

Seleção Genômica para Tolerância ao Déficit Hídrico em Milho

Kaio Olímpio das Graças Dias⁽¹⁾; Salvador Alejandro Gezan⁽²⁾; Claudia Teixeira Guimarães⁽³⁾; Roberto Willians Noda⁽³⁾; João Cândido de Souza⁽⁴⁾; Maria Marta Pastina⁽³⁾; Lauro José Moreira Guimarães⁽³⁾.

⁽¹⁾ Doutorando em Genética e Melhoramento de Plantas, Universidade Federal de Lavras (UFLA), Lavras, Minas Gerais, kaioolimpio@hotmail.com; ⁽²⁾ Universidade da Florida (UFL), Gainesville, Florida, Estados Unidos; ⁽³⁾ Embrapa Milho e Sorgo, Sete Lagoas, Minas Gerais; ⁽⁴⁾ Departamento de Biologia, Universidade Federal de Lavras (UFLA), Lavras, Minas Gerais.

RESUMO: Em milho, o déficit hídrico é uma das principais causas de perdas na produção de grãos. Devido às dificuldades do melhoramento convencional para tolerância a seca, o uso da seleção genômica pode resultar em maior eficiência de seleção. Este trabalho teve como objetivo avaliar a acurácia da seleção genômica para a predição dos valores genotípicos de híbridos simples de milho para tolerância ao déficit hídrico, utilizando modelos genético-estatísticos que incorporam o efeito da interação entre genótipos e ambientes (G x E), e também, efeitos genéticos aditivos e de dominância. Para isso, foram utilizados dados fenotípicos de produção de grãos obtidos para um conjunto de 310 híbridos simples de milho, avaliados em dois locais nos anos de 2010 e 2011, em experimentos com e sem estresse hídrico. Os genótipos dos híbridos foram inferidos com base nos dados genotípicos de marcadores SNP (*Single Nucleotide Polymorphism*) obtidos para seus genitores (linhagens) via genotipagem por sequenciamento. As análises de seleção genômica foram realizadas pelo método GBLUP (*Genomic Best Linear Unbiased Prediction*). A partir da incorporação do efeito de G x E, bem como dos efeitos aditivos e de dominância, nos modelos de seleção genômica, foram obtidas maiores acurácias para a predição dos valores genotípicos dos híbridos simples, principalmente na condição de estresse hídrico. Esses resultados contribuem para um melhor entendimento da arquitetura genética do caráter produção de grãos, e destacam a importância dos efeitos de dominância para a predição da performance de cruzamentos em programas de melhoramento para tolerância ao déficit hídrico em milho.

Termos de indexação: G x E, predições genômicas, efeitos aditivos e de dominância.

INTRODUÇÃO

O desenvolvimento de cultivares tolerantes ao déficit hídrico representa uma estratégia eficiente para o aumento da produtividade de milho sob condições tropicais, uma vez que o estresse hídrico é uma das principais causas de perdas na produção (Ribaut et al., 2009). No Brasil, as áreas de milho safrinha estão em plena expansão, sendo superiores em relação aos plantios de verão (CONAB, 2016). Como o cultivo de safrinha está mais sujeito à ocorrência de seca, o melhoramento tem um papel fundamental, pois pode auxiliar na produção de milho em épocas de escassez hídrica, dando certa segurança aos agricultores.

Recentemente, com o desenvolvimento de técnicas de genotipagem em larga escala a um custo reduzido e a disponibilidade de métodos estatísticos e recursos computacionais para o processamento de grande volume de dados, têm sido propostas novas abordagens de melhoramento assistido para caracteres de herança quantitativa, como a seleção genômica (Meuwissen et al., 2001), que utiliza informações de marcadores moleculares distribuídos ao longo do genoma para estimar os valores genéticos dos indivíduos.

