

# Seleção genômica para melhoramento vegetal com diferentes estruturas populacionais

Mágno Sávio Ferreira Valente<sup>(1)</sup>, José Marcelo Soriano Viana<sup>(1)</sup>, Marcos Deon Vilela de Resende<sup>(2)</sup>, Fabyano Fonseca e Silva<sup>(1)</sup> e Maria Teresa Gomes Lopes<sup>(3)</sup>

<sup>(1)</sup>Universidade Federal de Viçosa, Avenida Peter Henry Rolfs, s/nº, CEP 36570-000 Viçosa, MG, Brasil. E-mail: magnosavio@yahoo.com.br, jmsviana@ufv.br, fabyanofonseca@ufv.br <sup>(2)</sup>Embrapa Florestas, Estrada da Ribeira, Km 111, CEP 83411-000 Colombo, PR, Brasil. E-mail: marcos.deon@gmail.com <sup>(3)</sup>Universidade Federal do Amazonas, Avenida General Rodrigo Otávio, nº 6.200, CEP 69077-000 Manaus, AM, Brasil. E-mail: mtglopes@hotmail.com

Resumo – O objetivo deste trabalho foi avaliar a eficiência da seleção genômica em diferentes cenários de estrutura populacional em milho-pipoca, com estimação dos efeitos e uso de marcadores na própria população de referência e em populações não relacionadas, e determinar a influência do tamanho efetivo e das relações de parentesco na população de estimação sobre a acurácia da predição. Foram simuladas populações com diferentes desequilíbrios de ligação (LD) e variâncias aditivas, tendo-se considerado diferentes caracteres, densidades de marcadores, herdabilidades e gerações, no total de 144 cenários. Também foram simuladas populações estruturadas em progênies. A acurácia da predição dos valores genéticos aditivos foi obtida por meio da correlação entre os valores paramétricos e os valores estimados por RR-BLUP (ridge regression-best linear unbiased prediction). Em populações com baixo LD e menor variância aditiva, o uso de maiores densidades de SNP (10 SNP 0,1 cM<sup>-1</sup>) é indicado, e, além disso, o candidato à seleção deve ser relacionado à população de estimação, para que a acurácia de predição seja satisfatória. O uso de população de seleção na mesma geração da população de estimação reduz em pelo menos 8% a acurácia. A estruturação da população em progênies de maior relacionamento e menor tamanho efetivo aumenta a eficiência da seleção genômica.

Termos para indexação: *Zea mays*, acurácia da predição, eficiência da seleção genômica, predição de valor genético, relação de parentesco, tamanho efetivo.

## Genomic selection for plant breeding with different population structures

Abstract – The objective of this work was to evaluate the efficiency of genomic selection in different population structure scenarios in popcorn, by estimating the effects of and using markers in the same reference population and in populations not related to each other, and to determine the effect of the effective size and kin relationships on the estimation population over the prediction accuracy. Populations were simulated with different linkage disequilibria (LD) and additive variances, taking into account different characters, marker densities, heritabilities, and generations, totaling 144 scenarios. Populations structured in progenies were also simulated. The prediction accuracy of additive breeding values was obtained by correlating true breeding values and breeding values predicted by RR-BLUP (ridge regression-best linear unbiased prediction). In populations with low LD and lower additive variance, the use of higher SNP density (10 SNP 0.1 cM<sup>-1</sup>) is recommended; besides, the selection candidate should be related to the estimation population in order to obtain satisfactory prediction accuracy. The use of a selection population in the same generation of the estimation population reduces accuracy in at least 8%. Populations structured in progenies with stronger relationships and smaller effective size increase the efficiency of genomic selection.

Index terms: *Zea mays*, prediction accuracy, genomic selection efficiency, breeding value prediction, kin relationship, effective size.

