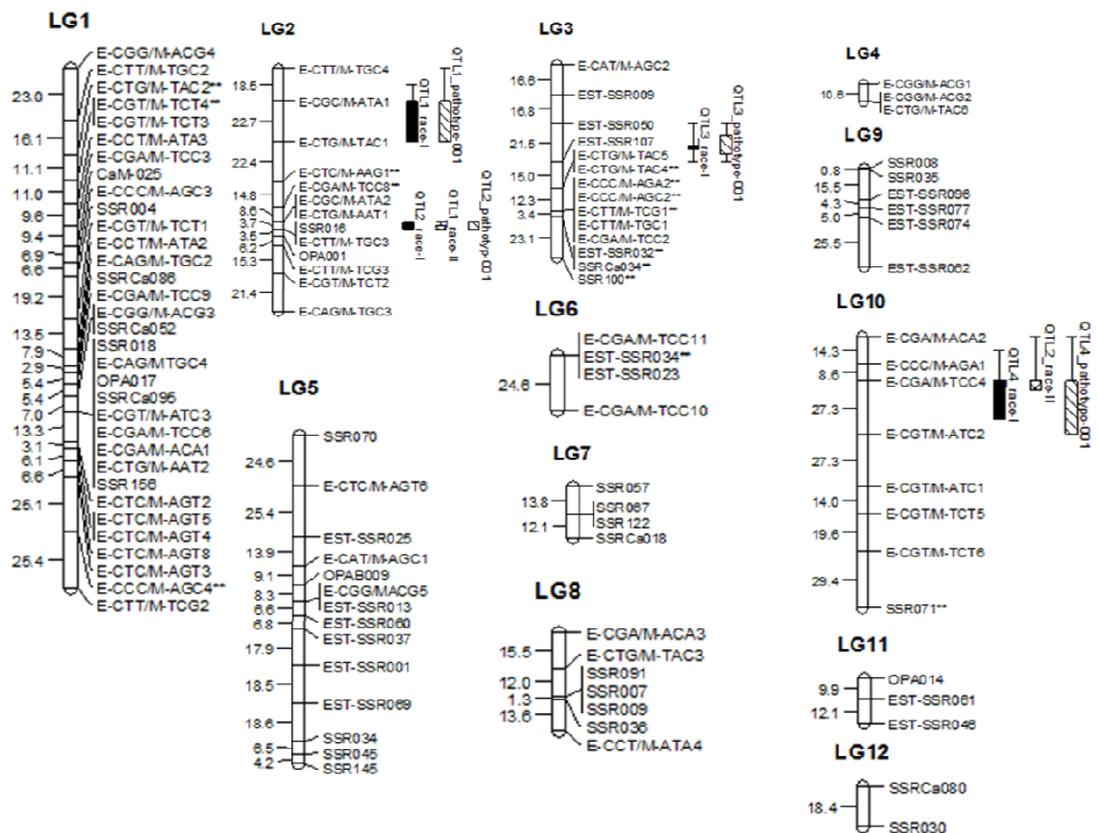


Figura 2 - Mapa de ligação de *C. arabica* contendo 111 marcadores. À direita da barra estão os SSR nos nomes dos marcadores moleculares e à esquerda, as distâncias em cM. As barras ao lado do nome dos marcadores indicam a posição dos QTLs para as raças I e II e o patótipo 001.

Dessa forma, foi possível identificar quatro QTLs para as raças I e o patótipo 001 localizados nos grupos de ligação (GL) 2, 3 e 10 e dois QTLs para a raça II

localizados nos GLs 2 e 10 (Tabela 3 e Figura 2). Todos os marcadores identificados associados ao(s) QTL(s) são alelos provenientes do progenitor resistente, Híbrido de Timor.

Para os QTLs identificados foram estimados: a sua posição no GL, os seus efeitos (efeito aditivo (a) e de dominância (d)), o tipo de interação alélica (gmd), a média de severidade da doença e quanto da variação fenotípica está sendo explicada pelo QTL ( $R^2$ ) (Tabela 3).

O grau médio de dominância (gmd) foi estimado segundo a fórmula:  $gmd = |d|/|a|$ , sendo considerado: AD = ausência de dominância quando  $gmd = 0$ ; DP = dominância parcial quando  $0 < gmd < 1,0$ ; DC = dominância completa quando  $gmd = 1,0$ ; e SD = sobredominância quando  $gmd > 1,0$ .

As estimativas dos efeitos genéticos e dos valores de gmd indicaram a ocorrência de diferentes tipos de interações alélicas para os marcadores identificados associados aos QTLs (Tabela 3). Para a raça I, as interações gênicas encontradas foram de SD (três marcadores) e DC (um marcador). Para a raça II, os dois marcadores identificados associados aos QTLs apresentaram interação alélica de SD. Para o patótipo 001, a interação gênica foi de DP (três marcadores) e SD (um marcador).

Por meio da análise da severidade média da doença (Tabela 3), foi possível observar que todos os marcadores identificados associados aos QTLs foram capazes de aumentar a resistência das plantas. Assim, a presença do alelo A1<sub>-</sub> (proveniente do progenitor resistente Híbrido de Timor) reduziu a severidade média da doença. Isso demonstra que todos os alelos de resistência identificados são provenientes do progenitor Híbrido de Timor.

Para a raça I, foi possível identificar quatro QTLs. Dois QTLs no GL 2 e um nomeado QTL1<sub>raça-I</sub> ficaram localizados a 22,70 cM do marcador E-CGC/M-ATA 1 (Figura 2), explicando 9,10% da variação fenotípica. O outro nomeado QTL2<sub>raça-I</sub> foi localizado a 3,01 cM do marcador E-CTG/M-AAT 1 (Figura 2), explicando 8,90% da variação fenotípica. O terceiro QTL (QTL3<sub>raça-I</sub>) foi identificado no GL 3 (Figura 2), a 13,01cM do marcador EST-SSR 050, explicando 5,70% da variação fenotípica. O quarto QTL (QTL4<sub>raça-I</sub>) foi identificado no GL 10 (Figura 2), localizado a 12,01 cM do E-CGA/M-TCC 4, explicando 7,80% da variação fenotípica.

Tabela 3 - Marcadores moleculares associados a QTLs para resistência do cafeeiro à ferrugem, identificados por intervalo simples por meio da análise de regressão (LOD>3)

GL	Marcador/ QTL	Pb	Pos. (cM)	LOD	R <sup>2</sup> (%)	M	a	d	gmd	Interaçã o alélica	Média de severidade		Redução na severidade média (%)	
											A <sub>1</sub> -	A <sub>2</sub> A <sub>2</sub>		
<b>Raça I</b>														
QTL 1	2	E-CGC/M-ATA 1	145	22,70	5,516	9,10	3,06	-0,49	-0,51	1,03	DC	2,782 8	3,250 0	14,38
QTL 2	2	E-CTG/M-AAT 1	670	3,01	4,941	8,90	3,60	0,39	-1,57	- 4,11	SD	2,785 0	3,295 5	15,49
QTL 3	3	EST-SSR 050	129	13,01	3,127	5,70	3,73	0,45	-1,87	- 4,19	SD	2,852 9	2,976 2	4,14
QTL 4	10	E-CGA/M-TCC 4	800	12,01	4,345	7,80	3,14	-0,11	-0,75	6,98	SD	2,771 7	3,177 4	12,77
<b>Raça II</b>														
QTL 1	2	E-CTG/M-AAT 1	670	3,70	3,147	5,80	2,75	-0,21	-0,42	2,00	SD	2,450 0	2,886 4	15,12
QTL 2	10	E-CCC/M-AGA 1	220	8,01	4,505	8,10	1,93	-0,96	1,16	- 1,21	SD	2,384 6	3,016 9	20,96
<b>Patótipo 001</b>														
QTL 1	2	E-CGC/M-ATA 1	145	17,01	8,071	13,40	2,53	-0,71	-0,56	0,79	DP	2,085 9	2,833 3	26,38
QTL 2	2	E-CTG/M-AAT 1	670	0,00	9,968	16,20	2,13	-0,90	0,29	- 0,33	DP	2,070 0	3,000 0	31,00
QTL 3	3	EST-SSR 050	129	12,01	4,229	7,50	3,43	0,40	-2,38	- 3,46	SD	2,152 0	2,619 0	17,83
QTL 4	10	E-CGA/M-TCC 4	800	0,00	8,351	13,80	2,00	-0,80	0,39	- 0,48	DP	2,005 4	2,903 2	30,92

**GL** - grupo de ligação; **Marcador/QTL**- marcadores identificados associados a QTLs das raças I e II e ao patótipo 001; **Pb**- pares de base do marcador molecular; **LOD**- logaritmo da base 10; e **R<sup>2</sup>**- coeficiente de determinação da regressão, que corresponde à porcentagem da variação fenotípica explicada pelo QTL; **m**- média da característica na população; **a**- efeito aditivo correspondente aos genótipos homocigotos, assume valores -1 e 1 para qq e QQ, respectivamente; **d**- efeito de dominância correspondente ao genótipo heterocigoto, Qq; **gmd**- grau médio de dominância, corresponde às estimativas dos efeitos genéticos e dos valores de d/a para cada QTL, indicando a ocorrência de diferentes tipos de interações alélicas; **AD**- ausência de dominância; **DP**- dominância parcial; e **SD**- sobredominância.

Dois QTLs foram identificados para a raça II. Um QTL (QTL1\_raça-II) localizado no GL 2, a 3,70 cM do marcador E-CTG/M-AAT 1, explicando 5,80% da variação fenotípica. O outro QTL (QTL2\_raça-II) localizado no GL 10 a 8,01 cM do marcador E-CCC/M-AGA1 explicou 8,10% da variação fenotípica (Figura 2).

Para o patótipo 001 foram identificados quatro QTLs (dois no GL 2, um no GL 3 e um no GL 4), Figura 2. No GL2, um dos QTLs (QTL1\_pathotype-001) identificados ficou localizado a 17,01 cM do marcador E-CGC/M-ATA 1, explicando 13,40% da variação fenotípica. O outro QTL (QTL2\_pathotype-001) no GL2 foi localizado a 0,00 cM do marcador E-CTG/M-AAT 1, explicando 16,20% da variação fenotípica. No GL 3, o QTL (QTL3\_pathotype-001) foi identificado localizado a 12,01 cM do marcador EST-SSR 050, explicando 7,50% e 8,80% da variação fenotípica. No GL10, o QTL (QTL4\_pathotype-001) identificado ficou localizado a 0,00 cM do marcador E-CGA/M-TCC 4, explicando 13,80% da variação fenotípica.

## 4. DISCUSSÃO

### 4.1. Avaliação da ferrugem

As análises fenotípicas não permitiram concluir se a resistência do Híbrido de Timor UFV 443-03 à ferrugem é condicionada por dois genes dominantes independentes ou três genes independentes, sendo dois dominantes e um recessivo, pois ambas as hipóteses não foram rejeitadas pelo teste  $\chi^2$  ( $P > 0,05$ ). Dessa forma, realizou-se a análise genômica para verificar a presença ou ausência do gene recessivo envolvido na resistência do Híbrido de Timor e, assim, auxiliar no entendimento da herança dessa importante fonte de resistência a diferentes raças de *H. vastatrix*.

Em outros acessos de Híbrido de Timor inoculados com diferentes raças, observou-se que a resistência era governada por um ou três genes dominantes. Pereira (1995), trabalhando com sete acessos de Híbrido de Timor com as raças II e XXV de *H. vastatrix*, verificou que em seis acessos a herança era governada por três genes dominantes e independentes e em apenas um por um gene dominante. Brito et al. (2010) encontraram um gene dominante conferindo resistência ao Híbrido de Timor UFV427-15 e à raça II de *H. vastatrix*.

### 4.2. Mapa de ligação

A construção do mapa genético de ligação foi dificultada pelo baixo nível de polimorfismo encontrado pelos marcadores moleculares. Do total de 373 pares de marcadores moleculares SSR analisados, apenas 58 foram polimórficos. O baixo

polimorfismo foi também identificado em outros trabalhos e justificado pela base genética estreita apresentada por *C. arabica*. Vieira et al. (2010), ao analisarem 127 marcadores SSR, encontraram apenas 22 polimórficos. Herrera et al. (2009) e Geleta et al. (2012) identificaram três e oito marcadores polimórficos, ao trabalharem com 13 e 12 SSR, respectivamente.

Os SSR são largamente utilizados para a construção de mapas de ligação, por serem, em geral, altamente polimórficos, informativos e servirem como âncoras para integração de mapas de ligação, além de requererem pequena quantidade de DNA para reação de PCR (CHISTIAKOV et al., 2006; BALL et al., 2010). No entanto, o baixo polimorfismo dos marcadores SSR encontrados neste trabalho inviabilizou a construção de um mapa genético apenas com esse marcador. Dessa forma, foram incorporados no mapa genético de ligação os marcadores moleculares AFLP e RAPD.

Os marcadores AFLP analisados apresentaram, em média, 4,35 bandas polimórficas/*primer*, enquanto nos RAPD apenas uma banda polimórfica/*primer*. O baixo polimorfismo também foi observado por Brito et al. (2010), Ky et al. (2000) e Pearl et al. (2004), que identificaram para AFLP 1,10; 6,5; e 1,34 bandas polimórficas/*primer*, respectivamente. Para os marcadores RAPD, Teixeira-Cabral et al. (2004) e Oliveira et al. (2007) encontraram 1,18 e 1,3 bandas polimórficas/*primer*, respectivamente.

Os marcadores AFLP e RAPD analisados, apesar de também apresentarem baixo polimorfismo, contribuíram para saturar o mapa genético de ligação construído. Os AFLP, por serem multilocos, permitiram a saturação do mapa genético com maior rapidez e eficiência.

Devido à poliploidia da espécie, outro fator que dificulta a construção de mapas genéticos é a distorção de segregação (DS). Do total de 137 marcadores utilizados neste trabalho, verificou-se a distorção em 21 (15,33%), valor inferior ao observado em alguns trabalhos sobre cafeeiro. A porcentagem de distorção de segregação dos marcadores depende não só da espécie em estudo, mas também do tipo de marcador e de população utilizados no mapeamento. Na construção de mapas genéticos do cafeeiro com marcadores RAPD, Oliveira et al. (2007), Teixeira-Cabral et al. (2004) e Paillard et al. (1996) verificaram 24,85%, 5,50% e 20,00% de segregação com distorção, respectivamente. O primeiro e o segundo trabalho utilizaram população de retrocruzamento derivada do Híbrido de Timor e o terceiro,

população de café robusta. Para marcadores AFLP, Ky et al. (2000) e Pearl et al. (2004) verificaram distorção na ordem de 30%, em diferentes espécies de *Coffea* sp. e 25% em população de *C. arabica*.

Geng et al. (2007) demonstraram que marcadores com distorção de segregação afetam a distância genética no mapa de ligação, causando superestimação da distância entre marcadores adjacentes. No entanto, Luo et al. (2005), mesmo partindo do pressuposto de que a distorção de segregação afeta a segregação mendeliana, não recomendaram a exclusão de marcadores com distorção de segregação em mapa genético de ligação, com o intuito de localizar QTLs. Segundo esses autores, a exclusão desses marcadores poderá causar perda de informação se o marcador com distorção de segregação estiver associado a qualquer QTL. Alguns melhoristas não utilizam marcadores com distorção de segregação para mapeamento de QTLs. Entretanto, esses marcadores têm sido rotineiramente utilizados com segurança para o mapeamento de QTL, sem efeito prejudicial sobre o resultado, em diversos trabalhos (TRUCO et al., 2007; COUTO et al., 2010; GUO et al., 2013).

Por meio de simulações, Zhang et al. (2010) e Wen e Zhang (2013) demonstraram que, se o marcador com segregação distorcida estiver intimamente ligado a qualquer QTL, a distorção não vai ter impacto significativo sobre o mapeamento do QTL. Dessa forma, neste estudo não foram retirados os marcadores com distorção de segregação para a construção do mapa genético de ligação.

O mapa genético construído apresentou distância média entre marcadores de 9,86 cM, semelhante à encontrada em outros trabalhos de mapeamento em cafeeiros. Pearl et al. (2004), Teixeira-Cabral et al. (2004) e Oliveira et al. (2007) encontraram distância média de 10,2 cM, 7,3 cM e 7,5 cM, respectivamente. Dessa forma, apesar de apresentar algumas regiões pouco saturadas, o mapa de ligação construído foi considerado apropriado para identificação de QTLs.

Uma vantagem do mapa de ligação obtido neste estudo em relação aos outros publicados para cafeeiros (PAILLARD et al., 1996; KY et al., 2000; LASHERMES et al., 2001; PEARL et al., 2004; TEIXEIRA-CABRAL et al., 2004; OLIVEIRA et al., 2007) é o maior número de indivíduos (247) analisados na população. População com número maior de indivíduos possibilita a geração de mapa genético de ligação mais acurado. Por meio de simulações, Ferreira et al. (2006) concluíram, a partir de populações F<sub>2</sub>, RILs e retrocruzamento, que populações com um total de 200 indivíduos são consideradas suficientes para a construção de mapas genéticos

razoavelmente precisos. Cruz (2006) concluiu que, em populações  $F_2$ , é inviável a utilização de amostras pequenas, como 50 indivíduos, mesmo com alto nível de saturação, sugerindo que amostras com tamanho mínimo de 100, 150 e 200 indivíduos podem ser utilizadas para recuperação completa do genoma. Bhering e Cruz (2008), por análise de simulação, verificaram que o tamanho ideal de populações de irmãos-completos, completamente informativas, é de 200 indivíduos. Para família de meios-irmãos, Good-God (2008) concluiu que não são recomendáveis populações com 50 e 100 indivíduos, mesmo com alto nível de informatividade, e que com 200 indivíduos é possível obter mapas com fidelidade.

Além dessas vantagens, o mapa construído possui 45 marcadores SSR que, por serem locos-específicos, podem ser utilizados como âncoras na integração de mapas parciais de cafeeiro, como alternativa para obtenção de um mapa genético saturado para essa cultura. Os SSR podem servir, também, para produção de sondas para FISH (*Fluorescent in Situ Hybridization*), com o intuito de integrar mapas genético e físico, possibilitando, assim, maior compreensão do genoma, sendo muito útil para mapeamento físico de genes.

### **4.3. Identificação de QTLs**

A identificação e localização de QTLs foram verificadas em mapa genético de ligação com alguns marcadores moleculares com distorção de segregação (DS) mendeliana neste trabalho. Esses marcadores foram considerados por aumentar a saturação do mapa de ligação e, portanto, evitar o descarte de marcadores ligados a QTLs, o que resultaria na perda de informação relevante. Todavia, neste trabalho os marcadores com distorção de segregação não foram associados a nenhum dos QTLs identificados.

Os marcadores moleculares identificados como associados aos QTLs explicaram a variação fenotípica de 5,70 a 16,20%. Valores semelhantes foram identificados para QTLs associados à resistência a outras doenças em diferentes espécies. Siviero et al. (2006) identificaram QTLs explicando a variação fenotípica entre 14 e 24% para a gomose do citrus. Risterucci et al. (2003) identificaram variação fenotípica de 7,5 a 12,4%, explicada pelos QTLs associados à resistência a *Phytophthora* spp. Em grão-de-bico, Sabbavarapu et al. (2013) identificaram dois

QTLs para resistência a *Fusarium wilt*, explicando 10,4 e 18,8% da variação fenotípica e seis QTLs para *Ascochyta blight*, sendo a maior variação fenotípica explicada por um dos QTLs de 31,9%. Esses autores relataram que a resistência a *F. wilt* é característica oligogênica e que os dois QTLs identificados correspondem a dois genes de efeito maior. No entanto, para *A. blight*, apenas o QTL que explicou 31,9% da variação fenotípica correspondeu a um QTL de efeito maior.

O valor da variação fenotípica explicado pela regressão ( $R^2$ ) pode ser afetado pela distância do marcador ao QTL e pela herdabilidade da característica. Por isso, é possível que os valores do coeficiente de determinação das associações dos marcadores com os respectivos QTLs encontrados neste trabalho tenham sido subestimados, uma vez que a distância dos marcadores aos QTLs varia de 0,0 a 22,70 cM. O refinamento da posição onde foram localizados os QTLs permitirá uma localização mais precisa dos seus limites, auxiliando ainda mais na identificação de genes.

Por meio da análise conjunta de QTLs para os três isolados estudados, foi possível observar que o QTL localizado no GL2 associado ao marcador E-CTG/M-AAT 1, como também o QTL localizado no GL10 associado ao marcador E-CCC/M-AGA 1, confere resistência às raças I e II e ao patótipo 001. O QTL, no GL2, associado ao marcador E-CGC/MATA 1, assim como o QTL localizado no GL3 entre os marcadores EST-SSR 050 e EST-SSR 107, confere resistência à raça I e ao patótipo 001. Dessa forma, os mesmos genes podem estar envolvidos na resistência à raça I e ao patótipo 001 ou a região contém genes próximos que conferem resistência a eles. De fato, genes de resistência estão frequentemente agrupados. O agrupamento pode ser resultante da duplicação de genes e processos de recombinação a partir de um ou alguns genes de ancestrais progenitores. Processos de recombinação e mutação entre os genes intimamente ligados ou muito próximos no cromossomo podem gerar novas especificidades de reconhecimento do patógeno, como outros genes de resistência (KEEN, 1990), importante para a coevolução planta-patógeno.

A capacidade de reconhecimento de genes de virulência de patógenos está diretamente associada aos genes de resistência da planta, ou seja, plantas resistentes possuem mais genes de resistência do que genes de virulência do patógeno. Assim, os genes de virulência ( $v_2$  e  $v_3$ ) da raça I são capazes de infectar plantas de cafeeiro que apresentam os genes de resistência  $S_{H5}$  ou  $S_{H2,5}$ . A raça II apresenta o gene de virulência  $v_5$ , capaz de suplantar a resistência em plantas que possuem apenas o gene

de resistência  $S_{H5}$ . De acordo com Capucho et al. (2009), o patótipo 001 possui o gene de virulência  $v_5$  e outro (s) não caracterizado (s), capazes de infectar plantas que contenham apenas o gene de resistência  $S_{H5}$  ou, além do  $S_{H5}$ , outro (s) não identificado (s) que seja (m) complementar (es) aos genes de virulência desse isolado. A raça I e o patótipo 001 apresentam mais de um gene de virulência, o que de fato possibilita identificar maior número de genes de resistência, como observado neste trabalho.

Os QTLs identificados no mapa de ligação possibilitaram inferir sobre a ausência do gene recessivo no Híbrido de Timor UFV 443-03, uma vez que não foi localizado QTL recessivo no mapa. Esse fato possibilita descartar a hipótese aceita no estudo de herança de que os três genes são dois dominantes e um recessivo, sugerindo que a resistência do Híbrido de Timor UFV 443-03 pode ser conferida por dois genes dominantes e independentes. Na raça II foi possível observar a presença de dois QTLs, o que confirma a presença de dois genes dominantes no Híbrido de Timor UFV 443-03, conferida à sua resistência a essa raça de *H. vastatrix*. No entanto, na raça I e no patótipo 001 foram identificados quatro QTLs, insinuando que a resistência do Híbrido de Timor UFV 443-03 a esses dois isolados pode ser condicionada por maior número de genes de resistência, em comparação com os resultados do estudo de herança.

Esses resultados indicaram que o genoma da fonte de resistência do Híbrido de Timor UFV 443-03 possui quatro importantes regiões envolvidas nos mecanismos de resistência à ferrugem. Duas regiões (GL2 e GL10) para as raças I e II e o patótipo 001 e outras duas regiões (GL2 e GL3) para a raça I e o patótipo 001. Dessa forma, a identificação dessas regiões por meio da seleção assistida por marcadores (SAM) possibilitará a introgressão desses genes/QTLs em variedades de interesse para os programas de melhoramento do cafeeiro que visam à obtenção de plantas com tolerância ou resistência durável a pelo menos três isolados (raça I e raça II e patótipo 001) de *H. vastatrix*.