Para tolerância ao déficit hídrico, a seleção genômica pode resultar em maior eficiência de seleção devido a natureza complexa do caráter, e também, as dificuldades do melhoramento convencional para a avaliação dos materiais genéticos em ambientes com déficit hídrico, que é demorado e laborioso, visto que as condições experimentais devem ser cuidadosamente

manejaadas. Estudos recentes têm relatado as vantagens do uso da seleção genômica para a tolerância ao déficit hídrico em milho (Zyomo & Bernardo, 2013; Beyene et al., 2015; Zhang et al., 2015). No entanto, esses estudos ficaram restritos a incorporação de efeitos aditivos nos modelos de seleção genômica, o que pode não ser realístico para espécies com alto nível de heterose, como é o caso do milho.

Em programas de melhoramento os genótipos são avaliados em múltiplos ambientes, ou seja, ao longo de diferentes anos e locais. Assim, para a identificação de genótipos com maior adaptabilidade e estabilidade, é necessário que o efeito de G x E seja considerado. Modelos genético-estatísticos que incorporam o efeito de G x E têm aumentado a acurácia da seleção genômica (Heslot et al., 2014). No entanto, são escassos, em milho e outras espécies, estudos que incorporam simultaneamente o efeito de G x E, os efeitos genéticos aditivos e de dominância nos modelos de seleção genômica.

Nesse contexto, o presente trabalho teve como objetivo avaliar a acurácia da seleção genômica para a predição dos valores genotípicos de híbridos simples de milho para tolerância ao déficit hídrico, utilizando modelos genético-estatísticos que incorporam o efeito de G x E, e também, efeitos genéticos aditivos e de dominância.

MATERIAL E MÉTODOS

Dados Fenotípicos

Foram avaliados 310 híbridos simples de milho (190 linhagens cruzadas com dois testadores, sendo algumas cruzadas com apenas um testador), em dois locais (Janaúba-MG e Teresina-PI) nos anos de 2010 e 2011, em experimentos com (WS, do inglês *water-stressed*) e sem (WW, do inglês *well-watered*) estresse hídrico. Em cada local, foram instalados três ensaios, cada um seguindo o delineamento de blocos completos casualizados, com três e duas repetições para WS e WW. Foram incluídas 4 testemunhas comuns entre os ensaios. As parcelas experimentais foram constituídas por duas linhas de quatro metros com espaçamento de 0,8 metros entre linhas. Em Janaúba, o estresse hídrico foi induzido a partir do corte da irrigação por um período de 30 com início aos 45 dias após plantio. Em Teresina, foi feito um manejo da quantidade de água no solo para indução do estresse hídrico. A combinação de locais, anos e regimes hídrico foram considerados como ambientes. O caráter avaliado foi produção de grãos ($\text{kg}\cdot\text{ha}^{-1}$). Os tratos culturais foram feitos conforme recomendações para a cultura do milho.

Dados Genotípicos

Para as 190 linhagens utilizadas como genitores dos híbridos simples avaliados, foram obtidos dados de marcadores SNP via genotipagem por sequenciamento (GBS, do inglês *Genotyping-by-Sequencing* - Elshire et al., 2011). Marcadores SNP monomórficos, ou seja, com frequência do alelo menos comum (MAF, do inglês *Minor Allele Frequency*) inferior a 5%, SNPs com porcentagem de dados perdidos superior a 20%, e/ou com mais de 5% de genótipos em heterozigose, foram descartados. Em seguida, o software NPUTE (Roberts et al., 2007) foi utilizado para a imputação de dados perdidos. Vale ressaltar que os genótipos em heterozigose restantes para alguns marcadores foram considerados como dados perdidos para o procedimento de imputação. Posteriormente, os genótipos dos híbridos simples foram inferidos a partir do genótipo de seus genitores (linhagens).

Modelos de Seleção Genômica

As análises de seleção genômica foram realizadas com base no método GBLUP usando o software Asreml-R v.3 (Butler et al., 2009), de acordo com o seguinte modelo:

$$y = \mu\mathbf{1} + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{r} + \mathbf{Z}_2\mathbf{a.s} + \mathbf{Z}_3\mathbf{d.s} + \mathbf{e}$$

