

ITAMARA BOMFIM GOIS

**SELEÇÃO GENÔMICA PARA CARACTERÍSTICAS DE FRUTOS E
RESISTÊNCIA À MANCHA MARROM DE ALTERNÁRIA EM *Citrus***

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

VIÇOSA
MINAS GERAIS – BRASIL
2019

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

T

G616s
2019

Gois, Itamara Bomfim, 1986-
Seleção genômica para características de frutos e resistência
à mancha marrom de alternária em *Citrus* / Itamara Bomfim
Gois. – Viçosa, MG, 2019.
ix, 51f. : il. ; 29 cm.

Orientador: Aluizio Borém de Oliveira.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Melhoramento genético. 2. Seleção precoce. 3. BLUP.
4. Marcadores genéticos. 5. Frutas cítricas. I. Universidade
Federal de Viçosa. Departamento de Fitotecnia. Programa de
Pós-Graduação em Genética e Melhoramento. II. Título.


CDD 22. ed. 631.52

ITAMARA BOMFIM GOIS

**SELEÇÃO GENÔMICA PARA CARACTERÍSTICAS DE FRUTOS E
RESISTÊNCIA À MANCHA MARROM DE ALTERNÁRIA EM *Citrus***

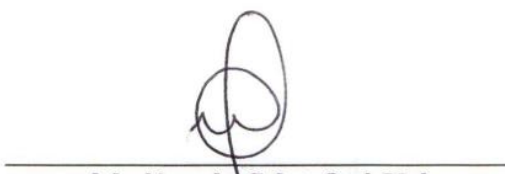
Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do programa de Pós-Graduação em Genética e Melhoramento, para a obtenção do título de *Doctor Science*.

Aprovada: 24 de abril de 2019


Maicon Nardino


Rusthon Magno Cortez dos Santos


Eveline Teixeira Caixeta


Mariângela Cristofani-Yaly
(Coorientadora)


Aluizio Borém de Oliveira
(Orientador)

“I learned that courage was not the absence of fear, but the triumph over it.”

Nelson Mandela

Aos meus pais, João e Maria.

Dedico

AGRADECIMENTOS

A Deus, por ter me dado forças para concluir mais esta etapa e por ter me presenteado com pessoas maravilhosas nesta jornada.

Aos meus pais, João e Maria, por todo amor, incentivo, ensinamentos e confiança. Devo tudo a vocês!

Aos meus irmãos, Camila e Itamar, por todo amor, apoio e amizade.

Ao meu orientador, Aluizio Borém, por ter aceitado me orientar (mais uma vez), pela confiança, ensinamentos e palavras de incentivo!

Ao meu coorientador, Marcos Deon Vilela de Resende, pelas valiosas contribuições para a execução desta pesquisa.

À minha coorientadora, Mariângela Cristofani-Yaly, por ter sido sempre tão solícita em todas as etapas desta pesquisa.

Aos membros da banca de defesa, Aluizio Borém, Eveline Caixeta, Mariângela Cristofani-yaly, Rusthon Cortez e Maicon Nardino, por toda a contribuição para a melhoria deste trabalho.

À Universidade Federal de Viçosa.

Ao Programa de Pós-graduação em Genética e Melhoramento e aos professores com os quais tive a oportunidade de adquirir valiosos conhecimentos.

Ao secretário do programa de pós-graduação em Genética e Melhoramento, Marco Túlio, por ser sempre tão prestativo e bem-humorado.

Ao Centro de Citricultura APTA Citros Sylvio Moreira/IAC, Cordeirópolis-SP, pela disponibilização dos dados para que esta pesquisa pudesse ser realizada.

As amigas que mesmo distantes se fizeram tão presentes: Adriana Uchôa, Lícia, Andrea Freire, Júlia, Ítala e Georgea.

Muito obrigada!

SUMÁRIO

LISTA DE TABELAS	vii
RESUMO	viii
ABSTRACT	ix
INTRODUÇÃO GERAL	1
REFERÊNCIAS BIBLIOGRÁFICAS	4
CAPÍTULO 1	6
SELEÇÃO GENÔMICA E CONTROLE GENÉTICO DE CARACTERES RELACIONADOS À QUALIDADE DE FRUTOS EM <i>Citrus</i>	6
RESUMO	7
ABSTRACT	8
INTRODUÇÃO.....	9
MATERIAL E MÉTODOS.....	11
<i>Caracterização da população e avaliação fenotípica</i>	11
<i>Análise dos dados fenotípicos</i>	12
<i>Genotipagem e controle de qualidade dos marcadores SNP</i>	13
<i>Predição dos valores genéticos genômicos dos indivíduos (VGG) com base nos efeitos aditivos e dominantes dos marcadores</i>	14
<i>Capacidade preditiva, vies da predição e acurácia da GWS</i>	15
<i>Eficiência da GWS</i>	16
RESULTADOS E DISCUSSÃO	18
REFERÊNCIAS BIBLIOGRÁFICAS	26
CAPÍTULO 2	31
SELEÇÃO GENÔMICA PARA RESISTÊNCIA À MANCHA MARROM DA ALTERNARIA EM <i>Citrus</i>	31
RESUMO	32
ABSTRACT	33

INTRODUÇÃO.....	34
MATERIAL E MÉTODOS.....	36
<i>População de estudo.....</i>	<i>36</i>
<i>Análise fenotípica para resistência a Alternaria alternata.....</i>	<i>36</i>
<i>Genotipagem e controle de qualidade dos marcadores.....</i>	<i>37</i>
<i>Predição dos valores genômicos dos indivíduos (VGG).....</i>	<i>38</i>
<i>Capacidade preditiva, viés da predição e acurácia da GWS.....</i>	<i>40</i>
RESULTADOS E DISCUSSÃO	41
CONCLUSÕES	46
REFERÊNCIAS BIBLIOGRÁFICAS	47
CONCLUSÕES GERAIS	51

LISTA DE TABELAS

CAPÍTULO 1	6
Tabela 1. Resultados experimentais da Seleção Genômica Ampla (GWS) para características de qualidade de frutos em uma população de melhoramento de <i>Citrus</i> . 19	
Tabela 2. Eficiência da GWS sobre a seleção fenotípica em <i>Citrus</i>	23
CAPÍTULO 2	31
Tabela 1. Estimativas dos parâmetros genéticos relacionados à resistência a mancha marrom de alternaria, <i>Alternaria alternata</i> , em uma população de <i>Citrus</i>	42
Tabela 2. Resultados experimentais da Seleção Genômica Ampla (GWS) para resistência à Mancha Marrom de alternaria, <i>Alternaria alternata</i> , em uma população de melhoramento de <i>Citrus</i>	43

RESUMO

GOIS, Itamara Bomfim, D.Sc., Universidade Federal de Viçosa, abril de 2019. **Seleção genômica para características de frutos e resistência à mancha marrom de alternária em *Citrus***. Orientador: Aluizio Borém de Oliveira. Coorientadores: Marcos Deon Vilela de Resende e Mariângela Cristofani-Yaly.

A seleção genômica ampla (*Genome Wide Selection* - GWS) é uma seleção realizada com base em marcadores moleculares e seus efeitos genéticos e foi proposta como uma alternativa para aumentar a eficiência dos programas de melhoramento. Este trabalho foi realizado com o objetivo de avaliar a aplicabilidade da GWS no melhoramento genético de duas populações de citros, a primeira avaliada para características relacionadas à qualidade de frutos e a segunda avaliada para resistência à mancha marrom de alternaria. Para tanto, ambas as populações foram genotipadas com marcadores SNP. Os valores genéticos genômicos dos indivíduos foram preditos por meio do método BLUP Genômico (G-BLUP), implementado via o método RKHS (*Reproducing Kernell Hilbert Spaces*) Bayesiano. Para as características relacionadas à qualidade de frutos, observou-se que a GWS pode ser superior à seleção fenotípica, em termos de ganho genético por unidade de tempo. Para a característica mancha marrom de alternaria (MMA), observou-se que o aumento da acurácia está diretamente relacionado com a utilização de marcadores em desequilíbrio de ligação com o gene que confere a resistência à MMA. A GWS apresenta potencial para ser utilizada como ferramenta no melhoramento genético de citros.

ABSTRACT

GOIS, Itamara Bomfim, D.Sc., Universidade Federal de Viçosa, April, 2019. **Genomic selection for fruit characteristics and *Alternaria* brown spot resistance of *Citrus*.** Adviser: Aluizio Borém de Oliveira. Co-advisers: Marcos Deon Vilela de Resende and Mariângela Cristofani-Yaly.

Genome Wide Selection (GWS) is a selection based on molecular markers and their genetic effects, and was proposed as an alternative to increase the efficiency of breeding programs. This work was carried out with the objective of evaluating the applicability of GWS to the genetic improvement of two populations of citrus. The first one was evaluated for traits related to fruit quality and the second one was evaluated for resistance to *Alternaria* brown spot (ABS). The populations were genotyped with SNP markers. The genomic breeding values of the individuals belonging to each population were predicted using the BLUP genomic method (G-BLUP), implemented via the Bayesian RKHS (Reproducing Kernell Hilbert Spaces) method. For the traits related to fruit quality, it was observed that GWS may be superior to the phenotypic selection in terms of genetic gain per unit of time. For the trait ABS it was observed that the increase of accuracy is directly related to the use of markers in linkage disequilibrium with the gene that confers resistance to MMA. The GWS presents potential to be used in the citrus breeding.

INTRODUÇÃO GERAL

A citricultura é um dos mais tradicionais setores da agricultura brasileira, tendo segmentos bem estruturados como o dos viveiristas, produtores, fornecedores de insumos e maquinário, fábricas para o processamento de suco e para o mercado de produtos frescos, entre outros (MACHADO et al., 2011). Além da organização que apresenta a produção de citros, o investimento em pesquisa tem proporcionado ao Brasil desempenho de destaque em nível mundial. O Brasil é o segundo maior produtor mundial de frutas cítricas (13,50%) e o maior produtor mundial de laranja (23,81%) (FAO, 2017).

Os citros compreendem um vasto número de espécies que pertencem ao gênero *Citrus* e gêneros afins, como *Fortunella* e *Poncirus*, além dos híbridos provenientes do cruzamento entre espécies do mesmo gênero e intergêneros. Embora exista uma alta diversidade genética dentro deste grupo, apenas um pequeno grupo de espécies e clones são utilizados como variedades comerciais (MACHADO et al., 2011).

A variabilidade genética de citros existente no Brasil tem sido conservada em Bancos Ativos de Germoplasma (BAG), como o observado na Embrapa Mandioca e Fruticultura Tropical e no Instituto Agrônomo de Campinas. Esta diversidade tem sido utilizada para dar suporte aos programas de melhoramento genético da cultura.

Quando se reporta ao melhoramento de citros é importante ressaltar que a planta cítrica é constituída de duas plantas, a parte aérea e o porta-enxerto, sendo necessários estudos direcionados para o melhoramento das duas plantas, como também para a interação entre elas (MACHADO, 1997). Consequentemente, a seleção fenotípica torna-se onerosa e demorada pois, por se tratar de uma espécie perene, apresenta ciclo de vida longo e demanda grandes áreas para as avaliações. Adicionalmente, o processo de seleção tende a ser mais demorado quando as características avaliadas são expressas em estádios tardios do desenvolvimento da planta (RESENDE JUNIOR et al., 2012).

A utilização de marcadores moleculares no melhoramento genético, com o intuito de aumentar os ganhos genéticos e facilitar o processo de seleção, foi proposta por Lande e Thompson (1990). A seleção assistida por marcadores (MAS) utiliza simultaneamente dados fenotípicos e de marcadores moleculares que estejam em ligação gênica com alguns locos controladores de características quantitativas (QTL) ou qualitativas. Estes marcadores são eleitos ou não como determinantes dos efeitos dos locos após modelagem estatística (RESENDE et al., 2012). A justificativa para a utilização de informações em nível de DNA no melhoramento de plantas é que estas podem proporcionar ganhos

genéticos com maior rapidez, maior eficiência seletiva e baixo custo quando comparada à seleção baseada somente em dados fenotípicos (MEUWISSEN, 2001; RESENDE et al., 2012).

A utilização de marcadores baseados em DNA possibilitou a construção de dezenas de mapas de ligação para a cultura dos citros. A construção destes mapas é a única forma de cobrir o genoma, sem o uso de sequenciamento, e é de suma importância para o entendimento dos mecanismos de herança para resistência ou susceptibilidade a fatores bióticos (ex. resistência ao vírus-da-tristeza-dos-citros, resistência à mancha marrom da alternaria, ao nematoide *Tylenchulus semipenetrans*, à gomose causada por *Phytophthora*) e abióticos (ex. tolerância a sais e ao frio), e para a localização de genes ligados às características que apresentam importância agrônômica (MACHADO et al., 2011; OLIVEIRA et al., 2014).

O desafio da seleção assistida por marcadores (SAM) advém da herança complexa dos caracteres quantitativos. Na cultura dos citros, esta estratégia tem sido adotada para diversas características de interesse econômico (ASINS et al., 2012; CURTOLO et al., 2017; GUSSEN et al., 2011; SIVIERO et al., 2002; SIVIERO et al., 2006; WEBER et al., 2003). No entanto, observa-se que a identificação de QTL ligados a essas características tem explicado apenas uma pequena fração da variação genética do caráter, o que inviabiliza o processo de seleção baseado nessas marcas.

Atualmente, a disponibilidade de marcadores moleculares que cobrem o genoma de maneira densa propiciou a introdução da seleção genômica ampla (GWS) no melhoramento genético (MISZTAL et al., 2009). A GWS foi proposta por Meuwissen et al. (2001) como uma alternativa para aumentar a eficiência dos programas de melhoramento genético, e consiste na seleção simultânea de centenas ou milhares de marcadores, de forma que todos os genes de um caráter quantitativo estejam em desequilíbrio de ligação com pelo menos uma parte dos marcadores (RESENDE, 2007).

Em programas de melhoramento genético, a GWS é uma alternativa promissora para a SAM, por permitir o estudo de características poligênicas, geralmente controlada por muitos genes de pequenos efeitos (KUMAR et al., 2012) e por permitir a estimação dos valores genéticos genômicos dos indivíduos sem informações fenotípicas (MOSER et al., 2009).

As propriedades que tornam a GWS promissora para o uso nos programas de melhoramento genético são: a não necessidade de mapeamento de QTL; a utilização de uma alta densidade de marcadores, o que aumenta a probabilidade dos marcadores

estarem em desequilíbrio de ligação com os QTL; utilização do real parentesco genético entre os indivíduos; o uso de um maior número de informações, como a fenotípica, a genotípica e a genealógica; possibilitar, de forma eficiente, a seleção de caracteres que apresentam baixa herdabilidade; permitir a melhor utilização dos recursos genéticos disponíveis por meio da seleção de cruzamentos de acordo com as informações genéticas; e, aumentar o ganho genético por unidade de tempo, pois propicia a seleção precoce direta (DAETWYLER et al., 2013; LEGARRA et al., 2008; MEUWISSEN et al., 2001; RESENDE et al., 2012).

