



GUILHERME DE JONG

**ANÁLISE DE ASSOCIAÇÃO GENÔMICA PARA
RESISTÊNCIA À PODRIDÃO DE ESPIGA
CAUSADA POR *Fusarium verticillioides***

**LAVRAS – MG
2016**

GUILHERME DE JONG

**ANÁLISE DE ASSOCIAÇÃO GENÔMICA PARA RESISTÊNCIA À
PODRIDÃO DE ESPIGA CAUSADA POR *Fusarium verticillioides***

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, para a obtenção do título de Mestre.

Prof. Dr. Renzo Garcia Von Pinho
Orientador

Prof. Dr. Marcio Balestre
Coorientador

**LAVRAS – MG
2016**

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

de Jong, Guilherme.

Análise de associação genômica para resistência à podridão de
espiga causada por *Fusarium verticillioides* / Guilherme de Jong. –
Lavras : UFLA, 2016.

103 p.

Dissertação (mestrado acadêmico)–Universidade Federal de
Lavras, 2016.

Orientador(a): Renzo Garcia Von Pinho.

Bibliografia.

1. Estudo de associação genômica. 2. Podridão de espiga. 3.
Bayesian Stochastic Search Variable. I. Universidade Federal de
Lavras. II. Título.

GUILHERME DE JONG

**ANÁLISE DE ASSOCIAÇÃO GENÔMICA PARA RESISTÊNCIA À
PODRIDÃO DE ESPIGA CAUSADA POR *Fusarium verticillioides***

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, para a obtenção do título de Mestre.

APROVADA em 31 de agosto de 2016.

Dr. Renzo Gracia Von Pinho UFLA

Dr. Marcio Balestre UFLA

Dra. Marcela Pedroso Mendes Resende UFG

Prof. Dr. Renzo Gracia Von Pinho
Orientador

Prof. Dr. Marcio Balestre
Coorientador

**LAVRAS – MG
2016**

Aos meus pais, Robert e Edenéia, pelo incentivo e confiança.
Aos meus irmãos, Gabriele e Gustavo, por todo o apoio.

DEDICO

AGRADECIMENTOS

A Deus.

À Universidade Federal de Lavras, ao Programa de Pós-Graduação em Genética e Melhoramento de Plantas, pela oportunidade de realização deste trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de mestrado.

Ao Professor Renzo Garcia Von Pinho, pela orientação, incentivo e ensinamentos.

Ao Professor Marcio Balestre, pela orientação, paciência, disponibilidade e conselhos.

À Marcela Pedroso Mendes Resende, pela disponibilidade em sempre me ajudar, pelo apoio e conselhos.

À Andrezza, pela grande ajuda, disponibilidade e apoio.

Aos Professores do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, pelos conhecimentos transmitidos ao longo do curso.

Aos meus pais, pela confiança depositada em mim, pela compreensão, pelos grandes ensinamentos e por toda a dedicação e carinho em me educar.

À minha família, por todo o apoio.

Aos amigos, Jéssica e Indalécio, por todo apoio, ajuda, amizade e pelos momentos de descontração.

Aos amigos no Núcleo de Estudos em Genética e Melhoramento de Plantas, pela aprendizagem e amizade.

A todos que, de alguma forma, me apoiaram e incentivaram, meus sinceros agradecimentos.

RESUMO

O milho é um dos cereais mais cultivados e seu amplo cultivo permite que uma extensa gama de patógenos ataquem a cultura. Os patógenos de maior importância são os agentes causais de podridões de espiga, pois afetam diretamente a produção e a qualidade dos grãos. Dentre os principais agentes causais da podridão de espiga, destaca-se a podridão causada pelo fungo *Fusarium verticillioides*. A identificação de regiões causais associadas com a resistência à doença pode ser uma ferramenta útil no melhoramento da cultura e também é uma forma de entender melhor os mecanismos de resistência. Neste trabalho, foi conduzido um estudo de associação genômica para identificar associações significativas entre os marcadores e as possíveis regiões causais relacionadas com a resistência à podridão de espiga causada pelo fungo *F. verticillioides*. Foram avaliadas 242 linhagens em dois locais (Lavras e Uberlândia), para a proporção de espigas com sintomas de podridão, ou seja, que apresentavam mais de 25% da espiga doente em relação ao número total de espigas. As associações foram testadas com 23.153 marcadores DArT, utilizando o modelo misto clássico de mapeamento e o modelo de seleção genômica BSSV (*Bayesian Stochastic Search Variable*). Foram identificados onze marcadores significativamente associados com a resistência, sendo quatro associados por ambos os métodos. Todos os marcadores foram identificados próximos a genes que possuem funções que podem ter relação com a resistência, sendo necessários estudos de rotas metabólicas.

Palavras-chave: Estudo de Associação Genômica. Podridão de Espiga. Bayesian Stochastic Search Variable.

ABSTRACT

Maize is one of the most cultivated cereals and its extensive cultivation allows a wide range of pathogens to attack the culture. The most important pathogens that attack maize culture are the agents of ear rot, which directly affect grain yield and quality of kernels. Among the main causative agents of ear rot, the rot caused by the fungus *Fusarium verticillioides* stands out. The identification of causal regions associated with this disease resistance may be a useful tool to improve maize breeding and is also a way to better understand the mechanisms of resistance. In this research, a genome-wide association study was conducted to identify significant associations between markers and possible causal regions related to the resistance of ear rot of corn caused by the fungus *F. verticillioides*. Two hundred forty-two inbred lines were evaluated in two locations (Lavras and Uberlândia) to the proportion of ears with rot symptoms, i.e., with more than 25% of the ear rotten in relation to the total number of ears. Associations were tested with 23,153 DArT markers, using the classic mixed model mapping and genomic selection model BSSV (Bayesian Stochastic Search Variable). Eleven markers significantly associated with ear rot resistance were identified, and four markers were associated by both methods. All markers were identified near genes which have functions that may be related to the resistance, requiring studies of their metabolic pathways.

Keywords: Genome-Wide Association Study. Ear Rot. Bayesian Stochastic Search Variable.

LISTA DE FIGURAS

- Figura 1 – Árvore filogenética construída pelo método Neighbor-Joining, mostrando os diferentes backgrounds das linhagens utilizadas no estudo de associação genômica. 54
- Figura 2 – Gráficos Manhattan mostrando as associações significativas para as análises conjunta, individual de Lavras e individual de Uberlândia, utilizando o método *Bayesian Stochastic Search Variable* (BSSV). 56
- Figura 3 – Gráficos Manhattan mostrando as associações significativas para as análises conjunta, individual de Lavras e individual de Uberlândia, utilizando o método de modelo misto (MM). 60
- Figura 4 – Declínio do desequilíbrio de ligação entre os marcadores ao longo do genoma. A linha em vermelho representa o valor médio. 66
- Figura 5 – Valores fenotípicos corrigidos e valores dos marcadores que foram significativamente associados com a resistência para a análise conjunta. As cores variam do branco ao vermelho, sendo que a coloração branca representa os menores valores, e a coloração vermelha, os maiores valores. 68
- Figura 6 – Valores fenotípicos corrigidos e valores dos marcadores que foram associados significativamente associados com a resistência para a análise de Lavras. As cores variam do branco ao vermelho, sendo que a coloração branca representa os menores valores, e a coloração vermelha, os maiores valores. 70
- Figura 7 – Valores fenotípicos corrigidos e valores dos marcadores que foram associados significativamente associados com a resistência para a análise de Uberlândia. As cores variam do branco ao vermelho, sendo que a coloração branca representa os menores valores, e a coloração vermelha, os maiores valores. 71
- Figura 8 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1 Mbp do cromossomo 1 para o marcador 4582295 (posição 78,131,634 bp), que foi significativamente associado na análise conjunta do método BSSV e

nas análises conjunta e de Uberlândia pelo método de modelo misto.
..... 90

Figura 9 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 0,9 Mbp do cromossomo 1 para o marcador 4577966 (posição 136,272,488 bp), que foi significativamente associado nas análises conjunta e de Lavras pelo método BSSV..... 91

Figura 10 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1,1 Mbp do cromossomo 7 para o marcador 4592042 (posição 120,994,839 bp), que foi significativamente associado na análise conjunta pelo método BSSV.
..... 92

Figura 11 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1 Mbp do cromossomo 7 para o marcador 4579702 (posição 158,219,730 bp), que foi significativamente associado na análise conjunta pelo método BSSV.
..... 93

Figura 12 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 0,6 Mbp do cromossomo 1 para o marcador 4767436 (posição 53,415,835 bp), que foi significativamente associado na análise de Lavras pelo método BSSV..... 94

Figura 13 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 0,9 Mbp do cromossomo 7 para o marcador 2518496 (posição 124,220,878 bp), que foi significativamente associado na análise de Lavras pelo método BSSV e nas análises conjunta de Lavras pelo método de modelo misto. 95

Figura 14 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 0,8 Mbp do cromossomo 8 para o marcador 4772985 (posição 136,137,749 bp), que foi significativamente associado na análise de Lavras pelo método BSSV..... 96

- Figura 15 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1 Mbp do cromossomo 4 para o marcador 4772929 (posição 36,878,586 bp), que foi significativamente associado na análise de Uberlândia pelo método BSSV. 97
- Figura 16 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1,6 Mbp do cromossomo 5 para o marcador 4592572 (posição 55,225,453 bp), que foi significativamente associado na análise de Uberlândia pelo método BSSV e na análise de Uberlândia pelo método de modelo misto. 98
- Figura 17 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1,1 Mbp do cromossomo 7 para o marcador 4581091 (posição 159,891,338 bp), que foi significativamente associado na análise de Lavras pelo método de modelo misto. 99
- Figura 18 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1,1 Mbp do cromossomo 1 para o marcador 2400968 (posição 289,504,622 bp), que foi significativamente associado na análise de Uberlândia pelo método de modelo misto. 100

LISTA DE TABELAS

- Tabela 1 – Estimativa dos componentes de variância genética aditiva (σ_a^2), componentes de variância residual (σ_e^2) e estimativa da herdabilidade (h^2) para as análises conjunta, individual de Lavras e individual de Uberlândia..... 53
- Tabela 2 – Cromossomo, posição em bp, valores de $-\log_{10}(\text{p-valor})$ e sequência dos DArTs que foram associados significativamente com a resistência à podridão de espiga pelo método *Bayesian Stochastic Search Variable* (BSSV)..... 57
- Tabela 3 – Cromossomo, posição em bp, valores de $-\log_{10}(\text{p-valor})$ e sequência dos DArTs que foram associados significativamente com a resistência à podridão de espiga pelo método de modelo misto (MM).
..... 61
- Tabela 4 – Genes identificados próximos dos DArTs associados à resistência, distância em bp e respectiva função por meio do método *Bayesian Stochastic Search Variable* (BSSV). 62
- Tabela 5 – Genes identificados próximos dos DArTs associados à resistência, distância em bp e respectiva função por meio do método de modelo misto (MM). 63
- Tabela 6 – Número de componentes principais que devem ser incluídos no modelo para explicar adequadamente sobre a estrutura populacional, baseado no modelo de seleção BIC. 67
- Tabela 7 – Efeitos dos DArTs e frequência do menor alelo dos marcadores nos diferentes métodos..... 72
- Tabela 8 – Posição e função dos genes que foram localizados dentro da janela genômica em que os marcadores foram significativamente associados com a resistência..... 101

SUMÁRIO

| | |
|---|-----------|
| PRIMEIRA PARTE..... | 15 |
| 1. INTRODUÇÃO | 15 |
| 2. REFERENCIAL TEÓRICO | 18 |
| 2.1. Histórico das podridões | 18 |
| 2.2. Agentes patogênicos, vias de infecção, sintomas e controle | 18 |
| 2.3. Micotoxinas..... | 21 |
| 2.4. Controle genético e mapeamento de QTL para resistência à podridão de espiga por <i>Fusarium</i> | 22 |
| 2.5. Estudo de associação genômica | 26 |
| 2.6. Seleção genômica | 29 |
| 2.7. Modelos utilizados no estudo de associação genômica | 30 |
| REFERÊNCIAS..... | 34 |
| | |
| SEGUNDA PARTE – ARTIGO | 41 |
| 1. INTRODUÇÃO | 43 |
| 2. MATERIAL E MÉTODOS..... | 46 |
| 2.1. Experimentos de campo | 46 |
| 2.2. Cultura do patógeno, inoculação e fenotipagem..... | 46 |
| 2.3. Caracterização das linhagens e análise da estrutura populacional | 47 |
| 2.4. Análise estatística | 48 |
| 2.5. Análise de associação via modelos mistos | 48 |
| 2.6. Análise de associação via <i>Bayesian Stochastic Search Variable</i> ... | 49 |
| 2.7. Análise pós-MCMC | 51 |
| 2.8. Estimativas do desequilíbrio de ligação e construção dos <i>heatmaps</i> | 51 |
| 2.9. Validação dos genes associados à resistência à podridão de espiga | 52 |
| 3. RESULTADOS..... | 53 |
| 3.1. Análises de fenotípicas para resistência à podridão de espiga causada por <i>Fusarium verticillioides</i> | 53 |
| 3.2. Análise de associação da resistência à podridão de espiga causada por <i>Fusarium verticillioides</i> por meio do método <i>Bayesian Stochastic Search Variable</i> (BSSV) | 55 |

| | | |
|------|--|----|
| 3.3. | Análise de associação da resistência à podridão de espiga causada por <i>Fusarium verticillioides</i> por meio do método de modelo misto | 59 |
| 3.4. | Identificação de genes candidatos associados por meio do método <i>Bayesian Stochastic Search Variable</i> (BSSV)..... | 62 |
| 3.5. | Identificação de genes candidatos associados por meio do método de modelo misto | 65 |
| 3.6. | Desequilíbrio de ligação entre os marcadores | 66 |
| 4. | DISCUSSÃO | 73 |
| 4.1. | Herdabilidade e estudo de associação | 73 |
| 4.2. | Genes candidatos | 80 |
| 5. | CONCLUSÃO | 84 |
| | APÊNDICE | 85 |

PRIMEIRA PARTE

1. INTRODUÇÃO

As podridões de espiga são doenças já conhecidas há um bom tempo e passaram a ter uma maior importância com a descoberta de metabólitos secundários produzidos pelos patógenos, as micotoxinas, as quais causam diversas doenças graves em humanos e também em animais. Atualmente, as doenças de espigas causadas pelos fungos do gênero *Fusarium*, como *F. graminearum*, *F. verticillioides* e *F. proliferatum*, são conhecidas em todas as regiões produtoras de milho (MESTERHÁZY; LEMMENS; REID, 2012; WHITE, 1999).

Os principais meios de entrada do patógeno no hospedeiro são através de feridas ou danos causados por insetos ou pássaros. Uma vez que os esporos chegam até o estilo/estigma, normalmente carregados pelo vento, a infecção ocorre tanto na espiga quanto nos grãos. O fungo então promove o crescimento de um micélio cotonoso de coloração mais rosada a avermelhada e, quando afeta diretamente os grãos, promove o aparecimento de estrias brancas no pericarpo (MESTERHÁZY; LEMMENS; REID, 2012; WHITE, 1999). Além disso, micotoxinas, que são metabólitos secundários, são produzidas pelo patógeno. O fungo *Fusarium verticillioides* produz micotoxinas chamadas de fumonisinas, das quais foram identificados os tipos B₁, B₂, B₃ e B₄ (MUSSER; PLATTNER, 1997; PROCTOR et al., 2006). As micotoxinas afetam diretamente e indiretamente tanto humanos quanto animais.

O controle do patógeno deve ser realizado de uma forma integrada, que incluem medidas que atuam desde o manejo da cultura até o uso de cultivares resistentes. O controle dos patógenos por meio do uso de cultivares resistente é o

mais recomendado, pois é mais simples, possui baixo impacto ambiental e apresenta um baixo custo (ALI; TAYLOR; JIE, 2005; LÖFFLER et al., 2010).

Apesar do uso da resistência ser vantajoso, não existem cultivares totalmente resistentes no mercado, principalmente pelo fato da resistência à podridão de espiga causada por *Fusarium* ser controlada por vários genes e também ser muito influenciada pelo ambiente (CLEMENTS et al., 2004; DE LEÓN; PANDEY, 1989; KING; SCOTT, 1981; NANKAM; PATAKY, 1996). Vários estudos foram conduzidos com diferentes populações biparentais e foi verificado que o controle genético da resistência à podridão de espiga causada por *Fusarium verticillioides* possui tanto efeitos aditivos quanto efeitos de dominância. Muitos desses estudos apresentaram estimativas altas de herdabilidade para a resistência, o que facilitaria a seleção fenotípica. Por outro lado, a avaliação fenotípica muitas vezes apresenta algumas dificuldades práticas, devido principalmente à necessidade de inoculações artificiais (CLEMENTS et al., 2004; ELLER, 2009). Uma alternativa eficiente para contornar as dificuldades que ocorrem na avaliação fenotípica da doença seria a utilização da seleção assistida por marcadores, na qual é necessário o mapeamento de QTL. Embora alguns estudos encontraram QTL associados com a resistência à podridão de espiga em todos os cromossomos, exceto no cromossomo 9, esta ferramenta é bastante limitada, pois sua eficácia na seleção de caracteres quantitativos depende da precisão da localização, da estimação dos efeitos do QTL e toda a variação genética do caráter em avaliação é limitada apenas à diferença entre os pais ou poucos genitores (HOLLAND, 2004; SINGH; SINGH, 2015).

Como alternativa às limitações do mapeamento de QTL, em 1996, Risch e Merikangas sugeriram a utilização de estudo de associação do genoma como forma de detecção de genes relacionados a doenças complexas. Diferentemente do mapeamento de QTL, o estudo de associação genômica explora a recombinação histórica e evolutiva ao nível da população e permite uma maior

resolução de mapeamento. O estudo de associação explora o desequilíbrio de ligação entre os marcadores e os genes ou QTL de interesse, a fim de identificar associações (SINGH; SINGH, 2015). O desenvolvimento das matrizes de genotipagem possibilitou o mapeamento de diversas características no milho.

Os principais métodos estatísticos que vêm sendo propostos para os estudos de associação genômica são os modelos mistos e modelos de seleção genômica que utilizam a abordagem bayesiana. Nas análises que utilizam modelos mistos, cada marca é testada por vez e os efeitos fixos dessas marcas são estimados; posteriormente, o p-valor de cada marca é calculado para avaliação do efeito de cada marcador (HESLOT et al., 2015). No estudo de associação genômica que utiliza o método de seleção genômica com a abordagem bayesiana, o efeito de todos os marcadores é estimado conjuntamente todos de uma vez só, o que evita o problema de múltiplos testes e, também, considera a estrutura populacional. Esses métodos têm sido bastante eficazes na descrição da arquitetura genética de características que são poligênicas.

2. REFERENCIAL TEÓRICO

2.1. Histórico das podridões

As podridões de espiga são conhecidas há um bom tempo, os primeiros relatos foram feitos no ano de 1923, no Canadá (BISBY; BAILEY, 1923). Inicialmente, não apresentavam muita importância, pois não causavam grandes perdas e, dessa forma, poucos esforços no melhoramento para a resistência às podridões de espiga foram despendidos.

Posteriormente, com a descoberta de componentes secundários tóxicos (micotoxinas) produzidos pelos fungos, os quais causam perdas econômicas indiretas, as podridões de espiga se tornaram mais relevantes. Como consequência, muitos países passaram a adotar regulamentações nos limites admissíveis de micotoxinas em alimentos e rações (MESTERHÁZY; LEMMENS; REID, 2012).

Os primeiros patógenos causadores das podridões de espiga relatados foram os fungos do gênero *Fusarium* spp. As primeiras espécies identificadas como agentes patogênicos foram *Fusarium graminearum*, que causa a podridão branca de espiga (SUTTON; BALIKO, 1981) e *Fusarium verticillioides* [syn. *Fusarium moniliforme*], que causa a podridão rosada da espiga (JUGENHEIMER, 1976; ULLSLRUP, 1953)

Atualmente, as podridões de espiga causadas por fungos do gênero *Fusarium* são doenças de espiga mais comuns nas principais regiões produtoras de milho no mundo (MESTERHÁZY; LEMMENS; REID, 2012; WHITE, 1999).

2.2. Agentes patogênicos, vias de infecção, sintomas e controle

Os principais agentes patogênicos causadores da podridão de espiga do gênero *Fusarium* spp. são *Fusarium graminearum*, *Fusarium verticillioides* (syn.

F. moniliforme) e *Fusarium proliferatum*. Existem também outras espécies que ocorrem no milho e, muitas vezes, também desempenham papel importante, causando podridões de espiga (MESTERHÁZY; LEMMENS; REID, 2012). Folcher et al. (2009), encontraram 12 espécies diferentes de *Fusarium* na França, sendo que algumas espécies eram mais encontradas na região norte do país, como *F. graminearum* e *F. culmorum*, enquanto em regiões mais ao sul, *F. verticillioides* e *F. proliferatum* eram as espécies mais presentes. Resultados semelhantes foram encontrados na Hungria e nos Estados Unidos (MESTERHÁZY; LEMMENS; REID, 2012; REID et al., 1999)

A podridão de espiga por *Fusarium* ou podridão de espiga rosada, causada pelo fungo *Fusarium verticillioides* (Sacc.) (antigamente *F. moniliforme* Sheldon), é uma das doenças de espiga mais comuns em todas as regiões produtoras de milho do mundo, e ocorre principalmente em lugares que apresentam um clima quente e seco (ELLER; HOLLAND; PAYNE, 2008; MESTERHÁZY; LEMMENS; REID, 2012; PARSONS, 2008; WHITE, 1999).

Os principais meios de entrada do patógeno no hospedeiro são através de feridas ou danos causados por insetos, pássaros ou até mesmo por meio de intempéries do tempo, os quais oferecem um ponto de entrada para o patógeno (SUTTON, 1982). Outra forma de entrada do patógeno é por meio do vento, o qual carrega os esporos até o estilo/estigma, onde germinam e crescem infectando a espiga e os grãos (KOEHLER, 1942). O fungo *F. verticillioides* também é sistêmico, podendo infectar espigas por meio de colmos infectados (MUNKVOLD; MCGEE; CARLTON, 1997).

A infecção pode iniciar em qualquer parte da espiga ou individualmente nos grãos, estando sempre associada a alguma injúria. O fungo promove desenvolvimento de seu micélio, que é cotonoso com coloração rósea clara a avermelhada, e em algumas situações os grãos podem apresentar estrias de coloração branca no pericarpo (ELLER, 2009; MESTERHÁZY; LEMMENS;

REID, 2012; WHITE, 1999). Se a infecção ocorrer através do pedúnculo da espiga, todos os grãos podem ser infectados, mas a infecção só desenvolverá naqueles que apresentarem alguma injúria no pericarpo (EMBRAPA, 2010). O metabolismo do fungo dentro dos grãos resulta em uma diminuição líquida no teor de matéria seca, densidade e rendimento (PRESELLO et al., 2008).

Além do seu estado hemibiotrófico, o patógeno também pode se apresentar em um estado assintomático, ou seja, em um estado que não apresenta sintomas visíveis, na maioria dos ciclos de crescimento do milho e em sementes (BACON; GLENN; YATES, 2008; YATES; BACON; HINTON, 1997). Sem causar sintomas visíveis, o patógeno cresce intracelularmente e absorve nutrientes no apoplasto, causando injúrias mínimas na planta hospedeira. A doença com os sintomas visíveis ocorre quando ocorrem alguns estresses bióticos e abióticos, fazendo com o fungo adote seu estado patogênico, causando grandes injúrias (BACON; GLENN; YATES, 2008; BACON; HINTON, 1996).

Embora *F. verticilloides* seja frequentemente isolado de sementes, esta não é a principal fonte de inoculo do patógeno. Os conídios do fungo podem sobreviver em restos culturais, uma vez que este apresenta uma fase saprofítica ativa. Além disso, o fungo também sobrevive no solo e em outros hospedeiros (EMBRAPA, 2010).