## Introdução

A seleção genômica ampla (genome wide selection, GWS) possibilita a identificação de indivíduos superiores a partir de predições de seus valores genéticos aditivos, obtidas pela análise de milhares

de dados de polimorfismo de DNA e de um reduzido conjunto de dados fenotípicos (Meuwissen et al., 2001). Seu fundamento reside na expectativa de ocorrência de desequilíbrio de ligação (LD) entre os marcadores moleculares e as regiões controladoras de características quantitativas (quantitative trait loci,

QTL). Em razão da alta densidade de marcadores utilizada no processo, espera-se que a seleção genômica seja capaz de capturar os efeitos de todos os QTLs determinantes do caráter e, dessa forma, melhorar a predição dos valores genéticos.

Muitos trabalhos têm comprovado que a seleção genômica é eficiente e pode ser utilizada no melhoramento vegetal (Daetwyler et al., 2013; De Los Campos et al., 2013). Embora Jonas & De Koning (2013) considerem que a utilização da seleção genômica no melhoramento vegetal não seja fácil, em razão da diversidade de procedimentos seletivos, a metodologia tem sido eficientemente utilizada na predição de valores genotípicos de híbridos simples de milho, trigo e arroz (Zhao et al., 2013, 2015; Xu et al., 2014); na predição de efeitos de capacidade geral de combinação de “testcrosses”, bem como de valores genotípicos de linhagens e linhas duplo-haploides de milho (Crossa et al., 2013; Albrecht et al., 2014); na predição de valores genotípicos de linhas puras, no melhoramento de autógamias (Crossa et al., 2014); e no melhoramento recorrente de alógamas (Yabe et al., 2013). Heslot et al. (2015) e Poland (2015) concordam que a seleção genômica proporcionará ganhos genéticos relevantes ao melhoramento vegetal, principalmente em caracteres de baixa herdabilidade, de difícil mensuração e de avaliação dispendiosa, uma vez que ela diminui a duração do ciclo de seleção.

Os modelos de predição, criados a partir da GWS com uso de informações de dados genotípicos e fenotípicos, permitem que somente informações genotípicas sejam usadas para prever fenótipos futuros. A maioria dos estudos sobre a eficiência da seleção genômica considera, em seus modelos de análise, apenas uma população, além de suas gerações subsequentes, uma vez que os modelos preditivos são usados principalmente em populações e ambientes específicos (Resende Jr et al., 2012). Todavia, o interesse da pesquisa pode estar relacionado à seleção de indivíduos de outras populações, em que o LD entre marcador e QTL pode ser diferente.

Os benefícios com o uso de populações de estimação e de seleção fortemente relacionadas são inegáveis. Entretanto, pouco se sabe sobre o uso dessas populações na ausência de parentesco ou de relacionamento, mas com valores similares de variância genética e LD, por exemplo. Além disso, outras questões também merecem ser esclarecidas, como o efeito do uso de

diferentes amostras como população de estimação e de seleção, sobre a perda mínima da acurácia esperada em populações relacionadas. Além disso, ainda que a queda do LD entre o QTL e o marcador molecular seja muitas vezes citada como a principal responsável pela redução da acurácia ao longo das gerações, outros fatores também podem exercer grande influência sobre essa variável e devem ser mais bem estudados, como a variância genética e o parentesco entre indivíduos.

O objetivo deste trabalho foi avaliar a eficiência da seleção genômica em diferentes cenários de estrutura populacional em milho-pipoca, com estimação dos efeitos e uso de marcadores na própria população de referência e em populações não relacionadas, e determinar a influência do tamanho efetivo e das relações de parentesco na população de estimação sobre a acurácia da predição.

## Material e Métodos

A eficiência da seleção genômica, com uso de diferentes populações de seleção e estimação, foi avaliada a partir da análise de dados simulados. Utilizou-se o programa Realbreeding, também utilizado em outros trabalhos de simulação (Viana et al., 2013, 2016), desenvolvido a partir do programa Realbasic. Consideraram-se dois caracteres de milho-pipoca, determinados por: 100 QTLs; três densidades de marcadores SNP (single nucleotide polymorphism); duas herdabilidades; e seis populações, representadas em duas gerações, o que totalizou 144 cenários. Para cada cenário, foram realizadas 30 simulações.