Diante do exposto, esta pesquisa foi realizada com o objetivo de verificar a aplicabilidade da GWS em características relacionadas à qualidade de frutos e à resistência à mancha marrom da alternaria em citros.

REFERÊNCIAS BIBLIOGRÁFICAS

ASINS, M.J.; FERNANDEZ-RIBACOBA, J.; BERNET, G.P.; GADEA, J.; CAMBRA, M.; GORRIS, M.T.; CARBONELL, E.A. The position of the major QTL for Citrus tristeza virus resistance is conserved among *Citrus grandis*, *C. aurantium* and *Poncirus trifoliata*. **Molecular Breeding** 29:575–587, 2012.

CURTOLO, M.; CRISTOFANI-YALY, M.; GAZAFFI, R.; TAKITA, M.A.; FIGUEIRA, A.; MACHADO, M.A. QTL mapping for fruit quality in *Citrus* using DArTseq markers. **BMC Genomics**, 18: 289, 2017.

DAETWYLER, A.D.; CALUS, M.P.L.; PONG-WONG, R.; DE LOS CAMPOS, G.; HICKEY, J.M. Genomic Prediction in Animals and Plants: Simulation of Data, Validation, Reporting, and Benchmarking. **Genetics** 193: 347–365, 2013.

FAO - **Food and Agriculture Organization of the United**. Disponível em: <http://faostat.fao.org/>; Acessado em: 08 de abril de 2019.

GUSSEN, O.; UZUN, A.; SEDAY, U.; KAFA, G. QTL analysis and regression model for estimating fruit setting in young *Citrus* trees based on molecular markers. **Scientia Horticulturae** 130(2), p. 418-424, 2011.

KUMAR, S.; BINK, M.C.A.M.; VOLZ, R.K.; BUS, V.G.M.; CHGNÉ, D. Towards genomic selection in apple (*Malus × domestica* Borkh.) breeding programmes: Prospects, challenges and strategies. **Tree Genetics & Genomes**, 8:1–14, 2012.

LANDE, R.; THOMPSON, R. Efficiency of marker-assisted selection in the improvement of quantitative traits. **Genetics** 124: 743-756, 1990.

LEGARRA, A.; ROBERT, GRANIÉ, C.; MANFREDI, E.; ELSSEN, J.M. Performance of genomic selection in mice. **Genetis** 180: 611-618, 2008.

MACHADO, M.A. Melhoramento de citros. In: Simpósio sobre atualização em genética e melhoramento de plantas. **Anais**, 1997, 297p.

MACHADO, M.A.; CRISTOFANI-YALY, M.; BASTIANEL, M. Breeding, genetic and genomic of citrus for disease resistance. **Rev. Bras. Frutic.**, Volume Especial: 158-172, 2011.

MEUWISSEN, T.H.E.; HAYES, B.J.; GODDARD, M.E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. **Genetics** n. 157, p. 1819–1829, 2001.

MISZTAL I, LEGARRA A, AGUILAR I. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. **J. Dairy Sci.** 92: 4648–4655, 2009.

MOSER, G.; TIER, B. CRUMP, R.E.; KHATKAR, M.S.; RAADSMA, H.W. A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide SNP markers. **Genetics Selection Evolution** 41(56): 1-16, 2009.

OLIVEIRA, A.P.; SOARES FILHO, W.S.; MACHADO, M.A.; FERREIRA, E.A.; SCIVITTARO, W.B.; GESTEIRA, A.S.; Melhoramento genético de plantas cítricas. Informe Agropecuário, 35(281): 22-29, 2014.

RESENDE MDV, RESENDE JUNIOR MFR, SANSALONI CP, PETROLI CD. Genomic selection for growth and wood quality in Eucalyptus: capturing the missing heritability and accelerating breeding for complex traits in forest trees. **New Phytologist**, 194: 116-128, 2012.

RESENDE JUNIOR MFR, MUÑOZ P, RESENDE MDV, GARRICK DJ. Accuracy of Genomic Selection Methods in a Standard Data Set of Loblolly Pine (*Pinus taeda* L.). **Genetics**, 190: 1503–1510, 2012.

RESENDE, M.D.V. **Seleção Genômica Ampla (GWS) e Modelos Lineares Mistos**. In: RESENDE, M.D.V. Matemática e Estatística na Análise de Experimentos. Colombo: Embrapa Florestas, 2007. 362p. Cap. 11. p. 517-533.

SIVIERO, A.; CRISTOFANI, M.; BOAVA, L.P.; MACHADO, M.A. Mapeamento de QTLs associados à produção de frutos e sementes em híbridos de *Citrus sunki* vs. *Poncirus trifoliata*. **Revista Brasileira de Fruticultura** 24: 741-743, 2002.

SIVIERO, A.; CRISTOFANI, M.; FURTADO, E.L.; GARCIA, A.A.F.; COELHO, A.S.G.; MACHADO, M.A. Identification of QTLs associated with citrus resistance to *Phitophthora gummosis* **Journal of Applied Genetics** 47(1), p. 23-28, 2006.

WEBER, C.A.; MOORE, G.A.; DENG, Z.; GMITTER JUNIOR, F.G. Mapping freeze tolerance quantitative trait loci in a *Citrus grandis* x *Poncirus trifoliata* F1 pseudo-testcross using molecular markers. **Journal Amer. Soc. Hort. Sci.** 128(4): 508-514, 2003.

CAPÍTULO 1
SELEÇÃO GENÔMICA E CONTROLE GENÉTICO DE CARACTERES
RELACIONADOS À QUALIDADE DE FRUTOS EM *Citrus*

Viçosa – MG
Abril, 2019

RESUMO

GOIS, Itamara Bomfim, D.Sc., Universidade Federal de Viçosa, Abril, 2019. **Seleção genômica e controle genético de caracteres relacionados à qualidade de frutos em *Citrus***. Orientador: Aluizio Borém de Oliveira. Coorientadores: Marcos Deon Vilela de Resende; Mariângela Cristofani-Yaly.

Os citros estão entre as fruteiras mais plantadas, consumidas e pesquisadas do mundo devido à sua importância econômica e social. Os programas de melhoramento de citros apresentam uma série de desafios de ordem genética e/ou botânica. Além disso, por serem perenes, demandam grandes áreas para a condução de experimentos e longos períodos para avaliação. A seleção genômica ampla (*Genome Wide Selection - GWS*) é uma seleção realizada com base em marcadores moleculares e seus efeitos genéticos, e foi proposta como uma alternativa para aumentar a eficiência dos programas de melhoramento. Este trabalho foi realizado com o objetivo de avaliar a eficiência da GWS no melhoramento genético de citros. Uma população de citros proveniente do cruzamento entre laranja Pêra (*Citrus sinensis* Osbeck) e Tangor “Murcott” (*Citrus sinensis* Osbeck X *Citrus reticulata* Blanco) foi avaliada para nove características quantitativas, relacionadas à qualidade de frutos, e genotipadas com 16.618 marcadores SNP. Os valores genéticos genômicos dos indivíduos foram preditos por meio do método BLUP Genômico (G-BLUP), implementado via o método RKHS (*Reproducing Kernell Hilbert Spaces*) Bayesiano. A capacidade preditiva da GWS variou de 0,33 a 0,48 e a acurácia da predição variou de 0,33 a 0,54. A fração da herdabilidade capturada pelos marcadores foi superior a 48%. As estimativas de acurácia da seleção fenotípica foram superiores àquelas obtidas por meio da GWS, no entanto, foi possível observar a superioridade da GWS em termos de ganho genético por unidade de tempo. Portanto, a GWS apresenta potencial para ser utilizada como ferramenta no melhoramento genético para qualidade de frutos em *Citrus*, devido à possibilidade de redução do ciclo seletivo, via seleção genômica.

Palavras-chave: ganho com a seleção; SNP; acurácia; seleção precoce.

ABSTRACT

GOIS, Itamara Bomfim, D.Sc., Universidade Federal de Viçosa, April, 2019. **Genomic selection and genetic control of fruit quality characters in *Citrus***. Adviser: Aluizio Borém de Oliveira. Co-advisers: Marcos Deon Vilela de Resende; Mariângela Cristofani-Yaly.

Citrus are among the world's most planted, consumed and researched fruit trees because of their economic and social importance. Citrus breeding programs present genetic and / or botanical challenges. In addition, because they are perennial, they require large areas for conducting experiments and long periods for evaluation. Genome Wide Selection (GWS) is a selection based on molecular markers and their genetic effects, and was proposed as an alternative to increase the efficiency of breeding programs. This work was carried out with the objective of evaluating the efficiency of GWS in the genetic improvement of citrus. A citrus population from the cross between orange Pera (*Citrus sinensis* Osbeck) and Tangor "Murcott" (*Citrus sinensis* Osbeck X *Citrus reticulata* Blanco) was evaluated for nine quantitative traits, related to fruit quality, and genotyped with 16,618 SNP markers. The genomic breeding values of the individuals were predicted by BLUP genomic method (G-BLUP), implemented via the Bayesian RKHS (Reproducing Kernell Hilbert Spaces) method. The predictive capacity of GWS ranged from 0.33 to 0.48 and the accuracy of prediction ranged from 0.33 to 0.54. The fraction of heritability captured by the markers was greater than 48%. The estimates of accuracy of the phenotypic selection were higher than those obtained through GWS, however, it was possible to observe the superiority of GWS, in terms of genetic gain per unit of time. Therefore, the GWS presents potential to be used as a tool in the genetic improvement for fruit quality in citrus, due to the possibility of reduction of the selective cycle, through genomic selection.

Key-words: gain from selection; SNP; accuracy; early selection.

INTRODUÇÃO

Os citros apresentam uma grande diversidade de espécies, variedades e clones e, devido à sua importância econômica e social, estão entre as fruteiras mais plantadas, consumidas e pesquisadas do mundo (MACHADO et al., 2011; OLIVEIRA et al., 2014). Neste cenário, o Brasil se destaca como o segundo maior produtor de frutas cítricas, sendo responsável por aproximadamente 15,5% da produção mundial, e como maior produtor de laranja, sendo responsável por aproximadamente 26,4% (FAO, 2016).

O melhoramento genético de citros tem sido realizado por diversas instituições de pesquisa no Brasil, dentre elas, o Instituto Agrônomo de Campinas (IAC), com o intuito de aumentar a produtividade e a qualidade dos produtos cítricos (GOIS et al., 2016). Vale salientar que o melhoramento genético de citros é realizado para as cultivares utilizadas como porta-enxerto e para aquelas utilizadas como copa, além da interação entre as mesmas (MACHADO, 1997; OLIVEIRA et al., 2014).

Os programas de melhoramento tradicionais de citros que utilizam hibridação e seleção recorrente apresentam uma série de dificuldades, seja de ordem genética (incompatibilidade; longo ciclo de seleção; baixa diversidade biológica presente nas populações de seleção; combinação interespecífica envolvendo porta-enxerto e copa) ou biológica (juvenilidade; apomixia) (MACHADO et al., 2011; TEROLL et al., 2007; TEROLL et al., 2008). Além disso, por serem perenes, os citros demandam grandes áreas para a condução de experimentos e longos períodos para avaliação. Frente a estas dificuldades, os programas de melhoramento de citros tornam-se morosos e dispendiosos, principalmente quando as características de interesse são expressas em estádios tardios do desenvolvimento da planta.

Os marcadores moleculares têm sido utilizados em diferentes abordagens visando o aumento da eficiência dos programas de melhoramento genético de espécies perenes. A justificativa para a utilização da genética molecular no melhoramento de plantas é que o uso de informação em nível de DNA permitirá ganhos mais rápidos e maior eficiência seletiva, quando comparado à seleção baseada apenas em dados fenotípicos (MEUWISSEN, 2001; RESENDE et al., 2012).

A disponibilidade de plataformas de genotipagem de alto rendimento vem modificando as estratégias utilizadas nos programas de melhoramento genético de fruteiras (IWATA et al., 2016; KUMAR et al., 2012). Em 2001, Meuwissen et al. propuseram a seleção genômica ampla (*Genome Wide Selection* - GWS) como uma

alternativa para aumentar a eficiência dos programas de melhoramento genético. A GWS consiste na seleção simultânea de centenas ou milhares de marcadores de forma que a maioria dos genes de um caráter quantitativo, ou qualitativo, estejam em desequilíbrio de ligação com pelo menos uma parte dos marcadores (CROS et al., 2015; MEUWISSEN et al., 2001; MEUWISSEN, 2007; RESENDE et al., 2008). Os marcadores em desequilíbrio de ligação com os QTL de grandes e pequenos efeitos explicarão quase a totalidade da variação genética de uma característica (CALUS et al., 2008; RESENDE et al., 2008).

A GWS enfatiza a estimação e a predição dos efeitos dos marcadores e dos valores genômicos dos indivíduos em avaliação sem o uso de testes de significância para a seleção de marcas. Assim, torna-se possível a explicação de maior parte da variância genética do caráter por evitar a chamada herdabilidade faltante, típica de estudos de ligação e associação (VIANA et al., 2017; ZAPATA-VALENZUELA et al., 2012). A herdabilidade faltante refere-se à diferença entre a herdabilidade obtida com base nos fenótipos e a herdabilidade explicada pelos marcadores moleculares. Além disso, a predição dos valores genéticos tende a ser mais acurada do que a seleção baseada apenas em fenótipos, pois a GWS explora com maior precisão a variação criada pela segregação mendeliana durante a formação dos gametas (RESENDE, 2007; ZAPATA-VALENZUELA et al. 2013).

A GWS apresenta elevado potencial para programas de melhoramento genético de espécies perenes, visto que estas demandam mais tempo e recursos financeiros para a experimentação e, portanto, métodos e estratégias mais eficientes para o processo de seleção das características que apresentam importância econômica. Diante do exposto, este trabalho foi realizado com o objetivo de avaliar a aplicabilidade da GWS no melhoramento genético de citros e a sua eficiência quando comparada com a seleção baseada em fenótipos.

MATERIAL E MÉTODOS

Caracterização da população e avaliação fenotípica

As estimativas dos efeitos genéticos dos marcadores foram obtidas para uma população de citros formada a partir do cruzamento entre Tangor “Muscott” (*Citrus sinensis* Osbeck X *Citrus reticulata* Blanco) e laranja Pêra (*Citrus sinensis* Osbeck). Para a formação da população, composta por 180 indivíduos (híbridos), foram utilizados dois genitores (Tamanho efetivo = 2).

A população avaliada está localizada no Centro de Citricultura Sylvio Moreira do Instituto Agrônomo de Campinas (IAC), em Cordeirópolis – SP. O experimento foi instalado em 2004 em delineamento inteiramente ao acaso (DIC), com três repetições clonais e espaçamento 6 x 3 m.