Os marcadores identificados tanto neste quanto em outros trabalhos (PRAKASH et al., 2004; MAHÉ et al., 2008; BRITO et al., 2010; DIOLA et al., 2011), associados à resistência a *H. vastatrix*, podem ser utilizados em programas de melhoramento por meio da SAM, possibilitando inferir, de forma rápida e eficiente, a presença dos genes de resistência pela presença de marcadores moleculares nas progênes sem a necessidade de avaliar o fenótipo da característica (CAIXETA et al.,

2009). Além disso, os marcadores moleculares podem ser utilizados para iniciar o mapeamento de alta resolução nas regiões onde foram detectados os QTLs, visando futura clonagem desses genes de resistência à ferrugem. Os marcadores SSR podem ser utilizados como âncoras na integração de mapas parciais de cafeeiro, como alternativa para obtenção de mapa genético saturado. Isso, de fato, possibilitará maior compreensão do genoma através de técnicas de biologia molecular avançada. Assim, tais informações deverão fornecer subsídios para futuros trabalhos de melhoramento visando obter genótipos mais resistentes e produtivos.

## 5. CONCLUSÃO

Com base nos resultados, pode-se inferir que a resistência do Híbrido de Timor UFV 443-03 à raça II de *H. vastatrix* pode ser condicionada por dois genes dominantes e independentes. No entanto, a resistência desse mesmo híbrido à raça I e ao patótipo 001 de *H. vastatrix* pode ser conferida por quatro genes dominantes e independentes. Os marcadores moleculares identificados ligados aos genes/QTLs para resistência à ferrugem podem ser úteis em programas de melhoramento do cafeeiro baseado em seleção assistida por marcadores moleculares que visam resistência durável a diferentes raças de *H. vastatrix*, uma vez que permitirão selecionar plantas com resistência a pelo menos três isolados do patógeno. Assim, tais informações deverão fornecer subsídios para futuros trabalhos de melhoramento visando obter genótipos mais resistentes e produtivos.

## 6. REFERÊNCIAS

ALVARENGA, S. M.; CAIXETA, E. T.; HUFNAGEL, B.; THIEBAUT, F.; MACIEL-ZAMBOLIM, E.; ZAMBOLIM, L.; SAKIYAMA, N. S. Molecular markers from coffee genome expressed sequences potentially involved in resistance to rust. **Pesquisa Agropecuária Brasileira**, v. 46, n. 8, p. 890-898, 2011. ISSN 0100-204X.

BALL, A. D.; STAPLEY, J.; DAWSON, D. A.; BIRKHEAD, T. R.; BURKE, T.; SLATE, J. A comparison of SNPs and microsatellites as linkage mapping markers: lessons from the zebra finch (*Taeniopygia guttata*). **BMC genomics**, v. 11, n. 1, p. 218, 2010. ISSN 1471-2164.

BARUAH, A.; NAIK, V.; HENDRE, P. S.; RAJKUMAR, R.; RAJENDRAKUMAR, P.; AGGARWAL, R. Isolation and characterization of nine microsatellite markers from *Coffea arabica* L., showing wide cross-species amplifications. **Molecular Ecology Notes**, v. 3, n. 4, p. 647-650, 2003. ISSN 1471-8286.

BERTHAUD, J.; CHARRIER, A. Genetic resources of *Coffea*. In: R.J., C. e R., M. (Ed.). **Coffee**. London, v.4, 1988. p.1-42.

BETTENCOURT, A. J.; CARVALHO, A. Melhoramento visando a resistência do cafeeiro à ferrugem. **Bragantia**, v. 27, n. 4, p. 35-68, 1968.

BETTENCOURT, A. J.; LOPES, J.; PALMA, S. Fatores genéticos que condicionam a resistência às raças de *Hemileia vastatrix* Berk. et Br. dos clones-tipo dos grupos 1, 2 e 3 de derivados de Híbrido de Timor. **Brotéria Genética**, v. 13, p. 185-194, 1992.

BETTENCOURT, A. J.; NORONHA-WAGNER, M. Genetic factors conditioning resistance of *Coffea arabica* L. to *Hemileia vastatrix* Berk. & Br. **Agronomia Lusitana**, v. 31, p. 285-292, 1971.

BETTENCOURT, A. J.; NORONHA-WAGNER, M.; LOPES, M. Factor genético que condiciona a resistência do clone 1343/269 (Híbrido de Timor) à *Hemileia vastatrix* Berk. and Br. **Brotéria Genética**, v. 1, n. LXXVI, 1980.

BETTENCOURT, A. J.; RODRIGUES JR, C. J. Principles and practice of coffee breeding for resistance to rust and other diseases. **Coffee agronomy**, v. 4, p. 199-234, 1988.

BHAT, P. R.; KRISHNAKUMAR, V.; HENDRE, P. S.; RAJENDRAKUMAR, P.; VARSHNEY, R. K.; AGGARWAL, R. K. Identification and characterization of expressed sequence tags-derived simple sequence repeats markers from robusta coffee variety 'CxR'(an interspecific hybrid of *Coffea canephora* × *Coffea congensis*). **Molecular Ecology Notes**, v. 5, n. 1, p. 80-83, 2005. ISSN 1471-8286.

BHERING, L. L.; CRUZ, C. D. Tamanho de população ideal para mapeamento genético em famílias de irmãos completos. **Pesquisa Agropecuária Brasileira**, v. 43, n. 3, p. 379-385, 2008.

BRITO, G. G.; CAIXETA, E. T.; GALLINA, A. P.; ZAMBOLIM, E. M.; ZAMBOLIM, L.; DIOLA, V.; LOUREIRO, M. E. Inheritance of coffee leaf rust resistance and identification of AFLP markers linked to the resistance gene. **Euphytica**, v. 173, n. 2, p. 255-264, 2010. ISSN 0014-2336.

CABRAL, P. G. C.; ZAMBOLIM, E. M.; ZAMBOLIM, L.; LELIS, T. P.; CAPUCHO, A. S.; CAIXETA, E. T. Identification of a new race of *Hemileia vastatrix* in Brazil. **Australasian Plant Disease Notes**, v. 4, n. 1, p. 129-130, 2009. ISSN 1833-928X.

CAIXETA, E. T.; OLIVEIRA, A. C. B.; BRITO, G. G. Tipos de marcadores moleculares. In: BORÉM, A. e CAIXETA, E. T. (Ed.). **Marcadores moleculares**. Viçosa, MG, 2009. p.11-93.

CAIXETA, E. T.; PESTANA, K. N. Marcadores moleculares no melhoramento genético do cafeeiro. In: GEAFIP (Ed.). **Melhoramento Genético no Manejo de Doenças de Plantas**. Visconde do Rio Branco-MG, 2013. cap. 6, p.130. ISBN 978-85-8179-044-2.

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. **Pesquisa Agropecuária Brasileira**, v. 44, n. 3, p. 276-282, 2009.

CARVALHO, A.; FAZUOLI, L. C.; COSTA, W. Melhoramento do cafeeiro: XLI. Produtividade do Híbrido de Timor, de seus derivados e de outras fontes de resistência a *Hemileia vastatrix*. **Bragantia**, v. 48, n. 1, p. 73-86, 1989.

CHALFOUN, S. M.; PEREIRA, M. C.; CARVALHO, G. R.; PEREIRA, A. A.; SAVIAN, T. V.; BOTELHO, D. M. S. Características sensoriais de cultivares de café (*Coffea arabica* L.) na região do Alto Paranaíba. **Coffee Science**, v. 8, n. 1, p. 43-52, 2013. ISSN 1984-3909.

CHISTIYAKOV, D. A.; HELLEMANS, B.; VOLCKAERT, F. A. M. Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. **Aquaculture**, v. 255, n. 1, p. 1-29, 2006. ISSN 0044-8486.

COMBES, M.-C.; ANDRZEJEWSKI, S.; ANTHONY, F.; BERTRAND, B.; ROVELLI, P.; GRAZIOSI, G.; LASHERMES, P. Characterization of microsatellite loci in *Coffea arabica* and related coffee species. **Molecular Ecology**, v. 9, n. 8, p. 1178-1180, 2000. ISSN 1365-294X.

COULIBALY, I.; REVOL, B.; NOIROT, M.; PONCET, V.; LORIEUX, M.; CARASCO-LACOMBE, C.; MINIER, J.; DUFOUR, M.; HAMON, P. AFLP and SSR polymorphism in a *Coffea* interspecific backcross progeny [(*C. heterocalyx* × *C. canephora*) × *C. canephora*]. **Theoretical and Applied Genetics**, v. 107, n. 6, p. 1148-1155, 2003. ISSN 0040-5752.

COUTO, K. R.; SANTOS, J. B.; RAMALHO, M. A. P.; SILVA, G. S. Identification of microsatellite markers associated with seed-coat darkening in common bean. **Pesquisa Agropecuária Brasileira**, v. 45, n. 11, p. 1268-1274, 2010. ISSN 0100-204X.

CRUZ, C. D. Programa GQMOL – Genética quantitativa e molecular. [www.ufv.br/dbg/gqmol/gqmol.htm](http://www.ufv.br/dbg/gqmol/gqmol.htm), 2008.

Genes: a software package for analysis in experimental statistics and quantitative genetics. **Acta Scientiarum. Agronomy**, v. 35, n. 3, p. 271-276, 2013. ISSN 1807-8621.

CRUZ, E. M. **Efeito da saturação e do tamanho de populações F2 e de retrocruzamento sobre a acurácia do mapeamento genético**. 2006. 119 (Doctor Scientiae). Biologia Geral, Universidade Federal de Viçosa, Viçosa-MG.

DINIZ, L. E. C.; SAKIYAMA, N. S.; LASHERMES, P.; CAIXETA, E. T.; OLIVEIRA, A. C. B.; ZAMBOLIM, E. M.; 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 TECHNOLOGY**, v. 5, n. 4, p. 387, 2005. ISSN 1518-7853.

DIOLA, V.; BRITO, G. G.; CAIXETA, E. T.; MACIEL-ZAMBOLIM, E.; SAKIYAMA, N. S.; LOUREIRO, M. E. High-density genetic mapping for coffee leaf rust resistance. **Tree Genetics & Genomes**, v. 7, n. 6, p. 1199-1208, 2011. ISSN 1614-2942.

ESKES, A. B. The use of leaf disk inoculations in assessing resistance to coffee leaf rust (*Hemileia vastatrix*). **Netherlands Journal of Plant Pathology**, v. 88, n. 4, p. 127-141, 1982. ISSN 0028-2944.

ESKES, A. B.; DA COSTA, W. M. 8. Characterization of incomplete resistance to *Hemileia vastatrix* in the icatu coffee population. **Euphytica**, v. 32, n. 2, p. 649-657, 1983. ISSN 0014-2336.

ESKES, A. B.; HOOGSTRATEN, J. G. J.; TOMA-BRAGHINI, M.; CARVALHO, A. Race-specificity and inheritance of incomplete resistance to coffee leaf rust in some Icatu coffee progenies and derivatives of Híbrido de Timor. **Euphytica**, v. 47, n. 1, p. 11-19, 1990. ISSN 0014-2336.

FERRÃO, L. F. V.; CAIXETA, E. T.; SOUZA, F. D. F.; ZAMBOLIM, E. M.; CRUZ, C. D.; ZAMBOLIM, L.; SAKIYAMA, N. S. Comparative study of different molecular markers for classifying and establishing genetic relationships in *Coffea canephora*. **Plant Systematics and Evolution**, v. 299, n. 1, p. 225-238, 2013. ISSN 0378-2697.

FERREIRA, A.; SILVA, M. F.; SILVA, L. D. C. E.; CRUZ, C. D. Estimating the effects of population size and type on the accuracy of genetic maps. **Genet Mol Biol**, v. 29, n. 1, p. 187-192, 2006.

GELETA, M.; HERRERA, I.; MONZÓN, A.; BRYNGELSSON, T. Genetic diversity of Arabica coffee (*Coffea arabica* L.) in Nicaragua as estimated by simple sequence repeat markers. **The Scientific World Journal**, v. 2012, 2012.

GENG, J.-F.; ZHU, C.-S.; ZHANG, X.-W.; CHENG, Y.; ZHANG, Y.-M.; HOU, X.-L. A genetic linkage map of nonheading Chinese cabbage. **Journal of the American Society for Horticultural Science**, v. 132, n. 6, p. 816-823, 2007. ISSN 0003-1062.

GOOD-GOD, P. I. V. **Mapeamento genético em famílias de meios-irmãos por simulação computacional**. 2008. 99 (Doctor Scientiae). Biologia Geral, Universidade Federal de Viçosa, Viçosa-MG.

GUO, L. L.; LIU, X. J.; LIU, X. C.; YANG, Z. M.; KONG, D. Y.; HE, Y. J.; FENG, Z. Y. The Construction of Molecular Genetic Map of Barley Using SRAP Markers. In: (Ed.). **Advance in Barley Sciences**: Springer, 2013. p.433-440. ISBN 9400746814.

HENDRE, P. S.; PHANINDRANATH, R.; ANNAPURNA, V.; LALREMRUATA, A.; AGGARWAL, R. K. Development of new genomic microsatellite markers from robusta coffee (*Coffea canephora* Pierre ex A. Froehner) showing broad cross-species transferability and utility in genetic studies. **BMC plant biology**, v. 8, n. 1, p. 51, 2008. ISSN 1471-2229.

HERRERA, J. C.; ALVARADO A, G.; CORTINA G, H. A.; COMBES, M.-C.; ROMERO G, G.; LASHERMES, P. Genetic analysis of partial resistance to coffee leaf rust (*Hemileia vastatrix* Berk & Br.) introgressed into the cultivated *Coffea arabica* L. from the diploid *C. canephora* species. **Euphytica**, v. 167, n. 1, p. 57-67, 2009. ISSN 0014-2336.

JOHNSON, R. Durable resistance: definition of, genetic control, and attainment in plant breeding. **Phytopathology**, v. 71, n. 6, p. 567-568, 1981. ISSN 0031-949X.

KEEN, N. T. Gene-for-gene complementarity in plant-pathogen interactions. **Annual review of genetics**, v. 24, n. 1, p. 447-463, 1990. ISSN 0066-4197.

KY, C.-L.; BARRE, P.; LORIEUX, M.; TROUSLOT, P.; AKAFFOU, S.; LOUARN, J.; CHARRIER, A.; HAMON, S.; NOIROT, M. Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp.). **Theoretical and Applied Genetics**, v. 101, n. 4, p. 669-676, 2000. ISSN 0040-5752.

LANDER, E. S.; BOTSTEIN, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, v. 121, n. 1, p. 185-199, 1989. ISSN 0016-6731.

LASHERMES, P.; COMBES, M.-C.; PRAKASH, N. S.; TROUSLOT, P.; LORIEUX, M.; CHARRIER, A. Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. **Genome**, v. 44, n. 4, p. 589-595, 2001. ISSN 0831-2796.

LUO, L.; ZHANG, Y. M.; XU, S. A quantitative genetics model for viability selection. **Heredity**, v. 94, n. 3, p. 347-355, 2005. ISSN 0018-067X.

MAHÉ, L.; COMBES, M.-C.; VÁRZEA, V. M. P.; GUILHAUMON, C.; LASHERMES, P. Development of sequence characterized DNA markers linked to leaf rust (*Hemileia vastatrix*) resistance in coffee (*Coffea arabica* L.). **Molecular Breeding**, v. 21, n. 1, p. 105-113, 2008. ISSN 1380-3743.

MAIA, T. A.; MACIEL-ZAMBOLIM, E.; CAIXETA, E. T.; MIZUBUTI, E. S. G.; ZAMBOLIM, L. The population structure of *Hemileia vastatrix* in Brazil inferred from AFLP. **Australasian Plant Pathology**, p. 1-10, 2013. ISSN 0815-3191.

MENDONÇA, L. M. V. L.; PEREIRA, R. G. F. A.; MENDES, A. N. G. Parâmetro bromatológicos de grãos crus e torrados de cultivares de café (*Coffea arabica* L.). **Sociedade Brasileira de Ciência e Tecnologia de Alimentos**, v. 25, 2005. ISSN 0101-2061.

MISSIO, R. F.; CAIXETA, E. T.; ZAMBOLIM, E. M.; PENA, G. F.; RIBEIRO, A. P.; ZAMBOLIM, L.; PEREIRA, A. A.; SAKIYAMA, N. S. Assessment of EST-SSR markers for genetic analysis on coffee. **Bragantia**, v. 68, n. 3, p. 573-581, 2009a. ISSN 0006-8705.

MISSIO, R. F.; CAIXETA, E. T.; ZAMBOLIM, E. M.; PENA, G. F.; ZAMBOLIM, L.; DIAS, L. A. S.; SAKIYAMA, N. S. Genetic characterization of an elite coffee germplasm assessed by gSSR and EST-SSR markers. **Genet Mol Res**, v. 10, p. 2366-2381, 2011.

MISSIO, R. F.; CAIXETA, E. T.; ZAMBOLIM, E. M.; ZAMBOLIM, L.; SAKIYAMA, N. S. Development and validation of SSR markers for *Coffea arabica* L. **Crop Breeding and Applied Biotechnology**, v. 9, n. 4, p. 361-371, 2009b. ISSN 1518-7853.

MONCADA, P.; MCCOUCH, S. Simple sequence repeat diversity in diploid and tetraploid *Coffea* species. **Genome**, v. 47, n. 3, p. 501-509, 2004. ISSN 0831-2796.

MONDEGO, J. M. C.; VIDAL, R. O.; CARAZZOLLE, M. F.; TOKUDA, E. K.; PARIZZI, L. P.; COSTA, G. G. L.; PEREIRA, L. F. P.; ANDRADE, A. C.; COLOMBO, C. A.; VIEIRA, L. G. E. An EST-based analysis identifies new genes and reveals distinctive gene expression features of *Coffea arabica* and *Coffea canephora*. **BMC plant biology**, v. 11, n. 1, p. 30, 2011. ISSN 1471-2229.

NORONHA-WAGNER, M.; BETTENCOURT, A. J. Genetic study of the resistance of *Coffea* sp to leaf rust 1. Identification and behavior of four factors conditioning disease reaction in *Coffea arabica* to twelve physiologic races of *Hemileia vastatrix*. **Canadian Journal of Botany**, v. 45, n. 11, p. 2021-2031, 1967. ISSN 0008-4026.

OLIVEIRA, A. C. B.; SAKIYAMA, N. S.; CAIXETA, E. T.; ZAMBOLIM, E. M.; RUFINO, R. J. N.; ZAMBOLIM, L. Partial map of *Coffea arabica* L. and recovery of the recurrent parent in backcross progenies. **Crop Breeding and Applied Biotechnology**, v. 7, n. 2, p. 196-203, 2007. ISSN 1518-7853.

OWUOR, J. B. O.; VAN DER VOSSSEN, H. A. M. Interspecific hybridization between *Coffea arabica* L. and tetraploid *Coffea canephora* Pierre: I: Fertility in F1 hybrids and backcrosses to *C. arabica*. **Euphytica**, v. 30, p. 861-866, 1981.

PAILLARD, M.; LASHERMES, P.; PÉTIARD, V. Construction of a molecular linkage map in coffee. **Theoretical and Applied Genetics**, v. 93, n. 1-2, p. 41-47, 1996. ISSN 0040-5752.

PEARL, H. M.; NAGAI, C.; MOORE, P. H.; STEIGER, D. L.; OSGOOD, R. V.; MING, R. Construction of a genetic map for arabica coffee. **Theoretical and Applied Genetics**, v. 108, n. 5, p. 829-835, 2004. ISSN 0040-5752.

PEREIRA, A. A. **Herança da resistência a *Hemileia vastatrix* Berk. et Br. em cafeeiros derivados do Híbrido de Timor**. 1995. 66 (Doctor Scientiae). Fitopatologia, Universidade Federal de Viçosa, Viçosa-MG.

PONCET, V.; HAMON, P.; MINIER, J.; CARASCO, C.; HAMON, S.; NOIROT, M. SSR cross-amplification and variation within coffee trees (*Coffea* spp.). **Genome**, v. 47, n. 6, p. 1071-1081, 2004. ISSN 0831-2796.

PRAKASH, N. S.; MARQUES, D. V.; VARZEA, V. M. P.; SILVA, M. C.; COMBES, M.-C.; LASHERMES, P. Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into *C. arabica* L. **Theoretical and Applied Genetics**, v. 109, n. 6, p. 1311-1317, 2004. ISSN 0040-5752.

RISTERUCCI, A. M.; PAULIN, D.; DUCAMP, M.; N'GORAN, J. A. K.; LANAUD, C. Identification of QTLs related to cocoa resistance to three species of *Phytophthora*. **Theoretical and Applied Genetics**, v. 108, n. 1, p. 168-174, 2003. ISSN 0040-5752.

RODRIGUES JR, C. J.; BETTENCOURT, A. J.; RIJO, L. Races of the pathogen and resistance to coffee rust. **Annual Review of Phytopathology**, v. 13, n. 1, p. 49-70, 1975. ISSN 0066-4286.

RODRIGUES JR, C. J.; VÁRZEA, V. M. P.; SILVA, M. C.; GUERRA-GUIMARÃES, L.; ROCHETA, M.; MARQUES, D. V. Recent advances on coffee leaf rust. **International Scientific Symposium on Coffee**, 2000. p.179-193.

ROVELLI, P.; METTULIO, R.; ANTHONY, F.; ANZUETO, F.; LASHERMES, P.; GRAZIOSI, G. Microsatellites in Coffea Arabica L. In: (Ed.). **Coffee biotechnology and quality**: Springer, 2000. p.123-133. ISBN 9048155657.

SABBAVARAPU, M. M.; SHARMA, M.; CHAMARTHI, S. K.; SWAPNA, N.; RATHORE, A.; THUDI, M.; GAUR, P. M.; PANDE, S.; SINGH, S.; KAUR, L. Molecular mapping of QTLs for resistance to Fusarium wilt (race 1) and Ascochyta blight in chickpea (*Cicer arietinum* L.). **Euphytica**, p. 1-13, 2013. ISSN 0014-2336.

SETOTAW, T. A.; CAIXETA, E. T.; PEREIRA, A. A.; OLIVEIRA, A. C. B.; CRUZ, C. D.; ZAMBOLIM, E. M.; ZAMBOLIM, L.; SUSSUMU, N. Coefficient of parentage in Coffea arabica L. cultivars grown in Brazil. 2013.

SIVIERO, A.; CRISTOFANI, M.; FURTADO, E. L.; GARCIA, A. A. F.; COELHO, A. S. G.; MACHADO, M. A. Identification of QTLs associated with citrus resistance to Phytophthora gummosis. **Journal of applied genetics**, v. 47, n. 1, p. 23-28, 2006. ISSN 1234-1983.

TAKAGI, H.; ABE, A.; YOSHIDA, K.; KOSUGI, S.; NATSUME, S.; MITSUOKA, C.; UEMURA, A.; UTSUSHI, H.; TAMIRU, M.; TAKUNO, S. QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. **The Plant Journal**, 2013. ISSN 1365-313X.

TAMAYO, P. J.; VALE, F. X. R.; ZAMBOLIM, L.; CHAVES, G. M.; PEREIRA, A. A. Resistência do Catimor à ferrugem e virulência de raças fisiológicas de Hemileia vastatrix Berk. & Br. **Fitopatologia Brasileira**, v. 20, p. 572-576, 1995.

TEIXEIRA-CABRAL, T. A.; SAKIYAMA, N. S.; ZAMBOLIM, L.; PEREIRA, A. A.; SCHUSTER, I. Single-locus inheritance and partial linkage map of Coffea arabica L. **Crop Breeding and Applied Biotechnology**, v. 4, p. 416-421, 2004.

TRUCO, M. J.; ANTONISE, R.; LAVELLE, D.; OCHOA, O.; KOZIK, A.; WITSENBOER, H.; FORT, S. B.; JEUKEN, M. J. W.; KESSELI, R. V.; LINDHOUT, P. A high-density, integrated genetic linkage map of lettuce (*Lactuca* spp.). **Theoretical and Applied Genetics**, v. 115, n. 6, p. 735-746, 2007. ISSN 0040-5752.

VÁRZEA, V. M. P.; MARQUES, D. V. Population variability of Hemileia vastatrix vs. coffee durable resistance. **Durable resistance to coffee leaf rust**, p. 53-74, 2005.

VIEIRA, E. S. N.; VON PINHO, É. V. R.; CARVALHO, M. G. G.; ESSELINK, D. G.; VOSMAN, B. Development of microsatellite markers for identifying Brazilian Coffea arabica varieties. **Genetics and molecular biology**, v. 33, n. 3, p. 507-514, 2010. ISSN 1415-4757.

VOORRIPS, R. E. MapChart: software for the graphical presentation of linkage maps and QTLs. **Journal of Heredity**, v. 93, n. 1, p. 77-78, 2002. ISSN 0022-1503.

