Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”

Identificação de QTLs em soja associados à resistência ao nematoide-das-lesões-
radiculares

José Maurício Terasawa

Tese apresentada para obtenção do título de Doutor em
Ciências. Área de concentração: Genética e
Melhoramento de Plantas

Piracicaba
2017

José Maurício Terasawa
Engenheiro Agrônomo

Identificação de QTLs em soja associados à resistência ao nematoide-das-lesões-radiculares

Orientador:
Prof. Dr. **JOSÉ BALDIN PINHEIRO**

Tese apresentada para obtenção do título de Doutor em
Ciências. Área de concentração: Genética e
Melhoramento de Plantas

Piracicaba
2017

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Terasawa, José Maurício

Identificação de QTLs em soja associados à resistência ao nematoide-das-lesões-radiculares/ José Maurício Terasawa. - - Piracicaba, 2017.

102p. Tese (Doutorado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. *Glycine max* 2. Fitonematoides 3. *Pratylenchus brachyurus* 4. Mapeamento de QTLs 5. Mapeamento multivariado 6. Marcador molecular 7. Melhoramento de plantas I. Título

DEDICATÓRIA

À minha amada esposa **Milene**,
pelo amor, companheirismo,
compreensão e incentivo,
dedico.

Aos meus pais, **Ana Cirte e Francisco**,
pelo amor incondicional,
pelo apoio e incentivo,
ofereço.

AGRADECIMENTOS

A Deus...

por permitir que eu superasse todos os obstáculos, por me dar forças para nunca desistir e por iluminar meu caminho para mais essa conquista;

À Família...

À minha querida e amada esposa, Milene Möller Terasawa, que sempre me incentivou e me apoiou em todos os momentos; Muito obrigado por todo amor, carinho, afeto e pelo companheirismo. Agradeço também pelas suas significativas contribuições que foram fundamentais para a finalização dessa tese;

À minha mãe, Ana Cirte Moro Terasawa, exemplo de vida, princípios, carinho, amor e dedicação infinita à família; Muito obrigado por todo o incentivo, apoio e suporte para concluir este sonho;

Ao meu pai, Francisco Terasawa, exemplo de pesquisador, perseverança e dedicação ao trabalho; sempre ao meu lado, me apoiando em todas as decisões; Muito obrigado por todo incentivo e suporte para chegar onde estou;

À minha amada irmã Daniela e meu cunhado Juliano por todo o carinho e apoio e por ter me presenteado com minha afilhada Gabriela, uma princesa que ilumina e alegria a minha vida; Muito obrigado pelo carinho e incentivo;

Ao meu irmão Francisco Júnior, pelo exemplo de dedicação ao trabalho e idoneidade; minha cunhada Lygia, minhas sobrinhas Luiza e Giulia (*in memorian*); Obrigado pelo carinho e incentivo;

À minha segunda família, que Deus me presenteou, minha sogra Márcia, meu sogro Gilson Möller, minha cunhada Bianca, meu cunhado Rodrigo e as minhas queridas e amadas sobrinhas Clara e Giovanna; Obrigado pelo carinho e incentivo;

Ao meu avô Humberto Moro, (*in memorian*), à minha avó Thereza Pilatti Moro (*in memorian*), ao meu avô Miguel Wataru Terasawa (*in memorian*) e à minha avó Carolina Terasawa (*in memorian*), que estarão sempre em minha memória e em meu coração;

Aos Mestres...

Ao meu orientador, Prof. Dr. José Baldin Pinheiro, pela oportunidade, ensinamentos, confiança e amizade;

A todos os membros da banca examinadora e suplentes, por disponibilizarem seus conhecimentos para essa defesa;

Agradeço também...

Ao Dr. João Ricardo Bachega Feijó Rosa, pela disponibilidade em me ajudar e pelas valiosas contribuições nas análises e conclusão da tese;

Aos colegas de trabalho do Laboratório de Diversidade Genética e Melhoramento, que, de alguma forma, contribuíram para alguma etapa deste trabalho;

À Escola Superior de Agricultura “Luiz de Queiroz” pela excelente qualidade de ensino e, em especial, ao Programa de Pós-Graduação em Genética e Melhoramento de Plantas, pela oportunidade em realizar o doutorado;

A todos os professores do Departamento de Genética da ESALQ - USP, pelos ensinamentos e conhecimentos compartilhados;

A todos os funcionários do Departamento de Genética, em especial à Léia, pela competência e compromisso com o trabalho; E aos funcionários da biblioteca central, em especial à Silvia e à Eliana, pela ajuda nessa etapa final;

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro concedido;

À FT Sementes, pelo apoio financeiro e logístico concedido para a realização do projeto; pela oportunidade de trabalho e crescimento profissional concedidos, e aos colaboradores da empresa, pelo apoio nessa etapa final;

A todos que de alguma forma contribuíram para o êxito deste trabalho, mesmo que aqui não mencionados, meus sinceros agradecimentos.

EPIGRAFE

“O Sucesso não tem a ver com o lugar de onde você veio, e sim com a confiança que você tem e o esforço que você está disposto a investir”.

Michele Obama

“Os que se encantam com a prática sem a ciência são como os timoneiros que entram no navio sem timão e nem bússola, nunca tendo certeza do seu destino”.

Leonardo da Vinci

SUMÁRIO

RESUMO	9
ABSTRACT	11
1. INTRODUÇÃO	13
2. REVISÃO BIBLIOGRÁFICA	15
2.1. SOJA	15
2.2. NEMATOIDE-DAS-LESÕES-RADICULARES.....	16
2.3. ESTRATÉGIAS DE MANEJO E CONTROLE DE <i>P. BRACHYURUS</i>	19
2.4. MECANISMOS DE RESISTÊNCIA DAS PLANTAS AOS NEMATOIDES.....	21
2.5. USO DO RNAI NO CONTROLE DOS NEMATOIDES	23
2.6. MARCADORES MOLECULARES NO MELHORAMENTO GENÉTICO	23
2.7. MAPEAMENTO DE QTLs	25
2.8. QTLs ASSOCIADOS À RESISTÊNCIA AOS NEMATOIDES	28
3. MATERIAL E MÉTODOS	31
3.1. CRUZAMENTO DOS GENITORES	31
3.2. CONDUÇÃO DA GERAÇÃO F ₁	32
3.3. OBTENÇÃO DA POPULAÇÃO F ₂	32
3.4. AVALIAÇÕES FENOTÍPICAS DA POPULAÇÃO F ₂	32
3.5. ANÁLISES FENOTÍPICAS DA POPULAÇÃO F ₂	34
3.6. EXTRAÇÃO DO DNA.....	35
3.7. DADOS GENOTÍPICOS DA POPULAÇÃO F ₂	35
3.8. CONSTRUÇÃO DO MAPA GENÉTICO	36
3.9. MAPEAMENTO DE QTLs	37
3.10. ANÁLISE DE SEQUÊNCIAS CANDIDATAS EM REGIÕES DE INTERESSE	39
4. RESULTADOS E DISCUSSÃO.....	41
4.1. ANÁLISES FENOTÍPICAS	41
4.2. MARCADORES SNP	45

4.3. MAPA GENÉTICO	46
4.1. MAPEAMENTO DE QTLs	61
4.2. ANÁLISE DAS SEQUÊNCIAS CANDIDATAS EM REGIÕES DE INTERESSE	72
5. CONSIDERAÇÕES FINAIS	77
REFERÊNCIAS.....	81
APÊNDICES	94

RESUMO

Identificação de QTLs em soja associados à resistência ao nematoide-das-lesões-radiculares

Soja (*Glycine max* (L.) Merrill) é uma das mais importantes oleaginosas no mundo e sua relevância pode ser também mensurada pela extensão da sua produção no Brasil, onde representa quase a metade do total da área plantada com grãos. A importância da soja está também associada à multiplicidade de utilização do grão, como por exemplo, na alimentação humana e animal, na indústria química e na geração de energia como biodiesel. A produtividade da soja tem sido frequentemente impactada devido ao ataque de pragas e doenças. Dentre as espécies de fitonematoides, o nematoide-das-lesões-radiculares (*Pratylenchus brachyurus* Godfrey), tem gerado significativas perdas econômicas aos produtores, variando entre 30 a 50%, dependendo da infestação da cultura. Este trabalho teve como objetivo identificar QTLs (*Quantitative Trait Loci*), a partir de um conjunto de fenótipos associados à resistência ao nematoide-das-lesões-radiculares utilizando a abordagem de mapeamento multivariado. Uma população de 174 indivíduos F₂, obtidos a partir do cruzamento entre duas linhagens de soja FTPG06A e FTPG12X (com baixo fator de reprodução do nematoide), foi utilizada para a obtenção dos valores genéticos preditos (BLUPs - *Best Linear Unbiased Predictions*) de cinco características estudadas e também para a genotipagem com o *SoySNP6k Bead Chip*. As características avaliadas foram: fator de reprodução (FR), peso fresco (PA) e comprimento (CA) da parte aérea, peso fresco (PR) e comprimento (CR) da raiz. Um total de 1.240 marcadores SNP (*Single Nucleotide Polymorphism*) foram mapeados nos 20 grupos de ligação (GLs) da soja. O comprimento total do mapa foi de 3.084,46 centiMorgans (cM), com intervalo médio de 2,54 cM entre marcadores adjacentes. A identificação de QTLs (*Quantitative Trait Loci*) para os caracteres fenotípicos foi realizada utilizando-se o mapeamento de intervalos múltiplos univariado (MIM) e multivariado (MT-MIM), com estimativa dos efeitos principais dos QTLs e análises de epistasia envolvendo pares de QTLs. Na abordagem MIM foram identificados três QTLs associados à variável CR, nos GLs (Grupos de Ligação) A2 e E (2 QTLs), explicando um total de 34,22% da variação fenotípica dessa variável para a população em estudo. Na abordagem MT-MIM, foram selecionados dois conjuntos de variáveis, de acordo com a correlação entre as mesmas. Três regiões genômicas foram reveladas, sendo estatisticamente significativas para as variáveis CR, FR e CA. A comparação dos seis QTLs identificados no presente estudo com o banco de dados de QTLs do Soybase forneceu evidências de que cinco QTLs, ainda não foram descritos na literatura. Uma busca por sequências candidatas em duas regiões de interesse, associadas à variável FR, foi realizada, com base na plataforma de dados genômicos *PlantGDB*. Várias sequências candidatas indicam relação com mecanismos importantes na resposta das plantas a estresses bióticos. Dessa forma, os resultados obtidos no presente estudo forneceram informações para auxiliar na melhor compreensão da arquitetura genética dos caracteres quantitativos analisados.

Palavras-chave: *Glycine max*; fitonematoides; *Pratylenchus brachyurus*; Mapeamento de QTLs; Mapeamento multivariado; Marcador molecular; Melhoramento de plantas

ABSTRACT

QTL identification in soybean related to root lesion nematode

Soybean (*Glycine max* (L.) Merrill) is one of the most important oil crop worldwide and its relevance can also be measured by the extension of soybean grain production in Brazil, where it represents almost half of the total planting area with grains crops. Importance of soybean is also related to a multiplicity of usage of the grain, for example, in human consumption, animal feed, chemical industry and for energy generation as biofuel. Soybean yield has been frequently reduced by occurrence of pest and diseases. Among phytonematodes species, root lesion nematode (*Pratylenchus brachyurus* Godfrey) has caused significant economic losses to farmers, ranging from 30 to 50%, depending on crop infestation. This research aimed to identify QTLs (Quantitative Trait Loci), from a set of phenotypes associated with resistance to root lesion nematode, using the multivariate multiple interval mapping. A mapping population of 174 F₂ plants derived from a by-parental cross between two soybean breeding lines FTPG06A and FTPG12X (with low nematode reproduction factor), was used for prediction of genetic values (BLUPs - Best Linear Unbiased Predictions) for five traits studied and also for genotyping with SoySNP6k Bead Chip. Traits evaluated were reproduction factor (FR), shoot weight (PA), shoot length (CA), root weight (PR) and root length (CR). A total of 1,240 SNP (Single Nucleotide Polymorphism) markers were mapped into 20 soybean linkage groups (LG). A total map length was 3,084.46 centiMorgans (cM) with an average of 2.54 cM between flanking markers. QTL mapping for those traits was performed using univariate (MIM) and multivariate (MT-MIM) multiple-interval mapping, with main QTL effects estimates and epistasis analysis between QTL pairs. MIM analysis identified three QTLs associated to CR trait at LG A2 and LG E (2 QTL), explaining 34,22% of phenotypic variation estimated for this mapping population. For MT-MIM analysis, two sets of traits were selected, according to the correlation among them. Three genomic regions statistically significant for CR, FR and CA traits were identified. Comparison between six identified QTL and QTL database at Soybase provided evidence that five QTL have not been published yet. Search for candidate sequences located in two regions associated with FR trait were further performed on the *PlantGDB* genomic data platform. Several candidate sequences indicate relationships with important plant response mechanisms to biotic stresses. Thus, results obtained in the present study provided information to improve knowledge of the genetic architecture of the analyzed quantitative traits.

Keywords: *Glycine max*; Phytonematodes, *Pratylenchus brachyurus*; QTL mapping; Multivariate analysis, Molecular marker, Plant breeding

1. INTRODUÇÃO

A soja é uma das oleaginosas de maior importância para o agronegócio mundial, sendo o Brasil um importante produtor e exportador do grão e seus derivados (CONAB 2017). O consumo de soja tem crescido anualmente devido a busca incessante e crescente por um meio de vida mais saudável. Esta mudança de hábito alimentar associada ao incremento da população mundial, tem impulsionado o consumo de soja, tanto para alimentação humana quanto animal.

No Brasil, o plantio de soja vem crescendo ano após ano desde o início da década de 1970. O avanço do cultivo da soja rumo as regiões centro-oeste, norte e nordeste coincidiu com o surgimento de novos desafios em relação à adaptabilidade das plantas às condições ecológicas e edafoclimáticas das baixas latitudes. Além disso, esta nova realidade, expôs os cultivos à pragas e doenças que anteriormente não afetavam a cultura da soja.

A ocorrência natural de fitonematoides, nos solos tropicais muito intemperizados, com destaque para o nematoide-das-lesões-radiculares *Pratylenchus brachyurus* Godfrey, tem gerado significativas perdas econômicas aos produtores (Asmus 2004; Sediyaama 2015). Na cultura da soja essa espécie é responsável por elevado dano, juntamente com os nematoides formadores de galha *Meloidogyne javanica* (Treub) Chitwood, e *M. incognita* (Kofoid; White) Chitwood, o nematoide de cisto da soja *Heterodera glycines* Ichinohe e o reniforme *Rotylenchus reniformis* (Linford; Oliveira) (Machado et al. 2006; Sediyaama 2015). *P. brachyurus* desenvolve-se com extrema facilidade em solos de textura arenosa por ser um organismo endoparasita migrador o que dificulta ainda mais o seu controle. A redução da produtividade de lavouras de soja devido à presença de *P. brachyurus* pode variar entre 30 a 50%, dependendo da infestação da lavoura (Schmitt and Barker 1981; Goulart 2008; Dias et al. 2010; Ribeiro et al. 2010; Debiasi et al. 2011). Estudos relatam redução média de 21% na produtividade em lavouras infestadas com 82 nematoides por grama de raiz (Antonio et al. 2012).

Diversas estratégias de manejo e controle de fitonematoides têm sido utilizadas. Contudo, as mesmas não têm apresentado eficiência no controle ou na erradicação do parasita. Dessa forma, uma estratégia de extrema relevância e caráter mais viável e duradouro para o controle dos fitonematoides consiste no desenvolvimento de cultivares de soja resistentes. Esta estratégia apresenta benefícios econômicos e ambientais, devido à redução do custo de produção e menor agressão ao meio ambiente através da menor utilização de nematicidas.

Tendo em vista a importância econômica da soja no Brasil e a patogenicidade do *Pratylenchus brachyurus*, o controle dessa praga é estratégico para elevar os patamares de produtividade da cultura no país. Até o momento não há relatos de identificação de QTLs para a

espécie *P. brachyurus* na cultura da soja. Dessa forma, o mapeamento de QTLs é uma técnica importante utilizada em programas de melhoramento genético para auxiliar melhoristas na identificação e seleção eficientes de plantas contendo os alelos desejáveis.

Assim, o presente trabalho teve como objetivo identificar QTLs em uma população F_2 , de soja, associados à resistência ao nematoide-das-lesões-radiculares.

2. REVISÃO BIBLIOGRÁFICA

2.1. Soja

A soja é uma planta autógama, pertencente à família Fabaceae (Leguminosae), gênero *Glycine*, espécie *Glycine max* (L.) Merrill ($2n=40$), sendo esta sua forma cultivada (NCBI 2017). Essa leguminosa é reconhecida como uma das mais antigas plantas cultivadas no planeta, sendo originária da Ásia onde foi domesticada e melhorada por cientistas chineses (Embrapa 2013). Esta cultura milenar constitui-se em uma das principais oleaginosas do mundo ((Paiva et al. 2006). No Brasil, foi introduzida na Bahia em 1882 e posteriormente no estado de São Paulo, onde apresentou melhor adaptação. No entanto, somente no final da década de 60 tornou-se economicamente importante quando a produção passou de 203 mil toneladas para 1 milhão de toneladas (Câmara 1998).

O melhoramento genético da soja permitiu que o cultivo se expandisse, do sul para praticamente todo o território nacional (Terasawa et al. 2017), tornando-se o produto agrícola de maior importância para a economia brasileira (Sediyama 2009, 2015; Silva et al. 2017), representando 49% da área plantada com grãos do país (CONAB 2017). Na safra 2016/17 foram produzidas cerca de 315 milhões de toneladas em 120,958 milhões de hectares da cultura no mundo (USDA 2017). Nesse aspecto, o Brasil aparece como o segundo maior produtor mundial, com produção total de mais de 113,9 milhões de toneladas na safra 2016/17, em uma área de quase 33,9 milhões de hectares (CONAB 2017). No Brasil, o complexo agroindustrial da soja que engloba a produção de grãos, farelo e óleo, tem grande peso na balança comercial, contribuindo com 14% das exportações totais do país, ou mais de US\$ 25,4 bilhões (SECEX 2017).

Na safra 2016/2017, o estado brasileiro com maior área de cultivo de soja foi o Mato Grosso, com 9,3 milhões de hectares, correspondendo a 28% da produção total do país, seguido do Rio Grande do Sul com 5,57 milhões de hectares e 17,6% da produção nacional e Paraná, 5,25 milhões de hectares, e 17%, respectivamente (CONAB 2017). A região que tem apresentado maior potencial de expansão da área cultivada com soja nos últimos anos é conhecida como MATOPIBA, representada pelos Estados do Maranhão, Tocantins, Piauí e Bahia.

A diversidade na utilização dos grãos de soja e seus subprodutos, deve-se ao seu uso na forma direta ou indireta na alimentação humana, na agroindústria, na indústria química, e têxtil e na geração de energia através do biodiesel (Miragaya 2005; Missão 2006). A composição do grão de soja é de aproximadamente 38% de proteínas, 19% de óleos e 27% de carboidratos (Dias et al. 2010).

A soja pode ser utilizada na alimentação humana na forma de grãos inteiros imaturos ou não, molho, bebida, e seus derivados. A lecitina obtida pela extração da goma do óleo de soja bruto é empregada como emulsificante na elaboração de produtos ricos em gorduras e óleos, como chocolate, margarina e produtos de panificação (Missão 2006).

O óleo de soja é rico em ácidos graxos poli-insaturados podendo ser usado como óleo de cozinha e para produção de maionese e margarinas. O óleo também apresenta aplicações industriais na fabricação de tinta de caneta, formulador de espumas, tinta de pintura em geral, xampu, sabões e detergentes (Dias et al. 2010; Ribeiro et al. 2011; Pípolo et al. 2015).

O farelo de soja é usado como suplemento alimentar rico em proteína para criação de gado, suínos, aves domésticas e peixes na aquicultura. A casca da soja é retirada durante o descascamento inicial dos grãos e contém material fibroso. É usada como forragem grossa e também como fonte de fibras dietéticas de cereais matinais e de certos lanches prontos (Missão 2006). A indústria têxtil tem feito uso da soja como alternativa na produção de fibras têxteis biodegradáveis (Alves et al. 2006).

Outra exploração recente da soja tem sido a utilização de seu óleo na geração de energia, para a produção de biodiesel. A utilização deste combustível resulta em impactos socioambientais favoráveis como a redução das emissões de CO₂ e de compostos poluentes como o óxido de enxofre e outros gases que contribuem para o efeito estufa (Costa Neto et al. 2000; Sentelhas et al. 2015).

A produtividade da soja tem sido frequentemente impactada devido ao ataque de pragas e doenças. O monocultivo em larga escala tem contribuído para o aumento da vulnerabilidade da cultura (Moura 1999; Boerma and Walker 2005). Pragas presentes em quase todas as regiões produtoras do país, como lagartas desfolhadoras, percevejos sugadores das vagens e fitonematoides, constituem os principais grupos de pragas que comumente ocasionam prejuízos econômicos às lavouras (Embrapa 2013; Abreu et al. 2014, 2016). Os danos causados por fitonematoides e as perdas que lhes são atribuídas têm sido causa de crescente preocupação entre produtores (Asmus 2004; Machado et al. 2006; Sediya 2015).

2.2. Nematoide-das-lesões-radiculares

O nematoide-das-lesões-radiculares pertencente à família *Pratylenchidae*, gênero *Pratylenchus* (NCBI 2017). Esse gênero é considerado mundialmente como o terceiro grupo de fitonematoides de importância econômica, sendo superado apenas pelo nematoide-de-galhas, do gênero *Meloidogyne* e nematoide de cisto *Heterodera* spp. (Asmus 2004; Castillo and Vovlas 2007; Inomoto and Silva 2011; Jones et al. 2013; Mainardi and Asmus 2015). Além de uma ampla

adaptação às diferentes condições ambientais, os danos ocasionados por *Pratylenchus* ssp. englobam espécies de importância econômica, incluindo cereais, legumes, forrageiras, hortaliças, vegetais, frutíferas e ornamentais (Smiley et al. 2005; Goulart 2008; Mainardi and Asmus 2015).

Mais de setenta espécies do gênero *Pratylenchus* já foram descritas mundialmente. No Brasil, as espécies de maior ocorrência são *P. brachyurus* (Godfrey, 1929) Filipjev & Schuurmans Stekhoven 1941, *P. coffeae* (Zimmermann, 1898) Filipjev & Schuurmans Stekhoven 1941, *P. jaehni* (Inserra et al. 2001); *P. penetrans* (Coob, 1917) Filipjev & Schuurmans Stekhoven, 1941, *P. vulnus* (Allen & Jensen, 1951) e *P. zeae* (Graham, 1951) (Monteiro et al. 2003; Gonzaga and Santos 2012).

A espécie *P. brachyurus* apresenta ampla distribuição mundial, provavelmente devido ao seu alto grau de polifagia e ação patogênica pronunciada em diferentes espécies anuais e perenes (Schmitt and Barker 1981; Ferraz 2006; Machado et al. 2006; Mainardi and Asmus 2015). No Brasil, essa espécie já foi observada em espécies como soja, batata, algodão, amendoim, milho, aveia, milheto, girassol, cana-de-açúcar, guandu e a maioria das plantas daninhas (Lordello 1988; Dias et al. 2010). Na cultura da soja essa espécie é responsável por dano significativo, juntamente com os nematoides formadores de galha *Meloidogyne javanica* (Treub) Chitwood, e *M. incognita* (Kofoid; White) Chitwood, o nematoide de cisto da soja *Heterodera glycines* Ichinohe e o reniforme *Rotylenchus reniformis* (Linford; Oliveira) (Machado et al. 2006; Sediyaama 2015).

A estimativa de perda anual no agronegócio brasileiro de soja, devido aos danos ocasionados por nematoides, é de aproximadamente R\$16,2 bilhões (Agrolink 2015). De acordo com estudos realizados no Brasil e nos Estados Unidos, a redução da produtividade de lavouras de soja devido à presença de *P. brachyurus* pode variar entre 30 a 50%, dependendo da infestação da lavoura (Schmitt and Barker 1981; Goulart 2008; Dias et al. 2010; Ribeiro et al. 2010; Debiasi et al. 2011). Estudos relatam redução média de 21% na produtividade em lavouras infestadas com 82 nematoides por grama de raiz (Antonio et al. 2012).

O ciclo de vida do *P. brachyurus* geralmente tem duração de três a seis semanas, sendo diretamente influenciado pela temperatura, umidade, textura do solo e planta hospedeira (Castillo and Vovlas 2007). O desenvolvimento desse nematoide compreende seis estádios: ovo, quatro estádios juvenis (J1 a J4) e a forma adulta. Cada fêmea deposita cerca de 70 a 80 ovos ao longo da vida, normalmente no interior dos tecidos das raízes ou no solo, próximo à superfície. A postura é isolada, sem a formação de massas de ovos e a reprodução geralmente ocorre por partenogênese, sendo raro observar machos nessa espécie (Mainardi and Asmus 2015). Os nematoides juvenis e adultos penetram nas raízes entre ou através das células do córtex, alimentando-se do conteúdo celular enquanto migram pelos tecidos. A penetração e migração são facilitadas pela combinação de um forte estilete e a liberação de substâncias enzimáticas que degradam as paredes celulares das

plantas (Goulart 2008). A movimentação desse patógeno também desorganiza o parênquima cortical, devido à destruição de numerosas células. Esses nematoides somente migram para o solo, em busca de outras plantas, quando as condições ambientais se tornam desfavoráveis (Dias et al. 2010).

Os sintomas mais observados pela ação desse endoparasita migratório obrigatório são relacionados à fisiologia e ao crescimento da planta, ocasionando reduções significativas da produtividade da cultura (Ferraz 1999). Além disso, a necrose e morte celular dos tecidos da raiz facilitam a entrada de outros patógenos do solo, como fungos e bactérias (Goulart 2008). A interação desses agentes resulta na redução das radículas e, em muitos casos, na perda da raiz pivotante. As interações mais frequentemente observadas são com fungos dos gêneros *Fusarium* e *Verticillium*, resultando em maiores danos à planta (Back et al. 2002; Castillo and Vovlas 2007). A parte aérea de plantas afetadas pode apresentar clorose ou murchamento durante a estação seca, e a desfolha total pode ocorrer quando o ataque é severo (Ferraz 2006). Importantes processos das plantas, como respiração, fotossíntese, absorção e transporte de nutrientes, relação hídrica e equilíbrio hormonal podem ser direta ou indiretamente afetados (Hussey and Williamson 1998). Ao contrário do nematoide de galhas que ocasiona a formação de tumores nas raízes, os sintomas observados na parte aérea, ocasionados por *P. brachyurus*, podem ser confundidos com os sintomas causados por outros patógenos, pela toxicidade de elementos do solo, deficiências nutricionais ou estresse hídrico, o que dificulta ainda mais a sua identificação (Goulart 2008). As plantas de soja infectadas por *P. brachyurus* apresentam, em geral, estatura reduzida, caule fino, má formação dos grãos, entrenós curtos e folhas de tamanho reduzido. O sistema radicular apresenta coloração marrom escura e a raiz principal apodrece, apresentando maior número de raízes secundárias (Ferraz 2006).

A disseminação de *P. brachyurus* pode ocorrer no campo pelo trânsito de implementos agrícolas contaminados, trabalhadores e animais. O escoamento de água em áreas de declive e água de irrigação também podem contribuir para a dispersão desse patógeno. A auto-disseminação é rara para o gênero, pois esses nematoides se locomovem somente em pequenas distâncias, formando reboleiras que aumentam de ano para ano. Nesses locais, pode-se observar gradientes crescentes de atrofiamento e clorose das plantas do centro para as bordas da reboleira (Dias et al. 2007).