Várias medidas devem ser tomadas de forma integrada para se ter um manejo eficiente da podridão de espiga, as quais são realizar rotação de culturas para reduzir o inóculo na área, controle de plantas daninhas que servem como hospedeiros, evitar plantios sucessivos de milho, utilizar sempre sementes saudáveis e com procedência, realizar o plantio na densidade recomendada, uso de cultivares resistentes e uso de fungicidas (EMBRAPA, 2010). O uso de fungicidas no controle da podridão de espiga ainda é motivo de várias dúvidas, pois não se tem conhecimento da eficiência, da época de aplicação e sua relação com a resistência

das cultivares; além disso, não existem fungicidas lançados recentemente para o controle do patógeno (EMBRAPA, 2010; LÖFFLER et al., 2010).

2.3. Micotoxinas

Os fungos do gênero *Fusarium* spp. produzem um grande número de micotoxinas (LOGRIECO et al., 2002). O fungo *Fusarium verticillioides* produz a micotoxina chamada de fumonisina, a qual é produzida durante a fase patogênica do fungo, mas, sabe-se também que, mesmo na fase assintomática, o fungo é capaz de produzir fumonisinas (BACON et al., 2001; BACON; GLENN; YATES, 2008). Diferentes tipos de fumonisinas foram identificadas, como B₁, B₂, B₃ e B₄. Dentre as micotoxinas sintetizadas por *F. verticillioides*, a fumonisina B₁ é a mais predominante e tóxica, e sua produção é bastante influenciada pelas condições do ambiente (MUSSER; PLATTNER, 1997; PROCTOR et al., 2006).

O consumo de grãos contaminados por fumonisinas está associado a várias doenças que afetam tanto humanos quanto animais, as quais incluem leucoencefalomalácia equina (KELLERMAN et al., 1990; MARASAS, 1996), edema pulmonar suíno (HARRISON et al., 1990), câncer de fígado em ratos (GELDERBLUM et al., 1988), defeitos do tubo neural em ratos (VOSS; GELINEAU-VAN WAES; RILEY, 2006), câncer de esôfago humano (RHEEDER et al., 1992; SYDENHAM et al., 1991) e retardo de crescimento em crianças (KIMANYA et al., 2010; SHIRIMA et al., 2013). Além de causar danos diretos e indiretos, esse fungo também pode causar danos na saúde dos trabalhadores que manipulam e processam os grãos (MESTERHÁZY; LEMMENS; REID, 2012).

Muito países adotaram regulamentações para limitar o máximo de fumonisina que pode estar presente nos alimentos, mas, muitas vezes, a fumonisina ainda é encontrada em milho moído e processado (ESCOBAR et al., 2013; SHEPHARD et al., 2013).

2.4. Controle genético e mapeamento de QTL para resistência à podridão de espiga por Fusarium

A resistência à podridão de espiga causada por *Fusarium* é controlada por vários genes e é muito influenciada pelo ambiente (CLEMENTS et al., 2004; DE LEÓN; PANDEY, 1989; KING; SCOTT, 1981; NANKAM; PATAKY, 1996). Nankam e Pataky (1996) avaliaram o comportamento de duas populações (IL125b x Ia 2256a e IL125b x IL783a) quanto à resistência à podridão de espiga e encontraram efeitos aditivos, de dominância e epistáticos significativos. Sendo que, para o cruzamento entre as linhagens IL125b e Ia 2256a, os efeitos de maior magnitude foram o aditivo e o de dominância. Para o cruzamento entre as linhagens IL125b e IL783a, somente o efeito aditivo foi de maior magnitude. Nesse estudo, as estimativas de herdabilidade variaram de 32% - 55% para infecção sintomática e de 24% - 46% para infecção assintomática.

Em um estudo com duas populações, Perez-Brito et al. (2001) encontraram tanto efeito aditivo quanto efeito de dominância no controle genético da resistência e não relataram sobre efeitos epistáticos no trabalho. As estimativas de herdabilidade encontradas por esses autores foi de 42% para a população 1 (linhagem 3 x linhagem 18) e de 26% para a população 2 (linhagem 5 x linhagem 18). No trabalho realizado por Robertson-Hoyt et al. (2006), não foi possível determinar se o controle genético era predominantemente aditivo ou de dominância, porque a população NCB era composta de linhagens recombinantes quase homozigotas e a população GEFR era composta por famílias derivadas de geração de primeiro retrocruzamento, em que os efeitos dominantes e aditivos de pai doador alelos são confundidos. As estimativas de herdabilidade encontrada por esses autores foram maiores que as herdabilidades encontradas em estudos anteriores. Para a população GEFR, a estimativa foi de 47% e, para a população NCB, a estimativa foi de 86%.

Em um estudo em dois locais e em anos diferentes, Ding et al., (2008), encontraram altas estimativas de herdabilidade para resistência à podridão de espiga, as quais variam entre 88% e 94% nas análises individuais e 85% e 86% para as análises conjuntas. Li et al. (2011), em seu estudo, afirmaram que o controle genético da resistência é aditivo e, no mesmo estudo, também encontraram estimativas de herdabilidade altas, que variaram entre 89% e 94% nas análises individuais e de 84% na análise conjunta. E, em um estudo posterior, Chen et al. (2012) encontraram que o controle genético era tanto aditivo quanto de dominância ou somente de efeito aditivo. As estimativas de herdabilidade encontradas por esses autores foram de menor magnitude quando comparadas com as estimativas de trabalhos anteriores, mas ainda foram estimativas altas. A herdabilidade para a resistência variou entre 80% e 88% para as análises individuais e foi de 65% na análise conjunta.

Apesar da herdabilidade para a resistência ser de alta magnitude, a avaliação fenotípica muitas vezes apresenta algumas dificuldades práticas, uma vez que a avaliação da doença somente é realizada no final do ciclo da cultura. Para a obtenção de resultados satisfatórios, existe a necessidade de uma ou duas inoculações artificiais, e também ocorre uma grande influência do ambiente no desenvolvimento da doença (CLEMENTS et al., 2004; ELLER, 2009).

Uma alternativa eficiente para contornar as dificuldades que ocorrem na avaliação fenotípica da doença seria a utilização da seleção assistida por marcadores. Para se ter sucesso na identificação e mapeamento de QTL, é necessário a criação de uma população de mapeamento por meio de cruzamentos biparentais, nos quais os pais devem ser divergentes para a característica em avaliação (ROBERTSON-HOYT et al., 2006; SINGH; SINGH, 2015). Posteriormente, a população de mapeamento é avaliada fenotipicamente em vários locais e anos, e os dois pais são testados com um grande número de marcadores, que cobrem todo o genoma. Depois, os indivíduos da população de

mapeamento são genotipados, utilizando os marcadores polimórficos previamente identificados em ambos o país. Os dados genotípicos são utilizados para construir um mapa de estrutura de ligação para a população, e, por fim, os dados genotípicos e os dados fenotípicos são analisados para detectar associação entre os marcadores e os QTL (SINGH; SINGH, 2015).

Uma das primeiras estratégias para estudar a arquitetura genética da resistência à podridão de espiga causado pelo fungo *Fusarium verticillioides* foi a utilização de populações biparentais. A seleção fenotípica juntamente com a identificação de QTL e o uso da seleção assistida por marcadores têm sido bastante utilizadas em programas de melhoramento, uma vez que essas ferramentas permitem estratégias mais eficientes de seleção (HOLLAND, 2004; ROBERTSON-HOYT et al., 2006).

Em um estudo para identificar QTL de resistência à podridão de espiga por *Fusarium*, Perez-Brito et al. (2001) utilizaram duas populações distintas, sendo as linhagens resistentes consideradas como fêmeas e uma linhagem suscetível em comum considerada como macho. Na população 1, encontraram 9 QTL, distribuídos nos cromossomos 1,2,3,4,6,7 e 10, já na população 2, foram identificados 7 QTL nos cromossomos 1,3,4,5,6 e 7. Somente os QTL identificados nos cromossomos 3 e 6 foram comuns nas populações, decorrente do uso de diferentes linhagens nos cruzamentos. Além disso, nesse estudo, houve um grande efeito do ambiente na expressão da resistência.

Em outro estudo visando identificar QTL para resistência à podridão de espiga e à contaminação por fumonisina em duas populações diferentes, Robertson-Hoyt et al. (2006) identificaram 7 QTL de resistência na população derivada da linhagem resistente GE440, os quais explicaram 47% da variância fenotípica e 99% da variância genética. Já na população originada da linhagem resistente NCB, foram identificados 5 QTL para a resistência à podridão de espiga, que explicaram 31% da variância fenotípica e 36% da variância genética.

Os resultados encontrados por Robertson-Hoyt et al. (2006) foram consistentes com os resultados encontrados por Perez-Brito et al. (2001), nos quais, em ambos os estudos, foram identificados QTL de pequeno efeito.

Ding et al. (2008) encontraram diferentes QTL associados à resistência em diferentes locais e anos. Nas análises individuais, foram identificados QTL para resistência nos cromossomos 3, 5, 8 e 10. Na análise conjunta, foram identificados 6 QTL, todos com efeito aditivo significativo e apenas os QTL identificados nos cromossomos 5, 8 e 10 apresentaram efeito de interação. O QTL identificado no cromossomo 3 foi o único que apresentou um efeito consistente, porque foi identificado nas análises individuais e conjunta. Além disso, esse QTL apresentou o maior efeito, o qual explicou 13 - 22% da variância fenotípica nas análises individuais e 25% da variância fenotípica na análise conjunta. Em comparação com os estudos anteriores realizados por Perez-Brito et al. (2001), houveram algumas consistências nas posições dos bins. Apesar disso, apenas o QTL identificado no cromossomo 3 foi consistente nos dois estudos, mas uma comparação melhor não poderia ser feita porque foram utilizados diferentes marcadores. Em comparação com o estudo feito por Robertson-Hoyt et al. (2006), não houve nenhum QTL identificado nas mesmas posições.

Li et al. (2011) identificaram 4 QTL de resistência nos cromossomos 3, 4, 5, e 6, os quais explicaram 22% da variância fenotípica, e a variância genótipo x ambiente explicou apenas 3% da variância fenotípica. Apesar de encontrarem alguns QTL nos mesmos cromossomos que estudos anteriores, nenhum foi identificado nas mesmas posições descritas pelos outros autores.

Em um estudo posterior, Chen et al. (2012), utilizando uma população de mapeamento diferente, mas com o mesmo parental resistente utilizado por Li et al. (2011), encontraram 3 QTL associados à resistência localizados nos cromossomos 4, 5 e 10. Os QTL identificados nos cromossomos 4 e 10 apresentaram tanto efeito aditivo quanto efeito de dominância, e o QTL

identificado no cromossomo 5 apresentou apenas efeito aditivo. No total, os efeitos aditivos explicaram 38% da variância fenotípica e os efeitos de dominância explicaram apenas 7% da variância fenotípica. Dos QTL identificados em estudos anteriores, houve somente coincidência do QTL identificado no cromossomo 4 com o estudo de Li et al. (2011). Sendo provavelmente o mesmo QTL, pois, nos dois estudos, a linhagem BT-1 foi utilizada como parental resistente.

Apesar da seleção assistida por marcadores ser uma boa ferramenta para o melhoramento, sua eficácia na seleção de caracteres quantitativos depende da precisão da localização e estimação dos efeitos do QTL, sendo que, para caracteres de baixa herdabilidade, a detecção e estimação são mais difíceis (BEAVIS, 1997; HOLLAND, 2004).

Apesar do mapeamento de QTL apresentar algumas vantagens, existem alguns fatores que são limitantes. Como a população de mapeamento é obtida a partir de um cruzamento biparental, toda a variação genética do caráter em avaliação é limitada apenas à diferença entre os pais, e a resolução do mapeamento também é baixa. No mapeamento, mesmo que seja identificado um QTL de grande efeito, é difícil identificar o gene que é responsável por este. Além disso, a criação de uma população de mapeamento requer esforço, tempo e, muitas vezes, não é aplicável para algumas espécies. Um QTL detectado em uma população não é necessariamente efetivo em outra população e o mesmo ocorre para os marcadores detectados, sendo necessário a validação em outras populações (SINGH; SINGH, 2015).

2.5. Estudo de associação genômica

A eficácia do mapeamento de QTL apresenta algumas limitações, uma vez que é bastante difícil detectar QTL de pequenos efeitos para uma dada característica em cruzamentos biparentais. Além disso, os eventos de recombinação em populações de linhagens recombinantes normalmente são

insuficientes para um mapeamento fino (HOLLAND, 2007; KORTE; FARLOW, 2013).

Uma alternativa para contornar as limitações do mapeamento de QTL e seleção assistida por marcadores é a utilização do estudo de associação genômica, que, apesar de se basear nos mesmos princípios da recombinação genética do mapeamento de QTL, esse analisa a herança compartilhada de vários indivíduos e que muitas vezes não têm a ascendência avaliada ou observada. Essencialmente, o estudo de associação genômica explora a recombinação histórica e evolutiva ao nível da população e, além disso, oferece uma maior amostragem dos alelos e também uma maior resolução de mapeamento, porque os marcadores que não estão intimamente ligados com os loci envolvidos na expressão da característica não geram sinais que são captados no estudo de associação genômica (YU; BUCKLER, 2006; ZHU et al., 2008).

De uma forma geral, o estudo de associação genômica utiliza o desequilíbrio de ligação entre os marcadores e os genes ou QTL de interesse para identificar associações, e isso é possível por meio da análise de desequilíbrio de ligação que utiliza todos os eventos de recombinação que teriam ocorrido entre o gene e o marcador no passado recente na população e que pode ser utilizado no estudo de associação genômica (SINGH; SINGH, 2015).

Duas estratégias podem ser utilizadas no estudo de associação genômica, uma é o mapeamento de genes candidatos e a outra é o mapeamento completo do genoma (GWAS – *Genome Wide Association Study*). No estudo de associação de genes candidatos, a análise é restrita à região genômica que possui o gene candidato ou o QTL para a característica de interesse. Um gene candidato é aquele gene que, com base no conhecimento prévio, sabe-se estar envolvido no controle do caráter de interesse, e as ferramentas utilizadas para a identificação de genes candidatos são genômica comparativa, anotação sequência do genoma, perfil de transcrição e análise de QTL (SINGH; SINGH, 2015). Uma limitação dessa

abordagem é que, algumas vezes, genes relacionados com a característica de interesse não são relacionados na lista de genes candidatos e, conseqüentemente, não podem ser avaliados no mapeamento. A seleção de genes candidatos é simples para rotas bioquímicas simples e rotas já bem caracterizadas, mas, para características mais complexas que envolvem bastantes genes, é mais difícil fazer a seleção de genes candidatos, pois todo o genoma serviria com candidato (YU; BUCKLER, 2006; ZHU et al., 2008).

A outra estratégia utilizada é o estudo de associação completo do genoma (GWAS), no qual os marcadores utilizados para a genotipagem são distribuídos, de uma forma densa e uniforme, ao longo de todo o genoma. Nessa abordagem, todos os loci envolvidos na expressão da característica podem ser avaliados (SINGH; SINGH, 2015). Uma questão importante que se deve levar em conta quando se realiza o GWAS é a densidade de marcadores a serem utilizados e o fator que determina isso é o decaimento do desequilíbrio de ligação sobre a distância física em uma população. Se o desequilíbrio de ligação da população em estudo decai rapidamente, serão necessários mais marcadores para que se cubra densamente o genoma para capturar os marcadores localizados perto o suficiente das regiões em que a característica de interesse se expressa (YU; BUCKLER, 2006).

A disponibilidade de uma sequência de genoma de referência para uma espécie permitiu o resequenciamento mais fácil dos indivíduos e o desenvolvimento de plataformas de genotipagem mais rentáveis (ZILA, 2014). Para o milho, a sequência do genoma de referência com base na linhagem B73 tornou-se disponível (SCHNABLE et al., 2009) e permitiu o desenvolvimento do primeiro HapMap milho com base na variação descoberta em 26 linhagens (GORE et al., 2009), e também possibilitou o desenvolvimento de uma matriz de genotipagem Illumina Infinium SNP para o milho, com quase 50.000 marcadores SNP (*Single Nucleotide Polymorphism*) (GANAL et al., 2011).

O desenvolvimento dessas matrizes de genotipagem possibilitou o mapeamento de diversas características no milho, as quais incluem época de florescimento, altura de planta e espiga (REMINGTON et al., 2001), coloração do endosperma (PALAISA et al., 2003), composição do grão (WILSON et al., 2004), época de florescimento e altura de planta (ANDERSEN et al., 2005), acúmulo de maysin e ácido clorogênico (SZALMA et al., 2005), fluxo em vias de biossíntese de carotenoides (HARJES et al., 2008), época de florescimento (BUCKLER et al., 2009), composição do grão (COOK et al., 2012) resposta de hipersensibilidade (OLUKOLU et al., 2013) e genes associados à resistência à podridão de espiga causada por *Fusarium verticillioides* (ZILA et al., 2013, 2014).

2.6. Seleção genômica

Diferentemente da análise de associação que mede a relação entre o marcador e o fenótipo ou QTL, a seleção genômica (*Genome Wide Selection – GWS*) compreende métodos que utilizam dados genotípicos para prever qualquer característica com uma precisão suficiente para permitir a seleção baseada somente nos dados preditos. A seleção genômica estima o efeito dos marcadores que cobrem todo o genoma para calcular os valores genéticos genômicos, ou seja, a seleção genômica busca explicar a variância genética total pela soma dos efeitos dos marcadores que cobrem todo o genoma para prever o valor genéticos dos indivíduos (DESTA; ORTIZ, 2014; MEUWISSEN; HAYES; GODDARD, 2001).

A principal etapa da seleção genômica é o cálculo dos valores genéticos genômicos estimados (*Genomic Estimated Breeding Values – GEBVs*) para os indivíduos que apresentam apenas os dados genotípicos usando um modelo que foi treinado com os indivíduos que apresentavam dados fenotípicos e genotípicos

(população de treinamento) (MEUWISSEN; HAYES; GODDARD, 2001). Posteriormente, os valores genéticos genômicos estimados são utilizados para o avanço no ciclo de melhoramento. Portanto, a seleção de um indivíduo sem os dados fenotípicos pode ser realizada utilizando-se um modelo para prever valores genéticos do indivíduo (HEFFNER; SORRELLS; JANNINK, 2009; MEUWISSEN; HAYES; GODDARD, 2001).

2.7. Modelos utilizados no estudo de associação genômica

Diferentes métodos estatísticos têm sido propostos para conduzir estudos de associação genômica, sendo esses dependentes da arquitetura genética estudada. Geralmente, têm sido utilizados modelos infinitesimais, os quais apresentam um bom poder preditivo e descrevem adequadamente a arquitetura genética (HESLOT et al., 2015). Nesse tipo de modelo, uma marca é testada por vez, e os efeitos fixos das marcas são estimados e, posteriormente, calcula-se o p-valor associado ao efeito de cada marcador.

Muitas vezes, a população com que se deseja realizar o estudo de associação apresenta uma estrutura populacional, o que pode fazer com que um locus não apresente qualquer relação ou efeito sobre a característica avaliada, porém apresente associações significativas, o que pode aumentar a taxa de erro do tipo I. Para controlar a taxa de falsos positivos causados pela ausência de estrutura populacional (erro tipo I), alguns métodos são utilizados, como associação estruturada (*Structured Association – SA*) e o controle genômico (*Genomic Control – GC*) (YU et al., 2006).

A análise com a associação estruturada (SA) utiliza marcadores aleatórios para estimar a estrutura populacional (Q), a qual é posteriormente incorporada na análise estatística. Por outro lado, na análise que utiliza o controle genômico (GC), um conjunto de marcadores aleatórios é utilizado para estimar e ajustar a inflação dos testes estatísticos causados pela estrutura da população, assumindo que

estrutura tem um efeito semelhante em todos os loci (YU et al., 2006; ZHU et al., 2008).

Apesar disso, para esse tipo de população, quando se leva em conta apenas a estrutura populacional, pode haver um controle inadequado dos falsos positivos ou perda do poder estatístico devido ao parentesco dentro da população. Diante disso, Yu et al. (2006) propuseram uma nova abordagem para o mapeamento associativo, a qual integra ferramentas genômicas com o modelo misto tradicional para descobrir a estrutura populacional e o parentesco.

A equação de modelo misto proposta por Yu et al. (2006) é expressa por

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{S}\boldsymbol{\alpha} + \mathbf{Q}\mathbf{v} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

em que \mathbf{y} é o vetor das observações fenotípicas; $\boldsymbol{\beta}$ é um vetor dos efeitos fixos não relacionados com os marcadores ou com a estrutura populacional; $\boldsymbol{\alpha}$ é um vetor dos efeitos dos marcadores; \mathbf{v} é um vetor dos efeitos populacionais; \mathbf{u} é um vetor dos efeitos poligênicos; \mathbf{e} é o vetor dos efeitos residuais; \mathbf{Q} é a matriz da estrutura populacional; e \mathbf{X} , \mathbf{S} e \mathbf{Z} são matrizes de incidência de 1s e 0s relacionadas a $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ e \mathbf{u} , respectivamente.

Nos métodos baseados em modelos lineares mistos, a estrutura da população se enquadra como um efeito fixo, ao passo que o parentesco entre os indivíduos é incorporado como a estrutura de variância-covariância do efeito aleatório para os indivíduos. Estudos de associação genômica exigem grandes amostras de indivíduos para obter poder estatístico suficiente para detectar alelos de pequenos efeitos, os quais normalmente estão relacionados a características complexas. Quando o conjunto de dados é muito grande, surgem problemas computacionais, pois o tempo para resolver um modelo linear misto é elevado ao cubo do número de indivíduos que se encaixam como um efeito aleatório. Além disso, o tempo computacional aumenta devido ao processo de iteração que é

necessário para estimar os parâmetros populacionais, como os componentes da variância, para cada marcador (ZHANG et al., 2010). Como uma alternativa para a redução do tempo computacional, Zhang et al. (2010) propuseram o modelo linear misto compresso, o qual reduz o tamanho dos efeitos aleatórios. Nessa metodologia, os indivíduos são alocados em grupos baseados no parentesco entre os indivíduos. Assim, o parentesco entre pares de grupos substitui o parentesco entre pares de indivíduos para o efeito aleatório de um modelo linear misto, e consequentemente, ocorre a redução do tempo computacional.

Além disso, estes autores desenvolveram um algoritmo chamado P3D (*Population Parameters Previously Determined*), com o qual não há necessidade das iterações para estimar os parâmetros populacionais, como a variância genética e a variância residual, para cada marcador. Dessa forma, as estimativas dos componentes da variância são obtidas uma única vez pela máxima verossimilhança residual (REML) (ZHANG et al., 2010).

Os modelos frequentistas, apesar do uso generalizado, possuem certas limitações, pois, normalmente, para verificar a existência da associação entre o marcador e o fenótipo de interesse, esses métodos se baseiam no cálculo do p-valor e verificação da hipótese nula (H_0) de não associação. Segundo Stephens e Balding (2009), é difícil quantificar o quanto uma associação entre o marcador e o fenótipo é confiável baseando-se apenas no p-valor. O mesmo p-valor calculado em diferentes marcadores ou em diferentes estudos pode ter diferentes implicações nos estudos de associação, sendo que a frequência do menor alelo (*Minor Allele Frequency* – MAF) e o tamanho da população em estudo podem afetar o poder do teste.

Outro método estatístico proposto para conduzir estudos de associação genômica são os modelos que utilizam uma abordagem bayesiana. Esses métodos podem ser mais eficazes na descrição da arquitetura genética quando a característica é poligênica, porque apresentam perfil poligênico e alta resolução

na identificação de genes com efeitos maiores (GIANOLA, 2013; HABIER et al., 2007). Apesar disso, os métodos bayesianos vêm sendo pouco utilizados nos estudos de associação genômica.

Nesse sentido, alguns métodos têm sido propostos com o objetivo de estimar o efeito de todos os marcadores conjuntamente e considerando a estrutura populacional.

REFERÊNCIAS

ALI, M. L.; TAYLOR, J. H.; JIE, L. Molecular mapping of QTLs for resistance to Gibberella ear rot, in corn, caused by *Fusarium graminearum*. **Genome**, v. 48, n. 3, p. 521–533, 2005.

ANDERSEN, J. R. et al. Validation of Dwarf8 polymorphisms associated with flowering time in elite European inbred lines of maize (*Zea mays* L.). **Theoretical and Applied Genetics**, v. 111, n. 2, p. 206–217, 2005.