As avaliações fenotípicas foram realizadas no ano de 2012. Uma amostra de cada repetição (clone), composta por cinco frutos, foi utilizada para a obtenção das seguintes características: massa do fruto (g), obtida em uma balança Filizola com capacidade para 15 kg e precisão de 5 g; diâmetros longitudinal e transversal dos frutos (cm), obtidos por leitura direta com o auxílio de uma régua graduada; número de gomos e de sementes, obtidos por meio da contagem direta; rendimento de suco (mL), determinado após esmagamento do fruto em extratora OIC modelo OTTO 1800 e calculado por meio da relação massa do suco/massa do fruto; teor de sólidos solúveis (°Brix), determinado por leitura direta em refratômetro B&S modelo RFM 330; *ratio*, calculado por meio da relação teor de sólidos solúveis/ acidez; e, número de frutos/caixa, obtido dividindo-se 40,8 Kg (capacidade da caixa) pelo peso médio dos frutos.

As nove características avaliadas são classificadas como de natureza quantitativa e foram selecionadas para este estudo devido à importância comercial. A fenotipagem foi realizada no Laboratório de Qualidade e Pós-Colheita do Centro APTA Citros Sylvio Moreira/IAC, Cordeirópolis-SP.

Análise dos dados fenotípicos

As estimativas da herdabilidade e da acurácia seletiva foram obtidas a partir da metodologia REML/BLUP, também denominada metodologia de modelos mistos (HENDERSON, 1973). O *software* Selegen REML/BLUP (RESENDE, 2016) foi utilizado para a realização das análises, sendo as estimativas obtidas por meio do modelo estatístico:

$$y = 1u + Zg + e$$

Em que, y é o vetor de dados fenotípicos; u é o escalar referente à média geral (efeito fixo); g é o vetor dos efeitos genotípicos (assumidos como aleatórios); e e é o vetor de resíduos (aleatórios).

O número 1 representa um vetor de “1s”; e a letra maiúscula Z representa a matriz de incidência para os efeitos genotípicos. As estruturas de médias e variâncias associadas ao modelo são: $y|u, V \sim N(1u, V)$; $g|G, \sigma_g^2 \sim N(0, G\sigma_g^2)$; $e|\sigma_e^2 \sim N(0, I\sigma_e^2)$; e $Cov(g, e) = 0$. Em que V é a matriz de covariância fenotípica, dada por $ZGZ' + I\sigma_e^2$; I é uma matriz identidade; σ_g^2 e σ_e^2 são as variâncias genotípica e residual, respectivamente.

As equações de modelos mistos para o BLUP (*Best Linear Unbiased Predictor*) dos valores genéticos individuais equivalem a:

$$\begin{bmatrix} \hat{u} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} 1'1 & 1'Z \\ Z'1 & Z'Z + G^{-1}\lambda \end{bmatrix}^{-1} \begin{bmatrix} 1'y \\ Z'y \end{bmatrix}$$

Em que, $\lambda = \sigma_e^2 / \sigma_g^2 = \frac{1-h^2}{h^2}$: parâmetro de penalização, associado ao fator de *shrinkage*; $h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2}$: herdabilidade individual no sentido amplo; e G : Matriz de correlação genotípica entre os indivíduos em avaliação.

Os componentes de variância foram estimados via REML (*Restricted Maximum Likelihood* – HENDERSON, 1973) a partir de iterações nas equações de modelos mistos. A herdabilidade ao nível de média de repetições foi calculada a partir da herdabilidade individual por meio da expressão (RESENDE et al., 2001): $h_{am}^2 = \frac{nh^2}{1+(n-1)h^2}$, em que n

equivale ao número de repetições; e a acurácia da predição dos valores genotípicos dos indivíduos avaliados pela expressão (RESENDE et al., 2001): $\hat{r}_{y\hat{y}} = \sqrt{\hat{h}_{am}^2}$.

Genotipagem e controle de qualidade dos marcadores SNP

A genotipagem da população de *Citrus* foi realizada por meio de marcadores SNP (*Single Nucleotide Polymorphism*). O DNA total de cada um dos 180 indivíduos avaliados, foi extraído, quantificado em Espectrofotômetro Nanodrop 8000, normalizado para 100 ng/ µl, distribuído em placas e enviado para a plataforma de genotipagem. Os marcadores SNP foram identificados por meio da plataforma de genotipagem *DArTseq* desenvolvida por *Diversity Arrays Technology* (Yarralumla, Austrália).

Um total de 16.618 marcadores SNP foram obtidos, com base no alinhamento de sequências do genoma de referência (*Citrus reticulata*). Os indivíduos genotipados foram codificados como 0, 1 e 2, em cada loco, de acordo com o número de cópias do alelo variante que eles possuíam. Quando o genótipo foi homocigoto para o alelo de referência, este foi codificado como 0; quando o genótipo foi heterocigoto, ou seja, apresentou o alelo de referência e o alelo variante, este foi codificado como 1; e, quando o genótipo foi homocigoto para o alelo variante, este foi codificado como 2.

O controle de qualidade dos marcadores SNP foi realizado por meio da eliminação dos locos com baixa taxa de atendimento na genotipagem dos indivíduos (*call rate* ≤ 95%); e de *loci* com baixo polimorfismo (*MAF* – *minor allele frequency* ≤ 5%). O nível crítico adotado para a *MAF* foi obtido por meio da equação $MAF = 1/\sqrt{2N}$ (Resende et al., 2017), em que N se refere ao número de indivíduos avaliados. A *MAF* é dada por: $MAF = \min[p, (1 - p)]$, em que *p* é a frequência do alelo de referência *i*. A *Call Rate* foi calculada por meio da expressão: $Call Rate = 1 - \frac{A}{P}$, em que *A* refere-se aos indivíduos com genótipos perdidos (*missing*) e *P* refere-se aos indivíduos com genótipos presentes (*not missing*) (RESENDE et al., 2014). Após o controle de qualidade, a imputação de valores perdidos (NA) foi realizada com o valor médio (2p). As análises de controle de qualidade foram conduzidas no *software* R (R Core Team, 2018).

Predição dos valores genéticos genômicos dos indivíduos (VGG) com base nos efeitos aditivos e dominantes dos marcadores

A parametrização das matrizes de incidência de marcadores foi realizada de acordo com Resende et al. (2014). Os valores da matriz W , associada aos efeitos genéticos aditivos (a), foram centrados em zero para que os efeitos dos marcadores sejam efeitos de substituição alélica (α) com média zero na população. Assim, a parametrização utilizada na matriz W foi: “ $2q$ ”, “ $q - p$ ” e “ $-2p$ ” para os genótipos 2, 1 e 0, respectivamente. A parametrização utilizada para a matriz S , associada aos efeitos genéticos devido aos desvios de dominância (d), foi: “ $-2q^2$ ”, “ $2pq$ ” e “ $-2p^2$ ”, para os valores 2, 1 e 0, respectivamente.

O método GBLUP (*Genomic Best Linear Unbiased Prediction*) via procedimento RKHS (*Reproducing Kernel Hilbert Spaces*) com algoritmo Bayesiano (GIANOLA, 2006) foi utilizado para a realização das análises de Seleção Genômica (GWS) para as nove características avaliadas.

O modelo adotado para predição dos valores genômicos individuais foi dado por: $y = 1\mu + Wa + Sd + e$, em que y é o vetor de observações fenotípicas; μ é a média geral; a é o vetor de efeitos genéticos aditivos dos indivíduos; d é o vetor de efeitos genéticos devido aos desvios de dominância dos indivíduos; e é o vetor de resíduos aleatórios; 1 é um vetor de 1's; W e S são as matrizes de incidência para a e d , respectivamente.

As equações de modelos mistos genômicas associadas ao método GBLUP/ GWS foram dadas por:

$$\begin{bmatrix} 1'1 & 1'W & 1'S \\ W'1 & W'W + G_a^{-1} \frac{\sigma_e^2}{\sigma_a^2} & W'S \\ S'1 & S'W & S'S + G_d^{-1} \frac{\sigma_e^2}{\sigma_d^2} \end{bmatrix} \begin{bmatrix} \hat{u} \\ \hat{a} \\ \hat{d} \end{bmatrix} = \begin{bmatrix} X'y \\ W'y \\ S'y \end{bmatrix}$$

Em que, $G_a = \frac{WW'}{\sum 2p_i q_i}$ e $G_d = \frac{SS'}{\sum (2p_i q_i)^2}$ são as matrizes de parentesco genômico aditivo e de dominância, respectivamente; e p_i e q_i são as frequências alélicas.

O método GBLUP equivale ao método RKHS (*Reproducing Kernel Hilbert Spaces*) quando a matriz dos *kernels* K é substituída pela matriz de parentesco genômico (G). Essa equivalência é demonstrada na sequência.

Na regressão RKHS a estrutura de covariância é proporcional a uma matriz de kernel K , dada por $\text{Cov}(g_i, g_j) \propto K_{\text{RKHS}}(w_i, w_j)$, em que w_i, w_j são vetores de genótipos marcadores para os indivíduos i e j , respectivamente, e $K(\dots)$ é uma função positiva definida avaliada nos genótipos marcadores.

O modelo RKHS pode ser também assim especificado: $y = 1u + K_h \alpha + e$, em que u é uma constante, K_h é a matriz positiva definida de *kernels*, dependente do parâmetro de suavização h ; α é um vetor contendo coeficientes não paramétricos que são assumidos com distribuição normal $\alpha_j \sim N(0, K_h^{-1} \sigma_\alpha^2)$, com σ_α^2 representando a recíproca do parâmetro de alisamento ($\sigma_\alpha^2 = \lambda^{-1}$). Os resíduos têm distribuição normal com matriz de covariância $R=I \sigma_e^2$. A solução para α é dada por $[\sigma_e^{-2} K_h + \sigma_\alpha^{-2} I] \hat{\alpha} = \sigma_e^{-2} y$.

Os fenótipos são preditos por $\hat{y} = \hat{u}1 + K_h^* \hat{\alpha}$, onde uma linha de K_h^* tem a forma $K_i^* = [K_h^*(w_i - w_j)]$, com $K_h^*(w_i - w_j)$ sendo o kernel entre o genótipo do indivíduo i no grupo de validação e o genótipo do indivíduo j no grupo de estimação.

O modelo individual univariado tradicional $y = Xb + Za + e$ pode ser expresso em termos de $y = g + e$ em que $g|0, K_{\text{RKHS}} \sigma_\alpha^2 \sim N(0, K_{\text{RKHS}} \sigma_\alpha^2)$, conduzindo ao estimador $[\sigma_e^{-2} I + \sigma_\alpha^{-2} K_{\text{RKHS}}^{-1}] \hat{g} = \sigma_e^{-2} y$ (Campos et al., 2009). Assim, substituindo-se K por G no método RKHS, tem-se RKHS = GBLUP.

Dessa forma, utilizou-se do método RKHS com algoritmo Bayesiano (GIANOLA, 2006) para a realização das análises de Seleção Genômica (GWS) para as nove características avaliadas. Para tanto, utilizou-se o pacote BGLR (PEREZ e DE LOS CAMPOS, 2014), implementado no *software* R (R Core Team, 2018), e a matriz K foi substituída pela matriz G .

Capacidade preditiva, viés da predição e acurácia da GWS

A capacidade preditiva da GWS, o viés da predição e a acurácia da seleção genômica foram computados a partir dos valores genômicos preditos, dos valores fenotípicos e da herdabilidade.

Para o cômputo da capacidade preditiva ($r_{y\hat{g}}$) foi utilizada a validação cruzada, por meio da reamostragem com reposição de um grupo de indivíduos via procedimento

k-fold. Os dados da população estudada foram divididos em dois subconjuntos, 90% dos indivíduos foram selecionados aleatoriamente para compor a população de treinamento, a partir da qual se desenvolveu as equações de predição, e o restante, 10%, foram usados para compor a população de validação. Esta análise foi repetida 10 vezes (10-fold) e, a cada vez os indivíduos utilizados para a predição e para a validação foram amostrados aleatoriamente e analisados independentemente. O número de iterações utilizado foi 100.000, com *burn in* de 10.000. A capacidade preditiva ($r_{y\hat{g}}$) foi estimada a partir da média dos coeficientes de correlação entre os valores fenotípicos e os valores genômicos preditos.

O viés da predição (\hat{b}) foi calculado por meio do coeficiente de regressão dos valores fenotípicos sobre os valores genômicos preditos. Quando b assume valores próximos a 1, indica que as avaliações não são viesadas e são efetivas em prever as reais magnitudes das diferenças entre os indivíduos em avaliação (RESENDE et al., 2012).

A estimativa da acurácia da seleção genômica foi obtida por meio da expressão: $r_{g\hat{g}} = \frac{r_{y\hat{g}}}{\sqrt{h^2}}$, em que $r_{y\hat{g}}$ equivale à capacidade preditiva da GWS e h^2 à herdabilidade do caráter (LEGARRA et al., 2008). A acurácia esperada foi calculada por meio da fórmula proposta por Resende (2008): $r_{ggE} = \sqrt{\frac{r_{mq}^2(Nh_m^2)}{1+[(N-1)h_m^2]}}$, em que, r_{mq}^2 é a proporção da variância genética explicada pelos marcadores dada por: $r_{mq}^2 = \frac{n_m}{n_m+M_e}$ (GODDARD et al. 2011); N é o número de indivíduos genotipados e fenotipados; h_m^2 é a herdabilidade individual de um loco, dada por: $h_m^2 = \frac{r_{mq}^2 h^2}{M_e}$; h^2 é a herdabilidade do caráter; e M_e é o número de segmentos cromossômicos independentes, dado por: $M_e = 2NeL$ (HAYES et al., 2009); Ne é o tamanho efetivo da população; e L é o tamanho do genoma em Morgans.

O número estimado de QTL controlando cada característica foi calculado por meio da expressão: $n_{QTL} = \frac{Nr_{mq}^2 h^2 (r_{mq}^2 - r_{gg}^2)}{r_{gg}^2}$. Esta expressão foi obtida a partir da fórmula de acurácia proposta por Resende et al. (2008), e o parâmetro “ r_{mq}^2 ” foi estimado a partir do número de marcadores utilizados na análise de GWS.

Eficiência da GWS

Para espécies perenes, a superioridade da GWS em relação à seleção baseada apenas em dados fenotípicos está, principalmente, relacionada com o aumento do ganho

com a seleção por unidade de tempo (RESENDE et al., 2012). Para verificar este fato, foi calculada a eficiência seletiva da GWS em relação à seleção baseada apenas em dados fenotípicos por meio da seguinte expressão: $Ef = \frac{r_{gg}t_f}{r_{yy}t_{GWS}}$, em que $r_{g\hat{g}}$ é a acurácia da GWS; r_{yy} é a acurácia da seleção com base apenas em fenótipos; t_f é o tempo médio requerido para o ciclo de seleção com base em fenótipos (Foi considerado um ciclo de seleção de 20 anos); e t_{GWS} é o tempo médio requerido para o ciclo de seleção com base na GWS (RESENDE et al., 2012).