em que: y é o vetor de observações fenotípicas; μ é a constante; $\mathbf{1}$ é um vetor de uns; $\boldsymbol{\beta}$ é o vetor de efeitos fixos de ambientes e ensaios dentro de ambientes; \mathbf{r} é o vetor de efeito aleatório de repetições dentro de ensaios, com $\mathbf{r} \sim \text{MVN}(\mathbf{0}, \mathbf{D}_r)$; $\mathbf{a.s}$ é o vetor aleatório de efeitos aditivos de híbridos dentro de ambientes, com $\mathbf{a.s} \sim \text{MVN}(\mathbf{0}, \mathbf{G} \otimes \mathbf{A})$; $\mathbf{d.s}$ é o vetor aleatório de efeitos de dominância de híbridos dentro de ambientes, com $\mathbf{d.s} \sim \text{MVN}(\mathbf{0}, \mathbf{G} \otimes \mathbf{D})$; \mathbf{e} é o vetor de resíduos, com, $\mathbf{e} \sim \text{MVN}(\mathbf{0}, \mathbf{D}_e)$. \mathbf{G} é uma matriz de variância e covariância com estrutura fator analítico de ordem 2 (FA_2), conforme proposto por Smith et al. (2001). \mathbf{D}_r e \mathbf{D}_e são matrizes diagonal, onde cada ambiente tem um componente de variância para as repetições e os resíduos, respectivamente. \mathbf{A} e \mathbf{D} são as matrizes de parentesco aditivo e de dominância, obtidas como descrito em VanRaden (2008) e Vitezica et al. (2013), respectivamente.

A acurácia preditiva dos modelos com a incorporação de apenas efeitos aditivos (modelo A) e de efeitos aditivos e de dominância (modelo AD) foram avaliadas a partir de um procedimento de validação cruzada, usando os 310 híbridos

avaliados em ambientes com e sem estresse hídrico como população de treinamento e validação. Tal procedimento de validação cruzada foi repetido cinco vezes. Ao final, a acurácia do modelo de seleção genômica foi verificada através da correlação entre o valor fenotípico observado e o valor genotípico predito para os híbridos simples.

RESULTADOS E DISCUSSÃO

O modelo AD permitiu a decomposição da variância genética em variância aditiva e de dominância, e apresentou menor valor de AIC quando comparado ao modelo A. Estimativas de herdabilidade no sentido amplo e no sentido restrito apresentaram menores valores nas condições de déficit hídrico (**Tabela 1**).

Quando o efeito de dominância foi incluído no modelo, a herdabilidade no sentido restrito diminuiu tanto em WS quanto em WW (**Tabela 1**). Essa redução pode ser devido a distribuição das frequências alélicas, ou seja, quando a frequência alélica está próxima a valores extremos, mesmo com a presença de efeitos não-aditivos, parte desses efeitos são contabilizados para estimar a variância aditiva (Hill et al., 2008). Em outras espécies, estudos recentes têm reportado essa mesma tendência de parte da variância não-aditiva ser estimada como variância aditiva, por exemplo em pinus e eucalipto (Muñoz et al., 2014). Esses resultados deixam evidente que o uso apenas de efeitos aditivos em modelos de seleção genômica pode superestimar a herdabilidade no sentido restrito, assim como o ganho predito com a seleção.

Tabela 1 – Estimativas dos parâmetros genéticos nos ambientes com (WS) e sem (WW) estresse hídrico para os modelos aditivo (A) e aditivo-dominância (AD).

	WS		WW	
	A	AD	A	AD
h^2	0,345	0,191	0,397	0,264
δ^2	-	0,140	-	0,193
H^2	-	0,331	-	0,457
ρ_A	0,678	0,411	0,670	0,547
ρ_D	-	0,589	-	0,399
AIC	4788,4	4567,9	3348,8	3342,1

h^2 herdabilidade no sentido restrito, δ^2 proporção da variância de dominância em relação a variância fenotípica, H^2 herdabilidade no sentido amplo, ρ_A correlação genética aditiva entre ambientes, ρ_D correlação genética de dominância entre ambientes.