BACON, C. W. et al. Biological control of *Fusarium moniliforme* in maize. **Environmental health perspectives**, v. 109 Suppl. 2, n. May, p. 325–332, 2001.

BACON, C. W.; GLENN, A. E.; YATES, I. E. Managing the Endophytic Association With Maize for Reduced Fumonisin Accumulation. **Toxin Reviews**, v. 27, n. 3–4, p. 411–446, 2008.

BACON, C. W.; HINTON, D. M. Symptomless endophytic colonization of maize by *Fusarium moniliforme*. **Canadian Journal of Botany**, v. 74, n. 8, p. 1195–1202, 1996.

BEAVIS, W. QTL Analyses. In: PATERSON, A. H. (Ed.). *Molecular Dissection of Complex Traits*. **CRC Press.**, 1997. p. 146–162.

BISBY, G. R.; BAILEY, D. L. Ear and root rots. **Fourth Annual Report of the Survey of the Prevalence of Plant Diseases in the Dominion of Canada**. Ottawa, Ontario, Canada, 1923.

BUCKLER, E. S. et al. The Genetic Architecture of Maize. **Science**, v. 325, n. 2009, p. 714–718, 2009.

CHEN, J. et al. Detection and verification of quantitative trait loci for resistance to *Fusarium* ear rot in maize. **Molecular Breeding**, v. 30, n. 4, p. 1649–1656, 2012.

CLEMENTS, M. J. et al. Sources of resistance to fumonisin accumulation in grain and fusarium ear and kernel rot of corn. **Phytopathology**, v. 94, n. 3, p. 251–260, 2004.

COOK, J. P. et al. Genetic Architecture of Maize Kernel Composition in the Nested Association Mapping and Inbred Association Panels. **Plant Physiology**,

v. 158, n. 2, p. 824–834, 2012.

DE LEÓN, C.; PANDEY, S. Improvement of Resistance to Ear and Stalk Rots and Agronomic Traits in Tropical Maize Gene Pools. **Crop Science**, v. 29, n. 1, p. 12–17, 1989.

DESTA, Z. A.; ORTIZ, R. Genomic selection: genome-wide prediction in plant improvement. **Trends in Plant Science**, v. 19, n. 9, p. 592–601, 2014.

DING, J. Q. et al. QTL mapping of resistance to Fusarium ear rot using a RIL population in maize. **Molecular Breeding**, v. 22, n. 3, p. 395–403, 2008.

ELLER, M. S. **Improving Resistance to Fusarium Ear Rot and Fumonisin Contamination and Increasing Yield with Exotic Maize Germplasm**. 2009. 176 p. Doctoral dissertation (Crop Science) – North Carolina State University, Raleigh, North Carolina, USA, 2009.

ELLER, M. S.; HOLLAND, J. B.; PAYNE, G. A. Breeding for Improved Resistance To Fumonisin Contamination in Maize. **Toxin Reviews**, v. 27, n. 3–4, p. 371–389, 2008.

EMBRAPA. **O Cultivo do Milho: Doenças**. Disponível em: <http://www.cnpms.embrapa.br/publicacoes/milho_6_ed/doencas.htm>. Acesso em: 2 jul. 2015.

ESCOBAR, J. et al. Occurrence and exposure assessment of Fusarium mycotoxins in maize germ, refined corn oil and margarine. **Food and Chemical Toxicology**, v. 62, p. 514–520, 2013.

FOLCHER, L. et al. Comparative activity of agrochemical treatments on mycotoxin levels with regard to corn borers and Fusarium mycoflora in maize (*Zea mays* L.) fields. **Crop Protection**, v. 28, n. 4, p. 302–308, 2009.

GANAL, M. W. et al. A large maize (*zea mays* L.) SNP genotyping array: Development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. **PLoS ONE**, v. 6, n. 12, 2011.

GELDERBLUM, W. C. et al. Fumonisin - novel mycotoxins with cancer-promoting activity produced by *Fusarium verticillioides*. **Applied Environmental Microbiology**, v. 54, n. 7, p. 1806–1811, 1988.

GIANOLA, D. Priors in whole-genome regression: The Bayesian alphabet

returns. **Genetics**, v. 194, n. 3, p. 573–596, 2013.

GORE, M. A. et al. A First-Generation Haplotype Map of Maize. **Science**, v. 326, n. 5956, p. 1115–1117, 2009.

HABIER, D.; FERNANDO, R. L.; DEKKERS, J. C. M. The impact of genetic relationship information on genome-assisted breeding values. **Genetics**, v. 177, n. 4, p. 2389–2397, 2007.

HARJES, C. E. et al. Natural Genetic Variation in Lycopene Epsilon Cyclase Tapped for Maize Biofortification. **Science**, v. 319, n. 5861, p. 330–333, 2008.

HARRISON, L. R. et al. Pulmonary Edema and Hydrothorax in Swine Produced by Fumonisin B1, a Toxic Metabolite of *Fusarium Moniliforme*. **Journal of Veterinary Diagnostic Investigation**, v. 2, n. 3, p. 217–221, 1990.

HEFFNER, E. L.; SORRELLS, M. E.; JANNINK, J.-L. Genomic Selection for Crop Improvement. **Crop Science**, v. 49, n. 1, p. 1–12, 2009.

HESLOT, N.; JANNINK, J.-L.; SORRELLS, M. E. Perspectives for Genomic Selection Applications and Research in Plants. **Crop Science**, v. 55, p. 1–12, 2015.

HOLLAND, J. B. Implementation of molecular markers for quantitative traits in breeding programs - challenges and opportunities. **Proceedings of the 4th International Crop Science Congress**, p. 1–13, 2004.

HOLLAND, J. B. Genetic architecture of complex traits in plants. **Current Opinion in Plant Biology**, v. 10, n. 2, p. 156–161, 2007.

JUGENHEIMER, R. W. **Corn: Improvement, seed production and uses**. New York: John Wiley & Sons Inc, 1976. 670 p.

KELLERMAN, T. S. et al. Leukoencephalomalacia in two horses induced by oral dosing of fumonisin B1. **The Onderstepoort journal of veterinary research**, v. 57, n. 4, p. 269–275, 1990.

KIMANYA, M. E. et al. Fumonisin exposure through maize in complementary foods is inversely associated with linear growth of infants in Tanzania. **Molecular Nutrition & Food Research**, v. 54, n. 11, p. 1659–1667, 2010.

KING, S. B.; SCOTT, G. E. Genotypic differences in maize to kernel infection by

Fusarium moniliforme. **Phytopathology**, v. 71, p. 1245–1247, 1981.

KOEHLER, B. Natural mode of entrance of fungi into corn ears and some symptoms that indicate infection. **J. Agric. Res.**, n. 64, p. 421–442, 1942.

KORTE, A.; FARLOW, A. The advantages and limitations of trait analysis with GWAS: a review. **Plant methods**, v. 9, n. 1, p. 29, 2013.

LI, Z. M. et al. A new QTL for resistance to Fusarium ear rot in maize. **Journal of Applied Genetics**, v. 52, n. 4, p. 403–406, 2011.

LÖFFLER, M. et al. Population parameters for resistance to Fusarium graminearum and Fusarium verticillioides ear rot among large sets of early, mid-late and late maturing European maize (*Zea mays* L.) inbred lines. **Theoretical and Applied Genetics**, v. 120, n. 5, p. 1053–1062, 2010.

LOGRIECO, A. et al. Toxigenic Fusarium species and mycotoxins associated with maize ear rot in Europe. **European Journal of Plant Pathology**, v. 108, n. 7, p. 597–609, 2002.

MARASAS, W. F. O. Fumonisin: History, World-Wide occurrence and impact. In: JACKSON, L. S.; DEVRIES, J. W.; BULLERMAN, L. B. (Eds.). **Fumonisin in Food. Advances in Experimental medicine and Biology**. Boston, MA: Springer US, 1996. v. 392p. 1–17.

MESTERHÁZY, Á.; LEMMENS, M.; REID, L. M. Breeding for resistance to ear rots caused by Fusarium spp. in maize - A review. **Plant Breeding**, v. 131, n. 1, p. 1–19, 2012.

MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, n. 4, p. 1819–1829, 2001.

MUNKVOLD, G. P.; MCGEE, D. C.; CARLTON, W. M. Importance of Different Pathways for Maize Kernel Infection by Fusarium moniliforme. **Phytopathology**, v. 87, n. 2, p. 209–217, 1997.

MUSSER, S. M.; PLATTNER, R. D. Fumonisin Composition in Cultures of Fusarium moniliforme, Fusarium proliferatum, and Fusarium nygami. **Journal of Agricultural and Food Chemistry**, v. 45, n. 4, p. 1169–1173, 1997.

NANKAM, C.; PATAKY, J. K. Resistance to kernel infection by Fusarium

moniliforme in the sweet corn inbred IL125b. **Plant Disease**, v. 80, n. 5, p. 593–598, 1996.

OLUKOLU, B. A. et al. A connected set of genes associated with programmed cell death implicated in controlling the hypersensitive response in maize. **Genetics**, v. 193, n. 2, p. 609–620, 2013.

PALAIISA, K. A et al. Applications of linkage disequilibrium and association mapping in crop plants. **The Plant Cell**, v. 15, n. August, p. 1795–1806, 2003.

PARSONS, M. W. **Biotic and abiotic factors associated with Fusarium ear rot of maize caused by Fusarium verticillioides**. 2008. 170 p. Doctoral dissertation (Plant Pathology and Microbiology) – Iowa State University, Ames, Iowa, USA 2008.

PEREZ-BRITO, D. et al. QTL mapping of Fusarium moniliforme ear rot resistance in high land maize, Mexico. **Agrociencia**, v. 35, p. 181–196, 2001.

PRESELLO, D. A. et al. Effect of disease severity on yield and grain fumonisin concentration of maize hybrids inoculated with Fusarium verticillioides. **Crop Protection**, v. 27, n. 3–5, p. 572–576, 2008.

PROCTOR, R. H. et al. Fumonisin Production in the Maize Pathogen Fusarium verticillioides: Genetic Basis of Naturally Occurring Chemical Variation. **Journal of Agricultural and Food Chemistry**, v. 54, n. 6, p. 2424–2430, 2006.

REID, L. M. et al. Interaction of Fusarium graminearum and F. moniliforme in Maize Ears: Disease Progress, Fungal Biomass, and Mycotoxin Accumulation. **Phytopathology**, v. 89, n. 11, p. 1028–1037, 1999.

REMYINGTON, D. L. et al. Structure of linkage disequilibrium and phenotypic associations in the maize genome. **Proceedings of the National Academy of Sciences of the United States of America**, v. 98, n. 20, p. 11479–11484, 2001.

RHEEDER, J. P. et al. Fusarium moniliforme and fumonisins in corn in relation to human esophageal cancer in Transkei. **The American Phytopathological Society**, v. 82, n. 3, p. 353 – 357, 1992.

RISCH, N.; MERIKANGAS, K. The future of genetic studies of complex human diseases. **Science**, v. 273, n. 5281, p. 1516–1517, 1996.

ROBERTSON-HOYT, L. A. et al. QTL mapping for fusarium ear rot and

fumonisin contamination resistance in two maize populations. **Crop Science**, v. 46, n. 4, p. 1734–1743, 2006.

SCHNABLE, P. S. et al. The B73 Maize Genome: Complexity, Diversity, and Dynamics. **Science**, v. 326, n. 5956, p. 1112–1115, 2009.

SHEPHARD, G. S. et al. The risk management dilemma for fumonisin mycotoxins. **Food Control**, v. 34, n. 2, p. 596–600, 2013.

SHIRIMA, C. P. et al. Dietary exposure to aflatoxin and fumonisin among Tanzanian children as determined using biomarkers of exposure. **Molecular Nutrition & Food Research**, v. 57, n. 10, p. 1874–1881, 2013.

SINGH, B. D.; SINGH, A. K. **Marker-Assisted Plant Breeding: Principles and Practices**. New Delhi: Springer India, 2015. 514p.

STEPHENS, M.; BALDING, D. J. Bayesian statistical methods for genetic association studies. **Nature Reviews Genetics**, v. 10, n. 10, p. 681–690, 2009.

SUTTON, J. C. Epidemiology of wheat head blight and maize ear rot caused by *Fusarium graminearum*. **Canadian Journal of Plant Pathology**, v. 4, n. 2, p. 195–209, 1982.

SUTTON, J. C.; BALIKO, W. Methods for quantifying partial resistance to Gibberellae in maize ears. **Canadian Journal of Plant Pathology**, v. 3, n. 1, p. 26–32, 1981.

SYDENHAM, E. W. et al. Fumonisin contamination of commercial corn-based human foodstuffs. **Journal of Agricultural and Food Chemistry**, v. 39, n. 11, p. 2014–2018, 1991.

SZALMA, S. J. et al. Association analysis of candidate genes for maysin and chlorogenic acid accumulation in maize silks. **Theoretical and Applied Genetics**, v. 110, n. 7, p. 1324–1333, 2005.

ULLSLRUP, A. J. Several Ear Rots of Corn. **Yearbook of Agriculture**. p. 390–392, 1953.

VOSS, K. A.; GELINEAU-VAN WAES, J. B.; RILEY, R. T. Fumonisin: current research trends in developmental toxicology. **Mycotoxin Research**, v. 22, n. 1, p. 61–69, 2006.

WHITE, D. G. **Compendium of Corn Diseases**. Third Edition. Ed. Amer Phytopathological Society, 1999. 78p.

WILSON, L. M. et al. Dissection of Maize Kernel Composition and Starch Production by Candidate Gene Association. **The Plant Cell**, v. 16, n. October, p. 2719–2733, 2004.

YATES, I. E.; BACON, C. W.; HINTON, D. M. Effects of Endophytic Infection by *Fusarium moniliforme* on Corn Growth and Cellular Morphology. **Plant Disease**, v. 81, n. 7, p. 723–728, 1997.

YU, J. et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. **Nature Genetics**, v. 38, n. 2, p. 203–208, 2006.

YU, J.; BUCKLER, E. S. Genetic association mapping and genome organization of maize. **Current Opinion in Biotechnology**, v. 17, n. 2, p. 155–160, 2006.

ZHANG, Z. et al. Mixed linear model approach adapted for genome-wide association studies. **Nature Genetics**, v. 42, n. 4, p. 355–360, 2010.

ZHU, C. et al. Status and Prospects of Association Mapping in Plants. **The Plant Genome Journal**, v. 1, n. 1, p. 5, 2008.

ZILA, C. T. et al. A Genome-Wide Association Study Reveals Genes Associated with *Fusarium* Ear Rot Resistance in a Maize Core Diversity Panel. **G3: Genes|Genomes|Genetics**, v. 3, n. 11, p. 2095–2104, 2013.

ZILA, C. T. et al. Genome-wide association study of *Fusarium* ear rot disease in the U.S.A. maize inbred line collection. **BMC Plant Biology**, v. 14, n. 1, p. 372, 2014.

SEGUNDA PARTE – ARTIGO

ARTIGO: ANÁLISE DE ASSOCIAÇÃO GENÔMICA PARA RESISTÊNCIA À PODRIDÃO DE ESPIGA CAUSADA POR *Fusarium verticillioides*

RESUMO

O milho é um dos cereais mais cultivados e seu amplo cultivo permite que uma extensa gama de patógenos ataquem a cultura. Os patógenos de maior importância são os agentes causais de podridões de espiga, pois afetam diretamente a produção e a qualidade dos grãos. Dentre os principais agentes causais da podridão de espiga, destaca-se a podridão causada pelo fungo *Fusarium verticillioides*. A identificação de regiões causais associadas com a resistência à doença pode ser uma ferramenta útil no melhoramento da cultura e também é uma forma de entender melhor os mecanismos de resistência. Neste trabalho, foi conduzido um estudo de associação genômica para identificar associações significativas entre os marcadores e as possíveis regiões causais relacionadas com a resistência à podridão de espiga causada pelo fungo *F. verticillioides*. Foram avaliadas 242 linhagens em dois locais (Lavras e Uberlândia), para a proporção de espigas com sintomas de podridão, ou seja, que apresentavam mais de 25% da espiga doente em relação ao número total de espigas. As associações foram testadas com 23.153 marcadores DArT, utilizando o modelo misto clássico de mapeamento e o modelo de seleção genômica BSSV (*Bayesian Stochastic Search Variable*). Foram identificados onze marcadores significativamente associados com a resistência, sendo quatro associados por ambos os métodos. Todos os marcadores foram identificados próximos a genes que possuem funções que podem ter relação com a resistência, sendo necessários estudos de rotas metabólicas.

Palavras-chave: Estudo de Associação Genômica. Podridão de espiga. Bayesian Stochastic Search Variable.

1. INTRODUÇÃO

O milho (*Zea mays* L.) é um dos cereais mais produzidos no mundo e sua importância econômica é caracterizada pela sua diversidade de utilização. A demanda mundial por milho vem aumentando principalmente estimulada pelo crescimento dos países asiáticos e pela utilização do milho para a produção de etanol pelos Estados Unidos. Além disso, o crescimento no setor de carnes, principalmente de aves e suínos, também tem aumentado a demanda (PAVÃO; FERREIRA FILHO, 2011).

O desenvolvimento de cultivares mais produtivas, a melhoria dos sistemas de cultivo, a fertilidade dos solos, dentre outros fatores têm contribuído para o aumento da produção do milho. Contudo, a ocorrência de plantios sucessivos e sem a rotação de culturas contribuiu para a mudança na dinâmica da população de patógenos, conseqüentemente, causando o aumento na incidência e severidade de doenças na cultura (COSTA et al., 2010; DENTI; REIS, 2003). O aumento na incidência e severidade de doenças na cultura do milho vem causando grande preocupação entre os produtores, pois as doenças são os principais fatores que provocam a redução na produtividade de grãos e na qualidade sanitária.

Dentre as principais doenças da cultura do milho no Brasil, destacam--se os agentes causais de podridões de espiga, causadoras de grãos ardidos. A ocorrência destes patógenos afeta principalmente a produção e a qualidade dos grãos. Dentre as podridões de espigas, uma que causa grande impacto econômico é a podridão rosada da espiga causada por *Fusarium verticillioides* (EMBRAPA, 2010). Além da redução na produção e na qualidade dos grãos, o patógeno ainda produz metabólitos tóxicos secundários, chamados de micotoxinas, como a fumonisina (BRODERS et al., 2007; MENDES, 2009). Essas toxinas apresentam efeitos tóxicos e agudos, podendo causar câncer tanto em humanos quanto em animais (BULLERMAN, 1996; MARASAS, 1996; MENDES, 2009).

Algumas práticas agrônômicas são recomendadas para reduzir o inóculo do patógeno na área de cultivo, como rotação de cultura, uso de sementes saudáveis, uso de híbridos transgênicos no controle de pragas, controle de plantas daninhas, uso de cultivares resistentes, plantio na densidade recomendada, entre outras (EMBRAPA, 2010; MUNKVOLD; HELLMICH; RICE, 1999; MUNKVOLD; HELLMICH; SHOWERS, 1997). Muitas vezes, essas práticas não são suficientes para o controle do patógeno, pois este tem vários hospedeiros e sobrevive em restos culturais. Dessa forma, o melhor controle do patógeno é realizado por meio da resistência, esta por ser mais eficiente, simples, de baixo impacto ambiental e custo (ALI; TAYLOR; JIE, 2005).

Sabe-se que a resistência à podridão de espiga por *Fusarium* tem o controle poligênico e é bastante influenciada pelo ambiente (KING; SCOTT, 1981; NANKAM; PATAKY, 1996; ROBERTSON et al., 2006). Apesar da vantagem de utilizar genótipos resistentes, ainda existem poucas cultivares resistentes no mercado devido à complexidade da resistência (BUSH et al., 2004; CLEMENTS et al., 2004; ELLER; HOLLAND; PAYNE, 2008).

Normalmente, no melhoramento para resistência a podridões de espiga, é realizada a seleção indireta, onde a resistência é avaliada por meio de características secundárias, como o número de espigas doentes e porcentagem de grãos ardidos (CLEMENTS et al., 2003). Levando em conta a baixa eficácia da seleção fenotípica para a seleção de genótipos resistentes, a utilização da seleção assistida por marcadores (MAS) surgiu como uma estratégia para melhorar a eficiência de seleção (HOLLAND, 2004; ROBERTSON-HOYT et al., 2006).

O uso da seleção genética via seleção auxiliada por marcadores tem sido muito aplicado na detecção de genes com efeitos maiores e na introgressão para melhorar certas características. No entanto, o uso da seleção assistida e o mapeamento de QTL tem mostrado algumas limitações, devido a longos ciclos de seleção e à busca de associações marcador-QTL significativos, sendo incapaz de

captar os efeitos de genes menores (DESTA; ORTIZ, 2014; HEFFNER; SORRELLS; JANNINK, 2009).

O estudo de associação é uma boa alternativa para superar algumas dessas limitações do mapeamento de QTL. De uma forma geral, o mapeamento associativo utiliza o desequilíbrio de ligação entre os marcadores e os genes ou QTL de interesse para identificar associações, e isso é possível por meio da análise de desequilíbrio de ligação que utiliza todos os eventos de recombinação que teriam ocorrido entre o gene e o marcador no passado na população para serem utilizados no mapeamento associativo (SINGH; SINGH, 2015).

Existem poucos estudos para a identificação de regiões causais associadas à resistência a *Fusarium verticillioides*, e a aplicação do mapeamento associativo pode ser uma estratégia eficaz para a identificação de genes de resistência. Zila et al. (2013) utilizaram o estudo de associação para a identificação de genes associados à resistência à podridão de espiga. Nesse estudo, os autores identificaram 3 SNPs associados a genes que teriam relação com a resistência à doença. E, em um estudo posterior, Zila et al. (2014), com um painel de linhagens e um conjunto de marcadores maior que o estudo anterior, identificaram 7 SNPs associados à resistência à podridão de espiga causada por *Fusarium verticillioides*. Assim, os objetivos desse estudo foram: i) utilizar o mapeamento associativo para identificar marcadores associados à resistência à podridão de espiga; ii) identificar possíveis genes relacionados com a resistência à doença e iii) verificar a capacidade de associação do método *Bayesian Stochastic Search Variable* (BSSV), que é um método de seleção genômica.

2. MATERIAL E MÉTODOS

2.1. Experimentos de campo

Os experimentos de avaliação de incidência de podridão de espiga foram realizados no ano agrícola 2012/2013 em dois locais, Lavras (910 m, 21°14'S e 45°00'W) e Uberlândia (863 m, 18°55'S e 48°16'), no estado de Minas Gerais.

Foram avaliadas 238 linhagens, das 447 genotipadas, mais 4 testemunhas resistentes. Os experimentos foram avaliados em um delineamento de blocos incompletos com tratamentos regulares e 3 repetições. As parcelas experimentais foram compostas por uma linha de 3 metros com espaçamento de 0,7 metros.

2.2. Cultura do patógeno, inoculação e fenotipagem

Os isolados do fungo (*Fusarium verticillioides*) foram obtidos e replicados no Laboratório de Fitopatologia de Sementes da UFLA. O inóculo do patógeno foi crescido em meio de cultura completo durante sete dias antes da inoculação. No dia da inoculação, foi feita uma suspensão de conídios que, posteriormente, foi ajustada para a concentração de $1,0 \times 10^6$ conídios mL⁻¹ por meio da contagem de esporos na câmara de Neubauer.

A metodologia de inoculação do patógeno utilizada foi a proposta por Clements et al. (2003), na qual foram feitas algumas modificações (foi inoculado 1 mL da suspensão de conídios ao invés de 5 mL da suspensão). A inoculação foi realizada 15 dias após 100% das plantas terem emitido os estilos/estigmas. Com uma pipeta, foi inoculado 1 mL da suspensão de conídios no interior de cada espiga.

A incidência da podridão de espiga foi avaliada por meio da proporção de espiga com sintomas de podridão. Nessa avaliação, foi considerada a proporção

de espigas que apresentaram mais de 25% de sintomas de podridão em relação ao número total de espigas da parcela.

2.3. Caracterização das linhagens e análise da estrutura populacional

Para a caracterização genética, foram genotipadas 447 com 23.153 DArTSeq™ obtidos pela *Diversity Arrays Technology Pty Ltd* (Yarralumla ACT, Australia). Mais informações sobre o método podem ser obtidas em Kilian et al. (2012).