RESULTADOS E DISCUSSÃO

A população de melhoramento de citros [(*Citrus sinensis* Osbeck X *Citrus reticulata* Blanco) X (*Citrus sinensis* Osbeck)], composta por 180 indivíduos, foi fenotipada para nove características e genotipada com 16.618 marcadores SNP (*Single Nucleotide Polymorphism*). Os marcadores que apresentaram MAF (*Minimum Allele Frequency*) $\geq 0,05$, e *Call Rate* $\geq 0,95$ foram empregados nas análises de seleção genômica. Após o controle de qualidade, 7.065 marcadores tiveram seus efeitos estimados sobre os fenótipos. Um total de 6.897 marcadores foram eliminados com o critério da MAF e 4.487 marcadores foram eliminados pela *Call rate*, evidenciando que alguns marcadores não atenderam aos dois critérios estabelecidos para o controle de qualidade.

A densidade de marcadores utilizada na análise foi de aproximadamente 1 marcador a cada 52kb, considerando o tamanho do genoma de *Citrus*, definido por Xu et al. (2013) para laranjeira valência. Meuwissen (2009), em um estudo com simulação considerando reduzido parentesco entre os indivíduos da população de treinamento e de seleção, demonstrou que para prever com alta acurácia ($\sim 0,9$) os valores genômicos dos indivíduos são necessários um total de $10N_eL$ marcadores informativos, sendo N_e o tamanho efetivo da população e L o tamanho do genoma em Morgans. Resende et al. (2008) afirmaram que um marcador a cada cM é suficiente para predições acuradas com a GWS, desde que tais marcadores sejam informativos, isto é, eles precisam estar em desequilíbrio de ligação com as variantes causais das características fenotípicas.

A magnitude esperada do desequilíbrio de ligação (DL) para o presente estudo é de 0,99. O desequilíbrio de ligação é uma medida de dependência entre dois ou mais locos e, quando assume valores próximos a 0, os alelos de diferentes genes são considerados independentes e, quando assume valores próximos a 1, os alelos de diferentes genes estão ligados (RESENDE et al., 2008). Assumindo que o DL entre os pares de marcadores é aproximadamente igual ao DL entre os marcadores e os QTL – *Quantitative Trait Loci* (ZHONG et al., 2009), a quantidade de marcadores utilizada neste estudo tende a proporcionar alto DL entre estes.

Na Tabela 1 são apresentadas as estimativas dos parâmetros relacionados à seleção genômica e à seleção com base apenas em fenótipo.

Tabela 1. Resultados experimentais da Seleção Genômica Ampla (GWS) para características de qualidade de frutos em uma população de melhoramento de *Citrus*.

Fenótipo	h^2	r_{yy}	h_g^2	ω_g	b_g	r_{yg}	r_{gg}	r_{ggE}	n_{QTL}
MF	0,81	0,90	0,46	0,57	1,22	0,35	0,39	0,82	223
DF	0,77	0,88	0,52	0,68	1,48	0,43	0,49	0,82	141
DT	0,76	0,87	0,42	0,55	1,31	0,33	0,38	0,81	218
NG	0,62	0,79	0,53	0,85	1,21	0,38	0,49	0,78	113
NS	0,80	0,89	0,38	0,48	0,91	0,29	0,33	0,82	286
RS	0,67	0,82	0,42	0,63	0,65	0,28	0,34	0,80	229
SS	0,76	0,87	0,37	0,49	1,64	0,35	0,40	0,81	200
Rt	0,81	0,90	0,61	0,75	1,25	0,48	0,54	0,82	121
F/C	0,68	0,82	0,51	0,75	1,42	0,40	0,48	0,80	129

h^2 : Herdabilidade no sentido amplo baseada em fenótipos; r_{yy} : Acurácia da seleção fenotípica com base no método REML/BLUP; h_g^2 : Herdabilidade baseada em marcadores obtida por meio da GWS; $\omega_g = h_g^2/h^2$: Proporção da herdabilidade baseada em fenótipos capturada pelos marcadores; b_g : viés da predição; r_{yg} : Capacidade preditiva da GWS; r_{gg} : Acurácia da GWS; r_{ggE} : Acurácia esperada da GWS; n_{QTL} : número estimado de QTL controlando cada característica. MF: Massa do fruto (g); DF: Diâmetro Longitudinal do fruto (cm); DT: Diâmetro Transversal do fruto (cm); NG: Número de gomos; NS: Número de sementes; RS: Rendimento de suco (mL); SS: Teor de sólidos solúveis; Rt: *Ratio*; F/C: Frutos por caixa.

A fração da herdabilidade capturada pelos marcadores (ω) variou de 0,48 (NS) a 0,85 (NG) (Tabela 1). O incompleto DL entre os alelos causadores da variação genética e os marcadores utilizados na predição é, provavelmente, a justificativa para a fração da herdabilidade não capturada pelos marcadores (Herdabilidade faltante). A perda de DL pode ocorrer quando os alelos causadores da variação genética apresentam menor MAF (*Minor Allele Frequency*) do que o marcador (KEMPER e GODDARD, 2012).

O viés da predição da GWS é obtido por meio do coeficiente de regressão do valor fenotípico observado sobre o valor genotípico predito. Quando os coeficientes de regressão apresentam valores próximos à unidade, as avaliações são consideradas não viesadas e são efetivas em predizer as reais magnitudes das diferenças entre os indivíduos que estão sendo avaliados (RESENDE et al., 2012a). As características que apresentaram maior viés da predição foram rendimento de suco e teor de sólidos solúveis. A variância genética da característica Rendimento de suco está superestimada, visto que o coeficiente de regressão foi inferior a 1; e subestimada para a característica teor de sólidos solúveis, pois o coeficiente de regressão foi superior a 1 (RESENDE et al., 2014).

A capacidade preditiva (r_{yg}) da GWS é fornecida pela correlação entre os valores fenotípicos observados e os valores preditos e reflete a consistência das informações moleculares em informar sobre o fenótipo (CAVALCANTI et al., 2012; RESENDE et al., 2012a). A capacidade preditiva variou de 0,28 (RS) a 0,48 (Rt) (Tabela 1). Legarra et

al. (2008) afirmaram que as estimativas da capacidade preditiva tendem a ser menores para as características que apresentam baixa herdabilidade. Este padrão não foi observado neste estudo.

As estimativas de acurácia com base nos fenótipos foram superiores às aquelas obtidas por meio da GWS para todas as características. Este fato indica que os marcadores falharam em capturar as segregações devidas à amostragem mendeliana na formação dos gametas (CROS et al., 2015). A acurácia de predição da GWS para as características avaliadas variou de 0,33 (NS) a 0,54 (Rt) (Tabela 1). A acurácia de predição é considerada intermediária ($0,3 \leq r < 0,7$) para todas as características avaliadas, considerando o critério adotado por Minamikawa et al. (2017).

A acurácia da GWS foi inferior à esperada para todas as características analisadas (Tabela 1). Os principais fatores apontados como determinantes da acurácia da GWS são: herdabilidade e arquitetura genética da característica sob seleção, estrutura genética e tamanho da população, tamanho efetivo populacional, densidade de marcadores, extensão do desequilíbrio de ligação (DL), sistema de avaliação genética das características e método utilizado para a predição dos valores genômicos dos indivíduos sob avaliação (GRATTAPAGLIA e RESENDE, 2011; KUMAR et al., 2012).

Os valores de acurácia observados sugerem que uma significativa fração da variação genética, provavelmente devido aos alelos com baixa frequência, não foi capturada com a densidade de SNP e o tamanho da população de treinamento utilizados neste estudo. Esta situação foi também observada por Kumar et al. (2012) estudando uma população de melhoramento de maçã (*Malus x domestica* Borkh.). O número reduzido de indivíduos utilizado neste estudo pode ser considerado um fator limitante para a obtenção de altas acurácias, uma vez que, quanto maior a população de treinamento mais acuradas serão as estimativas dos valores genômicos dos indivíduos – GEBV (GREVENHOF et al., 2012; HAYES et al., 2009).

O efeito do número de indivíduos utilizado para a predição genômica sobre a acurácia tem sido pesquisado por vários autores. Meuwissen et al. (2001), em estudo pioneiro, demonstraram que com a redução do número de indivíduos utilizados na população de treinamento, de 2.200 para 500 indivíduos, a acurácia reduziu de 17 a 61%, dependendo do método utilizado para a predição. Biscarini et al. (2017) pesquisaram a viabilidade da GWS no melhoramento genético da Pera e observaram que quanto maior o tamanho da população mais confiáveis (acurados) foram os resultados.

Minamikawa et al. (2017) fenotiparam 17 características relacionadas à qualidade de frutos em *Citrus* e genotiparam com 1.841 marcadores SNP uma população genitora composta por 111 variedades e uma população de melhoramento com 676 indivíduos pertencentes a 35 famílias de irmãos completos. Os autores obtiveram acurácias variando de baixa (~0,2 para a característica aparência) a alta (~0,9 para cor do fruto). Provavelmente, o número de indivíduos analisados (787) favoreceu os altos valores de acurácia observados no estudo, visto que o número de marcadores SNP utilizado foi inferior ao utilizado no presente estudo.

Habier et al. (2010) sugeriram que a redução da acurácia observada quando há a redução do tamanho da população de treinamento pode ser resultado do menor número de parentes nesta população. Os programas de melhoramento genético de *Citrus* apresentam uma estreita base genética (MACHADO et al., 2011) e, segundo Kumar et al. (2012), este fato garantiria uma alta acurácia devido ao parentesco genético aditivo entre os candidatos à seleção e a população de treinamento.

Para que a seleção genômica seja efetiva, especialmente através de gerações, é requerido que os marcadores estejam em suficiente DL com os QTL (MINAMIKAWA et al., 2017), o que pode ser alcançado por meio do aumento do tamanho da população de treinamento e da densidade de marcadores (CROS et al., 2015). Geralmente, o DL é menor em espécies alógamas heterozigotas, como os citros, do que em espécies autógamias homozigotas, sendo necessário um maior número de marcadores para cobrir o genoma de espécies alógamas.

Estudando duas populações de melhoramento de citros, Minamikawa et al. (2017) observaram uma ampla variação de DL ($0,25 < r^2 < 1\text{Mb}$), o que não era esperado visto que as populações de melhoramento de citros são altamente heterozigotas (MACHADO et al., 2011). Provavelmente, o valor de DL observado se deve ao fato de que os programas de melhoramento genético de citros são conduzidos com um reduzido número de genitores (MACHADO et al., 2011; OLIVEIRA et al., 2014), o que acarreta o chamado efeito de gargalo que influencia diretamente no aumento do DL.

Uma vez que a GWS é aplicada sobre todos os locos simultaneamente e é baseada na estimação e predição e não em teste de hipótese, ela explica muito mais da variância genética e previne a chamada herdabilidade perdida, típica de estudos de associação e ligação (VIANA et al., 2017). Com a utilização da GWS foi possível explicar uma maior proporção da variância genética do que com os métodos baseados na identificação de

QTL, os quais tem permitido a identificação de apenas alguns QTL com grandes efeitos para caracteres de interesse na cultura do citros.

Curtolo et al. (2017) fizeram o mapeamento de QTL, com marcadores DArT, para a população de citros analisada no presente estudo. A variância fenotípica explicada pelos marcadores variou de 0,04 a 9,7%. A explicação de apenas parte da variância genética do caráter permite confirmar a natureza quantitativa destas características, as quais são governadas por muitos genes de pequenos efeitos, podendo apresentar alguns com moderados a grandes efeitos, confirmando o modelo infinitesimal (KEMPER e GODDARD, 2012; MEUWISSEN et al, 2001; MEUWISSEN e GODDARD, 2010). A natureza quantitativa das características avaliadas pode ser comprovada pelo número de QTL estimado que variou de 113, para a característica número de gomos, a 286, para a característica número de sementes.

Constata-se que, mesmo para as características em que não foi possível detectar nenhum QTL, no estudo realizado por Curtolo et al. (2017), foi possível prever os valores genômicos dos genótipos com algum nível de acurácia, confirmando o potencial da GWS no melhoramento genético para características complexas, como as características relacionadas à qualidade e quantidade de frutos (IWATA, 2013).

A eficiência da GWS no melhoramento genético da população de citros analisada no presente estudo foi anteriormente confirmada por Gois et al. (2016) utilizando marcadores DArT. As acurácias da GWS estimadas no referido estudo variaram de 0,36 (Ratio) a 0,77 (Frutos por caixa). Constata-se que as acurácias preditas com marcadores DArT foram superiores às aquelas preditas com os marcadores SNP, fato este não esperado visto que os marcadores SNP são considerados mais informativos, devido à sua natureza codominante (CAIXETA et al., 2016). Os marcadores DArT utilizados foram mais informativos do que os SNPs, ou seja, provavelmente estavam em maior DL com as variantes causais das características analisadas.

O período juvenil consideravelmente longo, variando de 2 a 13 anos, observado em citros (MACHADO et al., 2005) demanda a adoção de métodos e estratégias eficientes para acelerar o processo de seleção. A juvenilidade refere-se à incapacidade de florescimento durante o período de desenvolvimento inicial da planta originada por semente (OLIVEIRA et al., 2014). Devido a esse longo período, ocorre o atraso do melhoramento genético da cultura, visto que neste período não há produção de frutos e, quando estes são produzidos, apresentam características distintas das plantas com completa maturação fisiológica (SPIEGEL-ROY e GOLDSCHMIDT, 2008).

A não adoção de métodos adequados no melhoramento genético de espécies perenes proporciona atraso do processo de hibridação e de seleção para as características desejadas. Além disso, há a necessidade de grandes áreas para o desenvolvimento dos híbridos obtidos, aumentando o custo de manutenção no campo e limitando o número de famílias e de indivíduos por família que podem ser avaliados (GMITTER et al., 2012; TALON e GMITTER, 2008). A GWS pode contribuir para a minimização destas consequências, pois, a seleção pode ser realizada em fases juvenis das plantas reduzindo o intervalo da geração, aumentando a intensidade de seleção e, conseqüentemente, o ganho por unidade de tempo e custo (GRATTAPAGLIA e RESENDE, 2011; HEFFNER et al., 2009; KEMPER e GODDARD, 2012; RESENDE et al., 2008; WONG e BERNARDO, 2008).

Na Tabela 2 é apresentada a eficiência da GWS em relação à seleção baseada apenas em dados fenotípicos.

Tabela 2. Eficiência da GWS sobre a seleção fenotípica em *Citrus*.