Os 100 QTLs foram distribuídos em 10 cromossomos, com 10 QTLs por cromossomo. Em cada cromossomo, foram inseridos 500, 51 ou 6 SNPs, em que a distância média entre SNPs adjacentes foi de 0,1, 1 e 10 cM, respectivamente. Assim, em cada cromossomo, os QTLs e SNPs foram distribuídos, em média, em 50 cM. Os dois caracteres de milho-pipoca foram simulados com valores genotípicos mínimos e máximos dos homocigotos de 20 e 200 g por planta, para a produção de grãos, e de 5 e 50 mL g<sup>-1</sup>, para a capacidade de expansão (CE). Assumiu-se dominância unidirecional positiva ( $0 < d/a \leq 1,2$ ), para a produção de grãos, e dominância bidirecional ( $-1,2 \leq d/a \leq 1,2$ ), para CE. As estimativas de herdabilidade ( $h^2$ ) foram de 0,3 e 0,7. Portanto, os valores de acurácia fenotípica [ $(h^2)^{0,5}$ ] foram de 0,55 e 0,84, respectivamente. O

tamanho amostral foi de 500 indivíduos genotipados e fenotipados.

Neste estudo, tomou-se como população de referência (população 0) um composto obtido de duas populações em equilíbrio de ligação, com diferença mínima de frequência alélica de 0,6, para QTLs e SNPs. Populações adicionais foram consideradas com o intuito de observar o impacto sobre a acurácia da predição, com o uso de diferentes populações de seleção. As populações adicionais incluíram populações com LD similar (populações 1, 2, 3), LD menor (população 4) e LD maior (população 5) do que o LD da população de referência. A estatística usada para medir o LD foi a diferença entre as frequências gaméticas observadas e esperadas com equilíbrio de ligação gênica (D). A média dos valores de D foram próximas de 0,09 (populações 0, 1, 2 e 3), 0,04 (4) e 0,15 (5), ao se levar em consideração a geração 0. De cada população, derivou-se uma nova população, após cinco gerações de acasalamentos ao acaso, sem seleção, mutação e migração. Na geração 5, os valores de D diminuíram e ficaram próximos de 0,05, 0,02 e 0,09, respectivamente às que acabamos de mencionar. Outros parâmetros também foram considerados, como variância aditiva e de dominância similar (populações 1, 4 e 5), variância aditiva menor (2) e os de variância aditiva maior (3) em relação à população de referência (Tabela 1).

Novo conjunto de dados foi tomado, a fim de se estudar o efeito do grau de parentesco entre indivíduos sobre a eficiência da seleção genômica. Estruturas de populações comumente usadas no melhoramento vegetal foram simuladas. Para tal, 500 indivíduos foram estruturados em progênies de meios-irmãos (FMI), de irmãos completos (FIC) e endógamos (progênies S1, FS1). A população de referência foi formada por 500 indivíduos não aparentados (OP). As análises consideraram 10 famílias, com 50 indivíduos cada, que geravam populações com tamanho efetivo ( $N_e$ ) de 38, 20 e 10, para FMI, FIC e FS1, respectivamente. Análises de 20 famílias, com 25 indivíduos, também foram consideradas e proporcionaram  $N_e$  de 72, 38 e 20, respectivamente. As estimativas de tamanho efetivo foram obtidas por meio das seguintes expressões (Resende & Barbosa, 2005):  $N_e = [(4N_f n) / (n + 3)]$  para meio-irmãos;  $N_e = [(2N_f n) / (n + 1)]$  para irmãos completos; e  $N_e = [(N_f n) / (n + 0,5)]$  para irmãos S1; em que  $n$  é o número de indivíduos por família, e  $N_f$