Os dados faltantes foram imputados utilizando a função *A.mat* (com o método *mean*) contida no pacote *rrBLUP* (ENDELMAN, 2011) do *software* R (R CORE TEAM, 2016). Os parentescos genômicos foram calculados usando a matriz de parentesco aditiva (**A**) proposta por VanRaden (2008) dada por

$$\mathbf{A} = \frac{\mathbf{W}_A \mathbf{W}'_A}{2 \sum \mathbf{p} \mathbf{q}}$$

em que **p** é a frequência do alelo favorável; **q** é a frequência do alelo desfavorável; **W_A** é a matriz de desvios dos marcadores centrada em **p** (média dos alelos favoráveis para um locus dado); e $2 \sum \mathbf{p} \mathbf{q}$ é a soma das variâncias dos loci.

Uma análise de agrupamento foi realizada por meio da função *hclust* do pacote base *stats* do *software* R (R CORE TEAM, 2016), calculada pelo método de Wald, utilizando uma matriz contendo as distâncias euclidianas dos elementos da matriz **A**.

Para construir a árvore filogenética, utilizou-se o método *Neighbor-Joining*, proposto por Saitou e Nei (1987), a partir de uma matriz contendo as distâncias euclidianas dos elementos da matriz de parentesco $\mathbf{A} = \frac{\mathbf{W}_A \mathbf{W}'_A}{2 \sum \mathbf{p} \mathbf{q}}$ (descrita anteriormente). A construção da árvore *neighbor-joining* foi realizada a partir da

função *nj* do pacote *ape* (PARADIS et al., 2004) do *software* R (R CORE TEAM, 2016).

A análise da estrutura populacional foi feita por meio do pacote GAPIT (LIPKA et al., 2012) do *software* R (R CORE TEAM, 2016), em que utilizou-se o modelo de seleção baseado no critério de informação bayesiano (BIC – *Bayesian Information Criterion*) para determinar o número ideal de componentes principais a serem incluídos no modelo.

2.4. Análise estatística

Utilizou-se um modelo misto para correção das observações de acordo com os efeitos de repetições, ambientes, interação genótipos x ambientes e resíduos. O modelo adotado foi:

$$y = X\beta + Tg + \Omega b + W\delta + \varepsilon \quad (1)$$

sendo y o vetor de observações; X a matriz incidência dos efeitos fixos (repetição dentro de ambiente); T a matriz incidência dos efeitos genéticos; Ω a matriz incidência dos efeitos aleatórios de bloco dentro de repetição; W a matriz incidência dos efeitos das interações genótipos x ambientes; β , g , b , δ os vetores dos efeitos relacionados a X , T , Ω e W , respectivamente; e ε o vetor dos resíduos. Assume-se que as distribuições dos efeitos g , b , δ e ε são $N(\mathbf{0}, \sigma_g^2)$, $N(\mathbf{0}, \sigma_b^2)$, $N(\mathbf{0}, \sigma_\delta^2)$ e $N(\mathbf{0}, \sigma_\varepsilon^2)$ respectivamente.

2.5. Análise de associação via modelos mistos

Utilizou-se o seguinte modelo misto

$$\tilde{\mathbf{y}} = \mathbf{j}\mathbf{u} + \mathbf{S}\boldsymbol{\tau} + \mathbf{Z}\mathbf{a} + \boldsymbol{\varepsilon} \quad (2)$$

em que : $\tilde{\mathbf{y}}$ é o vetor de observações corrigidas baseadas no modelo (1); \mathbf{j} é o vetor unitário correspondente à média; \mathbf{u} é a média amostral; \mathbf{S} é a matriz incidência correspondente aos marcadores; $\boldsymbol{\tau}$ é o vetor dos efeitos dos marcadores; \mathbf{Z} é a matriz incidência dos genótipos dos marcadores; \mathbf{a} é o vetor do efeito genético aditivo de cada marcador; e $\boldsymbol{\varepsilon}$ é o vetor dos efeitos residuais.

A matriz de variâncias fenotípicas \mathbf{V} é dada por

$$\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{I}\sigma_{\varepsilon}^2 = \sigma_{\varepsilon}^2(\mathbf{K}\boldsymbol{\lambda} + \mathbf{I})$$

sendo $\mathbf{G} = \mathbf{A}\sigma_a^2$ uma matriz de variância genética aditiva, $\mathbf{I}\sigma_e^2$ a matriz diagonal das variâncias residuais comuns e $\boldsymbol{\lambda} = \frac{\sigma_a^2}{\sigma_e^2}$.

A análise GWAS foi realizada com a função *GWAS* contida no pacote *rrBLUP* (ENDELMAN, 2011) do *software* R (R CORE TEAM, 2016). As estimativas dos BLUPs (*estimates of the best linear unbiased predictor*) e dos componentes da variância foram obtidas (uma única vez) pela máxima verossimilhança residual (REML) contida na função *GWAS* do pacote especificado acima, por meio do comando P3D (a terminologia P3D - *population parameters previously determined* – foi introduzida por Zhang et al, 2010).

2.6. Análise de associação via *Bayesian Stochastic Search Variable*

O modelo utilizado neste estudo foi BSSV (*Bayesian Stochastic Search Variable*) devido à sua capacidade para selecionar marcadores de grande efeito nos modelos com múltiplos marcadores. Uma nova abordagem foi utilizada, com ajustes no modelo original proposto por Yi et al. (2003), a fim de abranger todos

os efeitos do marcador. O modelo de múltiplos marcadores é adotado da seguinte forma

$$\mathbf{y} = \boldsymbol{\mu} + \sum_{j=1}^m \mathbf{z}_j \mathbf{a}_j + \boldsymbol{\varepsilon} \quad (3)$$

em que \mathbf{y} é um vetor de médias corrigidas baseadas no modelo (1); $\boldsymbol{\mu}$ é o intercepto; \mathbf{z}_j é o vetor dos genótipos do marcador j ; \mathbf{a}_j é o efeito aditivo do marcador j ; e $\boldsymbol{\varepsilon}$ é o vetor dos erros, seguindo distribuição $N(\mathbf{0}, \boldsymbol{\sigma}_e^2)$.

A aceitação do efeito do marcador depende da combinação de hipóteses a priori condicionadas a um conjunto de variáveis indicadoras. Portanto, podemos assumir que os efeitos aditivos dos marcadores a priori no modelo BSSV são: $a_j | \rho, \sigma_{a_j}^2, \delta \sim (1 - \rho)N(0, \sigma_{a_j}^2) + \rho N(0, \delta)$, $j = 1, \dots, m$ (4), sendo $\sigma_{a_j}^2$ e δ variâncias de alta e baixa magnitude no efeito genético do marcador, respectivamente.

Neste estudo, assumimos a priori que: $a_j | \nu, S^2 \sim \chi_{esc}^2(\nu = 1, S^2 = 10^{-8})$ (5), sendo os hiperparâmetros ν e S^2 relacionados ao método Bayes A descrito em Garrick e Fernando (2013), com $S^2 = \frac{\sigma_y^2 \cdot 0,005}{m \cdot (1 - \rho)} = 10^{-8}$; e $\delta = 10^{-13}$ (correspondente a S^2/m).

Assumiu-se que o hiperparâmetro ρ foi modelado por uma distribuição Beta $\rho | a, b \sim Beta(a = 1, b = 1)$, em vez de de 0,5, como descrito originalmente em Yi et al. (2003). A distribuição a priori para a média populacional foi assumida ser constante, e para a variância residual σ_e^2 foi assumida a mesma distribuição de $\sigma_{a_j}^2$.

2.7. Análise pós-MCMC

A significância dos efeitos dos marcadores foi determinada com o teste de Wald. As estatísticas deste teste W sob a hipótese nula seguem uma distribuição assintótica χ^2 com um grau de liberdade. Os valores do teste são obtidos em: $W = \frac{a_j^2}{\sigma_{a_j}^2}$ (6), sendo $\sigma_{a_j}^2 = \left[\emptyset - \emptyset \xi^{-1} \emptyset + \left(\sigma_{a_j}^2 \right)^{-1} \right]^{-1}$. O valor crítico para aceitar o marcador foi dado por $\chi_{tab}^2 = 3,84$, considerando um nível de significância de 5%, que é equivalente a $-\log_{10}(\text{p-valor}) = 1,30$.

O valor crítico adotado neste trabalho foi diferente do valor utilizado em modelos mistos normalmente sugerido para GWAS, já que, neste modelo utilizado, o efeito do marcador é submetido a um processo duplo shrinkage (priori do efeito nulo com probabilidade igual a ρ e variância próxima de zero e encolhimento a priori dado por $(v_{a_j}^2)^{-1}$). Em um modelo clássico GWAS, cada genótipo de marcador fixo é ajustado pelo tempo e, aqui, utilizou-se a correção de p-valores baseada em Barsh et al. (2012).

2.8. Estimativas do desequilíbrio de ligação e construção dos *heatmaps*

O cálculo do desequilíbrio de ligação (r^2) entre os DArTs foi realizado por meio da função *LD.Measures* do pacote *LDcorSV* (MANGIN et al., 2012) do *software* R (R CORE TEAM, 2016), e o *heatmap* referente ao desequilíbrio de ligação dos marcadores dentro de uma janela genômica foi construído pela função *LDheatmap* do pacote *LDheatmap* (SHIN et al., 2006) do mesmo *software*.

Uma análise de agrupamento foi realizada pela função *hclust* do pacote *base stats* do *software* R (R CORE TEAM, 2016), calculada pelo método de Wald, utilizando uma matriz contendo as distâncias euclidianas dos elementos da matriz **A**. O *heatmap* relacionado a essa análise de agrupamento, juntamente com o

dendrograma, foi construído por meio da função *heatmap* do mesmo pacote base deste *software*.

2.9. Validação dos genes associados à resistência à podridão de espiga

A caracterização dos genes que foram localizados próximos dos DArTs que foram significativamente associados com a resistência à doença foi feita utilizando o genoma de referência da linhagem B73 (SCHNABLE et al., 2009), por meio do *Maize Genetics and Genomics Database* (MAIZEGDB) (LAWRENCE et al., 2008), utilizando o algoritmo *Basic Local Alignment Search Tool* (BLAST) (ALTSCHUL et al., 1990). A função de cada gene bem como suas vias metabólicas foram determinadas utilizando o Maizecyc (MONACO et al., 2013).

3. RESULTADOS

3.1. Análises de fenotípicas para resistência à podridão de espiga causada por *Fusarium verticillioides*

Foram obtidas as estimativas da variância genética aditiva, variância residual e a herdabilidade. A proporção da variância total explicada pela variância genética para a análise conjunta foi de 46,74%. Para a análise individual de Lavras, a proporção da variância total explicada pela variância genética foi de 51,21% e, para a análise de Uberlândia, foi 36,85% (Tabela 1).

Tabela 1 – Estimativa dos componentes de variância genética aditiva (σ_a^2), componentes de variância residual (σ_e^2) e estimativa da herdabilidade (h^2) para as análises conjunta, individual de Lavras e individual de Uberlândia.

| | σ_a^2 | σ_e^2 | h^2 |
|------------|--------------|--------------|--------|
| CONJUNTA | 0.0371 | 0.0423 | 0.4674 |
| LAVRAS | 0.0360 | 0.0343 | 0.5121 |
| UBERLÂNDIA | 0.0584 | 0.1000 | 0.3685 |

Para a caracterização das linhagens, foi construída uma árvore filogenética pelo método *neighbor-joining*. Por meio da árvore filogenética, foi possível observar que as linhagens utilizadas no estudo de associação foram agrupadas em diferentes subgrupos, sendo que isso ocorreu porque as linhagens apresentam diferentes *backgrounds*. É possível observar uma clara distinção entre as linhagens pertencentes aos diferentes subgrupos. Na parte superior da árvore filogenética, é possível observar o agrupamento das linhagens derivadas do grupo *Stiff Stalk Synthetic* (SSS-B73-B17) e linhagens pertencentes ao grupo UNR. Na parte mediana da árvore, observa-se uma sobreposição dos grupos de linhagens pertencentes ao subgrupo temperado, o qual inclui as linhagens pertencentes aos subgrupos FDK-Arg, *Non Stiff Stalk Synthetic* (NSSS-PG84,TAT), *Lancaster*,

MNK-ARG e *Iodent*. E, na parte inferior da figura, encontram-se as linhagens dos grupos *Tropical Dent*, *Tropical Flint*, *Suwan* e *Amarilo Dent*, as quais pertencem ao subgrupo tropical de linhagens (Figura 1).

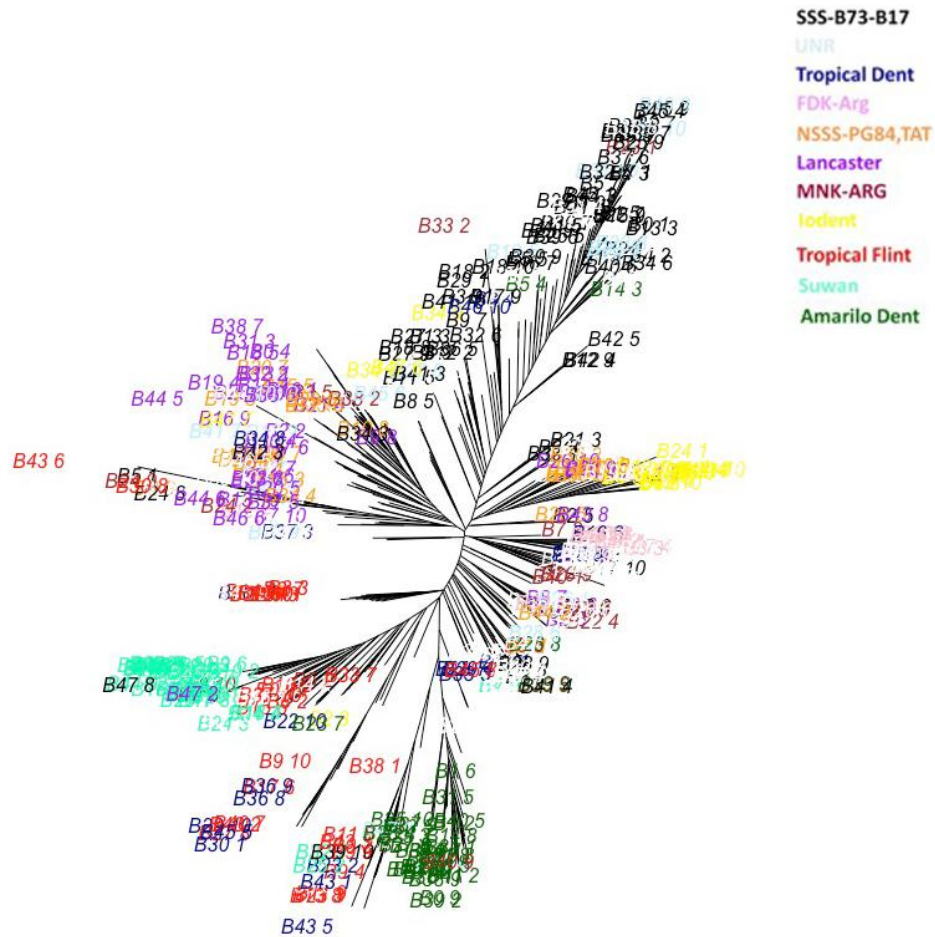


Figura 1 – Árvore filogenética construída pelo método Neighbor-Joining mostrando os diferentes *backgrounds* das linhagens utilizadas no estudo de associação genômica.

3.2. Análise de associação da resistência à podridão de espiga causada por *Fusarium verticillioides* por meio do método *Bayesian Stochastic Search Variable* (BSSV)

Nas análises com o método *Bayesian Stochastic Search Variable* (BSSV), na análise conjunta, foram identificados cinco DArTs significativamente associados à resistência à podridão de espiga. Estes DArTs foram identificados nos cromossomos 1 (posição 78,131,634 bp e posição 136,272,488 bp), 7 (posição 120,994,839 bp e posição 158,219,730 bp) e no grupo sem cromossomo identificado (Ni). Para Lavras, cinco DArTs foram associados significativamente nos cromossomos 1 (posição 53,415,853 bp e posição 136,272,488 bp), 7 (na posição 124,220,838 bp), 8 (na posição 136,137,749 bp) e no grupo sem cromossomo identificado (Ni). Para Uberlândia, dois DArTs foram significativamente associados à resistência, sendo um identificado no cromossomo 4 (posição 38,878,586 bp) e outro no cromossomo 5 (posição 55,225,453 bp) (Tabela 2 e Figura 2). Do total de 12 DArTs significativamente associados à resistência por meio do método BSSV, dois DArTs foram associados tanto na análise conjunta quanto na análise de Lavras.

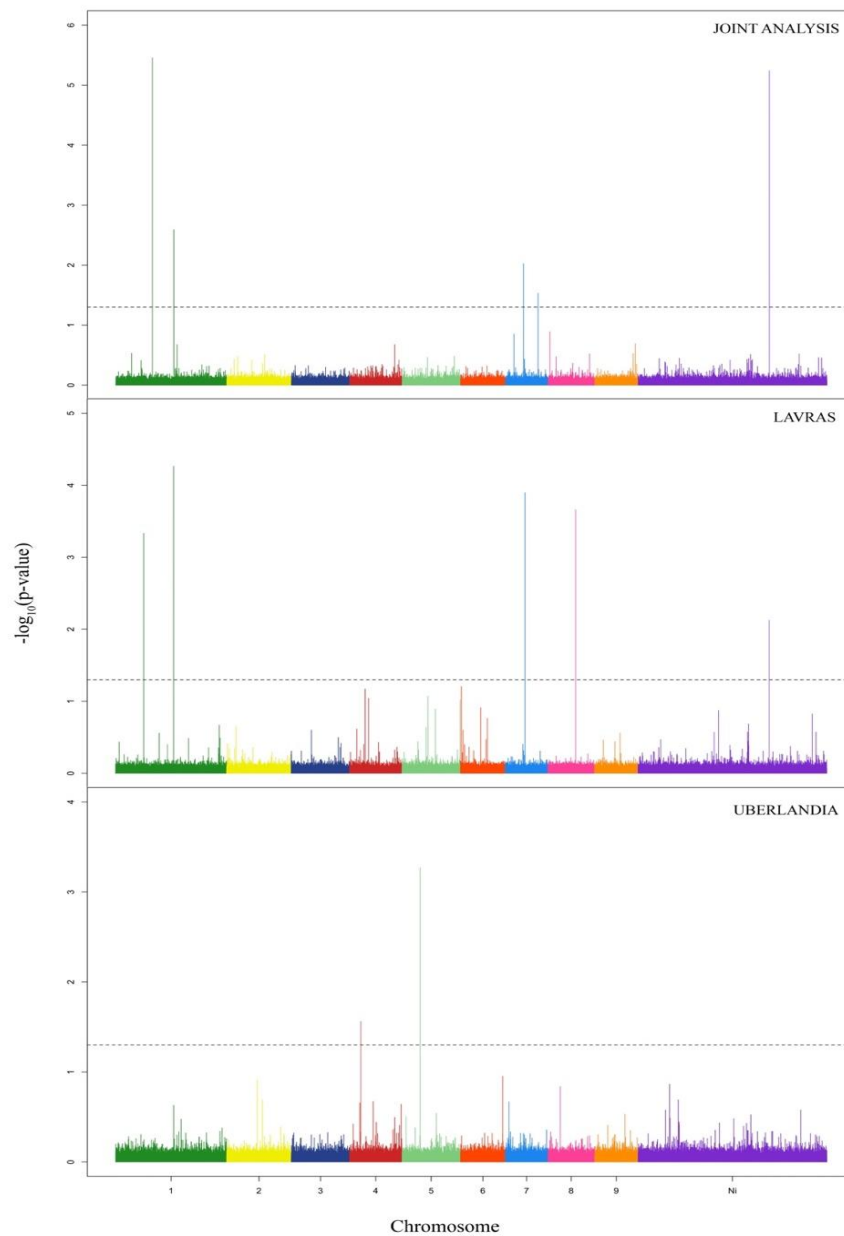


Figura 2 – Gráficos Manhattan mostrando as associações significativas para as análises conjunta, individual de Lavras e individual de Uberlândia utilizando o método *Bayesian Stochastic Search Variable* (BSSV).

Tabela 2 – Cromossomo, posição em bp, valores de $-\log_{10}(\text{p-valor})$ e sequência dos DArTs que foram associados significativamente com a resistência à podridão de espiga pelo método Bayesian Stochastic Search Variable (BSSV) (Continua).

| | Cromossomo | Posição (bp) | $-\log_{10}(\text{p-valor})$ | Sequência |
|----------|------------|--------------|------------------------------|---|
| CONJUNTA | 1 | 78,131,634 | 5.457281 | TGCAGAGAAGGGTGTGGCTTCTTA TGGGGATGGGGCTGACTCAACTGG GATTGTTATTAGAAGCCGTTA TGCAGCTGCCTCGCCCGTTTGGAT |
| | 1 | 136,272,488 | 2.591074 | GCACATGAGATCGTAAGAGCGGTT CAGCAGGAATGCCGAGACCGA TGCAGGCAAAAGAGCTAGCCACTA |
| | 7 | 120,994,839 | 2.026296 | GAACAAAGCTAGACTTCGACACTT CAAGGCGAAACTGGGGGGATA TGCAGCTGCGTGTGGAGTTGCCAG |
| | 7 | 158,219,730 | 1.534253 | CGGCCAGCAACAAAGGTGGCGGA GCTAGCTAGCCCCGCATGACAA |
| | Ni | - | 5.241185 | - |
| LAVRAS | 1 | 53,415,853 | 3.332977 | TGCAGGTAAAGGTGTGCATCAAAT CCGGTCCGGAGCAACAAAATACG CCACGTGAACAATGTGGCGCG TGCAGCTGCCTCGCCCGTTTGGAT |
| | 1 | 136,272,488 | 4.264869 | GCACATGAGATCGTAAGAGCGGTT CAGCAGGAATGCCGAGACCGA TGCAGCAGGTACAGAATCACTATC |
| | 7 | 124,220,838 | 3.897082 | GCATTCGCGAGCACATTATTGTGC ACGATATGATCAGCCATGGGA |

Tabela 2 – Cromossomo, posição em bp, valores de $-\log_{10}(\text{p-valor})$ e sequência dos DArTs que foram associados significativamente com a resistência à podridão de espiga pelo método Bayesian Stochastic Search Variable (BSSV) (Conclusão).

| | Cromossomo | Posição (bp) | $-\log_{10}(\text{p-valor})$ | Sequência |
|------------|------------|--------------|------------------------------|---|
| UBERLÂNDIA | 8 | 136,137,749 | 3.659298 | TGCAGATATTTGGAAGAGAAGACA GAAATTAATTCAGGTCAGCTACAC ATGAGATCGGAAGAGCGGTTC |
| | Ni | - | 2.124613 | - |
| | 4 | 36,878,586 | 1.561804 | TGCAGATATTTCAAGCCGGAACTT TCTTCTCTTTATAATAAATCTGCCA GCTAACATGAGATCGGAAGA |
| | 5 | 55,225,453 | 3.268957 | TGCAGAATATCAATTGTTGTTGATT GTTCTGAAGAGACGTCATTTCTTA GTTGGCAATTGGGAGAACCT |

3.3. Análise de associação da resistência à podridão de espiga causada por *Fusarium verticillioides* por meio do método de modelo misto

Nas análises com o método de modelo misto, na análise conjunta, três DArTs foram associados significativamente à resistência. Os DArTs foram identificados no cromossomo 1 (posição 78,131,634 bp), no cromossomo 7 (posição 124,220,838 bp) e a outro DArT foi associado a um cromossomo não identificado (Ni). Para Lavras, três DArTs foram identificados significativamente associados à resistência, sendo dois no cromossomo 7 (posição 124,220,838 bp e posição 159,891,338 bp) e um DArT no grupo sem cromossomo identificado (Ni). Para Uberlândia, também foram identificados três DArTs associados à resistência, dois no cromossomo 1 (posição 78.131.634 bp e posição 289,504,622 bp) e outro no cromossomo 5 (posição 55,225,543 bp) (Tabela 3 e Figura 3). Do total de DArTs significativamente associados por meio dos métodos de modelo misto, três foram coincidentes nas análises.