Para se estabelecer um sistema eficiente de manejo é fundamental identificar as espécies de nematoides presentes na lavoura, bem como suas densidades populacionais. Para uma adequada amostragem do campo deve-se coletar entre 20 a 25 amostras a cada 10 hectares (ha) (Abreu and Agnes 2015). Além do solo, deve-se coletar a raiz, que é o local de maior concentração de *P.*

brachyurus. As plantas amostradas devem apresentar aspecto de sanidade, preferencialmente nas bordas das reboleiras onde geralmente há uma maior concentração de nematoides. Como não existem métodos eficientes de controle químico para os nematoides é importante a identificação dos gêneros presentes no campo para a adoção de medidas de manejo adequadas para a reduzir a população abaixo dos níveis de danos.

Apesar da grande importância econômica dos nematoides *Pratylenchus* ssp., existem poucos estudos envolvendo esse gênero. Alguns estudos moleculares envolvendo esse gênero referem-se à determinação de espécies e estudos filogenéticos envolvendo o sequenciamento do DNA ribossômico (Al-Banna et al. 1997, 2004; Bert et al. 2008; Subbotin et al. 2008; Holterman et al. 2009). Para as espécies *P. zeae*, *P. thornei*, *P. coffeae* e *P. vulnus* foram desenvolvidos *primers* espécie-específicos, com base nas sequências de DNA dos ribossomos, permitindo a identificação das mesmas (Machado et al. 2005; Berry et al. 2008; Yan et al. 2012).

Uma forma de verificar a resistência de plantas aos nematoides consiste na avaliação da capacidade ou taxa de reprodução dos parasitas, denominado fator de reprodução (FR) (Oostenbrink 1966). Para a determinação do FR geralmente são realizados trabalhos experimentais, em condições controladas (casa-de-vegetação, inoculação artificial, irrigação e temperatura controlada, etc), calculando-se o quociente entre a população final do nematoide (Pf) e a população inicial (Pi). Valores menores do que 1 ($FR < 1$) indicam que a cultivar e/ou espécie apresenta efeito redutor na população de nematoide presente no solo, sendo considerada resistente. No caso de $FR > 1$, ocorre efeito multiplicador, portanto, deve-se verificar a presença de danos, indicando suscetibilidade ou ausência de danos indicando tolerância (Canto-Saenz 1985).

Geralmente, as lavouras afetadas apresentam alto índice populacional desse nematoide. Portanto, o uso de cultivares de soja com menor FR, em rotação com uma espécie vegetal não hospedeira, é uma estratégia para reduzir ou até mesmo eliminar esses patógenos (Dias et al. 2010).

2.3. Estratégias de manejo e controle de *P. brachyurus*

O aumento da incidência desse nematoide nas áreas agrícolas, principalmente da região Centro-Oeste do Brasil, tem preocupado os agricultores (Ribeiro et al. 2010, 2011). A expansão das fronteiras agrícolas para áreas com textura arenosa e média, o cultivo contínuo de espécies hospedeiras e a adoção do sistema de plantio direto são fatores que contribuem para o desenvolvimento desse patógeno (Ribeiro et al. 2009; Embrapa 2013).

Um levantamento realizado na safra 2008/2009, nas principais regiões produtoras de soja do Mato Grosso, demonstrou a presença de *P. brachyurus* em 96% das 411 amostras coletadas (Ribeiro et al. 2010). Abreu et al. (2014) identificou a presença desse nematoide em 94% das

amostras coletadas na região dos Chapadões. Resultado semelhante foi verificado na safra 2013/2014, em amostras provenientes do Mato Grosso do Sul, Mato Grosso, Goiás e Pará, sendo que 97% apresentaram a presença dessa espécie de nematoide (Agnes et al. 2014). Um levantamento realizado na safra 2015/2016 demonstrou a presença desse nematoide em 86% das 322 amostras, provenientes do Mato Grosso, Mato Grosso do Sul e Goiás (Abreu et al. 2016).

O manejo de nematoides deve ser planejado para integrar diferentes métodos de controle, como: aceleração do processo de degradação das raízes de soja ou outras plantas hospedeiras, com um cuidado especial com tigueras e daninhas, flexibilização do plantio direto, cuidado com maquinários, trânsito de pessoas em áreas infestadas, aplicação de nematicidas, rotação de culturas com o uso de plantas não hospedeiras e correção do solo (Ribeiro et al. 2010; Debiasi et al. 2011; Inomoto and Silva 2011; Antonio et al. 2012; Grigolli and Asmus 2014; Mainardi and Asmus 2015; Motter et al. 2015). Segundo Ribeiro et al. (2011), equipamentos agrícolas podem transportar até 1.200 nematoides por 200 cm³ de solo aderido. Dessa forma, áreas infestadas devem ser sempre as últimas a serem trabalhadas e, após as operações agrícolas, os maquinários devem ser devidamente lavados. O fato dessa espécie de nematoide ser polífaga dificulta a escolha de espécies vegetais para a rotação com a soja. No entanto, algumas espécies de adubo verde podem interferir contribuindo para a redução da população dos nematoides (FR<1), como *Crotalaria spectabilis*, *C. breviflora* e *C. ochroleuca* e algumas cultivares de milho (Inomoto and Silva 2011). Apesar de oferecerem uma maior vantagem para o controle desse nematoide, a crotalaria e o milho não são opções que geram renda ao agricultor, sendo geralmente plantadas na entressafra. Portanto, não há tempo suficiente para a redução da população de nematoides, que é capaz de sobreviver no solo, sem a presença do hospedeiro, por até 8 meses (Valentini 2014).

As estratégias adotadas não têm se mostrado totalmente eficientes para a erradicação do parasita. Dessa forma, a utilização de cultivares resistentes ou tolerantes seria a medida mais eficiente e viável para o controle desse nematoide. Não há relatos de cultivares de soja resistentes ao *P. brachyurus* (FR < 1). No entanto, avaliações em casa de vegetação demonstram que algumas cultivares diferem bastante com relação à capacidade de reprodução do nematoide (Ribeiro et al. 2009, 2011) e de tolerância ao mesmo (Dias et al 2008). As cultivares que tem apresentado os menores valores de FR são: FTS TRIUNFO RR, FTS IBYARA RR, BRSGO 204, BRSGO Chapadões e BRS Aurora, BRS Favorita RR, M-SOY 8378RR, M-SOY 8360RR, M-SOY 8585RR, M-SOY 8045RR, M-SOY 8757RR, CD 219RR, TMG 103RR, A7002 (Alves 2008; Rocha et al. 2008). Portanto, o uso dessas cultivares em locais de alta incidência se faz necessário como forma de reduzir a população local desse patógeno. Embora algumas cultivares americanas também tenham sido identificadas com baixo FR (Cook, Davis, Deltapine 417, Essex, McNair 600, NK

Coker 338, Ransom) ou valores intermediários de FR (Hutton, NK Coker 136, Pickett, Asgrow A5979, Bragg), o uso de cultivares adaptadas às condições brasileiras se torna mais interessante em programas de melhoramento genético (Dias et al. 2007).

Embora diversas alternativas de manejo e controle têm sido utilizadas, as mesmas não têm apresentado eficácia para a erradicação do parasita. A contínua expansão da soja e a forma de rotação de cultura utilizada pela maioria dos agricultores reforça a hipótese de que a forma mais eficaz para o controle desses nematoides seria o desenvolvimento de cultivares resistentes ao aumento da incidência desse nematoide nas áreas agrícolas, principalmente da região Centro-Oeste do Brasil, tem preocupado os agricultores (Ribeiro et al. 2010, 2011). A expansão das fronteiras agrícolas para áreas com textura arenosa e média, o cultivo contínuo de espécies hospedeiras e a adoção do sistema de plantio direto são fatores que contribuem para o desenvolvimento desse patógeno (Ribeiro et al. 2009; Embrapa 2013).

2.4. Mecanismos de resistência das plantas aos nematoides

As plantas possuem um complexo mecanismo de defesa capaz de detectar moléculas efetoras específicas dos patógenos (Jones and Dangl 2006; Dodds and Rathjen 2010; Rehman et al. 2016; Gillet et al. 2017). A primeira linha de defesa da planta consiste em receptores extracelulares, capazes de reconhecer os efetores dos patógenos, resultando em modificações da parede celular, produção de espécies reativas de oxigênio ou nitrogênio ou até mesmo metabólitos tóxicos aos patógenos (Melillo et al. 2006; Rehman et al. 2016; Gillet et al. 2017). Algumas espécies de plantas, especialmente as leguminosas, podem sintetizar rapidamente compostos químicos denominados fitoalexinas, como isoflavonóides, poliacetilenos, terpenóides e esteroides, em respostas à penetração dos nematoides (Huang 1985). No entanto, alguns patógenos possuem proteínas efetoras, denominadas *Avr*, capazes de modificar diretamente os receptores das plantas e ocasionar virulência. Como resposta à infecção, algumas espécies de plantas são capazes de ativar uma segunda forma de defesa, na qual uma proteína de resistência específica, denominada proteína R, é capaz de reconhecer a proteína *Avr*. Como resposta à infecção ocorre uma reação de hipersensibilidade, caracterizada pela morte rápida dos tecidos infectados da planta (Jones and Dangl 2006; Williamson and Roberts 2010). O primeiro estudo envolvendo morte celular associado à presença de nematoides foi conduzido por (Veech and Endo 1970), demonstrando que as células gigantes produzidas em cultivares de soja, em resposta à infecção por *M. incógnita*, sofriam colapso e morriam.

O sequenciamento genômico e o estudo de sequências expressas (EST – *Expressed Sequence Tags*) tem permitido a identificação e caracterização de diferentes efetores na defesa das plantas

(Dodds and Rathjen 2010; Thakur and Sohal 2013; Bektas and Eulgem 2015; Kushalappa et al. 2016). Grande parte dos efetores consiste em enzimas capazes de degradar a parede celular vegetal, e provavelmente foram adquiridas por transferência horizontal de bactérias ou fungos (Haegeman et al. 2011). No entanto, efetores com diferentes funções já foram descritos, como no caso de efetores capazes de suprimir ou mesmo interferir no sistema de defesa da planta, ou até induzir o sistema hormonal da planta a produzir um órgão de alimentação para o nematoide, como no caso dos efetores do gênero *Meloidogyne* (Nguyen et al. 2014).

A resistência pode ser descrita como uma coevolução entre hospedeiro e parasita. Nematoides semi-endoparasitas, como *Meloidogyne* spp. Goeldi, *Heterodera* spp. Schmidt e *Rotylenchulus* spp. Linford e Oliveira, são considerados parasitas mais especializados, resultando na presença de raças, patótipos ou biótipos. No caso dos nematoides *Pratylenchus* spp. Filipjev, o melhoramento vegetal para a identificação de espécies resistentes se torna mais difícil pelo fato de ser uma espécie polífaga, pouco especializada e de hábito endoparasita migrador, não se fixando na planta hospedeira. O fato dos nematoides sedentários serem mais especializados resulta em uma maior pressão de seleção e, conseqüentemente, uma maior frequência de genes R, em relação aos nematoides de hábito migrador (Komatsu et al. 2005; Matsuo et al. 2012). Além disso, o modo de reprodução dos nematoides também apresenta grande relevância por influenciar diretamente na evolução da resistência da população (Matsuo et al. 2012). Por exemplo, *H. glycines* apresenta diferentes biótipos pelo fato de possuir reprodução sexuada. Já as espécies de *Meloidogyne* apresentam menor variação por se reproduzirem por partenogênese (reprodução assexuada).

O modo como a resistência a nematoides é herdada define a estratégia a ser adotada para incorporar esta resistência em cultivares comerciais (Boerma and Hussey 1992). Ela pode ser classificada de acordo com número de genes que controlam a característica, sendo (i) monogênica, no caso de um único gene; (ii) oligogênica quando poucos genes controlam a característica; e (iii) poligênica quando múltiplos genes de resistência estão envolvidos. Do ponto de vista epidemiológico, há ainda a divisão em resistência vertical e horizontal, sendo a primeira específica às raças e qualitativa e a segunda efetiva para todas as variantes do patógeno e quantitativa (Mundt 2015). No geral, a resistência de natureza quantitativa tende a se manter estável sob condições de alta pressão de seleção do patógeno (Roberts 2002).

O grau de nutrição da planta hospedeira também afeta a taxa de reprodução dos nematoides nos tecidos infectados. No caso do nematoide de galhas, as fêmeas são responsáveis pela reprodução e ocasionam maiores danos nos tecidos infectados. De acordo com Huang (1985), a limitação de nutrientes para a planta favorece o aumento da relação machos/fêmeas, reduzindo os danos ocasionados à planta.

2.5. Uso do RNAi no controle dos nematoides

A análise do transcriptoma de espécies do gênero *Pratylenchus* tem se mostrado uma importante ferramenta para a identificação de genes candidatos em estudos envolvendo o uso de RNA de interferência (RNAi). Estudos realizados a partir de sequências expressas (EST - *Expressed Sequence Tags*), para as espécies *P. penetrans* (Mitrevva et al. 2004; Vieira et al. 2015), *P. coffeae* (Haegeman et al. 2011) e *P. thornei* (Nicol et al. 2012), e comparação com o banco de dados de anotações de genes e proteínas para *Caenorhabditis elegans* e outros nematoides permitiram a identificação de genes candidatos relacionados aos processos de crescimento, desenvolvimento e parasitismo (genes efetores) dos mesmos. Nicol et al. (2012) identificaram genes responsáveis pela codificação de enzimas de degradação de paredes celulares em plantas, e sugerem que provavelmente esses genes foram adquiridos por transferência horizontal. O silenciamento de genes envolvidos na movimentação dos nematoides, *psy-10* e *unc-87*, nas espécies *P. coffeae*, *P. thornei* e *P. zeae*, se mostrou uma importante ferramenta na estratégia de controle desses parasitas (Joseph et al. 2012; Tan et al. 2013). A inativação desses genes, empregando-se RNA dupla-fita (RNAs), permitiu uma redução superior a 81% na reprodução desses nematoides (Tan et al. 2013).

A abordagem do silenciamento gênico com a utilização de RNAi no melhoramento de plantas é promissora, pois não depende da raça, além de permitir a escolha de sequências alvo que conferem resistência a um grande número de pragas e patógenos.

2.6. Marcadores moleculares no melhoramento genético

Marcadores SSR (*Simple Sequence Repeat*) e SNP (*Single Nucleotide Polymorphism*) vem sendo amplamente empregados em estudos de mapeamento em soja (Cregan et al. 1999; Song et al. 2004; Choi et al. 2007; Hwang et al. 2009) resultando em um mapa consenso (Consensus Map 4.0) contendo 5.500 marcadores genéticos (www.soybase.org) (Hyten et al. 2010).

Estima-se que, no genoma da soja, existam 4 a 5 milhões de SNPs baseado na taxa de 280 SNPs observados nos 76 Kpb estudados por (Zhu et al. 2003). Diversos métodos de alto desempenho estão sendo designados para genotipar rapidamente um grande número de marcas no genoma de espécies diferentes. O ensaio Infinium (Illumina Inc., San Diego, CA) utiliza uma plataforma capaz de gerar dados de qualidade, apresentando também uma reprodutibilidade consistente. A natureza *high throughput* do ensaio torna a plataforma interessante para a genotipagem de populações biparentais usadas para o mapeamento de QTL. O chip BARCSoySNP6k foi configurado para incluir 6.000 SNPs que foram pré-selecionados para ter alta probabilidade de polimorfismo no germoplasma alvo, de maneira a produzir número suficiente de marcadores

polimórficos na população de mapeamento de QTL, visando a obtenção de uma cobertura satisfatória do genoma para realizar estudos de QTL em soja. O BARCSoySNP6k foi desenvolvido pelo Soybean Genomics and Improvement Lab, pertencente ao Beltsville Agricultural Research Center (BARC), para a utilização em pesquisas com soja. A genotipagem de um elevado número de marcadores distribuídos por todo o genoma auxilia na identificação de marcadores mais próximos a QTLs, favorecendo a sua utilização em estudos de seleção genômica ou mesmo na seleção assistida por marcadores moleculares (SAM).

No melhoramento genético de plantas, busca-se cada vez mais a seleção assistida por marcadores moleculares, visando obter maior eficiência na transferência e posterior seleção de genes favoráveis. Marcadores moleculares ligados a diferentes características de importância econômica têm sido desenvolvidos, permitindo a seleção indireta de caracteres desejáveis em gerações segregantes precoces. Esta estratégia reduziria tempo, recursos e energia, necessários não só para se obter grandes populações segregantes por várias gerações, mas também para estimar parâmetros usados na seleção indireta (Borém and Caixeta 2009). Dentre as várias aplicações, os marcadores moleculares podem ser usados no processo de retrocruzamento para antecipar a recuperação do genoma do genitor recorrente. Somente as linhagens que possuem o alelo de resistência idêntico ao genitor recorrente são utilizadas nos próximos ciclos de retrocruzamento, diminuindo dessa forma o trabalho e o espaço no desenvolvimento do estudo (Alzate-Marin et al. 2005). A SAM (Seleção Assistida por Marcadores) tem se apresentado como importante ferramenta na seleção de genótipos de soja superiores e resistentes a doenças e pragas, entre elas o nematoide do cisto, aumentando dessa forma a eficiência de seleção (Machado et al. 2006; Santana et al. 2014).

No caso de caracteres quantitativos, o uso da seleção genômica para a identificação de genótipos superiores tem sido frequente para diferentes espécies (Lorenz et al. 2012; Massman et al. 2013; Jarquín et al. 2014; Xavier et al. 2016). A grande disponibilidade de marcadores obtidos com a genotipagem em larga escala, aliada à redução do custo de genotipagem, têm contribuído para a utilização da seleção genômica em programas de melhoramento.

Em soja, as abordagens GBS (*Genotype by Sequencing*) (Elshire et al. 2011) e chips de SNP (Song et al. 2013) têm sido muito utilizadas em estudos de GWAS (*Genome Wide Association Study*) para diferentes finalidades, como para a avaliação de caracteres agronômicos (Sonah et al. 2015; Zhang et al. 2015b, Contreras-Soto et al. 2017; Fang et al. 2017), estudos envolvendo a análise da composição da semente (Leamy et al. 2017), bem como para a avaliação da resistência à podridão vermelha da raiz (Zhang et al. 2015a) e da resistência ao nematoide de cisto (Zhang et al. 2016).

Basicamente, um conjunto de indivíduos é utilizado como população de treinamento, para a obtenção dos dados genotípicos e fenotípicos. A metodologia consiste em estimar o valor

genético genômico (VGG) desses indivíduos para estimar o VGG de uma população de validação, que possui apenas os dados genotípicos (Meuwissen et al. 2001). Estudos tem demonstrado uma maior acurácia em modelos de seleção genômica com a utilização de QTLs, previamente identificados em estudos de mapeamento, como efeito fixo, juntamente com os demais marcadores como efeito aleatório (Bernardo 2014; Spindel et al. 2015; Arruda et al. 2016)

2.7. Mapeamento de QTLs

O advento dos marcadores moleculares aliado às técnicas de genotipagem em larga escala permitem a obtenção de mapas de ligação saturados, que são de grande relevância para o mapeamento de QTLs (Ferreira and Grattapaglia 1998). A grande maioria dos caracteres de importância econômica em plantas é classificada como de natureza quantitativa. Tais caracteres são aqueles cuja expressão fenotípica apresenta uma variação contínua, atribuída à segregação simultânea de muitos genes distribuídos pelo genoma. As regiões cromossômicas referentes à sua localização são definidas como QTLs (*Quantitative Trait Loci*). Com o advento dos marcadores moleculares tornou-se possível mapear tais regiões (QTLs) que afetam esses caracteres quantitativos (Toledo et al. 2008).

O mapeamento de QTLs baseia-se na identificação da associação não-aleatória, ou preferencial, entre os marcadores moleculares flanqueando o gene de interesse, sendo que os genótipos dos marcadores fornecem informações referentes ao QTL. As informações da segregação dos marcadores moleculares e os valores de médias fenotípicas de cada indivíduo permitem estimar o número, a posição, os efeitos, as interações entre QTLs, bem como as interações entre QTL e ambientes (Jiang and Zeng 1995; Ferreira and Grattapaglia 1998; Zeng 2001).

Para a identificação de QTLs neste contexto é necessário a obtenção de um mapa de ligação genética satisfatório. Diversos fatores podem afetar a qualidade do mapa genético, dentre eles o tipo de marcador molecular utilizado, bem como a variabilidade genética existente para a característica estudada. Outro fator que impacta significativamente no resultado de mapeamento é o tamanho da amostra ou população a ser analisada, interferindo diretamente na capacidade de se determinar a correta sequência dos marcadores no mapa genético. A distância entre os marcadores é estimada com base na frequência de recombinação entre os mesmos (Schuster and Cruz 2004). Portanto, a correta e adequada estimativa das distâncias genéticas entre os marcadores é primordial para a acurácia e precisão do ordenamento dos mesmos e para a definição dos grupos de ligação, de acordo com o número básico de cromossomos da espécie.

Em programas de melhoramento genético, o conhecimento e o entendimento da base genética dos caracteres quantitativos tem importância capital pois permite que as corretas estratégias de seleção sejam definidas e devidamente implementadas, de acordo com o número de genes influenciando a característica de interesse, bem como a inter-relação entre os mesmos (Falconer and Mackay 1996).

A existência de desequilíbrio gamético devido à ligação física é o princípio fundamental das análises de mapeamento de QTLs, sendo decorrente da redução da frequência de recombinação entre genes situados em regiões próximas entre si ao longo de determinado cromossomo (Lynch and Walsh 1998; Coelho 2000). Por meio de análises estatísticas adequadas é possível estimar os efeitos quantitativos associados ao marcador.

Diferentes metodologias para o mapeamento de QTLs foram desenvolvidas, com sucessivo avanço tanto em relação ao número de informações obtidas como em relação à precisão das mesmas. O mapeamento por intervalo ou IM (*Interval Mapping*) (Lander and Botstein 1989) fundamenta-se na busca por um possível QTL entre um par de marcadores adjacentes. A procura por este QTL ocorre de forma condicionada a certo intervalo. Entretanto, o poder estatístico do IM é relativamente baixo, o que pode ser explicado parcialmente pela não independência dos diferentes intervalos na presença de mais de um QTL (Doerge 2002).

Posteriormente, a abordagem de mapeamento por intervalo composto ou CIM (*Composite Interval Mapping*) foi proposta por Zeng (1993), mostrando-se mais eficiente que o IM por combinar o mapeamento por intervalos e regressão linear múltipla. A utilização dos marcadores como co-fatores contribui para a redução da variação genética resultante dos demais QTLs não incluídos no modelo (Liu 1998). Assim, o mapeamento CIM possibilita maior precisão nas estimativas dos efeitos e da localização do QTL, pois inclui um maior número de marcadores no modelo, reduzindo, dessa forma, o resíduo do mesmo. Além disso, essa abordagem permite o controle de falsos positivos, pois os marcadores utilizados como co-fatores bloqueiam os efeitos de possíveis QTLs pertencentes às regiões adjacentes ao intervalo analisado (Doerge 2002). Apesar do seu maior poder estatístico, o mapeamento CIM considera um modelo com um único QTL a cada posição genômica testada, ou seja, o efeito aditivo de cada QTL é estimado considerando um único QTL por vez. Múltiplos testes são realizados e o *threshold* é determinado por teste de permutação (Churchill and Doerge 1994; Chen and Storey 2006).

No caso da abordagem de mapeamento por intervalos múltiplos ou MIM (*Multiple Interval Mapping*), proposta por Kao et al. (1999), o modelo considera a existência de QTLs segregando conjuntamente na população. A inclusão do efeito de mais de um QTL no modelo aumenta o poder dos testes para identificação de QTLs, reduz a variância residual e aumenta a precisão na

estimativa de posição dos QTLs. Além disso, o mapeamento MIM corrobora para a maior compreensão da arquitetura genética dos caracteres quantitativos, possibilitando estimar os efeitos de epistasia (interação) entre os QTLs (Zeng 2001). Além disso, através do mapeamento MIM é possível estimar o valor genotípico dos indivíduos, os quais podem ser diretamente utilizados em programas de melhoramento genético (Zeng 2001).

Como a abordagem de mapeamento univariado possibilita a identificação de genes que tenham efeito em uma única variável e, geralmente os estudos de mapeamento de QTLs envolvem mais de uma característica, Jiang and Zeng (1995) propuseram uma extensão do mapeamento por intervalo composto para o caso multivariado, denominado de mapeamento para intervalo composto para múltiplas características ou MT-CIM (*Multiple Trait Composite Interval Mapping*).

Posteriormente, (Silva et al. 2012) expandiram o MIM para a análise conjunta de fenótipos correlacionados denominado mapeamento por intervalos múltiplos para múltiplas características ou MT-MIM (*Multiple Trait Multiple Interval Mapping*).

O maior poder de detecção de regiões genômicas, associadas a características quantitativas, no caso das metodologias multivariadas, refere-se à utilização dos padrões de correlação representados na matriz de variâncias/covariâncias, resultando na identificação de QTLs não detectados nas análises univariadas. Segundo Jiang and Zeng (1995) e Silva et al. (2012), pode-se obter incremento no poder de detecção de QTLs considerando a estrutura de correlação entre múltiplos fenótipos. Além disso, é possível obter maior precisão da localização e das estimativas dos parâmetros genéticos em relação às análises de fenótipos individuais.

Na abordagem multivariada, é possível estimar tanto a interação entre QTLs quanto a interação QTL x Ambiente, possibilitando um maior entendimento da sua arquitetura genética (Silva and Zeng 2010).

A relevância da abordagem de mapeamento multivariado no melhoramento de plantas é fundamentada na constante busca pela obtenção de genótipos superiores que reúnem alelos favoráveis para diversas características, simultaneamente (Walker et al. 2010). Deste modo, a identificação de regiões genômicas que controlam um conjunto de variáveis contribui de forma significativa para promover a seleção de genótipos superiores, possibilitando ganhos de seleção para diversas características concomitantemente.

Na literatura, somente o trabalho de Möller (2017) foi encontrado utilizando a abordagem multivariada em soja. No entanto, essa mesma abordagem vem sendo utilizada em estudos de mapeamento realizados em milho, para características fenotípicas associadas à absorção de fósforo (Gianotto 2015), e na área médica, para mapeamento de genes associados à hipertensão e obesidade

(Duarte 2007). Todos esses trabalhos ressaltam a superioridade da abordagem multivariada na identificação de QTLs, considerando variáveis correlacionadas, em relação à abordagem univariada.

2.8. QTLs associados à resistência aos nematoides

Diversos QTLs para resistência ao nematoide do cisto da soja (NCS) (*Heterodera glycines*) têm sido identificados em quase todos os cromossomos dessa espécie (Concibido et al. 2004; Brucker et al. 2005; Guo et al. 2006a, b; Winter et al. 2007; Wu et al. 2009; Kazi et al. 2010; Vuong et al. 2010; Ferreira et al. 2011; Chang et al. 2011; Jiao et al. 2015;). Em tais estudos, (Ferreira et al. 2011) encontraram interações epistáticas detectadas entre locos, o que indica resistência às raças 9 e 14 e também verificaram a presença de QTLs de maior e menor efeito para a resistência ao NCS.

Alguns estudos para identificação de QTLs de resistência para o gênero *Pratylenchus* têm sido efetuados em cevada (Sharma et al. 2011), em trigo (Sheedy and Thompson 2009; Linsell et al. 2013), em banana (Sundararaju 2010) e em alfafa-lipulina (*Medicago littoralis*) (Oldach et al. 2014). Linsell et al. (2013) uma população de 150 linhagens duplo-haplóides de trigo para genotipagem pelo método *Diversity Array Technology* (DArT), e identificou seis QTLs significativos para resistência ao *Pratylenchus thornei*, utilizando o método do mapeamento por intervalo. Cinco destes QTLs identificados explicavam de 30 a 48% da variação fenotípica e foram utilizados para o desenvolvimento de marcadores SSR e SNP para a seleção assistida no melhoramento de trigo. Oldach et al. (2014) utilizou uma população de 138 linhagens de *M. littoralis* e identificou um QTL significativo explicando 49% da variação fenotípica de resistência a *P. neglectus*. Análises comparativas da sequência da região identificada com o genoma com outras espécies de leguminosas permitiram identificar que o intervalo desse QTL também está presente em *Lotus japonicus*, feijão (*Phaseolus vulgaris*), soja (*Glycine max*) e grão de bico (*Cicer arietinum*). Além disso, os autores avaliaram os 55 genes presentes na região do QTL identificado e concluíram que cinco deles são genes candidatos responsáveis pela resistência. De Luca et al. (2011) realizaram um estudo para identificação filogenética entre diferentes espécies de *Pratylenchus*, e o alinhamento das sequências ITS (*Internal Transcribed Spacer*) identificou uma maior proximidade entre as espécies *P. brachyurus* e *P. neglectus*. Tal resultado evidencia a relevância da investigação do QTL associado à resistência a *P. neglectus* (Oldach et al. 2014), para a confirmação da presença do mesmo em outras leguminosas, bem como a validação da sua eficácia para conferir resistência aos nematoides do gênero *Pratylenchus*.