é o número de genitores femininos. Para a população de polinização aberta, o tamanho efetivo considerado foi de 1.000 indivíduos. Os dois caracteres (capacidade de expansão e produção de grãos) foram determinados por 100 QTLs, distribuídos em 10 cromossomos (10 QTLs por cromossomo), duas herdabilidades (0,3 e 0,7) e 2.000 marcadores SNPs, com distância média entre marcadores adjacentes de 0,1 cM. Para cada cenário, realizaram-se 10 simulações. As simulações foram conduzidas para que se obtivesse magnitude similar de LD entre as populações. Assim, os valores de LD variaram aproximadamente de 0,03 a 0,04.

Vários métodos de predição de valores aditivos, por meio de valores genômicos, têm sido empregados em diferentes estudos. Em geral, eles têm apresentado resultados diferentes, de acordo com a estrutura da população e natureza do caráter avaliado. Em razão de sua alta eficiência, fácil implementação e ampla utilização, adotou-se o método RR-BLUP no presente trabalho. Para a obtenção dos valores genômicos por meio deste método, utilizou-se o pacote rrBLUP (Endelman, 2011) do pacote estatístico R (R Core Team, 2013). Assumiu-se o modelo aditivo, e a acurácia de predição dos valores aditivos foi obtida

**Tabela 1.** Média e variâncias aditiva ( $\sigma_A^2$ ) e de dominância ( $\sigma_D^2$ ), para a capacidade de expansão e produção de grãos, em seis populações e duas gerações de milho-pipoca.

Geração	Capacidade de expansão (mL g <sup>-1</sup> )			Produção de grãos (g por planta)		
	Média	$\sigma_A^2$	$\sigma_D^2$	Média	$\sigma_A^2$	$\sigma_D^2$
População 0						
0	29,22	2,56	0,59	143,02	35,78	17,17
5	29,22	2,54	0,58	143,02	35,33	15,55
População 1						
0	29,36	2,45	0,55	144,34	33,07	17,07
5	29,36	2,45	0,55	144,34	33,03	15,48
População 2						
0	30,44	1,76	0,46	148,06	25,79	22,11
5	30,44	1,85	0,47	148,06	27,19	19,90
População 3						
0	28,03	3,38	0,55	142,75	51,57	18,48
5	28,03	3,27	0,54	142,75	49,69	16,60
População 4						
0	30,37	2,47	0,51	143,90	33,11	8,17
5	30,37	2,48	0,51	143,90	33,23	7,85
População 5						
0	27,83	2,45	0,64	140,09	38,76	37,62
5	27,83	2,55	0,64	140,09	40,22	32,57

pela correlação entre os valores paramétricos, gerados pelo REALbreeding, e os valores genômicos, obtidos por RR-BLUP.

O modelo linear misto utilizado foi:  $y = 1u + Zg + \varepsilon$ , em que:  $y$  é um vetor coluna de fenótipos;  $1$  é um vetor coluna que contém o número 1 representado  $N$  (número de indivíduos) vezes;  $Z$  é a matriz de incidência que aloca o genótipo de cada loco marcador a cada indivíduo;  $u$  é um escalar que contém a média populacional;  $g$  é um vetor dos efeitos aleatórios de marcadores; e  $\varepsilon$  é um vetor de erros aleatórios.

As equações de modelo misto equivalem à expressão:

$$\begin{bmatrix} 1'1 & 1'Z \\ Z'1 & Z'Z + I \frac{\sigma_e^2}{\sigma_g^2} \end{bmatrix} \begin{bmatrix} \hat{u} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} 1'y \\ Z'y \end{bmatrix}$$

A variação genética contribuída por um loco marcador  $\sigma_g^2$  foi dada por

$$\sigma_g^2 = \sigma_a^2 / [2 \sum_{i=1}^n p_i (1 - p_i)],$$

em que:  $\sigma_a^2$  é a variância genética aditiva do caráter, e  $p_i$  é a frequência alélica no loco marcador  $i$ .