Na comparação entre os dois métodos, BSSV e modelo misto, quatro DArTs em comum entre os dois métodos foram associados significativamente à podridão de espiga causada pelo fungo *Fusarium verticillioides*. Os DArTs em comum foram identificados no cromossomo 1 (posição 78,131,34 bp), no cromossomo 5 (posição 55,225,453 bp), no cromossomo 7 (posição 124,220,838 bp) e no grupo sem cromossomo identificado (Ni).

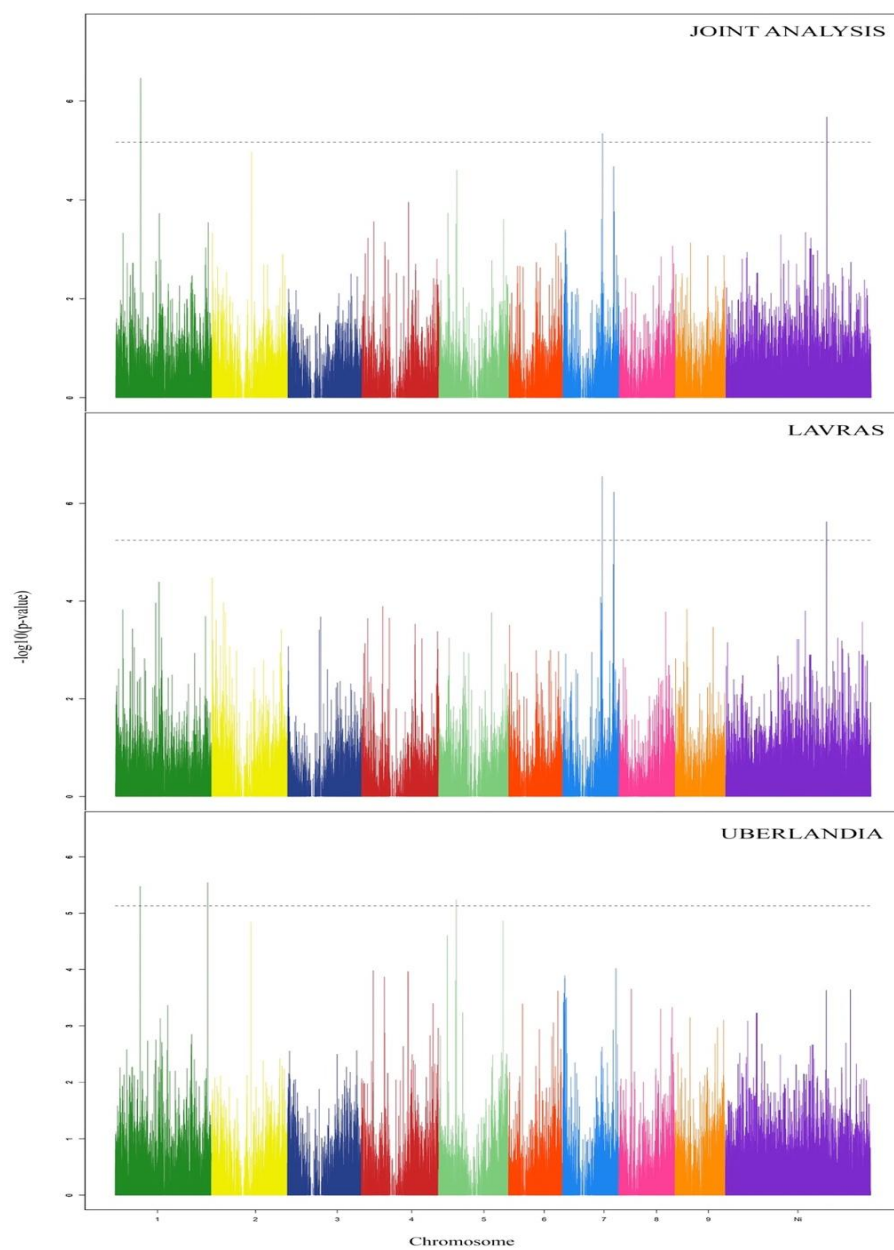


Figura 3 – Gráficos Manhattan mostrando as associações significativas para as análises conjunta, individual de Lavras e individual de Uberlândia utilizando o método de modelo misto (MM).

Tabela 3 – Cromossomo, posição em bp, valores de $-\log_{10}(\text{p-valor})$ e sequência dos DArTs que foram associados significativamente com a resistência à podridão de espiga pelo método de modelo misto (MM).

| | Cromossomo | Posição (bp) | $-\log_{10}(\text{p-valor})$ | Sequência |
|------------|------------|--------------|------------------------------|---|
| CONJUNTA | 1 | 78,131,634 | 6.462276 | TGCAGAGAAGGGTGTGGCTTCTTATGGGGATGGGG CTGACTCAACTGGGATTGTTATTAGAAGCCGTTA |
| | 7 | 124,220,838 | 5.342481 | TGCAGCAGGTACAGAATCACTATCGCATTTCGCGAG CACATTATTGTCGACGATATGATCAGCCATGGGA |
| | Ni | - | 5.677778 | - |
| LAVRAS | 7 | 124,220,838 | 6.553781 | TGCAGCAGGTACAGAATCACTATCGCATTTCGCGAG CACATTATTGTCGACGATATGATCAGCCATGGGA |
| | 7 | 159,891,338 | 6.230421 | TGCAGAGCCGGCGATGCCAGGTTCCGCCGCTCGTGG GGAAGAGATGGAGCACAGAAGAGTGAAGGATCAG |
| | Ni | - | 5.623650 | - |
| UBERLÂNDIA | 1 | 78,131,634 | 5.477389 | TGCAGAGAAGGGTGTGGCTTCTTATGGGGATGGGG CTGACTCAACTGGGATTGTTATTAGAAGCCGTTA |
| | 1 | 289,504,622 | 5.542544 | TGCAGCAGCAAAGGCGGAGGAGTACACGCCAGTG ACGACCGCATGAGCCGAGCCGGGCGGAGCACCCA G |
| | 5 | 55,225,453 | 5.236968 | TGCAGAATATCAATTGTTGTTGATTGTTCTGAAGA GACGTCATTTCTAGTTGGCAATTGGGAGAACCT |

3.4. Identificação de genes candidatos associados por meio do método *Bayesian Stochastic Search Variable (BSSV)*

Para a identificação das regiões causais associadas significativamente à resistência à podridão de espiga, foi utilizado o *Maize Genetics and Genomics Database (MaizeGDB)* (LAWRENCE et al., 2008), o qual utiliza o genoma da linhagem B73 como referência (SCHNABLE et al., 2009). Além disso, para verificar a resolução do mapeamento e também para o desequilíbrio de ligação entre os DArTs, foi elaborado um gráfico do tipo *heatmap*, a partir de uma janela genômica contendo o marcador que foi associado significativamente à resistência à podridão de espiga.

Tabela 4 – Genes identificados próximos dos DArTs associados à resistência, distância em bp e respectiva função por meio do método *Bayesian Stochastic Search Variable (BSSV)* (Continua).

| | Crom. | Posição DArT (bp) | Gene mais próximo | Dist. DART – gene | Função |
|----------|-------|-------------------|-------------------|-------------------|--------------------------------|
| CONJUNTA | 1 | 78,131,634 | GRMZM2 G445187 | 45 kbp | ORF |
| | 1 | 136,272,488 | GRMZM2 G072415 | 239 kbp | Transporte de proteínas |
| | 7 | 120,994,839 | GRMZM2 G147966 | 53 kbp | Processo de oxidação e redução |
| | 7 | 158,219,730 | GRMZM2 G107565 | 20 kbp | Atividade catalítica |
| LAVRAS | 1 | 53,415,853 | GRMZM2 G031001 | 3 kbp | Regulação da transcrição |
| | 1 | 136,272,488 | GRMZM2 G072415 | 239 kbp | Transporte de proteínas |
| | 7 | 124,220,838 | GRMZM2 G121649 | 28 kbp | ORF |
| | 8 | 136,137,749 | GRMZM2 G088196 | 56 kbp | Transporte transmembrana |

Tabela 4 – Genes identificados próximos dos DArTs associados à resistência, distância em bp e respectiva função por meio do método Bayesian Stochastic Search Variable (BSSV) (Conclusão).

| | Crom. | Posição DArT (bp) | Gene mais próximo | Dist. DART – gene | Função |
|------------|-------|-------------------|-------------------|-------------------|-------------------------------------|
| UBERLÂNDIA | 4 | 36,878,586 | GRMZM2 G046743 | 31 kbp | Transporte de aminoácidos |
| | 5 | 55,225,453 | GRMZM2 G162184 | 18 kbp | Atividade da Adenosilhomocisteinase |

Tabela 5 – Genes identificados próximos dos DArTs associados à resistência, distância em bp e respectiva função por meio do método de modelo misto (MM).

| | Crom. | Posição DArT (bp) | Gene mais próximo | Dist. DART – gene | Função |
|------------|-------|-------------------|-------------------|-------------------|-------------------------------------|
| CONJUNTA | 1 | 78,131,634 | GRMZM2 G445187 | 45 kbp | ORF |
| | 7 | 124,220,838 | GRMZM2 G121649 | 28 kbp | ORF |
| LAVRAS | 7 | 124,220,838 | GRMZM2 G121649 | 28 kbp | ORF |
| | 7 | 159,891,338 | GRMZM2 G092669 | 3 kbp | Atividade hormonal |
| UBERLÂNDIA | 1 | 78,131,634 | GRMZM2 G445187 | 45 kbp | ORF |
| | 1 | 289,504,622 | GRMZM2 G171315 | Íntron | ORF |
| | 5 | 55,225,453 | GRMZM2 G162184 | 18 kbp | Atividade da Adenosilhomocisteinase |

No total de doze DArTs associados significativamente à resistência à podridão de espiga, apenas um DArT foi localizado dentro de um gene

identificado no mapa físico do milho, sendo um localizado dentro de uma região não codificante do gene (íntron). Os restantes dos DArTs foram identificados em regiões próximas a alguns genes (Tabela 4 e 5).

Na análise conjunta por meio do método BSSV, nenhum dos quatro DArTs que foram associados significativamente estavam localizados dentro de um gene. Todos os marcadores associados com a resistência ficaram localizados próximos a genes que possuem funções básicas na célula. Os dois DArT identificados no cromossomo 7 (posição 120,994,839 e posição 158,219,730 bp) foram localizados próximos (53 e 20 kbp, respectivamente) dos genes GRMZM2G147966 (posição 121,043,853-121,048,231 bp) e GRMZM2G107565 (posição 158,194,386-158,199,300 bp), que são responsáveis pelo processo de oxidação e redução e pela atividade catalítica na célula, respectivamente, e podem estar relacionados com a resistência à podridão de espiga (Tabela 4).

Na análise individual de Lavras com o método BSSV, nenhum dos DArTs que foram significativamente associados à doença foram identificados dentro de regiões codantes de genes. O DArT identificado no cromossomo 1 (posição 53,415,853 bp), apesar de não estar dentro de nenhum gene previsto no mapa físico do milho, foi localizado muito próximo (3 kbp) de um gene que é responsável pela regulação da transcrição (GRMZM2G031001, posição 53,419,144-53,421,347 bp). Os demais DArTs identificados foram localizados próximos a genes que são responsáveis por funções básicas ou próximos a ORFs (*Open Reading Frame* – Fase de Leitura Aberta), que são sequências de nucleotídeos sem um códon de terminação (Tabela 4).

Com o método BSSV, para a análise individual de Uberlândia, dois DArTs foram associados significativamente à resistência à doença. Um DArT foi identificado no cromossomo 4 (posição 36,878,586 bp) e não foi localizado dentro de nenhum gene. Apesar disso, o DArT identificado está localizado próximo (31 kbp) do gene (GRMZM2G046743, posição 36,910,205-36,913,985 bp) que é

responsável pelo transporte de aminoácidos. O outro DArT identificado no cromossomo 5 também não é localizado dentro de nenhum gene, mas foi localizado próximo (18 kbp) ao gene responsável pela atividade da adenosilhomocisteinase (GRMZM2G162184, posição 55,243,566-55,261,337 bp) (Tabela 4).

3.5. Identificação de genes candidatos associados por meio do método de modelo misto

Na análise conjunta, pelo método de modelo misto, dois DArTs significativamente associados à resistência foram identificados. Esses DArTs foram também identificados nas análises individuais de Lavras e de Uberlândia. Os DArTs identificados no cromossomo 1 (posição 78,131,634 bp) e no cromossomo 7 (posição 124,220,838 bp) não foram localizados próximos a nenhum gene que poderia estar relacionado com algum mecanismo de resistência. Ambos os marcadores foram localizados próximos a ORFs (Tabela 5).

Por meio do método de modelo misto, na análise individual de Lavras, dois DArTs, ambos no cromossomo 7, foram associados significativamente à resistência à doença. O DArT identificado na posição 78,131,643 bp também foi associado significativamente na análise conjunta. O segundo DArT associado significativamente no cromossomo 7 (159,891,338 bp) foi localizado próximo (3 kbp) a um gene que regula a atividade hormonal (GRMZM2G092669, posição 159,895,056-159,898,945 bp) (Tabela 5).

Na análise individual de Uberlândia, pelo método de modelo misto, foram identificados três DArTs associados significativamente à resistência à doença, sendo apenas um localizado dentro de uma região intrônica de um gene. Dois DArTs destes também foram identificados nas outras análises. O DArT identificado no cromossomo 1 (posição 78,131,634 bp) foi associado significativamente tanto na análise conjunta pelo método BBSV quanto na análise

conjunta pelo método de modelo misto. O DArT identificado no cromossomo 5 (posição 55,225,453 bp) também foi associado significativamente na análise de Uberlândia por meio do método BSSV. E o outro DArT identificado no cromossomo 1 (posição 289,504,622 bp) foi localizado na região intrônica do gene (GRMZM2G171315, posição 289,433,159 – 289,530,335 bp), que uma ORF (Tabela 5).

3.6. Desequilíbrio de ligação entre os marcadores

De modo geral, em espécies alógamas, o desequilíbrio de ligação (r^2) decai muito rapidamente. Como pode ser observado (Figura 4), o desequilíbrio de ligação cai rapidamente para valores abaixo de $r^2 = 0,2$ em uma distância muito pequena.

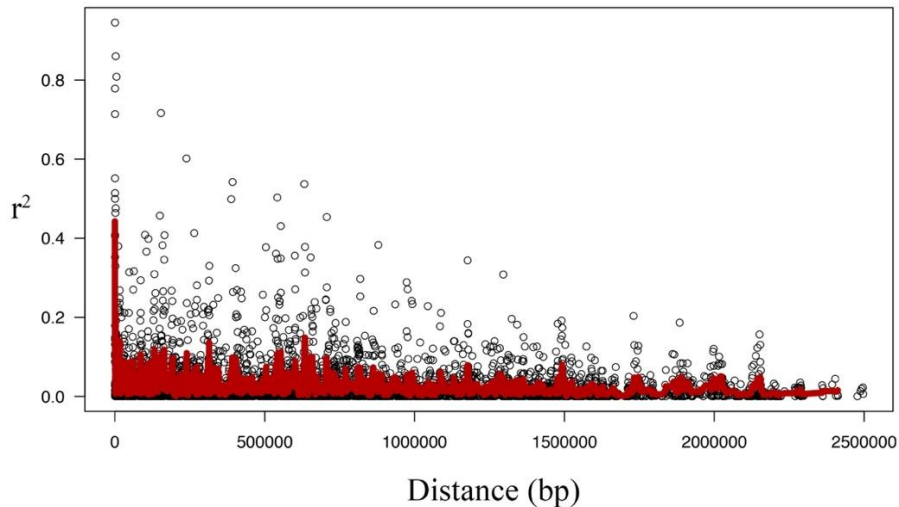


Figura 4 – Declínio do desequilíbrio de ligação entre os marcadores ao longo do genoma. A linha em vermelho representa o valor médio.

A partir de uma janela genômica contendo o DArT que foi significativamente associado com a resistência, foram construídos gráficos do tipo *heatmap* para verificar o desequilíbrio de ligação entre os DArTs associados significativamente com a resistência à podridão de espiga e os DArTs adjacentes, e, também, para verificar a existência de blocos de ligação nos quais os DArTs associados com a doença poderiam se encontrar. Em todas as análises, foi possível observar o baixo desequilíbrio de ligação entre os marcadores associados com a resistência com os demais marcadores dentro da janela genômica. Além disso, foi observado a existência de pequenos blocos gênicos com estimativas de desequilíbrio de ligação mais elevadas (Figuras 8 a 18). A função e a localização dos genes apresentados nos gráficos do tipo *heatmap* das Figuras 8 a 18 estão presentes na Tabela 8 (APÊNDICE).

Pela análise de Critério de Informação Bayesiana (BIC), foi determinado o número de componentes principais que seriam necessários para que a estrutura populacional fosse explicada adequadamente. Tanto para a análise conjunta quanto para as análises individuais de Lavras e de Uberlândia, o número de componentes principais necessários foi 0 (Tabela 6).

Tabela 6 – Número de componentes principais que devem ser incluídos no modelo para explicar adequadamente sobre a estrutura populacional, baseado no modelo de seleção BIC.

| | Número de Componentes Principais | Critério de Informação Bayesiana |
|------------|----------------------------------|----------------------------------|
| CONJUNTA | 0 | 2.5811 |
| LAVRAS | 0 | 22.4603 |
| UBERLÂNDIA | 0 | -89.1539 |

Nas Figuras 5, 6 e 7, é possível analisar o valor dos marcadores que foram significativamente associados à resistência à podridão de espiga para todas as linhagens avaliadas e compará-los com os valores fenotípicos corrigidos da avaliação de proporção de espigas doentes.

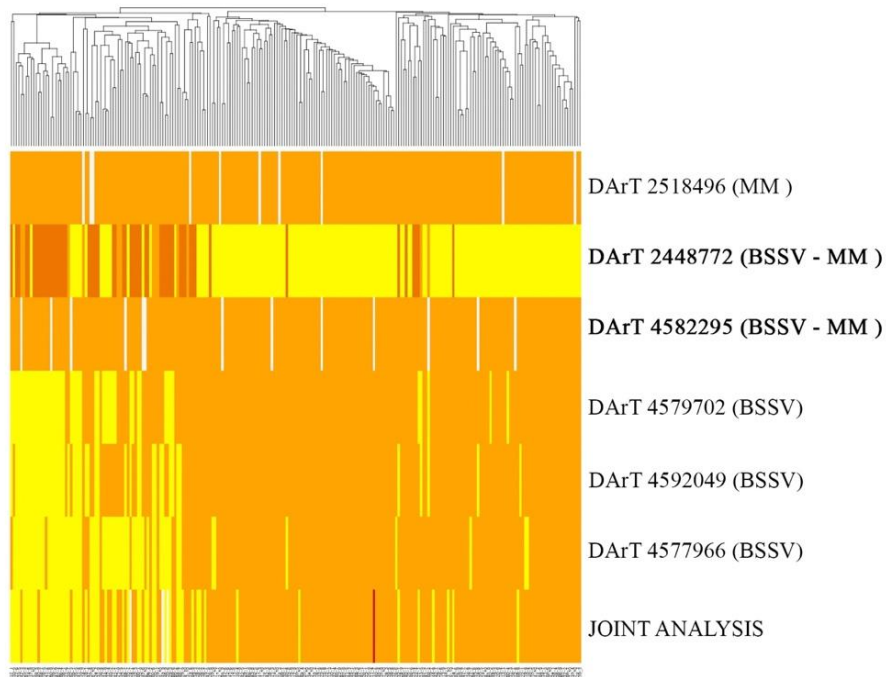


Figura 5 – Valores fenotípicos corrigidos e valores dos marcadores que foram significativamente associados com a resistência para a análise conjunta. As cores variam do branco ao vermelho, sendo que a coloração branca representa os menores valores, e a coloração vermelha, os maiores valores.

Na análise conjunta (Figura 5), foi observado que os DArTs associados significativamente pelo método BSSV apresentaram um comportamento semelhante ao fenótipo, ou seja, os marcadores praticamente copiaram o comportamento fenotípico das linhagens. Para os DArTs 4577966, 4592049 e 4579702, que foram associados pelo método BSSV, foi observado que os maiores valores dos marcadores foram apresentados pelas linhagens que apresentaram os maiores valores de proporção de espigas com sintomas da doença. Já para o DArT 2448772, foi observado um comportamento ao contrário do que foi observado no

fenótipo das linhagens, ou seja, as linhagens que apresentaram os maiores valores fenotípicos apresentaram valores do marcador menores. Para o DArT 4582295, que foi associado à resistência pelo método BSSV e pelo método de modelo misto, e para o DArT 2518496, que foi associado somente pelo método de modelo misto, não houveram coincidências entre os valores dos marcadores com os valores fenotípicos corrigidos.

Para a análise de Lavras (Figura 6), também foi possível observar que os marcadores associados à resistência pelo método BSSV (DArTs 4767436 e 4577966), por ambos os métodos (DArT 2448772) e o marcador que foi associado somente pelo método de modelo misto copiaram o comportamento dos fenótipos, mas não de forma tão clara como na análise conjunta. Os demais DArTs (4772985, 2518496 e 4581091) não apresentaram nenhuma coincidência com o comportamento dos fenótipos.

Na análise de Uberlândia (Figura 7), apenas o DArT 4772929 identificado pelo método BSSV apresentou um desempenho coincidente com o fenótipo, no qual os marcadores com os menores valores coincidiram com os menores valores de proporção de espigas doentes das linhagens. Os demais DArTs (4592572, 4582295 e 2400968) não apresentaram comportamento consistente com os fenótipos.

Essa observação está de acordo com os valores dos efeitos dos marcadores, pois os DArTs 4585610 e 4767973 apresentaram efeitos positivos, enquanto o DArT 4775329 apresentou um efeito alélico negativo (Tabela 7).

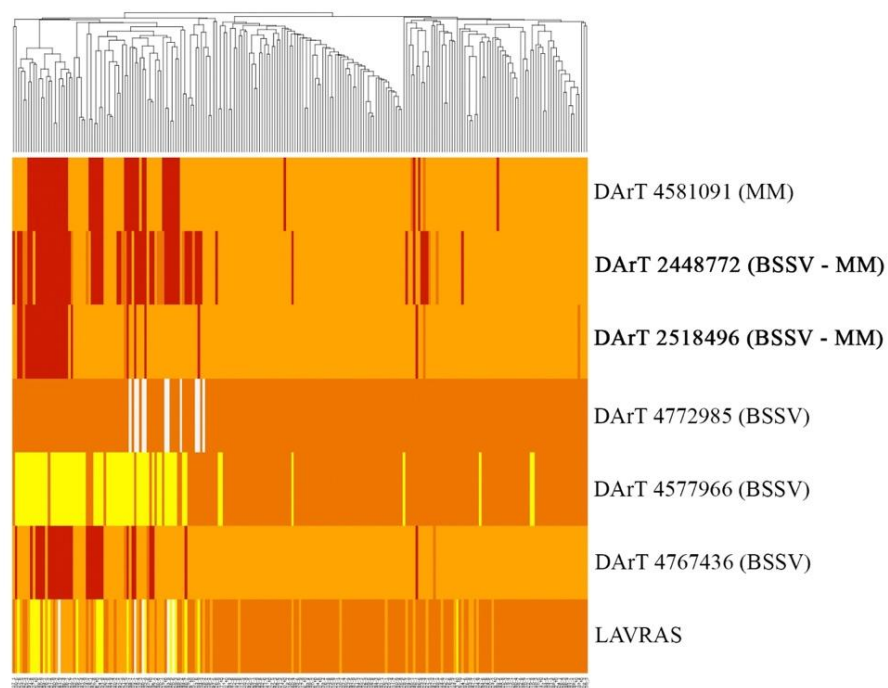


Figura 6 – Valores fenotípicos corrigidos e valores dos marcadores que foram associados significativamente associados com a resistência para a análise de Lavras. As cores variam do branco ao vermelho, sendo que a coloração branca representa os menores valores, e a coloração vermelha, os maiores valores.