WAGNER, M.; BETTENCOURT, A. J. Inheritance of reaction to *Hemileia vastatrix* Berk & Br. **Coffea arabica**, p. 1960-1965, 1965.

WEN, J.; ZHANG, Y.-M. Multi-QTL mapping for quantitative traits using distorted markers. **Molecular Breeding**, v. 31, n. 2, p. 395-404, 2013. ISSN 1380-3743.

ZAMBOLIM, L.; VALE, F. X. R.; PEREIRA, A. A.; CHAVES, G. M. Manejo integrado das doenças do cafeeiro. **Produção de café com qualidade. Viçosa: UFV**, p. 134-215, 1999.

ZHANG, L.; WANG, S.; LI, H.; DENG, Q.; ZHENG, A.; LI, S.; LI, P.; LI, Z.; WANG, J. Effects of missing marker and segregation distortion on QTL mapping in F2 populations. **Theoretical and Applied Genetics**, v. 121, n. 6, p. 1071-1082, 2010. ISSN 0040-5752.

## **CAPÍTULO 2**

### **MAPA DE LIGAÇÃO INTEGRADO PARA *Coffea arabica***

## RESUMO

Mapas genéticos parciais de *Coffea arabica* foram construídos. No entanto, ainda não se dispõe de um mapa saturado e de referência para essa espécie. Para obtenção desse mapa de referência, foram construídos dois mapas parciais para posterior integração em um único mapa. O mapa parcial 1 foi desenvolvido com um total de 137 marcadores moleculares utilizados para a genotipagem de 247 plantas F<sub>2</sub> da população 1. Esse mapa ficou composto por 12 grupos de ligação (GL), contendo 111 marcadores e cobrindo uma distância de 976,30 cM (centi-Morgan) do genoma, e a distância média entre marcadores foi de 9,86 cM. Outro mapa parcial (mapa 2) foi construído com 217 plantas F<sub>2</sub> da população 2, genotipadas com 141 marcadores moleculares. Desses, 101 marcadores foram distribuídos em 12 GL, cobrindo 1.158,60 cM e uma distância média entre marcadores de 13,01 cM. A existência de 18 locos em comum (marcadores âncoras) permitiu a integração dos dois conjuntos de dados. Foi possível combinar oito GL do mapa 1 e sete GL do mapa 2. A análise resultou em um mapa final de ligação com 16 GL, e sete deles correspondem à integração. O mapa integrado ficou composto por 191 marcadores, cobrindo 1.423,40 cM do genoma de *C. arabica*. A distância média entre marcadores foi de 8,13 cM, e apenas 12% dos intervalos ultrapassaram 20 cM. Esse mapa foi usado para identificar QTLs envolvidos na resistência à ferrugem (*Hemileia vastatrix*). Para essa análise, plantas de ambas as populações foram inoculadas com a raça II desse patógeno. Dois QTLs foram identificados, os quais correspondem aos dois QTLs encontrados no mapa parcial em trabalho anterior. A integração dos mapas parciais provou ser estratégia promissora para obtenção do mapa de referência para *C. arabica*, que pode ser utilizado para investigação do genoma dessa espécie, como a

análise de QTL, clonagem posicional de genes e seleção assistida por marcadores moleculares no programa de melhoramento do cafeeiro, bem como para comparação de genomas entre as espécies de café e futura integração dos mapas genético e físico.

## 1. INTRODUÇÃO

A busca por cultivares resistentes e produtivas e a necessidade de informações detalhadas sobre o genoma da espécie *Coffea arabica* têm demandado o desenvolvimento e uso de análise genética refinada. Essa análise permite aumentar o conhecimento sobre essa espécie e gerar informações básicas para manipulação gênica. A manipulação dos genes pode ser auxiliada por meio de mapas genéticos de ligação. Os mapas genéticos determinam a ordem, posição linear e distâncias relativas entre os genes ou marcadores genéticos ao longo dos cromossomos (GRIFFITHS, 2008). Além disso, fornecem informações sobre as taxas de recombinação em relação ao comprimento total do genoma, assim como a visão sobre rearranjos genéticos entre espécies próximas dentro e entre cromossomos. Uma das considerações importantes na construção de mapas genéticos de ligação são o número e tipo de marcador molecular a ser utilizado.

Os SSR (*Simple Sequence Repeats*) são largamente utilizados para construção de mapas de ligação por serem, em geral, altamente polimórficos, informativos, locos específicos, codominantes, multialélicos, altamente reproduzíveis e de grande distribuição ao longo dos cromossomos, além de servirem como âncoras para integração de mapas de ligação, como também requerer pequena quantidade de DNA para reação de PCR (GONZALO et al., 2005; CHISTIYAKOV et al., 2006; BALL et al., 2010; NOVELLI et al., 2013). Além dessas vantagens, os marcadores SSR são transferíveis entre espécies do mesmo gênero (CUBRY et al., 2008). A transferência de SSR entre espécies do gênero *Coffea* demonstrou-se bastante eficiente (RAZAFINARIVO et al., 2013).

Grandes quantidades de marcadores são requeridas para construção de mapas de ligação saturados. Mapas de ligação saturados são construídos com o intuito de aumentar o número de marcadores disponíveis ao longo do cromossomo e são necessários para a eficiente utilização de marcadores na análise genômica e no melhoramento de plantas (DIAZ et al., 2011). Além de possibilitarem a localização de regiões genômicas que controlam caracteres de importância e a quantificação dos efeitos dessas regiões na característica estudada, esses mapas permitem posicionar e clonar precisamente um gene de interesse entre duas marcas proximamente ligadas (CAIXETA; PESTANA, 2013).

Em vista disso, a obtenção de um mapa de ligação saturado para *C. arabica* será um referencial para nele e, por meio dele, serem localizados os genes de interesse agrônômico e elucidar as relações de ligação de tais genes. Contudo, a construção de mapas saturados em *C. arabica* tem sido um desafio devido ao longo ciclo da cultura, à dificuldade em manter grandes populações no campo, ao baixo nível de polimorfismos no germoplasma disponível, especialmente os de interesse para o melhoramento, a ausência de grande número de marcadores de DNA polimórficos e a sequência genômica para essa espécie e complicações advindas da poliploidia, que incluem o grande número de grupos de ligação (PAILLARD et al., 1996). Em virtude das dificuldades citadas até o momento, os mapas de ligação construídos para a espécie *C. arabica* são parciais.

O primeiro mapa de ligação construído para *C. arabica* foi com a utilização de marcadores AFLP (PEARL et al., 2004). Esse mapa consistiu de 456 marcadores distribuídos em 31 grupos de ligação. O comprimento total do mapa foi de 1.802,8 cM, com uma distância média de 10,2 cM entre marcadores adjacentes. Outros dois mapas foram construídos com marcadores RAPD. Um deles foi construído a partir de uma população de retrocruzamento derivada do Híbrido de Timor CIFC 2570 e Mundo Novo IAC 464-18. Foram mapeados 82 marcadores RAPD em oito grupos de ligação, cobrindo uma distância de 560,6 cM do genoma (TEIXEIRA-CABRAL et al., 2004). O outro mapa também obtido de população de retrocruzamento, originado do cruzamento de outro Híbrido de Timor (UFV 445-46) com Catuaí Amarelo IAC 30, resultou em 11 grupos de ligação, com 117 marcadores RAPD, cobrindo 803,2 cM (OLIVEIRA et al., 2007).

Para minimizar as dificuldades de construção de um mapa saturado para *C. arabica*, alternativa viável seria a construção de mapas parciais utilizando diferentes

populações segregantes e diferentes tipos de marcadores de DNA e a sua posterior integração. Mapas integrados podem ser utilizados para mapear genes qualitativos importantes, bem como locos que codificam para características quantitativas (*Quantitative Trait Loci* – QTLs) (DIAZ et al., 2011).

Mapas genéticos integrados foram construídos em diversas culturas, incluindo uva (*Vitis vinifera* L.), alface (*Lactuca* spp.), milho (*Zea mays*), trevo-vermelho (*Trifolium pratense* L.), pimenta (*Capsicum* spp.), nabo (*Brassica rapa*), melão (*Cucumis melo* L.), azevém (*Lolium multiflorum*), entre outros (FALQUE et al., 2005; TRUCO et al., 2007; VEZZULLI et al., 2008; ISOBE et al., 2009; LEE et al., 2009; STUDER et al., 2010; XU et al., 2010). Esses mapas foram, geralmente, construídos com os objetivos de determinar a posição relativa de marcadores transferíveis e aumentar o número de marcadores de DNA disponível, obtendo-se, assim, mapas saturados para estudo de comparação da localização de locos controladores de características quantitativas (QTL) e genes de interesse para o germoplasma.

Da mesma forma, a construção de um mapa de ligação integrado para *C. arabica* deve permitir determinar a estabilidade das posições dos locos aleatórios nos germoplasmas de cafeeiro, bem como o aumento do número de locos no mapa de ligação. Análise de polimorfismo do genoma tem sido um foco recente na detecção de QTL e seleção assistida por marcadores em tentativa de aproveitar a diversidade genômica de uma espécie-alvo (GUPTA et al., 2005).

Mapas integrados também são usados em estudos comparativos entre espécies e em análise comparativas com base em mapas genéticos, citogenéticos e físicos. Além disso, as informações da sequência do DNA fornecerão novos conhecimentos sobre a evolução de genomas, e o mapeamento de fatores que controlam características de interesse tornar-se-á possível à clonagem e caracterização funcional de novos genes, que funcionam no desenvolvimento da planta, na adaptação ao estresse abiótico e biótico ou na formação de outros caracteres agronômicas (GEBHARDT et al., 2005).

A chave para integração de mapas distintos é, contudo, a presença de marcas comuns entre os mapas, de forma a servirem de “âncoras” para interligá-los. Somente quando há um número mínimo de âncoras disponível entre os diferentes mapas, estes podem ser representados juntos (SALGADO et al., 2011).

Em vista disso, o objetivo deste trabalho foi a construção de um mapa genético integrado para *C. arabica* a partir de duas populações F<sub>2</sub>, com o intuito de gerar um mapa- referência para essa espécie, assim como identificar genes e QTLs associados à resistência à ferrugem (*Hemileia vastatrix*).

## **2. MATERIAL E MÉTODOS**

### **2.1. Material vegetal das populações de mapeamento para o mapa integrado**

Duas populações foram utilizadas para construção do mapa integrado. A denominada população 1, constituída por 247 indivíduos  $F_2$  originada da autofecundação controlada da planta  $F_1$  H511-1, é resultante do cruzamento artificial da cv. suscetível à ferrugem (Catuaí Amarelo IAC 64 (UFV 2148-57)) com o progenitor resistente Híbrido de Timor (UFV 443-03). A outra população, denominada população 2, composta por 217 indivíduos  $F_2$  originada da autofecundação controlada da planta  $F_1$  H419-1, é oriunda do cruzamento artificial entre a cultivar suscetível à ferrugem Catuaí Amarelo IAC 30 (UFV 2143-235 EL7) com o progenitor resistente à ferrugem Híbrido de Timor – CIFIC 2570 (UFV 445-46).

### **2.2. Genotipagem das populações**

#### **2.2.1. Extração de DNA**

A extração de DNA dos progenitores e dos indivíduos  $F_2$  das duas populações de mapeamento foi realizada utilizando a metodologia descrita por Diniz et al. (2005). A qualidade do DNA foi avaliada em gel de agarose (1%) e sua quantificação, em espectrofotômetro NanoDrop 2000 Spectrophotometer Thermo

Scientific. As amostras de DNA foram padronizadas em 25 ng/ $\mu$ L e armazenadas a -20 °C.

### **2.2.2. Marcadores SSR**

Os marcadores SSR polimórficos entre os progenitores de cada população de mapeamento foram utilizados para construção do mapa de ligação. Nas populações 1 e 2, foram utilizados 58 e 41 *primers* SSR, respectivamente, para construção do mapa de ligação (Tabela 1). As reações de PCR foram realizadas de acordo com as metodologias descritas por Missio et al. (2009a).

Tabela 1 - Descrição dos *primers* SSR utilizados para genotipagem das populações 1 e 2 de mapeamento

<b>Primers</b>	<b>Forward primer (5'→3')</b>	<b>Reverse primer (5'→3')</b>	<b>Tm °C</b>	<b>Tam.</b>	<b>Literatura</b>
EST-SSR 001 <sup>a</sup>	GAAGACCAAGCACCTCAAC	ACACCAACTACGGGCAGACA	59.4	151	Souza et al. (2013)
EST-SSR 002	GAAGGGACAAAGACGCCTAA	CGACAGATGCAGGAATAAACT	58.4	184	Souza et al. (2013)
EST-SSR 006 <sup>a</sup>	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	55.3	227	Não publicado
EST-SSR 009	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	52.9	160	Não publicado
EST-SSR 013	GCCTTGCTCATATCTGCTGTCT	GATCCTTCAACTGAGCCAAA	49.6	141	Missio et al. (2011)
EST-SSR 023 <sup>a</sup>	GCCATTTCACAATCTCACCTC	AGACCCAGCAGACAACAACA	57.3	207	Missio et al. (2011)
EST-SSR 025 <sup>a</sup>	AGATACCCACCGCCTAATCCT	GCAACAACCTTCTGCTCATCC	51.7	224	Missio et al. (2011)
EST-SSR 032	CGGAGCAAGTGAATTGAACAGA	AAAGGGAAAGGAAGAAGGAG	49.6	109	Missio et al. (2011)
EST-SSR 034 <sup>a</sup>	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	51	174	Não publicado
EST-SSR 035	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	49.6	139	Não publicado
EST-SSR 037	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	49.6	164	Não publicado
EST-SSR 041	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	58.4	176	Não publicado
EST-SSR 046	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	52.3	119	Não publicado
EST-SSR 050 <sup>a</sup>	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	52.3	129	Não publicado
EST-SSR 060	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	51.7	140	Não publicado
EST-SSR 061	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	49.6	155	Não publicado
EST-SSR 062 <sup>a</sup>	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	51.7	197	Não publicado
EST-SSR 067	CGCCCGAAGATCAAACAA	TTATATCCCGCGGCAAGTCC	53.7	100	Ferrão et al. (2013)
EST-SSR 069	TGAGCTAACCAAGACCAGTTCC	CAACAGGAAATCACCGCCTA	51.7	101	Missio et al. (2011)
EST-SSR 074 <sup>a</sup>	GCATCCTACCGAGTACATAAA	TCCATCAACAACAACCGAAG	49.6	259	Souza et al. (2013)
EST-SSR 077 <sup>a</sup>	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	54.2	151	Não publicado
EST-SSR 096 <sup>a</sup>	GTGAACCTCCCTTTCCTTG	ACTGGTCTCTCGTCTGTGAA	57.3	152	Souza et al. (2013)
EST-SSR 107	XXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXX	52.9	159	Não publicado
SSR 004 <sup>a,b</sup>	AGGCCTTCATCTCAAAAACC	AGCGTACTTGAGGCAAAGA	60-55	217	Rovelli et al. (2000)
SSR 007 <sup>a</sup>	TGACATAGGGGGCTAAATTG	TTAATGGTGACGCTTTGATG	60-55	215	Rovelli et al. (2000)
SSR 008 <sup>a</sup>	CACTGGCATTAGAAAGCACC	GGCAAAGTCAATGATGACTC	60-55	198	Rovelli et al. (2000)
SSR 009 <sup>a,b</sup>	ATGCCAAGTCGGAAAAGAA	GGCAAGCTCTAGCCTTTGA	60-55	136	Rovelli et al. (2000)
SSR 016 <sup>a,b</sup>	ACCCGAAAGAAAGAACCAAG	CCACACAACCTCTCCTCATTC	60-55	140-146	Combes et al. (2000)
SSR 018 <sup>a,b</sup>	GGCTCGAGATATCTGTTTAG	TTTAATGGGCATAGGGTCC	60-55	132-166	Combes et al. (2000)
SSR 030 <sup>a,b</sup>	ATGGGGCCAACTTGAATATG	CAGGGCATCTATCTACTTCTCT	60-55	220	Poncet et al. (2004)
SSR 034 <sup>a</sup>	GGAGACGCAGGTGGTAGAAG	TCGAGAAGTCTTGGGGTGTT	58	294	Poncet et al. (2004)
SSR 035 <sup>a</sup>	CTGGCATTAGAAAGCACCTTG	GCTTGGCTCACTGTAGGACTG	60-55	158	Poncet et al. (2004)
SSR 036 <sup>a,b</sup>	ATGCCAAGTCGGAAAAGAA	CACTAGCATAATTGCGTGGA	55-50	157	Poncet et al. (2004)

Continua...

Tabela 1 - Cont.

Primers	Forward primer (5'→3')	Reverse primer (5'→3')	Tm °C	Tam.	Literatura
SSR 040 <sup>a,b</sup>	AAAGGAAAATTGTTGGCTCTGA	TCCACATACATTTCCAGCA	60–55	114	Poncet et al. (2004)
SSR 045 <sup>a</sup>	CTGAGCGCATGGAAGGAGTA	GGAGACGCAGGTGGTAGAAG	58	194	Poncet et al. (2004)
SSR 056 <sup>a,b</sup>	AGCTATCTTTATCTCACACACAC	GTTAGTGTTCGATTTGGTACT	60–55	158	Poncet et al. (2004)
SSR 057 <sup>a,b</sup>	CTCGCTTTCACGCTCTCTCT	CGGTATGTTCCCTCGTTCCTC	60–55	102	Coulibaly et al. (2003)
SSR 067 <sup>a,b</sup>	CGTCTCGTTTCACGCTCTCT	GATCTGCATGTACTGGTGCTTC	60–55	237	Poncet et al. (2004)
SSR 070 <sup>a,b</sup>	GTAACCACCACCTCCTCTGC	TGGAGGTAACGGAAGCTCTG	60	167–230	Baruah et al. (2003)
SSR 071 <sup>a</sup>	GCTAAGTTCAATTGCCCTGT	GGGTTAATTTGATTGCGTGA	59	210–232	Baruah et al. (2003)
SSR 072 <sup>b</sup>	GCCAATTGTGCAAAGTGCT	ATTCATGGGGCCTTTGTCTT	60	159–183	Baruah et al. (2003)
SSR 085 <sup>a,b</sup>	CTTCTCAGCTTTAGGTTCACTTT	TTTTGAATACTGGCTCGTGAAC	55	291	Moncada & Mccouch
SSR 091 <sup>a,b</sup>	TTGCTTCGTTTTAAATGTGAGG	AATTGTCTATAAAAAAGAGAG	55	135	Moncada & Mccouch
SSR 095 <sup>a</sup>	TAAGAAGCCACGTGACAAGTAA	TATGGCCCTTCTCGCTTTAGTT	55	352	Moncada & Mccouch
SSR-100 <sup>a,b</sup>	ACCCTTTACTACTTATTTACTCTC	ACATCCCCTTGCCATTTCTTC	55	157	Moncada & Mccouch
SSR 122 <sup>a,b</sup>	CGTCTCGTTTCACGCTCTCT	GATCTGCATGTACTGGTGCTTC	60–55	237	Poncet et al. (2004)
SSR 133 <sup>a,b</sup>	GATACAACCTATCAAACGCATC	CTGTAGGATTGGGTCATTTCC	60–55	103	Poncet et al. (2004)
SSR 137 <sup>a</sup>	AGCCTTGATATGCGGAAGAC	AGGGTTTTGGACACAAGTAGG	60–55	167	Poncet et al. (2004)
SSR 145 <sup>a</sup>	CTGCGAGGAGGAGTTAAAGATA	GCCGGGAGTCTAGGGTTCTGTG	57	138–150	Bhat et al. (2005)
SSR 157 <sup>b</sup>	AAACTAGCGGTACATTCCAACA	TGCCACATCTTGCTATTTTT	55	312	Moncada & Mccouch
SSRCa 002 <sup>a</sup>	CTGTCCCACCAACCAAAA	CTTCAACCCCAACACAC	57	258	Missio et al. (2009b)
SSRCa 012 <sup>b</sup>	TCTCCTCTATTTCGCTGTTCTC	TCTGTGCTCGTTTTTTTTCAC	56	595	Missio et al. (2009b)
SSRCa 018 <sup>a,b</sup>	GTCTCGTTTCACGCTCTCTC	ATTTTTGGCACGGTATGTTC	57	115	Missio et al. (2009b)
SSRCa 023 <sup>a,b</sup>	GACCCTTGCCTTTTGTG	GCCATTCATCCATTCATTC	56	259	Missio et al. (2009b)
SSRCa 034 <sup>a,b</sup>	TGGACAAGAAATTGAAGTGG	GGGTTTAAATTATCGGGTGT	55	257	Missio et al. (2009b)
SSRCa 045 <sup>b</sup>	GACTTGTTGCATTCCCCTA	GCGCATGTGAAGAGAAAGT	56	303	Missio et al. (2009b)
SSRCa 052 <sup>a,b</sup>	GATGGAAACCCAGAAAGTTG	TAGAAGGGCTTTGACTGGAC	57	129	Missio et al. (2009b)
SSRCa 080 <sup>a,b</sup>	GTTCTTTCCGCCGTCAAT	GAGAAGAGAGAGGAAGGGAA	57	250	Missio et al. (2009b)
SSRCa 086 <sup>a,b</sup>	AGAGAGAAGCCATGATTTGA	TCAGTCCCAGAGAATAAGGA	54	105	Missio et al. (2009b)
SSRCa 091 <sup>b</sup>	CGTCTCGTATCACGCTCTC	TGTTCCCTCGTTCCTCTCTCT	56	110	Missio et al. (2011)
SSRCa 095 <sup>a</sup>	GAGAGAGCCGAGTGAAGAGA	GAGAGAGAAGCCATGATTTGA	57	185	Missio et al. (2011)
CaM 025 <sup>a</sup>	TCCATCTTCCTTCATTTCTGCTGC	CCTTCACCCCTTTGCACTTCCT	57	186	Hendre et al. (2008)
CARF 005 <sup>a</sup>	GGACATCAACACCAACCTC	ATCCCTACCATCCACTTCAAC		400	Alvarenga et al. (2011)
CARF 014 <sup>b</sup>	AGTCTGGGAGTAAATGGACAG	GGTTTCAAGGGACACAAGG			Alvarenga et al. (2011)

<sup>a</sup>Primers genotipados na população 1 (proveniente da autofecundação controlada do híbrido F<sub>1</sub> H511-1); <sup>b</sup>Primers genotipados na população 2 (proveniente da autofecundação controlada do híbrido F<sub>1</sub> H511-1); <sup>a,b</sup>Primers genotipados nas populações 1 e 2; e Tam.= tamanho.

### 2.2.3. Marcadores AFLP

Os *primers* AFLP foram genotipados apenas na população 1. A técnica de AFLP foi realizada conforme descrito por Brito et al. (2010), utilizando 17 combinações de *primers* seletivos resultantes de *EcoRI* e *MseI* (E-CAG/M-TGC; E-CAT/M-AGT; E-CCT/M-ATA; E-CCC/M-AGC; E-CCC/MAGA; E-CGA/M-ACA; E-CGG/M-ACG; E-CGT/M-TCT; E-CGA/M-TCC; E-CGC/M-ATA; E-CGT/M-ATC; E-CTG/M-AAT; E-CTG/M-TAC; E-CTT/M-TGC; E-CTC/M-AGT; E-CTC/M-AAG; e E-CTT/MTCG).

### 2.2.4. Marcadores RAPD

Para a população do mapeamento 2, foram utilizados 67 *primers* RAPD, enquanto para a população 1, apenas quatro desse marcador (OPA 001, OPA 014, OPA 017 e OPAB 009). As reações de PCR foram realizadas de acordo com a metodologia descrita por Ferrão et al. (2013).

### 2.2.5. Marcador *primer*-específico

As duas populações de mapeamento foram analisadas com *primers* específicos para regiões RGA (*Disease Resistance Gene Analog*). Os *primers* utilizados foram o CARF 005 (população 1) e o CARF 014 (população 2). O *primer* CARF 005 amplifica uma região de 400 pb de um potencial gene de resistência a doenças. Esse *primer* foi desenvolvido a partir de sequência minerada do Projeto Brasileiro do Genoma Café e corresponde a uma porção do gene de resistência com anotação “gi|24459841| em b|CAC82597.1| *disease resistance-like protein*” de *C. arabica* (ALVARENGA et al., 2011). O *primer* CARF 014 corresponde a uma porção do gene de resistência com anotação gi|28555894| em b|CAD45029.1| NBS-LRR *disease resistance protein homologue* [*Hordeum vulgare*] (ALVARENGA et al., 2011).