Depois de estimados os efeitos dos marcadores, os valores aditivos genômicos dos indivíduos foram obtidos por  $\hat{a} = Z\hat{g}$ . A matriz de incidência  $Z$  contém os valores  $Z_{ij}$  iguais a 0, 1 e 2, para os marcadores dos tipos mm, Mm e MM, respectivamente (Resende et al., 2012).

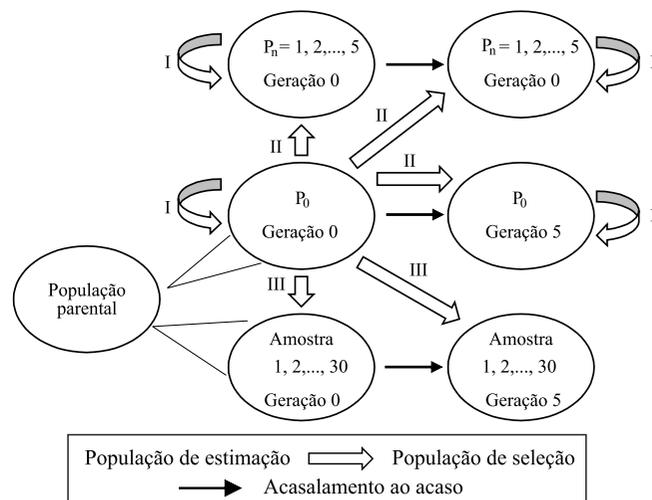
A eficiência da seleção genômica foi obtida sob diferentes cenários (Figura 1): estimacão e uso dos efeitos dos marcadores na própria população de referência; estimacão e uso em populações não relacionadas; e estimacão e uso em diferentes amostras da população de referência. De acordo com o cenário de análise considerado, as populações foram usadas como população de referência, tendo sido genotipadas e fenotipadas; ou população de seleçao, caso em que foram apenas genotipadas.

A avaliacão do efeito do grau de parentesco sobre a estimacão do valor genético (GBV) foi obtida, tendo-se assumido estimacão e uso dos efeitos dos marcadores na própria população. O procedimento de estimacão e uso de marcadores genéticos na mesma população é pouco interessante em termos práticos, em razão de seu alto custo de genotipagem e fenotipagem, e pelo fato de, muitas vezes, tal procedimento gerar valores superestimados de acurácia, quando o objetivo é a utilizacão do modelo em outra amostra. No entanto, esse

procedimento pode ser uma ferramenta interessante em estudos de simulacão, uma vez que é possível isolar fatores, em comparacão com a utilizacão de múltiplas populações. Os valores de acurácia obtidos nesse tipo de análise devem ser entendidos como acurácia máxima de predicão para o conjunto de dados em estudo.

## Resultados e Discussão

Quando os efeitos dos SNPs foram estimados e usados na própria população, independentemente do caráter em estudo, o aumento da acurácia de predicão foi proporcional ao aumento da herdabilidade, embora esta nem sempre tenha estado associada a altos incrementos (Tabela 2). O aumento da acurácia variou de 8 a 37%, valores abaixo do aumento da acurácia do valor fenotípico (53%), que passou de 0,548 para 0,837. Com a reduçao da herdabilidade, menor proporçao da variância genética aditiva é explicada por marcadores genéticos e, portanto, menores acurácias são obtidas (Goddard & Hayes, 2007). A proporcionalidade entre herdabilidade ( $h^2$ ) e acurácia ( $r_{gg}$ ) da GWS pode ser



**Figura 1.** Esquema representativo da avaliacão da eficiência da seleçao genômica, em diferentes cenários. I, estimacão e avaliacão dos efeitos dos marcadores na própria população de referência; II, estimacão e avaliacão em populações diferentes que não compartilham parentesco em comum; e III, estimacão e avaliacão em diferentes amostras da população de referência.

constatada na expressão  $r_{\text{gg}} = \sqrt{(Nh^2/n_Q)/[1+(Nh^2/n_Q)]}$ , em que N equivale ao número de indivíduos usados na população de estimação e  $n_Q$  é o número de locos (Daetwyler et al., 2008).