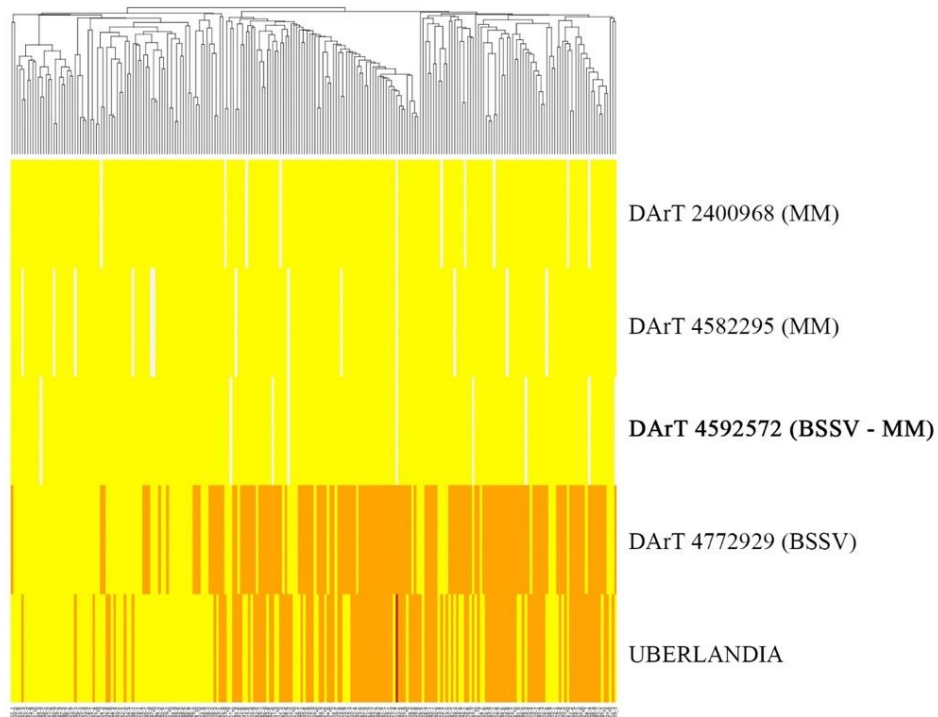


Figura 7 – Valores fenotípicos corrigidos e valores dos marcadores que foram associados significativamente associados com a resistência para a análise de Uberlândia. As cores variam do branco ao vermelho, sendo que a coloração branca representa os menores valores, e a coloração vermelha, os maiores valores.

Tabela 7 – Efeitos dos DArTs e frequência do menor alelo dos marcadores nos diferentes métodos.

| | Método | DArT | Efeito alélico | Frequência menor alelo |
|------------|---------|---------|----------------|------------------------|
| CONJUNTA | BSSV | 4582295 | -0.2035 | 0.0631 |
| | | 4577966 | 0.0691 | 0.2935 |
| | | 4592049 | 0.0368 | 0.2225 |
| | | 4579702 | 0.0317 | 0.2323 |
| | | 2448772 | -0.1485 | 0.2549 |
| | MM | 4582295 | -0.3230 | 0.4685 |
| 2518496 | | -0.3514 | 0.0575 | |
| 2448772 | | -0.2445 | 0.1274 | |
| LAVRAS | BSSV | 4767436 | -0.0888 | 0.1338 |
| | | 4577966 | 0.1020 | 0.2974 |
| | | 2518496 | -0.1146 | 0.1121 |
| | | 4772985 | 0.1380 | 0.0533 |
| | | 2448772 | -0.0306 | 0.2583 |
| MM | 2518496 | -0.3711 | 0.0582 | |
| | 4581091 | -0.2546 | 0.0910 | |
| | 2448772 | -0.2256 | 0.1291 | |
| UBERLÂNDIA | BSSV | 4772929 | 0.0468 | 0.3739 |
| | | 4592572 | -0.2262 | 0.0446 |
| | MM | 4582295 | -0.4406 | 0.4685 |
| | | 2400968 | -0.5057 | 0.4731 |
| | | 4592572 | -0.5074 | 0.4776 |

4. DISCUSSÃO

4.1. Herdabilidade e estudo de associação

As estimativas de herdabilidade observadas neste estudo (Tabela 1) foram consistentes com as estimativas de estudos que utilizaram populações biparentais, como os de Perez-Brito et al. (2001) e Robertson-Hoyt et al. (2006), e, também, com outra pesquisa que utilizou algumas linhagens europeias para o estudo da resistência à podridão de espiga (BOLDUAN et al., 2009). Em estudos de mapeamento associativo realizado por Zila et al. (2013) e Zila et al. (2014), as estimativas de herdabilidade também foram coincidentes, apesar de serem um pouco mais elevadas que as deste.

No estudo de associação genômica, foram identificados onze DArTs associados significativamente à resistência à podridão de espiga. Desses onze DArTs, cinco foram identificados somente por meio do método BSSV e dois foram identificados somente pelo método de modelo misto. E quatro DArTs foram identificados tanto no método BSSV quanto no método de modelo misto (Tabelas 2 e 3).

Os marcadores que foram significativamente associados à resistência à podridão de espiga neste estudo foram diferentes dos marcadores identificados em outros estudos. Zila et al. (2013) avaliaram 267 linhagens com 47.445 SNPs, e identificaram quatro SNPs associados à resistência, sendo três em uma análise individual e um na análise conjunta. Os SNPs foram localizados nos cromossomos 1 (posição 63,540,590 bp), 5 (posição 30,997,717 bp) e 9 (posição 151,295,233 bp, análise individual e conjunta). Nenhum dos marcadores identificados neste estudo foram localizados próximos dos marcadores identificados por Zila et al. (2013). Os marcadores mais próximos foram localizados a uma distância de 14 Mbp. Não foi possível identificar os mesmos marcadores em ambos os estudos,

pois nenhum dos marcadores utilizados por Zila et al. (2013) estavam presentes no conjunto de marcadores utilizados neste estudo. Apesar dos marcadores serem diferentes, alguns marcadores utilizados neste estudo foram localizados bem próximos dos marcadores que Zila et al. (2013) identificaram associados com a resistência à podridão.

Em um estudo posterior, com um painel de linhagens muito maior (1.687 linhagens) e com mais marcadores (200.978 SNPs), Zila et al. (2014) identificaram um total de sete SNPs significativamente associados à resistência à podridão de espiga. Desse total de marcadores associados, dois foram identificados quando foi feita a análise com todas as linhagens (1.689 linhagens) e cinco foram identificados quando a análise foi realizada somente com as linhagens que foram testadas em três anos (737 linhagens). Na análise total, os SNPs foram localizados nos cromossomos 5 (posição 64,771,372 bp) e 9 (19,532,465 bp). Os marcadores associados neste estudo também não foram os mesmos identificados por Zila et al. (2014), entretanto, não foram localizados tão distantes uns dos outros. No cromossomo 5, os marcadores associados nos estudos ficaram 9 Mbp distantes e, no cromossomo 9, os marcadores ficaram a uma distância de 7 Mbp. Na análise com as linhagens que foram avaliadas em três anos, Zila et al. (2014) identificaram 5 SNPs associados (posição 7,566,354; 7,618,125; 7,618,284; 9,353,851 e 124,930,996 bp), todos no cromossomo 4. Apesar de, neste estudo, ter sido identificado 1 DArT (posição 18,591,461 bp) associado no cromossomo 4, este foi diferente e localizado em uma posição distante dos SNPs identificados por Zila et al. (2014).

A diferença entre os marcadores associados neste estudo com os estudos anteriores (ZILA et al., 2013, 2014) pode ser devido às diferentes linhagens utilizadas nos estudos e aos tamanhos de amostra utilizados, bem como devido ao número e aos diferentes marcadores utilizados, de modo que podem ocorrer

diferenças nos alelos amostrados e diferenças nos efeitos dos marcadores nos diferentes estudos.

Entre os marcadores associados significativamente à resistência, para o método BSSV, houve coincidência entre os DArTs associados nas análises individuais com a análise conjunta. Os DArTs do cromossomo 1 (posição 136,272,488 bp) e em um grupo de cromossomo não identificado (Ni) foram associados significativamente à resistência à podridão de espiga tanto na análise individual de Lavras quanto na análise conjunta (Tabela 2). Essa coincidência entre os DArTs que foram associados significativamente nas análises conjunta e individuais mostra coerência nos resultados. Embora alguns DArTs tenham sido coincidentes nas análises, os demais DArTs não foram coincidentes entre as análises, sendo que essas diferenças nos resultados dos DArTs significativamente associados podem ser devido ao grande efeito da interação, pois os ambientes onde as linhagens foram avaliadas são bastante distintos. Além disso, esse caráter tem um feito poligênico (CLEMENTS et al., 2004; DE LEÓN; PANDEY, 1989; KING; SCOTT, 1981; NANKAM; PATAKY, 1996) e as variantes causais podem ter um efeito muito pequeno, dificultando, assim, a detecção de marcadores significativamente associados à resistência à podridão de espiga (ZILA et al., 2013, 2014).

As análises realizadas pelo método de modelo misto apresentaram consistência entre os DArTs associados nas análises, pois alguns DArTs associados com resistência à doença identificados na análise conjunta também foram associados nas análises individuais de Lavras e Uberlândia. Os DArTs que foram associados à resistência no cromossomo 7 (posição 124,220,838 bp) e no grupo de cromossomo não identificado (Ni) foram identificados tanto na análise individual de Lavras quanto na análise conjunta. E o DArT identificado no cromossomo 1 (posição 78,131,634 bp) foi associado significativamente nas análises de Uberlândia e na conjunta. Esse fato mostra uma consistência nos

resultados obtidos, pois muitos DArTs foram associados significativamente à resistência à podridão de espiga tanto na análise conjunta quanto nas análises individuais.

Além disso, alguns DArTs foram associados à resistência à doença tanto pelo método BSSV quanto pelo método de modelo misto. Nas análises conjuntas, os DArTs do cromossomo 7 (posição 78,131,634 bp) e de um grupo de cromossomo não identificado foram associados significativamente por ambos os métodos. Para as análises de Lavras, também dois DArTs foram associados à resistência pelos dois métodos, sendo os DArTs identificados nos cromossomos 7 (posição 124,220,838 bp) e em um grupo de cromossomo não identificado (Ni). Nas análises de Uberlândia, apenas um DArT identificado no cromossomo 5 (posição 55,225,453 bp) foi associado pelos dois métodos. Esse fato mostra que tanto o método BSSV quanto o método de modelo misto são capazes de identificar marcadores associados com a resistência à podridão de espiga, e também nos mostra que ambos os métodos podem ser utilizados de uma forma que um complemente o outro.

Existe uma grande dificuldade na avaliação de doenças em regiões de clima tropical, sendo isso um grande desafio para os melhoristas. Esse fato mostra a importância de se estudar a interação genótipos x ambientes, pois a interação dificulta as análises e pode contribuir na redução da habilidade de detecção de marcadores significativos no mapeamento. Nesses estudos, a forte interação pode ser causada pelas diferenças climáticas entre os dois ambientes que foram utilizados na avaliação das linhagens. O clima em Lavras é classificado como clima temperado úmido com inverno seco e verão quente com uma maior pluviosidade, enquanto o clima de Uberlândia é classificado como clima tropical com estação seca, geralmente no inverno. Pereira et al. (2015) utilizaram as mesmas linhagens de milho avaliadas neste trabalho para verificar a relação da produção de espigas com doenças foliares e doenças de espiga. Esses autores

observaram efeito de interação genótipos por ambientes significativo, o qual influenciou o comportamento dos patógenos nos diferentes locais e concluíram que a interação foi devido à diferença climática entre os dois ambientes (Lavras e Uberlândia). Dessa forma, a diferença entre as análises ressalta a importância da realização das análises individuais de mapeamento quando se tem uma forte influência da interação genótipos x ambiente.

Embora os estudos anteriores realizados por Zila et al. (2013, 2014) tenham utilizado um maior número de linhagens e marcadores nos estudos de associação, eles identificaram um menor número de regiões gênicas associadas com a resistência à doença. Neste estudo, foram identificados marcadores associados com a resistência à podridão de espiga em quase todos os cromossomos do milho.

Outro fator que pode influenciar na identificação de marcadores associados à resistência à doença é a frequência do menor alelo (MAF – *Minor Allele Frequency*). Na literatura, é recomendado fazer a correção pela frequência do menor alelo, pois loci com alelos de menor frequência são mais difíceis de serem detectados devido à baixa replicabilidade, e também têm maior probabilidade de resultarem em associações espúrias (TABANGIN; WOO; MARTIN, 2009). Dessa forma, a frequência do menor alelo poderia afetar a probabilidade de uma associação não verdadeira atingir o nível de significância sob a hipótese de nulidade. Além disso, os alelos de menor frequência tendem a ser associados com a estrutura populacional, uma vez que são limitados às subpopulações. Porém, quando se faz a correção para a estrutura da população, o potencial de detecção dos alelos de menor frequência é ainda menor (ZILA et al., 2014).

Neste estudo, não foi realizada a remoção dos DArTs que apresentaram frequência do menor alelo menores que 0,05 em nenhum dos métodos utilizados para realizar o mapeamento associativo para resistência. Foi decidido não realizar

a remoção dos marcadores com frequências de alelos menores pois, quando foi realizada uma análise preliminar com o método de modelo misto com a retirada dos alelos de menor frequência, algumas análises não apresentaram associações significativas. Outra justificativa para a realização das análises com os DArTs com alelos de menor frequência é que esses alelos raros podem fazer parte da arquitetura genética da resistência à podridão de espiga. Assim, a retirada dos marcadores com alelos raros faria com que associações que deveriam ser significativas fossem descartadas. Pelos valores das frequências do menor alelo, foi observado que, mesmo não fazendo a retirada dos alelos com frequência menor que 0,05, o método BSSV foi capaz de associar marcadores com a resistência.

Já o método de modelo misto associou marcadores com a frequência do menor alelo próximas a 0,05 (Tabela 4). No caso das análises que utilizaram o BSSV, a retirada dos alelos de menor frequência pouco influenciou na significância de DArTs associados com a resistência à doença. Porém, no caso do modelo misto, as frequências do menor alelo podem ter influenciado na associação dos marcadores, resultando em associações espúrias. No futuro, serão necessários mais estudos e também novas ferramentas para se ter uma maior confiabilidade nos resultados

O principal fator do mapeamento associativo é o desequilíbrio de ligação entre os loci funcionais e os marcadores, e, conseqüentemente, tem grande impacto no mapeamento, pois a alta resolução de mapeamento é dependente do desequilíbrio de ligação no genoma (YU et al., 2006). Neste estudo, foi observado um rápido declínio do desequilíbrio de ligação entre os DArTs (Figura 4). Esse fato está de acordo com trabalhos realizados por outros autores que trabalharam com diferentes populações e avaliaram as mesmas características. Remington et al. (2001) utilizaram diversas linhagens de milho e demonstraram que o desequilíbrio de ligação decai rapidamente em pequena distância ($r^2 < 0,1$ em 1500 bp). Romay et al. (2013) utilizaram um grande número de acessos de linhagens de

milho e também observaram que o desequilíbrio cai muito rapidamente em uma pequena distância ($r^2 \leq 0,2$ em 1 kbp). Samayoa et al. (2015) avaliaram várias características relacionadas à resistência à broca do milho e também observaram o rápido decaimento do desequilíbrio de ligação entre as marcas associadas à resistência e as marcas adjacentes ($r^2 \leq 0,1$).

Esse fato também foi observado quando foram feitos os gráficos do tipo *heatmap* para os DArTs que foram associados significativamente com a resistência à podridão de espiga. Foi observado que existiam alguns blocos gênicos, os quais apresentavam certo desequilíbrio de ligação, mas, de fato, não foi encontrado nenhum DArT com alto desequilíbrio de ligação com outro marcador e que dentro desse intervalo estivessem contidos alguns genes que poderiam ter alguma relação com a resistência à doença.

A estrutura populacional pode ser também um fator limitante para o estudo de associação quando essa não é levada em conta na análise. Quando a população é estruturada e os alelos das subpopulações apresentam frequências muito diferentes, pode ocorrer que alguns loci que não apresentam qualquer relação ou efeito sobre a característica avaliada apresentem associações significativas (ERSOZ; YU; BUCKLER, 2007). Embora as linhagens utilizadas neste estudo apresentem certa estrutura populacional (DOS SANTOS et al., 2016) (Figura 1), quando foi realizada uma análise para determinar o número de componentes principais que explicassem adequadamente a estrutura populacional, o número de componentes principais para incluir no modelo de GWAS foi 0 (Tabela 6). A não inclusão de um componente principal na análise de mapeamento indica que a estrutura da população utilizada no trabalho não influencia na análise e, provavelmente, a estrutura populacional observada por dos Santos et al. (2016) é devido aos diferentes *backgrounds* genéticos das linhagens.

Na comparação dos valores fenotípicos com os valores dos marcadores, foi possível observar que, com os marcadores associados pelo método BSSV, os

valores dos marcadores foram consistentes com os valores fenotípicos apresentados pelas linhagens. Já para os marcadores associados pelo método de modelo misto, os valores dos marcadores não foram consistentes com os valores fenotípicos das linhagens. Somente quando os valores fenotípicos eram extremos, apenas alguns marcadores que foram identificados pelo método de modelo misto foram mais consistentes com os valores fenotípicos das linhagens (Figuras 5, 6 e 7).

Além disso, por meio do dendrograma foi possível observar a formação de dois grupos, um grupo com as linhagens que apresentaram os menores valores de proporção de espigas doentes (linhagens resistentes) e outro grupo com as linhagens com os maiores valores de proporção (linhagens suscetíveis). Foi observado que os marcadores associados significativamente à resistência, de modo geral, também foram divididos em dois grupos. Os marcadores que foram associados significativamente por meio do método de modelo misto não apresentam esse comportamento, possivelmente porque as análises por meio do método de modelo misto conseguiram captar apenas associações de marcadores com grandes efeitos.

Apesar do método BSSV parecer mais consistente devido ao comportamento semelhante ao dos fenótipos, é necessária certa cautela, pois, nesse método, não é considerada a estrutura da população. Assim, o comportamento dos marcadores pode ser somente devido ao diferente *background* das linhagens avaliadas neste estudo.

4.2. Genes candidatos

Para a identificação de genes candidatos responsáveis pela resistência à podridão de espiga causada pelo fungo *Fusarium verticillioides*, foi utilizado o genoma de referência da linhagem B73 por meio do *MaizeGDB*

(LAWRENCE et al., 2008). Do total de onze DArTs que foram associados significativamente à resistência, apenas um foi localizado dentro de um gene. Os demais DArTs não foram localizados em regiões codificantes, mas estavam localizados bem próximos a alguns genes (Tabela 4 e 5).

Pelo método BSSV na análise conjunta, os DArTs associados foram identificados nos cromossomos 1 e 7. Esses DArTs não foram localizados dentro de nenhum gene. Embora os marcadores não tenham sido localizados dentro de uma região codante, três destes foram encontrados próximos de genes funcionais, como o gene responsável pelo transporte de proteínas (cromossomo 1) e os genes responsáveis pelo processo de oxidação e redução e pela atividade catalítica na célula (cromossomo 7). Apesar de dois DArTs identificados no estudo de associação não estarem próximos de genes que poderiam ter alguma função direta na resistência, todos foram localizados próximos a genes que possuem funções vitais para o funcionamento celular e que estes, de alguma forma, direta ou indiretamente, podem atuar na resistência ao patógeno. O DArT identificado no cromossomo 7 foi localizado bem próximo do gene responsável pelo processo de oxidação e redução na célula e esse gene pode ter um papel fundamental na resistência à podridão de espiga (Tabela 4).

Na análise de Lavras pelo método BSSV, foram identificados DArTs associados à resistência nos cromossomos 1, 7 e 8 (Tabela 4). Nenhum desses marcadores foram localizados dentro de genes, mas, também, foram localizados próximos a genes com funções conhecidas, exceto o DArT identificado no cromossomo 7, que foi localizado próximo a uma ORF. Os DArTs identificados nos cromossomos 1 e 8 foram localizados próximos aos genes responsáveis pela regulação da transcrição, transporte de proteínas e transporte transmembrana. Mesmo que não tenham funções diretas relacionadas com a resistência, esses genes podem ter alguma importância na resistência à doença. O DArT identificado no cromossomo 7 foi localizado próximo a um gene responsável pela

atividade hormonal. Na análise de Uberlândia, os DArTs identificados nos cromossomos 4 e 5 foram localizados próximos a genes que possuem funções básicas na célula, como transporte de aminoácidos e atividade da adenosilhomocisteinase (Tabela 4).

Nas análises realizadas pelo método de modelo misto, três DArTs associados à resistência foram localizados próximos a genes e um DArT foi localizado na região intrônica de um gene. Todos os DArTs identificados por esse método foram localizados próximos a ORFs, que são genes com pouca ou nenhuma informação sobre. Além disso, todos os marcadores que foram associados à resistência à doença por meio desse método não apresentaram nenhum desequilíbrio de ligação com os demais marcadores (Tabela 5).

Todos os marcadores que foram associados significativamente pelo método BSSV foram localizados próximos a genes que têm funções gerais e também próximos a genes com alguma função que possa ter alguma relação com a resistência à podridão de espiga. Além disso, alguns marcadores que foram associados com a resistência à doença foram identificados em desequilíbrio de ligação com marcadores mais distantes, sendo que a região entre os marcadores continha genes que poderiam estar relacionados com a resistência. Os marcadores identificados pelo método de modelo misto não permitiram fazer nenhuma inferência sobre os genes que foram associados significativamente com a resistência à podridão, e também os marcadores não estavam em desequilíbrio com nenhuma outra marca.

Apesar de que as vias metabólicas e da genética da resistência à podridão de espiga causada por *Fusarium verticillioides* sejam ainda completamente desconhecidas, o mapeamento associativo tem-se mostrado uma ótima ferramenta para auxiliar na identificação de genes que possam estar envolvidos na resistência. Neste trabalho, vários genes foram identificados e alguns podem ter relação direta com a resistência à podridão de espiga. Um entendimento melhor da genética da

resistência assim como estudos mais aprofundados nas vias metabólicas dos genes que estão envolvidos na resistência serão necessários.

5. CONCLUSÃO

Foram identificados 11 marcadores associados à resistência à doença tanto pelo método *Bayesian Stochastic Search Variable* (BSSV) quanto pelo modelo clássico de estudo de associação (modelo misto). Além disso, 4 DArTs foram associados significativamente tanto pelo método BSSV quanto pelo método de modelo misto. Foram identificados alguns genes que podem estar relacionados com a resistência, mas estudos mais aprofundados sobre as vias metabólicas da resistência são necessários no futuro. O método BSSV, que é um método utilizado na seleção genômica, foi eficiente no estudo de associação genômica, podendo ser utilizado como uma ferramenta complementar ao estudo de associação que utiliza o método de modelo misto.

REFERÊNCIAS

ALI, M. L.; TAYLOR, J. H.; JIE, L. Molecular mapping of QTLs for resistance to Gibberella ear rot, in corn, caused by *Fusarium graminearum*. **Genome**, v. 48, n. 3, p. 521–533, 2005.

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

BARSH G. S., COPENHAVER G. P., GIBSON G., WILLIAMS S. M. Guidelines for Genome-Wide Association Studies. **PLoS Genet**, v.8, n. 7, p. 7-8, 2012.

BOLDUAN, C. et al. Genetic variation for resistance to ear rots and mycotoxins contamination in early European maize inbred lines. **Crop Science**, v. 49, n. 6, p. 2019–2028, 2009.

BRODERS, K. D. et al. Evaluation of *Fusarium graminearum* Associated with Corn and Soybean Seed and Seedling Disease in Ohio. **Plant Disease**, v. 91, n. 9, p. 1155–1160, 2007.

BULLERMAN, L. B. Occurrence of *Fusarium* and fumonisins in food grains and in foods. In: JACKSON, L. S.; DEVRIES, J. W.; BULLERMAN, L. B. (Eds.). **Fumonisin in Food. Advances in Experimental medicine and Biology**. Boston, MA: Ed: Springer US. v. 392, 1996. p. 27–38.

BUSH, B. J. et al. Infection and Fumonisin Production by *Fusarium verticillioides* in Developing Maize Kernels. **Phytopathology**, v. 94, n. 1, p. 88–93, 2004.