### 2.3. Construção de mapas parciais

Os dados dos marcadores para cada população de mapeamento foram codificados e analisados no *Software* GQMOL (CRUZ, 2008). A representação gráfica dos marcadores em mapa genético de ligação foi efetuada em *software* MapChart 2.2 (VOORRIPS, 2002). A segregação em cada marcador foi avaliada por meio do teste qui-quadrado ( $\chi^2$ ) ( $P = 0,05$ , grau de liberdade = 1), usando-se a proteção de Bonferroni. A hipótese de segregação mendeliana foi de 3:1 para cada um dos marcadores (AFLP, RAPD, SSR e RGA). Quanto aos marcadores SSR, apesar de serem codominantes, foi utilizado apenas o alelo do progenitor resistente, em vista da não disposição de *softwares* que detectam QTLs em mapas genéticos de ligação para marcadores codominantes e dominantes conjuntamente. Os grupos de ligação foram formados e ordenados utilizando LOD *score* mínimo de 3,0 e máxima recombinação de 30%. As frequências de recombinação estimadas foram convertidas em distância genética (centiMorgans – cM). O mapa da população 1 foi denominado mapa 1 e o da população 2, mapa 2.

### 2.4. Integração de mapas

Os grupos de ligação das duas populações de mapeamento contendo marcadores comuns foram selecionados para serem integrados. A integração foi efetuada utilizando o *Software* GQMOL, pela metodologia proposta por Salgado et al. (2011). A diagramação do mapa foi realizada no *software* MapChart 2.2 (VOORRIPS, 2002).

De acordo com a metodologia proposta por Salgado et al. (2011), a integração envolveu quatro etapas: identificação dos marcadores comuns entre as duas populações, os quais foram identificados *a priori* como âncoras e, em seguida, para associar os correspondentes grupos de ligação do mapa integrado; obtenção de um mapa alinhado, sem enfatizar distâncias nem o posicionamento mais apropriado; estabelecimento do melhor ordenamento e geração do mapa ordenado; e obtenção do mapa integrado efetivo, com as posições relativas dos marcadores-âncora apropriadamente calculadas para melhor posicioná-los em um único grupo de

ligação. A precisão da posição dos marcadores-âncora depende do tamanho do intervalo do marcador-âncora no mapa parcial.

## **2.5. Análise fenotípica para identificação de QTLs**

Os uredósporos da raça II de *H. vastatrix* foram inoculados de acordo com a metodologia descrita por Capucho et al. (2009) nas duas populações de mapeamento, em três épocas diferentes. A avaliação da resistência/suscetibilidade seguiu a escala de Tamayo et al. (1995), que se baseia na ausência ou presença de uredósporos. De acordo com essa escala, são consideradas resistentes (notas 1 a 3), a ausência de sintomas (nota 1), as lesões cloróticas pequenas (nota 2) e as lesões cloróticas grandes sem esporulação (nota 3). A suscetibilidade (notas 4 a 6) é atribuída a lesões cloróticas grandes com poucos uredósporos, ocupando menos de 25% da área foliar (nota 4), lesões com esporulação ocupando de 25 a 50% da área (nota 5) e lesões com esporulação, ocupando mais de 50% da área com uredósporos (nota 6).

A identificação de QTLs associados à resistência do cafeeiro à *H. vastatrix* foi realizada com o auxílio do *Software* GQMOL (CRUZ, 2008), utilizando-se a metodologia de intervalo simples, por meio da análise de regressão. Os possíveis QTLs foram verificados segundo a metodologia descrita por Lander e Botstein (1989), em que o coeficiente de determinação do QTL corresponde ao pico de maior significância estatística do QTL, sendo adotado um valor de LOD score superior a 3,0. A posição do possível QTL foi declarada quando os valores de LOD excederam os valores críticos de corte em cada grupo de ligação.

### 3. RESULTADOS

#### 3.1. Genotipagem das populações

Inicialmente, foram feitos testes de detecção de polimorfismo entre os progenitores das duas populações de mapeamento. Do total de 373 marcadores SSR testados, 58 (15,54%) e 41 (10,99%) foram polimórficos para os progenitores das populações 1 e 2, respectivamente. Das 17 combinações de AFLP utilizadas para a população 1, foram obtidas 74 bandas polimórficas, média de 4,35 bandas polimórficas/*primer*. Os quatro *primers* RAPD utilizados na população 1 apresentaram apenas uma banda polimórfica/*primer*. Para a população 2, dos 67 *primers* RAPD utilizados, obtiveram-se 99 (média de 1,48 banda/*primer*) bandas polimórficas. Com base em todos os marcadores utilizados, um total de 137 (58 SSR, 74 AFLP, 4 RAPD e 1 RGA) marcadores moleculares foram utilizados para as populações 1 e 141 (99 RAPD, 41 SSR e 1 RGA) para a população 2.

#### 3.2. Construção de mapas parciais

Para obtenção do mapa de ligação integrado, primeiramente construíram-se dois mapas parciais originados das duas populações F<sub>2</sub> de *C. arabica*. Os mapas parciais de ligação foram chamados de mapa 1 (população 1) e mapa 2 (população 2).

O mapa de ligação 1 constituiu-se de 111 marcadores alocados em 12 grupos de ligação (GL), cobrindo 976,30 cM do genoma. Vinte e seis marcadores (12 AFLP, 13 de SSR e o *primer*-específico) não se ligaram a nenhum dos GL formados. A

distância entre dois marcadores adjacentes variou de 0 (isto é, dois marcadores completamente ligados) a 29,39 cM. Os tamanhos individuais dos GL variaram de 10,60 cM (GL 4) a 234,60 cM (GL 1). A maior parte dos intervalos (84,85 %) não excedeu a 20 cM. Dentro de cada grupo, a distância média entre marcadores variou de 10,22 cM (GL 9) a 24,60 cM (GL 6). E para todos os grupos de ligação a distância média foi de 9,86 cM (Tabela 2).

No mapa de ligação 2, um total de 101 marcadores moleculares foram distribuídos em 12 GLs, cobrindo 1.158,60 cM do genoma. Quarenta marcadores (24 RAPD e 16 SSR) não se ligaram a nenhum dos GLs formados. Os tamanhos individuais dos GLs variaram de 4,10 cM (GL 5) a 343,30 cM (GL 1). A distância entre dois marcadores adjacentes variou de 0 a 30,0 cM. A maior parte dos intervalos (70,78%) não excedeu a 20 cM. Dentro de cada grupo, a distância média entre marcadores variou de 4,10 cM (GL 5) a 28,60 cM (GL 12). E para todos os grupos de ligação a distância média foi de 13,01 cM (Tabela 2).

### **3.3. Integração de mapas**

Do total de marcadores SSR utilizados para construção dos mapas parciais, 38 comuns foram genotipados nas duas populações. Desses, 18 marcadores foram localizados em ambos os mapas e, portanto, serviram como âncora para integração. Com base nesses marcadores, foi possível integrar oito GLs do mapa 1 com sete GLs do mapa 2 (Figuras 1 a 7). Os demais grupos de ligação dos dois mapas parciais não foram integrados devido à ausência de marcadores que permitissem o processo de ancoragem e consequente integração.

Dessa forma, um mapa integrado composto por 16 GLs foi construído (Figura 8), dos quais sete grupos correspondem à integração entre os mapas parciais. O mapa integrado ficou constituído por 191 (77 RAPD, 62 AFLP, 51 SSR e 1 RGA) marcadores, cobrindo 1.423,40 cM do genoma de *C. arabica*. A distância média interlocos foi de 8,13 cM, e apenas 12% dos intervalos entre marcadores excederam 20 cM. Os tamanhos individuais dos GLs variaram de 4,10 cM (GL 12) a 355,0 cM (GL 1). Dentro de cada grupo, a distância média entre marcadores variou de 4,10 cM (GL 12) a 28,60 cM (GL 16) (Tabela 2 e Figura 8).

Considerando apenas os resultados dos GLs utilizados para integração dos mapas parciais, pode-se verificar na Tabela 3 que a distância média entre marcadores foi de 9,14 cM para o mapa 1 e 12,41 cM para o mapa 2, respectivamente. Avaliando apenas os resultados dos sete GLs integrados, verificou-se que a distância média entre marcadores foi de 6,98 cM e que 92,71% dos intervalos entre marcadores não excederam 20 cM (Tabela 3).

### **3.4. Identificação de QTLs**

O mapa integrado foi utilizado para identificar QTLs envolvidos na resistência à ferrugem do cafeeiro (*H. vastatrix*). Para essa análise, plantas originadas de ambas as populações foram inoculadas com a raça II de *H. vastatrix*. Dois QTLs foram identificados, nomeados de QTL1 e QTL2 (Figura 2- GL 2 e Figura 8- GLs 2 e 11).

Tabela 2 - Resumo dos mapas parciais de ligação e mapa integrado

GL	Mapa 1				Mapa 2				Mapa Integrado			
	Comp. (cM)	Nº	Dist. média (cM)	Dist. min. e max. (cM)	Comp. (cM)	Nº	Dist. média (cM)	Dist. min. e max. (cM)	Comp. (cM)	Nº	Dist. média (cM)	Dist. min. e max. (cM)
1	234,60	33	7,33	0,0-25,39	343,30	33	10,73	0,00-25,89	355,00	67	5,38	0,30-23,00
2	137,10	13	11,42	0,0-22,69	187,00	14	14,38	0,00-30,00	143,60	23	6,52	0,80-18,60
3	108,70	14	8,36	0,0-23,10	219,40	19	12,18	0,00-27,19	207,80	29	7,42	0,00-27,19
4	10,60	3	5,30	0,0-10,59	143,60	12	13,05	2,79-28,19	224,00	26	8,96	0,00-24,60
5	160,40	14	12,34	0,0-25,39	4,10	2	4,10	4,09	43,20	5	10,80	4,69-17,43
6	24,60	4	8,20	0,0-24,60	35,60	4	11,87	8,59-17,60	61,40	6	12,28	8,92-25,00
7	25,90	4	8,63	0,0-13,80	34,20	3	17,10	11,59-22,60	20,30	2	20,30	0,00-10,59
8	42,40	7	7,06	0,0-15,50	65,60	5	16,40	11,80-25,00	10,60	3	5,30	0,00-24,60
9	51,10	6	10,22	0,80-25,60	20,20	2	20,20	20,19	24,60	4	8,20	0,80-25,60
10	140,50	8	20,07	8,59-29,39	22,80	2	22,80	22,80	51,10	6	2,04	8,59-29,39
11	22,00	3	11,00	9,90-12,09	54,20	3	13,55	25,60-28,60	140,50	8	2,87	18,39
12	18,40	2	18,40	18,39	28,60	2	28,60	28,60	4,10	2	4,10	4,09
13	-	-	-	-	-	-	-	-	34,20	3	8,55	11,59-22,60
14	-	-	-	-	-	-	-	-	20,20	2	20,20	20,19
15	-	-	-	-	-	-	-	-	54,20	3	13,55	25,60-28,60
16	-	-	-	-	-	-	-	-	28,60	2	28,60	28,60
Geral	976,30	111	9,86	0,00-29,39	1158,60	101	13,01	0,00-30,00	1423,40	191	8,13	0,00-29,39
NL		26				40				-		
Distorc.		21				28						

GL- grupo de ligação; Comp.- comprimento em cM; cM- unidade em centimorgan; Nº- número de marcadores; Dist. média (cM)- distância média entre marcadores; Dist. min. e max- distância mínima e máxima entre marcadores; NL- marcadores não ligados; e Distorc.- marcadores distorcidos.

Tabela 3 - Resumo dos grupos de ligação dos mapas parciais que foram utilizados para integração e apenas grupo de ligação integrado

GL	Mapa 1				Mapa 2				Mapa integrado			
	Comp. (cM)	Nº	Dist. média (cM)	Dist. min. e max. (cM)	Comp. (cM)	Nº	Dist. média (cM)	Dist. min. e max. (cM)	Comp. (cM)	Nº	Dist. média (cM)	Dist. min. e max. (cM)
1	234,60	33	7,33	0,0-25,39	343,30	33	10,73	0,00-25,89	355,00	67	5,38	0,30-23,00
2	137,10	13	11,42	0,0-22,69	187,00	14	14,38	0,00-30,00	143,60	23	6,52	0,80-18,60
3	108,70	14	8,36	0,0-23,10	219,40	19	12,18	0,00-27,19	207,80	29	7,42	0,00-27,19
4	-	-	-	-	143,60	12	13,05	2,79-28,19	224,00	26	8,96	0,00-24,60
5	160,40	14	12,34	0,0-25,39	-	-	-	-	43,20	5	10,80	4,69-17,43
6	-	-	-	-	35,60	4	11,87	8,59-17,60	61,40	6	12,28	8,92-25,00
7	25,90	4	8,63	0,0-13,80	-	-	-	-	20,30	2	20,30	0,00-10,59
8	42,40	7	7,06	0,0-15,50	65,60	5	16,40	11,80-25,00	-	-	-	-
9	-	-	-	-	-	-	-	-	-	-	-	-
10	-	-	-	-	22,80	2	22,80	22,80	-	-	-	-
11	22,00	3	11,00	9,90-12,09	-	-	-	-	-	-	-	-
12	18,40	2	18,40	18,39	-	-	-	-	-	-	-	-
13	-	-	-	-	-	-	-	-	-	-	-	-
14	-	-	-	-	-	-	-	-	-	-	-	-
15	-	-	-	-	-	-	-	-	-	-	-	-
16	-	-	-	-	-	-	-	-	-	-	-	-
Geral	749,50	90	9,14	0,00-25,39	1017,30	89	13,01	0,0-30,00	1055,30	158	6,58	0,00-27,19

GL- grupo de ligação; Comp.- comprimento em cM; cM- unidade em centiMorgan; Nº- número de marcadores; Dist. média (cM)- distância média entre marcadores; e Dist. min.; e max- distâncias mínima e máxima entre marcadores.

## Mapa Integrado

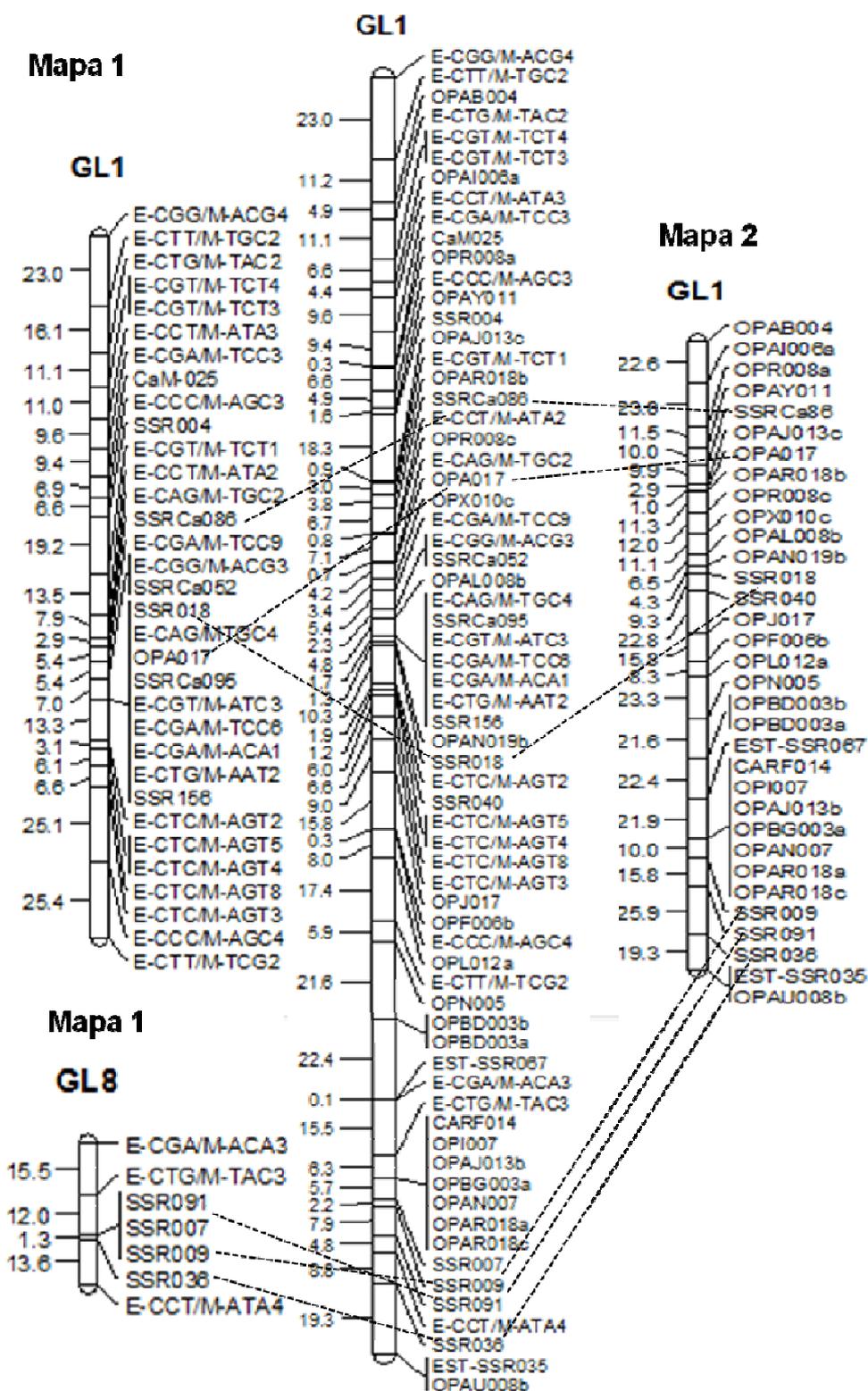


Figura 1 - Grupo de ligação (GL) 1 do mapa integrado, proveniente da integração dos GLs 1 e 8 do mapa 1 (população 1) e GL 1 do mapa 2 (população 2).

Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais.

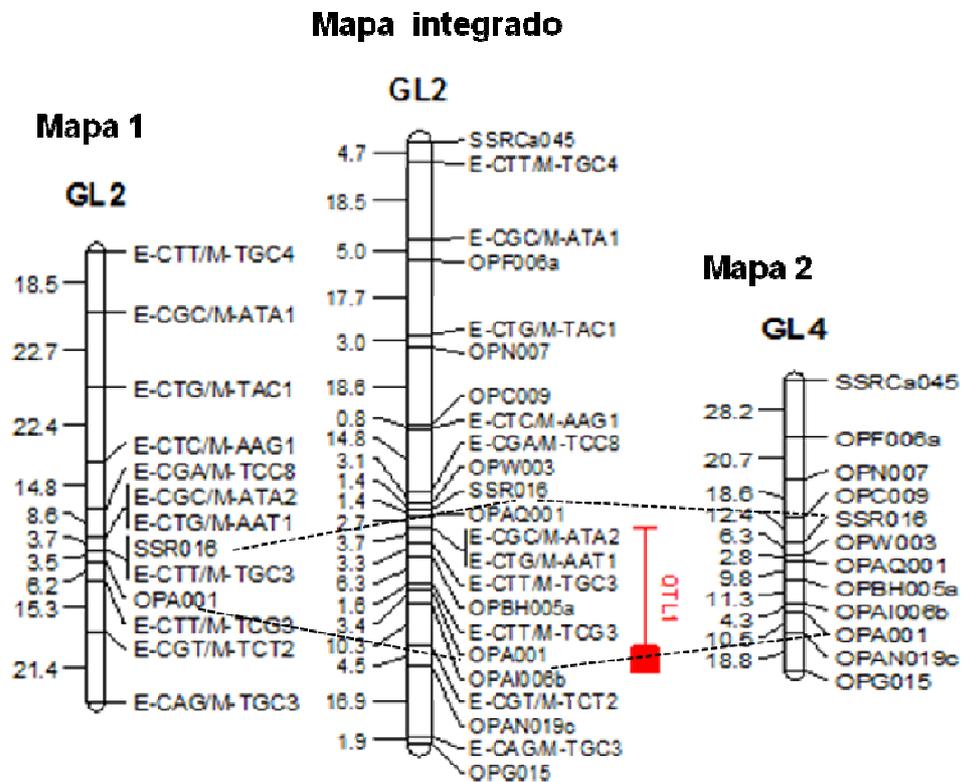


Figura 2 - Grupo de ligação (GL) 2 do mapa integrado, proveniente da integração do GL 2 do mapa 1 (população 1) e GL 4 do mapa 2 (população 2). Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais. A barra vermelha ao lado do GL 2 indica a posição do QTL1 identificado associado à resistência do cafeeiro à raça II de *H. vastatrix*.

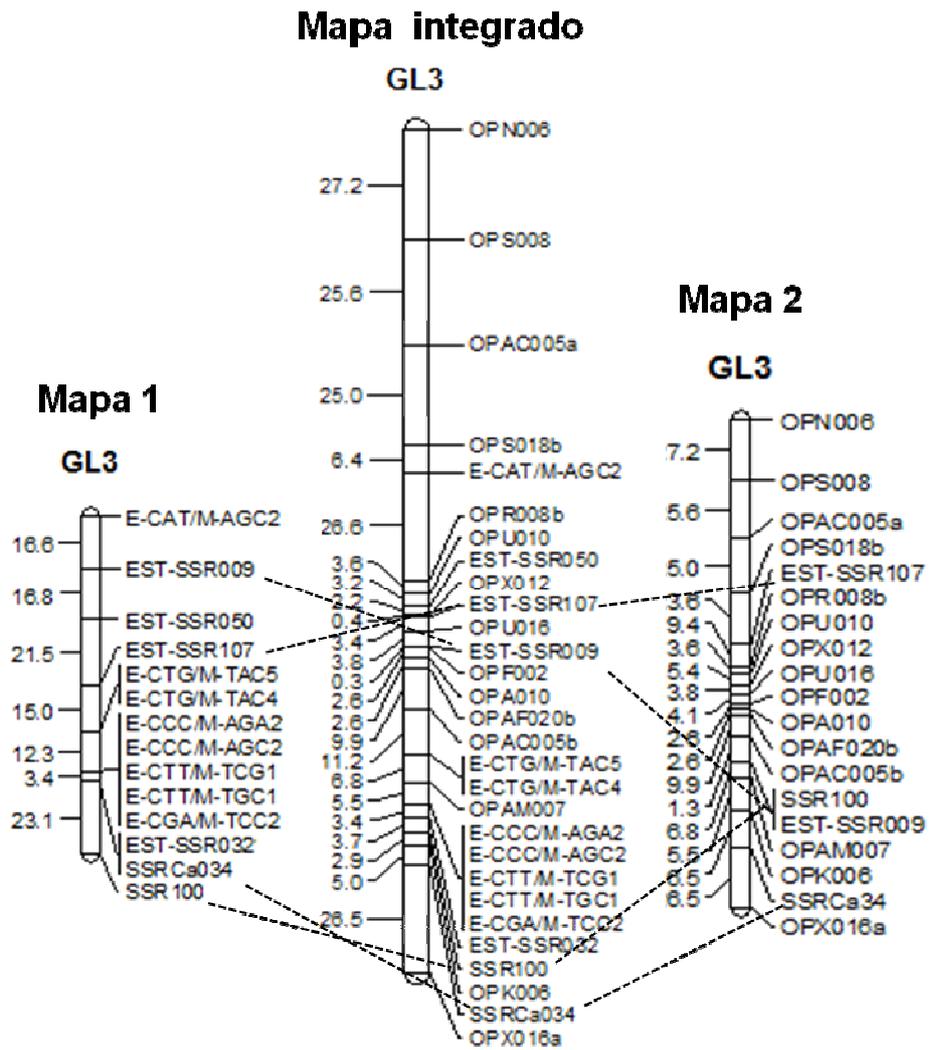


Figura 3 - Grupo de ligação (GLs) 3 do mapa integrado, proveniente da integração do GL 3 do mapa 1 (população 1) e GL3 do mapa 2 (população 2). Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais.