Considerando os resultados do modelo AD, maior G x E foi encontrada na condição de estresse hídrico para os efeitos aditivos (**Tabela 1**). Já para os efeitos de dominância, foi observada um menor G x E. Modelos que levam em consideração a G x E são essenciais para a seleção de genótipos estáveis nas condições com e sem estresse hídrico. Os modelos usados no presente trabalho levam em conta a existência de heterogeneidade de variâncias genéticas e as correlações genéticas entre ambientes. Além disso, esses modelos permitem trabalhar com dados desbalanceados, que são comuns na rotina de um programa de melhoramento, uma vez que os conjuntos de materiais avaliados são distintos ao longo dos anos. Zhang et al. (2015) demonstraram que a acurácia de predição foi maior quando modelos que levem em conta a G x E foram usados para seleção genômica em populações biparentais de milho.

Diferenças na acurácia de predição entre os modelos A e AD foram encontradas para produção de grãos (**Figura 1**). No entanto, essa diferença entre os modelos foi mais evidente na condição de estresse hídrico, onde as acurácias foram de 0.245 e 0.508 com os modelos A e AD, respectivamente. Essas diferenças destacam a importância dos efeitos de dominância para a predição da performance de híbridos para o caráter produção de grãos em condições de déficit hídrico. Existem duas possíveis explicações para esses resultados: (i) a herdabilidade foi maior na condição sem estresse hídrico quando comparada a condição com estresse hídrico. De acordo com Da et al. (2014), modelos AD tendem a não trazer vantagens para características de alta herdabilidade; (ii) a variância de dominância foi maior em condições de déficit hídrico. Almeida Filho et al. (2016) mostrou que, em situações onde os efeitos de dominância são pequenos, não existem vantagens em usar modelos AD.

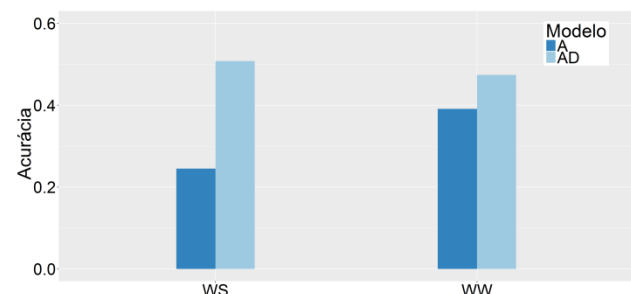


Figura 1. Acurácias de predição para produção de grãos em condições com (WS) e sem (WW) estresse hídrico a partir de um modelo genético-estatístico que leva em conta o efeito de G x E, e também, efeitos genéticos aditivos (A) e aditivos

e de dominância (AD).

Embora estudos recentes têm relatado o uso da seleção genômica para tolerância ao déficit hídrico em milho (Zyomo & Bernardo, 2013; Beyene et al., 2015; Zhang et al., 2015), em geral, esses estudos não levaram em conta efeitos de dominância nos modelos de predição, assim como efeitos de dominância e efeitos de G x E simultaneamente. Assim, os resultados do presente trabalho enfatizam a importância dos efeitos de dominância para a predição da performance de híbridos simples para o caráter produção de grãos em ambientes com e sem estresse hídrico. Além disso, esses resultados contribuem para um melhor entendimento da arquitetura genética da produção de grãos sob déficit hídrico, com base em modelos mais adequados para a realidade dos programas de melhoramento, onde os dados fenotípicos são geralmente desbalanceados, avaliados em múltiplos anos e locais.

Os resultados descritos no presente trabalho mostram que modelos que incorporam simultaneamente os efeitos de G x E e os efeitos genéticos aditivos e de dominância devem ser considerados para a identificação de híbridos superiores com maior estabilidade em ambientes específicos. Além disso, modelos AD podem ser usados para a seleção genômica em outras espécies com o objetivo de explorar a heterose.

CONCLUSÕES

Foram observadas diferenças nas acurácias de predição entre os modelos A e AD. No entanto, essas diferenças foram mais evidentes na condição de déficit hídrico. Esses resultados contribuem para um melhor entendimento da arquitetura genética do caráter produção de grãos em ambientes com estresse hídrico, e destaca a importância dos efeitos de dominância para a predição da performance de híbridos simples de milho.

AGRADECIMENTOS

Os autores agradecem às agências de fomento: FAPEMIG, CNPq e CAPES pelo auxílio financeiro.