CLEMENTS, M. J. et al. Evaluation of Inoculation Techniques for *Fusarium* Ear Rot and Fumonisin Contamination of Corn. **Plant Disease**, v. 87, n. 2, p. 147–153, 2003.

CLEMENTS, M. J. et al. Sources of resistance to fumonisin accumulation in grain and fusarium ear and kernel rot of corn. **Phytopathology**, v. 94, n. 3, p. 251–260, 2004.

COSTA, R. V. DA et al. Incidência de *Colletotrichum graminicola* em colmos de genótipos de milho. **Summa Phytopathologica**, v. 36, n. 2, p. 122–128, 2010.

DE LEÓN, C.; PANDEY, S. Improvement of Resistance to Ear and Stalk Rots and Agronomic Traits in Tropical Maize Gene Pools. **Crop Science**, v. 29, n. 1, p. 12–17, 1989.

DENTI, E. A.; REIS, E. M. Levantamento de fungos associados às podridões do colmo e quantificação de danos em lavouras de milho do Planalto Médio gaúcho e dos Campos Gerais do Paraná. **Fitopatologia Brasileira**, v. 28, n. 6, p. 585–590, 2003.

DESTA, Z. A.; ORTIZ, R. Genomic selection: genome-wide prediction in plant improvement. **Trends in Plant Science**, v. 19, n. 9, p. 592–601, 2014.

DOS SANTOS, J. P. R. et al. Genomic selection to resistance to *Stenocarpella maydis* in maize lines using DArTseq markers. **BMC Genetics**, v. 17, n. 86, p. 1–10, 2016.

ELLER, M. S.; HOLLAND, J. B.; PAYNE, G. A. Breeding for Improved Resistance To Fumonisin Contamination in Maize. **Toxin Reviews**, v. 27, n. 3–4, p. 371–389, 2008.

EMBRAPA. **O Cultivo do Milho: Doenças**. Disponível em: <http://www.cnpms.embrapa.br/publicacoes/milho_6_ed/doencas.htm>. Acesso em: 2 jul. 2016.

ENDELMAN, J. B. Ridge regression and other kernels for genomic selection with R package rrBLUP. **Plant Genome**, v.4, n.3, p. 250–255, 2011.

ERSOZ, E. S.; YU, J.; BUCKLER, E. S. Applications of Linkage Disequilibrium and Association Mapping in Crop Plants. In: VARSHNEY, R. K.; TUBEROSA, R. (Eds.). **Genomics-Assisted Crop Improvement: Genomics Approaches and Platforms**. Ed: Springer Netherlands, 2007, v. 1, p. 97–119.

GARRICK, D. J.; FERNANDO, R. L. Implementing a QTL Detection Study (GWAS) Using Genomic Prediction Methodology. In: GONDRO, C.; VAN DER WERF, J.; HAYES, B. (Eds.). **Genome-Wide Association Studies and Genomic Prediction**. Ed: Humana Press, 2013. p. 275–298.

HEFFNER, E. L.; SORRELLS, M. E.; JANNINK, J.-L. Genomic Selection for Crop Improvement. **Crop Science**, v. 49, n. 1, p. 1–12, 2009.

HOLLAND, J. B. Implementation of molecular markers for quantitative traits in breeding programs - challenges and opportunities. **Proceedings of the 4th International Crop Science Congress**, p. 1–13, 2004.

KILIAN, A. et al. Diversity Arrays Technology: A Generic Genome Profiling Technology on Open Platforms. In: POMPANON, F.; BONIN, A. (Eds). **Data Production and Analysis in Population Genomics: Methods and Protocols**. New York, NY: Ed: Human Press. v. 888, 2012, p. 67–89

KING, S. B.; SCOTT, G. E. Genotypic differences in maize to kernel infection by *Fusarium moniliforme*. **Phytopathology**, v. 71, p. 1245–1247, 1981.

LAWRENCE, C. J. et al. MaizeGDB: The Maize Model Organism Database for Basic, Translational, and Applied Research. **International Journal of Plant Genomics**, v. 2008, p. 1–10, 2008.

LIPKA, A. E. et al. GAPIT: genome association and prediction integrated tool. **Bioinformatics**, v. 28, n. 18, p. 2397–2399, 2012.

MANGIN et al. Novel Measures of linkage Disequilibrium that correct for the bias due to population structure and relatedness. **Heredity**, v. 108, p. 285–291, 2012.

MARASAS, W. F. O. Fumonisin: History, World-Wide occurrence and impact. In: JACKSON, L. S.; DEVRIES, J. W.; BULLERMAN, L. B. (Eds.). **Fumonisin in Food. Advances in Experimental medicine and Biology**. Boston, MA: Ed. Springer US. v. 392, 1996, p. 1–17.

MENDES, M. C. **Micotoxinas, aspectos químicos e bioquímicos relacionados a grãos ardidos em híbridos de milho**. 2009. 106 p. Tese (Doutorado em Agronomia – Fitotecnia) – Universidade Federal de Lavras, Lavras, 2009. Disponível em: <<http://repositorio.ufla.br/jspui/handle/1/3373>>

MONACO, M. K. et al. Maize Metabolic Network Construction and Transcriptome Analysis. **Plant Genome**, v. 6, n. 1, p. 12, 2013.

MUNKVOLD, G. P.; HELLMICH, R. L.; RICE, L. G. Comparison of Fumonisin Concentrations in Kernels of Transgenic Bt Maize Hybrids and Nontransgenic Hybrids. **Plant Disease**, v. 83, n. 2, p. 130–138, 1999.

MUNKVOLD, G. P.; HELLMICH, R. L.; SHOWERS, W. B. Reduced fusarium ear rot and symptomless infection in kernels of maize genetically engineered for European corn borer resistance. **Phytopathology**, v. 87, n. 10, p. 1071–1077, 1997.

NANKAM, C.; PATAKY, J. K. Resistance to kernel infection by *Fusarium moniliforme* in the sweet corn inbred IL125b. **Plant Disease**, v. 80, n. 5, p. 593–598, 1996.

PARADIS E, CLAUDE J, STRIMMER K. APE: analyses of phylogenetics and evolution in R language. **Bioinformatics**, v. 20, n. 2, p. 289-290, 2004.

PAVÃO, A. R.; FERREIRA FILHO, J. B. DE S. Impactos econômicos da introdução do milho Bt11 no Brasil: uma abordagem de equilíbrio geral inter-regional. **Revista de Economia e Sociologia Rural**, v. 49, n. 1, p. 81–108, 2011.

PEREIRA, G. S. et al. Indirect selection for resistance to ear rot and leaf diseases in maize lines using biplots. **Genetics and Molecular Research**, v. 14, n. 3, p. 11052–11062, 2015.

PEREZ-BRITO, D. et al. QTL mapping of *Fusarium moniliforme* ear rot resistance in high land maize, Mexico. **Agrociencia**, v. 35, p. 181–196, 2001.

POWELL J. E., VISSCHER P. M., GODDARD M. E. Reconciling the analysis of IBD and IBS in complex trait studies. **Nature Reviews Genetics**, v.11, n.11, p. 800-805, 2010.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2016. Disponível em: <<http://www.R-project.org/>>.

REMINGTON, D. L. et al. Structure of linkage disequilibrium and phenotypic associations in the maize genome. **Proceedings of the National Academy of Sciences of the United States of America**, v. 98, n. 20, p. 11479–11484, 2001.

ROBERTSON-HOYT, L. A. et al. QTL mapping for fusarium ear rot and fumonisin contamination resistance in two maize populations. **Crop Science**, v. 46, n. 4, p. 1734–1743, 2006.

ROBERTSON, L. A. et al. Heritabilities and correlations of fusarium ear rot resistance and fumonisin contamination resistance in two maize populations. **Crop Science**, v. 46, n. 1, p. 1420, 2006.

ROMAY, M. C. et al. Comprehensive genotyping of the USA national maize inbred seed bank. **Genome Biology**, v. 14, n. 6 (R5), p. 1-18, 2013.

SAITOU N, NEI M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, v. 4, n. 4, p. 406-425, 1987.

SAMAYOA, L. F. et al. Genome-wide association study reveals a set of genes associated with resistance to the Mediterranean corn borer (*Sesamia nonagrioides* L.) in a maize diversity panel. **BMC Plant Biology**, v. 15, n. 1, p. 35, 2015.

SCHNABLE, P. S. et al. The B73 Maize Genome: Complexity, Diversity, and Dynamics. **Science**, v. 326, n. 5956, p. 1112–1115, 2009.

SHIN J-H, BLAY S, McNENEY B, GRAHAM J. LDheatmap: An R Function for Graphical Display of Pairwise Linkage Disequilibria Between Single Nucleotide Polymorphisms. **Journal of Statistical Software**, v. 16, n. October, p. 1-9, 2006.

SINGH, B. D.; SINGH, A. K. **Marker-Assisted Plant Breeding: Principles and Practices**. Ed: New Delhi: Springer India. 2015. 514 p.

TABANGIN, M. E.; WOO, J. G.; MARTIN, L. J. The effect of minor allele frequency on the likelihood of obtaining false positives. **BMC Proceedings**, v. 3 n. Suppl 7 (S41), p. 1-4, 2009.

VAN RADEN P. M. Efficient Methods to Compute Genomic Predictions. **J. Dairy Sci.**, v. 91, n.418, p. 4414-4423, 2008.

YI, N.; GEORGE V.; ALISSON, D. B. Stochastic search variable selection for identifying multiple quantitative trait loci. **Genetics**, v. 164, n. 3, p. 1129–1138, 2003.

YU, J. et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. **Nature Genetics**, v. 38, n. 2, p. 203–208, 2006.

ZILA, C. T. et al. A Genome-Wide Association Study Reveals Genes Associated with Fusarium Ear Rot Resistance in a Maize Core Diversity Panel. **G3: Genes|Genomes|Genetics**, v. 3, n. 11, p. 2095–2104, 2013.

ZILA, C. T. et al. Genome-wide association study of Fusarium ear rot disease in the U.S.A. maize inbred line collection. **BMC Plant Biology**, v. 14, n. 1, p. 372, 2014.

APÊNDICE

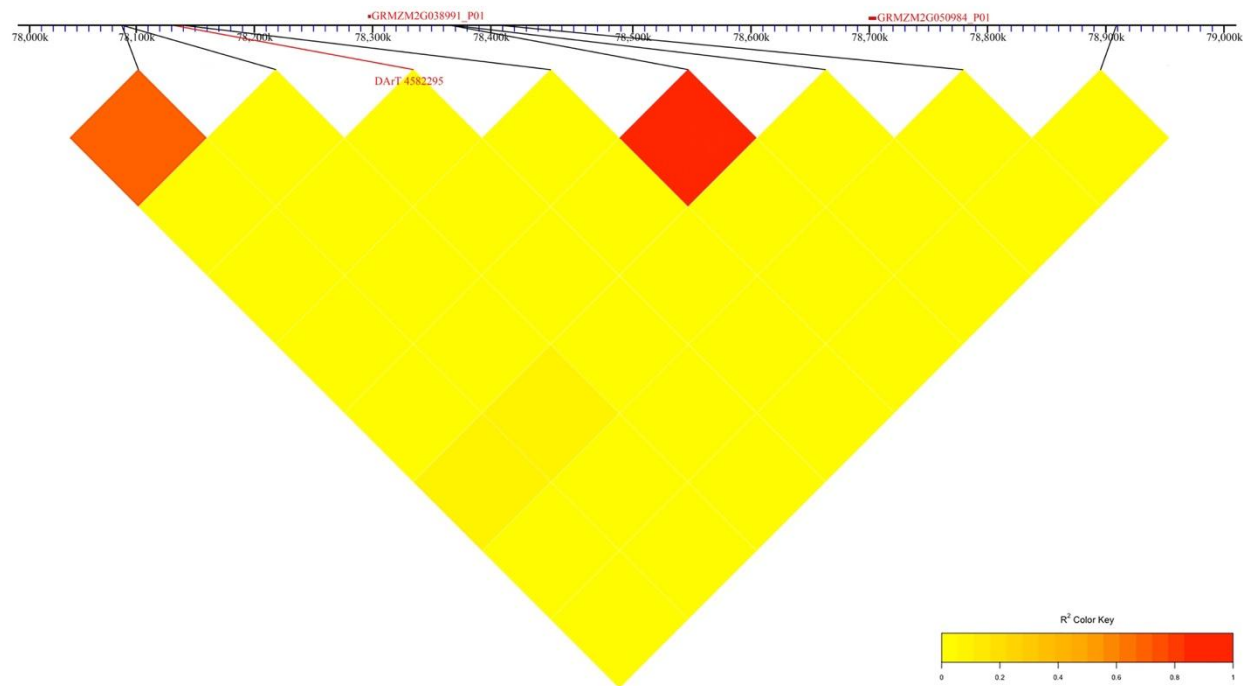


Figura 8 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1 Mbp do cromossomo 1 para o marcador 4582295 (posição 78,131,634 bp), que foi significativamente associado na análise conjunta do método BSSV e nas análises conjunta e de Uberlândia pelo método de modelo misto.

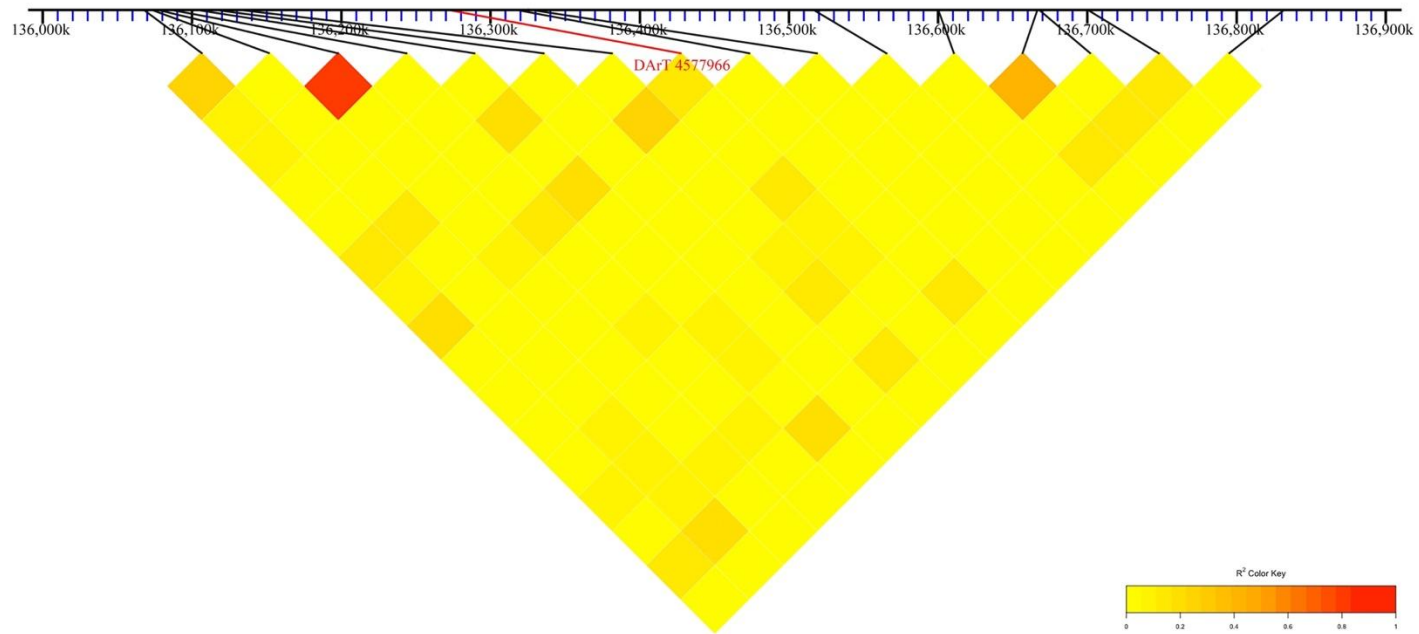


Figura 9 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 0,9 Mbp do cromossomo 1 para o marcador 4577966 (posição 136,272,488 bp), que foi significativamente associado nas análises conjunta e de Lavras pelo método BSSV.

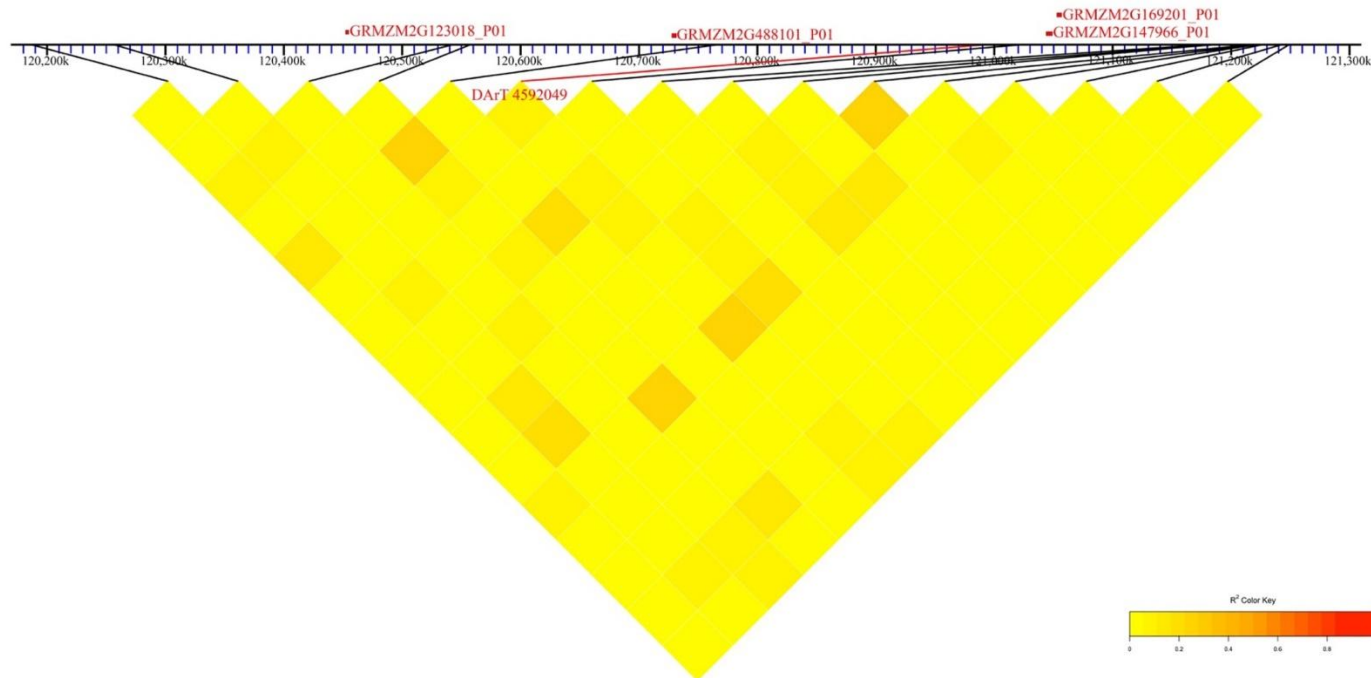


Figura 10 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1,1 Mbp do cromossomo 7 para o marcador 4592042 (posição 120,994,839 bp), que foi significativamente associado na análise conjunta pelo método BSSV.

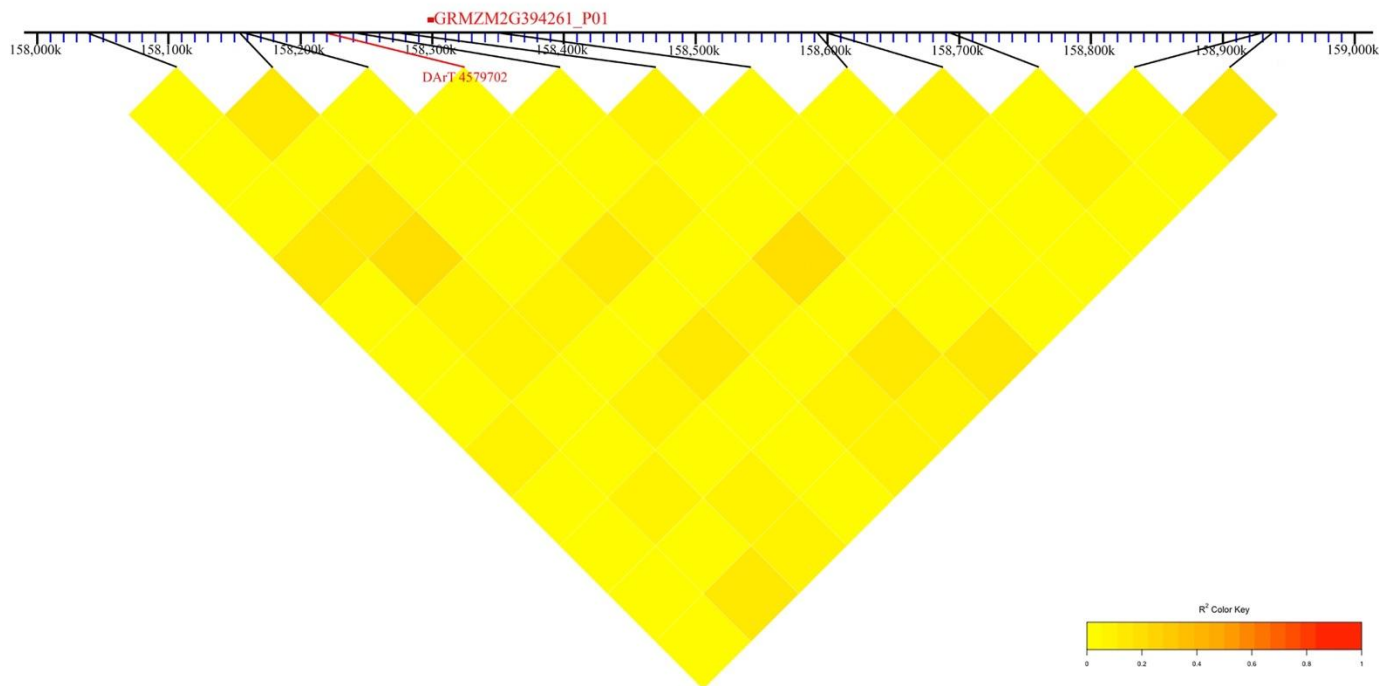


Figura 11 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1 Mbp do cromossomo 7 para o marcador 4579702 (posição 158,219,730 bp), que foi significativamente associado na análise conjunta pelo método BSSV.

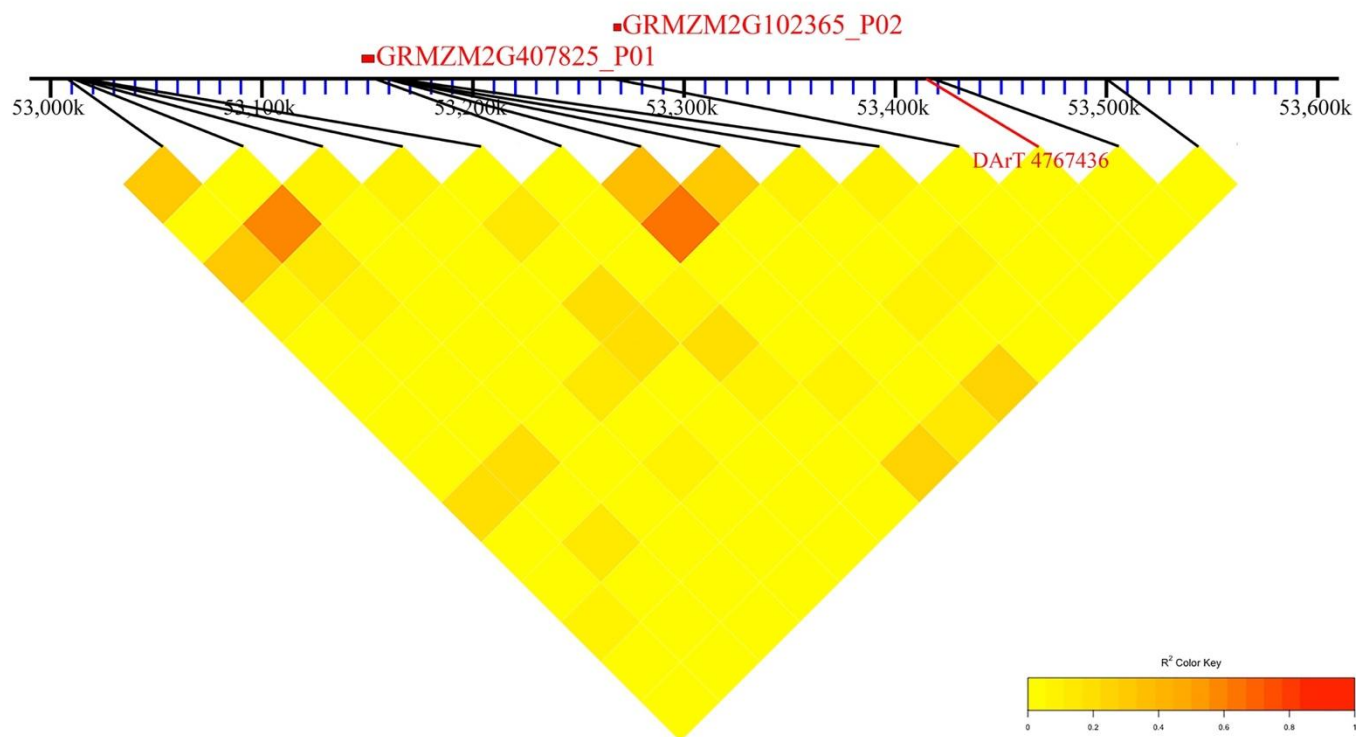


Figura 12 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 0,6 Mbp do cromossomo 1 para o marcador 4767436 (posição 53,415,835 bp), que foi significativamente associado na análise de Lavras pelo método BSSV.