Até o momento não há relatos de identificação de QTLs para a espécie *P. brachyurus* na cultura da soja. Dessa forma, a identificação de QTLs que controlam a resistência da soja ao *P. brachyurus* é extremamente relevante para o conhecimento do mecanismo de controle da

enfermidade e irá auxiliar em trabalhos futuros de diversos grupos de pesquisa, facilitando a identificação e/ou, o processo de seleção assistida de genótipos resistentes ao nematoide-das-lesões-radiculares.

3. MATERIAL E MÉTODOS

3.1. Cruzamento dos genitores

A escolha dos genitores utilizados para os cruzamentos foi realizada, primeiramente, considerando-se o critério de resistência e suscetibilidade ao nematoide-das-lesões-radiculares, para a obtenção de uma população F_2 segregante para essa característica de interesse. Posteriormente, foram considerados atributos agronômicos desejáveis, em soja, como potencial produtivo e resistência ao acamamento.

Duas linhagens desenvolvidas pela FTS Sementes S.A. foram utilizadas. A linhagem FTPG06A não apresenta resistência ao nematoide *P. brachyurus*. No entanto, possui resistência à 11 raças de nematoide de cisto, sendo resistente às raças 1, 2, 3, 4, 5, 6, 9, 10 e 14, e moderadamente resistente às raças 4+ e 14+. Essa cultivar pertence ao grupo de maturidade (GM) 8.1, possui grande capacidade de engalhamento, alto potencial produtivo e ampla adaptabilidade. Já a linhagem FTPG12X pertence ao GM 8.0 e apresenta baixo fator de reprodução do nematoide, sendo considerado moderadamente resistente ao *P. brachyurus*. Além disso, essa linhagem apresenta alto potencial produtivo, resistência ao acamamento e ampla adaptabilidade.

Os cruzamentos entre os genitores foram realizados diariamente, entre dezembro de 2013 a fevereiro de 2014, no Centro de Pesquisa e Desenvolvimento de Cultivares de Soja, da FTS Sementes S.A., localizado em Ponta Grossa, Paraná. A área situa-se a 960 metros de altitude, latitude 25°10'11,28"S e longitude 50°8'55,53"W. A semeadura dos dois genitores foi realizada na estufa, manualmente, em linhas de 3 metros. Foram realizadas quatro datas de semeadura (25/10/2013, 15/11/2013, 6/12/2013 e 27/12/2013), para garantir a disponibilidade de flores viáveis e botões, de ambos os genitores, durante todo o período. O solo da estufa foi revolvido antes do plantio e adubado com 300 kg.ha⁻¹ da fórmula 00-20-20. Os tratos culturais utilizados foram os recomendados para a cultura da soja.

Foram realizados cruzamentos artificiais recíprocos entre os genitores. No momento do cruzamento foi realizada a emasculação do genitor feminino e posterior contato do pólen, proveniente do genitor masculino, com o estigma do botão floral, do genitor feminino. Cada botão floral foi devidamente etiquetado com a data de realização do cruzamento, e a identificação do genitor masculino.

3.2. Condução da geração F₁

Após a colheita das vagens foram obtidas 128 sementes F₁, provenientes do cruzamento entre os genitores FTPG06A e FTPG12X. As plantas F₁ foram conduzidas no campo experimental do Centro de Pesquisa e Desenvolvimento de Cultivares de Soja, da FTS Sementes S.A., localizado em Ponta Grossa, Paraná. Inicialmente, as sementes foram germinadas em rolo de papel *germitest*, e após 10 dias foram transplantadas em uma área com cerca elétrica para evitar o ataque de lebres e outros animais. No total, foram obtidas 13 plantas F₁, utilizando-se como genitor feminino a cultivar FTPG06A, e 34 plantas F₁, provenientes do genitor feminino FTPG12X. O espaçamento utilizado entre as linhas de semeadura e entre as plantas de cada linha foi de um metro. O solo da área foi revolvido antes do plantio e adubado com 300 kg.ha⁻¹ da fórmula 00-20-20. Os tratamentos culturais utilizados foram os recomendados para a cultura da soja.

3.3. Obtenção da população F₂

As sementes da população F₂ foram obtidas por autofecundação natural das plantas F₁. Ao final do ciclo as plantas foram manualmente colhidas, com tesouras de poda. As vagens contendo sementes F₂ foram manualmente debulhadas e armazenadas em pacotes individuais, considerando as 47 plantas F₁. Para a avaliação da resistência ao nematoide-das-lesões-radiculares foram selecionadas 200 sementes da população F₂. Do total de sementes selecionadas, 52 foram provenientes de plantas F₁ cujo genitor feminino corresponde à linhagem FTPG12X e 148 sementes foram resultantes de plantas F₁ cujo genitor feminino corresponde à linhagem FTPG06A.

3.4. Avaliações fenotípicas da população F₂

O experimento para obtenção dos dados fenotípicos, da população F₂, foi realizado em casa de vegetação, com ambiente controlado, no centro de análises da Agrolab, em Primavera do Leste, Mato Grosso, Brasil. A temperatura na casa de vegetação foi mantida entre 26 e 30° Celsius, sendo que os aquecedores eram acionados quando a temperatura atingia 26° Celsius. A irrigação foi realizada manualmente, para evitar respingos e, conseqüentemente, a contaminação entre os tubetes.

No total foram avaliadas 10 repetições de cada genitor (FTPG06A e FTPG12X) e das testemunhas consideradas com alto (TMG 132 RR) e moderado (FTS IBYARA RR) fator de reprodução ao *P. brachyurus*, além dos 188 indivíduos da população F₂, em delineamento inteiramente casualizado. As características avaliadas foram: comprimento da parte aérea (CA) e da raiz (CR), peso de matéria fresca da parte aérea (PA) e da raiz (PR) e fator de reprodução (FR),

obtido após a divisão do número de espécimes observadas no final do ciclo pelo número de espécimes inoculadas.

As sementes foram semeadas em substrato composto de 50% de areia autoclavada e 50% substrato comercial. A composição de 50% de areia foi utilizada para permitir a remoção de substrato a partir do sistema radicular na hora da colheita. Os lotes do substrato comercial e a areia foram previamente analisados para garantir a ausência de nematoides. A mistura foi adicionada aos tubetes plásticos, de 6 cm de diâmetro e 19 cm de altura, totalizando 500 mL de solo. Posteriormente, os tubetes foram alocados em bandejas, delimitando uma distância de 15 cm entre os mesmos. Considerando a disponibilidade de sementes dos genitores e da testemunha, foram semeadas três sementes por tubete, sendo o desbaste realizado alguns dias após a germinação, deixando uma única planta por tubete.

A inoculação foi realizada 10 dias após o plantio, utilizando-se inóculo obtido a partir de uma população pura de *P. brachyurus*, multiplicada em plantas de quiabo, na casa de vegetação. Para o preparo do inóculo, a extração dos espécimes foi realizada conforme a metodologia de (Coolen and D’Herd 1972). Em seguida, a suspensão foi recolhida em béquero e a contagem dos ovos foi realizada utilizando-se microscópio composto com o auxílio da câmara de *Peters*, para ajustar a solução à concentração desejada. Foram inoculados 800 juvenis por planta, em um orifício de aproximadamente 1,5 cm, próximo ao caule de cada planta, posteriormente coberto com solo esterilizado.

A avaliação foi realizada após 50 dias da inoculação dos nematoides, retirando-se com cuidado a planta do tubete para separação da parte aérea do sistema radicular. As raízes foram lavadas em água corrente para a eliminação de partículas e solo. Foram realizadas medições e pesagens tanto da parte aérea quanto da raiz de cada planta. Posteriormente, para o cálculo do FR, as raízes foram processadas pelo método Coolen and D’Herd (1972). Após serem trituradas em um liquidificador, por 30 segundos, com solução de água e hipoclorito de sódio, na proporção de 3 para 1, a solução foi peneirada utilizando-se a peneira de 200 *mesh* para a limpeza da amostra. Em seguida, todo o material foi peneirado utilizando-se as peneiras de 60 *mesh*, 200 *mesh* e 500 *mesh*. Por fim, os ovos retidos na peneira de 500 *mesh* foram coletados em um becker com a ajuda de uma pisseta contendo água. O número de ovos foi contabilizado e o FR foi estimado de acordo com a equação 1.

$$\text{FR} = \text{População final} / \text{População inicial} \quad (1)$$

Segundo o critério inicialmente proposto por Oostenbrink (1966), indivíduos com FR inferior e superior a 1 são considerados como genótipos resistentes (não multiplicadores) e suscetíveis (multiplicadores) ao nematoide, respectivamente. No entanto, estimativas de FR superiores a 1 têm sido frequentemente observadas em outros estudos envolvendo a avaliação de *P. brachyurus* em soja (Costa and Ferraz 1998; Alves 2008; Machado and Araújo Filho 2016), mostrando que o critério qualitativo de Oostenbrink (1966) pode ser rigoroso e pouco real na interação complexa envolvendo a soja e os nematoides. Este argumento pode ser substancialmente válido se considerarmos os estudos recentes de mapeamento genético em soja para resistência a nematoides, nos quais múltiplos locos foram detectados explicando proporção significativa da variação, em específico para FR (seção 2.8).

3.5. Análises Fenotípicas da população F₂

Inicialmente, os dados das características avaliadas foram analisados de forma exploratória, por meio da construção de histogramas, linhas de tendência entre as características e o cálculo das correlações de Pearson entre as mesmas. Tais análises foram realizadas utilizando-se o pacote *Performance Analytics* (Peterson et al. 2015), disponível no software estatístico R (R Development Core Team 2017).

Posteriormente, os dados das características foram analisados por meio da abordagem de modelos mistos, através do pacote *nlme* (Pinheiro and Bates 2000), também disponível no software R. O seguinte modelo estatístico foi ajustado para cada característica avaliada:

$$Y_{ij} = \mu + G_j + \varepsilon_{ij} \quad (2)$$

em que,

Y_{ij} é a observação fenotípica aleatória da característica para o indivíduo j ($j = 1, 2, \dots, 204$) na repetição i ($i = 1, 2, \dots, 10$);

μ é o efeito fixo da média geral do experimento;

G_j é o efeito aleatório do j -ésimo indivíduo;

ε_{ij} é o erro associado à observação Y_{ij}

O ajuste do modelo estatístico foi realizado considerando resíduos identicamente distribuídos e independentes entre si, tendo-se homogeneidade de variâncias residuais (σ^2) e ausência de covariâncias e correlações residuais ($\varepsilon_{ij} \sim MVN(0, \sigma^2)$). Então, no presente ajuste, considerou-se uma estrutura mais simples $\sigma^2 \mathbf{I}$ para a matriz residual $\mathbf{R}$ de variâncias e covariâncias.

A partir das análises realizadas, foram estimadas as variâncias genética ($\hat{\sigma}_G^2$), residual ($\hat{\sigma}_E^2$), e fenotípica ($\hat{\sigma}_F^2$) para cada característica, possibilitando a subsequente estimação do coeficiente de herdabilidade no sentido amplo ($\hat{h}^2$). Além disso, foram obtidos os valores genéticos preditos (BLUPs, do inglês *Best Linear Unbiased Predictions*) de 183 genótipos da população F₂, que tiveram dados fenotípicos disponíveis e que foram posteriormente utilizados na etapa de mapeamento de QTLs, dos genitores da população e das testemunhas. Todas essas estimativas foram obtidas utilizando-se o método da máxima verossimilhança restrita (REML, do inglês *Restricted Maximum Likelihood*), proposta por Patterson and Thompson (1971).

3.6. Extração do DNA

A coleta de trifólios jovens dos indivíduos F₂, dos genitores e das testemunhas foi realizada na casa de vegetação da Agrolab, em Primavera do Leste, Mato Grosso do Sul, no dia anterior à avaliação do experimento. As folhas coletadas foram armazenadas em tecido *voil* e colocadas diretamente em nitrogênio líquido para evitar sua degradação. Os trifólios foram macerados em nitrogênio líquido e a extração foi realizada utilizando-se o método CTAB (Doyle and Doyle 1990), com modificações. A qualidade e concentração do DNA extraído foram avaliadas por eletroforese em gel de agarose 1% tratado com SYBR Safe, por comparação com o DNA-padrão do *fago lambda*.

3.7. Dados genotípicos da população F₂

O DNA de 174 indivíduos da população F₂, dos genitores e das duas testemunhas foram utilizados para a genotipagem com o *SoySNP6k Bead Chip*, um subconjunto do *SoySNP50K*, na plataforma Infinium da Illumina (Song et al. 2013). A metodologia utilizada foi a mesma descrita em Fan et al. (2003) e Fan et al. (2006) e a genotipagem foi realizada pela Neogen/Deoxi, no laboratório Neogen's GeneSeek and Igenity Genomics Operation, localizado em Lincoln, Nebraska, Estados Unidos. O *Illumina BeadStation 500G* foi utilizado para determinar diferenças genotípicas referentes à plataforma *Golden Gate*. O programa *GenCall* foi utilizado para verificar a intensidade da fluorescência Cy3 e Cy5, e assim determinar a presença de polimorfismo entre os genitores. Os genótipos da população F₂ foram avaliados no programa *Genome Studio* (Illumina Inc.) e os marcadores SNPs gerados foram filtrados de acordo com os seguintes critérios: i) eliminação dos SNPs com dados perdidos nos genitores; ii) eliminação dos SNPs com genótipos heterozigotos nos genitores; e iii) eliminação dos SNPs não polimórficos entre os genitores; e iv) análise manual

dos SNPs eliminados dentro do programa *Genome Studio*, a fim de confirmar os passos anteriores realizados ou recuperar SNPs previamente descartados.

Para a construção do mapa genético foram utilizados marcadores polimórficos entre os genitores. Os dados da genotipagem original desses marcadores, obtidos através do Illumina, foram recodificados para o mapeamento de acordo com o seguinte critério: “A” para indivíduos contendo o alelo do genitor resistente em homozigose, “B” para os indivíduos contendo o alelo do genitor susceptível em homozigose e “H” para os indivíduos heterozigotos para o marcador. Para os indivíduos que apresentarem falha na amplificação ou dúvida quanto à intensidade de fluorescência o sinal “-“ foi utilizado.

3.8. Construção do mapa genético

Um total de 174 indivíduos da população F₂ de mapeamento, oriunda do cruzamento entre os genitores FTPG06A e FTPG12X, e 1.351 marcadores do tipo SNP, selecionados previamente de acordo com a seção anterior, foram utilizados para a construção do mapa genético de ligação. Inicialmente, todos os marcadores foram submetidos ao teste do qui-quadrado a 5% de probabilidade de erro, seguido da correção de Bonferroni para múltiplos testes, com o objetivo de investigar a segregação esperada. Com base nos marcadores SNPs que apresentaram segregação Mendeliana esperada, foi possível iniciar a construção do mapa genético dentro da versão 2.0-4 do pacote *OneMap* (Margarido et al. 2007), disponível no software estatístico R (R Development Core Team 2017). A primeira etapa foi estimar a fração de recombinação entre cada possível par de SNPs, utilizando-se abordagem de dois pontos. A segunda etapa foi ordenar os SNPs dentro de cada um dos 20 grupos de ligação (GLs) da soja, que foram previamente estabelecidos com base no seu genoma de referência (Schmutz et al. 2010).

Para ordenação dos SNPs dentro de cada possível GL, utilizou-se abordagem multiponto, baseada em cadeias de Markov ocultas (*HMM*, do inglês *Hidden Markov Models*). Inicialmente, todas as possíveis ordens envolvendo cinco SNPs iniciais foram investigadas (função *order.seq*), e as frações de recombinação de dois pontos, estimadas na primeira etapa, foram atualizadas no contexto multiponto, considerando a informação de todos os SNPs de determinado GL. A ordem mais provável entre os cinco SNPs iniciais foi detectada com base na função de máxima verossimilhança do modelo. Posteriormente, os SNPs remanescentes do GL foram testados e incluídos, um a um, na respectiva ordem mais provável (função *try.seq*), e possíveis inversões na ordenação foram investigadas a cada três SNPs (função *ripple*). Durante todo o processo, gráficos de *heatmaps* envolvendo as frações de recombinação multiponto e os valores de LOD (*logarithm of the odds*) score foram utilizados como diagnóstico de crescimento do mapa genético. Ao final, os

marcadores com desvios de segregação Mendeliana, desconsiderados inicialmente, foram testados e incluídos em seus respectivos GLs, e as distâncias genéticas foram estimadas utilizando a função de mapeamento Kosambi (Kosambi 1943). Uma probabilidade de erro de 5% foi utilizada na etapa final de atualização das distâncias genéticas, visto que superestimativas das frações de recombinação poderiam estar sendo consideradas em função de erros de genotipagem e sequenciamento.

A representação do mapa genético final foi realizada utilizando-se o software MapChart (Voorrips 2002).

3.9. Mapeamento de QTLs

O mapeamento de QTLs foi realizado utilizando-se o modelo de mapeamento intervalos múltiplos univariado (MIM – *Multiple Interval Mapping*) e multivariado (MT-MIM – *Multiple Trait Multiple Interval Mapping*). Na abordagem MIM (Kao and Zeng 1997; Kao et al. 1999; Zeng et al. 1999), o seguinte modelo de mapeamento foi empregado para cada uma das variáveis analisadas:

$$y_i = \mu + \sum_{r=1}^m \beta_{1r} x_{ir} + \sum_{r=1}^m \beta_{2r} z_{ir} + \varepsilon_i \quad (3)$$

em que,

y_i é o valor genético predito (BLUP) do i -ésimo indivíduo ($i = 1, 2, \dots, 174$);

μ é o intercepto ou a média geral;

β_{1r} é o coeficiente da regressão que mede o efeito aditivo do r -ésimo QTL incluído no modelo ($r = 1, 2, \dots, m$);

x_{ir} é uma variável explanatória, definida de acordo com o modelo genético de Cockerham (Kao et al. 1999; Kao and Zeng 2002). Para cada indivíduo i da população F_2 , a variável x_{ir} assume valor de 1, 0 ou -1 se o r -ésimo QTL possui genótipo homozigoto dominante (QQ), heterozigoto (Qq) ou homozigoto recessivo (qq), respectivamente. As probabilidades dos genótipos do QTL foram calculadas de acordo com cada intervalo de marcadores no mapa genético de ligação, condicionalmente às informações dos marcadores flanqueadores;

β_{2r} é o coeficiente da regressão que mede o efeito de dominância do r -ésimo QTL incluído no modelo ($r = 1, 2, \dots, m$);

z_{ir} é uma variável explanatória, definida de acordo com o modelo genético de Cockerham (Kao et al. 1999; Kao and Zeng 2002). Para cada indivíduo i da população F_2 , a variável z_{ir} assume valor de $1/2$ se o r -ésimo QTL tem genótipo heterozigoto (Qq) ou $-1/2$ caso contrário.

ε_i é o resíduo do i -ésimo indivíduo. Assume-se que os resíduos são independentes entre si e identicamente distribuídos ($\varepsilon_i \sim MVN(0, \sigma_e^2)$).

O mapeamento de QTLs multivariado (MT-MIM) realizado para dois grupos de variáveis correlacionadas (*Multi1*: FR, PA, CA e CR; e *Multi2*: PA, PR e CA), foi realizado empregando-se o seguinte modelo de mapeamento (Silva and Zeng 2010; Silva et al. 2012).

$$y_{ti} = \mu_t + \sum_{r=1}^m \beta_{1tr} x_{ir} + \sum_{r=1}^m \beta_{2tr} z_{ir} + \varepsilon_{ti} \quad (4)$$

em que,

y_{ti} é o valor genético predito (*BLUP*) do i -ésimo indivíduo ($i = 1, 2, \dots, 174$) para a t -ésima característica ($t = 1, 2, \dots, T$);

μ_t é o intercepto ou a média geral da t -ésima característica;

β_{1tr} é o coeficiente da regressão que mede o efeito aditivo do r -ésimo QTL incluído no modelo ($r = 1, 2, \dots, m$) para a t -ésima característica;

x_{ir} é uma variável explanatória, definida de acordo com o modelo genético de Cockerham (Kao et al. 1999; Kao and Zeng 2002). Para cada indivíduo i , da população F_2 , a variável x_{ir} assume valor de 1, 0 ou -1 se o r -ésimo QTL possui genótipo homozigoto dominante (QQ), heterozigoto (Qq) ou homozigoto recessivo (qq), respectivamente. As probabilidades dos genótipos do QTL foram calculadas de acordo com cada intervalo de marcadores no mapa genético de ligação, condicionalmente às informações dos marcadores flanqueadores;

β_{2tr} é o coeficiente da regressão que mede o efeito de dominância do r -ésimo QTL incluído no modelo ($r = 1, 2, \dots, m$);

z_{ir} é uma variável explanatória, definida de acordo com o modelo genético de Cockerham (Kao et al. 1999; Kao and Zeng 2002). Para cada indivíduo i , da população F_2 , a variável z_{ir} assume valor de $1/2$ se o r -ésimo QTL tem genótipo heterozigoto (Qq) ou $-1/2$ caso contrário.

ε_{ti} é o resíduo do i -ésimo indivíduo na t -ésima característica. Assume-se que os resíduos são independentes entre si e identicamente distribuídos ($\varepsilon_{ti} \sim MNV(0, \Sigma_e)$).

Inicialmente, a abordagem MIM foi realizada para cada característica avaliada. A partir de um modelo nulo ajustado (sem a presença de QTL), uma busca *forward* foi realizada para a inclusão dos efeitos principais – aditivos e de dominância – de QTLs para explicar proporção da variação genética. Para tanto, o modelo da Equação 3 foi ajustado utilizando-se a aproximação de Haley-Knott (Haley and Knott 1992) a cada 1 cM no mapa genético, considerando um tamanho de janela de 5 cM tanto à direita quanto à esquerda dos marcadores flanqueadores. A presença de um possível

QTL foi baseada na metodologia de reamostragem *score* (Zou et al. 2004; Silva and Zeng 2010; Laurie et al. 2014), com um nível de significância de 10%. Para os efeitos principais dos QTLs, um nível de significância marginal de 5% foi utilizado com base na distribuição de qui-quadrado. O refinamento das posições e efeitos de QTLs foi realizado a cada novo ciclo de busca ao longo do genoma. Ao final, um procedimento de eliminação *backward* foi realizado a partir do modelo completo (com a presença de QTL), utilizando-se os mesmos critérios descritos para o procedimento *forward*.

Na sequência, a abordagem MT-MIM foi realizada para dois grupos de características estatisticamente correlacionadas: i) Multi 1: FR, PA, CA e CR; e ii) Multi 2: PA, PR e CA. A estratégia de busca por QTLs foi a mesma adotada para a abordagem MIM.

Uma vez que os efeitos principais dos possíveis QTLs foram incorporados nos modelos de mapeamento (MIM e MT-MIM), análises de epistasia envolvendo pares de QTLs foram adicionalmente realizadas. Para tanto, dois procedimentos foram aqui adotados: i) busca por efeitos de epistasia (interação) entre pares de QTLs já incluídos nos modelos (aditivos) de mapeamento (**0-D**); e ii) busca por efeitos de epistasia entre um QTL incluído no modelo e outro possível QTL ainda não detectado, com a investigação de cada posição genômica que não evidenciou, antes, a presença de um suposto QTL (**1-D**). Para ambas as abordagens (MIM e MT-MIM), a significância dos efeitos de interação foi verificada utilizando-se a estatística *score* (Zou et al. 2004; Silva and Zeng 2010; Laurie et al. 2014), com um nível de significância de 1% de probabilidade de erro.

Todas as análises de QTLs foram realizadas utilizando-se a versão 1.0-2 do pacote *Oneqtl*, escrito para o software estatístico R (R Development Core Team 2017). O pacote foi desenvolvido pelo Laboratório de Genética Estatística do Departamento de Genética da ESALQ e será disponibilizado, em breve, no Github e na rede de arquivos do R (CRAN).

3.10. Análise de Sequências Candidatas em Regiões de Interesse

Inicialmente, uma determinada região genômica foi selecionada a partir de certa posição de interesse. Tal região foi estabelecida caminhando-se tanto à direita quanto à esquerda da posição correspondente ao QTL ou da posição alvo com um suposto “QTL”, a partir de um único marcador (QTL com posição coincidente) ou de dois marcadores flanqueadores (QTL em posição intervalar). O tamanho da janela estabelecida, para cada região, foi de 500 Kb para ambos os lados, levando em consideração os padrões extensivos do desequilíbrio de ligação (DL) no genoma da soja, considerando dados de SNPs (Hao et al. 2012). Com base na plataforma de dados genômicos *PlantGDB* (Dong et al. 2004; Duvick et al. 2008), todas as sequências candidatas presentes na região, correspondendo às sequências transcritas preditas possivelmente únicas (*PUTs*, *PlantGDB-derived*

Unique Transcripts), foram selecionadas e investigadas na plataforma de bioinformática *Blast2GO* (Conesa et al. 2005; Conesa and Götz 2008).

A análise das sequências candidatas foi realizada em três etapas: i) alinhamento local de nucleotídeos contra sequências de proteínas (*BLASTx*, *Basic Local Alignment Search Tool*), considerando a base de dados não redundante (nr) do NCBI (*National Center for Biotechnology Information*); ii) mapeamento dos termos da base *Gene Ontology* (GO) associados aos *bits* obtidos na etapa anterior, alinhados contra as sequências transcritas; e iii) anotação funcional e automática de tais sequências, com a atribuição específica dos termos de GO. Para o alinhamento, foram utilizados os seguintes parâmetros: 20 *bits* de alinhamento, valor esperado (*E-Value*) de E^{-5} como *threshold*, *Word Size* de 3 e ausência de filtragem para menor complexidade. Para o mapeamento e anotação, foram utilizadas as opções estabelecidas como padrão pelo *Blast2GO*.

4. RESULTADOS E DISCUSSÃO

4.1. Análises Fenotípicas

As análises exploratórias envolvendo as cinco características avaliadas são mostradas graficamente na Figura 1. É possível observar a distribuição dos dados de cada característica na diagonal principal e as linhas de tendência e correlações de Pearson entre duas características quaisquer fora da diagonal principal. As características que apresentaram as maiores correlações positivas e estatisticamente significativas ($p < 0,001$) foram: i) PA e CA (0,76); ii) PA e PR (0,60); e iii) PR e CA (0,57). Tais resultados sugerem que os desenvolvimentos das partes aérea e de raiz estão forte e positivamente correlacionados em soja, indicando que genes em comum podem estar envolvidos na formação destas estruturas. No entanto, a característica CR mostrou estar fraca e estatisticamente correlacionada, positiva ou negativamente, com as características PA (0,19; $p < 0,01$), CA (0,19; $p < 0,01$) e PR (-0,29; $p < 0,001$), indicando que a relação entre os desenvolvimentos das partes aérea e de raiz pode ser, na verdade, complexa em soja. Liang et al. (2010) também observaram correlação positiva e estatisticamente significativa ($p < 0,01$) entre as variáveis PA e PR (0,74), corroborando com os resultados obtidos neste trabalho. Entretanto, para as variáveis CR e PR, ao contrário do observado neste trabalho, Liang et al. (2010) e Liang et al. (2014) identificaram correlação positiva, de 0,47 e 0,61, respectivamente.