Fenótipo	h^2	r_{yy}	r_{gg}	Redução no ciclo seletivo		
				25%	50%	75%
MF	0,81	0,90	0,39	0,57	0,86	1,71
DL	0,77	0,88	0,49	0,74	1,11	2,22
DT	0,76	0,87	0,39	0,58	0,87	1,74
NG	0,62	0,79	0,48	0,83	1,24	2,48
NS	0,80	0,89	0,33	0,49	0,73	1,46
RS	0,67	0,82	0,23	0,55	0,83	1,66
°B	0,76	0,87	0,40	0,62	0,93	1,85
Rt	0,81	0,90	0,54	0,79	1,19	2,38
F/C	0,68	0,82	0,48	0,78	1,17	2,34

h^2 : Herdabilidade baseada em fenótipos; r_{yy} : Acurácia da seleção fenotípica com base no método REML/BLUP; r_{gg} : Acurácia da GWS obtida por meio do modelo aditivo-dominante. MF: Massa do fruto (g); DL: Diâmetro Longitudinal do fruto (cm); DT: Diâmetro Transversal do fruto (cm); NG: Número de gomos; NS: Número de sementes; RS: Rendimento de suco (mL); SS: Teor de sólidos solúveis; Rt: *Ratio*; F/C: Frutos por caixa.

Com a redução do ciclo seletivo em 25%, a eficiência seletiva variou de 0,49 (NS) a 0,83 (NG). Com a redução do ciclo seletivo em 50%, a eficiência seletiva variou de 0,73 (NS) a 1,24 (NG). As características diâmetro longitudinal, número de gomos, *ratio* e frutos por caixa apresentaram superioridade à seleção com base em fenótipos variando de 11 a 24%. Com a redução do ciclo seletivo em 75%, para todas as características avaliadas, a GWS foi superior à seleção baseada em fenótipos, com superioridade variando de 46 a 148%.

A superioridade da GWS em relação à seleção baseada exclusivamente em fenótipos observada para a característica número de gomos confirma a importância da GWS para características que apresentam baixa herdabilidade. Observa-se que mesmo quando acurácias moderadas são obtidas com a GWS, é possível a obtenção de ganhos genéticos superiores à seleção fenotípica devido à redução do ciclo de seleção (HEFFNER et al, 2009; BISCARINI et al., 2017). A GWS é, também, uma ferramenta promissora para o melhoramento genético quando as características de interesse econômico são difíceis de serem mensuradas, ou apresentam alto custo para avaliação (CROS et al., 2015; GRATTAPAGLIA e RESENDE, 2011).

CONCLUSÃO

A GWS apresenta potencial para ser utilizada como ferramenta no melhoramento genético para qualidade de frutos em citros devido à possibilidade de redução do ciclo seletivo e, conseqüentemente, a obtenção de maior ganho genético por unidade de tempo.

REFERÊNCIAS BIBLIOGRÁFICAS

BISCARINI, F.; NAZZICARI, N.; BINK, M.; ARÚS, P.; ARANZANA, M.J.; VERDE, I. MICALI, S.; PASCAL, T.; QUILOT-TURION, B.; LAMBERT, P.; LINGE, C.S.; PACHECO, I.; BASSI, D. STELLA, A.; ROSSINI, L. Genome-enabled predictions for fruit weight and quality from repeated records in European peach progenies. **BMC Genomics**, 18:432, 2017.

CAIXETA, E.T.; OLIVEIRA, A.C.B.; BRITO, G.G.; SAKIYAMA, N.S. Tipos de marcadores moleculares. In: BORÉM, A.; CAIXETA, E.T. Ed. Marcadores moleculares. 1ª edição, Viçosa – MG, Ed. UFV, p.9-93, 2016.

CALUS, M.P.L.; MEUWISSEN, T.H.E.; ROOS, A.P.W.; VEERKAMP, R.F. Accuracy of Genomic Selection Using Different Methods to Define Haplotypes. **Genetics** 178: 553–561, 2008.

CAVALCANTI, J.J.V.; RESENDE, M.D.V.; SANTOS, F.H.C.; PINHEIRO, C.R. Predição simultânea dos efeitos de marcadores moleculares e seleção genômica ampla em cajueiro. **Revista Brasileira de Fruticultura** 34, 3: 840-846, 2012.

CAMPOS, G. De Los; GIANOLA, D.; ROSA, G.J.M. Reproducing kernel Hilbert spaces regression: a general framework for genetic evaluation. **Journal of Animal Science**, 87, 1883-1887, 2009.

CROS, D.; DENIS, M.; SANCHEZ, L.; COCHARD, B.; FLORI, A.; DURAND-GASEELIN, T.; NOUY, B.; OMORÉ, A.; POMIÈS, V.; RIOU, V.; SURYANA, E.; BOUVET, J. Genomic selection prediction accuracy in a perennial crop: case study of oil palm (*Elaeis guineensis* Jacq.). **Theor. Appl. Genet.**, 128:397–410, 2015.

CURTOLO, M.; CRISTOFANI-YALY, M.; GAZAFFI, R.; TAKITA, M.A.; FIGUEIRA, A.; MACHADO, M.A. QTL mapping for fruit quality in *Citrus* using DArTseq markers. **BMC Genomics**, 18: 289, 2017.

FAO - **Food and Agriculture Organization of the United**. Disponível em: <http://faostat.fao.org/>; Acessado em: 10 de junho de 2016.

GIANOLA, D. Genomic-Assisted Prediction of Genetic Value With Semiparametric Procedures. **Genetics**, 173:1761–1776, 2006.

GMITTER JUNIOR, F.G.; CHEN, C. MACHADO, M.A.; SOUZA, A.A.; OLLITRAULT, P.; FROEHLICHER, Y.; SHIMIZU, T. Citrus genomics. **Tree Genetics & Genomes** 8: 611–626, 2012.

GOIS, I.B.; BORÉM, A.; CRISTOFANI-YALY, M.; RESENDE, M.D.V.; AZEVEDO, C.F.; BASTIANEL, M.; NOVELLI, V.M.; MACHADO, M.A. Genome wide selection in *Citrus* breeding. *Genetics and Molecular Research*, 15 (4): 1-14, 2016.

GRATTAPAGLIA, D.; RESENDE, M.D.V. Genomic selection in forest tree breeding. **Tree Genetics & Genomes** v. 7, p. 241-255, 2011.

GREVENHOF, E.M.V.; ARENDONK, J.A.M.V.; BIJMA, P. Response to genomic selection: The Bulmer effect and the potential of genomic selection when the number of phenotypic records is limiting. **Genetics Selection Evolution**, 44:26, 2012.

GUSSEN, O.; UZUN, A.; SEDAY, U.; KAFA, G. QTL analysis and regression model for estimating fruit setting in young *Citrus* trees based on molecular markers. **Scientia Horticulturae** 130(2), p. 418-424, 2011.

HAYES, B.J.; CHAMBERLAIN, A.J.; GODDARD, M.E. Genomic selection in dairy cattle: Progress and Challenges. *Journal of Dairy Science*, 2009.

HABIER, D.; TETENS, J.; SEEFRIED, F.R.; LICHTNER, P.; THALLER, G. The impact of genetic relationship information on genomic breeding values in German Holstein cattle. **Genet Sel Evol**, 42:5, 2010.

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

HENDERSON, C.R. **Maximum likelihood estimation of variance components**. Unpublished manuscript, Department of Animal Science, Cornell University, 1973.

IWATA H, HAYASHI T, TERAOKAMI S, TAKADA N, et al. (2013). Potential assessment of genome-wide association study and genomic selection in Japanese pear ***Pyrus pyrifolia***. **Breeding Science** 63: 125–140.

JARRELL, D.C.; ROOSE, M.L.; TRAUGH, S.N.; KUPPER, R.S. A genetic map of citrus based on the segregation of isozymes and RFLPs in an intergeneric cross. **Theory Applied Genetics** 84:49–56, 1992.

KEMPER, K.E.; GODDARD, M.E. Understanding and predicting complex traits: knowledge from cattle. **Human Molecular Genetics** 21:45-51, 2012.

KUMAR, S.; BINK, M.C.A.M.; VOLZ, R.K.; BUS, V.G.M.; CHGNÉ, D. Towards genomic selection in apple (*Malus × domestica* Borkh.) breeding programmes: Prospects, challenges and strategies. **Tree Genetics & Genomes**, 8:1–14, 2012.

LEGARRA, A.; ROBERT, GRANIÉ, C.; MANFREDI, E.; ELSSEN, J.M. Performance of genomic selection in mice. **Genetis** 180: 611-618, 2008.

LYNCH, M.; WALSH, B. **Genetics and analysis of quantitative traits**. Sunderland: Sinauer Associates, Inc., 1997, 980p.

MACHADO, M.A. Melhoramento de citros. In: Simpósio sobre atualização em genética e melhoramento de plantas. **Anais**, 1997, 297p.

MACHADO, M.A. **Genética, melhoramento e biotecnologia de citros**. In: MATTOS JUNIOR, D. (Ed.). *Citros*. Campinas: IAC: FUNDAG, p. 221-277, 2005.

MACHADO, M.A.; CRISTOFANI-YALY, M.; BASTIANEL, M. Breeding, genetic and genomic of citrus for disease resistance. **Rev. Bras. Frutic.**, Volume Especial: 158-172, 2011.

MEUWISSEN, T. Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. **Genetics Selection Evolution**, 41:35, 2009.

MEUWISSEN, T.; GODDARD, M. Accurate prediction of genetic values for complex traits by whole-genome resequencing. **Genetics**.185(2): 623–31, 2010.

MEUWISSEN, T.H.E.; HAYES, B.J.; GODDARD, M.E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. **Genetics** n. 157, p. 1819–1829, 2001.

MEUWISSEN, T.H.E. Genomic selection: marker assisted selection on a genome wide scale. **J. Anim. Breed. Genet.** 124: 321–322, 2007.

MINAMIKAWA, M.F.; NONAKA, K.; KAMINUMA, E.; KAJIYA-KANEGAE, H.; ONOGI, A.; GOTO, S.; YOSHIOKA, T.; IMAI, A.; HAMADA, H.; HAYASHI, T.; MATSUMOTO, S.; KATAYOSI, Y.; TOYODA, A.; HAMADA, H.; FUJIYAMA, A.; NAKAMURA, Y.; SHIMIZU, T.; IWATA, H. Genome-wide association study and genomic prediction in citrus: Potential of genomics-assisted breeding for fruit quality traits. **Scientific RepoRts**, 7: 4721, 2017.

OLIVEIRA, A.P.; SOARES FILHO, W.S.; MACHADO, M.A.; FERREIRA, E.A.; SCIVITTARO, W.B.; GESTEIRA, A.S.; Melhoramento genético de plantas cítricas. **Informe Agropecuário**, 35(281): 22-29, 2014.

PEREZ, P.; DE LOS CAMPOS, G. Genome-Wide Regression and Prediction with the BGLR Statistical Package. **Genetics**, 198:483–495, 2014.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2018. <http://www.R-project.org>.

RESENDE, M. D. V. **Genética biométrica e estatística no melhoramento de plantas perenes**. Brasília: Embrapa Informação Tecnológica, 2002. 975p.

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

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

RESENDE, M.D.V. **Seleção Genômica Ampla (GWS) e Modelos Lineares Mistos**. In: RESENDE, M.D.V. Matemática e Estatística na Análise de Experimentos. Colombo: Embrapa Florestas, 2007. 362p. Cap. 11. p. 517-533.

RESENDE, M.D.V.; AZEVEDO, C.F.; SILVA, F.F. **Seleção Genômica**. In: RESENDE, M.D.V.; SILVA, F.F.; AZEVEDO, C.F. Estatística Matemática, Biométrica e Computacional. Visconde do Rio Branco: Suprema, 2014. Cap. 14. p. 627-768.

RESENDE, M.D.V.; AZEVEDO, C. F.; SILVA, F.F.; GOIS, I.B. **Atualidades da biometria no melhoramento de plantas perenes**. In: LUDCK, W.H.; ANDRADE, A.C.B.; VOLPATO, L.; et al. (Org.). Desafios biométricos no melhoramento genético. 1ªed. Belo Horizonte - MG: eDOC Brasil, 2017, v. 1, p. 67-84.

RESENDE, M.D.V.; FURLANI-JUNIOR, E.; MORAES, M.L.T.; FAZUOLI, L.C. Estimativas de parâmetros genéticos e predição de valores genotípicos no melhoramento do cafeeiro pelo procedimento REML/BLUP. **Bragantia** 60(3): 185-193, 2001.

RESENDE, M.D.V.; LOPES, P.S.; SILVA, R.L.; PIRES, I.E. Seleção genômica ampla (GWS) e maximização da eficiência do melhoramento genético. **Pesquisa Florestal Brasileira** 56: 63-77, 2008.

RESENDE, M.D.V.; SILVA, F.F.; LOPES, P.S.; AZEVEDO, C.F. **Seleção Genômica Ampla (GWS) via Modelos Mistos (REML/BLUP), Inferência Bayesiana (MCMC), Regressão Aleatória Multivariada e Estatística Espacial**. Viçosa: universidade Federal de Viçosa/ Departamento de Estatística. 2012a. 291p.

SPIEGEL-ROY, P.; GOLDSCHMIDT, E.E. **Biology of citrus**. 2.ed. Cambridge: Cambridge University, 2008. 230p.

SVED, J.A. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. **Theor Popul Biol**, 2: 125–141, 1971.

TALON, M.; GMITTER JUNIOR, F.G. Citrus genomics. **International Journal of Plant Genomics**, 2008.

TEROLL, J.; CONESA, A; COLMENERO, J.M.; CERCOS, M.; TADEO, F.; AGUSTÍ, J.; ALÓS, E.; ANDRES, F.; SOLER, G.; BRUMOS, J.; IGLESIAS, D.J.; GOTZ, S.; LEGAZ, F.; ADOURT, X.; COURTOIS, B.; OLLITRAULT, P.; DOSSAT, C.; WINCKER, P.; MORILLON, R.; TALON, M. Analysis of 13000 unique *Citrus* clusters associated with fruit quality, production and salinity tolerance. **BMC Genomics**, n. 8, v. 31, p. 1-22, 2007.

TEROLL, J.; NARANJO, M.A.; OLLITRAULT, P.; TALON, M. Development of genomic resources for *Citrus clementina*: Characterization of three deep-coverage BAC libraries and analysis of 46,000 BAC end sequences. **BMC Genomics**, n. 9, v. 423, p. 1-12, 2008.

VIANA, A.P.; SILVA, F.H.L.; GLÓRIA, L.S.; RIBEIRO, R.M.; KRAUSE, W.; BOECHAT, M.S.B. Implementing genomic selection in sour passion fruit Population. **Euphytica**, 213:228, 2017.

WONG, C.K.; BERNARDO, R. Genome wide selection in oil palm: increasing selection gain per unit time and cost with small populations. **Theor Appl Genet**. 116:815–824, 2008.

ZAPATA-VELENZUELA, J.; WHETTEN, R.W.; NEALE, D.; MCKEAND, S.; ISIK, F. Genomic Estimated Breeding Values Using Genomic Relationship Matrices in a Cloned Population of Loblolly Pine. **Genes, genome, Genetics** 3: 909-916, 2013.