A eficiência da seleção genômica, em comparação à seleção fenotípica, foi inversamente proporcional à herdabilidade. A seleção genômica conseguiu ser superior à seleção fenotípica em até 48,5%, quando se

**Tabela 2.** Acurácia de predição dos valores aditivos, em diferentes cenários, e com intervalo de confiança "bootstrap" de 95%, com base no método RR-BLUP, para seis populações de milho-pipoca, com os efeitos dos marcadores estimados e utilizados na própria população.

Geração	Densidade de SNP	Acurácia do valor fenotípico [(h <sup>2</sup> ) <sup>0,5</sup> ]			
		0,548 (h <sup>2</sup> = 0,3)		0,837 (h <sup>2</sup> = 0,7)	
		Capacidade de expansão	Produção de grãos	Capacidade de expansão	Produção de grãos
População 0					
0	0,1	0,692 (0,641:0,748)	0,666 (0,591:0,725)	0,817 (0,791:0,842)	0,778 (0,750:0,806)
	1,0	0,663 (0,616:0,718)	0,643 (0,575:0,691)	0,761 (0,730:0,786)	0,733 (0,697:0,772)
	10	0,565 (0,509:0,619)	0,548 (0,477:0,614)	0,634 (0,590:0,679)	0,612 (0,566:0,651)
5	0,1	0,646 (0,574:0,700)	0,621 (0,553:0,682)	0,815 (0,784:0,842)	0,787 (0,754:0,820)
	1,0	0,604 (0,520:0,667)	0,578 (0,486:0,644)	0,758 (0,707:0,791)	0,724 (0,682:0,764)
	10	0,445 (0,291:0,533)	0,430 (0,349:0,496)	0,536 (0,458:0,589)	0,525 (0,464:0,571)
População 1					
0	0,1	0,664 (0,597:0,713)	0,645 (0,558:0,705)	0,801 (0,772:0,832)	0,768 (0,730:0,804)
	1,0	0,627 (0,548:0,678)	0,614 (0,528:0,679)	0,737 (0,695:0,772)	0,712 (0,639:0,760)
	10	0,504 (0,420:0,574)	0,501 (0,423:0,567)	0,579 (0,533:0,629)	0,577 (0,524:0,619)
5	0,1	0,640 (0,558:0,711)	0,607 (0,535:0,682)	0,812 (0,778:0,845)	0,781 (0,736:0,814)
	1,0	0,591 (0,503:0,669)	0,564 (0,482:0,644)	0,745 (0,681:0,787)	0,721 (0,669:0,761)
	10	0,415 (0,342:0,493)	0,399 (0,297:0,532)	0,508 (0,432:0,571)	0,498 (0,428:0,569)
População 2					
0	0,1	0,553 (0,450:0,614)	0,507 (0,418:0,567)	0,745 (0,662:0,788)	0,657 (0,552:0,710)
	1,0	0,490 (0,392:0,567)	0,467 (0,357:0,549)	0,640 (0,565:0,690)	0,583 (0,504:0,643)
	10	0,364 (0,267:0,445)	0,346 (0,267:0,428)	0,448 (0,375:0,513)	0,416 (0,344:0,481)
5	0,1	0,651 (0,600:0,699)	0,590 (0,503:0,680)	0,828 (0,794:0,858)	0,750 (0,616:0,808)
	1,0	0,601 (0,546:0,660)	0,538 (0,457:0,617)	0,761 (0,710:0,800)	0,686 (0,573:0,759)
	10	0,397 (0,321:0,462)	0,328 (0,249:0,411)	0,478 (0,388:0,534)	0,423 (0,318:0,509)
População 3					
0	0,1	0,765 (0,721:0,800)	0,760 (0,720:0,799)	0,858 (0,835:0,876)	0,844 (0,817:0,863)
	1,0	0,748 (0,706:0,778)	0,741 (0,700:0,788)	0,823 (0,793:0,849)	0,819 (0,788:0,836)
	10	0,659 (0,609:0,711)	0,656 (0,606:0,710)	0,714 (0,671:0,744)	0,712 (0,652:0,743)