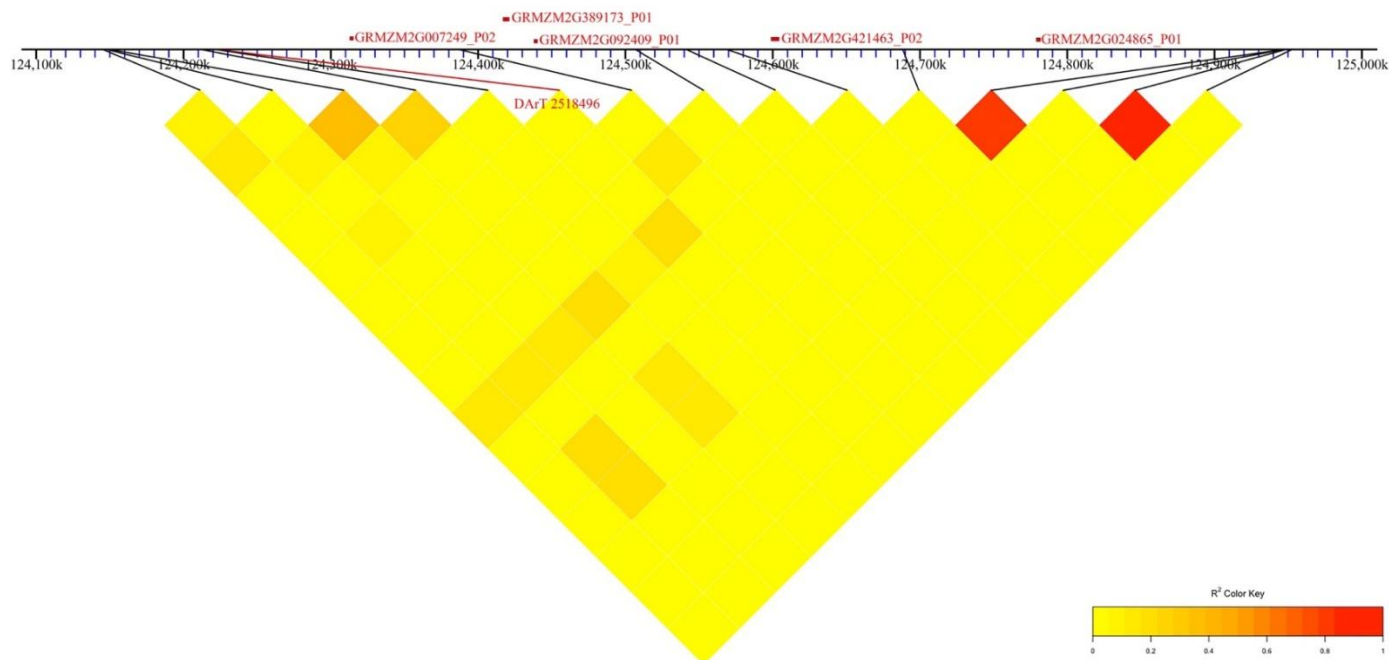


Figura 13 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 0,9 Mbp do cromossomo 7 para o marcador 2518496 (posição 124,220,878 bp), que foi significativamente associado na análise de Lavras pelo método BSSV e nas análises conjunta de Lavras pelo método de modelo misto.

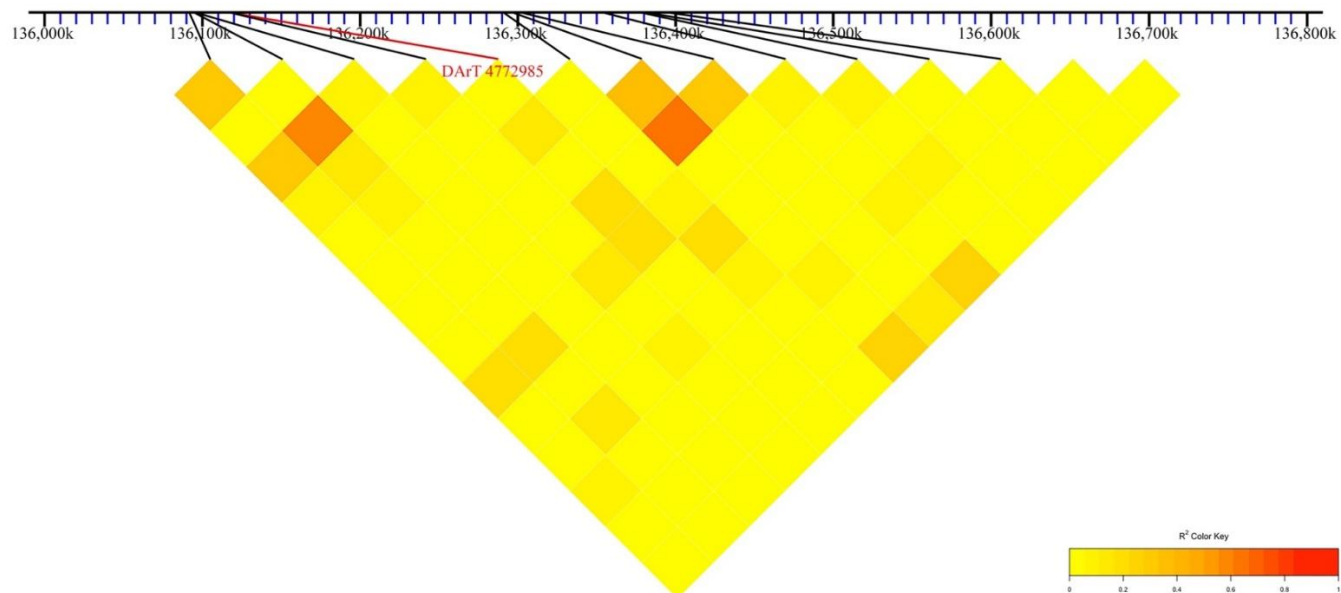


Figura 14 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 0,8 Mbp do cromossomo 8 para o marcador 4772985 (posição 136,137,749 bp), que foi significativamente associado na análise de Lavras pelo método BSSV.

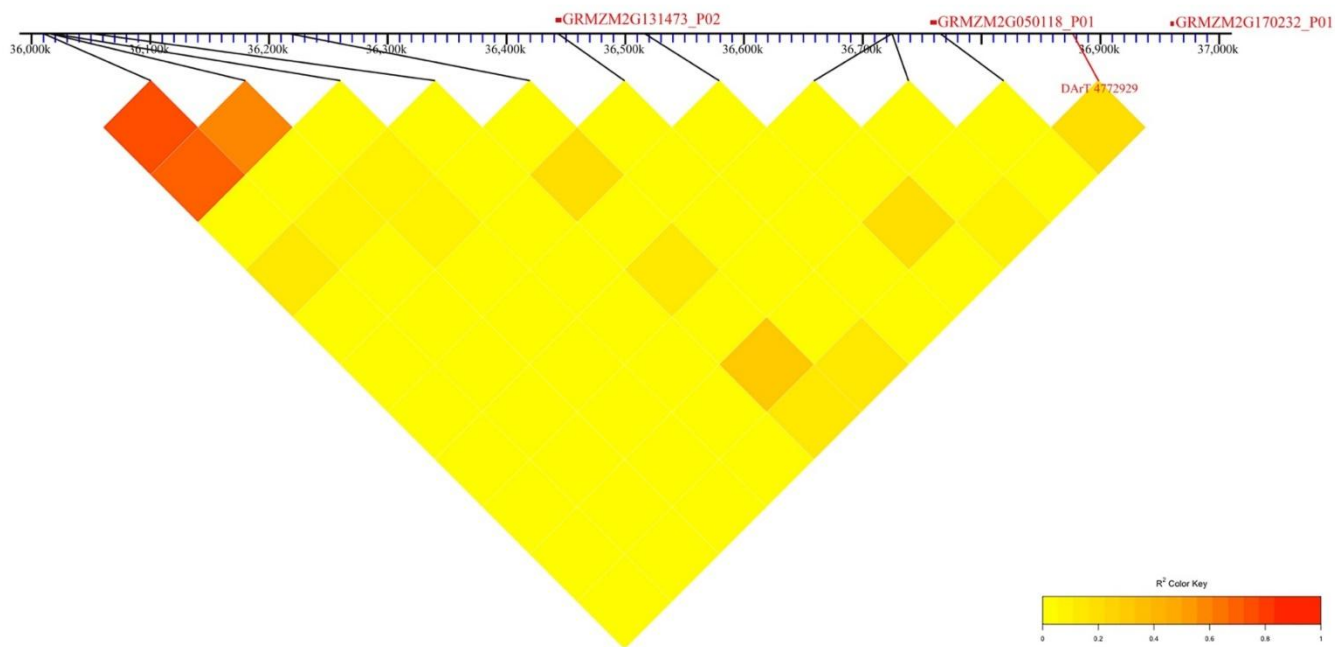


Figura 15 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1 Mbp do cromossomo 4 para o marcador 4772929 (posição 36,878,586 bp), que foi significativamente associado na análise de Uberlândia pelo método BSSV.

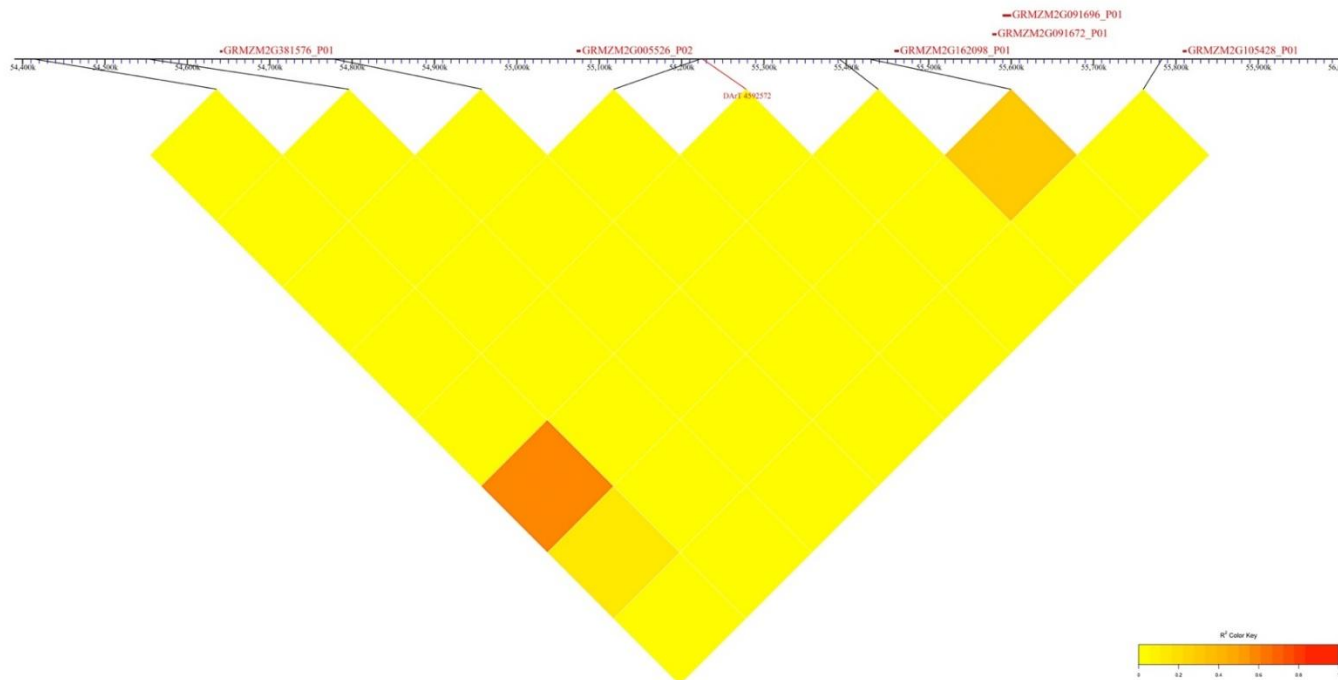


Figura 16 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1,6 Mbp do cromossomo 5 para o marcador 4592572 (posição 55,225,453 bp), que foi significativamente associado na análise de Uberlândia pelo método BSSV e na análise de Uberlândia pelo método de modelo misto.

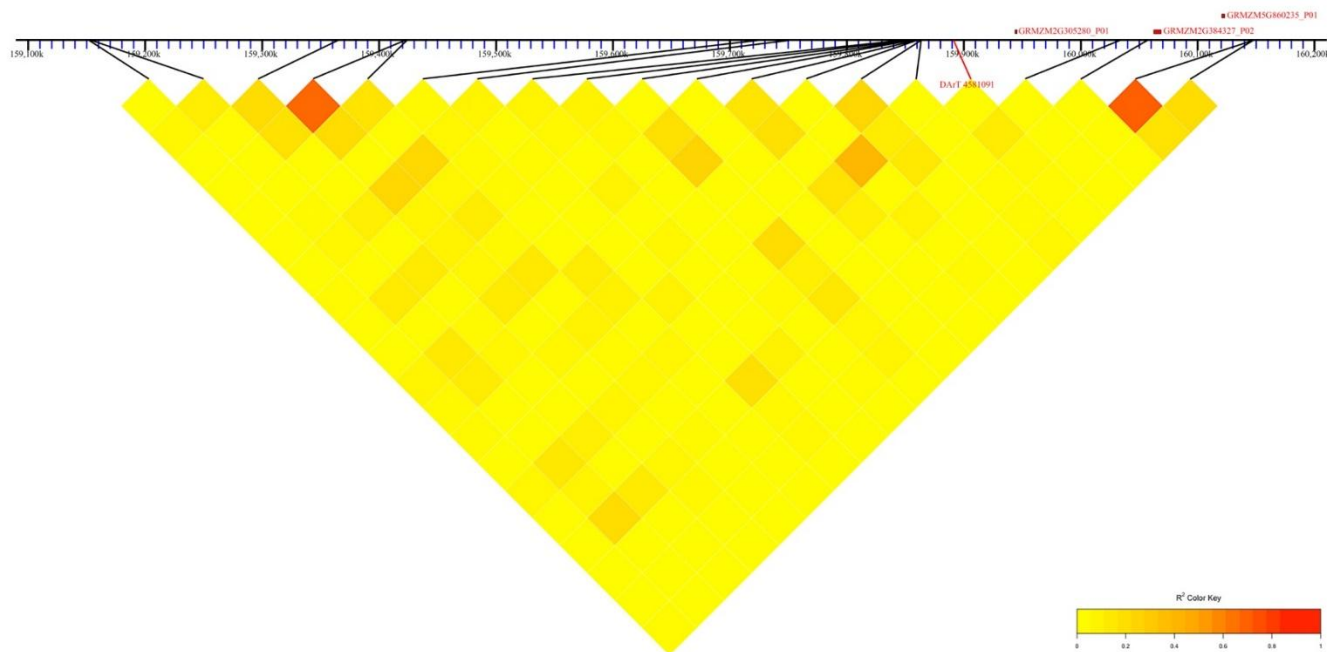


Figura 17 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1,1 Mbp do cromossomo 7 para o marcador 4581091 (posição 159,891,338 bp), que foi significativamente associado na análise de Lavras pelo método de modelo misto.

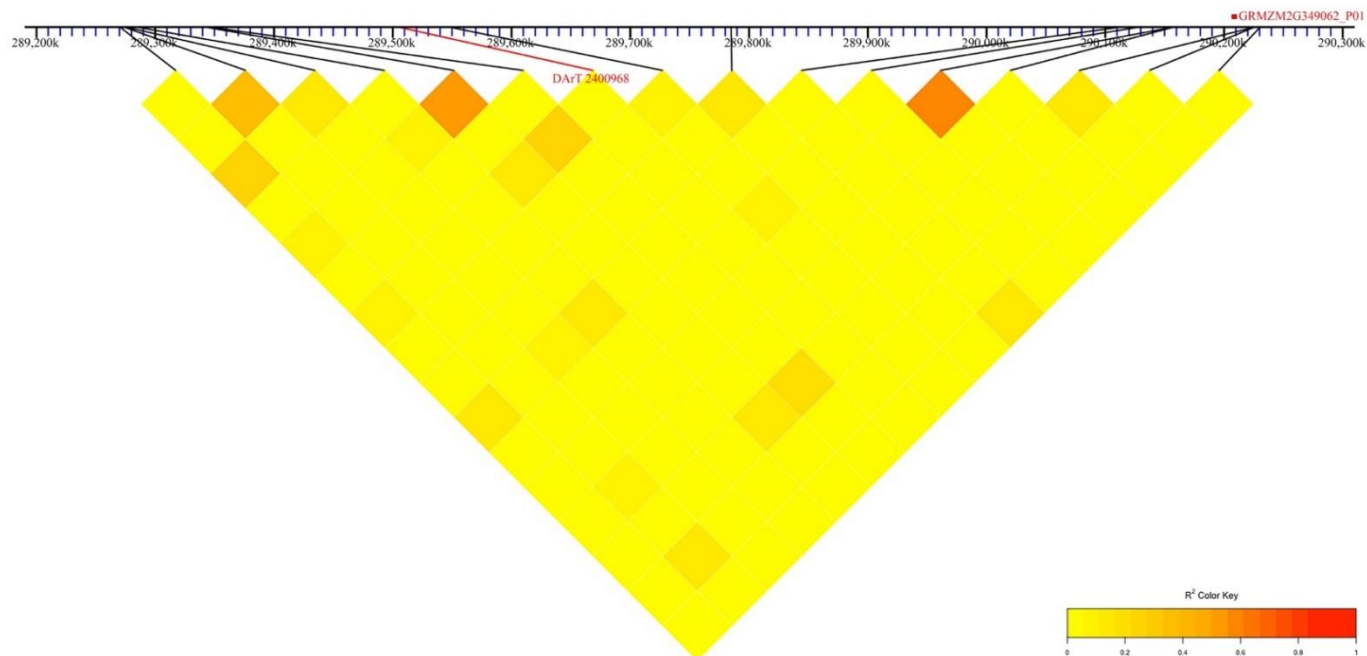


Figura 18 – Gráfico do tipo *heatmap* para o desequilíbrio de ligação entre os marcadores em uma região genômica de 1,1 Mbp do cromossomo 1 para o marcador 2400968 (posição 289,504,622 bp), que foi significativamente associado na análise de Uberlândia pelo método de modelo misto.

Tabela 8 – Posição e função dos genes que foram localizados dentro da janela genômica em que os marcadores foram significativamente associados com a resistência (Continua).

| Marcador | Posição | Cromossomo | Gene | Função | Localização (bp) |
|----------|-------------|------------|-------------------|--|---------------------------|
| 4582295 | 78,131,634 | 1 | GRMZM2G038991_P01 | G-protein coupled receptor signaling pathway | 78,297,267 – 78,298,351 |
| | | | GRMZM2G050984_P01 | Oxidation-reduction process | 78,709,345 – 78,715,745 |
| 4577966 | 136,272,488 | 1 | - | - | - |
| 4767436 | 53,415,853 | 1 | GRMZM2G407825_P01 | Metabolic process | 53,148,300 – 53,153,278 |
| | | | GRMZM2G102365_P02 | Cell redox homeostasis | 53,267,616 – 53,270,157 |
| 2400968 | 289,504,622 | 1 | GRMZM2G349062_P01 | Oxidation-reduction process | 290,206,796 – 290,209,672 |
| 4772929 | 36,878,586 | 4 | GRMZM2G131473_P02 | Proteolysis | 36,441,320 – 36,445,344 |
| | | | GRMZM2G050118_P01 | Heat shock protein binding | 36,757,143 – 36,761,457 |
| | | | GRMZM2G170232_P01 | Heat shock protein binding | 36,959,864 – 36,961,776 |
| 4592572 | 55,225,453 | 5 | GRMZM2G381576_P01 | G-protein coupled receptor signaling pathway | 54,640,245 – 54,642,479 |

Tabela 8 – Posição e função dos genes que foram localizados dentro da janela genômica em que os marcadores foram significativamente associados com a resistência (Continuação).

| Marcador | Posição | Cromossomo | Gene | Função | Localização (bp) |
|----------|-------------|------------|-------------------|---|---------------------------|
| | | | GRMZM2G005526_P02 | Superoxide metabolic process; Oxidation-reduction process | 55,072,828 – 55,077,252 |
| | | | GRMZM2G162098_P01 | Apoptotic process; Defense response | 55,459,467 – 55,463,724 |
| | | | GRMZM2G091672_P01 | Apoptotic process; Defense response | 55,578,326 – 55,582,323 |
| | | | GRMZM2G091696_P01 | Apoptotic process; Defense response | 55,590,775 – 55,599,910 |
| | | | GRMZM2G105428_P01 | Apoptotic process; Defense response | 55,809,843 – 55,813,411 |
| | | | GRMZM2G123018_P01 | Cell wall modification | 120,453,188 – 120,454,951 |
| 4592049 | 120,994,839 | 7 | GRMZM2G488101_P01 | Proteolysis | 120,727,788 – 120,730,433 |
| | | | GRMZM2G147966_P01 | Oxidation-reduction process | 121,043,853 – 121,048,231 |
| | | | GRMZM2G169201_P01 | Oxidation-reduction process | 121,052,889 – 121,056,040 |
| 4579702 | 158,219,730 | 7 | GRMZM2G394261_P01 | Innate immune response | 158,297,463 – 158,301,857 |

Tabela 8 – Posição e função dos genes que foram localizados dentro da janela genômica em que os marcadores foram significativamente associados com a resistência (Conclusão).

| Marcador | Posição | Cromossomo | Gene | Função | Localização (bp) |
|----------|-------------|------------|-------------------|--|---------------------------|
| 2518496 | 124,220,838 | 7 | GRMZM2G007249_P02 | Oxidation-reduction process | 124,313,854 – 124,315,851 |
| | | | GRMZM2G389173_P01 | Cell redox homeostasis | 124,417,173 – 124,422,099 |
| | | | GRMZM2G092409_P01 | Type I hypersensitivity | 124,438,515 – 124,440,381 |
| | | | GRMZM2G421463_P02 | Cell wall macromolecule catabolic process | 124,599,410 – 124,605,593 |
| | | | GRMZM2G024865_P01 | Proteolysis | 124,779,165 – 124,781,770 |
| 4581091 | 159,891,338 | 7 | GRMZM2G305280_P01 | Type I hypersensitivity | 159,943,753 – 159,944,792 |
| | | | GRMZM2G384327_P02 | Protein coupled receptor signaling pathway | 160,062,430 – 160,068,019 |
| | | | GRMZM5G860235_P01 | Type I hypersensitivity | 160,121,913 – 160,123,965 |
| 4772985 | 136,137,749 | 8 | - | - | - |