ZAPATA-VALENZUELA, J.; ISIK, F.; MALTECCA, C.; WEGRZYN, J.; NEALE, D.; MCKEAND, S.; WHETTEN, R. SNP markers trace familial linkages in a cloned population of *Pinus taeda*—prospects for genomic selection. **Tree Genetics & Genomes**, 8:1307–1318, 2012.

ZHONG, S.; DEKKERS, J.C.M.; FERNANDO, R.L.; JANNINK, J.L. Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: A barley cases study. **Genetics**, 182: 355–364, 2009.

CAPÍTULO 2
SELEÇÃO GENÔMICA PARA RESISTÊNCIA À MANCHA MARROM DA
ALTERNARIA EM *Citrus*

Viçosa – MG
Abril, 2019

RESUMO

GOIS, Itamara Bomfim, D.Sc., Universidade Federal de Viçosa, Abril, 2019. **Seleção genômica para resistência a *Alternaria alternata* em *Citrus***. Orientador: Aluizio Borém de Oliveira. Coorientadores: Marcos Deon Vilela de Resende; Mariângela Cristofani-Yaly.

A mancha marrom de alternaria (MMA) é uma doença causada pelo fungo *Alternaria alternata*, cuja susceptibilidade tem sido observada em tangerinas, pomelos, tangores e seus híbridos. A alta incidência desta doença em pomares causa a redução significativa da produção e, a utilização de genótipos resistentes é a melhor estratégia para o seu controle. Este trabalho foi realizado com o objetivo de avaliar 235 híbridos provenientes do cruzamento entre laranja Pêra e o híbrido H163 (tangor Murcott (*C. reticulata* x *C. sinensis*) e laranja Pêra (*C. sinensis*)) com relação à resistência à MMA e a aplicabilidade da seleção genômica (GWS) no melhoramento genético para essa característica. Para a avaliação da resistência à MMA, foram coletadas três folhas jovens de cada híbrido e dos genitores. Foi empregado o delineamento inteiramente ao acaso, sendo considerados resistentes os genótipos que não apresentaram sintomas da doença. Para a análise de seleção genômica foram utilizados marcadores SNP e a predição dos valores genômicos foi realizada por meio do método GBLUP (*Genomic Best Linear Unbiased Prediction*) via procedimento RKHS (*Reproducing Kernel Hilbert Spaces*) com algoritmo Bayesiano. Observou-se que 75,3% dos híbridos foram classificados como resistentes à MMA. As estimativas da herdabilidade no sentido amplo e da acurácia foram consideradas de alta magnitude (>0,9). Na análise da GWS, a fração da herdabilidade capturada pelos marcadores (ω) foi igual a 0,49 e a capacidade preditiva igual a 0,41. A estimativa da acurácia com base nos fenótipos foi superior àquela obtida por meio da GWS (0,42). A resistência à MMA é uma característica governada por mais de um gene e a utilização de marcadores em alto desequilíbrio de ligação (DL) com estes genes aumentará a acurácia da GWS.

Palavras-chave: Característica qualitativa; SNP; acurácia; seleção precoce.

ABSTRACT

GOIS, Itamara Bomfim, D.Sc., Universidade Federal de Viçosa, April, 2019. **Genomic selection for resistance to *Alternaria alternata* in *Citrus***. Orientador: Aluizio Borém de Oliveira. Coorientadores: Marcos Deon Vilela de Resende e Mariângela Cristofani-Yaly.

The alternaria brown spot (ABS) is a disease caused by the fungus *Alternaria alternata*, whose susceptibility is observed in tangerines, grapefruits, tangores and their hybrids. The high incidence of this disease in orchards causes a significant reduction of the production, and the use of resistant genotypes is the best strategy to control it. This work was carried out with the objective of evaluating 235 hybrids from the cross between orange Pêra and H163 (Murcott (*C. reticulata* x *C. sinensis*) e orange Pêra (*C. sinensis*)) in relation to resistance to ABS and the applicability of genomic selection (GWS) in the genetic improvement of this characteristic. For the evaluation of resistance to ABS, three young leaves of each hybrid and of the parents were collected. The design was completely randomized, being considered resistant the genotypes that did not present symptoms of the disease. For the genomic selection analysis, SNP markers were used and the prediction of the genomic values was computed using the GBLUP (Genomic Best Linear Unbiased Prediction) method using a RKHS (Reproducing Kernel Hilbert Spaces) procedure with Bayesian algorithm. It was observed that 75.3% of the hybrids were classified as resistant to ABS. Estimates of heritability in the broad sense and accuracy were considered as high magnitude (> 0.9). In the GWS analysis, the fraction of heritability captured by the markers (ω) was equal to 0.49 and the predictive capacity equal to 0.41. The estimate of accuracy based on the phenotypes was higher than that obtained by GWS (0.42). Resistance to ABS is a characteristic governed by more than one gene and the use of markers in high linkage disequilibrium (LD) with these genes will increase the accuracy of GWS.

Key-words: Qualitative characteristic; SNP; accuracy; early selection.

INTRODUÇÃO

A base genética da citricultura brasileira é bastante restrita, fato que a torna vulnerável a vários problemas fitossanitários. Dentre as doenças que ocorrem nos pomares pode-se citar a mancha marrom de alternaria (MMA), que é uma doença fúngica causada pela *Alternaria alternata*, cuja suscetibilidade tem sido observada em tangerinas, pomelos, tangores e alguns de seus híbridos (Azevedo et al., 2010; Azevedo et al., 2015; Pacheco et al., 2012).

Os principais sintomas observados quando a planta é infectada por este fungo são lesões necrosadas em ramos, frutos, tecidos verdes imaturos ou em fase de crescimento, podendo ocorrer um desfolhamento durante os períodos de maior infecção e redução na produção devido à queda de frutos (Azevedo et al., 2010; Stuart et al., 2009; Timmer et al., 2003). Além disso, as manchas presentes nos frutos afetam a comercialização dos mesmos (Aktimisu et al., 2003).

A alta susceptibilidade à MMA das principais variedades plantadas no estado de São Paulo torna a produção pouco rentável, visto que a aplicação de fungicidas é requerida para o controle da doença (Pacheco et al., 2012). Dependendo do clima da região e da susceptibilidade da cultivar, a aplicação de fungicida de quatro a dez vezes por ano é realizada para produzir frutos com qualidade (Cuenca et al., 2016). Observa-se que a adoção desta estratégia é morosa, dispendiosa e gera preocupações com o meio ambiente e com a saúde humana.

A seleção de genótipos resistentes à MMA é a melhor estratégia para contornar estas situações e, com o advento de marcadores moleculares é possível a realização da seleção em estádios iniciais do desenvolvimento da planta. Meuwissen et al. (2001) afirmaram que a justificativa para a utilização da genética molecular no melhoramento de plantas é que o uso de informação em nível de DNA permitirá ganhos mais rápidos, maior eficiência seletiva e baixo custo quando comparado à seleção baseada apenas em dados fenotípicos.

A seleção genômica ampla (GWS) é um tipo de seleção precoce que foi proposta como uma alternativa para aumentar a eficiência dos programas de melhoramento genético. A GWS consiste na predição simultânea, sem o uso de testes de significância para marcas individuais, dos efeitos genéticos dos marcadores de DNA dispersos no genoma, de forma a capturar os efeitos dos locos e explicar toda a variação genética de um caráter quantitativo (Resende et al., 2013).

A GWS emprega marcadores moleculares amplamente distribuídos no genoma, que tiveram seus efeitos estimados em uma população de “treinamento”. Os marcadores são utilizados para construir modelos de predição que serão aplicados aos candidatos à seleção, os quais possuem apenas os dados genotípicos (de marcadores), sendo que os valores genéticos genômicos serão preditos por meio dos dados genômicos. Os modelos de predição são validados em uma população de “validação”, sendo que o modelo que apresentar maior correlação entre os valores observados e os preditos é selecionado para ser utilizado na fase de seleção (Grattappaglia et al., 2018; Li et al., 2018).

Este trabalho foi realizado com o objetivo de avaliar uma população de híbridos provenientes do cruzamento entre laranja Pêra e o híbrido H163 (tangor Murcott (*C. reticulata* x *C. sinensis*) e laranja Pêra (*C. sinensis*)) com relação à resistência à MMA e a aplicabilidade da GWS no melhoramento genético para esta característica.

MATERIAL E MÉTODOS

População de estudo

A fenotipagem e genotipagem foram realizadas em uma população de 235 híbridos provenientes do cruzamento entre laranja Pêra de Abril (LPA) (*Citrus sinensis*) e o híbrido H163. Este híbrido é resultante do cruzamento entre tangor Murcott (*C. reticulata* x *C. sinensis*) e laranja Pêra (*C. sinensis*).

A população foi obtida no ano de 2010 e é mantida em casa de vegetação no Centro de Citricultura Sylvio Moreira do Instituto Agrônomo de Campinas (CCSM-IAC), Cordeirópolis, SP.

Análise fenotípica para resistência a Alternaria alternata

Os isolados de *A. alternata* foram obtidos a partir de tecidos lesionados de tangor Murcott, coletados em plantas mantidas em campo no Centro de Citricultura Sylvio Moreira em São Paulo. Para a avaliação da resistência à mancha marrom da alternaria (*A. alternata*), foram coletadas três folhas jovens, com cerca de 2-3 cm, de todos os híbridos e dos genitores. Como controle, foram utilizadas folhas de tangerina Dancy e de tangor Murcott, que são variedades suscetíveis, e folhas de tangerina Fremont, que é uma variedade considerada resistente. A laranja Pêra é resistente à MMA.

As inoculações *in vitro* foram realizadas no laboratório de Melhoramento e Qualidade da fruta do Centro de Citricultura Sylvio Moreira do Instituto Agrônomo de Campinas, em câmara BOD (FANEM, 347CGD), a 27°C, com fotoperíodo de 12 horas, seguindo protocolo descrito por Peever et al. (1999) e Campos et al. (2017).

Os híbridos foram avaliados com relação à resistência a *A. alternata*, sendo considerados resistentes aqueles que não apresentaram sintomas da doença e suscetíveis aqueles que apresentaram. Foi empregado o delineamento inteiramente ao acaso.

A análise dos dados foi realizada a partir da metodologia de modelos mistos (Henderson, 1973), utilizando o *software* Selegen REML/BLUP (Resende, 2016). O modelo estatístico utilizado equivale a: $y = 1\mu + Zg + e$, em que y é o vetor de dados, μ é a média geral (fixo), g é o vetor de efeitos genotípicos totais (aleatórios), e é o vetor de erros (aleatórios), 1 é um vetor de 1's. As letras maiúsculas representam as matrizes de incidência para os referidos efeitos. A estrutura de médias e variâncias associadas ao modelo são: $y|u, V \sim N(1u, V)$; $g|\sigma_g^2 \sim N(0, I\sigma_g^2)$; $e|\sigma_e^2 \sim N(0, I\sigma_e^2)$; $Cov(g, e) = 0$. Em

que, V é a matriz de covariância fenotípica; 1 é um vetor de uns e I é uma matriz identidade. Nesse caso, I é adequada para indivíduos irmãos com o mesmo grau de parentesco genealógico entre eles, como no caso de indivíduos da família de irmãos germanos usada no presente estudo.

As equações de modelos mistos para o BLUP (*Best Linear Unbiased Predictor*) dos efeitos genotípicos equivalem a:

$$\begin{bmatrix} 1'1 & 1'Z \\ Z'1 & Z'Z + I^{-1}\lambda_1 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} 1'y \\ Z'y \end{bmatrix}$$

Em que, $\lambda_1 = \sigma_e^2 / \sigma_g^2$ é o parâmetro de penalização para o efeito genotípico, σ_g^2 é a variância dos efeitos genotípicos e σ_e^2 é a variância residual.

Os componentes de variância foram estimados via REML (*Restricted Maximum Likelihood*; Patterson e Thompson, 1971), por meio de iterações nas equações de modelos mistos. A acurácia da predição dos valores genotípicos foi calculada a partir da inversa da matriz de coeficientes das equações de modelos mistos (Resende, 2002).

Genotipagem e controle de qualidade dos marcadores

A genotipagem foi realizada por meio de marcadores SNP (*Single Nucleotide Polymorphism*). O DNA total de cada um dos 237 híbridos foi extraído, quantificado em Espectrofotômetro Nanodrop 8000, normalizado para 100 ng/ μ l e enviado para a plataforma de genotipagem. Os marcadores SNP foram identificados por meio da plataforma de genotipagem *DArTseq* desenvolvida por *Diversity Arrays Technology* (Yarralumla, Austrália).

Após o processo de genotipagem, foram obtidos 7.639 marcadores SNP, os quais foram identificados a partir do alinhamento de sequências com o genoma de referência, *Citrus reticulata*. Os genótipos foram codificados como 0, 1 e 2, em cada loco, de acordo com o número de cópias do alelo variante que eles possuíam. Quando o genótipo foi homocigoto para o alelo de referência, este foi codificado como 0; quando o genótipo foi heterocigoto, ou seja, apresentou o alelo de referência e o alelo variante, este foi codificado como 1; e, quando o genótipo foi homocigoto para o alelo variante, este foi codificado como 2.

O controle de qualidade dos marcadores SNP foi realizado por meio da eliminação dos locos com baixa taxa de atendimento na genotipagem dos indivíduos (*call rate* $\leq 95\%$); e de locos com baixo polimorfismo (*MAF* – *minor allele frequency* $\leq 5\%$). O nível crítico (ponto de corte) adotado para a MAF foi obtido por meio da equação $MAF = 1/\sqrt{2N}$ (Resende et al., 2017), em que N se refere ao número de indivíduos avaliados. A MAF é dada por: $MAF = \min[p, (1 - p)]$, em que p é a frequência do alelo de referência i ; a *Call Rate* foi calculada por meio da expressão: $Call Rate = 1 - \frac{A}{P}$, em que A refere-se aos indivíduos com genótipos perdidos (*missing*) e P refere-se aos indivíduos com genótipos presentes (*not missing*) (Resende et al., 2014). Após o controle de qualidade, a imputação de valores perdidos (NA) de marcadores foi realizada com o valor médio ($2p$). As análises de controle de qualidade foram conduzidas no *software* R (R Core Team, 2018).

Predição dos valores genômicos dos indivíduos (VGG)

A parametrização das matrizes de incidência de marcadores foi realizada de acordo com Resende et al. (2014). Os valores da matriz W , associada aos efeitos genéticos aditivos (a), foram centrados em zero para que os efeitos dos marcadores sejam efeitos de substituição alélica (α) com média zero na população. Assim, a parametrização utilizada na matriz W foi: “ $2q$ ”, “ $q - p$ ” e “ $-2p$ ” para os genótipos 2, 1 e 0, respectivamente. A parametrização utilizada para a matriz S , associada aos efeitos genéticos devido aos desvios de dominância (d), foi: “ $-2q^2$ ”, “ $2pq$ ” e “ $-2p^2$ ”, para os genótipos 2, 1 e 0, respectivamente.