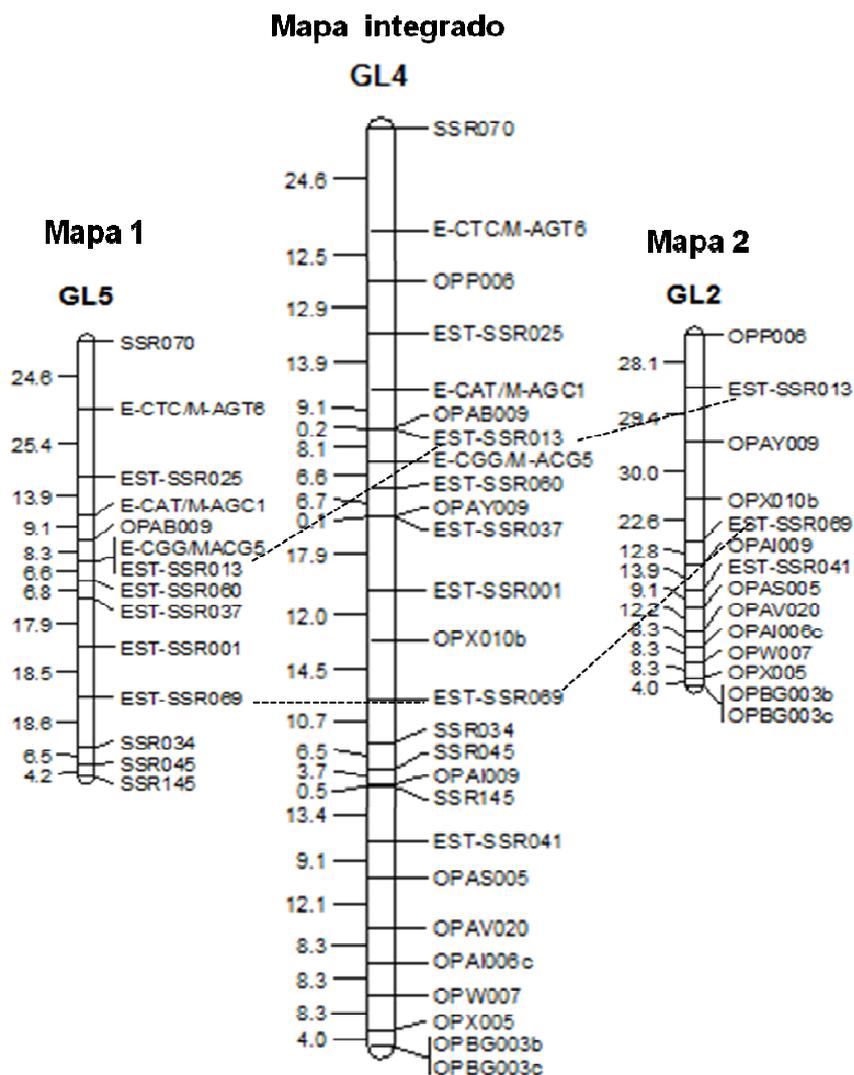


Figura 4 - Grupo de ligação (GL) 4 do mapa integrado, proveniente do GL 6 do mapa 1 (população 1) e GL 2 do mapa 2 (população 2). Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais.

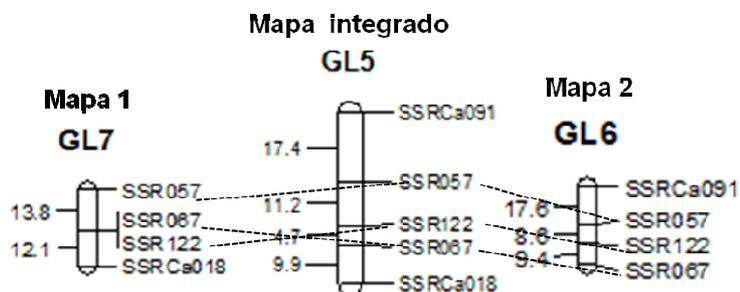


Figura 5 - Grupo de ligação (GL) 5 do mapa integrado, resultante da integração do GL 7 do mapa 1 e GL 6 do mapa 2. Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais.

### Mapa integrado

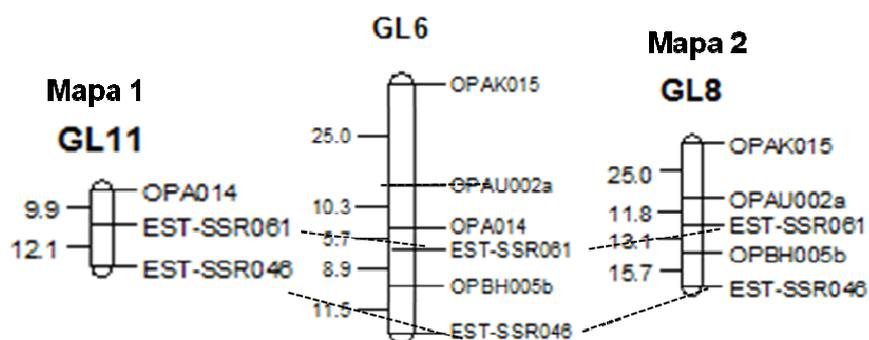


Figura 6 - Grupo de ligação (GL) 6 do mapa integrado, proveniente da integração do GL11 do mapa 1 e do GL 8 do mapa 2. Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais.

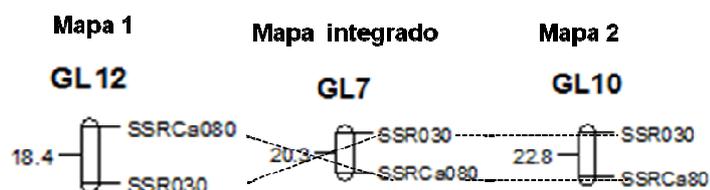


Figura 7 - Grupo de ligação (GL) 7 do mapa integrado, resultante da integração do GL 12 do mapa 1 e GL 10 do mapa 2. Linhas tracejadas indicam a posição dos marcadores âncoras nos mapas parciais.

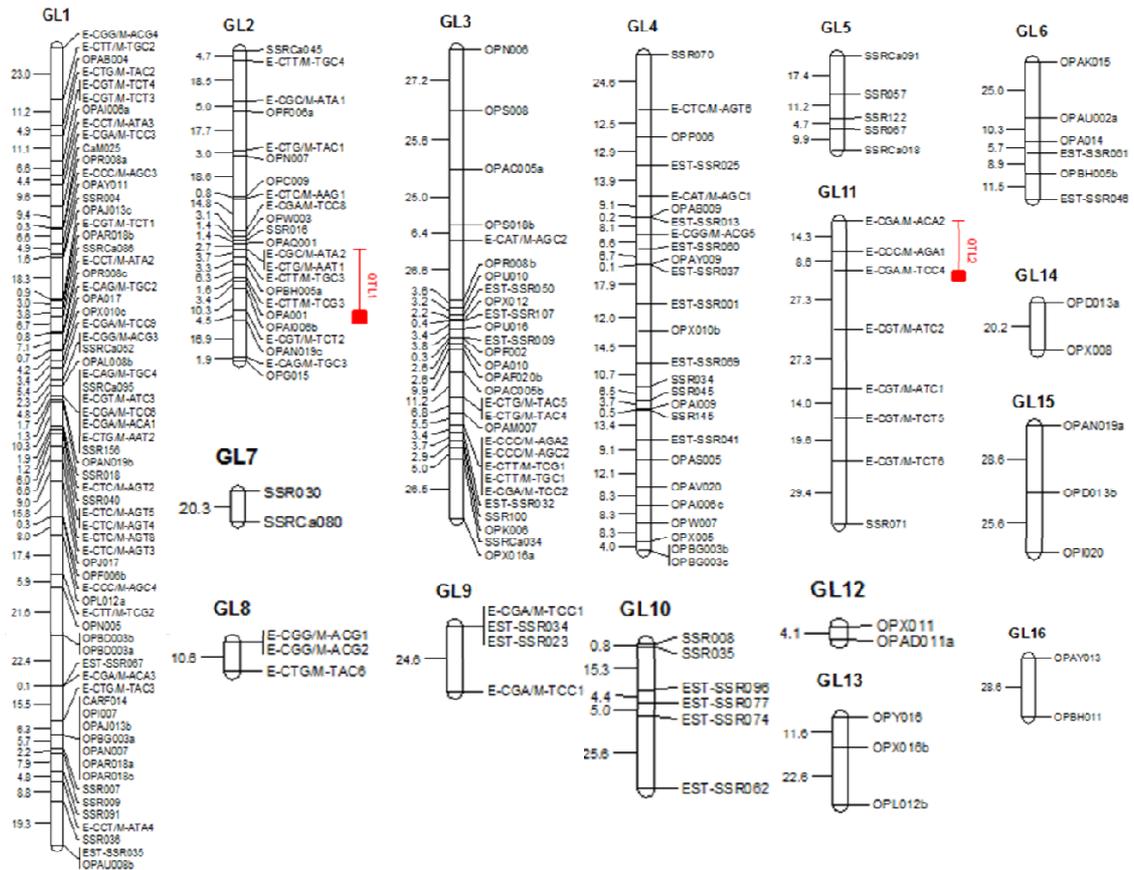


Figura 8 - Demais grupos de ligação dos mapas parciais não resultantes do processo de integração GL 8 a 16 correspondem a GL não integrados. A barra vermelha ao lado dos GLs 2 e 11 indica a posição dos QTLs 1 e 2 identificados para a raça II de *H. vastatrix*, respectivamente.

## 4. DISCUSSÃO

### 4.1. Genotipagem das populações

Do total de 373 marcadores SSR utilizados neste trabalho, apenas 58 foram polimórficos na população 1 e 41 na população 2. O baixo polimorfismo encontrado pode ser justificado pelo fato de as cultivares brasileiras e o material genético de interesse para o melhoramento da espécie *C. arabica* serem bastante aparentados. Setotaw et al. (2013) observaram que apenas sete ancestrais contribuíram com 97,55% da base genética das principais variedades plantadas no Brasil. Diferentes trabalhos com marcadores SSR têm demonstrado esse baixo polimorfismo em *C. arabica*. Vieira et al. (2010), ao analisarem 127 marcadores SSR, encontraram apenas 22 polimórficos. Herrera et al. (2009) e Geleta et al. (2012) identificaram três e oito marcadores polimórficos, ao trabalharem com 13 e 12 SSR, respectivamente.

Devido ao baixo polimorfismo observado, as populações estudadas neste trabalho só foram genotipadas com 58 (população 1) e 41 (população 2) marcadores SSR.

### 4.2. Construção de mapas parciais

O baixo polimorfismo encontrado pelos marcadores SSR inviabilizou a construção de mapas parciais apenas com esse tipo de marcador. Dessa forma, foram incorporados nos mapas genéticos parciais de ligação outros tipos de marcadores moleculares (RAPD, AFLP e um RGA). Os marcadores AFLP e RAPD analisados, apesar de apresentarem também baixo polimorfismo (média de 4,35 bandas polimórficas/*primers* AFLP para a população 1; 1,0; e 1,47 bandas

polimórficas/*primer* RAPD para as populações 1 e 2, respectivamente), contribuíram para aumentar a saturação dos mapas parciais. Esses marcadores foram escolhidos por serem amplamente utilizados em estudos do cafeeiro da espécie *C. arabica*. Maluf et al. (2005) relataram que os AFLPs identificaram maior número de loco, mas menor quantidade de fragmento polimórfico. Em contraste, os marcadores RAPD e SSR produziram menor número de loco, no entanto maior número de alelos polimórficos, em comparação com os AFLP. Para os marcadores AFLP, esse fato também foi observado neste trabalho, tendo em vista que cada combinação de *primer* AFLP detecta em torno de 60 a 70 locos, no entanto apenas dois locos polimórficos foram detectados pela maioria das combinações de AFLP analisadas. Apesar do baixo polimorfismo encontrado nos diferentes trabalhos com cafeeiro, os marcadores AFLP têm sido utilizados para construção de mapas genéticos de ligação (KY et al., 2000; LASHERMES et al., 2001; PEARL et al., 2004).

Para a espécie *C. arabica*, até o momento foram publicados três mapas genéticos de ligação. Pearl et al. (2004) construíram um mapa parcial com 60 indivíduos da população pseudo-F<sub>2</sub> derivada do cruzamento entre as cultivares ‘Tall Mokka’ e ‘Catimor’, genotipadas com 456 marcadores AFLP. O comprimento total do mapa foi de 1.802,8 cM, com uma distância média de 10,2 cM entre marcadores adjacentes. Teixeira-Cabral et al. (2004) construíram um mapa com 104 indivíduos a partir de uma população de retrocruzamento derivada do Híbrido de Timor CIFC 2570 e Mundo Novo IAC 464-18. Neste trabalho foram mapeados 82 marcadores RAPD em oito grupos de ligação, cobrindo uma distância de 560,6 cM do genoma. Outro mapa também obtido a partir de 59 indivíduos da população de retrocruzamento, originado do cruzamento do Híbrido de Timor (UFV 445-46) com Catuaí Amarelo IAC 30, resultou em 11 grupos de ligação, com 117 marcadores RAPD, cobrindo 803,2 cM (OLIVEIRA et al., 2007).

Os mapas parciais de ligação construídos neste trabalho apresentam vantagens em relação aos outros publicados. Uma vantagem é o tamanho da população de mapeamento. O mapa 1 foi construído por meio da análise de 247 indivíduos F<sub>2</sub> e o mapa 2, de 217. Por meio de simulações, Ferreira et al. (2006) concluíram, a partir de populações F<sub>2</sub>, RILs e retrocruzamento, que o tamanho mínimo da população considerado suficiente para a construção de mapas genéticos precisos é de 200 indivíduos. Os mapas obtidos neste trabalho foram os primeiros a serem construídos com mais de 200 indivíduos. Outra vantagem são os tipos de

marcadores utilizados, e os primeiros mapas foram obtidos apenas com marcadores RAPD e AFLP. Nos mapas parciais construídos neste trabalho, foram incorporados marcadores SSR. Esses marcadores, por serem locos específicos, contribuem para a integração de mapas genéticos parciais em um único mapa integrado, permitindo maior cobertura do genoma da espécie. Além dessas vantagens, utilizaram-se dois marcadores específicos RGA (*Disease Resistance Gene Analog*), um para cada população de mapeamento. Os *primers* RGA amplificam regiões de um potencial gene de resistência a doenças (ALVARENGA et al., 2011).

### 4.3. Integração de mapas

A utilização de todos os mapas publicados para *C. arábica* (PEARL et al., 2004; TEIXEIRA-CABRAL et al., 2004; OLIVEIRA et al., 2007) na integração de um único mapa foi impossibilitada pela falta de marcadores comuns entre eles. Com o intuito de obter um mapa mais saturado para essa espécie, foram construídos dois mapas parciais apresentando marcadores comuns (âncoras), entre eles para posterior integração em um único mapa.

O mapa integrado construído neste trabalho será útil como um mapa-referência de *C. arábica*, uma vez que contém número (191 marcadores) maior de marcadores que qualquer outro relatado até o momento. Além disso, o mapa integrado apresenta comprimento total e distância média de 1.423,40 e 8,13 cM, respectivamente. Esse fato demonstra maior densidade de marcador e menor distância entre marcadores do que os mapas parciais obtidos (976,30 e 9,86 cM para a população 1 e 1.158,60 e 13,01 cM para a população 2).

Observando apenas os resultados dos GLs dos mapas parciais utilizados na integração (oito GLs do mapa 1 e sete GLs do mapa 2) e dos sete GLs integrados do mapa final, pode-se constatar aumento na sua resolução. A distância média obtida com esses GLs no mapa 1 foi de 9,14 cM e no mapa 2 de 12,41 cM, enquanto no mapa integrado se observou redução da distância média para 6,98 cM. Essa menor distância entre marcadores demonstra a importância da utilização da integração entre mapas parciais em *C. arábica*, resultando em maior saturação. Além disso, a combinação de vários tipos de marcadores em um único mapa, proveniente de diferentes *backgrounds* genéticos, aumenta as chances de identificar marcadores

polimórficos em determinada região do genoma em diferentes populações de interesse de melhoramento. Dessa forma, a integração facilita a análise de sintenia e seleção assistida por marcadores, bem como a análise genética (análise de QTL) e a clonagem de características importantes para a agricultura baseada em mapas (TRUCO et al., 2007).

Assim, o mapa integrado construído neste trabalho pode servir como ferramenta fundamental para o melhoramento molecular de *C. arabica* e espécies afins, além de fornecer uma base para estudos de evolução e organização do genoma. Além disso, os fragmentos mapeados nos grupos de ligação podem também servir como sondas para trabalhos de hibridização *in situ*, contribuindo para a integração do mapa genético, mapa físico e mapa citogenético. Analisando as Figuras 1 a 7, pode-se observar que a ordem dos marcadores nos GLs dos mapas parciais e o mapa integrado foram bem conservados, demonstrando que o mapa construído pode ser utilizado para essa finalidade.

Os marcadores SSR presentes no mapa integrado podem, por serem locos específicos, ser utilizados para diferentes populações de mapeamento. Embora o mapa integrado possua algumas regiões pouco saturadas e pequenos grupos de ligação, os marcadores SSR facilitarão estudos de sintenia em *C. arabica* com outras espécies de importância econômica.

A construção de mapas genéticos de alta densidade é crucial para a integração de mapas genéticos e físicos, ferramenta eficiente para vincular a variação fenotípica e genotípica (ZHAO et al., 2013). Mapas genéticos integrados com mapas físicos são atualmente bastante utilizados nas pesquisas científicas, com o intuito de obter um mapa saturado como referência para a cultura, além de facilitar a montagem do genoma com rapidez e eficiência, genômica comparativa, clonagem de genes importantes, atribuição dos grupos de ligação aos cromossomos, e para a pesquisa de cromossomos sexuais (STRATIKOPOULOS et al., 2008; YU et al., 2009; FEBRER et al., 2010; XU et al., 2010; HAN et al., 2011; LI et al., 2011; NINWICHIAN et al., 2012; ZHAO et al., 2013). Os mapas físicos, genéticos e citogenéticos integrados com montagens de sequência do genoma inteiro melhoram a precisão e facilitam diretamente o isolamento de genes (FEBRER et al., 2010).

A integração aumenta a possibilidade de encontrar marcadores informativos em alguma parte do genoma e, em último plano, diferenças genéticas inferindo sobre a diversidade genética da espécie. Além disso, mapas genéticos de alta densidade são

importantes para estudos de locos controladores de características quantitativas (QTL) (DIAZ et al., 2011).

#### **4.4. Identificação de QTLs**

Dois QTLs foram identificados associados à raça II de *H. vastatrix*. Os QTLs correspondem aos mesmos QTLs encontrados no mapa de ligação 1 para essa mesma raça do patógeno (dados não mostrados). Apesar de o QTL2 ser identificado em um grupo de ligação não resultante da integração e, sim, proveniente da população 1, os dados fenotípicos foram derivados das duas populações, e isso de fato poderá justificar que o QTL2 identificado seja de grande importância para as duas populações de mapeamento. Os dois QTLs identificados podem corresponder a genes maiores de resistência presentes no Híbrido de Timor. Os marcadores que flanqueiam os QTLs podem ser utilizados para seleção assistida por marcadores moleculares, com o intuito de selecionar indivíduos resistentes em fases iniciais do melhoramento do cafeeiro. Além disso, o posicionamento dos QTLs no mapa integrado irá facilitar análises comparativas entre QTLs de diferentes populações, tendo em vista proporcionarem conhecimentos sobre o controle genético da diversidade fenotípica nas várias cultivares de café arábica.

## 5. CONCLUSÃO

O mapa integrado construído neste trabalho deve ser considerado o mapa de referência para *C. arabica*. Além disso, os marcadores SSR utilizados são potencialmente transferíveis para ampla gama de estudos genéticos, como a integração dos mapas físicos e genéticos, a análise de colinearidade, a clonagem de genes baseados em mapas e a seleção assistida por marcadores moleculares. Os QTLs identificados fornecem informações importantes sobre análise de genes candidatos para resistência/suscetibilidade à raça II de *H. vastatrix*.

## 6. REFERÊNCIAS

ALVARENGA, S. M.; CAIXETA, E. T.; HUFNAGEL, B.; THIEBAUT, F.; MACIEL-ZAMBOLIM, E.; ZAMBOLIM, L.; SAKIYAMA, N. S. Molecular markers from coffee genome expressed sequences potentially involved in resistance to rust. **Pesquisa Agropecuária Brasileira**, v. 46, n. 8, p. 890-898, 2011. ISSN 0100-204X.

BALL, A. D.; STAPLEY, J.; DAWSON, D. A.; BIRKHEAD, T. R.; BURKE, T.; SLATE, J. A comparison of SNPs and microsatellites as linkage mapping markers: lessons from the zebra finch (*Taeniopygia guttata*). **BMC genomics**, v. 11, n. 1, p. 218, 2010. ISSN 1471-2164.

BARUAH, A.; NAIK, V.; HENDRE, P. S.; RAJKUMAR, R.; RAJENDRAKUMAR, P.; AGGARWAL, R. Isolation and characterization of nine microsatellite markers from *Coffea arabica* L., showing wide cross-species amplifications. **Molecular Ecology Notes**, v. 3, n. 4, p. 647-650, 2003. ISSN 1471-8286.

BHAT, P. R.; KRISHNAKUMAR, V.; HENDRE, P. S.; RAJENDRAKUMAR, P.; VARSHNEY, R. K.; AGGARWAL, R. K. Identification and characterization of expressed sequence tags-derived simple sequence repeats markers from robusta coffee variety 'CxR'(an interspecific hybrid of *Coffea canephora* × *Coffea congensis*). **Molecular Ecology Notes**, v. 5, n. 1, p. 80-83, 2005. ISSN 1471-8286.

BRITO, G. G.; CAIXETA, E. T.; GALLINA, A. P.; ZAMBOLIM, E. M.; ZAMBOLIM, L.; DIOLA, V.; LOUREIRO, M. E. Inheritance of coffee leaf rust resistance and identification of AFLP markers linked to the resistance gene. **Euphytica**, v. 173, n. 2, p. 255-264, 2010. ISSN 0014-2336.

CAIXETA, E. T.; PESTANA, K. N. Marcadores moleculares no melhoramento genético do cafeeiro. In: GEAFIP (Ed.). **Melhoramento Genético no Manejo de Doenças de Plantas**. Visconde do Rio Branco-MG, 2013. cap. 6, p.130. ISBN 978-85-8179-044-2.

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. **Pesquisa Agropecuária Brasileira**, v. 44, n. 3, p. 276-282, 2009.

CHISTIAKOV, D. A.; HELLEMANS, B.; VOLCKAERT, F. A. M. Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. **Aquaculture**, v. 255, n. 1, p. 1-29, 2006. ISSN 0044-8486.

COMBES, M.-C.; ANDRZEJEWSKI, S.; ANTHONY, F.; BERTRAND, B.; ROVELLI, P.; GRAZIOSI, G.; LASHERMES, P. Characterization of microsatellite loci in *Coffea arabica* and related coffee species. **Molecular Ecology**, v. 9, n. 8, p. 1178-1180, 2000. ISSN 1365-294X.

COULIBALY, I.; REVOL, B.; NOIROT, M.; PONCET, V.; LORIEUX, M.; CARASCO-LACOMBE, C.; MINIER, J.; DUFOUR, M.; HAMON, P. AFLP and SSR polymorphism in a *Coffea* interspecific backcross progeny [(*C. heterocalyx* × *C. canephora*) × *C. canephora*]. **Theoretical and Applied Genetics**, v. 107, n. 6, p. 1148-1155, 2003. ISSN 0040-5752.

CRUZ, C. D. **Programa GQMOL – Genética quantitativa e molecular**. [www.ufv.br/dbg/gqmol/gqmol.htm](http://www.ufv.br/dbg/gqmol/gqmol.htm), 2008.

CUBRY, P.; MUSOLI, P.; LEGNATÉ, H.; POT, D.; DE BELLIS, F.; PONCET, V.; ANTHONY, F.; DUFOUR, M.; LEROY, T. Diversity in coffee assessed with SSR markers: structure of the genus *Coffea* and perspectives for breeding. **Genome**, v. 51, n. 1, p. 50-63, 2008.

DIAZ, A.; FERGANY, M.; FORMISANO, G.; ZIARSOLO, P.; BLANCA, J.; FEI, Z.; STAUB, J. E.; ZALAPA, J. E.; CUEVAS, H. E.; DACE, G. A consensus linkage map for molecular markers and quantitative trait loci associated with economically important traits in melon (*Cucumis melo* L.). **BMC plant biology**, v. 11, n. 1, p. 111, 2011. ISSN 1471-2229.