A característica FR, diretamente relacionada ao *P. brachyurus*, mostrou correlação negativa e estatisticamente significativa com as características PA (-0,22; $p < 0,01$), CA (-0,23; $p < 0,001$) e CR (-0,24; $p < 0,001$). Esses resultados mostram que maiores incidências de *P. brachyurus*, em tecidos radiculares, podem afetar significativamente o desenvolvimento das partes aérea e de raiz em soja. Então, embora *P. brachyurus* se reproduza especificamente nas raízes da soja, bem como nas raízes de outras culturas (Nicol et al. 2011; Santana-Gomes et al. 2016), sendo um nematoide que ataca e causa lesões nestes tecidos (Jones and Fosu-Nyarko 2014; Rodrigues et al. 2014; Machado et al. 2015; Fosu-Nyarko and Jones 2016; Machado and Araújo Filho 2016), sua ação pode gerar grandes impactos em outras estruturas de crescimento. Isso seria biologicamente esperado, pois a morte (necrose) de todo o sistema radicular, ou de parte dele, causada por *P. brachyurus*, deveria refletir diretamente na disponibilidade de água e nutrientes (Lima et al. 2015) para outras porções radiculares e outras estruturas morfológicas de crescimento. Consequentemente, o desenvolvimento da soja poderia ser comprometido, causando clorose, nanismo das plantas (Costa and Ferraz 1998) e redução da produtividade (Ferraz 1999; Castillo and Vovlas 2007). Além disso, as lesões ocasionadas pelo nematoide, na raiz, favorecem o desenvolvimento de outros fungos

(*Fusarium virguliforme* e *Verticillium* sp.) e bactérias presentes no solo (Castillo and Vovlas 2007). No entanto, como as correlações aqui obtidas foram fracamente negativas, sugere-se que diferentes genótipos de soja comportam-se de maneira distinta quando expostos a elevadas incidências de *P. brachyurus*.

Figura 1. Correlação fenotípica entre cinco características avaliadas em soja: FR: fator de reprodução; PA: peso de matéria fresca da parte aérea; PR: peso de matéria fresca da raiz; CA: comprimento da parte aérea; CR: comprimento da raiz. Significância estatística baseada na distribuição de probabilidade do Teste *t-student* (* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$)

Ao avaliar os efeitos da interação entre *P. brachyurus* e *Meloidogyne javanica*, na cultivar de soja FT Cristalina, (Ferraz 1995) detectou reduções estatisticamente significativas (Teste de Tukey; $p < 0,05$) dos pesos de matéria fresca das raízes (PFR) e matéria seca da parte aérea (PSPA) com o aumento gradativo do inóculo de *P. brachyurus* (1.000, 5.000 e 10.000 nematoides/planta). Tais reduções não foram observadas com a incorporação e o aumento do inóculo de *Meloidogyne javanica* (1.000 e 5.000 nematoides/planta), mostrando que os danos causados na cultivar FT Cristalina foram provenientes do ataque de *P. brachyurus*. Esses resultados reforçam as correlações negativas e estatisticamente significativas obtidas entre FR e características das partes aérea (PA e CA) e de

raiz (CR) do presente estudo. No entanto, os resultados de Ferraz (1995) foram gerados a partir de análise de um único genótipo, de modo que as suas evidências podem não ocorrer, com a mesma intensidade, para outros genótipos de soja.

Dois exemplos claros deste fato são os estudos realizados por Goulart et al. (2013) e Mainardi and Asmus (2015). No primeiro caso, os autores, ao inocularem diferentes níveis populacionais de *P. brachyurus* (300, 600, 1.200, 2.400 e 4.800 nematoides juvenis e adultos/planta) nas raízes de outra cultivar de soja (EMGOPA-316), verificaram, paradoxalmente, um aumento nos pesos da parte aérea e de raiz com o aumento no número de nematoides inoculados. No segundo caso, para a cultivar Anta 82, os autores não detectaram quaisquer diferenças significativas nas produções de matéria fresca das raízes e matéria seca da parte aérea, bem como em outras características avaliadas (altura de plantas, teor de clorofila, taxa fotossintética e número de lesões no sistema radicular) com o aumento da densidade populacional de *P. brachyurus* em soja e demais espécies vegetais (algodão, milho, mombaça e marandu). Tais constatações foram baseadas em dois diferentes experimentos, com densidades populacionais bastante variadas de *P. brachyurus*: i) Experimento 1: 0, 333, 1.000, 3.000 e 9.000 espécimes/planta; e ii) Experimento 2: 0, 1.500, 4.500, 13.500 e 40.500 espécimes/planta. No entanto, (Mainardi and Asmus 2015) observaram danos primários mais intensos no sistema radicular da soja, mostrando que *P. brachyurus* pode gerar grandes impactos nas lavouras comerciais desta cultura. Tal constatação evidencia a relevância do presente trabalho.

Outro estudo interessante sobre a dinâmica populacional de *P. brachyurus* foi realizado por Lima et al. (2015), em campo comercial da cultivar de soja M9144RR. Esses autores realizaram dois diferentes experimentos e observaram correlações negativas e estatisticamente significativas entre densidade de *P. brachyurus* e as características altura de planta (-0,78; $p < 0,001$), número de vagens (-0,48; $p < 0,03$), peso de raiz (-0,45; $p < 0,05$) e produtividade de grãos (-0,48; $p < 0,03$). Embora não tenha sido detectada uma correlação significativa entre FR e PR no presente estudo, os resultados de Lima et al. (2015) claramente sugerem, em condições naturais de campo, que há correlação negativa e estatisticamente significativa entre *P. brachyurus* e peso de raiz em soja. Mais do que isso, há correlação negativa entre densidade de *P. brachyurus* e diversas características relacionadas ao crescimento da planta, reforçando as discussões anteriores.

Os resultados das análises estatísticas realizadas para as cinco características de interesse encontram-se na Tabela 1. Claramente, é possível notar a diferenciação entre os valores genéticos preditos (*BLUPs*) dos genitores e das testemunhas para todas as características, com exceção de PR, que apresentou valores bem inferiores e reduzida variação. Para a característica FR, o genitor FTPG12X apresentou a menor reprodução de *P. brachyurus* (1,95) em comparação com o genitor

FTPG06A (3,14) e as testemunhas FTS IBYARA RR (2,69) e TMG 132 RR (6,15). Tais proporções devem ter sido decisivas para um maior desenvolvimento das partes aérea e de raiz do genitor FTPG12X, o que pode ser mais evidentemente notado para as características de comprimento (CA e CR). De qualquer forma, a diferenciação apresentada pelos genitores, sobretudo em relação a FR, CA e CR, permitiu a obtenção de variações genéticas importantes na população F₂, o que pode ser observado pelas maiores amplitudes dos *BLUPs* e pelas elevadas herdabilidades estimadas (0,98; 0,99; e 0,99, respectivamente).

Tabela 1. Valores genéticos preditos (BLUP), variâncias fenotípica (σ_F^2), genética (σ_G^2) e residual (σ_E^2), coeficiente de herdabilidade no sentido amplo (h^2) e coeficiente de variação ambiental (CVE%) para as características fator de reprodução (FR), peso de matéria fresca da parte aérea (PA), peso de matéria fresca da raiz (PR), comprimento da parte aérea (CA) e comprimento da raiz (CR).

Característica	P1	P2	T1	T2	F ₂ (min – max)	F ₂ (1º – 3º quartis)
FR	3,14	1,95	2,69	6,15	3,45 (1,16 – 8,92)	(2,27 – 4,49)
PA	3,23	3,77	3,61	2,12	3,63 (2,02 – 4,74)	(3,43 – 3,78)
PR	3,20	3,84	3,63	2,26	3,40 (0,93 – 5,08)	(3,00 – 3,89)
CA	42,67	47,78	45,01	39,62	45,81 (36,03 – 48,99)	(44,00 – 47,99)
CR	17,59	24,78	21,53	15,51	16,50 (10,02 – 24,97)	(15,00 – 18,00)

Característica	σ_F^2	σ_G^2	σ_E^2	h^2	CVE%
FR	2,87	2,83	0,04	0,98	5,80
PA	0,21	0,18	0,03	0,88	4,80
PR	0,81	0,80	0,01	0,98	3,00
CA	6,06	6,04	0,02	0,99	0,31
CR	4,64	4,63	0,02	0,99	0,85

P1: FTPG06A; P2: FTPG12X; T1: FTS IBYARA RR; T2: TMG 132 RR.

Os valores genéticos preditos (*BLUPs*) referentes à FR variaram de 1,16 a 8,92 na população F₂. Considerando que o critério inicialmente proposto por Oostenbrink (1966) tem se mostrado rigoroso e pouco real na interação complexa envolvendo a soja e os nematoides, como

já discutido na seção 3.4, um critério de natureza quantitativa (Moura and Régis 1987), englobando diferentes graus de resistência e suscetibilidade, foi adicionalmente utilizado, no presente trabalho, para melhor classificar os indivíduos quanto ao ataque de *P. brachyurus* em suas raízes, com base nas estimativas obtidas do FR. Considerando que o indivíduo mais suscetível a *P. brachyurus* mostrou um FR estimado de 8,92 na população F₂, foram observadas as seguintes atribuições na população (Moura and Régis 1987): i) 15 (8,20%) indivíduos mostraram redução de 0 a 25% (altamente suscetível) no FR em relação ao FR de 8,92 do indivíduo mais suscetível; ii) 37 (20,22%) indivíduos apresentaram redução de 26 a 50% (suscetível); iii) 114 (62,30%) indivíduos tiveram redução de 51 a 75% (pouco resistente); iv) 17 (9,30%) indivíduos mostraram redução de 76 a 95% (moderadamente resistente); e v) nenhum (0%) indivíduo apresentou redução de 96 a 99% (resistente) e 100% (altamente resistente). Os genitores da população mostraram redução de 64,80% (FTPG06A) e 78,14% (FTPG12X) no FR em relação ao FR de 8,92 do indivíduo mais suscetível, sendo considerados pouco resistente e moderadamente resistente, respectivamente. Da mesma forma, as testemunhas utilizadas apresentaram redução de 69,84% (FTS IBYARA RR) e 31,05% (TMG 132 RR) no FR em relação ao FR de 8,92 do indivíduo mais suscetível, sendo consideradas pouco resistente e suscetível, respectivamente.

As estimativas de herdabilidade no sentido amplo (h^2) foram bastante elevadas para todas as características. Tais resultados indicam que a magnitude da correlação entre genótipo e fenótipo foi consideravelmente alta em todos os casos, de forma que grande parte da variância fenotípica estimada foi de natureza genética. Sendo um dos grandes objetivos do melhoramento genético, esta correspondência elevada foi obtida como resultado de um experimento bem delineado e conduzido, possibilitando a redução dos efeitos residuais (estimativas próximas a zero) e o aumento da precisão e qualidade experimental (baixos valores de CV_E %). Machado and Araújo Filho (2016), ao analisarem o FR em dois experimentos com 42 genótipos de soja, em casa de vegetação, obtiveram estimativas de herdabilidade próximas (95,77 e 86,89) àquelas detectadas no presente estudo. No entanto, há poucos relatos na literatura sobre a estimação de parâmetros genéticos e o estudo da herança de características relacionadas a *P. brachyurus* em soja. Dessa forma, os resultados aqui obtidos poderão fornecer um grande avanço para o entendimento da base genética envolvida.

4.2. Marcadores SNP

De um total de 6.000 SNPs disponíveis (*SoySNP6k Bead Chip*), 1.313 (21,88%) foram previamente selecionados a partir dos critérios i, ii e iii, apresentados na seção 3.6. do presente trabalho. Dos 4.687 (78,12%) SNPs inicialmente descartados, 1.720 (35,99%) foram reconsiderados para uma investigação manual e específica dentro do *Genome Studio*, de acordo com

o critério iv da referida seção. Tais marcadores reconsiderados mostraram, para a grande maioria dos indivíduos, uma pontuação na atribuição genotípica (*GenCall Score*) superior a 0,15 no processo de genotipagem, seguindo as recomendações de qualidade da Illumina Inc. No total, 38 (2,21%) SNPs foram manualmente recuperados dentro do *Genome Studio*, que, somados aos 1.313 SNPs previamente selecionados, resultaram em 1.351 (22,52%) marcadores polimórficos disponíveis para a construção do mapa genético de ligação.

Destes 1.351 marcadores polimórficos, 695 (51,44%) apresentaram desvios de segregação mendeliana, de acordo com as proporções esperadas para uma população F_2 (1:2:1). Dessa forma, apenas 656 (48,56%) SNPs estiveram efetivamente disponíveis para iniciar a construção do mapa genético, corroborando com a quantidade de marcadores utilizada por Akond et al. (2013). Esses autores, ao genotiparem uma população de soja formada por 92 linhagens endogâmicas recombinantes (RILs, do inglês *Recombinant Inbred Lines*), com o mesmo conjunto de SNPs do presente trabalho (*SoySNP6k Bead Chip*), construíram um mapa genético a partir de 657 SNPs polimórficos. Então, em ambos os casos, uma proporção significativa de SNPs (> 50%), com desvios da segregação esperada, foi excluída do conjunto total de marcadores polimórficos, algo que pode ter ocorrido devido à fatores como tamanho populacional e erros no processo de sequenciamento e/ou genotipagem (Wall et al. 2014; Maruki and Lynch 2015; Goodwin et al. 2016).

De acordo com o estudo realizado por Wall et al. (2014), erros no processo de genotipagem normalmente se reduzem com o aumento da pontuação de qualidade genotípica (GQ). No entanto, esta relação não parece ocorrer de forma linear especificamente para dados da plataforma Illumina, visto que houve uma perda da especificidade a partir de um GQ superior a 60. Além disso, também para dados de Illumina, Wall et al. (2014) observaram que as taxas de erro em processos de sequenciamento se elevaram com o aumento da profundidade de cobertura, mostrando que é preciso ter cuidado na interpretação dos resultados obtidos a partir de tais dados. Embora a plataforma Illumina seja menos suscetível a erros durante o processo, mostrando uma acurácia global superior a 99%, algumas atribuições podem ser pouco confiáveis para regiões ricas em AT (Adenina-Timina) e GC (Guanina e Citosina), havendo uma tendência de ocorrer erros de substituição (Goodwin et al. 2016).

4.3. Mapa genético

Do total de 1.351 marcadores polimórficos, 1.240 (91,78%) SNPs puderam ser mapeados com sucesso e constituíram o presente mapa de ligação, abrangendo os 20 GLs da soja, com uma distância total de 3.084,46 cM. A caracterização e representação do mapa genético de ligação estão mostradas nas Tabela 2 e Figura 2, respectivamente. Com exceção dos GLs da soja denominados

D1a, M e I, todos os demais foram subdivididos em dois ou mais subgrupos de ligação. Essas subdivisões foram realizadas quando foram observadas a presença de *gaps* maiores do que 30cM. No entanto, a informação da distância entre os marcadores no genoma de referência da soja (ConsensusMap 4.0) foi utilizada como um critério biológico no caso dos *gaps* mantidos em alguns GLs. A presença de *gaps* na construção do mapa já era esperada, considerando que o *SoySNP6k Bead Chip* não foi desenvolvido especificamente para a população em estudo. Após a etapa de filtragem dos dados, a eliminação de grande parte dos 4.649 SNPs (77,48% dos dados iniciais) ocorreu em blocos, como podemos verificar na Figura 2. Tal fato pode ter favorecido a presença de *gaps* e a falta de conexão entre os subgrupos obtidos. Jun et al. (2012) também obtiveram GLs subdivididos na construção do mapa genético para uma população de RILs em soja. Esses autores encontraram um total de 32 subgrupos envolvendo 516 marcadores SNPs, com um intervalo médio de 5,2 cM entre dois possíveis marcadores. Embora a quantidade de subgrupos detectada no presente estudo tenha sido maior, totalizando 45 subgrupos, a distância média entre dois SNPs foi relativamente menor (2,54 cM, com amplitude entre 0,17 e 7,20 cM), sugerindo maior saturação e cobertura do genoma da soja.

Em contrapartida, Lee et al. (2015) encontraram os exatos 20 GLs na construção de um mapa genético para 357 RILs de soja, com base em 2.545 marcadores SNPs. Esses autores utilizaram o conjunto denominado *BARCSoySNP6k*, que, segundo os próprios, não é o mesmo conjunto utilizado por Akond et al. (2013) e pelo presente estudo (*SoySNP6k Bead Chip*). De qualquer forma, Lee et al. (2015) obtiveram um mapa genético bastante saturado, contendo de 86 a 162 locos por cromossomo e uma distância média entre marcadores de 1,1 cM (entre 0,9 e 1,3 cM), considerando um comprimento total de 2.346 cM. No entanto, apesar dos autores também terem realizado uma filtragem inicial dos dados, é pouco provável que todos os SNPs polimórficos (2.545, do total de 2.585 SNPs, representando 98,45%) apresentaram segregação mendeliana esperada de 1:1, mesmo com um tamanho de população muito superior ao deste estudo. Isso porque erros no processo de genotipagem têm sido muito recorrentes nas estratégias baseadas em sequenciamento (Wall et al. 2014; Maruki and Lynch 2015; Goodwin et al. 2016), que podem resultar em atribuições equivocadas nas populações, gerando maiores ou menores proporções das classes genotípicas esperadas.

Neste trabalho, embora 695 SNPs tenham apresentado desvios da segregação esperada, um total de 584 (84,03%) marcadores pôde ser efetivamente incorporado no mapa genético prévio (656 SNPs). Tal mapeamento sugeriu que, para a grande maioria dos SNPs, os desvios de segregação não foram fortes o bastante a ponto de não permitirem a sua alocação por entre os demais SNPs dos GLs, o que prejudicaria a extensão do mapa genético. Contudo, pelo menos uma

subdivisão de cada GL apresentou todos os SNPs (100,00%) com desvios de segregação, exceto os GLs D1a, M e I, sugerindo que os marcadores com tais alterações tenderam a permanecer, em muitos casos, unidos dentro de certos “blocos”. Foi o que aconteceu para todas as subdivisões destacadas com um asterisco na Tabela 2, nos cromossomos contendo as letras B (3, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18 e 19), C (2, 4, 5, 6, 9, 14 e 16) e D (9).

Tabela 2. Caracterização do mapa genético de ligação da população F₂ de soja obtida a partir do cruzamento entre os genitores FTPG06A e FTPG12X.

Cromossomo	Grupo de Ligação (GL)	Número de SNPs no Mapa	Número de SNPs Com DSM¹	Proporção de SNPs Com DSM	Distância Total (cM)	Distância Média (cM)
1	D1a	32	10	31,25%	60,85	1,96
2 ^a	D1b	15	13	86,66%	73,43	5,25
2B	D1b	13	7	53,85%	66,05	5,51
2C*	D1b	25	25	100,00%	56,88	2,37
3A	N	76	4	5,26%	127,77	1,70
3B*	N	18	18	100,00%	19,52	1,15
4A	C1	36	2	5,55%	118,33	3,38
4B	C1	3	1	33,33%	1,54	0,77
4C*	C1	6	6	100,00%	7,20	1,44
5A	A1	12	0	0,00%	43,14	3,92
5B	A1	3	0	0,00%	3,02	1,51
5C*	A1	11	11	100,00%	9,53	0,95
6A	C2	31	13	41,94%	71,18	2,37
6B	C2	22	5	22,73%	77,48	3,69
6C*	C2	28	28	100,00%	34,36	1,27
7	M	63	14	22,22%	141,14	2,28
8A	A2	78	36	46,15%	224,93	2,92
8B*	A2	13	13	100,00%	36,28	3,02
9A	K	18	1	5,55%	53,40	3,14
9B	K	7	0	0,00%	13,99	2,33
9C*	K	5	5	100,00%	4,67	1,17
9D*	K	6	6	100,00%	25,14	5,03
10A	O	16	5	31,25%	108,01	7,20

(Continuação)						
10B*	O	33	33	100,00%	79,24	2,48
11A	B1	33	13	39,39%	198,15	6,19
11B*	B1	14	14	100,00%	19,64	1,51
12A	H	54	11	20,37%	150,41	2,84
12B*	H	24	24	100,00%	21,00	0,91
13A	F	91	40	43,96%	224,98	2,50
13B*	F	23	23	100,00%	43,70	1,99
14A	B2	35	7	20,00%	153,40	4,51
14B*	B2	30	30	100,00%	18,04	0,62
14C*	B2	3	3	100,00%	8,00	4,00
15A	E	32	10	31,25%	124,76	4,03
15B*	E	33	33	100,00%	31,66	0,99
16A	J	35	7	20,00%	91,44	2,69
16B*	J	4	4	100,00%	0,52	0,17
16C*	J	3	3	100,00%	2,96	1,48
17A	D2	53	10	18,87%	77,92	1,50
17B*	D2	7	7	100,00%	10,46	1,74
18A	G	30	22	73,33%	73,22	2,53
18B*	G	21	21	100,00%	30,96	1,55
19A	L	72	6	8,33%	152,12	2,14
19B*	L	7	7	100,00%	4,57	0,76
20	I	66	27	40,91%	189,48	2,92
Total	–	1.240	584	–	3.084,46	2,54

¹DSM: Desvio da Segregação Mendeliana; *: Grupos de ligação completamente formados por SNPs com DSM.

Acredita-se que tais “blocos” tenham sido gerados a partir de erros sistemáticos no processo de genotipagem e sequenciamento. Possivelmente, quando um erro de sequenciamento ocorreu em determinado momento, para certa região do genoma, erros sequenciais em cadeia foram gerados, obtendo-se “blocos” de SNPs com proporções genotípicas inesperadas na população. No entanto, é possível que alguns destes “blocos”, ou parte deles, sejam o resultado de fatores genéticos ocorridos evolutivamente, em função do processo histórico de domesticação e eventos recorrentes de seleção. É o que pode ter ocorrido para os GLs 14A e 14 B do presente

mapa genético. Song et al. (2013) observaram duas regiões genômicas com evidências de seleção ($F_{ST} \geq 0,6$; sendo F_{ST} o coeficiente de endogamia por subdivisão populacional) no cromossomo 14 da soja, considerando análise a partir de indivíduos selvagens e domesticados. Segundo os autores, essas regiões estiveram compreendidas entre duas regiões genômicas, sendo elas: (i) 13,85 e 31,85 mega pares de bases (Mbp); e (ii) 40,85 e 49,15 Mbp. Ao analisarmos os subgrupos 14A e 14B, que apresentaram marcadores com desvios de segregação, foi possível notar a presença de SNPs sequenciais compreendidos na primeira região genômica (i), mencionada acima (14A: Gm14_27937142_C_T; 14B: Gm14_17569578_A_G, Gm14_18515404_A_G, Gm14_18927676_T_C, Gm14_19030405_G_A, Gm14_19103544_T_C, Gm14_19373649_A_G, Gm14_26132219_T_C, Gm14_28004374_T_C, Gm14_28158698_T_C, Gm14_28262597_T_C, Gm14_28361034_A_G, Gm14_30024382_T_C, Gm14_31348204_T_C e Gm14_31829614_G_T), e na segunda região genômica (ii) (14A: Gm14_43270248_C_T, Gm14_43348217_C_T, Gm14_44146640_T_C, Gm14_44193109_G_A, Gm14_44393026_T_C, Gm14_44495960_G_A, Gm14_44854103_A_G, Gm14_45186718_A_G, Gm14_45234122_T_C, Gm14_45457682_A_G, Gm14_47354923_T_C, Gm14_47418167_T_C, Gm14_47475993_G_A e Gm14_47643269_A_G; 14B: Gm14_40852336_C_T, Gm14_40986359_T_C, Gm14_41187024_T_C e Gm14_41251974_T_C). Estes SNPs não estão necessariamente ordenados no mapa genético estimado (Figura 2), visto que diferenciações quanto ao posicionamento dos marcadores foram obtidas em comparação com o genoma de referência (Schmutz et al. 2010).

Figura 1. Representação do mapa genético de ligação da população F₂ de soja, obtida a partir do cruzamento entre os genitores FTPG06A e FTPG12X. À direita, os nomes dos marcadores ancorados no mapa. À esquerda, as posições dos marcadores em cada grupo de ligação da soja, em centimorgans, utilizando-se a função de mapeamento de Kosambi.

GL N (Chr 3A)

GL C1 (Chr 4A)

(Chr 4B)

(Chr 4C)

(Chr 3B)

Figura 1 (continuação) Representação do mapa genético de ligação da população F₂ de soja, obtida a partir do cruzamento entre os genitores FTPG06A e FTPG12X. À direita, os nomes dos marcadores ancorados no mapa. À esquerda, as posições dos marcadores em cada grupo de ligação da soja, em centimorgans, utilizando-se a função de mapeamento de Kosambi.

GL A1 (Chr 5A)

(Chr 5B)

(Chr 5C)

GL C2 (Chr 6A)

(Chr 6B)

(Chr 6C)

Figura 1 (continuação) Representação do mapa genético de ligação da população F₂ de soja, obtida a partir do cruzamento entre os genitores FTPG06A e FTPG12X. À direita, os nomes dos marcadores ancorados no mapa. À esquerda, as posições dos marcadores em cada grupo de ligação da soja, em centiMorgans, utilizando-se a função de mapeamento de Kosambi.

GL M (Chr 7)

GL A2 (Chr 8A)

(Chr 8B)

Figura 1 (continuação) Representação do mapa genético de ligação da população F₂ de soja, obtida a partir do cruzamento entre os genitores FTPG06A e FTPG12X. À direita, os nomes dos marcadores ancorados no mapa. À esquerda, as posições dos marcadores em cada grupo de ligação da soja, em centiMorgans, utilizando-se a função de mapeamento de Kosambi.

GL K (Chr 9A)

(Chr 9B)

(Chr 9C)

(Chr 9D)

GL O (Chr 10A)

(Chr 10B)

Figura 1 (continuação) Representação do mapa genético de ligação da população F₂ de soja, obtida a partir do cruzamento entre os genitores FTPG06A e FTPG12X. À direita, os nomes dos marcadores ancorados no mapa. À esquerda, as posições dos marcadores em cada grupo de ligação da soja, em centiMorgans, utilizando-se a função de mapeamento de Kosambi.

Figura 1 (continuação) Representação do mapa genético de ligação da população F₂ de soja, obtida a partir do cruzamento entre os genitores FTPG06A e FTPG12X. À direita, os nomes dos marcadores ancorados no mapa. À esquerda, as posições dos marcadores em cada grupo de ligação da soja, em centiMorgans, utilizando-se a função de mapeamento de Kosambi.

GL F (Chr 13A)

(Chr 13B)

GL B2 (Chr 14A)

(Chr 14B)

(Chr 14C)

Figura 1 (continuação) Representação do mapa genético de ligação da população F₂ de soja, obtida a partir do cruzamento entre os genitores FTPG06A e FTPG12X. À direita, os nomes dos marcadores ancorados no mapa. À esquerda, as posições dos marcadores em cada grupo de ligação da soja, em centiMorgans, utilizando-se a função de mapeamento de Kosambi.

Figura 1 (continuação) Representação do mapa genético de ligação da população F_2 de soja, obtida a partir do cruzamento entre os genitores FTPG06A e FTPG12X. À direita, os nomes dos marcadores ancorados no mapa. À esquerda, as posições dos marcadores em cada grupo de ligação da soja, em centiMorgans, utilizando-se a função de mapeamento de Kosambi.

GL D2 (Chr 17A)

(Chr 17B)

GL G (Chr 18A)

(Chr 18B)

Figura 1 (continuação) Representação do mapa genético de ligação da população F₂ de soja, obtida a partir do cruzamento entre os genitores FTPG06A e FTPG12X. À direita, os nomes dos marcadores ancorados no mapa. À esquerda, as posições dos marcadores em cada grupo de ligação da soja, em centiMorgans, utilizando-se a função de mapeamento de Kosambi.

Figura 1 (continuação) Representação do mapa genético de ligação da população F₂ de soja, obtida a partir do cruzamento entre os genitores FTPG06A e FTPG12X. À direita, os nomes dos marcadores ancorados no mapa. À esquerda, as posições dos marcadores em cada grupo de ligação da soja, em centiMorgans, utilizando-se a função de mapeamento de Kosambi.