O método GBLUP (*Genomic Best Linear Unbiased Prediction*) via procedimento RKHS (*Reproducing Kernel Hilbert Spaces*) com algoritmo Bayesiano (GIANOLA, 2006) foi utilizado para a realização das análises de Seleção Genômica (GWS) para as características avaliadas.

O modelo adotado para predição dos valores genômicos individuais foi dado por: $y = 1\mu + Wa + Sd + e$, em que y é o vetor de observações fenotípicas; μ é a média geral; a é o vetor de efeitos genéticos aditivos dos indivíduos; d é o vetor de efeitos genéticos devido aos desvios de dominância dos indivíduos; e é o vetor de resíduos aleatórios; 1 é um vetor de 1's; W e S são as matrizes de incidência para a e d , respectivamente.

As equações de modelos mistos genômicas associadas ao método GBLUP/ GWS são dadas por:

$$\begin{bmatrix} 1'1 & 1'W & 1'S \\ W'1 & W'W + G_a^{-1} \frac{\sigma_e^2}{\sigma_a^2} & W'S \\ S'1 & S'W & S'S + G_d^{-1} \frac{\sigma_e^2}{\sigma_d^2} \end{bmatrix} \begin{bmatrix} \hat{u} \\ \hat{a} \\ \hat{d} \end{bmatrix} = \begin{bmatrix} X'y \\ W'y \\ S'y \end{bmatrix}$$

Em que, $G_a = \frac{WW'}{\sum 2p_i q_i}$ e $G_d = \frac{SS'}{\sum (2p_i q_i)^2}$ são as matrizes de parentesco genômico aditivo e devido aos desvios de dominância, respectivamente; e p_i e q_i são as frequências alélicas.

O método GBLUP equivale ao método RKHS (*Reproducing Kernel Hilbert Spaces*) quando a matriz dos *kernels* K é substituída pela matriz de parentesco genômico (G).

Na regressão RKHS, a estrutura de covariância é proporcional a uma matriz de kernel K , dada por $\text{Cov}(g_i, g_j) \propto K_{\text{RKHS}}(w_i, w_j)$, em que w_i, w_j são vetores de genótipos marcadores para os indivíduos i e j , respectivamente, e $K(\dots)$ é uma função positiva definida avaliada nos genótipos marcadores.

O modelo KRHS pode ser também assim especificado: $y = 1u + K_h \alpha + e$, em que u é uma constante, K_h é a matriz positiva definida de *kernels*, dependente do parâmetro de suavização h ; α é um vetor contendo coeficientes não paramétricos que são assumidos com distribuição normal $\alpha_j \sim N(0, K_h^{-1} \sigma_\alpha^2)$, com σ_α^2 representando a recíproca do parâmetro de alisamento ($\sigma_\alpha^2 = \lambda^{-1}$). Os resíduos têm distribuição normal com matriz de covariância $R=I \sigma_e^2$. A solução para α é dada por $[\sigma_e^{-2} K_h + \sigma_\alpha^{-2} I] \hat{\alpha} = \sigma_e^{-2} y$.

Os fenótipos são preditos por $\hat{y} = \hat{u}1 + K_h^* \hat{\alpha}$, onde uma linha de K_h^* tem a forma $K_t^* = [K_h^*(w_i - w_j)]$, com $K_h^*(w_i - w_j)$ sendo o kernel entre o genótipo do indivíduo i no grupo de validação e o genótipo do indivíduo j no grupo de estimação.

O modelo individual univariado tradicional $y = Xb + Za + e$ pode ser expresso em termos de $y = g + e$, em que $g|0, K_{\text{RKHS}} \sigma_\alpha^2 \sim N(0, K_{\text{RKHS}} \sigma_\alpha^2)$, conduzindo ao estimador $[\sigma_e^{-2} I + \sigma_\alpha^{-2} K_{\text{RKHS}}^{-1}] \hat{g} = \sigma_e^{-2} y$ (Campos et al., 2009). Assim, substituindo-se K por G no método RKHS, tem-se $\text{RKHS} = \text{GBLUP}$.

Dessa forma, utilizou-se o método RKHS com algoritmo Bayesiano (Gianola, 2006) para a realização das análises de GWS por meio do pacote BGLR (Perez e De Los Campos, 2014), implementado no *software* R (R Core Team, 2018).

Capacidade preditiva, viés da predição e acurácia da GWS

Para o cômputo da capacidade preditiva ($r_{y\hat{g}}$) foi utilizada a validação cruzada, por meio da reamostragem com reposição de um grupo de indivíduos via procedimento k-fold. Os dados da população estudada foram divididos em dois subconjuntos, 90% dos indivíduos foram selecionados aleatoriamente para compor a população de treinamento e o restante, 10%, foram usados para compor a população de validação. Esta análise foi repetida 10 vezes (10-fold) e, a cada vez, os indivíduos utilizados para a predição e para a validação foram amostrados aleatoriamente e analisados independentemente. O número de iterações utilizado foi 100.000, com *burn in* de 10.000. A capacidade preditiva ($r_{y\hat{g}}$) foi estimada a partir da média dos coeficientes de correlação entre os valores fenotípicos e os valores genômicos preditos (Daetwyler et al., 2013).

O viés da predição foi calculado por meio do coeficiente de regressão (\hat{b}) dos valores fenotípicos sobre os valores genômicos preditos. Quando b assume valores próximos a 1, indica que as avaliações não são viesadas e são efetivas em predizer as reais magnitudes das diferenças entre os indivíduos em avaliação (RESENDE et al., 2012a).

A estimativa da acurácia da seleção genômica foi obtida a partir da expressão: $r_{g\hat{g}} = \frac{r_{yg}}{\sqrt{h^2}}$, em que $r_{y\hat{g}}$ equivale à capacidade preditiva da GWS e h^2 é a herdabilidade do caráter (LEGARRA et al., 2008). A acurácia esperada foi calculada por meio da fórmula proposta por Resende (2008): $r_{ggE} = \sqrt{\frac{r_{mq}^2(Nh_m^2)}{1+[(N-1)h_m^2]}}$, em que, r_{mq}^2 é a proporção da variância genética explicada pelos marcadores dada por: $r_{mq}^2 = \frac{n_m}{n_m + M_e}$ (GODDARD et al., 2011); N é o número de indivíduos genotipados e fenotipados; h_m^2 é a herdabilidade individual de um loco, dada por: $h_m^2 = \frac{r_{mq}^2 h^2}{M_e}$; h^2 é a herdabilidade do caráter; e M_e é o número de segmentos cromossômicos independentes, dado por: $M_e = 2NeL$ (HAYES et al., 2009); Ne é o tamanho efetivo da população; e L é o tamanho do genoma em Morgans.

RESULTADOS E DISCUSSÃO

Os 235 híbridos estudados e os parentais foram avaliados para a característica resistência à mancha marrom da alternaria (MMA), e genotipados com 7.639 marcadores SNP (*Single Nucleotide Polymorphism*).

Um total de 177 híbridos (75,3%) foram classificados como resistentes à MMA por não apresentarem qualquer sintoma dessa doença no decorrer do teste. A herança da resistência à alternaria em *Citrus* tem sido classificada como qualitativa, sendo controlada por um alelo recessivo (Kohmoto, 1991). Desta forma, o alelo dominante é responsável pela expressão do genótipo suscetível e o alelo recessivo pela expressão do genótipo resistente.

Os parentais utilizados para a obtenção dos híbridos foram assintomáticos. A laranja Pêra é resistente à MMA (Stuart et al., 2009) e o Híbrido 163 foi obtido a partir do cruzamento entre tangor Murcott (*C. reticulata* x *C. sinensis*), classificado como genótipo heterozigoto suscetível (Cuenca et al., 2013), e laranja Pêra. Observa-se que os resultados obtidos não estão de acordo com o esperado para uma característica controlada por um alelo recessivo. Assim, outro gene pode estar envolvido no controle desta característica.

Campos et al. (2017) avaliaram os híbridos do presente estudo para a característica área abaixo da curva de progresso da doença (AACPD) mancha marrom da alternaria. Nos híbridos que apresentaram sintomas da doença, foi observada uma grande amplitude para os valores de AACPD, os quais variaram de 14,4 a 2.387,6. A amplitude observada é um indicativo de que mais de um gene está controlando a resistência à MMA. Esta medida, AACPD, é utilizada nos estudos de doenças policíclicas, especialmente àquelas relacionadas com a resistência quantitativa (Jeger e Viljanen-Rollinson, 2001). Diferentes graus de resistência/ suscetibilidade entre híbridos sugerem que mais de um gene deve estar envolvido no controle genético do caráter (Peever et al., 2000).

As estimativas dos parâmetros genéticos para a característica avaliada estão apresentadas na Tabela 1.

Tabela 1. Estimativas dos parâmetros genéticos relacionados à resistência a mancha marrom da alternaria, *Alternaria alternata*, em uma população de *Citrus*.

Característica	h_g^2	$CV_{gi}\%$	$CV_e\%$	CV_r	r_{gg}	Média
Resistência à alternaria	0,96	61,15	8,63	7,08	0,99	0,72

h_g^2 : herdabilidade individual no sentido amplo; $CV_{gi}\%$: coeficiente de variação genotípica; $CV_e\%$: coeficiente de variação residual; CV_r : coeficiente de variação relativo; r_{gg} : acurácia da seleção.

A estimativa de herdabilidade no sentido amplo, ou seja, considerando os efeitos aditivos e devido aos desvios de dominância, foi igual a 0,96 (Tabela 1). Este parâmetro expressa a proporção da variância fenotípica que é explicada por fatores genéticos e apresenta valor considerado de alta magnitude (Resende, 2002). Quanto maior o valor da herdabilidade menor é a influência do ambiente na expressão do caráter, sendo este governado por um ou pouco genes (natureza qualitativa).

As estimativas obtidas indicam que a seleção pode ser realizada com alta confiabilidade, fato que pode ser comprovado pelos altos valores do coeficiente de determinação dos efeitos genéticos (h_g^2) e da acurácia da seleção (Tabela 1). A estimativa da acurácia é considerada de alta magnitude (0,99) (Resende e Duarte, 2007), sendo considerada satisfatória para inferir sobre os valores genotípicos dos indivíduos em avaliação (Brasileiro et al., 2013).

O coeficiente de variação relativo (CV_r) refere-se à razão entre o coeficiente de variação genética e o coeficiente de variação residual. Vencovsky e BARRIGA (1992) afirmaram que quando a estimativa do CV_r é superior a 1, a variável apresenta elevado controle genético e é pouco influenciada pelo ambiente. Este resultado corrobora com a natureza qualitativa da resistência à mancha marrom de alternaria.

Na análise de GWS foram utilizados os marcadores que apresentaram MAF (*Minimum Allele Frequency*) $\geq 0,05$ e *Call Rate* $\geq 0,95$, um total de 6.637 marcadores. Considerando o tamanho do genoma de *Citrus* igual a 367Mb (Xu et al., 2013), a densidade de marcadores utilizada neste estudo foi de aproximadamente um marcador para 56 kb. A utilização de um marcador a cada cM é suficiente para a obtenção de predições acuradas com a GWS, desde que os marcadores estejam em desequilíbrio de ligação (DL) com os variantes causais das características fenotípicas (Resende et al., 2008).

As estimativas dos parâmetros referentes à seleção genômica estão apresentados na Tabela 2.

Tabela 2. Resultados experimentais da Seleção Genômica Ampla (GWS) para resistência à Mancha Marrom da alternaria, *Alternaria alternata*, em uma população de melhoramento de *Citrus*.

Característica	h^2	r_{yy}	h_{GWS}^2	ω	r_{yg}	b	r_{gg}	r_{gge}
Resistência à alternaria	0,95	0,99	0,46	0,49	0,41	1,21	0,42	0,87

h^2 : Herdabilidade baseada em fenótipos; r_{yy} : Acurácia da seleção fenotípica com base no método REML/BLUP fenotípico; h_{GWS}^2 : Herdabilidade baseada em marcadores obtida por meio da GWS; r_{yg} : Capacidade preditiva da GWS; r_{gg} : Acurácia da GWS; b_g : indicador do viés da predição; $\omega = h_{GWS}^2/h^2$: Proporção da herdabilidade baseada em fenótipos capturada pelos marcadores; $r_{g\hat{e}}$: Acurácia esperada da GWS.

A fração da herdabilidade capturada pelos marcadores (ω) foi igual a 0,49 (Tabela 3). Nota-se que a fração da herdabilidade não capturada pelos marcadores foi igual a 51%. Este fato indica que os alelos causadores da variação genética não estão em completo DL com os marcadores SNP utilizados. O DL é uma medida de dependência entre dois ou mais locos (Resende et al., 2008) e, seu valor reduz quando a distância entre os locos aumenta (Goddard e Hayes, 2007).

A capacidade preditiva ($r_{y\hat{g}}$) é calculada por meio da correlação entre os valores observados e os valores preditos, fornecendo, desta forma, a capacidade dos marcadores em informar sobre o fenótipo (Resende et al., 2012). As estimativas da capacidade preditiva tendem a ser maiores para características que apresentam alta herdabilidade (Legarra et al., 2008), desde que os marcadores estejam em DL com as variantes causais da resistência à MMA. Este padrão não foi observado, o que pode ser constatado pelo valor da estimativa da capacidade preditiva ($r_{y\hat{g}}= 0,41$).

O viés da predição da GWS é obtido por meio do coeficiente de regressão (b) do valor fenotípico observado sobre o valor predito, e foi igual a 1,21. Com base no que foi descrito por Resende et al. (2014), a variância genética para a característica resistência à mancha marrom de alternaria está subestimada, uma vez que, o valor de b foi superior a 1 (1,21).

A estimativa da acurácia com base nos fenótipos (0,99) foi superior àquela obtida por meio da GWS (0,42) (Tabela 3). A estimativa da acurácia da GWS é considerada intermediária ($0,3 \leq r < 0,7$) para a característica resistência à mancha marrom de alternaria, com base no critério apresentado por Minamikawa et al. (2017), e foi inferior a acurácia esperada, com base na fórmula proposta por Resende (2008).

A herdabilidade do caráter e o número de *loci* que o controlam são alguns dos fatores que influenciam a acurácia da seleção genômica (Grattapaglia e Resende, 2011). Além da arquitetura genética do caráter, o tamanho da população e o número de marcadores utilizados contribuem para o aumento da acurácia da GWS (Li et al., 2018). O tamanho da população utilizado é considerado reduzido quando comparado aos sugeridos para estudos de GWS (Meuwissen et al., 2001). No entanto, quando estão sendo estudadas características qualitativas, populações pequenas podem ser utilizadas para a detecção de marcadores associados à característica (Cao et al., 2016).