DINIZ, L. E. C.; SAKIYAMA, N. S.; LASHERMES, P.; CAIXETA, E. T.; OLIVEIRA, A. C. B.; ZAMBOLIM, E. M.; 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 TECHNOLOGY**, v. 5, n. 4, p. 387, 2005. ISSN 1518-7853.

FALQUE, M.; DÉCOUSSET, L.; DERVINS, D.; JACOB, A.-M.; JOETS, J.; MARTINANT, J.-P.; RAFFOUX, X.; RIBIÈRE, N.; RIDEL, C.; SAMSON, D. Linkage mapping of 1454 new maize candidate gene loci. **Genetics**, v. 170, n. 4, p. 1957-1966, 2005. ISSN 0016-6731.

FEBRER, M.; GOICOECHEA, J. L.; WRIGHT, J.; MCKENZIE, N.; SONG, X.; LIN, J.; COLLURA, K.; WISSOTSKI, M.; YU, Y.; AMMIRAJU, J. S. An integrated physical, genetic and cytogenetic map of *Brachypodium distachyon*, a model system for grass research. **PloS one**, v. 5, n. 10, p. e13461, 2010. ISSN 1932-6203.

FERRÃO, L. F. V.; CAIXETA, E. T.; SOUZA, F. D. F.; ZAMBOLIM, E. M.; CRUZ, C. D.; ZAMBOLIM, L.; SAKIYAMA, N. S. Comparative study of different molecular markers for classifying and establishing genetic relationships in *Coffea canephora*. **Plant Systematics and Evolution**, v. 299, n. 1, p. 225-238, 2013. ISSN 0378-2697.

FERREIRA, A.; SILVA, M. F.; SILVA, L. D. C. E.; CRUZ, C. D. Estimating the effects of population size and type on the accuracy of genetic maps. **Genet Mol Biol**, v. 29, n. 1, p. 187-192, 2006.

GEBHARDT, C.; SCHMIDT, R.; SCHNEIDER, K. Plant genome analysis: the state of the art. **International review of cytology**, v. 247, p. 223-284, 2005. ISSN 0074-7696.

GELETA, M.; HERRERA, I.; MONZÓN, A.; BRYNGELSSON, T. Genetic diversity of Arabica coffee (*Coffea arabica* L.) in Nicaragua as estimated by simple sequence repeat markers. **The Scientific World Journal**, v. 2012, 2012.

GONZALO, M. J.; OLIVER, M.; GARCIA-MAS, J.; MONFORT, A.; DOLCET-SANJUAN, R.; KATZIR, N.; ARUS, P.; MONFORTE, A. J. Simple-sequence repeat markers used in merging linkage maps of melon (*Cucumis melo* L.). **Theoretical and Applied Genetics**, v. 110, n. 5, p. 802-811, 2005. ISSN 0040-5752.

GRIFFITHS, A. J. F. **Introdução à genética**. Rio de Janeiro, Guanabara Koogan: 2008. ISBN 8527714973.

GUPTA, P. K.; RUSTGI, S.; KULWAL, P. L. Linkage disequilibrium and association studies in higher plants: present status and future prospects. **Plant molecular biology**, v. 57, n. 4, p. 461-485, 2005. ISSN 0167-4412.

HAN, Y.; ZHANG, Z.; HUANG, S.; JIN, W. An integrated molecular cytogenetic map of *Cucumis sativus* L. chromosome 2. **BMC genetics**, v. 12, n. 1, p. 18, 2011. ISSN 1471-2156.

HENDRE, P. S.; PHANINDRANATH, R.; ANNAPURNA, V.; LALREMURATA, A.; AGGARWAL, R. K. Development of new genomic microsatellite markers from robusta coffee (*Coffea canephora* Pierre ex A. Froehner) showing broad cross-species transferability and utility in genetic studies. **BMC plant biology**, v. 8, n. 1, p. 51, 2008. ISSN 1471-2229.

HERRERA, J. C.; ALVARADO A, G.; CORTINA G, H. A.; COMBES, M.-C.; ROMERO G, G.; LASHERMES, P. Genetic analysis of partial resistance to coffee leaf rust (*Hemileia vastatrix* Berk & Br.) introgressed into the cultivated *Coffea arabica* L. from the diploid *C. canephora* species. **Euphytica**, v. 167, n. 1, p. 57-67, 2009. ISSN 0014-2336.

ISOBE, S.; KÖLLIKER, R.; HISANO, H.; SASAMOTO, S.; WADA, T.; KLIMENKO, I.; OKUMURA, K.; TABATA, S. Construction of a consensus linkage

map for red clover (*Trifolium pratense* L.). **BMC plant biology**, v. 9, n. 1, p. 57, 2009. ISSN 1471-2229.

KY, C.-L.; BARRE, P.; LORIEUX, M.; TROUSLOT, P.; AKAFFOU, S.; LOUARN, J.; CHARRIER, A.; HAMON, S.; NOIROT, M. Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp.). **Theoretical and Applied Genetics**, v. 101, n. 4, p. 669-676, 2000. ISSN 0040-5752.

LANDER, E. S.; BOTSTEIN, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, v. 121, n. 1, p. 185-199, 1989. ISSN 0016-6731.

LASHERMES, P.; COMBES, M.-C.; PRAKASH, N. S.; TROUSLOT, P.; LORIEUX, M.; CHARRIER, A. Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. **Genome**, v. 44, n. 4, p. 589-595, 2001. ISSN 0831-2796.

LEE, H.-R.; BAE, I.-H.; PARK, S.-W.; KIM, H.-J.; MIN, W.-K.; HAN, J.-H.; KIM, K.-T.; KIM, B.-D. Construction of an integrated pepper map using RFLP, SSR, CAPS, AFLP, WRKY, rRAMP, and BAC end sequences. **Molecules and cells**, v. 27, n. 1, p. 21-37, 2009. ISSN 1016-8478.

LI, W.; ZHANG, J.; MOU, Y.; GENG, J.; MCVETTY, P.; HU, S.; LI, G. Integration of Solexa sequences on an ultradense genetic map in *Brassica rapa* L. **BMC genomics**, v. 12, n. 1, p. 249, 2011. ISSN 1471-2164.

MALUF, M. P.; SILVESTRINI, M.; RUGGIERO, L. M. D. C.; GUERREIRO FILHO, O.; COLOMBO, C. A. Genetic diversity of cultivated *Coffea arabica* inbred lines assessed by RAPD, AFLP and SSR marker systems. **Scientia Agricola**, v. 62, n. 4, p. 366-373, 2005. ISSN 0103-9016.

MISSIO, R. F.; CAIXETA, E. T.; ZAMBOLIM, E. M.; PENA, G. F.; RIBEIRO, A. P.; ZAMBOLIM, L.; PEREIRA, A. A.; SAKIYAMA, N. S. Assessment of EST-SSR markers for genetic analysis on coffee. **Bragantia**, v. 68, n. 3, p. 573-581, 2009a. ISSN 0006-8705.

MISSIO, R. F.; CAIXETA, E. T.; ZAMBOLIM, E. M.; PENA, G. F.; ZAMBOLIM, L.; DIAS, L. A. S.; SAKIYAMA, N. S. Genetic characterization of an elite coffee germplasm assessed by gSSR and EST-SSR markers. **Genet Mol Res**, v. 10, p. 2366-2381, 2011.

MISSIO, R. F.; CAIXETA, E. T.; ZAMBOLIM, E. M.; ZAMBOLIM, L.; SAKIYAMA, N. S. Development and validation of SSR markers for *Coffea arabica* L. **Crop Breeding and Applied Biotechnology**, v. 9, n. 4, p. 361-371, 2009b. ISSN 1518-7853.

MONCADA, P.; MCCOUCH, S. Simple sequence repeat diversity in diploid and tetraploid *Coffea* species. **Genome**, v. 47, n. 3, p. 501-509, 2004. ISSN 0831-2796.

NINWICHIAN, P.; PEATMAN, E.; LIU, H.; KUCUKTAS, H.; SOMRIDHIVEJ, B.; LIU, S.; LI, P.; JIANG, Y.; SHA, Z.; KALTENBOECK, L. Second-generation genetic linkage map of catfish and its integration with the BAC-based physical map. **G3: Genes| Genomes| Genetics**, v. 2, n. 10, p. 1233-1241, 2012. ISSN 2160-1836.

NOVELLI, V. M.; CRISTOFANI-YALY, M.; BASTIANEL, M.; PALMIERI, D. A.; MACHADO, M. A. Screening of Genomic Libraries. In: (Ed.). **Microsatellites**: Springer, 2013. p.17-24. ISBN 1627033882.

OLIVEIRA, A. C. B.; SAKIYAMA, N. S.; CAIXETA, E. T.; ZAMBOLIM, E. M.; RUFINO, R. J. N.; ZAMBOLIM, L. Partial map of *Coffea arabica* L. and recovery of the recurrent parent in backcross progenies. **Crop Breeding and Applied Biotechnology**, v. 7, n. 2, p. 196-203, 2007. ISSN 1518-7853.

PAILLARD, M.; LASHERMES, P.; PÉTIARD, V. Construction of a molecular linkage map in coffee. **Theoretical and Applied Genetics**, v. 93, n. 1-2, p. 41-47, 1996. ISSN 0040-5752.

PEARL, H. M.; NAGAI, C.; MOORE, P. H.; STEIGER, D. L.; OSGOOD, R. V.; MING, R. Construction of a genetic map for arabica coffee. **Theoretical and Applied Genetics**, v. 108, n. 5, p. 829-835, 2004. ISSN 0040-5752.

PONCET, V.; HAMON, P.; MINIER, J.; CARASCO, C.; HAMON, S.; NOIROT, M. SSR cross-amplification and variation within coffee trees (*Coffea* spp.). **Genome**, v. 47, n. 6, p. 1071-1081, 2004. ISSN 0831-2796.

RAZAFINARIVO, N. J.; GUYOT, R.; DAVIS, A. P.; COUTURON, E.; HAMON, S.; CROUZILLAT, D.; RIGOREAU, M.; DUBREUIL-TRANCHANT, C.; PONCET, V.; DE KOCHKO, A. Genetic structure and diversity of coffee (*Coffea*) across Africa and the Indian Ocean islands revealed using microsatellites. **Annals of botany**, v. 111, n. 2, p. 229-248, 2013. ISSN 0305-7364.

ROVELLI, P.; METTULIO, R.; ANTHONY, F.; ANZUETO, F.; LASHERMES, P.; GRAZIOSI, G. Microsatellites in *Coffea Arabica* L. In: (Ed.). **Coffee biotechnology and quality**: Springer, 2000. p.123-133. ISBN 9048155657.

SALGADO, C. C.; CRUZ, C. D.; NASCIMENTO, M.; BARRERA, C. F. S. O uso da variância como metodologia alternativa para integração de mapas genéticos. **Pesquisa Agropecuária Brasileira**, v. 46, n. 1, p. 66-73, 2011. ISSN 1678-3921.

SETOTAW, T. A.; CAIXETA, E. T.; PEREIRA, A. A.; OLIVEIRA, A. C. B.; CRUZ, C. D.; ZAMBOLIM, E. M.; ZAMBOLIM, L.; SUSSUMU, N. **Coefficient of parentage in *Coffea arabica* L. cultivars grown in Brazil**. 2013.

STRATIKOPOULOS, E. E.; AUGUSTINOS, A. A.; PETALAS, Y. G.; VRAHATIS, M. N.; MINTZAS, A.; MATHIOPOULOS, K. D.; ZACHAROPOULOU, A. An integrated genetic and cytogenetic map for the Mediterranean fruit fly, *Ceratitis capitata*, based on microsatellite and morphological markers. **Genetica**, v. 133, n. 2, p. 147-157, 2008. ISSN 0016-6707.

STUDER, B.; KÖLLIKER, R.; MUYPLE, H.; ASP, T.; FREI, U.; ROLDÁN-RUIZ, I.; BARRE, P.; TOMASZEWSKI, C.; MEALLY, H.; BARTH, S. EST-derived SSR markers used as anchor loci for the construction of a consensus linkage map in ryegrass (*Lolium* spp.). **BMC plant biology**, v. 10, n. 1, p. 177, 2010. ISSN 1471-2229.

TAMAYO, P. J.; VALE, F. X. R.; ZAMBOLIM, L.; CHAVES, G. M.; PEREIRA, A. A. Resistência do Catimor à ferrugem e virulência de raças fisiológicas de *Hemileia vastatrix* Berk. & Br. **Fitopatologia Brasileira**, v. 20, p. 572-576, 1995.

TEIXEIRA-CABRAL, T. A.; SAKIYAMA, N. S.; ZAMBOLIM, L.; PEREIRA, A. A.; SCHUSTER, I. Single-locus inheritance and partial linkage map of *Coffea arabica* L. **Crop Breeding and Applied Biotechnology**, v. 4, p. 416-421, 2004.

TRUCO, M. J.; ANTONISE, R.; LAVELLE, D.; OCHOA, O.; KOZIK, A.; WITSENBOER, H.; FORT, S. B.; JEUKEN, M. J. W.; KESSELI, R. V.; LINDHOUT, P. A high-density, integrated genetic linkage map of lettuce (*Lactuca* spp.). **Theoretical and Applied Genetics**, v. 115, n. 6, p. 735-746, 2007. ISSN 0040-5752.

VEZZULLI, S.; TROGGIO, M.; COPPOLA, G.; JERMAKOW, A.; CARTWRIGHT, D.; ZHARKIKH, A.; STEFANINI, M.; GRANDO, M. S.; VIOLA, R.; ADAM-BLONDON, A.-F. A reference integrated map for cultivated grapevine (*Vitis vinifera* L.) from three crosses, based on 283 SSR and 501 SNP-based markers. **Theoretical and Applied Genetics**, v. 117, n. 4, p. 499-511, 2008. ISSN 0040-5752.

VIEIRA, E. S. N.; VON PINHO, É. V. R.; CARVALHO, M. G. G.; ESSELINK, D. G.; VOSMAN, B. Development of microsatellite markers for identifying Brazilian *Coffea arabica* varieties. **Genetics and molecular biology**, v. 33, n. 3, p. 507-514, 2010. ISSN 1415-4757.

VOORRIPS, R. E. MapChart: software for the graphical presentation of linkage maps and QTLs. **Journal of Heredity**, v. 93, n. 1, p. 77-78, 2002. ISSN 0022-1503.

XU, J.; QIAN, X.; WANG, X.; LI, R.; CHENG, X.; YANG, Y.; FU, J.; ZHANG, S.; KING, G.; WU, J. Construction of an integrated genetic linkage map for the A genome of *Brassica napus* using SSR markers derived from sequenced BACs in *B. rapa*. **BMC genomics**, v. 11, n. 1, p. 594, 2010. ISSN 1471-2164.

YU, Q.; TONG, E.; SKELTON, R. L.; BOWERS, J. E.; JONES, M. R.; MURRAY, J. E.; HOU, S.; GUAN, P.; ACOB, R. A.; LUO, M.-C. A physical map of the papaya genome with integrated genetic map and genome sequence. **BMC genomics**, v. 10, n. 1, p. 371, 2009. ISSN 1471-2164.

ZHAO, L.; ZHANG, Y.; JI, P.; ZHANG, X.; ZHAO, Z.; HOU, G.; HUO, L.; LIU, G.; LI, C.; XU, P. A Dense Genetic Linkage Map for Common Carp and Its Integration with a BAC-Based Physical Map. **PLoS one**, v. 8, n. 5, p. e63928, 2013. ISSN 1932-6203.

## **CAPITULO 3**

### **SELEÇÃO GENÔMICA AMPLA EM *Coffea arabica***

## RESUMO

O desenvolvimento de diferentes tipos de marcadores moleculares tornou disponíveis várias ferramentas para selecionar indivíduos superiores no melhoramento genético de plantas. Uma das ferramentas que podem auxiliar os melhoristas é a seleção genômica ampla (GWS). Essa técnica utiliza informação direta do DNA na seleção, o que permite alta eficiência seletiva e rapidez na obtenção de ganhos genéticos em relação à tradicional seleção baseada apenas em dados fenotípicos. Em vista disso, os objetivos deste trabalho foram verificar a capacidade preditiva da GWS em prever os valores genômicos dos cafeeiros para resistência à ferrugem, bem como comparar os marcadores identificados pela GWS com os identificados associados à resistência a três isolados de *H. vastatrix* (raças I e II e patótipo 001), por meio do mapa genético de ligação de *Coffea arabica*. Para isso, foram utilizados 245 indivíduos F<sub>2</sub> resultantes do cruzamento da cv. suscetível Catuaí Amarelo IAC 64 (UFV 2148-57) e da fonte de resistência ao Híbrido de Timor UFV 443-03, genotipados com 137 (74 AFLP, 58 SSR, 4 RAPD e 1 *primer*-específico) marcadores moleculares. As estimativas dos parâmetros genéticos foram obtidas via BLASSO. A capacidade preditiva (CP) foi medida utilizando todos os indivíduos da população como população de treinamento e validação, como também dividindo aleatoriamente a população em 200 indivíduos para população de treinamento e 45 indivíduos para população de validação. Para o primeiro caso, a maior CP preditiva foi para o patótipo 001 (0,78). No segundo caso, a maior CP para população de validação foi para a raça I (0,49). Quando se utilizaram apenas os marcadores de maiores efeitos para cada isolado, a maior CP na população de validação foi para o patótipo 001 (0,49). Ainda foi possível observar que as herdabilidades estimadas a partir de

marcadores moleculares para a resistência aos três isolados foram inferiores às herdabilidades baseadas em valores fenotípicos. Isso demonstra maior predição dos valores genômicos dos indivíduos na seleção genômica em comparação com a seleção fenotípica. A maioria dos marcadores identificados associados aos QTLs no mapa genético de ligação foi identificada também pela GWS, confirmando a acurácia da metodologia utilizada. Além disso, a GWS permitiu prever, com sucesso, os valores genômicos dos indivíduos.

## 1.INTRODUÇÃO

Ao longo dos anos, vários métodos de melhoramento surgiram com o objetivo de otimizar o processo de seleção de indivíduos que apresentam caracteres de interesse agrônomico. Contudo, a estratégia básica utilizada até hoje é a de prever o valor genético do indivíduo, baseado em informações fenotípicas. No entanto, com o desenvolvimento de marcadores moleculares e o avanço em técnicas de biologia molecular, surgiu a expectativa de que informações genotípicas, uma vez correlacionadas com características fenotípicas de interesse, possam ser amplamente utilizadas na identificação e seleção de indivíduos, com maiores valores genéticos (RESENDE JR. et al., 2013)

Uma técnica proposta para acelerar o melhoramento genético a partir de marcadores moleculares é a seleção genômica (GS), ou seleção genômica ampla (*Genome Wide Selection* – GWS), proposta por Meuwissen et al. (2001). Essa técnica utiliza informação direta do DNA na seleção e não exige conhecimento prévio das posições (mapa) dos QTLs (RESENDE, 2008), como no caso da seleção assistida por marcadores (SAM). Nos métodos tradicionais de seleção, o fenótipo é utilizado para inferir sobre o efeito do genótipo. Na GWS, o genótipo é utilizado para prever o fenótipo a ser expresso nos candidatos à seleção (POLAND et al., 2012; RESENDE et al., 2012).

O modelo de predição do fenótipo via GWS é estabelecido, preliminarmente, com base em dados genotípicos e fenotípicos de indivíduos pertencentes a uma amostra da população de seleção, conhecida como população de treinamento. Essa população é utilizada para estimar os parâmetros do modelo que será posteriormente utilizado para calcular os valores genômicos estimados dos indivíduos candidatos à

seleção, possuindo apenas os dados genotípicos. A população deve ser representativa dos candidatos à seleção (HEFFNER et al., 2009). A predição e a seleção podem ser realizadas na fase inicial do melhoramento de plantas, acelerando, assim, o processo de melhoramento genético (RESENDE, 2007).

Para a GWS, vários métodos têm sido propostos, que diferem entre si pelo tipo de suposição sobre o modelo genético associado ao caráter quantitativo. Entre eles, pode-se citar o BLASSO (bayesian LASSO) (LEGARRA et al., 2011).

O método BLASSO, de Park e Casella (2008), ajusta uma variância separada para cada marcador e força os estimadores para zero, como no caso do RR-BLUP. No entanto, efetivamente permite que alguns estimadores sejam identicamente iguais a zero, realizando simultaneamente o procedimento de *shrinkage* e seleção de covariáveis (De Los CAMPOS et al., 2009). Segundo Resende et al. (2010), o LASSO é uma boa opção quando se usa número limitado de marcadores, contudo o melhor método é aquele que reflete melhor a natureza biológica do caráter poligênico em questão, em termos de efeitos gênicos.

A principal limitação da aplicação da seleção genômica em *C. arabica* tem sido o grande número de marcadores necessários. Populações de melhoramento de *C. arabica* apresentam baixa taxa de polimorfismo, quando analisadas com marcadores moleculares (HERRERA et al., 2009; VIEIRA et al., 2010; GELETA et al., 2012), o que dificulta avanços no melhoramento molecular dessa espécie.

Resende et al. (2010) relataram que as capacidades em prever os valores genéticos dos indivíduos diminuem com o aumento do número de marcadores, além de um número ótimo. Isso foi demonstrado por Fritsche-Neto et al. (2012) e Cavalcanti et al. (2012), os quais provaram que o incremento no número de marcadores não aumenta linearmente a capacidade preditiva da GWS.

Outro aspecto a ser considerado é o tipo de marcador molecular a ser utilizado, uma vez que a GWS requer grande número de marcadores genéticos dispersos em todo o genoma de um organismo. O desenvolvimento de marcadores moleculares tipo SNPs (*Single Nucleotide Polymorphisms*) permitiu a automação do processo de genotipagem, que passou a ser realizado em larga escala. Apesar da utilidade de marcadores SNP para a GWS, sua utilização requer duas a três vezes maior densidade de marcadores do que SSR. Isso se deve à natureza bialélica dos SNPs e multialélica dos SSR (SOLBERG et al., 2008). Os marcadores moleculares do tipo SSR podem também ser usados na GWS, por serem codominantes,

multialélicos e apresentarem alta transferabilidade entre indivíduos e espécies (RESENDE, 2008).

Marcadores SSR foram utilizados em milho por Fritsche-Neto et al. (2012) para estudos de seleção genômica ampla. Além de marcadores SSR, Cavalcanti et al. (2012) utilizaram marcadores AFLP e ISSR para seleção genômica em cajueiro. Marcadores dominantes tipo DArT foram utilizados em eucalipto por Resende et al. (2012). De acordo com Cavalcanti et al. (2012), todos esses marcadores podem ser usados na GWS, desde que o modelo biométrico seja adequadamente estabelecido. Em vista disso, utilizaram-se neste trabalho marcadores moleculares do tipo AFLP, SSR, RAPD e RGA.

Assim, o objetivo deste trabalho foi avaliar a capacidade preditiva da seleção genômica ampla (GWS) em prever os valores genômicos dos indivíduos (GEBVs), bem como comparar os marcadores moleculares de maior efeito genômico via GWS com os marcadores identificados associados à QTLs para as raças I e II e o patótipo 001 de *H. vastatrix*, em mapa genético de ligação.

## **2. MATERIAL E MÉTODOS**

### **2.1. Material vegetal**

Para este estudo, utilizou-se uma população F<sub>2</sub>, constituída por 245 indivíduos oriundos da autofecundação controlada do híbrido F<sub>1</sub> H511-1. Este híbrido originou-se do cruzamento da cv. suscetível Catuaí Amarelo IAC 64 (UFV 2148-57) e com o progenitor resistente Híbrido de Timor (UFV 443-03).

### **2.2. Extração de DNA**

A extração do DNA dos progenitores e dos indivíduos da população F<sub>2</sub> foi realizada de acordo com o protocolo descrito por Diniz et al. (2005).

### **2.3. Genotipagem da população**

A genotipagem foi efetuada com 137 (74 AFLP, 58 SSR, 4 RAPD e 1 *primer*-específico) marcadores moleculares nos progenitores e nos 245 indivíduos da população F<sub>2</sub>.