Figura 2. Representação gráfica da disposição dos SNPs excluídos do mapa genético obtido, para a população F_2 , de acordo com os 20 GLs do genoma de referência da soja (ConsensusMap 4.0).

4.1. Mapeamento de QTLs

Os QTLs identificados neste trabalho receberam a notação proposta por Gianotto (2015) $F@Chr_P$, em que ' F ' representa o fenótipo no modelo de mapeamento univariado, sendo posteriormente substituído por " $Multi$ " no mapeamento multivariado; ' Chr ' representa o cromossomo da soja onde o QTL se localiza; e ' P ' representa a posição (em cM) do QTL, considerando o mapa de ligação construído para a população F_2 em estudo.

O mapeamento de intervalos múltiplos (MIM) para os fenótipos individuais foi capaz de detectar três QTLs associados ao comprimento de raiz (CR@8B_17,00; CR@15B_19,67 e CR@15B_28,00) (Tabela 3 e Figura 3). Juntos, esses três QTLs explicam 34,22% da variação do fenótipo de CR. Dos três QTLs identificados para CR, o QTL CR@15B_19,67 é o único que apresenta efeito aditivo estatisticamente significativo sobre o fenótipo avaliado individualmente. Para este QTL, o alelo do genitor FTPG12X contribui com o aumento de 2,531 desvios padrões da média da variável CR, o que corresponde a 16,15% da variação desse fenótipo. Os outros dois QTLs identificados para CR, CR@8B_17,00 e CR@15B_28,00, apresentam efeito de dominância estatisticamente significativo. Por se tratar de uma espécie autógama, o efeito aditivo representa a porção da variância genética que será transmitida para a próxima geração, estando diretamente relacionada com o ganho de seleção (Bernardo 2010). Dessa forma, a identificação de QTLs que

apresentam efeitos aditivos de maior magnitude refletem diretamente em uma maior eficiência de seleção para a característica de interesse. Como *P. brachyurus* causa lesões nas raízes das plantas, a identificação de QTLs associados à arquitetura e ao desenvolvimento das mesmas, tal como CR, é de fundamental importância para a seleção de genótipos que apresentam maior eficiência no controle desse nematoide.

O valor obtido de R^2 , no modelo MIM, foi inferior ao valor estimado de herdabilidade (h^2) para a variável CR (0,99) (Tabela 1). Não foram encontrados, na literatura, estudos de mapeamento de QTLs, em soja, associados à variável CR, utilizando a abordagem multivariada. Assim, não é possível comparar diretamente os valores de R^2 deste estudo com valores disponíveis na literatura. Entretanto, esses resultados estão em consonância com diversos estudos de mapeamento de QTLs envolvendo características associadas à arquitetura e ao desenvolvimento da raiz, utilizando os modelos de mapeamento IM (*Interval Mapping*), CIM (*Composite Interval Mapping*) e MQM (*Multiple QTL Mapping*) (Cui and Yu 2007; Liang et al. 2010, 2014; Zhou et al. 2011; Williams et al. 2012; Manavalan et al. 2015; Prince et al. 2015). Liang et al. (2010) identificaram 13 QTLs associados à arquitetura da raiz, em um conjunto de 106 RILs (*Recombinant Inbred Lines*), explicando de 10,9 a 18,7% da variação fenotípica das características avaliadas. Os dois QTLs identificados pelos autores, associados à variável CR, explicam 14,4% (GL B1) e 16,8% (GL D1a) da variação fenotípica observada. Williams et al. (2012) utilizaram um conjunto de 94 RILs e identificaram 12 QTLs associados a características relacionadas à raiz. Três dos 12 QTLs identificados afetam a variável CR, localizados no GL C2 (2 QTLs) e N, explicando de 17 a 33% da variação fenotípica observada. No GL A2 (Chr 8), o mesmo GL onde foi identificado um QTL para CR no presente trabalho, os autores também identificaram quatro QTLs associados às variáveis peso seco e fresco da raiz. Em um outro estudo utilizando 447 RILs, Liang et al. (2014) identificaram cinco QTLs, associados ao comprimento da raiz, localizados nos GL B1, L (2 QTLs), N e O, explicando de 7,05 a 13,18% da variação fenotípica observada.

Estudos de mapeamento de QTLs associados à arquitetura da raiz também têm sido conduzidos em populações interespecíficas, a partir do cruzamento entre *G. max* e *G. soja* (Manavalan et al. 2015; Prince et al. 2015). Manavalan et al. (2015) avaliaram 251 BILs (*Backcross Inbred Lines*) e identificaram cinco QTLs, associados a quatro caracteres relacionados à arquitetura da raiz, sendo eles: comprimento da raiz principal e das raízes secundárias, número de raízes secundárias e peso seco da raiz total. Os QTLs identificados localizam-se nos GL A2 (4 QTLs) e H, explicando de 7,77 a 20,77% da variação fenotípica das variáveis estudadas. Dentre esses QTLs, um QTL, localizado no GL A2, afeta a variável CR, explicando 12,18% da variação fenotípica observada, resultado esse semelhante ao obtido no presente estudo para esse GL. Um QTL para

cada uma das outras três variáveis analisadas também foram identificados nessa mesma região do GL A2, o que ressalta a relevância da mesma no controle de variáveis associadas à raiz. Prince et al. (2015) identificaram quatro QTLs, nos GL C2 (2 QTLs) e M (2 QTLs), explicando de 10 a 15% da variação fenotípica observada. O QTL associado à variável CR, localizado no GL C2, explica 11% da variação dos fenótipos estudados do conjunto de 160 RILs.

Outros estudos envolvendo mapeamento de QTLs, em soja, associados à arquitetura da raiz, identificaram QTLs explicando entre 4,8 e 17% (Cui and Yu 2007) e entre 10,5 e 26,1% (Zhou et al. 2011) da variação fenotípica das variáveis associadas à estrutura e desenvolvimento da raiz, o que corrobora com os resultados obtidos neste estudo. Não foram identificados QTLs, na literatura, associados à arquitetura da raiz, no GL E (Chr 15). Portanto, os resultados sugerem que dois novos QTLs, afetando a variável CR, foram identificados nesse GL E (Chr 15), para a população em estudo. Dessa forma, essas regiões que contemplam tais QTLs devem ser melhor investigadas, utilizando-se outras populações de mapeamento, para confirmar se esses QTLs, no GL E, não resultam de *background* genético específico. Por se tratar de uma característica quantitativa, os resultados obtidos corroboram com a hipótese de vários genes controlando a mesma, sendo que cada gene contribui para uma pequena porção do fenótipo.

Tabela 3. QTLs identificados para a variável comprimento de raiz (CR) utilizando-se a análise MIM. QTLs destacados em negrito foram identificados somente em tal abordagem. QTLs destacados em cinza correspondem a QTLs novos, ainda não descritos na literatura.

Nome QTL ¹	Marcadores Adjacentes ²		LOD ³	Efeito Aditivo (a) ⁴ Dominância (d) ⁵	% Var ⁶	R ^{2 7}
CR@8B_17,00	Gm08_11560988	Gm08_13459269	4,29	d = 2,276***	11,74	
CR@15B_19,67	Gm15_3160423	Gm15_3160423	6,26	a = -2,531*	16,15	34,22
CR@15B_28,00	Gm15_49318218	Gm15_50536934	2,34	d = -1,765*	6,33	

¹Fenótipo@*Chr_posição(cM)*; ²Marcadores adjacentes ao QTL identificado; ³LOD marginal significativo com base na distribuição de $\chi^2=0,835$ (5%) ⁴Efeito aditivo (a) medido como desvio em relação à média; ⁵Efeito de dominância; Significância dos efeitos baseada em *p*-valor estimado por reamostragem baseada na estatística *score* (**p*<0,05; ***p*<0,01; ****p*<0,001); ⁶Razão entre a variância do efeito do QTL sobre a variância fenotípica, vezes 100; ⁷R² calculado com o modelo contendo apenas os efeitos marginais significativos.

Figura 3. Perfis de LOD *score* ao longo dos GLs A2 (subgrupo 8B) e E (subgrupo 15B) utilizando-se a abordagem univariada (MIM) para o fenótipo comprimento de raiz (CR). O mapa genético estimado para os GL A2 e E é representado no eixo x, em cM, e os marcadores genéticos que os constituem estão ilustrados por triângulos de cor azul escuro. Os triângulos invertidos indicam a posição dos QTLs incluídos no modelo final selecionado

Os modelos de mapeamento multivariado (MT-MIM) utilizados para a análise dos dois grupos de variáveis (*Multi1* e *Multi2*) permitiram a identificação de três regiões genômicas com efeito em, pelo menos, uma das variáveis avaliadas (Tabela 4, Figuras 4 e 5). O modelo selecionado no grupo *Multi1* permitiu a identificação de duas posições genômicas, explicando 7,30% e 30,63% da variação fenotípica para FR e CR, respectivamente.

A região genômica identificada no GL K (*Multi1@9B_0,00*) apresentou efeitos estatisticamente significativos para as variáveis FR e CR. Para a variável FR os efeitos aditivo e de dominância foram significativos, com base na estatística *score*. No entanto, o efeito aditivo explicou uma maior percentagem da variação do fenótipo em estudo. O alelo do genitor FTPG12X é responsável pela redução de 0,468 desvios padrões da média da variável FR, o que corresponde a 4,54% da variação desse fenótipo. Embora não tenham sido observados indivíduos na população F₂ com FR menor do que 1, o genitor FTPG12X mostrou-se moderadamente resistente à reprodução de *P. brachyurus*, de acordo com a metodologia de Moura and Régis (1987). Considerando-se o fato de se tratar de uma característica complexa, provavelmente influenciada por um grande número de genes, pode-se sugerir que o QTL identificado no GL K provavelmente está associado a genes que interferem na reprodução desse nematoide, afetando diretamente a variável FR estudada. No caso da variável CR, o efeito de dominância foi estatisticamente significativo, representando 3,34% da variação do fenótipo. Não foram identificados, na literatura, QTLs afetando as variáveis FR e CR, no GL K. Como já discutido anteriormente, *P. brachyurus* afeta diretamente a raiz das plantas, e as variáveis CR e FR apresentaram correlação negativa. Portanto, os resultados sugerem que essa região identificada, afetando as duas variáveis, seja melhor investigada, considerando a relevância das mesmas na avaliação da resistência das cultivares a esse nematoide.

A posição genômica identificada no GL E (*Multi1@15B_20,00*) corresponde à mesma região genômica identificada no modelo de mapeamento univariado (CR@15B_19,67). No entanto, as estimativas de LOD *score* e do valor do efeito do QTL foram superiores na abordagem multivariada, corroborando a tendência do modelo MT-MIM apresentar maior poder estatístico na detecção de QTLs e sugerir efeitos mais precisos quando comparado ao modelo MIM.

A região genômica identificada no GL A2 (*Multi2@8A_175,37*) apresentou efeito aditivo estatisticamente significativo para a variável CA, com base na estatística *score*. O alelo do genitor FTPG12X contribui com o aumento de 0,660 desvios padrões da média da variável CA, representando 4,026% da variação desse fenótipo. Cinco QTLs, afetando a variável CA, encontram-se descritos no Soybase (www.soybase.org), nos GL C1, D2, F e L (2 QTLs), explicando entre 6,9 e 20,2% da variação fenotípica observada (Kuroda et al. 2013). Portanto, os

resultados sugerem que o QTL identificado neste estudo, afetando a variável CA, corresponde a um novo QTL ainda não descrito na literatura, no GL A2. A presença do nematoide *P. brachyurus* em tecidos radiculares afeta significativamente o desenvolvimento da parte aérea e da raiz, em soja (Figura 1). Portanto, além de ocasionar lesões radiculares, esses nematoides interferem na absorção de nutrientes e água, o que prejudica e afeta diretamente o crescimento e desenvolvimento das plantas (Jones and Fosu-Nyarko 2014; Rodrigues et al. 2014; Machado et al. 2015; Fosu-Nyarko and Jones 2016; Machado and Araújo Filho 2016).

Duas regiões com a presença de QTLs em pelo menos uma das características estudadas somente foram reveladas com o emprego da abordagem multivariada (*Multi1@9B_0,00* e *Multi2@8A_175,37*) (Tabela 4, Figuras 4 e 5). Essas duas regiões genômicas, que não foram identificadas nos modelos de fenótipos individuais, consistem em QTLs de menor efeito, explicando menos do que 5% da variação do fenótipo. Os resultados sugerem maior poder estatístico da abordagem multivariada em identificar QTLs de menor efeito em relação ao modelo MIM, no caso das variáveis estudadas. Estudos de mapeamento de QTLs associados à resistência a percevejos em soja (Möller 2017), à absorção de fósforo em milho (Gianotto 2015) e à variáveis que influenciam hipertensão e obesidade em ratos (Duarte 2007) também ressaltaram o maior poder estatístico da abordagem multivariada em relação à univariada. Para todos esses estudos, a utilização da abordagem de mapeamento multivariado também permitiu a identificação de QTLs de menor efeito, que não haviam sido detectados no modelo univariado. Além disso, os autores também relatam a obtenção de estimativas de LOD *score* superiores para a abordagem multivariada, no caso de QTLs identificados tanto no modelo univariado quanto multivariado.

Os modelos MT-MIM também apresentaram valores de R^2 inferiores aos valores de herdabilidade das variáveis analisadas. Nenhum trabalho foi encontrado na literatura utilizando a abordagem de mapeamento multivariado (MT-MIM), em soja, para essas características estudadas. Portanto, não será possível comparar diretamente os valores de R^2 obtidos com valores de R^2 disponíveis na literatura científica. No entanto, Möller (2017) utilizou a abordagem MT-MIM, em soja, para 15 variáveis agrônômicas e associadas à resistência aos percevejos e verificou que os valores de R^2 observados se aproximavam dos valores relatados em literatura com a estratégia CIM de mapeamento de QTLs.

Dois QTLs somente foram revelados com a abordagem univariada (*CR@8B_17,00* e *CR@15B_28,00*). No GL A2 (15B), apesar do perfil de LOD *score* no modelo multivariado apresentar valores mais altos do que os modelos univariados, o pico apresentado no final do GL não foi estatisticamente significativo, de acordo com o critério baseado na estatística *score*. De acordo com uma simulação realizada por Silva et al. (2012), os modelos MT-MIM podem

apresentar menor poder de detecção em regiões com efeito restrito a apenas um subconjunto das características avaliadas. Essa pode ser uma explicação aceitável para esses dois QTLs, revelados somente pelo modelo univariado, não terem sido detectados pelo modelo MT-MIM. Outra possível explicação, no caso do GL E, seria de que o modelo MT-MIM acomodou um outro QTL associado à CR, próximo à essa região, o que pode ter dificultado a detecção de mais um QTL neste mesmo GL, pela abordagem multivariada. De forma semelhante, Gianotto (2015) e Möller (2017) também relataram a identificação de QTLs que foram revelados somente na abordagem univariada, apesar do perfil de LOD *score* ter sido superior no modelo multivariado em relação aos modelos univariados. Os resultados obtidos neste trabalho ressaltam a importância da estratégia de se aliar o mapeamento univariado com o multivariado, permitindo a identificação de diferentes regiões genômicas controlando as variáveis em estudo.

Tabela 4. QTLs identificados usando a análise MT-MIM para as características do grupo *Multi1*: FR, PA, CA e CR e *Multi2*: PA, CA e PR. Os QTLs destacados em negrito foram identificados somente em tal abordagem. QTLs destacados em cinza correspondem a QTLs novos, ainda não descritos na literatura

Nome QTL ¹	Marcadores Adjacentes ²	LOD	FR		PA		CA		CR		PR	
			Efeito Aditivo ³	% Var ⁵	Efeito Aditivo ³	% Var ⁵	Efeito Aditivo ³	% Var ⁵	Efeito Aditivo ³	% Var ⁵	Efeito Aditivo ³	% Var ⁵
			Dominância ⁴		Dominância ⁴		Dominância ⁴		Dominância ⁴		Dominância ⁴	
Multi1@9B_0,00	Gm09_44855340 Gm09_44855340	8,90	-0,468**	4,54	0,045	-	-0,321	-	-0,131	3,34	N/A	
			0,567*	2,76	0,020		0,675		0,753**			
Multi1@15B_20,00	Gm15_3160423 Gm15_14361475	9,99	0,637	-	0,020	-	0,973	-	-3,465*	27,29	N/A	
			0,067		-0,052		-0,468		-0,966			
Multi2@8A_175,37	Gm08_42292068 Gm08_42292068	5,88	N/A		0,031	-	-0,660**			0,053		
					0,000		-0,445	4,03	N/A		0,107	-
R ² ⁶			7,30		4,03		30,63					

¹*Multi@Chr_posição(cM)*; ²Marcadores adjacentes ao QTL identificado; ³Efeito aditivo medido como desvio em relação à média; ⁴Efeito de dominância; Significância dos efeitos baseada em *p*-valor estimado por reamostragem baseada na estatística *score* (**p* < 0,05; ***p* < 0,01; ****p* < 0,001); ⁵Razão entre a variância do efeito do QTL sobre a variância fenotípica, vezes 100;

⁶R² calculado com o modelo contendo apenas os efeitos marginais significativos

Figura 4. Representação gráfica do perfil de LOD *score* obtido para as variáveis do grupo *Multi1* para o modelo de mapeamento univariado e multivariado, nos GLs K (9B) e E (15B). A cor azul clara representa o resultado obtido com o mapeamento multivariado. As demais cores estão indicadas na legenda e representam os resultados do mapeamento univariado para as características: FR: fator de reprodução; PA: peso fresco da parte aérea; CA: comprimento da parte aérea; e CR: comprimento da raiz. O mapa genético estimado para os GL K e E é representado no eixo x, em cM, e os marcadores genéticos que os constituem estão ilustrados por triângulos de cor azul escuro. Os triângulos invertidos indicam a posição dos QTLs incluídos no modelo final selecionado

Figura 5. Representação gráfica do perfil de LOD *score* obtido para as variáveis do grupo *Multi2* para o modelo de mapeamento univariado e multivariado. A cor azul clara representa o resultado obtido com o mapeamento multivariado. As demais cores estão indicadas na legenda e representam os resultados do mapeamento univariado para as características: PA: peso fresco da parte aérea; PR: peso fresco da raiz; e CA: comprimento da parte aérea. O mapa genético estimado para o GL A2 é representado no eixo x, em cM, e os marcadores genéticos que os constituem estão ilustrados por triângulos de cor azul escuro. Os triângulos invertidos indicam a posição dos QTLs incluídos no modelo final selecionado

Figura 6.

Nenhum QTL foi identificado, no presente estudo, afetando as variáveis PA e PR. Podemos verificar na Tabela 1 que essas variáveis foram as que apresentaram os menores valores de variância genética para essa população em estudo, correspondendo a 0,18 e 0,80 para PA e PR, respectivamente. Provavelmente, a variação genética obtida para essas variáveis, para a população em estudo, não foi suficiente para a identificação de QTLs. Outra questão relevante é o tamanho populacional utilizado, que pode não ter sido suficiente para identificar a variabilidade existente para cada característica analisada.

Para a variável FR, com base na estatística *score*, não há evidências de QTL no GL C2 (Chr 6A), utilizando-se o modelo univariado (Figura 6). No entanto, o pico obtido de LOD *score* se

aproxima de 3,45, que é superior ao pico de LOD *score* do QTL identificado para FR, no GL K (Chr 9B), na análise multivariada. Embora não possamos dizer que se trata de um QTL efetivamente, os resultados sugerem que essa seria uma região alvo importante para investigações posteriores, pois pode conter genes relacionados à resistência ao *P. brachyurus*, influenciando na reprodução desse nematoide.

Figura 7. Perfis de LOD *score* ao longo do GL C2 (subgrupo 6A) utilizando-se a abordagem univariada (MIM) para o fenótipo fator de reprodução (FR). O mapa genético estimado para o GL C2 é representado no eixo x, em cM, e os marcadores genéticos que os constituem estão ilustrados por triângulos de cor azul escuro. Os triângulos invertidos indicam a posição dos QTLs incluídos no modelo final selecionado

Uma investigação no banco de dados de QTLs, identificados para a cultura da soja, no Soybase (www.soybase.org), foi realizada nos GL A2, E e K para identificar QTLs já descritos conferindo resistência a outros nematoides ou doenças radiculares em soja. Diversos QTLs foram identificados, conferindo resistência ao nematoide de cisto (*Heterodera glycines*), no GL A2 (Concibido et al. 2004; Brucker et al. 2005; Guo et al. 2006a, b; Kazi et al. 2010; Vuong et al. 2010; Ferreira et al. 2011), no GL E (Chang et al. 2011; Jiao et al. 2015) e GL K (Winter et al. 2007; Wu et al. 2009). Também foram identificados QTLs associados à resistência à podridão vermelha da raiz (*Fusarium virguliforme*), no GL A2 (Abdelmajid et al. 2012; Swaminathan et al. 2016), no GL E (Ellis et al. 2012) e no GL K (Anderson et al. 2015; Swaminathan et al. 2016). Em relação à podridão

radicular de fitóftora (*Phytophthora sojae*), foram identificados QTLs afetando a resistência da soja a esse fungo que ataca o sistema radicular da planta, no GL A2 (Li et al. 2010; Wang et al. 2012) e GL E (Lee et al. 2013) e no GL K (Vuong et al. 2008). QTLs associados à resistência da soja ao mofo branco (*Sclerotinia sclerotiorum*) também foram identificados no GL A2 (Guo et al. 2008; Vuong et al. 2008) no GL E (Guo et al. 2008) e no GL K (Vuong et al. 2008; Manavalan et al. 2015). A presença desses QTLs associados à resistência a diversos patógenos que afetam tanto o sistema radicular como o desenvolvimento das plantas de soja, nos mesmos GL onde foram identificados QTLs, no presente estudo, (GL A2, K e E), ressalta a relevância dos resultados obtidos.

Em relação as análises de epistasia (interação) realizadas, os efeitos de interação entre pares de QTLs não foram estatisticamente significativos a 1% de probabilidade de erro, com base na estatística *score*, para ambos os procedimentos de busca. Esses resultados sugerem que, neste estudo, efeitos epistáticos não explicaram proporção genética da variação fenotípica para todas as características, tanto para o MIM quanto o MT-MIM.

4.2. Análise das sequências candidatas em regiões de interesse

Neste trabalho, 2.563 sequências transcritas candidatas foram investigadas considerando-se sete regiões genômicas de interesse para soja (Tabela 5). Tais regiões foram selecionadas a partir dos cinco QTLs identificados, utilizando-se os modelos MIM e MT-MIM, e a posição alvo com um suposto “QTL”, no GL C2. Para o QTL identificado no GL A2 (8B) foram consideradas duas regiões alvo de interesse (Região 3 e 4), pois os dois marcadores flanqueando o QTL apresentaram posição muito distante no genoma de referência da soja (ConsensusMap 4.0). Dentre tais regiões, duas (Região 1 e 5) estiveram diretamente relacionadas à resistência ao *P. brachyurum*, com base na característica FR e, dessa forma, foram selecionadas para uma discussão mais detalhada das sequências candidatas. Para as demais regiões identificadas neste estudo (Região 2, 3, 4, 6 e 7), associadas às variáveis CR e CA, as sequências candidatas identificadas encontram-se descritas nos Apêndices A a E.

No caso das Regiões 1 e 5, grande parte das sequências candidatas apresentou alinhamento satisfatório contra as sequências de proteínas do banco de dados não-redundante do NCBI, algo que também foi observado para as demais regiões. No caso da região 1, 390 (67,01%) sequências candidatas puderam ser satisfatoriamente alinhadas, das quais 320 (82,05%) foram posteriormente mapeadas e anotadas para os termos de GO, dentro do Blast2GO. Para a região 5, 247 (74,18%) sequências candidatas puderam ser alinhadas, das quais 208 (84,21%) foram mapeadas e anotadas para tais termos.

Tabela 5. Alinhamento, mapeamento e anotação das sequências transcritas candidatas, selecionadas em regiões genômicas de interesse, utilizando-se a plataforma *Blast2GO*.

Região	GL (Cromossomo)	Intervalo (pb)	Sequências Candidatas	Alinhadas, Mapeadas e Anotadas	Alinhadas e Mapeadas	Alinhadas	Não Alinhadas
1 ¹	C2 (6A)	[46.177.140; 49.082.951]	582	320	3	67	192
2	A2 (8A)	[41.792.068; 42.792.068]	262	131	7	34	90
3	A2 (8B)	[12.959.269; 13.959.269]	301	165	5	29	102
4	A2 (8B)	[39.566.964; 40.566.964]	147	91	2	20	34
5	K (9B)	[44.098.626; 45.098.626]	333	208	7	32	86
6	E (15B)	E (15B)	390	245	7	21	117
7	E (15B)	E (15B)	548	295	3	73	177
Total	–	–	2.563	1.455	34	276	798

¹Região genômica alvo com a possível presença de genes relacionados à resistência ao *Pratylenchus bachiurus*, embora não tenha havido evidência de QTL para FR, com base nas abordagens MIM e MT-MIM.

Uma análise do alinhamento realizado mostrou, para as regiões 1 e 5, que várias sequências candidatas estiveram relacionadas a componentes importantes na resposta das plantas a estresses bióticos (Tabelas 6 e 7). Estes componentes correspondem a fatores de transcrição de famílias conhecidas, enzimas que atuam em processos bioquímicos de certos hormônios e proteínas com domínios característicos, que são conhecidas por atuarem, especificamente, nos mecanismos de defesa das plantas contra agentes patogênicos.

No total, quatro famílias de fatores de transcrição foram aqui observadas: i) *Ethylene-Responsive-Element-Binding Factors* (ERF); ii) *Basic-Domain Leucine-Zipper* (bZIP); iii) WRKY; e iv) MYB. Os fatores de transcrição ERF, que são assim chamados por apresentarem um domínio conservado ERF, podem se ligar a promotores de genes relacionados à resistência e ativar a produção de etileno (ET) (Singh et al. 2002), que é um dos hormônios fundamentais nos mecanismos de resposta induzida das plantas ao ataque de patógenos (Singh et al. 2002; Loake and Grant 2007; Balbi and Devoto 2008; Grant and Jones 2009; Pieterse et al. 2009; Nahar et al. 2011; Robert-Seilantantz et al. 2011; Beneventi et al. 2013; Denancé et al. 2013). Os fatores de transcrição com domínio bZIP também atuam em defesa das plantas a estresses bióticos (Alves et al. 2013; Llorca et al. 2014), e embora os membros mais conhecidos de tal família estejam presentes em *Arabidopsis* (Singh et al. 2002), estes fatores já foram observados em outras espécies de plantas, tais

como arroz, sorgo, soja e milho (Llorca et al. 2014). Os fatores de transcrição WRKY, que contêm um ou dois domínios WRKY e uma região com motivo zinc-finger, mostram maior expressão e/ou atividade de ligação ao DNA com a indução de diversos agentes patogênicos e sinais de defesa das plantas, desempenhando um importante papel nos mecanismos de resistência (Eulgem et al. 2000; Singh et al. 2002). E, por fim, os fatores de transcrição MYB também desempenham importante papel na resposta das plantas a estresses bióticos (Ambawat et al. 2013).

A enzima *Jasmonate O-Methyltransferase*, descrita para uma das sequências candidatas da região 5, parece estar diretamente envolvida em via de sinalização do hormônio ácido jasmônico ou jasmonato (JA), que é um importante e conhecido indutor de defesa em plantas (Wasternack 2007; Bari and Jones 2009; Nahar et al. 2011; Robert-Seilanianz et al. 2011; Carvalhais et al. 2013; Yan et al. 2013; Yan and Xie 2015; Larrieu and Vernoux 2016). A biossíntese do JA ocorre a partir do chamado ácido linolênico, podendo seguir o caminho das composições ativa (Jasmonato-Isoleucina) ou volatilizada (Metil Jasmonato). Diversos estudos mostram que concentrações de JA se elevam nos locais do dano em resposta à infecção do patógeno ou ferimento, e aplicações exógenas deste hormônio induzem a expressão de genes relacionados aos mecanismos de defesa (Wasternack 2007; Grant and Jones 2009; Santner et al. 2009; Coerini 2014).