No presente estudo, a utilização de um maior número de marcadores não é um fator determinante para o aumento da acurácia seletiva da GWS, a menos que estes marcadores estejam em completo DL com os genes responsáveis pela resistência à mancha marrom de alternaria. Assim, a acurácia da seleção genômica é também dependente do padrão de DL, visto que a abordagem GWS é baseada no DL entre os marcadores e os QTL (Mewissen et al., 2001).

Jannink et al. (2010) afirmaram que a acurácia da seleção genômica não depende apenas do número e do tamanho do QTL, mas também do DL entre os marcadores e os variantes causais, uma vez que, a variação genética explicada por cada marcador é determinada pelo nível de DL entre os marcadores e os QTL ligados aos marcadores.

O aumento do número de marcadores pode ser útil para a identificação de variantes causais (Li et al., 2018), uma vez que, o parâmetro chave da acurácia da GWS é a proporção da variância do caráter explicada pelos marcadores, o qual é dependente do DL entre o QTL e o marcador ou uma combinação linear de marcadores (Goddard, 2009).

A estimativa de acurácia obtida corrobora com o fato de que os marcadores SNP utilizados no presente estudo não são tão informativos para a característica analisada. Como a característica analisada é de natureza qualitativa (Kohmoto, 1991), a obtenção de alta acurácia está diretamente relacionada com a utilização de marcadores SNP em alto DL com os genes responsáveis pela resistência à mancha marrom de *Alternaria alternata*.

Cuenca et al. (2016) fizeram o mapeamento de ligação de uma progênie de citros proveniente do cruzamento entre um genótipo heterozigoto suscetível e um resistente à MMA, para identificar genes candidatos à resistência à MMA e desenvolver marcadores para serem utilizados na seleção assistida por marcadores (SAM). Os autores identificaram nove marcadores SNP relacionados à resistência a MMA. O SNP 8 (cromossomo 3, posição 25862085), mapeado a 0,4 cM do gene de resistência à MMA, foi selecionado para ser utilizado na seleção de genótipos resistentes. Os nove SNP

identificados no estudo de Cuenca et al. (2016), relacionados à resistência, não estão presentes nos marcadores SNP utilizados no presente estudo.

Para características que apresentam variantes causais com maiores efeitos responsáveis pela sua expressão é primordial a utilização de SNP em alto DL com estes para a obtenção de altas acurácias para GWS (Kumar, 2012). Li et al. (2018) observaram que a utilização de SNP significativamente associados a tolerância à seca em *Cicer arietinum* L. promoveu a melhoria da predição da acurácia da GWS.

CONCLUSÕES

A resistência à MMA é uma característica qualitativa (alta herdabilidade) que deve ser governada por mais de um gene.

A utilização de marcadores em alto DL com os genes responsáveis pela resistência à MMA, ou de marcadores significativamente associados, proporcionará o aumento da acurácia da GWS.

REFERÊNCIAS BIBLIOGRÁFICAS

- AKIMITSU, K.; PEEVER, T.L.; TIMMER, L.W. Molecular, ecological and evolutionary approaches to understanding *Alternaria* diseases of citrus. **Molecular Plant Pathology**. 4(6), 435–446, 2003.
- AZEVEDO, F.A.; POLYDORO, D.A.; BASTIANEL, M.; KUPPER, K.C.; STUART, R.M.; COSTA, F.P.; PIO, R.M. Resposta de diferentes genótipos de tangerinas e seus híbridos à inoculação in vitro e in vivo de *Alternaria alternata*. **Revista Brasileira de Fruticultura**, Jaboticabal, v.32, p.1-10, 2010.
- AZEVEDO, F. A.; MARTELLI, I. B.; POLYDORO, D. A.; PACHECO, C. A.; SCHINOR, E. H.; BASTIANEL, M. Positive relationship between citrus leaf miner and *alternaria* brown spot. **Ciência Rural**, 45(7), 1160-1163. 2015.
- BISCARINI, F.; NAZZICARI, N.; BINK, M.; ARÚS, P.; ARANZANA, M.J.; VERDE, I. MICALI, S.; PASCAL, T.; QUILOT-TURION, B.; LAMBERT, P.; LINGE, C.S.; PACHECO, I.; BASSI, D. STELLA, A.; ROSSINI, L. Genome-enabled predictions for fruit weight and quality from repeated records in European peach progenies. **BMC Genomics**, 18:432, 2017.
- BRASILEIRO, B.P; SILVA, S.A.; SOUZA, D.R.; SANTOS, P.A. OLIVEIRA, R.S.; LYRA, D.H. Genetic diversity and selection gain in the physic nut (*Jatropha curcas*). **Genetics and Molecular Research** 12(3): 2341-2350, 2013.
- CAMPOS, K.A.F.; AZEVEDO, F.A.; BASTIANEL, M.; CRISTOFANI-YALY, M. Resistance to *alternaria* brown spot of new citrus hybrids. **Rev. Bras. Frutic.**, 39(5), 2017.
- CAMPOS, G. De Los; GIANOLA, D.; ROSA, G.J.M. Reproducing kernel Hilbert spaces regression: a general framework for genetic evaluation. **Journal of Animal Science**, 87, 1883-1887, 2009.
- CAO, K.; ZHOU, Z.; WANG, Q.; GUO, J.; ZHAO, P.; ZHU, G.; FANG, W.; CHEN, C.; WANG, X.; WANG, X.; TIAN, Z.; WANG, L. Genome-wide association study of 12 agronomic traits in peach. **Nature Communications**. 2016.
- CUENCA, J.; ALEZA, P.; VICENT, A.; BRUNEL, D.; OLLITRAULT, P.; NAVARRO, L. Genetically Based Location from Triploid Populations and Gene Ontology of a 3.3-Mb Genome Region Linked to *Alternaria* Brown Spot Resistance in Citrus Reveal Clusters of Resistance Genes. **Plos One**. 8(10), 2013.
- CUENCA, J.; ALEZA, P.; GARCI-LOR, A.; OLLITRAULT, P.; NAVARRO, L. Fine Mapping for Identification of Citrus *Alternaria* Brown Spot Candidate Resistance Genes and Development of New SNP Markers for Marker-Assisted Selection. **Front. Plant Sci.** 7, 2016.
- DAETWYLER, A.D.; CALUS, M.P.L.; PONG-WONG, R.; DE LOS CAMPOS, G.; HICKEY, J.M. Genomic Prediction in Animals and Plants: Simulation of Data, Validation, Reporting, and Benchmarking. **Genetics** 193: 347–365, 2013.

GIANOLA, D. Genomic-Assisted Prediction of Genetic Value With Semiparametric Procedures. **Genetics**, 173:1761–1776, 2006.

GODDARD, M.E. Genomic selection: prediction of accuracy and maximization of long term response. **Genetica**, 136(2), 245-257, 2009.

GODDARD, M.E.; HAYES, B.J. Genomic Selection. **Journal of Animal Breeding and Genetics**. 124, 323-330, 2007.

GODDARD ME, HAYES BJ, MEUWISSEN THE. Using the genomic relationship matrix to predict the accuracy of genomic selection. *J. anim. Breed. Genet* 128(6): 409-421, 2011.

GRATTAPAGLIA, D.; RESENDE, M.D.V. Genomic selection in forest tree breeding. **Tree Genetics & Genomes**, v. 7, p. 241-255, 2011.

GRATTAPAGLIA, D.; SILVA-JUNIOR, O.; RESENDE, R.T.; CAPPA, E.P.; MULLER, B.S.F.; TAN, B.; ISIK, F.; RATCLIFFE, B.; EL-KASSABY, Y.A. Quantitative genetics and genomics converge to accelerate forest tree breeding. **Front. Plant Sci.**, 2018.

HAYES, B.J.; CHAMBERLAIN, A.J.; GODDARD, M.E. Genomic selection in dairy cattle: Progress and Challenges. **Journal of Dairy Science**, 2009.

JANNINK, J-L; LORENZ, A.J.; IWATA, H. Genomic selection in plant breeding: from theory to practice. **Briefings in Functional Genomics**. 9(2), 166-177, 2010.

JARRELL, D.C.; ROOSE, M.L.; TRAUGH, S.N.; KUPPER, R.S. A genetic map of citrus based on the segregation of isozymes and RFLPs in an intergeneric cross. **Theory Applied Genetics** 84:49–56, 1992.

JEGER, M.; VILJANEN-ROLLINSON, S. The use of the area under the disease-progress curve (AUDPC) to assess quantitative disease resistance in crop cultivars. *Theoretical and Applied Genetics*, v. 102, n. 1, p. 32-40, 2001.

KOHMOTO, K.; AKIMITSU, K.; OTANI, H. Correlation of resistance and susceptibility of *Citrus* to *Alternaria alternata* with sensitivity to host-specific toxin. *Physiology and Biochemistry*. v.81 (7), 1991.

KUMAR S, BINK MCAM, VOLZ RK, BUS VGM. Towards genomic selection in apple (*Malus × domestica* Borkh.) breeding programmes: Prospects, challenges and strategies. **Tree Genetics & Genomes** 8:1–14, 2012.

LEGARRA, A.; ROBERT, GRANIÉ, C.; MANFREDI, E.; ELSESEN, J.M. Performance of genomic selection in mice. **Genetics** 180: 611-618, 2008.

Li, Y.; RUPERAO, P.; BATLEY, J.; EDWARDS, D.; KHAN, T.; COLMER, T.D.; PANG, J.; SIDDIQUE, K.H.M.; SUTTON, T. Investigating Drought Tolerance in

Chickpea Using Genome-Wide Association Mapping and Genomic Selection Based on Whole-Genome Resequencing Data. **Frontier Plant Science**. 9, 1-12, 2018.

MARTELLI, I. B. **Manejo da mancha marrom de alternária em citros: poda de limpeza e correlação com a lagarta minadora**, Dissertação (Mestrado em Tecnologia da Produção Agrícola) Instituto agrônômico de Campinas, Campinas, SP. p.41, 2011.

MEWUISSEN, T.H.E.; HAYES, B.J.; GODDARD, M.E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. **Genetics**, n. 157, p. 1819–1829, 2001.

MINAMIKAWA, M.F.; NONAKA, K.; KAMINUMA, E.; KAJIYA-KANEGAE, H.; ONOGI, A.; GOTO, S.; YOSHIOKA, T.; IMAI, A.; HAMADA, H.; HAYASHI, T.; MATSUMOTO, S.; KATAYOSI, Y.; TOYODA, A.; HAMADA, H.; FUJIYAMA, A.; NAKAMURA, Y.; SHIMIZU, T.; IWATA, H. Genome-wide association study and genomic prediction in citrus: Potential of genomics-assisted breeding for fruit quality traits. **Scientific Reports**, 7: 4721, 2017.

PACHECO, C.A.; MARTELLI, I. B.; POLYDORO, D. A.; SCHINOR, E. H.; PIO, R. M.; KUPPER, K. C.; AZEVEDO, F. A. Resistance and susceptibility of mandarins and their hybrids to *Alternaria alternata*. *Scientia Agricola*, 69(6), 386-392, 2012.

PATTERSON HD, THOMPSON R. Recovery of inter-block information when block sizes are unequal. **Biometrika** 58: 545-554, 1971.

PEEVER, T.L.; CANILHOS, Y.; OLSEN, L.; IBÁÑEZ, A.; LIU, Y.C.; TIMMER, L.W. Population genetic structure and host specificity of *Alternaria* spp causing brown spot of *Minneola tangelo* and rough lemon in Florida. **Phytopathology**, v.89, p.851-860, 1999.

PEEVER, T.L.; OLSEN, L.; IBÁÑEZ, A. & TIMMER, L.W. Genetic differentiation and host specificity among populations of *Alternaria* spp. causing brown spot of grapefruit and tangerine x grapefruit hybrids in Florida. **Phytopathology**, v.90, p.407-414, 2000.

PEREZ, P.; DE LOS CAMPOS, G. Genome-Wide Regression and Prediction with the BGLR Statistical Package. *Genetics*, 198:483–495, 2014.

R Core Team. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2018. <http://www.R-project.org>.

RESENDE, M. D. V. **Genética biométrica e estatística no melhoramento de plantas perenes**. Brasília: Embrapa Informação Tecnológica, 2002. 975p.

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

Resende MDV, Duarte JB. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesq. Agropec.Trop.** 37: 182-194, 2007.

RESENDE, M.D.V.; AZEVEDO, C.F.; SILVA, F.F. **Seleção Genômica**. In: RESENDE, M.D.V.; SILVA, F.F.; AZEVEDO, C.F. Estatística Matemática, Biométrica e Computacional. Visconde do Rio Branco: Suprema, 2014. Cap. 14. p. 627-768.

RESENDE MDV, LOPES OS, SILVA RL, PIRES IE. Seleção genômica ampla (GWS) e maximização da eficiência do melhoramento genético. **Pesq. Flor. Bra.** 56: 63-77, 2008.

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

RESENDE, M.D.V.; SILVA, F.F.; RESENDE JUNIOR, M.F.R. **Seleção Genômica Ampla**. In: Biotecnologia aplicada ao melhoramento de plantas. BORÉM, A.; FRITSCHÉ-NETO, R. Suprema, 2013, 336p.

RESENDE, M.D.V.; AZEVEDO, C.F.; SILVA, F.F. **Seleção Genômica**. In: RESENDE, M.D.V.; SILVA, F.F.; AZEVEDO, C.F. Estatística Matemática, Biométrica e Computacional. Visconde do Rio Branco: Suprema, 2014. Cap. 14. p. 627-768.

RESENDE, M.D.V. Software Selegen-REML/BLUP: a useful tool for plant breeding. *Crop Breeding and Applied Biotechnology*, 16(4), 330-339, 2016.

STUART, R.M.; BASTIANEL, M.; AZEVEDO, F.A.; MACHADO, M.A. Alternaria Brown Spot. **Laranja**, Cordeirópolis, v.30, n.1, p.29-44, 2009.

TIMMER, L.W.; PEEVER, T.L.; SOLEIL, Z.; AZUYA, K.; KIMITSU, A. Alternaria diseases of citrus-novel pathosystems. **Phytopathologia Mediterranea**, Firenze, v.42, p.99-112, 2003.

VENCOVSKY, R.; BARRIGA, P. **Genética biométrica no fitomelhoramento**. Ribeirão Preto: Sociedade Brasileira de Genética, 1992. 496p.

CONCLUSÕES GERAIS

A GWS é uma estratégia de seleção que apresenta potencial para ser utilizada em programas de melhoramento genético de *Citrus*, visto que há a possibilidade de redução do ciclo seletivo e, conseqüentemente, ganho por unidade de tempo.

A obtenção de altas acurácias na GWS para a seleção de características classificadas como qualitativas, características que sofrem pouca influência do ambiente e é governada por um pequeno número de genes, irá depender diretamente da utilização de marcadores moleculares que estejam em alto desequilíbrio de ligação com os genes responsáveis por tais características.