REFERÊNCIAS

ALMEIDA FILHO, J. E.; GUIMARÃES, J. F. R.; SILVA, F. F.; RESENDE, M. V.; MUÑOZ, P.; KIRST, M. RESENDE JR, M. F. R. The contribution of dominance to phenotype prediction in a pine breeding and simulated population. *Heredity*, v. 117, n. 1, p. 33-41, 2016.

BEYENE, Y.; SEMAGN, K.; MUGO, S.; TAREKEGNE, A.; BADU, R.; MEISEL, B.; SEHABIAGUE, P.; MAKUMBI, D.; MAGOROKOSHO, C.; OIKEH, S.; GAKUNGA, J.; VARGAS, M.; OLSEN, M.; PRASANNA, B. M.;

BANZIGER, M.; CROSSA, J. Genetic gains in grain yield through genomic selection in eight bi-parental maize populations under drought stress. *Crop Science*, v. 55, p.154-163, 2015.

BUTLER, D. G; CULLIS, B. R; GILMOUR, A. R; GOGEL, B. J. ASReml-R reference manual, release 3. 160 technical report, Queensland Department of Primary Industries. 2009.

CONAB - Companhia Nacional de Abastecimento. 4º Levantamento de Grãos Safra 2015/2016. Janeiro 2016.

DA, Y.; WANG, C.; WANG, S.; HU, G. Mixed model methods for genomic prediction and variance component estimation of additive and dominance effects using SNP markers. *PLoS one*. v. 9, p. 1-11, 2014.

ELSHIRE, R. J.; GLAUBITZ, J. C.; SUN, Q.; POLAND, J. A.; KAWAMOTO, K.; BUCKLER, E.; MITCHELL, S. E. A Robust Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *Plos One*, v. 6, p. 1-10, 2011.

HESLOT, N.; AKDEMIR, D.; SORRELLS, M. E.; JANNINK, J. L. Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theoretical and Applied Genetics*, v. 127, p. 463-489, 2014.

HILL, W.; GODDARD, M.; VISSCHER, P. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics*, v. 4, p. 1-10, 2008.

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

MUÑOZ, P. R.; RESENDE JR, M. F. R.; GEZAN, S. A.; RESENDE, M. D. V.; DE LOS CAMPOS, G.; KIRST, M.; HUBER, D.; PETER, G. R. Unraveling additive from nonadditive effects using genomic relationship matrices. *Genetics*, v. 198, p. 1759-1768, 2014.

RIBAUT, J.M.; BETRAN, J.; MONNEVEUX, P.; SETTER, T. (2009). Drought tolerance in maize. In: BENNETZEN, J.L.; HAKE, S.C. (Ed.). Handbook of maize: its biology, p 311-344.

ROBERTS, A.; MCMILLAN, L.; WANG, W.; PARKER, J.; RUSYN, I.; THREADGILL, D. Inferring missing genotypes in large SNP panels using fast nearest-neighbor searches over sliding windows. *Bioinformatics*, v. 23, p. 401-407, 2007.

SMITH, A.; CULLIS, B. R.; THOMPSON, R. Analyzing variety by environment data using multiplicative mixed models and adjustment for spatial field trend. *Biometrics*, v. 57, p. 1138-47, 2001.

VANRADEN, P. M. Efficient methods to compute genomic predictions. **Journal of Dairy Science**, v. 91, p. 4414-4423, 2008.

VITEZICA, Z. G.; VARONA, L.; LEGARRA, L. On the additive and dominant variance and covariance of individuals within the genomic selection scope. **Genetics** v. 195, p. 1223-1230, 2013.

ZIYOMO, C.; BERNARDO, R. Drought tolerance in maize-indirect selection through secondary traits versus genome wide selection. **Crop Science** v. 52, p. 1269-1275, 2013.

ZHANG, X.; PÉREZ-RODRÍGUEZ, P.; SEMAGN, K.; BEYENE, Y.; BABU, R.; LÓPEZ-CRUZ, M. A.; SAN VICENTE, F.; OLSEN, M.; BUCKLER, E.; JANNINK, J. L.; PRASANNA, B. M.; CROSSA, J. Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. **Heredity**, v. 114, p. 291-299, 2015.