### 2.3.1. Marcadores SSR

Inicialmente, foram efetuados testes de polimorfismo dos *primers* SSR utilizando o DNA dos genitores e de 10 indivíduos da população F<sub>2</sub>. Foram testados 373 pares de *primers* SSR de cafeeiro (COMBES et al., 2000; ROVELLI et al., 2000; BARUAH et al., 2003; COULIBALY et al., 2003; MONCADA; MCCOUCH, 2004; PONCET et al., 2004; BHAT et al., 2005; HENDRE et al., 2008; MISSIO et al., 2009b; MISSIO et al., 2011). As reações de PCR foram realizadas de acordo com as metodologias descritas por Missio et al. (2009a). A detecção do polimorfismo dos marcadores foi realizada em gel de poliacrilamida desnaturante 6%, corado com nitrato de prata, conforme protocolo descrito por Brito et al. (2010) e em sequenciador automático *ABI 3130 xl Genetic Analyzer (Applied Biosystems)*. Para análise no sequenciador, um dos *primers* do par que amplifica cada loco microssatélite, o *forward*, foi marcado com um fluoróforo específico (6-FAM, azul, ou NED, amarelo, ou HEX, verde). As reações de PCR para cada *primer* foram efetuadas separadamente, e a reação de três fluoróforos específicos foi misturada para análise, caracterizando um “falso multiplex”. Cada reação de PCR foi diluída na proporção de 1:10, e 1 µL de cada reação diluída foi transferido a uma placa específica para o sequenciador, contendo 6,8 µL de formamida HI-DI Applied Biosystems e 0,2 µL do Rox- 500 (padrão de peso molecular marcado com fluoróforo RED). A reação foi desnaturada em termociclador a 95 °C, por 3 min. Em seguida, as amostras foram eletroinjetadas no sequenciador, utilizando-se capilares de sílica de 50 cm de comprimento e 50 µm de diâmetro preenchidos com polímero POP 7, conforme recomendação do fabricante. Os fragmentos foram analisados por meio do *software GeneMapper V4.1 Applied Biosystems*.

### 2.3.2. Marcadores AFLP

A técnica de AFLP foi realizada conforme descrito por Brito et al. (2010), com algumas modificações. Os DNAs dos progenitores (Híbrido de Timor UFV 443-03 e Catuaí Amarelo UFV 2148-57) e dos indivíduos F<sub>2</sub> resultantes desse cruzamento foram digeridos com as enzimas de restrição *EcoRI* e *MseI* por 8 h, a 37 °C. Os fragmentos gerados foram ligados aos adaptadores específicos, utilizando-se a

enzima *T4 DNA* ligase por 14 h, a 4 °C. O material digerido e ligado foi diluído em 1:10 em TE pH 8,0 e armazenado a -20 °C. A amplificação pré-seletiva foi realizada com *primers* complementares aos adaptadores, contendo um nucleotídeo adicional (N) na extremidade 3'. O produto pré-amplificado foi diluído em 1:40 em TE pH 8,0 e armazenado a -20 °C. Para a reação de amplificação, foram utilizadas 17 combinações de *primers* seletivos resultantes de *EcoRI* e *MseI* (E-CAG/M-TGC; E-CAT/M-AGT; E-CCT/M-ATA; E-CCC/M-AGC; E-CCC/MAGA; E-CGA/M-ACA; E-CGG/M-ACG; E-CGT/M-TCT; E-CGA/M-TCC; E-CGC/M-ATA; E-CGT/M-ATC; E-CTG/M-AAT; E-CTG/M-TAC; E-CTT/M-TGC; E-CTC/M-AGT; E-CTC/M-AAG; e E-CTT/MTCG). Aos produtos amplificados foram adicionados 8 µL de formamida (formamida 98%, EDTA pH 8,0, 10 mM, *xylene cyanol* 1 mg/mL e *bromophenol blue* 1 mg/mL). O polimorfismo dos marcadores foi observado em gel de poliacrilamida desnaturante (6%), corado com nitrato de prata.

### 2.3.3. Marcadores RAPD

Conforme a metodologia descrita por Oliveira et al. (2007), foram utilizados quatro *primers* RAPD da Operon Technologies (OPA 001, OPA 014, OPA 017 e OPAB 009). Cada reação de amplificação dos *primers* RAPD consistiu de um volume total de 25 µl, contendo os seguintes componentes: 25 ng de DNA genômico, 1 U de Taq DNA polimerase, 0,1 mM de cada dNTP, 0,2 µM de *primer*, 50 mM de KCl, 10 mM de Tris HCl pH 8,3 e 2 mM de MgCl<sub>2</sub> e o volume final completado com água milli-Q estéril.

### 2.3.4. Marcador primer-específico

O *primer* específico CARF 005 foi utilizado de acordo com a metodologia descrita por Alvarenga et al. (2011). Esse *primer* amplifica uma região RGA (*Disease Resistance Gene Analog*) de 400 pb de um potencial gene de resistência a doenças. O CARF 005 foi desenvolvido a partir de sequência minerada do Projeto Brasileiro do Genoma Café e corresponde a uma porção do gene de resistência com

anotação “gi|24459841| emb|CAC82597.1| *disease resistance-like protein*” de *C. arabica*.

#### **2.4. Codificação dos marcadores**

Dados de marcadores dominantes em aproximação (alelo proveniente do progenitor resistente Híbrido de Timor UFV443-03) foram codificados -1 para a presença e 1 para a ausência da banda. Nos marcadores dominantes em repulsão (alelo proveniente do progenitor suscetível Catuaí Amarelo UFV 2148-57), as codificações foram 1 e -1 para a presença e ausência da banda, respectivamente. Os marcadores codominantes foram codificados como 0 para o heterozigoto, -1 para bandas provenientes do progenitor resistente e 1 para bandas provenientes do progenitor suscetível.

#### **2.5. Seleção genômica ampla (*Genome Wide Selection - GWS*) – Lasso bayesiano (BLASSO)**

As análises de GWS para prever os valores genômicos dos indivíduos foram efetuadas primeiramente com todos os 245 indivíduos da população  $F_2$  para população de treinamento e validação. Posteriormente, os indivíduos foram sorteados, dividindo-se a população em duas, ou seja, do total de 245 indivíduos  $F_2$ , 200 (77,50%) foram utilizados para população de treinamento e 45 (22,5%) para população de validação.

Para população de treinamento (utilizando todos os indivíduos ou dividindo a população), foram consideradas: a) a genotipagem de todos os marcadores e a fenotipagem para resistência às raças I e II e ao patótipo 001; b) a genotipagem apenas dos marcadores de maior efeito e a fenotipagem dos três isolados de *H. vastatrix*.

Na população de validação, as equações de predição dos valores genômicos dos indivíduos foram testadas primeiramente com todos os marcadores e, posteriormente, com os marcadores de maior efeito, para verificar a capacidade preditiva da GWS.

A capacidade preditiva é dada pela fórmula  $r_{y\hat{y}}^2 = r_{g\hat{g}}^2 h$ , em que  $r_{g\hat{g}}$  é a acurácia de seleção e  $h$  é a raiz quadrada da herdabilidade individual (RESENDE JR. et al., 2012). Como na validação dos resultados, o valor fenotípico é conhecido e a capacidade preditiva da seleção genômica foi avaliada ao calcular a correlação do valor genético predito com o fenótipo observado nos indivíduos.

As análises foram efetuadas no *software* R (TEAM, 2012), por meio de *scripts* disponíveis no *software* GENES (CRUZ, 2013).

Para predição dos efeitos dos marcadores, foi utilizada a implementação bayesiana do método LASSO (*Least Absolute Shrinkage and Selection Operator*), proposto por De Los Campos et al. (2009). Conforme descrito por Almeida et al. (2012), a implementação impõe um produto de densidades exponenciais duplas como distribuição *a priori* dos "p" coeficientes de regressão, de forma que  $p(\beta|\lambda) = \prod_{j=1}^p \frac{\lambda}{2} \exp(-\lambda|\beta_j|)$ . Na distribuição *a priori* conjunta construída, a densidade atribuída aos coeficientes de regressão regularizados por LASSO será  $\prod_{j=1}^p N(\beta_{j1}|0, \sigma_e^2 \tau_j^2)$ , resultando em variâncias específicas para cada coeficiente de regressão. Por sua vez, a distribuição *a priori* para o parâmetro de escala  $\tau_j^2$  será representada por:  $\prod_{j=1}^p \exp(\tau_j^2|\lambda)$ , no qual o parâmetro de suavização  $\lambda$  influencia o ajuste dos coeficientes de regressão. Para estimação dos parâmetros, foi considerado o método MCMC, gerando-se  $10^6$  simulações com *burn-in* de  $10^5$ , tendo sido os parâmetros amostrados a cada cinco observações. Os valores genômicos foram estimados conforme De Los Campos et al. (2009), e a taxa de erro aparente na predição dos fenótipos dos indivíduos utilizados para validação foi calculada, considerando-se notas e a classificação binária resistente e suscetível, de acordo com a escala de notas.

## 2.6. Análise fenotípica e mapeamento de QTLs associados a *H. vastatrix*

Os uredósporos das raças I e II e do patótipo 001 de *H. vastatrix* foram multiplicados e inoculados de acordo com a metodologia descrita por Capucho et al. (2009). A inoculação foi efetuada nos progenitores, no híbrido F<sub>1</sub> e em todos os 245 indivíduos da população F<sub>2</sub>. A avaliação da resistência/suscetibilidade foi efetuada de acordo com Capucho et al. (2009), seguindo-se a escala de Tamayo et al. (1995), que

se baseia na ausência ou presença de uredósporos. De acordo com essa escala, são consideradas resistentes (notas 1 a 3); a ausência de sintomas (nota 1), as lesões cloróticas pequenas (nota 2); e as lesões cloróticas grandes, sem esporulação (nota 3). A suscetibilidade (notas 4 a 6) é atribuída a lesões cloróticas grandes com poucos uredósporos, ocupando menos de 25% da área foliar (nota 4), lesões com esporulação ocupando de 25 a 50% da área (nota 5) e lesões com esporulação ocupando mais de 50% da área com uredósporos (nota 6).

### 3. RESULTADOS E DISCUSSÃO

#### 3.1. Genotipagem dos marcadores

Inicialmente, foram analisados 373 pares de *primers* SSR nos progenitores e em 10 indivíduos F<sub>2</sub> da população em estudo. Desses, 58 foram polimórficos (15,55%), 271 monomórficos (72,65%) e 44 não amplificaram (11,80%). A não amplificação de alguns marcadores pode ser explicada pela falta de homologia de alguns *primers* no genoma de *C. arabica*, por terem sido construídos para *C. canephora*. Os 58 marcadores SSR polimórficos foram analisados em toda a população F<sub>2</sub>. O baixo polimorfismo encontrado pode ser justificado pelo fato de as cultivares brasileiras da espécie *C. arabica* serem bastante aparentadas. Setotaw et al. (2013) observaram que apenas sete ancestrais contribuíram com 97,55% da base genética das principais variedades plantadas no Brasil. Esse baixo polimorfismo tem sido identificado em outros trabalhos e justificado pela base genética estreita apresentada pela espécie *C. arabica*, tanto no Brasil quanto em outros países (HERRERA et al., 2009; VIEIRA et al., 2010; GELETA et al., 2012).

Na genotipagem com marcadores AFLP, mais de um loco polimórfico foi verificado em cada combinação de *primer*. Assim, das 17 combinações AFLP utilizadas foram obtidas 74 bandas polimórficas, com média de 4,35 bandas polimórficas/*primer*. Esse baixo polimorfismo com marcadores AFLP também foi observado por Brito et al. (2010), Ky et al. (2000) e Pearl et al. (2004), autores esses que identificaram 1,10; 6,5; e 1,34 bandas polimórficas/*primer*, respectivamente.

### 3.2. Capacidade preditiva da seleção genômica ampla

A capacidade preditiva (CP) em prever os valores genômicos dos indivíduos na seleção com todos os 137 marcadores e todos os 245 indivíduos da população F<sub>2</sub> como população de treinamento e validação é apresentada na Tabela 1. Observa-se, nessa tabela, que a maior CP (0,78) foi para o patótipo 001. Esse resultado pode ser justificado pelo fato de esse isolado apresentar maior herdabilidade estimada por marcadores moleculares, uma vez que a capacidade preditiva é diretamente proporcional à raiz quadrada da herdabilidade ( $r_{y\hat{y}}^2 = r_{g\hat{g}}^2 h$ ) (RESENDE JR. et al., 2012).

Tabela 1 - Capacidade preditiva da seleção genômica ampla utilizando os mesmos indivíduos para população de treinamento e validação para estimação dos valores genômicos dos indivíduos da população de seleção, para três isolados de *H. vastatrix* em *C. arabica*.

Resistência ao isolado	$h_1^2$	$h_2^2$	$r_{y\hat{y}}^2$	Nº indivíduos
Raça I	0,193918	0,51473	0,66	245
Raça II	0,154665	0,523466	0,62	245
Patótipo 001	0,362894	0,500455	0,78	245

$h_1^2$  = herdabilidade estimada por meio dos marcadores moleculares;  $h_2^2$  = herdabilidade da característica considerando dados fenotípicos e o delineamento genético de análise de gerações P1, P2, F1 e F2; e  $r_{y\hat{y}}^2$  = capacidade preditiva.

Por meio da seleção genômica, observou-se que as herdabilidades ( $h_1^2$ ) estimadas a partir de marcadores moleculares para a resistência aos três isolados foram inferiores às herdabilidades ( $h_2^2$ ) baseadas em valores fenotípicos. A menor herdabilidade ( $h_1^2$ ) pode ser devida à natureza dos marcadores utilizados, que por vezes não correspondem aos genes causais da herança, à quantidade de marcadores utilizados ou, ainda, ao tamanho da população de estudo. Contudo, as herdabilidades ( $h_1^2$ ) utilizadas para calcular a CP da seleção genômica ( $r_{y\hat{y}}^2 = r_{g\hat{g}}^2 h$ ) foram capazes de identificar valores de capacidade preditiva para a resistência aos três isolados superiores às herdabilidades ( $h_2^2$ ) baseadas em valores fenotípicos. Isso demonstra maior predição dos valores genômicos dos indivíduos na seleção genômica, em comparação com a seleção fenotípica.

Observou-se ainda que as maiores e as menores CP foram obtidas para a resistência ao patótipo 001 e à raça II, respectivamente. Apesar de aparentemente baixos, os valores de capacidade preditiva observados para as características de resistência aos três isolados foram superiores às descritas por outros autores. Dada à escassa disponibilidade até o momento de trabalhos focando a seleção genômica em café, torna-se dificultada a comparação dos parâmetros genéticos estimados. No entanto, têm sido relatadas em outras culturas estimativas de parâmetros inferiores às obtidas neste trabalho e, mesmo assim, isso não tem sido impedimento para a seleção e melhoramento das respectivas espécies.

As capacidades preditivas observadas nos três isolados apresentados na Tabela 1 foram, assim, superiores às encontradas por Fritsche-Neto et al. (2012) em milho e Cavalcanti et al. (2012) em cajueiro, pelo método RR-BLUP, utilizando uma mesma população, tanto para treinamento quanto para validação com todos os marcadores. Fritsche-Neto et al. (2012) utilizaram uma população composta por 41 combinações híbridas e 80 SSR (214 alelos). Avaliando vários caracteres fenotípicos, eles identificaram valores de CP variando de -0,23 a 0,61. Cavalcanti et al. (2012), com 74 indivíduos e 238 (224 AFLP, 12 SSR e dois ISSR) marcadores, verificaram capacidade preditiva de 0,05 para peso da amêndoa em cajueiro. Baixa capacidade preditiva também foi observada para duas características de resistência à ferrugem em pinheiro (RESENDE JR. et al., 2012). Eles identificaram capacidade preditiva de 0,28 e 0,24 para *Rust\_bin* e *Rust\_gall\_vol*, respectivamente, com herdabilidade da ordem de 0,21 e 0,12, utilizando 951 indivíduos e 4.853 SNPs, pelo método BLASSO. Esses autores relataram que a CP foi linearmente correlacionada com a herdabilidade da característica e, ainda, que para caracteres de menor herdabilidade menor será a predição na GWS. No entanto, neste trabalho, para a raça II, com herdabilidade ( $h_1^2$ ) de 0,15, a CP foi de 0,62, com menor número de indivíduos (245) e marcadores (137).

Os resultados de análise da seleção genômica ampla, em que a população F<sub>2</sub> foi dividida em duas populações (200 indivíduos para população de treinamento e 45 indivíduos para população de validação), estão apresentados na Tabela 2.

Tabela 2 - Capacidade preditiva da seleção genômica ampla com indivíduos selecionados de forma aleatória, para constituir a população de treinamento e validação para estimação dos valores genéticos genômicos dos indivíduos da população de seleção, para três isolados de *H. vastatrix* em *C. arabica*.

<b>Resistência ao isolado</b>	<b>População de treinamento (<math>r_{\hat{y}y}^2</math>)</b>	<b>População de validação (<math>r_{\hat{y}y}^2</math>)</b>
Raça I	0,66	0,49
Raça II	0,68	0,16
Patótipo 001	0,81	0,38
Nº de indivíduos	200	45

De acordo com a Tabela 2, pode-se observar que os valores de CP da população de treinamento foram poucos superiores ao verificado na Tabela 1, em que a população de treinamento foi utilizada com todos os indivíduos da população. Contudo, a capacidade preditiva da população de validação foi inferior à observada na população de treinamento, o que de fato é justificável pelo número (45) de indivíduos inferior ao constatado na população de treinamento (200). Contudo, analisando as Tabelas 1 e 2, não houve diferença pronunciada entre os valores de CP da população de treinamento, o que possibilita inferir que o número de indivíduos utilizados nessa população não interferiu na análise estatística. Contudo, necessita-se de mais indivíduos na população de treinamento e validação para que se tenha maior precisão das estimativas dos valores genômicos dos indivíduos da população de seleção.

Duas populações de eucalipto da empresa CENIBRA (780 indivíduos) e FIBRIA (920 indivíduos) foram utilizadas para seleção genômica ampla, avaliando quatro caracteres, com 3.000 marcadores DArT. Para os mesmos caracteres, os resultados de capacidade preditiva para a população da CENIBRA foram de 0,51 a 0,60 e de 0,38 a 0,55 para a população da FIBRIA, tomando-se dados da população de validação (RESENDE et al., 2012).

Em geral, a seleção dos indivíduos superiores poderá tornar-se de alto custo se houver necessidade de genotipá-los com todos os marcadores disponíveis. Porém, existe a possibilidade de utilizar apenas marcadores de maiores efeitos para inferir sobre os valores genômicos de indivíduos superiores, reduzindo custo e tempo. A

possibilidade de utilizar apenas os marcadores de maior efeito também foi efetuada neste trabalho, de acordo com os resultados apresentados na Tabela 3.

Tabela 3 - Capacidade preditiva da seleção genômica ampla com indivíduos selecionados de forma aleatória, para constituir a população de treinamento e validação para estimação dos valores genético-genômicos dos indivíduos da população de seleção, para três isolados de *H. vastatrix* em *C. arabica*. Selecionando marcadores de maior efeito.

<b>Resistência ao isolado</b>	<b>População de treinamento (<math>r^2_{y\hat{y}}</math>)</b>	<b>População de validação (<math>r^2_{y\hat{y}}</math>)</b>	
Nº de indivíduos	200	45	Nº marcadores
Raça I	0,50	0,42	10
Raça II	0,49	0,37	10
Patótipo 001	0,66	0,49	6

Nessa tabela, pode-se verificar que houve pouca variação dos valores de capacidade preditiva na população de treinamento, comparada com as Tabelas 1 e 2. Dessa forma, pode-se inferir que os marcadores de maiores efeitos podem ser utilizados para prever os valores genômicos dos indivíduos candidatos à seleção, sem haver a necessidade de genotipá-los com todos os marcadores.

Comparando os resultados apresentados nas Tabelas 2 e 3 para a população de validação, pode-se observar que para a raça I a CP foi um pouco inferior à observada na Tabela 2, para todos os marcadores. Contudo, para a raça II e o patótipo 001 os resultados de CP, utilizando-se apenas marcadores de maior efeito, foi um pouco superior ao observado com os resultados da utilização de todos os marcadores. Isso justifica o emprego de apenas os marcadores de maiores efeitos para prever os valores genômicos de indivíduos superiores. O uso de menor número de marcadores é de grande interesse para o melhorista, uma vez que a obtenção da informação de indivíduos superiores será dada com menor tempo e custo.

O número de marcadores que maximizaram a capacidade preditiva foi verificado em milho e cajueiro, por Fritsche-Neto et al. (2012) e Cavalcanti et al. (2012), respectivamente. Contudo, esses autores dividiram o total de marcadores moleculares em vários grupos. A divisão dos grupos foi dada de forma sequencial,

retirando-se da análise os marcadores de menor efeito até atingir o extremo de apenas um marcador (de maior efeito), formando um subgrupo.

Para milho, Fritsche-Neto et al. (2012) relataram que houve grande variação entre os caracteres e o número de marcas que maximizaram a capacidade preditiva. O número de marcadores que maximizaram a CP variou de 1 a 50. Contudo, a característica que apresentou maior capacidade preditiva (0,68) foi a eficiência de absorção com apenas cinco marcadores. Todavia, em cajueiro a maior capacidade preditiva observada para peso de amêndoa foi de 0,79, com 70 marcadores. De acordo com Resende et al. (2010), as capacidades preditivas diminuem com o aumento no número de marcadores, além de um número ótimo. Fritsche-Neto et al. (2012) e Cavalcanti et al. (2012) relataram ainda que, quando foram consideradas todas as marcas, as CP foram muito baixas, ou seja, o aumento no número de marcadores não aumenta linearmente a capacidade preditiva, sendo fundamental para cada caráter a ser selecionado estabelecer o número ótimo de marcadores que maximizem a capacidade preditiva na população de validação.

Resende Jr. et al. (2012) também identificaram, para resistência à doença em pinheiro (*Rust\_bin* e *Rust\_gall\_vol*), que a capacidade preditiva máxima era atingida com número menor de marcadores (110 a 590) do que com 3.000 DArT, conforme relatado anteriormente. Contudo, o número ótimo de marcadores que maximize a capacidade preditiva na população de validação não foi verificado neste trabalho. No entanto, os marcadores de maior efeito apresentaram bons resultados, os quais poderão ser utilizados para selecionar indivíduos resistentes.

A capacidade da GWS em prever os valores genômicos dos indivíduos foi realizada por meio de um cálculo denominado taxa de erro na população de validação para as raças I e II e o patótipo 001 (Tabelas 4, 5 e 6). Para todos os três isolados, observou-se que a taxa de erro em prever os valores genômicos foi bem maior quando se usaram as notas fenotípicas variando de 1 a 6 do que quando foram utilizadas as notas binárias resistentes (R) e suscetíveis (S). Isso pode ser devido, entre outros fatores, às dificuldades inerentes ao processo de fenotipagem, uma vez que esta é subjetiva e dependente do grau de experiência do avaliador. Além disso, na população F<sub>2</sub> utilizada foram observados poucos indivíduos com notas variando de 4 a 6 e muitos indivíduos com notas variando de 1 a 3, o que torna a média viesada. Assim, a subtração da média dos valores fenotípicos observados da população de validação pela média dos valores genômicos dos indivíduos preditos

pela GWS resulta em constante de baixo valor. Essa característica dificulta a identificação dos indivíduos que apresentam notas elevadas (4, 5 e 6) pela GWS, como pode ser observado nas Tabelas 4, 5 e 6 para as raças I e II e o patótipo 001, respectivamente. Apesar de tal fato, pode-se inferir sobre a possibilidade de praticar a seleção genômica ampla em indivíduos de interesse, tomando-se os marcadores de maiores efeitos para três isolados de *H. vastatrix*.