O hormônio JA, juntamente com os hormônios ET e ácido salicílico (SA), desempenham papel fundamental nos mecanismos de defesa induzido das plantas (Loake and Grant 2007; Balbi and Devoto 2008; Bari and Jones 2009; Pieterse et al. 2009; Nahar et al. 2011). A ação destes hormônios é bastante conhecida no processo de defesa das plantas contra patógenos, sendo que, em geral, JA e ET estão normalmente associados à defesa contra os agentes necrotróficos, que desenvolvem-se em células mortas do hospedeiro, e SA está normalmente associado à defesa contra os agentes biotróficos, que dependem das células vivas do hospedeiro (Glazebrook 2005).

Outros hormônios conhecidos que têm se mostrado envolvidos com a defesa das plantas a estresses bióticos são a auxina (AUX), o ácido abscísico (ABA) e a giberelina (GA). Embora o papel de cada um deles ainda não seja bem conhecido neste contexto (Bari and Jones 2009), duas sequências transcritas candidatas (PUT-169a-Glycine_max-115652 e PUT-169a-Glycine_max-172245), presentes na região 1, apresentaram similaridades acima de 90% com sequências relacionadas à AUX. Beneventi et al. (2013) verificaram que os hormônios AUX e GA desempenharam papel crucial na defesa de plantas de soja ao ataque do nematoide do gênero *Meloidogyne*, mostrando a importância dos resultados aqui obtidos envolvendo o hormônio AUX.

Uma das sequências transcritas candidatas (PUT-169a-Glycine_max-213493), presente na região 1 (cromossomo 6A), apresentou similaridade acima de 80% com os domínios TIR, NBS e LRR, que estão diretamente relacionados com os mecanismos de defesa. Em plantas, a grande

maioria das proteínas codificadas por genes de resistência a doenças (genes R) apresenta repetições em tandem ricas em leucina (*Leucine-Rich Repeats*, LRR), que possuem papel decisivo no reconhecimento específico de patógenos (Jones 2001; McDowell and Woffenden 2003). Essas repetições correspondem a um motivo composto por aproximadamente 24 aminoácidos em comprimento (Kobe and Deisenhofer 1994; Bent 1996). Possuem leucinas ou outros resíduos hidrofóbicos em intervalos regulares e, também, podem conter prolinas e asparaginas regularmente espaçadas em sua estrutura (Bent 1996). No entanto, os membros da família de proteínas R, mais extensamente representada, também possuem um sítio de ligação nucleotídica (*Nucleotide-Binding Site*, NBS), sendo que alguns deles também podem conter um domínio TIR, que apresenta homologia com uma superfamília de receptores de imunidade (McDowell and Woffenden 2003).

Tabela 6. Alinhamento e anotação de sequências candidatas selecionadas para a região 1, no GL C2 (6A).

Sequência Candidata	Descrição do Alinhamento	Comprimento (pb)	Valor Esperado	Similaridade (%)	Termos <i>GO</i>
PUT-169a-Glycine_max-115652	<i>Auxin-induced 6B-like</i>	495	9,78E ⁻⁶²	94,00	0009733
PUT-169a-Glycine_max-172245	<i>Auxin-induced X10.A</i>	589	9,70E ⁻⁶⁹	93,25	0009734
PUT-169a-Glycine_max-72191	<i>bZIP Family Transcription Factor</i>	388	9,34E ⁻⁷⁵	82,30	0003700, 0006355, 0043565
PUT-169a-Glycine_max-213493	<i>Disease Resistance (TIR-NBS-LRR)</i>	413	1,43E ⁻⁸⁸	83,05	0007165, 0043531
PUT-169a-Glycine_max-49030	<i>Ethylene-responsive Transcription Factor 9-like (ERF)</i>	980	1,39E ⁻¹⁴⁹	69,90	0003677, 0003700, 0005634, 0006351, 0006355
PUT-169a-Glycine_max-252603	<i>Peroxidase 3</i>	1327	0,0	89,15	0004601, 0005576, 0006979, 0009505, 0009664, 0020037, 0042744, 0046872, 0055114, 0098869
PUT-169a-Glycine_max-96042	<i>MYB-like Transcription Factor ETC1</i>	1246	1,35E ⁻¹¹	87,25	0003677
PUT-169a-Glycine_max-87593	<i>Homeobox-Leucine Zipper HDG2-like Isoform X1</i>	243	7,83E ⁻⁴¹	98,90	0005634, 0006355, 0008289, 0043565
PUT-169a-Glycine_max-211259	<i>Probable WRKY Transcription Factor 31</i>	604	7,27E ⁻¹³⁵	72,30	0003700, 0006355, 0043565

Duas sequências transcritas candidatas (PUT-169a-Glycine_max-87593 e PUT-169a-Glycine_max-218747), presentes nas regiões 1 (cromossomo 6A) e 5 (cromossomo 9B), apresentaram similaridade de 98,90 e 89,0%, respectivamente, com estruturas que possuem sequências espaçadas de leucina (*Leucine Zipper*, LZ). Tais sequências podem fazer parte da estrutura das proteínas que contêm os domínios LRR e NBS, permanecendo entre o N-terminal e o domínio NBS (Bent 1996). A presença de LZ em proteínas que contêm os domínios LRR e NBS

foi constatada em *Arabidopsis* (Bent et al. 1994; Mindrinos et al. 1994; Grant et al. 1995) e tomate (Salmeron et al. 1996).

Tabela 7. Alinhamento e anotação de sequências candidatas selecionadas para a região 5, no GL K (9B).

Sequência Candidata	Descrição do Alinhamento	Comprimento (pb)	Valor Esperado	Similaridade (%)	Termos <i>GO</i>
PUT-169a-Glycine_max-257862	Jasmonate O-Methyltransferase-like	641	2,90E ⁻¹⁵²	86,10	0030795, 0032259
PUT-169a-Glycine_max-42234	WRKY Transcription Factor	1314	0,0	82,35	0003700, 0006355, 0043565
PUT-169a-Glycine_max-155093	1-Aminocyclopropane-1-carboxylate synthase	637	1,92E ⁻¹⁴⁴	94,65	0016847, 0030170, 0042218
PUT-169a-Glycine_max-68448	Peroxisomal Membrane 11C	1100	1,94E ⁻¹⁶⁷	95,45	0005779, 0016559, 0044375
PUT-169a-Glycine_max-207314	Probable Methyltransferase PMT13	308	4,21E ⁻⁹	97,20	0005768, 0005774, 0005802, 0016021, 0032259
PUT-169a-Glycine_max-218747	Homeobox-Leucine Zipper	221	9,69E ⁻¹¹	89,0	0005634, 0006355, 0008289, 0043565

5. CONSIDERAÇÕES FINAIS

A utilização da abordagem multivariada no melhoramento de plantas tem se mostrado uma ferramenta promissora, no caso de mapeamento de QTLs para características complexas e que apresentam correlação entre si. Os resultados obtidos neste trabalho sugerem que a utilização das duas abordagens, MIM univariada e multivariada, foram satisfatórias e complementares na identificação de QTLs. A abordagem MT-MIM apresentou maior precisão das estimativas dos parâmetros genéticos e na localização dos QTLs, em relação às análises de fenótipos individuais. No entanto, a abordagem MIM permitiu revelar alguns QTLs particulares para cada variável analisada. Dessa forma, ambas as abordagens contribuíram para auxiliar no entendimento da arquitetura genética das variáveis estudadas.

A análise univariada permitiu a identificação de três QTLs associados à variável CR, sendo que um desses QTLs (CR@15B_19,67), identificado no GL E, apresentou efeito aditivo estatisticamente significativo sobre o fenótipo avaliado individualmente. Os resultados sugerem que esse QTL deve ser melhor investigado, para sua validação em diferentes *backgrounds* genéticos e ambientes. A comparação dos QTLs identificados com o banco de QTLs do Soybase permitiu verificar que diversos outros QTLs, associados à variável CR, foram já descritos em outros GLs da soja. No entanto, nenhum QTL associado a essa variável foi identificado no GL E, o que sugere que um novo QTL foi identificado para CR nessa população em estudo. Por se tratar de uma característica quantitativa, os resultados obtidos corroboram com a hipótese de vários genes controlando a mesma, sendo que cada gene contribui para uma pequena porção da variação do fenótipo. Para a população em estudo, o QTL identificado explica 16,35% variação fenotípica observada, sendo um valor significativamente relevante para sua utilização em futuros estudos de seleção genômica.

A análise multivariada permitiu a identificação de três posições genômicas com efeito em, pelo menos, uma das variáveis avaliadas sendo uma das regiões coincidente com o QTL identificado no GL E, na análise univariada. A região genômica identificada no GL K (Multi1@9B_0,00) apresentou efeitos estatisticamente significativos para as variáveis FR e CR, explicando 7,30% e 3,34% da variação fenotípica observada para essas variáveis, respectivamente. Considerando-se que se tratam de características complexas, provavelmente influenciadas por um grande número de genes, pode-se sugerir que a região genômica identificada no GL K provavelmente está associada a genes que interferem tanto na reprodução do nematoide como no desenvolvimento da raiz em soja. Não foram identificados na literatura QTLs afetando as variáveis FR e CR, no GL K. Portanto, os resultados sugerem que essa região identificada, afetando as duas

variáveis, seja melhor investigada, considerando a relevância das mesmas na avaliação da resistência das cultivares ao nematoide-das-lesões-radiculares.

A região genômica identificada no GL A2 (*Multi2@8A_175,37*) apresentou efeito aditivo estatisticamente significativo para a variável CA, com base na estatística *score*, sendo responsável por 4,03% da variação fenotípica para essa variável, na população em estudo. Apesar de o banco de dados de QTLs em soja apresentar cinco QTLs associados à variável CA, nenhum QTL foi descrito no GL A2, sugerindo se tratar de um novo QTL. Embora o nematoide *P. brachyurus* interfira diretamente no desenvolvimento da raiz, o crescimento da parte aérea da planta também é comprometido pelo desbalanceamento da disponibilidade de água e nutrientes. Assim, a região identificada no GL A2 também deve ser considerada para futuras investigações considerando a relevância dessa variável na identificação de genótipos resistentes.

Como a população de mapeamento utilizada no presente estudo refere-se a geração F_2 , avaliada em um único ambiente, os QTLs e regiões genômicas identificados no presente estudo devem ser melhor investigados e avaliados em outras populações de mapeamento ou mesmo em painéis de mapeamento associativo, para a validação dos mesmos em outros *backgrounds* genéticos e ambientes. Esses QTLs candidatos poderiam ser utilizados em estudos de seleção genômica, em soja, em uma etapa de pré-melhoramento, para a escolha de genitores com maior número de alelos favoráveis para as características de interesse. Estudos tem demonstrado uma maior acurácia em modelos de seleção genômica com a utilização de QTLs, previamente identificados em estudos de mapeamento, como efeito fixo, juntamente com os demais marcadores como efeito aleatório (Bernardo 2014; Spindel et al. 2015; Arruda et al. 2016). Após a identificação e validação dos SNPs é realizada a etapa de associação à característica de interesse. Dessa forma, pode-se criar um banco de dados genotípicos do germoplasma disponível, com base nos haplótipos identificados, para a escolha dos genitores de interesse de acordo com às características desejadas.

A análise das sequências transcritas candidatas, associadas às regiões de resistência ao *P. brachyurus*, mostrou uma alta relação a componentes importantes na resposta das plantas a estresses bióticos. As quatro famílias de fatores de transcrição observadas: i) Ethylene-Responsive-Element-Binding Factors (ERF); ii) Basic-Domain Leucine-Zipper (bZIP); iii) WRKY; e iv) MYB já foram descritas em estudos de expressão diferencial para diferentes espécies de plantas, associadas aos mecanismos de defesa das mesmas a patógenos.

O mapeamento de QTLs permite a identificação da região genômica que contem genes de interesse, mas geralmente o intervalo de confiança para essa região é constituída por milhares de pares de bases, podendo conter muitos genes. Dessa forma, a comparação das sequências transcritas das regiões de interesse com anotações gênicas é uma ferramenta que auxilia no

direcionamento e seleção de genes associados à característica de interesse. Os resultados obtidos neste trabalho têm o potencial de ajudar na compreensão da arquitetura genética dos caracteres fenotípicos avaliados e levantam hipóteses interessantes para ser testadas por meio de experimentos posteriores de estudo de expressão gênica.

REFERÊNCIAS

- Abdelmajid KM, Ramos L, Leandro L, et al (2012) The “ PI 438489B ” by “ Hamilton ” SNP-Based Genetic Linkage Map of Soybean [*Glycine max* (L .) Merr .] Identified Quantitative Trait Loci that Underlie Seedling SDS Resistance. J Plant Genome Sci 1:18–30. doi: 10.5147/jpgs.2012.0053
- Abreu ABL, Agnes DCA (2015) Nematoides: amostragem, armazenamento e transporte de amostras. In: Safra 2014/2015.
- Abreu ABL, Agnes DCA, Borges EP, et al (2016) Incidência de nematóides em amostras de soja na safra 2015/16 enviadas a Fundação Chapadão. In: XXXV Reunião de Pesquisa de Soja. Anais da XXXV Reunião de Pesquisa de Soja, Londrina, pp 158–159
- Abreu ABL, Agnes DCA, Borges EP, et al (2014) Incidência de nematoides na região dos Chapadões. In: 47 Congresso Brasileiro de Fitopatologia. Anais do 47 Congresso Brasileiro de Fitopatologia, Londrina, pp 18–23
- Agnes DCA, Abreu ABL, Borges EP, et al (2014) Incidência de nematoides em amostras de soja na safra 2013/14 enviadas à Fundação Chapadão. In: XXXIV Reunião de Pesquisa de soja. Embrapa, Londrina, pp 245–247
- Agrolink (2015) Por ano, nematoides causam prejuízos de R\$ 35 bilhões ao agronegócio nacional. https://www.agrolink.com.br/noticias/por-ano--nematoides-causam-prejuizos-de-r--35-bilhoes-ao-agronegocio-nacional_343212.html. Accessed 1 Jan 2015
- Akond M, Liu S, Schoener L, et al (2013) A SNP-Based Genetic Linkage Map of Soybean Using the SoyS - NP6K Illumina Infinium BeadChip Genotyping Array. Plant Genet Genomics, Biotechnol 1:80–89. doi: 10.5147/jpgs.2013.0090
- Al-Banna L, Ploeg AT, Williamson VM, Kaloshian I (2004) Discrimination of six *Pratylenchus* species using PCR and species-specific primers. J Nematol 36:142–146.
- Al-Banna L, Williamson VM, Gardner SL (1997) Phylogenetic analysis of nematodes of the genus *Pratylenchus* using nuclear 26 S rDNA. Mol Phylogenet Evol 7:94–102.
- Alves GJS, Raphaelli N, Fangueiro R (2006) Desenvolvimento sustentável na indústria têxtil: estudo de propriedades e características de malhas produzidas com fibras biodegradáveis. XXII CNTT Congr Nac Técnicos Têxteis - VIII Fenatêxtil 25–30.
- Alves MS, Dadalto SP, Gonçalves AB, et al (2013) Plant bZIP transcription factors responsive to pathogens: A review. Int J Mol Sci 14:7815–7828. doi: 10.3390/ijms14047815
- Alves TCU (2008) Reação de cultivares de soja ao nematoide das lesões radiculares *Pratylenchus brachyurus*. Universidade Federal de Mato Grosso, Faculdade de Agronomia e Medicina Veterinária, Dissertação de Mestrado
- Alzate-Marin AL, Cervigni GDL, Moreira MA, Barros EG (2005) Seleção assistida por marcadores moleculares visando ao desenvolvimento de plantas resistentes a doenças, com ênfase em feijoeiro e soja. Fitopatol Bras 30:333–342. doi: 10.1590/S0100-41582005000400001
- Ambawat S, Sharma P, Yadav NR, Yadav RC (2013) MYB transcription factor genes as regulators for plant responses: An overview. Physiol Mol Biol Plants 19:307–321. doi: 10.1007/s12298-013-0179-1
- Anderson J, Akond M, Kassem MA, et al (2015) Quantitative trait loci underlying resistance to sudden death syndrome (SDS) in MD96-5722 by “Spencer” recombinant inbred line population of soybean. 3 Biotech 5:203–210. doi: 10.1007/s13205-014-0211-3

- Antonio S, Mendes F, Franchini J, et al (2012) Perdas de produtividade da soja em área infestada por nematoide das lesões radiculares em Vera-MT. In: 6 Congresso brasileiro de soja. Anais do 6 Congresso brasileiro de soja, Cuiabá, pp 125–129
- Arruda MP, Lipka AE, Brown PJ, et al (2016) Comparing genomic selection and marker-assisted selection for *Fusarium head* blight resistance in wheat (*Triticum aestivum* L.). Mol Breed. doi: 10.1007/s11032-016-0508-5
- Asmus GL (2004) Ocorrência de nematóides fitoparasitos em algodoeiro no Estado de Mato Grosso do Sul. Nematol Bras 28:77–86.
- Back MA, Haydock PPJ, Jenkinson P (2002) Disease complexes involving plant parasitic nematodes and soilborne pathogens. Plant Pathol 51:683–697. doi: 10.1046/j.1365-3059.2002.00785.x
- Balbi V, Devoto A (2008) Jasmonate signalling network in *Arabidopsis thaliana*: Crucial regulatory nodes and new physiological scenarios. New Phytol 177:301–318. doi: 10.1111/j.1469-8137.2007.02292.x
- Bari R, Jones JDG (2009) Role of plant hormones in plant defence responses. Plant Mol Biol 69:473–488. doi: 10.1007/s11103-008-9435-0
- Bektas Y, Eulgem T (2015) Synthetic plant defense elicitors. Front Plant Sci 5:1–17. doi: 10.3389/fpls.2014.00804
- Beneventi M, da Silva O, de Sá ME, et al (2013) Transcription profile of soybean-root-knot nematode interaction reveals a key role of phytohormones in the resistance reaction. BMC Genomics 14:322. doi: 10.1186/1471-2164-14-322
- Bent AF (1996) Plant Disease Resistance Genes: Function Meets Structure. Plant Cell 8:1757–1771. doi: 10.1105/tpc.8.10.1757
- Bent AF, Kunkel BN, Dahlbeck D, et al (1994) RPS2 of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science (80-) 265:1856–1860.
- Bernardo R (2010) Breeding for quantitative traits in plants. Stemma Press, Woodbury
- Bernardo R (2014) Genomewide selection when major genes are known. Crop Sci 54:68–75. doi: doi:10.2135/cropsci2013.05.0315
- Berry S, Fargette M, Vaughan W, et al (2008) Detection and quantification of root-knot nematode (*Meloidogyne javanica*), lesion nematode (*Pratylenchus zeae*) and dagger nematode (*Xiphinema elongatum*) parasites of sugarcane using real-time PCR. Mol Cell Probes 22:168–176.
- Bert W, Leliaerte F, Vierstraete A, et al (2008) Molecular phylogeny of the *Tylenchina* and evolution of the female gonoduct (Nematoda: *Rhabditida*). Mol Phylogenet Evol 48:728–744.
- Boerma HR, Hussey RS (1992) Breeding plants for resistance to nematodes. J Nematol 24:242–252.
- Boerma HR, Walker DR (2005) Discovery and utilization of QTLs for insect resistance in soybean. Genetica 123:181–189.
- Borém A, Caixeta ET (2009) Marcadores Moleculares, 2ed. Universidade Federal de Viçosa, Brasília - DF
- Brucker E, Carlson S, Wright E, et al (2005) Rhg1 alleles from soybean PI 437654 and PI 88788 respond differentially to isolates of *Heterodera glycines* in the greenhouse. Theor Appl Genet 111:44–49.
- Câmara GM (1998) Origem, difusão geográfica e importância da soja. Piracicaba
- Canto-Saenz M (1985) The nature of resistance to *Meloidogyne incognita*. In: An advanced treaties on *Meloidogyne*. North Caroline State University, Raleigh, pp 225–231
- Carvalhais LC, Dennis PG, Badri D V., et al (2013) Activation of the Jasmonic Acid Plant Defence Pathway Alters the Composition of Rhizosphere Bacterial Communities. PLoS One 8:1–5. doi: 10.1371/journal.pone.0056457

- Castillo P, Vovlas N (2007) *Pratylenchus* (Nematoda: Pratylenchidae): diagnosis, biology, pathogenicity and management, 6th edn. Brills, London
- Chang W, Dong L, Wang Z, et al (2011) QTL underlying resistance to two HG types of *Heterodera glycines* found in soybean cultivar “L-10”. BMC Genomics 12:233. doi: 10.1186/1471-2164-12-233
- Chen L, Storey JD (2006) Relaxed significance criteria for linkage analysis. Genetics 173:2371–2381. doi: 10.1534/genetics.105.052506
- Choi I-Y, Hyten DL, Matukumalli LK, et al (2007) A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. Genetics 176:685–96. doi: 10.1534/genetics.107.070821
- Churchill G, Doerge R (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971.
- Coelho A (2000) Considerações gerais sobre a análise de QTL's. In: Análise de QTL no melhoramento de plantas. Funape, Goiânia, pp 1–36
- Coerini LF (2014) Expressão de genes das vias de jasmonato e etileno na resposta de plantas de citros às bactérias Candidatus *Liberibacter* spp., causadoras do Huanglongbing. Universidade Estadual de São Paulo - Escola Superior de Agricultura Luiz de Queiroz, Dissertação de Mestrado
- CONAB (2017) Companhia Nacional de Abastecimento. <http://www.conab.gov.br>. Accessed 10 Jul 2017
- Concibido VC, Diers BW, Arelli PR (2004) A decade of QTL mapping for cyst nematode resistance in soybean. 1131:1121–1131.
- Conesa A, Götz S (2008) Blast2GO: A comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics. doi: 10.1155/2008/619832
- Conesa A, Götz S, García-Gomez JM, et al (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676.
- Contreras-Soto RI, Mora F, De Oliveira MAR, et al (2017) A genome-wide association study for agronomic traits in soybean using SNP markers and SNP-Based haplotype analysis. PLoS One 12:1–22. doi: 10.1371/journal.pone.0171105
- Coolen WA, D'Herdt CJ (1972) A method for the quantitative extraction of nematodes from plant tissue. State Nematology and Entomology Research Station, Ghent
- Costa D da C, Ferraz S (1998) Avaliação da resistência de cultivares e linhagens de soja a *Pratylenchus brachyurus*. An da Esc Agron e Veterinária 28:67–76.
- Costa Neto PR, Rossi LFS, Zagonel GF, Ramos LP (2000) Produção de biocombustível alternativo ao óleo diesel através da transesterificação de óleo de soja usado em frituras. Quim Nova 23:531–537. doi: 10.1590/S0100-40422000000400017
- Cregan PB, Jarvik T, Bush a. L, et al (1999) An Integrated Genetic Linkage Map of the Soybean Genome. Crop Sci 39:1464. doi: 10.2135/cropsci1999.3951464x
- Cui S, Yu D (2007) QTL mapping of phosphorus deficiency tolerance in soybean (*Glycine max* L.) during seedling stage. Acta Agron Sin 33:378–383. doi: 10.3321/j.issn
- Debiasi H, Moraes M, Franchini J, et al (2011) Monitoramento da fertilidade do solo e da ocorrência do nematoide das lesões radiculares em soja no Mato Grosso. In: Congresso brasileiro de ciência do solo. Anais eletrônicos do congresso brasileiro de ciência do solo, Uberlândia, pp 122–127
- Denancé N, Sánchez-Vallet A, Goffner D, Molina A (2013) Disease resistance or growth: the role of plant hormones in balancing immune responses and fitness costs. Front Plant Sci 4:1–12. doi: 10.3389/fpls.2013.00155

- Dias W, Silva J, Garcia A, Carneiro G (2007) Nematóides de importância para a soja no Brasil. Bol Pesqui da soja 11:173–183.
- Dias WP, Soja E, Grosso EM, Carneiro GE de S (2010) Nematóides em Soja: Identificação e Controle 76. 1–8.
- Dodds PN, Rathjen JP (2010) Plant immunity: towards an integrated view of plant–pathogen interactions. Nat Rev Genet 11:539–548. doi: 10.1038/nrg2812
- Doerge RW (2002) Multifactorial Genetics mapping and Analysis of Quantitative Trait Loci in Experimental Populations. Nat Rev Genet 3:43–52. doi: 10.1038/nrg703
- Dong Q, Schlueter SD, Brendel V (2004) PlantGDB, plant genome database and analysis tools. Nucleic Acids Res 32:D354–9. doi: 10.1093/nar/gkh046
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus (Madison) 12:13–15.
- Duarte N (2007) Análise multivariada no mapeamento genético de traços quantitativos. Universidade de São Paulo - USP, Dissertação de Mestrado
- Duvick J, Fu A, Muppirala U, et al (2008) PlantGDB: A resource for comparative plant genomics. Nucleic Acids Res 36:959–965.
- Ellis ML, Wang H, Paul PA, et al (2012) Identification of soybean genotypes resistant to *Fusarium graminearum* and genetic mapping of resistance quantitative trait loci in the cultivar Conrad. Crop Sci 52:2224–2233. doi: 10.2135/cropsci2011.11.0624
- Elshire RJ, Glaubitz JC, Sun Q, et al (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6:e19379. doi: 10.1371/journal.pone.0019379
- Embrapa (2013) Tecnologias de produção de soja – Região central do Brasil 2014, 1ed. Embrapa Soja, Londrina
- Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. Trends Plant Sci 5:199–206. doi: 10.5423/PPJ.2010.26.2.110
- Falconer D, Mackay T (1996) Introduction to quantitative genetics. Edinburgh, Longman
- Fan J-B, Gunderson K, Bibikova M, et al (2006) Illumina universal bead arrays. Methods Enzymol 410:57–73.
- Fan J-B, Oliphant A, Shen R, et al (2003) Highly parallel SNP genotyping. Cold Spring Harb Symp Quant Biol 68:69–78. doi: 10.1101/sqb.2003.68.69
- Fang C, Ma Y, Wu S, et al (2017) Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. Genome Biol 18:161. doi: 10.1186/s13059-017-1289-9
- Ferraz L (2006) O nematóide *Pratylenchus brachyurus* e a soja sob plantio direto. Rev Plantio Direto 95:23–27.
- Ferraz L (1999) Gênero *Pratylenchus* – os nematoides das lesões radiculares. Revisão Anu Patol plantas 7:57–195.
- Ferraz LCCB (1995) Interações entre *Pratylenchus brachyurus* e *Meloidogyne javanica* em soja. Sci Agric 52:305–309.
- Ferreira ME, Grattapaglia D (1998) Introdução ao uso de marcadores moleculares em análises genéticas, 3.ed. Embrapa-CENARGEN, Brasília - DF
- Ferreira MF da S, Cervigni GDL, Ferreira A, et al (2011) QTLs for resistance to soybean cyst nematode, races 3, 9, and 14 in cultivar Hartwig. Pesqui Agropecu Bras 46:420–428. doi: 10.1590/S0100-204X2011000400012
- Fosu-Nyarko J, Jones MGK (2016) Advances in Understanding the Molecular Mechanisms of Root Lesion Nematode Host Interactions. Annu Rev Phytopathol 54:253–278. doi: 10.1146/annurev-phyto-080615-100257
- Gianotto AC (2015) Desenvolvimento de modelos de causalidade com informações de QTLs para estudo do relacionamento de caracteres fenotípicos relativos à absorção de fósforo em milho. Universidade Estadual de São Paulo, Tese de Doutorado

- Gillet F-X, Bournaud C, de Souza Júnior JDA, Fatima Grossi-de-Sa M (2017) Plant-parasitic nematodes: towards understanding molecular players in stress responses. *Ann Bot* mcw260. doi: 10.1093/aob/mcw260
- Glazebrook J (2005) Contrasting Mechanisms of Defense Against Biotrophic and Necrotrophic Pathogens. *Annu Rev Phytopathol* 43:205–227. doi: 10.1146/annurev.phyto.43.040204.135923
- Gonzaga V, Santos JM dos (2012) Chave ilustrada para a identificação das seis espécies de *Pratylenchus* mais comuns no Brasil. <http://nematologia.com.br/wp-content/uploads/2012/08/chavigo.pdf>.
- Goodwin S, McPherson JD, McCombie WR (2016) Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* 17:333–351. doi: 10.1038/nrg.2016.49
- Goulart A (2008) Aspectos Gerais sobre nematóides das lesões radiculares (gênero *Pratylenchus*). Embrapa Cerrados, Planaltina
- Goulart MMP, Carmo EL, Santos CB, et al (2013) Avaliação do efeito de diferentes populações de *Pratylenchus brachyurus* na cultura da soja. *Glob Sci Technol* 6:8–14.
- Grant MR, Godiard L, Straube E, et al (1995) Structure of the Arabidopsis RPM1 gene enabling dual specificity disease resistance. *Science* (80-) 269:843–846.
- Grant MR, Jones JDG (2009) Hormone (Dis)harmony moulds plant health and disease. *Sci mag* 324:750–753.
- Grigolli JFJ, Asmus GL (2014) Manejo de Nematoides na Cultura da Soja. *Nematoides da Soja* 3:194–203.
- Guo B, Sleper D a, Lu P, et al (2006a) QTLs Associated with Resistance to Soybean Cyst Nematode in Soybean: Meta-Analysis of QTL Locations. *Crop Sci* 46:595–602. doi: 10.2135/cropsci2005.04-0036-2
- Guo B, Sleper DA, Nguyen HT, et al (2006b) Quantitative trait loci underlying resistance to three soybean cyst nematode populations in soybean PI 404198A. *Crop Sci* 46:224–233. doi: 10.2135/cropsci2004.0757
- Guo X, Wang D, Gordon S, et al (2008) Genetic mapping of QTLs underlying partial resistance to *Sclerotinia sclerotiorum* in Soybean PI 391589A and PI 391589B. *Crop Sci* 48:1129–1139.
- Haegeman A, Joseph S, Gheysen G (2011) Analysis of the transcriptome of the root lesion nematode *Pratylenchus coffeae* generated by 454 sequencing technology. *Mol Biochem Parasitol* 178:7–14. doi: 10.1016/j.molbiopara.2011.04.001
- Haley CS, Knott SA (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity (Edinb)* 69:315–324.
- Hao D, Cheng H, Yin Z, et al (2012) Identification of single nucleotide polymorphisms and haplotypes associated with yield and yield components in soybean (*Glycine max*) landraces across multiple environments. *Theor Appl Genet* 124:447–458. doi: 10.1007/s00122-011-1719-0
- Holterman M, Karssen G, van den Elsen S, et al (2009) Small subunit rDNA-based phylogeny of the *Tylenchida* sheds light on relationships among some high-impact plant-parasitic nematodes and the evolution of plant feeding. *Phytopathology* 99:227–235. doi: 10.1094/PHYTO-99-3-0227
- Huang JS (1985) Mechanisms of resistance to root-knot nematodes. In: *An advanced treatise on Meloidogyne: biology and control*. North Carolina State University, Raleigh, pp 11–17
- Hussey R, Williamson V (1998) Physiological and molecular aspects of nematode parasitism. In: *Plant and nematodes interactions*. American society of agronomy, Madison, pp 87–108
- Hwang T-Y, Sayama T, Takahashi M, et al (2009) High-density integrated linkage map based on SSR markers in soybean. *DNA Res* 16:213–25. doi: 10.1093/dnares/dsp010

- Hyten DL, Choi IY, Song Q, et al (2010) A high density integrated genetic linkage map of soybean and the development of a 1536 universal soy linkage panel for quantitative trait locus mapping. *Crop Sci* 50:960–968. doi: 10.2135/cropsci2009.06.0360
- Inomoto MM, Silva RA (2011) Importância dos nematoides da soja e influência da sucessão de cultura. *Bol. Pesqui. Soja* 15:392–399.
- Jarquín D, Kocak K, Posadas L, et al (2014) Genotyping by sequencing for genomic prediction in a soybean breeding population. 1–10.
- Jiang C, Zeng Z (1995) Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* 140:1111–1127.
- Jiao Y, Vuong TD, Liu Y, et al (2015) Identification and evaluation of quantitative trait loci underlying resistance to multiple HG types of soybean cyst nematode in soybean PI 437655. *Theor Appl Genet* 128:15–23. doi: 10.1007/s00122-014-2409-5
- Jones JDG (2001) Putting knowledge of plant disease resistance genes to work. *Curr Opin Plant Biol* 4:281–287. doi: 10.1016/S1369-5266(00)00174-6
- Jones JDG, Dangl JL (2006) The plant immune system. *Nature* 444:323–329. doi: 10.1038/nature05286
- Jones JT, Haegeman A, Danchin EGJ, et al (2013) Top 10 plant-parasitic nematodes in molecular plant pathology. *Mol Plant Pathol* 14:946–961. doi: 10.1111/mpp.12057
- Jones MGK, Fosu-Nyarko J (2014) Molecular biology of root lesion nematodes (*Pratylenchus* spp.) and their interaction with host plants. *Ann Appl Biol* 164:163–181. doi: 10.1111/aab.12105
- Joseph S, Gheysen G, Subramaniam K (2012) RNA interference in *Pratylenchus coffeae*: Knock down of Pc-pat-10 and Pc-unc-87 impedes migration. *Mol Biochem Parasitol* 186:51–59. doi: 10.1016/j.molbiopara.2012.09.009
- Jun TH, Rouf Mian MA, Michel AP (2012) Genetic mapping revealed two loci for soybean aphid resistance in PI 567301B. *Theor Appl Genet* 124:13–22. doi: 10.1007/s00122-011-1682-9
- Kao CH, Zeng Z-B (1997) General formulas for obtaining the MLEs and the asymptotic variance-covariance matrix in mapping quantitative trait loci when using the EM algorithm. *Biometrics* 53:653–665.
- Kao CH, Zeng Z-B (2002) Modeling epistasis of quantitative trait loci using Cockerham's model. *Genetics* 160:1243–1261.
- Kao CH, Zeng ZB, Teasdale R (1999) Multiple interval mapping for quantitative trait loci. *Genetics* 152:1203–1216.
- Kazi S, Shultz J, Afzal J, et al (2010) Iso-lines and inbred-lines confirmed loci that underlie resistance from cultivar “Hartwig” to three soybean cyst nematode populations. *Theor Appl Genet* 120:633–644. doi: 10.1007/s00122-009-1181-4
- Kobe B, Deisenhofer J (1994) The leucine-rich repeat: a versatile binding motif. *Trends Biochem Sci* 19:415–421. doi: 10.1016/0968-0004(94)90090-6
- Komatsu K, Okuda S, Takahashi M, et al (2005) QTL Mapping of Antibiosis Resistance to Common Cutworm (*Fabricius*) in Soybean. *Crop Sci* 45:2044. doi: 10.2135/cropsci2004.0744
- Kosambi DD (1943) The estimation of map distance from recombination values. *Eugenics* 12:172–175.
- Kuroda Y, Kaga A, Tomooka N, et al (2013) QTL affecting fitness of hybrids between wild and cultivated soybeans in experimental fields. *Ecol Evol* 3:2150–2168. doi: 10.1002/ece3.606
- Kushalappa AC, Yogendra KN, Karre S (2016) Plant Innate Immune Response: Qualitative and Quantitative Resistance. *CRC Crit Rev Plant Sci* 35:38–55. doi: 10.1080/07352689.2016.1148980
- Lander E, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185–199.

- Larrieu A, Vernoux T (2016) Q&A: How does jasmonate signaling enable plants to adapt and survive? BMC Biol 14:79. doi: 10.1186/s12915-016-0308-8
- Laurie C, Wang S, Carlini-Garcia LA, Zeng ZB (2014) Mapping epistatic quantitative trait loci. BMC Genet 15:112. doi: 10.1186/s12863-014-0112-9
- Leamy LJ, Zhang H, Li C, et al (2017) A genome-wide association study of seed composition traits in wild soybean (*Glycine soja*). BMC Genomics 18:18. doi: 10.1186/s12864-016-3397-4
- Lee S, Freewalt KR, McHale LK, et al (2015) A high-resolution genetic linkage map of soybean based on 357 recombinant inbred lines genotyped with BARCSoySNP6K. Mol Breed 35:1–7. doi: 10.1007/s11032-015-0209-5
- Lee S, Mian MAR, McHale LK, et al (2013) Novel quantitative trait loci for partial resistance to *Phytophthora sojae* in soybean PI 398841. Theor Appl Genet 126:1121–1132. doi: 10.1007/s00122-013-2040-x
- Li X, Hang Y, Teng W, et al (2010) Pyramided QTL underlying tolerance to *Phytophthora* root rot in mega-environments from soybean cultivars “Conrad” and “Hefeng 25.” Theor Appl Genet 121:651–658.
- Liang H, Yu Y, Yang H, et al (2014) Inheritance and QTL mapping of related root traits in soybean at the seedling stage. Theor Appl Genet 127:2127–2137. doi: 10.1007/s00122-014-2366-z
- Liang Q, Cheng X, Mei M, et al (2010) QTL analysis of root traits as related to phosphorus efficiency in soybean. Ann Bot 106:223–234. doi: 10.1093/aob/mcq097
- Lima FS de O, Santos GR dos, Nogueira SR, et al (2015) Population dynamics of the root lesion nematode, *Pratylenchus brachyurus*, in soybean fields in tocanins state and its effect to soybean yield. Nematropica 45:170–177.
- Linsell KJ, Riley IT, Davies KA, Oldach KH (2013) Characterization of resistance to *Pratylenchus thornei* (Nematoda) in wheat (*Triticum aestivum*); attraction, penetration, motility and reproduction. Phytopathology 104:174–187. doi: 10.1094/PHYTO-12-12-0345-R.
- Liu H (1998) Statistical genomics: linkage, mapping and QTL analysis. CRC Press, Boca Raton
- Llorca CM, Potschin M, Zentgraf U (2014) bZIPs and WRKYs: two large transcription factor families executing two different functional strategies. Front Plant Sci 5:1–14. doi: 10.3389/fpls.2014.00169
- Loake G, Grant M (2007) Salicylic acid in plant defence-the players and protagonists. Curr Opin Plant Biol 10:466–472. doi: 10.1016/j.pbi.2007.08.008
- Lordello L (1988) Nematóides das plantas cultivadas, 6ed. Nobel, São Paulo
- Lorenz AJ, Smith KP, Jannink J-L (2012) Potential and optimization of genomic selection for fusarium head blight resistance in six-row barley. Crop Sci 52:1609–1621.
- Luca FD, Reyes A, Troccoli A, Castillo P (2011) Molecular variability and phylogenetic relationships among different species and populations of *Pratylenchus* (Nematoda: Pratylenchidae) as inferred from the analysis of the ITS rDNA. Eur J Plant Pathol 130:415–426. doi: 10.1007/s10658-011-9763-9
- Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland
- Machado ACZ, Araújo Filho J V. (2016) Broad-sense heritability and variance component estimates for *Pratylenchus brachyurus* resistance in Brazilian soybean genotypes. Trop Plant Pathol 41:390–396. doi: 10.1007/s40858-016-0113-5
- Machado ACZ, Beluti DB, Silva RA, et al (2006) Avaliação de Danos Causados por *Pratylenchus brachyurus* em Algodoeiro. Delta 31:11–16.
- Machado ACZ, Ferraz LCCB, Oliveira CMG De (2005) Development of a Species-Specific Reverse Primer for the Molecular Diagnostic of *Pratylenchus Brachyurus*. Nematropica 37:249–258.

- Machado ACZ, Siqueira KMS, Ferraz LCCB, et al (2015) Characterization of Brazilian populations of *Pratylenchus brachyurus* using morphological and molecular analyses. Trop Plant Pathol 40:102–110. doi: 10.1007/s40858-015-0013-0
- Mainardi JT, Asmus GL (2015) Danos e potencial reprodutivo de *Pratylenchus brachyurus* em cinco espécies vegetais. 38–47.
- Manavalan LP, Prince SJ, Musket TA, et al (2015) Identification of Novel QTL Governing Root Architectural Traits in an Interspecific Soybean Population. PLoS One. doi: 10.1371/journal.pone.0120490
- Margarido GRA, Souza AP, Garcia AAF (2007) OneMap: Software for genetic mapping in outcrossing species. Hereditas 144:78–79. doi: 10.1111/j.2007.0018-0661.02000.x
- Maruki T, Lynch M (2015) Genotype-frequency estimation from high-throughput sequencing data. Genetics 201:473–486. doi: 10.1534/genetics.115.179077
- Massman JM, Jung HJG, Bernardo R (2013) Genomewide selection versus marker-assisted recurrent selection to improve grain yield and stover-quality traits for cellulosic ethanol in maize. Crop Sci 53:58–66. doi: 10.2135/cropsci2012.02.0112
- Matsuo É, Sediyaama T, Arc RD, et al (2012) Characterization of type and genetic diversity among soybean cyst nematode differentiators. 69:147–151.
- McDowell JM, Woffenden BJ (2003) Plant disease resistance genes: Recent insights and potential applications. Trends Biotechnol 21:178–183. doi: 10.1016/S0167-7799(03)00053-2
- Melillo MT, Leonetti P, Bongiovanni M, et al (2006) Modulation of reactive oxygen species activities and H₂O₂ accumulation during compatible and incompatible tomato – root-knot nematode interactions. New Phytol 170:501–512. doi: 10.1111/j.1469-8137.2006.01724.x
- Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.
- Mindrinis M, Katagiri F, Yu G-L, Ausubel FM (1994) The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. Cell 78:1089–1099.
- Miragaya JCG (2005) Biodiesel: Tendências no mundo e no Brasil. Inf Agropecuário 26:7–13.
- Missão MR (2006) Soja: origem, classificação, utilização e uma visão abrangente do mercado. Rev Ciências Empres 3:7–15.
- Mitreva M, Elling AA, Dante M, et al (2004) A survey of SL-1 spliced transcripts from the root lesion nematode *Pratylenchus penetrans*. Mol Genet genomics 272:138–148.
- Möller M (2017) Identificação de QTLs em soja associados à resistência aos percevejos e a caracteres agrônômicos utilizando a abordagem de mapeamento multivariado. Universidade de São Paulo, Tese de Doutorado
- Monteiro A, Ferraz L, Inomoto M (2003) Apontamentos de nematologia em plantas. ESALQ, Piracicaba
- Motter P, Goelzer H, Valle D, et al (2015) A tecnologia que revolucionou a agricultura brasileira. Plantio Direto 144.
- Moura MF (1999) Avaliação de linhagens de soja (*Glycine max* (L.) Merrill) visando adaptação às condições de cerrado e resistência a insetos. Universidade Federal de Goiás, Dissertação de Mestrado
- Moura RM, Régis EMO (1987) Reação de cultivares de feijoeiro comum (*Phaseolus vulgaris*) em relação ao parasitismo de *Meloidogyne javanica* e *M. incognita*. Nematol Bras 11:215–225.
- Mundt CC (2015) Pathogens and Pests. 446–455. doi: 10.1016/j.meegid.2014.01.011.Durable
- Nahar K, Kyndt T, De Vleeschauwer D, et al (2011) The Jasmonate Pathway Is a Key Player in Systemically Induced Defense against Root Knot Nematodes in Rice. Plant Physiol 157:305–316. doi: 10.1104/pp.111.177576

- NCBI (2017) National Center for Biotechnology Information, Taxonomy. <https://www.ncbi.nlm.nih.gov/>. Accessed 11 Jul 2017
- Nguyen PV, Bellaifiore S, Petitot A-S, et al (2014) *Meloidogyne incognita* - rice (*Oryza sativa*) interaction: a new model system to study plant-root-knot nematode interactions in monocotyledons. *Rice* 7:23. doi: 10.1186/s12284-014-0023-4
- Nicol JM, Turner SJ, Coyne DL, et al (2011) Genomics and Molecular Genetics of Plant-Nematode Interactions. doi: 10.1007/978-94-007-0434-3
- Nicol P, Gill R, Fosu-Nyarko J, Jones MGK (2012) De novo analysis and functional classification of the transcriptome of the root lesion nematode, *Pratylenchus thornei*, after 454 GS FLX sequencing. *Int J Parasitol* 42:225–237. doi: 10.1016/j.ijpara.2011.11.010
- Oldach KH, Peck DM, Nair RM, et al (2014) Genetic analysis of tolerance to the root lesion nematode *Pratylenchus neglectus* in the legume *Medicago littoralis*. *BMC Plant Biol* 14:100. doi: 10.1186/1471-2229-14-100
- Oostenbrink M (1966) Major characteristics of the relations between nematodes and plants. *Meded Landbouwhogeschool* 66:1–46.
- Paiva BM, Alves RM, Hel NM (2006) Aspectos socioeconômicos da soja. 26:7–14.
- Patterson H, Thompson R (1971) Recovery of inter-block information when block sizes are unequal. *Biometrika* 58:545–554.
- Peterson BG, Carl P, Boudt K, et al (2015) Econometric tools for performance and risk analysis - PerformanceAnalytics. 193p.
- Pieterse CMJ, Leon-Reyes A, Van der Ent S, Van Wees SCM (2009) Networking by small-molecule hormones in plant immunity. *Nat Chem Biol* 5:308–316. doi: 10.1038/nchembio.164
- Pinheiro JC, Bates DM (2000) Mixed-effects models in S and S-PLUS. Springer, New York
- Pípolo EA, Hungria M, Franchini JC, et al (2015) Teores de óleo e proteína em soja: fatores envolvidos e qualidade para a indústria. Londrina, PR
- Prince SJ, Song L, Qiu D, et al (2015) Genetic variants in root architecture-related genes in a *Glycine soja* accession, a potential resource to improve cultivated soybean. *BMC Genomics* 16:132. doi: 10.1186/s12864-015-1334-6
- R Development Core Team (2017) R: A language and environment for statistical computing.
- Rehman S, Gupta VK, Goyal AK (2016) Identification and functional analysis of secreted effectors from phytoparasitic nematodes. *BMC Microbiol* 16:48. doi: 10.1186/s12866-016-0632-8
- Ribeiro N, Bezerra F, Silveira T, et al (2009) Avaliação da resistência de genótipos de milho (*Zea mays*) ao nematóide *Pratylenchus brachyurus*. In: 2 International congress of tropical nematology. Anais do 2 International congress of tropical nematology, Maceió, p 41
- Ribeiro N, Dias W, Santos J (2010) Distribuição de fitonematoides em regiões produtoras de soja no estado de Mato Grosso. *Bol Pesqui da soja* 14:289–295.
- Ribeiro N, Miranda D, Favoreto L (2011) Nematoides um desafio constante. *Bol Pesqui da soja* 400–414.
- Robert-Seilanianz A, Grant M, Jones JDG (2011) Hormone Crosstalk in Plant Disease and Defense: More Than Just Jasmonate-salicylate Antagonism. *Annu Rev Phytopathol* 49:317–343. doi: 10.1146/annurev-phyto-073009-114447
- Roberts PA (2002) Concepts and consequences of resistance. In: Plant resistance to parasitic nematodes. CABI, Wallingford, pp 23–41

- Rocha MR, Santos LC, Teixeira RA, et al (2008) Reação de 33 cultivares de a *Pratylenchus brachyurus* em área naturalmente infestada. In: REUNIÃO DE PESQUISA DE SOJA DA REGIÃO CENTRAL DO BRASIL. Embrapa/CNPso, Rio Verde, GO, pp 1410–1441
- Rodrigues DB, Dias-Arieira CR, Vilvert Vedoveto MV, et al (2014) Crop rotation for *Pratylenchus brachyurus* control in soybean. *Nematropica* 44:146–151.
- Salmeron JM, Oldroyd GED, Rommens CMT, et al (1996) Tomato *Pto* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. *Cell* 86:123–133. doi: 10.1016/S0092-8674(00)80083-5
- Santana-Gomes S de M, Hernandez I, Machado ACZ, et al (2016) Host status and phenotypic diversity of rice genotypes in relation to *Pratylenchus brachyurus* resistance. *Nematology* 18:15–21. doi: 10.1163/15685411-00002938
- Santana FA, Silva MF da, Guimarães JKF, et al (2014) Marker-assisted selection strategies for developing resistant soybean plants to cyst nematode. *Crop Breed Appl Biotechnol* 14:180–186. doi: 10.1590/1984-70332014v14n3a27
- Santner A, Calderon-Villalobos LIA, Estelle M (2009) Plant hormones are versatile chemical regulators of plant growth. *Nat Chem Biol* 5:301–307. doi: 10.1038/nchembio.165
- Schmitt DP, Barker KR (1981) Damage and Reproductive Potentials of *Pratylenchus brachyurus* and *P. penetrans* on soybean. *J Nematol* 13:327–332.
- Schmutz J, Cannon SB, Schlueter J, et al (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463:178–83. doi: 10.1038/nature08670
- Schuster I, Cruz C (2004) Estatística genômica aplicada a populações derivadas de cruzamentos controlados. Viçosa, MG
- SECEX (2017) Ministério da indústria, comércio exterior e serviços. <http://portal.siscomex.gov.br/informativos/noticias-orgaos/secex>.
- Sediyama T (2015) Melhoramento genético da soja. Mecnas, Londrina
- Sediyama T (2009) Tecnologia de produção e uso da soja. Mecnas, Londrina
- Sentelhas PC, Battisti R, Câmara GMS, et al (2015) The soybean yield gap in Brazil – magnitude, causes and possible solutions for sustainable production. *J Agric Sci* 153:1394–1411. doi: 10.1017/S0021859615000313
- Sharma S, Sharma S, Kopisch-Obuch FJ, et al (2011) QTL analysis of root-lesion nematode resistance in barley: 1. *Pratylenchus neglectus*. *Theor Appl Genet* 122:1321–1330.
- Sheedy JG, Thompson JP (2009) Resistance to the root-lesion nematode *Pratylenchus thornei* of Iranian landrace wheat. *Australas Plant Pathol* 38:478–489. doi: 10.1071/AP09030
- Silva F, Borém A, Sediyama T, Ludke W (2017) Melhoramento da Soja. UFV, Viçosa, MG
- Silva L, Wang S, Zeng Z-B (2012) Multiple trait multiple interval mapping of quantitative trait loci from inbred line crosses. *BMC Genet* 13:67. doi: 10.1186/1471-2156-13-67
- Silva L, Zeng Z-B (2010) Current progress on statistical methods for mapping quantitative trait loci from inbred line crosses. *J Biopharm Stat* 20:454–481.
- Singh KB, Foley RC, Oñate-Sánchez L (2002) Transcription factors in plant defense and stress responses. *Curr Opin Plant Biol* 5:430–436. doi: 10.1016/S1369-5266(02)00289-3
- Smiley RW, Whittaker RG, Gourlie JA, Easley SA (2005) *Pratylenchus thornei* Associated with Reduced Wheat Yield in Oregon. *J Nematol* 37:45–54.

- Sonah H, O'Donoghue L, Cober E, et al (2015) Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. *Plant Biotechnol J* 13:211–221. doi: 10.1111/pbi.12249
- Song Q, Hyten DL, Jia G, et al (2013) Development and Evaluation of SoySNP50K, a High-Density Genotyping Array for Soybean. *PLoS One* 8:1–12. doi: 10.1371/journal.pone.0054985
- Song QJ, Marek LF, Shoemaker RC, et al (2004) A new integrated genetic linkage map of the soybean. *Theor Appl Genet* 109:122–8. doi: 10.1007/s00122-004-1602-3
- Spindel J, Begum H, Akdemir D, et al (2015) Genomic selection and association mapping in rice (*Oryza sativa*): effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. *PLoS Genet* 11:e1004982. doi: 10.1371/journal.pgen.1004982
- Subbotin SA, Ragsdale EJ, Mullens T, et al (2008) A phylogenetic framework for root lesion nematodes of the genus *Pratylenchus* (Nematoda): Evidence from 18S and D2-D3 expansion segments of 28S ribosomal RNA genes and morphological characters. *Mol Phylogenet Evol* 48:491–505. doi: 10.1016/j.ympev.2008.04.028
- Sundararaju P (2010) Identification of Nematode Resistant Gene Sources against Root-Lesion Nematode (*Pratylenchus coffeae*) in Banana. *Indian J Nematol* 40:48–54.
- Swaminathan S, Abeysekara NS, Liu M, et al (2016) Quantitative trait loci underlying host responses of soybean to *Fusarium virguliforme* toxins that cause foliar sudden death syndrome. *Theor Appl Genet* 129:495–506. doi: 10.1007/s00122-015-2643-5
- Tan JACH, Jones MGK, Fosu-Nyarko J (2013) Gene silencing in root lesion nematodes (*Pratylenchus* spp.) significantly reduces reproduction in a plant host. *Exp Parasitol* 133:166–178. doi: 10.1016/j.exppara.2012.11.011
- Terasawa F, Terasawa JM, Terasawa MM (2017) A FT Sementes e a Expansão da Soja no Brasil. In: Silva F, Borém A, Sedyama T, Ludke W (eds) *Melhoramento da Soja*. UFV, Viçosa, MG, pp 30–43
- Thakur M, Sohal BS (2013) Role of Elicitors in Inducing Resistance in Plants against Pathogen Infection: A Review. *ISRN Biochem* 2013:1–10. doi: 10.1155/2013/762412
- Toledo ER De, Leandro RA, Souza Junior CL De, Souza AP De (2008) Mapeamento de QTLS: uma abordagem Bayesiana. *Rev Bras Biometria* 26:107–114.
- USDA (2017) United States Department of Agriculture. <https://www.usda.gov/>.
- Valentini M (2014) Culturas sucessoras a Soja, Milho e Algodão, que promovem a redução do nematoide-das-lesões-radiculares. *Agrisus - Chapadão do Sul-Mato Grosso do Sul*
- Veech JA, Endo BY (1970) Comparative morphology and enzyme histochemistry of root-knot resistant and susceptible soybean. *Phytopathology* 60:90–96.
- Vieira AJD, De Oliveira DA, Soares TCB, et al (2006) Use of the QTL approach to the study of soybean trait relationships in two populations of recombinant inbred lines at the F7 and F8 generations. *Brazilian J Plant Physiol* 18:281–290. doi: 10.1590/S1677-04202006000200004
- Vieira P, Eves-Van Den Akker S, Verma R, et al (2015) The *Pratylenchus penetrans* transcriptome as a source for the development of alternative control strategies: Mining for putative genes involved in parasitism and evaluation of in planta RNAi. *PLoS One* 10:1–25. doi: 10.1371/journal.pone.0144674
- Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93:77–78.