Tabela 4 - Estimativa do valor genômico do indivíduo para a raça I de *H. vastatrix*, considerando-se apenas os marcadores de maiores efeitos. Dividindo a população F<sub>2</sub> em dois grupos, amostrados aleatoriamente, sendo o primeiro grupo formado por 200 indivíduos para população de treinamento e o segundo por 45 indivíduos para população de validação. Ind = indivíduo da população F<sub>2</sub>; F = fenotípico; e G = genotípico

Marcador	Efeito	Ind.	Valores		Notas	
			F	G	F	G
E-CCC/M-AGA1	0,1573	22	<b>3</b>	<b>4</b>	<b>R</b>	<b>S</b>
E-CGA/M-TCC4	0,1240	49	<b>3</b>	<b>4</b>	<b>R</b>	<b>S</b>
E-CTG/M-TAC5	0,1753	13	3	3	R	R
E-CAG/M-TGC1	0,0543	99	<b>5</b>	<b>3</b>	<b>S</b>	<b>R</b>
E-CTG/M-AAT1	0,1998	206	3	3	R	R
E-CGC/M-ATA1	0,04639	165	3	3	R	R
SSR035	-0,0330	208	<b>4</b>	<b>3</b>	<b>S</b>	<b>R</b>
SSR008	-0,0531	239	<b>4</b>	<b>3</b>	<b>S</b>	<b>R</b>
EST-SSR067	-0,1330	111	<b>4</b>	<b>3</b>	<b>S</b>	<b>R</b>
EST-SSR096	-0,1831	83	3	3	R	R
		209	3	3	R	R
		112	3	3	R	R
		42	3	3	R	R
		53	3	3	R	R
		18	3	3	R	R
		229	<b>2</b>	<b>3</b>	R	R
		119	3	3	R	R
		201	<b>2</b>	<b>3</b>	R	R
		194	3	3	R	R
		85	<b>2</b>	<b>3</b>	R	R
		221	<b>2</b>	<b>3</b>	R	R
		25	3	3	R	R
		60	3	3	R	R
		79	3	3	R	R
		232	<b>2</b>	<b>3</b>	R	R
		19	3	3	R	R
		141	<b>2</b>	<b>3</b>	R	R
		26	3	3	R	R
		121	3	3	R	R
		215	<b>4</b>	<b>3</b>	<b>S</b>	<b>R</b>
		236	3	3	R	R

Continua...

Tabela 4 - Cont.

Marcador	Efeito	Ind.	Valores		Notas	
			F	G	F	G
		44	3	3	R	R
		177	3	3	R	R
		40	3	3	R	R
		10	3	3	R	R
		47	<b>2</b>	<b>3</b>	R	R
		193	3	3	R	R
		6	<b>2</b>	<b>3</b>	R	R
		14	3	3	R	R
		63	3	3	R	R
		222	<b>2</b>	<b>3</b>	R	R
		154	<b>2</b>	<b>3</b>	R	R
		233	3	3	R	R
		217	<b>2</b>	<b>3</b>	R	R
		58	<b>2</b>	<b>2</b>	R	R
Taxa de erro (%)			40,0		15,5	

Tabela 5 - Estimativa do valor genômico do indivíduo para a raça II de *H. vastatrix*, considerando-se apenas os marcadores de maiores efeitos. Dividindo a população F<sub>2</sub> em dois grupos, amostrados aleatoriamente, sendo o primeiro grupo formado por 200 indivíduos para população de treinamento e o segundo por 45 indivíduos para população de validação. Ind = indivíduo da população F<sub>2</sub>; F = fenotípico; e G = genotípico

Marcador	Efeito	Ind	Valores		Notas	
			F	G	F	G
E-CTG/M-TAC5	0.2056	60	<b>3</b>	<b>2</b>	R	R
E-CTG/M-AAT1	0.1337	85	<b>3</b>	<b>2</b>	R	R
E-CCC/M-AGA1	0.1250	10	<b>1</b>	<b>2</b>	R	R
E-CGA/M-TCC4	0.1112	99	<b>2</b>	<b>3</b>	R	R
E-CGC/M-ATA1	0.0435	40	<b>3</b>	<b>2</b>	R	R
E-CAG/M-TGC1	0.0344	79	<b>3</b>	<b>2</b>	R	R
SSR035	0.0137	193	2	2	R	R
SSR007	-0.0260	25	<b>1</b>	<b>2</b>	R	R
EST-SSR067	-0,1423	232	<b>1</b>	<b>2</b>	R	R
EST-SSR096	-0,2073	206	3	3	R	R
		154	2	2	R	R
		217	2	2	R	R
		239	3	3	R	R
		6	<b>1</b>	<b>2</b>	R	R
		119	<b>2</b>	<b>3</b>	R	R
		165	3	3	R	R
		236	<b>3</b>	<b>2</b>	R	R
		222	<b>3</b>	<b>2</b>	R	R
		22	<b>2</b>	<b>4</b>	<b>R</b>	<b>S</b>
		42	3	3	R	R

Continua...

Tabela 5 - Cont.

Marcador	Efeito	Ind	Valores		Notas	
			F	G	F	G
		47	<b>3</b>	<b>2</b>	R	R
		19	<b>3</b>	<b>2</b>	R	R
		44	2	2	R	R
		141	<b>3</b>	<b>2</b>	R	R
		13	3	3	R	R
		233	<b>3</b>	<b>2</b>	R	R
		111	3	3	R	R
		177	2	2	R	R
		58	2	2	R	R
		208	3	3	R	R
		112	3	3	R	R
		53	3	3	R	R
		209	3	3	R	R
		26	2	2	R	R
		63	2	2	R	R
		215	2	2	R	R
		121	<b>3</b>	<b>2</b>	R	R
		83	<b>2</b>	<b>3</b>	R	R
		229	<b>2</b>	<b>3</b>	R	R
		194	3	3	R	R
		14	2	2	R	R
		49	3	3	R	R
		201	<b>2</b>	<b>3</b>	R	R
		221	<b>3</b>	<b>2</b>	R	R
		18	3	3	R	R
Taxa de erro (%)			48,90		2,22	

Tabela 6 - Estimativa do valor genômico do indivíduo para o patótipo 001 de *H. vastatrix*, considerando-se apenas os marcadores de maiores efeitos. Dividindo a população F<sub>2</sub> em dois grupos, amostrados aleatoriamente, sendo o primeiro grupo formado por 200 indivíduos para população de treinamento e o segundo grupo por 45 indivíduos para população de validação. Ind = indivíduo da população F<sub>2</sub>; F = fenotípico; e G = genotípico.

Marcador	Efeito	Ind.	Valores		Notas	
			F	G	F	G
E-CTG/M-AAT1	0,3141	99	3	3	R	R
E-CGA/M-TCC4	0,2722	123	3	3	R	R
E-CTG/M-TAC5	0,3367	138	<b>2</b>	<b>3</b>	R	R
E-CCC/M-AGA1	0,2215	71	3	3	R	R
SSR016	0,2724	31	3	3	R	R
E-CTC/M-AAG3	-0,0312	18	3	3	R	R
		101	<b>2</b>	<b>3</b>	R	R

Continua...

Tabela 6 - Cont.

Marcador	Efeito	Ind.	Valores		Notas	
			F	G	F	G
		244	3	3	R	R
		194	<b>2</b>	<b>3</b>	R	R
		56	<b>2</b>	<b>3</b>	R	R
		136	3	3	R	R
		24	3	3	R	R
		13	2	2	R	R
		34	2	2	R	R
		230	2	2	R	R
		145	<b>3</b>	<b>2</b>	R	R
		38	2	2	R	R
		130	2	2	R	R
		53	<b>3</b>	<b>2</b>	R	R
		162	<b>1</b>	<b>2</b>	R	R
		40	2	2	R	R
		190	2	2	R	R
		79	2	2	R	R
		142	<b>3</b>	<b>2</b>	R	R
		20	2	2	R	R
		57	2	2	R	R
		5	<b>1</b>	<b>2</b>	R	R
		63	2	2	R	R
		72	2	2	R	R
		220	<b>3</b>	<b>2</b>	R	R
		15	<b>1</b>	<b>2</b>	R	R
		61	2	2	R	R
		137	<b>1</b>	<b>2</b>	R	R
		148	2	2	R	R
		170	2	2	R	R
		9	<b>1</b>	<b>2</b>	R	R
		241	<b>2</b>	<b>1</b>	R	R
		89	<b>2</b>	<b>1</b>	R	R
		91	1	1	R	R
		42	1	1	R	R
		128	<b>2</b>	<b>1</b>	R	R
		109	1	1	R	R
		192	<b>3</b>	<b>1</b>	R	R
		182	<b>2</b>	<b>1</b>	R	R
		78	<b>2</b>	<b>1</b>	R	R
Taxa de erro (%)			42,22		0	

Diante dos resultados, pode-se inferir que é possível praticar a seleção genômica ampla em indivíduos de interesse, tomando os marcadores de maiores efeitos para três isolados de *H. vastatrix*.

Comparando os resultados de marcadores moleculares de maior efeito por meio da técnica da GWS com os marcadores do mapa genético de ligação associado

aos QTLs (Capítulo 1) (Tabela 7), pode-se observar para a raça I que apenas o marcador EST-SSR050 identificado no mapa não foi identificado pela GWS. Contudo, outros marcadores foram identificados pela seleção genômica ampla.

Tabela 7 - Marcadores moleculares identificados associados às raças I e II e ao patótipo 001 de *H. vastatrix* em mapa genético de ligação e pelo método de seleção genômica ampla/BLASSO, considerando-se todos os indivíduos para população de treinamento e validação.

Mapa de ligação			BLASSO	
GL	QTL	Marcador	GL	Marcador
Raça I				
2	QTL1	E-CGC/M-ATA1	10	E-CGA/M-TCC4
2	QTL2	E-CTG/M-AAT1	10	E-CCC/M-AGA1
3	QTL3	EST-SSR050	2	E-CTG/M-AAT1
10	QTL4	E-CGA/M-TCC4	2	E-CGC/M-ATA1
			NL	E-CGG/M-ACG6
			6	EST-SSR023
Raça II				
2	QTL 1	E-CTG/M-AAT 1	10	E-CCC/M-AGA1
10	QTL 2	E-CCC/M-AGA 1	10	E-CGA/M-TCC4
			2	E-CTG/M-AAT1
			3	E-CTG/M-TAC5
			2	SSR016
			2	E-CGC/M-ATA1
Patotipo001				
2	QTL 1	E-CGC/M-ATA 1	2	E-CTG/M-AAT1
2	QTL 2	E-CTG/M-AAT 1	10	E-CGA/M-TCC4
3	QTL 3	EST-SSR 050	3	E-CTG/M-TAC5
10	QTL 4	E-CGA/M-TCC 4	10	E-CCC/M-AGA1
			2	SSR016
			NL	E-CGA/M-TCC5
			NL	EST-SSR002
			1	E-CGA/M-TCC3
			2	E-CGC/M-ATA1

GL- grupo de ligação; e NL- marcador molecular não ligado a nenhum dos GLs formados no mapa genético de ligação.

Para a raça II, os dois marcadores identificados em mapa genético foi observado pela GWS. No entanto, outros marcadores também foram identificados. O marcador SSR016, localizado no grupo de ligação (GL) 2, está muito próximo do E-CTG/M-AAT1, como também o marcador E-CGA/M-TCC4 está muito próximo do marcador E-CCC/M-AGA1 no GL10. Dos marcadores identificados para o patótipo 001 em mapa genético de ligação, apenas o EST-SSR050 não foi identificado pela

GWS. Contudo, o marcador E-CTG/M-TAC5 está próximo desse marcador no GL3. Em vista disso, os marcadores identificados por meio das duas técnicas (mapa de ligação e GWS) poderão ser utilizados com maior precisão para seleção precoce de indivíduos superiores.

## 4. CONCLUSÃO

A análise da população avaliada permitiu a estimação dos valores genômicos dos indivíduos, assim como predição dos fenótipos de resistência.

Os marcadores de maiores efeitos selecionados podem ser utilizados na predição dos valores genômicos dos indivíduos candidatos à seleção, eliminando a necessidade de genotipagem dos indivíduos com todos os marcadores.

A maioria dos marcadores identificados associados aos QTLs no mapa genético de ligação foi identificada também pela GWS, confirmando a acurácia da metodologia utilizada.

## 5. REFERÊNCIAS

ALMEIDA, Í. F.; CRUZ, C. D.; ALMEIDA, R. V. D.; MENDES, F. R. MÉTODOS DE ESTIMAÇÃO E VALIDAÇÃO NA SELEÇÃO GENÔMICA-DOI: 10.12971/2179-5959. v03n02a01. **Revista Agrotecnologia**, v. 3, n. 2, p. 1-11, 2012. ISSN 2179-5959.

ALVARENGA, S. M.; CAIXETA, E. T.; HUFNAGEL, B.; THIEBAUT, F.; MACIEL-ZAMBOLIM, E.; ZAMBOLIM, L.; SAKIYAMA, N. S. Molecular markers from coffee genome expressed sequences potentially involved in resistance to rust. **Pesquisa Agropecuária Brasileira**, v. 46, n. 8, p. 890-898, 2011. ISSN 0100-204X.

BARUAH, A.; NAIK, V.; HENDRE, P. S.; RAJKUMAR, R.; RAJENDRAKUMAR, P.; AGGARWAL, R. Isolation and characterization of nine microsatellite markers from *Coffea arabica* L., showing wide cross-species amplifications. **Molecular Ecology Notes**, v. 3, n. 4, p. 647-650, 2003. ISSN 1471-8286.

BHAT, P. R.; KRISHNAKUMAR, V.; HENDRE, P. S.; RAJENDRAKUMAR, P.; VARSHNEY, R. K.; AGGARWAL, R. K. Identification and characterization of expressed sequence tags-derived simple sequence repeats markers from robusta coffee variety 'CxR'(an interspecific hybrid of *Coffea canephora* × *Coffea congensis*). **Molecular Ecology Notes**, v. 5, n. 1, p. 80-83, 2005. ISSN 1471-8286.

BRITO, G. G.; CAIXETA, E. T.; GALLINA, A. P.; ZAMBOLIM, E. M.; ZAMBOLIM, L.; DIOLA, V.; LOUREIRO, M. E. Inheritance of coffee leaf rust resistance and identification of AFLP markers linked to the resistance gene. **Euphytica**, v. 173, n. 2, p. 255-264, 2010. ISSN 0014-2336.

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. **Pesquisa Agropecuária Brasileira**, v. 44, n. 3, p. 276-282, 2009.

CAVALCANTI, J. J. V.; RESENDE, M. D. V. D.; SANTOS, F. H. C. D.; PINHEIRO, C. R. Simultaneous prediction of the effects of molecular markers and

genome wide selection in cashew. **Revista Brasileira de Fruticultura**, v. 34, n. 3, p. 840-846, 2012. ISSN 0100-2945.

COMBES, M.-C.; ANDRZEJEWSKI, S.; ANTHONY, F.; BERTRAND, B.; ROVELLI, P.; GRAZIOSI, G.; LASHERMES, P. Characterization of microsatellite loci in *Coffea arabica* and related coffee species. **Molecular Ecology**, v. 9, n. 8, p. 1178-1180, 2000. ISSN 1365-294X.

COULIBALY, I.; REVOL, B.; NOIROT, M.; PONCET, V.; LORIEUX, M.; CARASCO-LACOMBE, C.; MINIER, J.; DUFOUR, M.; HAMON, P. AFLP and SSR polymorphism in a *Coffea* interspecific backcross progeny [(*C. heterocalyx* × *C. canephora*) × *C. canephora*]. **Theoretical and Applied Genetics**, v. 107, n. 6, p. 1148-1155, 2003. ISSN 0040-5752.

CRUZ, C. D. Genes: a software package for analysis in experimental statistics and quantitative genetics. **Acta Scientiarum. Agronomy**, v. 35, n. 3, p. 271-276, 2013. ISSN 1807-8621.

DE LOS CAMPOS, G.; NAYA, H.; GIANOLA, D.; CROSSA, J.; LEGARRA, A.; MANFREDI, E.; WEIGEL, K.; COTES, J. M. Predicting quantitative traits with regression models for dense molecular markers and pedigree. **Genetics**, v. 182, n. 1, p. 375-385, 2009. ISSN 0016-6731.

DINIZ, L. E. C.; SAKIYAMA, N. S.; LASHERMES, P.; CAIXETA, E. T.; OLIVEIRA, A. C. B.; ZAMBOLIM, E. M.; 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 TECHNOLOGY**, v. 5, n. 4, p. 387, 2005. ISSN 1518-7853.

FRITSCHÉ-NETO, R.; RESENDE, M. D. V.; MIRANDA, G. V.; DOVALE, J. C. Seleção genômica ampla e novos métodos de melhoramento do milho. **Rev. Ceres, Viçosa-MG**, v. 28, p. 794-802, 2012.

GELETA, M.; HERRERA, I.; MONZÓN, A.; BRYNGELSSON, T. Genetic diversity of Arabica coffee (*Coffea arabica* L.) in Nicaragua as estimated by simple sequence repeat markers. **The Scientific World Journal**, v. 2012, 2012.

HEFFNER, E. L.; SORRELLS, M. E.; JANNINK, J.-L. Genomic selection for crop improvement. **Crop Science**, v. 49, n. 1, p. 1-12, 2009. ISSN 1435-0653.

HENDRE, P. S.; PHANINDRANATH, R.; ANNAPURNA, V.; LALREMURATA, A.; AGGARWAL, R. K. Development of new genomic microsatellite markers from robusta coffee (*Coffea canephora* Pierre ex A. Froehner) showing broad cross-species transferability and utility in genetic studies. **BMC plant biology**, v. 8, n. 1, p. 51, 2008. ISSN 1471-2229.

HERRERA, J. C.; ALVARADO A, G.; CORTINA G, H. A.; COMBES, M.-C.; ROMERO G, G.; LASHERMES, P. Genetic analysis of partial resistance to coffee leaf rust (*Hemileia vastatrix* Berk & Br.) introgressed into the cultivated *Coffea*

arabica L. from the diploid *C. canephora* species. **Euphytica**, v. 167, n. 1, p. 57-67, 2009. ISSN 0014-2336.

KY, C.-L.; BARRE, P.; LORIEUX, M.; TROUSLOT, P.; AKAFFOU, S.; LOUARN, J.; CHARRIER, A.; HAMON, S.; NOIROT, M. Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp.). **Theoretical and Applied Genetics**, v. 101, n. 4, p. 669-676, 2000. ISSN 0040-5752.

LEGARRA, A.; ROBERT-GRANIÉ, C.; CROISEAU, P.; GUILLAUME, F.; FRITZ, S. Improved Lasso for genomic selection. **Genetics research**, v. 93, n. 1, p. 77, 2011. ISSN 0016-6723.

MEUWISSEN, T. H.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, p. 1819-1829, 2001.

MISSIO, R. F.; CAIXETA, E. T.; ZAMBOLIM, E. M.; PENA, G. F.; RIBEIRO, A. P.; ZAMBOLIM, L.; PEREIRA, A. A.; SAKIYAMA, N. S. Assessment of EST-SSR markers for genetic analysis on coffee. **Bragantia**, v. 68, n. 3, p. 573-581, 2009a. ISSN 0006-8705.

MISSIO, R. F.; CAIXETA, E. T.; ZAMBOLIM, E. M.; PENA, G. F.; ZAMBOLIM, L.; DIAS, L. A. S.; SAKIYAMA, N. S. Genetic characterization of an elite coffee germplasm assessed by gSSR and EST-SSR markers. **Genet Mol Res**, v. 10, p. 2366-2381, 2011.

MISSIO, R. F.; CAIXETA, E. T.; ZAMBOLIM, E. M.; ZAMBOLIM, L.; SAKIYAMA, N. S. Development and validation of SSR markers for *Coffea arabica* L. **Crop Breeding and Applied Biotechnology**, v. 9, n. 4, p. 361-371, 2009b. ISSN 1518-7853.

MONCADA, P.; MCCOUCH, S. Simple sequence repeat diversity in diploid and tetraploid *Coffea* species. **Genome**, v. 47, n. 3, p. 501-509, 2004. ISSN 0831-2796.

OLIVEIRA, A. C. B.; SAKIYAMA, N. S.; CAIXETA, E. T.; ZAMBOLIM, E. M.; RUFINO, R. J. N.; ZAMBOLIM, L. Partial map of *Coffea arabica* L. and recovery of the recurrent parent in backcross progenies. **Crop Breeding and Applied Biotechnology**, v. 7, n. 2, p. 196-203, 2007. ISSN 1518-7853.

PARK, T.; CASELLA, G. The bayesian lasso. **Journal of the American Statistical Association**, v. 103, n. 482, p. 681-686, 2008. ISSN 0162-1459.

PEARL, H. M.; NAGAI, C.; MOORE, P. H.; STEIGER, D. L.; OSGOOD, R. V.; MING, R. Construction of a genetic map for arabica coffee. **Theoretical and Applied Genetics**, v. 108, n. 5, p. 829-835, 2004. ISSN 0040-5752.

POLAND, J.; ENDELMAN, J.; DAWSON, J.; RUTKOSKI, J.; WU, S.; MANES, Y.; DREISIGACKER, S.; CROSSA, J.; SÁNCHEZ-VILLEDA, H.; SORRELLS, M. Genomic selection in wheat breeding using genotyping-by-sequencing. **The Plant Genome**, v. 5, n. 3, p. 103-113, 2012. ISSN 1940-3372.

PONCET, V.; HAMON, P.; MINIER, J.; CARASCO, C.; HAMON, S.; NOIROT, M. SSR cross-amplification and variation within coffee trees (*Coffea* spp.). **Genome**, v. 47, n. 6, p. 1071-1081, 2004. ISSN 0831-2796.

RESENDE JR, M. F. R.; ALVES, A. A.; SÁNCHEZ, C. F. B.; RESENDE, M. D. V.; CRUZ, C. D. Seleção Genômica Ampla. In: CRUZ, C. D.;SALGADO, C. C., *et al* (Ed.). **Genômica Aplicada**. Viçosa-MG, 2013. p.375-424.

RESENDE JR, M. F. R.; MUÑOZ, P.; RESENDE, M. D. V.; GARRICK, D. J.; FERNANDO, R. L.; DAVIS, J. M.; JOKELA, E. J.; MARTIN, T. A.; PETER, G. F.; KIRST, M. Accuracy of genomic selection methods in a standard data set of loblolly pine (*Pinus taeda* L.). **Genetics**, v. 190, n. 4, p. 1503-1510, 2012. ISSN 0016-6731.

RESENDE, M. D. V. Matemática e estatística na análise de experimentos e no melhoramento genético. **Embrapa Florestas, Colombo-PR**, v. 1, p. 561p, 2007.

RESENDE, M. D. V. Genômica quantitativa e seleção no melhoramento de plantas perenes e animais. **Embrapa Florestas, Colombo-PR**, p. 330p, 2008.

RESENDE, M. D. V.; JÚNIOR, M. F. R. R.; AGUIAR, A. M.; ABAD, J. I. M.; MISSIAGGIA, A. A.; SANSALONI, C.; PETROLI, C.; GRATTAPAGLIA, D. **Computação da seleção genômica ampla (GWS)**. 2010.

RESENDE, M. D. V.; RESENDE, M. F. R.; SANSALONI, C. P.; PETROLI, C. D.; MISSIAGGIA, A. A.; AGUIAR, A. M.; ABAD, J. M.; TAKAHASHI, E. K.; ROSADO, A. M.; FARIA, D. A. Genomic selection for growth and wood quality in Eucalyptus: capturing the missing heritability and accelerating breeding for complex traits in forest trees. **New Phytologist**, v. 194, n. 1, p. 116-128, 2012. ISSN 1469-8137.

ROVELLI, P.; METTULIO, R.; ANTHONY, F.; ANZUETO, F.; LASHERMES, P.; GRAZIOSI, G. Microsatellites in *Coffea Arabica* L. In: (Ed.). **Coffee biotechnology and quality**: Springer, 2000. p.123-133. ISBN 9048155657.

SETOTAW, T. A.; CAIXETA, E. T.; PEREIRA, A. A.; OLIVEIRA, A. C. B.; CRUZ, C. D.; ZAMBOLIM, E. M.; ZAMBOLIM, L.; SUSSUMU, N. **Coefficient of parentage in *Coffea arabica* L. cultivars grown in Brazil**. 2013.

SOLBERG, T. R.; SONESSON, A. K.; WOOLLIAMS, J. A. Genomic selection using different marker types and densities. **Journal of animal science**, v. 86, n. 10, p. 2447-2454, 2008. ISSN 0021-8812.

TAMAYO, P. J.; VALE, F. X. R.; ZAMBOLIM, L.; CHAVES, G. M.; PEREIRA, A. A. Resistência do Catimor à ferrugem e virulência de raças fisiológicas de *Hemileia vastatrix* Berk. & Br. **Fitopatologia Brasileira**, v. 20, p. 572-576, 1995.

TEAM, R. D. C. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**. Vienna, Austria, 2012. ISSN 3-900051-07-0. Disponível em: <<http://www.R-project.org/>>.

VIEIRA, E. S. N.; VON PINHO, É. V. R.; CARVALHO, M. G. G.; ESSELINK, D. G.; VOSMAN, B. Development of microsatellite markers for identifying Brazilian *Coffea arabica* varieties. **Genetics and molecular biology**, v. 33, n. 3, p. 507-514, 2010. ISSN 1415-4757.