- Vuong T, Diers BW, Harman G (2008) Identification of QTL for Resistance to Sclerotinia Stem Rot in Soybean Plant Introduction 194639. *Crop Sci* 48:2209–2214.
- Vuong TD, Slepier DA, Shannon JG, Nguyen HT (2010) Novel quantitative trait loci for broad-based resistance to soybean cyst nematode (*Heterodera glycines* Ichinohe) in soybean PI 567516C. *Theor Appl Genet* 121:1253–1266.
- Walker D, Monteros M, Yates J (2010) Molecular breeding. In: Genetics, genomics and breeding of soybean. Science Publishers, Enfield, NH, pp 123–167
- Wall JD, Tang LF, Zerbe B, et al (2014) Estimating genotype error rates from high-coverage next-generation sequence data. *Genome Res* 24:1734–1739. doi: 10.1101/gr.168393.113
- Wang H, St.Martin S, Dorrance A (2012) Comparison of Phenotypic Methods and Yield Contributions of Quantitative Trait Loci for Partial Resistance to *Phytophthora sojae* in Soybean. *Crop Sci* 52:609–622.
- Wasternack C (2007) Jasmonates: An update on biosynthesis, signal transduction and action in plant stress response, growth and development. *Ann Bot* 100:681–697. doi: 10.1093/aob/mcm079
- Williams B, Kantartzi SK, Meksem K, et al (2012) Genetic Analysis of Root and Shoot Traits in the “Essex” By “Forrest” Recombinant Inbred Line (RIL) Population of Soybean [*Glycine max* (L.) Merr.]. *J Plant Genome Sci* 1:1–9. doi: 10.5147/jpgs.2012.0051
- Williamson VM, Roberts PA (2010) Mechanisms and genetics of resistance. In: Root-knot nematodes. CAB International, Wallingford, pp 301–325
- Winter SMJ, Shelp BJ, Anderson TR, et al (2007) QTL associated with horizontal resistance to soybean cyst nematode in *Glycine soja* PI464925B. *Theor Appl Genet* 114:461–472. doi: 10.1007/s00122-006-0446-4
- Wu X, Blake S, Slepier D, et al (2009) QTL, additive and epistatic effects for SCN resistance in PI 437654. *Theor Appl Genet* 118:1093–1105.
- Xavier A, Muir WM, Rainey KM (2016) Assessing Predictive Properties of Genome-Wide Selection in Soybeans. 6:2611–2616. doi: 10.1534/g3.116.032268
- Yan C, Xie D (2015) Jasmonate in plant defence: Sentinel or double agent? *Plant Biotechnol J* 13:1233–1240. doi: 10.1111/pbi.12417
- Yan G, Smiley RW, Okubara P a (2012) Detection and quantification of *Pratylenchus thornei* in DNA extracted from soil using real-time PCR. *Phytopathology* 102:14–22. doi: 10.1094/PHYTO-03-11-0093
- Yan Y, Borrego E, Kolomiets M V (2013) Jasmonate biosynthesis, perception and function in plant development and stress responses. In: Baez R V (ed) *Lipid Metabolism*. InTech, pp 393–442
- Zeng Z-B, Kao CH, Basten CJ (1999) Estimating the genetic architecture of quantitative traits. *Genet Res (Camb)* 74:279–289.
- Zeng ZB (2001) Statistical methods formapping quantitative trait loci. Raleigh, North Carolina
- Zeng ZB (1993) Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proc Natl Acad Sci* 90:10972–10976.
- Zhang J, Singh A, Mueller DS, Singh AK (2015a) Genome-wide association and epistasis studies unravel the genetic architecture of sudden death syndrome resistance in soybean. *Plant J* 84:1124–1136. doi: 10.1111/tpj.13069
- Zhang J, Song QJ, Cregan PB, et al (2015b) Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. *BMC Genomics* 16:217. doi: 10.1186/s12864-015-1441-4

- Zhang J, Song QJ, Cregan PB, Jiang G-L (2016) Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine max*). Theor Appl Genet 129:117–130. doi: 10.1007/s00122-015-2614-x
- Zhou R, Chen H-F, Wang XZ, et al (2011) Analysis of QTLs for Root Traits at Seedling Stage in Soybean. Acta Agron Sin 37:1151–1158.
- Zhu YL, Song QJ, Hyten DL, et al (2003) Single-nucleotide polymorphisms in soybean. Genetics 163:1123–34.
- Zou F, Fine JP, Hu J, Lin DY (2004) An efficient resampling method for assessing genome-wide statistical significance in mapping quantitative trait loci. Genetics 168:2307–2316. doi: 10.1534/genetics.104.031427

APÊNDICES

APÊNDICE A.

Alinhamento e anotação de sequências candidatas selecionadas para a região 2, no

GL A2 (8A).

Sequência Candidata	Descrição do Alinhamento	Comprimento (pb)	Valor Esperado	Similaridade (%)	Termos <i>GO</i>
PUT-169a-Glycine_max-161045	Glutathione S-Transferase F9	214	1,48E ⁻²⁶	98,35	0004364, 0005737, 0006749
PUT-169a-Glycine_max-128262	Glutathione S-Transferase F9	427	1,11E ⁻⁴⁴	93,75	0016740
PUT-169a-Glycine_max-199918	Aspartic Ase Nepenthesin-2	667	2,60E ⁻⁴⁰	81,15	0003735, 0003860, 0004190, 0005840, 0006412, 0006508, 0016740, 0030163
PUT-169a-Glycine_max-138410	3-Hydroxyisobutyryl-hidrolase 5	1540	0,0	93,20	0003735, 0003860, 0005840, 0006412, 0016853
PUT-169a-Glycine_max-200062	Vacuoleless1	264	8,68E ⁻³⁹	98,20	0032889, 0005768, 0035542, 0003779, 0007033, 0016197, 0006886, 0005765, 0030897
PUT-169a-Glycine_max-183757	Callose Synthase 7	246	1,16E ⁻⁴⁸	98,50	0003843, 0006075, 0016021, 0000148
PUT-169a-Glycine_max-24649	GDSL Esterase Lipase At4g10955-like	442	1,76E ⁻³⁰	84,85	0006629, 0016787
PUT-169a-Glycine_max-325271727	Mitogen-activated Kinase 10	1275	2,13E ⁻⁸⁸	97,65	0005524, 0000165, 0004707, 0005622
PUT-169a-Glycine_max-237959	Probable Galacturonosyltransferase 15	867	0,0	93,00	0047262, 0016021, 0045489, 0000139, 0071555
PUT-169a-Glycine_max-45050	Agamous-like MADS-box AGL1 Isoform X3	135	2,53E ⁻⁸	100,00	0005634, 0003677, 0003700, 0006351, 0006355, 0046983
PUT-169a-Glycine_max-215008	Zinc Finger BED Domain-containing RICESLEEPER 2-like	817	9,88E ⁻¹²⁸	95,70	0005634, 0003700, 0000977, 0046983, 0006357
PUT-169a-Glycine_max-142944	Serine Threonine Phosphatase 5	697	5,77E ⁻⁵⁹	99,40	0005634, 0004721, 0005737, 0006470, 0016021
PUT-169a-Glycine_max-157938	Citrate Glyoxysomal-like	217	2,00E ⁻⁶	97,33	0016746, 0006099, 0016740, 0046912
PUT-169a-Glycine_max-75907	Transcription Elongation Factor B Polypeptide 3	477	2,11E ⁻¹⁴	96,85	0005634, 0016021, 0006355
PUT-169a-Glycine_max-172201	Spermidine Hydroxycinnamoyl Transferase-like	508	1,23E ⁻¹¹⁹	87,00	0047672, 0047172
PUT-169a-Glycine_max-129168	Ataxin-2 Homolog Isoform X1	908	2,05E ⁻⁶⁹	94,50	0034063, 0010494, 0010603
PUT-169a-Glycine_max-91329	Retrovirus-related Pol Poly LINE-1	72	1,30E ⁻⁷	95,55	0003964, 0006278
PUT-169a-Glycine_max-61538	Heat Shock 70 kDa 16	451	1,99E ⁻¹⁰⁴	91,90	0005524
PUT-169a-Glycine_max-63605	Dipeptidyl Peptidase Family Member 6	272	8,03E ⁻⁵⁶	97,60	0006508, 0008236
PUT-169a-Glycine_max-113224	Hypothetical Protein Glysoja 042752	612	3,92E ⁻¹³⁸	95,55	0000902

APÊNDICE B.

GL A2 (8B).

Alinhamento e anotação de sequências candidatas selecionadas para a região 3, no

Sequência Candidata	Descrição do Alinhamento	Comprimento (pb)	Valor Esperado	Similaridade (%)	Termos <i>GO</i>
PUT-169a-Glycine_max-122199	ALA-interacting Subunit 3	1461	0,00	92,40	0016021
PUT-169a-Glycine_max-138626	Ent-kaur-16-ene Chloroplasic Isoform X2	463	1,12E ⁻⁹⁹	92,20	0000287, 0008152, 0009899
PUT-169a-Glycine_max-271400	50S Ribosomal L7 L12	295	1,83E ⁻¹³	99,55	0003735, 0005840, 0006412
PUT-169a-Glycine_max-336271735	Zinc Transporter 1	238	2,55E ⁻³²	96,40	0016021, 0005385, 0071577, 0005886
PUT-169a-Glycine_max-21271729	Metal Tolerance 10-like	332	1,76E ⁻⁷⁰	96,70	0010043, 0061088, 0016021, 0005385, 0071577
PUT-169a-Glycine_max-1795271735	Phosphoglycerate Cytosolic	1590	0,00	96,85	0005524, 0004618, 0006096
PUT-169a-Glycine_max-162527	Phosphoglycerate Chloroplasic	254	1,17E ⁻³³	96,50	0005524, 0004618, 0006096
PUT-169a-Glycine_max-181763	Serine Threonine Kinase CTR1	900	8,42E ⁻¹⁴⁹	97,85	0005524, 0000165, 0000186, 0004709, 0005622
PUT-169a-Glycine_max-182131	Probable Serine Threonine Kinase DDB_G0282963 Isoform X1	1362	0,00	92,70	0005524, 0000165, 0000186, 0004709, 0005622
PUT-169a-Glycine_max-13984	Dual Specificity Kinase splB-like	260	6,37E ⁻⁴²	99,45	0005524, 0000165, 0000186, 0004709, 0005622
PUT-169a-Glycine_max-210989	Translation Elongation Factor EF-2 Subunit	237	1,30E ⁻³⁶	99,80	0005525, 0003746, 0003924, 0006414, 0005622
PUT-169a-Glycine_max-155880	UDP-Glycosyltransferas e TURAN Isoform X1	870	7,55E ⁻¹¹⁸	95,40	0004578, 0097502, 0006486, 0005783
PUT-169a-Glycine_max-49221	Probable BOI-related E3 Ubiquitin Ligase 3	66	4,32E ⁻⁶	100,00	0004842, 0005634, 0008270, 0031347, 0016567
PUT-169a-Glycine_max-31690	Probable BOI-related E3 Ubiquitin Ligase 3	95	1,60E ⁻⁹	100,00	0005634, 0004842, 0008270, 0031347, 0016567
PUT-169a-Glycine_max-106195	Probable BOI-related E3 Ubiquitin Ligase 3	99	1,33E ⁻⁸	100,00	0005634, 0004842, 0008270, 0031347, 0016567
PUT-169a-Glycine_max-77665	Probable BOI-related E3 Ubiquitin Ligase 3	94	7,07E ⁻⁷	100,00	0005634, 0004842, 0008270, 0031347, 0016567
PUT-169a-Glycine_max-145242	Probable BOI-related E3 Ubiquitin Ligase 3	95	1,21E ⁻⁶	100,00	0005634, 0004842, 0008270, 0031347, 0016567
PUT-169a-Glycine_max-256632	PAT1 like 1	662	4,62E ⁻⁴⁹	91,65	0003723, 0000290, 0033962, 0000932, 0016853
PUT-169a-Glycine_max-108735	NAC Domain-containing 21 22	556	2,56E ⁻⁴⁸	95,80	0005634, 0003677, 0006351, 0006355

Continuação Apêndice B					
PUT-169a- Glycine_max-184021	D-Box ATP- dependent RNA helicase D 12	124	5,81E ⁻²¹	100,00	0003676, 0005524, 0030529, 0019013, 0004386
PUT-169a- Glycine_max- 1588271727	Photosystem II 10 kDa Chloroplastic	755	2,07E ⁻⁹¹	95,20	0016021, 0009654, 0015979, 0009535
PUT-169a- Glycine_max-94919	DNAJ Homolog DNJ-5-like	391	4,58E ⁻⁸⁰	96,95	0016021
PUT-169a- Glycine_max- 1080271732	DNAJ Homolog DNJ-5-like	729	6,29E ⁻¹³⁶	96,95	0016021

APÊNDICE C.

Alinhamento e anotação de sequências candidatas selecionadas para a região 4, no GL K (9B).

Sequência Candidata	Descrição do Alinhamento	Comprimento (pb)	Valor Esperado	Similaridade (%)	Termos <i>GO</i>
PUT-169a-Glycine_max-99566	Transcription Factor DIVARICATA-like	383	3,31E ⁻⁸⁸	94,80	0005634, 0003677, 0006351, 0006355
PUT-169a-Glycine_max-2156271729	LSM12 Homolog	1106	2,02E ⁻¹⁵³	93,40	0016021
PUT-169a-Glycine_max-253726	Cyclin-D1-1	457	0,00	89,50	0005634
PUT-169a-Glycine_max-1598	Retrovirus-related Pol Poly LINE-1	414	1,46E ⁻⁶⁸	87,80	0003964, 0006278
PUT-169a-Glycine_max-11301	Retrovirus-related Pol Poly LINE-1	251	1,18E ⁻⁴⁴	73,00	0008152
PUT-169a-Glycine_max-128097	TIR-NBS-LRR Type Disease Resistance	155	8,45E ⁻²⁶	86,45	0008152, 0016021, 0016787
PUT-169a-Glycine_max-226897	Probable Inactive Purple Acid Phosphatase 1 Isoform X1	533	1,05E ⁻¹¹⁹	95,70	0046872, 0016311, 0003993
PUT-169a-Glycine_max-225123	IAA-Amino Acid Hydrolase ILR1-like 4	304	1,05E ⁻¹⁶	92,43	0008152, 0016787
PUT-169a-Glycine_max-31891	Transcription Factor MYB1R1-like	676	3,30E ⁻⁹⁸	79,55	0035066, 0005634, 0003677, 0004402, 0003700, 0008270, 0006338, 0003713, 0006351, 0016573, 0003682, 0006357
PUT-169a-Glycine_max-198663	Phospholipid-transporting ATPase 3	638	3,87E ⁻¹⁴⁶	99,35	0000287, 0005524, 0005802, 0045332, 0016021, 0000139, 0004012, 0048194, 0005886
PUT-169a-Glycine_max-202975	DNAJ Homolog Subfamily B Member 4-like	723	1,28E ⁻¹²³	90,20	0051082, 0006457
PUT-169a-Glycine_max-190969	Probable AMP Deaminase	455	6,35E ⁻²⁷	94,90	0003876, 0032264, 0016021
PUT-169a-Glycine_max-101382	Pentatricopeptide Repeat-containing At5g40405	272	6,13E ⁻⁴⁵	90,60	0008270, 0008568
PUT-169a-Glycine_max-133707	NRT1 PTR Family-like	975	2,10E ⁻⁶⁴	93,05	0016021, 0006810, 0005215

APÊNDICE D. Alinhamento e anotação de sequências candidatas selecionadas para a região 6, no GL E (15B).

Sequência Candidata	Descrição do Alinhamento	Comprimento (pb)	Valor Esperado	Similaridade (%)	Termos <i>GO</i>
PUT-169a-Glycine_max-224140	Zinc Transporter ZTP 29	465	3,02E ⁻²⁰	98,35	0016021, 0005385, 0071577
PUT-169a-Glycine_max-28821	Xanthine Dehydrogenase	567	1,00E ⁻¹²¹	96,90	0004854, 0016614, 0004855, 0005506, 0016903, 0050660, 0009055, 0005829, 0009115, 0055114, 0051537
PUT-169a-Glycine_max-115986	Polyadenylate-binding RBP47B-like Isoform X2	228	9,57E ⁻⁸	99,20	0000166, 0003676
PUT-169a-Glycine_max-102224	Origin of Replication Complex Subunit 1A-like	229	5,13E ⁻⁴⁶	95,80	0005524, 0008270, 0003682
PUT-169a-Glycine_max-68441	Transcription Elongation Factor B Polypeptide 1	717	5,80E ⁻⁶⁴	95,95	0032968, 0004842, 0003746, 0005829, 0070449, 0006414, 0042787
PUT-169a-Glycine_max-68370	Transcription Elongation Factor B Polypeptide 1	774	1,35E ⁻⁶³	95,95	0032968, 0004842, 0003746, 0005829, 0070449, 0006414, 0042787
PUT-169a-Glycine_max-68368	Transcription Elongation Factor B Polypeptide 1	438	1,16E ⁻⁶⁵	95,95	0032968, 0004842, 0003746, 0005829, 0070449, 0006414, 0042787
PUT-169a-Glycine_max-1379271735	TOC75-Chloroplastic	1000	4,28E ⁻¹⁶²	96,55	0071806, 0015450, 0006886, 0005622, 0019867
PUT-169a-Glycine_max-163094	Pumilio Homolog 2-like Isoform X1	734	1,58E ⁻³²	99,00	0003723
PUT-169a-Glycine_max-112159	Scarecrow 14	968	0,00	91,55	0005634, 0003700, 0006351, 0006355, 0043565
PUT-169a-Glycine_max-91329	Retrovirus-related Pol Poly LINE-1	72	1,30E ⁻⁷	95,55	0003964, 0006278
PUT-169a-Glycine_max-128097	TIR-NBS-LRR Type Disease Resistance	155	8,41E ⁻²⁶	86,45	0008152, 0016021, 0016787
PUT-169a-Glycine_max-126710	Factor of DNA Methylation 4-like	210	6,03E ⁻²¹	92,69	0090502, 0004526, 0031047, 0005655
PUT-169a-Glycine_max-263457	Triosephosphate Cytosolic	103	3,40E ⁻¹³	98,20	0003988, 0006096, 0004807
PUT-169a-Glycine_max-1053271727	Outer Envelope Pore Chloroplastic	831	7,20E ⁻⁹⁸	91,10	0009707, 0005744, 0031305, 0006811, 0015266, 0015288, 0042803, 0030150, 0046930, 0045037, 0034426
PUT-169a-Glycine_max-238223	U-box Domain-containing 33	462	1,64E ⁻³³	95,20	0004842, 0005524, 0004672, 0006468, 0016567
PUT-169a-Glycine_max-119342	E3 Ubiquitin Ligase XBAT31	1954	0,00	93,75	0008270, 0016874

Continuação Apêndice D					
PUT-169a-Glycine_max-207662	Nucleolar Complex 2 Homolog	477	3,04E ⁻¹¹¹	93,85	0004601, 0016021, 0098869
PUT-169a-Glycine_max-34746	Histone H2A	221	2,95E ⁻²⁵	99,85	0003677, 0000790, 0046982, 0006342, 0000786
PUT-169a-Glycine_max-391271727	RAS-related RABE 1c	596	3,46E ⁻⁷²	94,30	0005525, 0004767, 0007264, 0006904, 0016021, 0009306, 0005622
PUT-169a-Glycine_max-220198	Zinc Finger ZAT12	235	1,65E ⁻²³	84,89	0005634, 0003700, 0046872, 0044212, 0006355, 0043565, 0010200
PUT-169a-Glycine_max-78793	Methyl-Binding Domain-containing 7	585	7,44E ⁻²⁰	88,56	0000122, 0005634, 0008327, 0006346
PUT-169a-Glycine_max-81659	Nuclease HARBI1	108	5,43E ⁻¹⁴	93,80	0005524, 0006270, 0043565, 0006275
PUT-169a-Glycine_max-21255	Hypothetical Protein Glysoja_010579	89	1,99E ⁻⁶	95,00	0016020, 0016021
PUT-169a-Glycine_max-171505	60S Ribosomal L23A	895	3,50E ⁻¹⁰⁰	95,35	0000166, 0003735, 0022625, 0000027, 0006412
PUT-169a-Glycine_max-136728	Nodulin Homeobox Isoform X1	244	6,76E ⁻³¹	99,25	0003677
PUT-169a-Glycine_max-150295	Nutrient Reservoir	286	4,52E ⁻⁴⁵	97,65	0045735
PUT-169a-Glycine_max-162123	60S Ribosomal L35	869	2,64E ⁻⁷⁴	97,75	0000463, 0003735, 0003729, 0022625, 0006412
PUT-169a-Glycine_max-162124	60S Ribosomal L35	561	2,90E ⁻⁷⁷	97,75	0000463, 0003735, 0003729, 0022625, 0006412
PUT-169a-Glycine_max-217650	Kinesin KIN-5D	250	5,67E ⁻⁵⁰	100,00	0005524, 0005871, 0008574, 0007018, 0005874, 0008017
PUT-169a-Glycine_max-116	Nuclear Factor related to Kappa-B-binding	111	2,50E ⁻¹⁵	98,40	0003677, 0031011
PUT-169a-Glycine_max-196554	Ethylene-responsive Transcription Factor RAP2-7-like Isoform X1	659	1,07E ⁻¹⁶²	95,55	0005634, 0003677, 0003700, 0032440, 0006351, 0007275, 0006355, 0055114

APÊNDICE E. Alinhamento e anotação de sequências candidatas selecionadas para a região 7, no GL E (15B).

Sequência Candidata	Descrição do Alinhamento	Comprimento (pb)	Valor Esperado	Similaridade (%)	Termos <i>GO</i>
PUT-169a-Glycine_max-240182	50S Ribosomal L7 L12	929	3,46E ⁻¹⁵	91,30	0003735, 0005840, 0006412
PUT-169a-Glycine_max-130416	ALA-interacting Subunit 3-like	83	3,81E ⁻⁷	100,00	0016021
PUT-169a-Glycine_max-38648	MYB Family Transcription Factor APL-like Isoform X3	896	2,86E ⁻¹⁵	98,00	0005634, 0003677, 0006351, 0006355
PUT-169a-Glycine_max-125728	Peroxidase 7-like	574	7,96E ⁻¹⁰⁰	93,50	0004601, 0006979, 0046872, 0098869, 0009664, 0055114, 0005576, 0009505, 0020037, 0042744
PUT-169a-Glycine_max-143391	50S Ribosomal L15-like	264	3,93E ⁻⁵⁵	98,00	0003735, 0005762, 0006412
PUT-169a-Glycine_max-55441	GMP Synthase [Glutamine-Hydrolyzing]	940	2,32E ⁻⁶⁰	97,75	0008725, 0006284
PUT-169a-Glycine_max-262240	Nuclear-pore Anchor	250	6,51E ⁻³⁹	96,40	0007094, 0006606, 0005487, 0005643
PUT-169a-Glycine_max-91892	YTH Domain Family 2	411	5,36E ⁻⁹	99,00	0003723
PUT-169a-Glycine_max-38201	Carboxyl-terminal Peptidase	150	1,29E ⁻¹⁰	100,00	0016021
PUT-169a-Glycine_max-31101	Carboxyl-terminal Peptidase	134	9,85E ⁻¹⁰	100,00	0016021
PUT-169a-Glycine_max-233254	Myosin-12 Isoform X2	208	6,54E ⁻¹⁹	97,00	0016459, 0005524, 0003779, 0003917, 0003774
PUT-169a-Glycine_max-1613271728	ATP-Citrate Synthase Alpha Chain 1-like	500	4,74E ⁻¹¹⁴	97,30	0005524, 0003878, 0016829
PUT-169a-Glycine_max-249045	40S Ribosomal SA	1240	0,00	90,35	0006407, 0000461, 0003735, 0022627, 0000447, 0000028, 0030686, 0006412
PUT-169a-Glycine_max-147271727	Major Facilitator Superfamily Domain-containing 12-like Isoform X1	1162	3,34E ⁻³³	95,60	0016021
PUT-169a-Glycine_max-185996	AAA-ATPase At4g25835-like	1190	2,41E ⁻¹⁶⁷	92,50	0005524
PUT-169a-Glycine_max-232503	PREDICTED: uncharacterized protein LOC100776077	724	4,30E ⁻⁵⁰	90,30	0016021
PUT-169a-Glycine_max-232501	PREDICTED: uncharacterized protein LOC100776077	635	1,60E ⁻⁵⁰	90,30	0016021
PUT-169a-Glycine_max-121887	Retrovirus-related Pol Poly from Transposon TNT 1-94	919	1,99E ⁻⁸²	76,60	0005488

Continuação Apêndice E					
PUT-169a-Glycine_max-109379	Citrate Mitochondrial	798	2,85E ⁻¹⁰³	95,70	0006099
PUT-169a-Glycine_max-200469	Homeobox-Leucine Zipper HOX11-like	940	4,52E ⁻¹⁴³	84,10	0005634, 0003700, 0006355, 0043565
PUT-169a-Glycine_max-201351	Proline-Rich Receptor Kinase PERK11 Isoform X2	691	4,33E ⁻⁷²	89,15	0017053, 0007049
PUT-169a-Glycine_max-96926	DNA-directed RNA Polymerases IV and V Subunit 8B-like	423	7,98E ⁻¹⁴	93,80	0001056, 0005736, 0001055, 0001054, 0006383, 0006360, 0006366, 0005666, 0005665
PUT-169a-Glycine_max-2292	DNA-binding Bromodomain	1219	0,00	81,90	0003677
PUT-169a-Glycine_max-127150	Probable Serine Threonine-Kinase DDB_G0280111	796	5,38E ⁻¹³¹	95,45	0005524, 0005737, 0016021, 0030100, 0004674, 0006468
PUT-169a-Glycine_max-126567	LRR Receptor-like Serine Threonine-Kinase	669	2,67E ⁻²⁷	88,50	0005524, 0007165, 0016021, 0004672, 0043531, 0006468, 0006952
PUT-169a-Glycine_max-136837	DUF246 Domain-containing	1121	4,96E ⁻¹⁷⁰	96,30	0016757, 0005737, 0016021
PUT-169a-Glycine_max-243882	Vicianin Hydrolase	142	6,95E ⁻²²	90,25	0016052, 0050392, 0102483
PUT-169a-Glycine_max-85469	30S Ribosomal Chloroplastic	90	1,30E ⁻¹¹	98,00	0000166, 0003723, 0005840
PUT-169a-Glycine_max-86220	30S Ribosomal Chloroplastic	97	1,28E ⁻¹⁰	98,00	0000166, 0003723, 0005840
PUT-169a-Glycine_max-1530271731	Clustered Mitochondria-like Isoform X1	2024	0,00	92,85	0003743, 0005737, 0048312, 0006413
PUT-169a-Glycine_max-909271733	SURP and G-Patch Domain-containing 1 Isoform X1	1022	4,60E ⁻¹⁴⁶	97,10	0003723, 0006396
PUT-169a-Glycine_max-263434	Zinc Finger A20 and AN1 Domain-containing Stress-Associated 5-like	101	1,27E ⁻⁹	95,85	0003677, 0008270
PUT-169a-Glycine_max-41767	Transmembrane	457	2,98E ⁻¹⁹	83,05	0006508, 0016021, 0004252
PUT-169a-Glycine_max-267516	ADP,ATP Carrier Mitochondrial-like	412	4,55E ⁻⁶³	97,70	0055085, 0003735, 0016021, 0005743, 0005215, 0006412
PUT-169a-Glycine_max-115723	ATP Synthase Subunit Mitochondrial	244	2,07E ⁻³⁴	97,95	0009507, 0008270, 0005886, 0015986, 0005753, 0046933, 0050897
PUT-169a-Glycine_max-238497	Telomerase Reverse Transcriptase	875	0,00	94,80	0003721, 0070034, 0042162, 0016021, 0007004, 0000333
PUT-169a-Glycine_max-96271733	MLP 43	689	1,07E ⁻¹¹⁰	84,50	0009607, 0006952

Continuação Apêndice E					
PUT-169a-Glycine_max-174973	Detoxification 43	574	1,20E ⁻¹²⁹	95,65	0015137, 0006855, 0015297, 0016021, 0015746, 0005886, 0015238
PUT-169a-Glycine_max-58210	Zinc Finger CCH Domain-containing 30-like	778	2,45E ⁻¹⁶⁸	90,75	0046872, 0016740
PUT-169a-Glycine_max-70171	Cellulose Synthase A Catalytic Subunit 3 [UDP-forming]-like	906	8,40E ⁻¹¹⁹	99,20	0008270, 0016021, 0016760, 0030244, 0071555, 0005886
PUT-169a-Glycine_max-153808	Photosystem II 5 kDa Chloroplastic-like	281	6,84E ⁻⁹	95,90	0016021
PUT-169a-Glycine_max-716271734	DNA Topoisomerase 2	340	5,10E ⁻⁶⁹	98,50	0009295, 0005634, 0003677, 0005524, 0000819, 0009330, 0000712, 0003918, 0006265, 0044774, 0006268, 0006312
PUT-169a-Glycine_max-211667	Primary Amine Oxidase-like	822	2,55E ⁻¹⁵⁷	95,50	0009308, 0005507, 0048038, 0008131, 0016021, 0055114
PUT-169a-Glycine_max-123527	Chloroplastic Isoform X3	245	2,25E ⁻⁴⁷	97,10	0030976, 0009063, 0070204, 0009234