5	0,1	0,684 (0,639:0,749)	0,686 (0,645:0,727)	0,839 (0,811:0,864)	0,819 (0,788:0,850)
	1,0	0,653 (0,613:0,723)	0,662 (0,636:0,706)	0,798 (0,770:0,832)	0,784 (0,752:0,817)
	10	0,515 (0,415:0,593)	0,511 (0,427:0,566)	0,602 (0,541:0,663)	0,590 (0,531:0,648)
População 4					
0	0,1	0,567 (0,514:0,629)	0,568 (0,497:0,629)	0,757 (0,726:0,793)	0,760 (0,706:0,787)
	1,0	0,483 (0,425:0,550)	0,492 (0,404:0,538)	0,631 (0,580:0,677)	0,636 (0,574:0,675)
	10	0,317 (0,229:0,385)	0,333 (0,234:0,407)	0,404 (0,335:0,442)	0,424 (0,364:0,495)
5	0,1	0,594 (0,535:0,650)	0,592 (0,525:0,653)	0,803 (0,769:0,836)	0,801 (0,768:0,827)
	1,0	0,512 (0,444:0,577)	0,515 (0,439:0,576)	0,697 (0,644:0,747)	0,709 (0,665:0,758)
	10	0,281 (0,182:0,381)	0,291 (0,202:0,409)	0,382 (0,306:0,444)	0,399 (0,345:0,466)
População 5					
0	0,1	0,814 (0,780:0,849)	0,776 (0,706:0,820)	0,896 (0,876:0,913)	0,855 (0,825:0,885)
	1,0	0,808 (0,772:0,841)	0,770 (0,709:0,819)	0,881 (0,858:0,898)	0,845 (0,816:0,874)
	10	0,742 (0,697:0,775)	0,716 (0,655:0,768)	0,798 (0,776:0,826)	0,775 (0,748:0,808)
5	0,1	0,792 (0,765:0,824)	0,732 (0,670:0,813)	0,889 (0,873:0,911)	0,852 (0,823:0,878)
	1,0	0,775 (0,745:0,816)	0,716 (0,638:0,803)	0,869 (0,846:0,893)	0,833 (0,799:0,862)
	10	0,624 (0,581:0,691)	0,585 (0,511:0,679)	0,699 (0,660:0,740)	0,679 (0,627:0,729)

considerou  $h^2 = 0,3$ . Para  $h^2 = 0,7$ , a sua superioridade não ficou tão evidente, uma vez que foi necessária alta densidade de marcadores para que se obtivessem acurácias superiores – no máximo 7% – às obtidas com a seleção fenotípica, no caso de populações com alto LD (população 5) ou alta variância aditiva (população 3).

Na comparação de populações com valores equivalentes de LD (populações 0, 1, 2 e 3), observou-se que o aumento da acurácia foi proporcional ao aumento da variância aditiva, e a população de maior variância aditiva (população 3) apresentou as maiores acurácias. Em trabalho de simulação com populações de bovinos, Habier et al. (2007) sugeriram, como populações de genotipagem, as mestiças, com maior variabilidade genética, ao invés de populações homogêneas ou de raça pura.

A redução da distância entre marcadores adjacentes de 1 para 0,1 cM não causou aumento relevante da acurácia de predição (8%) e, na maioria dos casos, os valores de acurácia estiveram contemplados em ambos os intervalos de confiança, independentemente da geração (Tabela 2). Entretanto, o aumento da distância de 0,1 para 10 cM resultou em uma perda média da acurácia de 39%, com variação de 10 a 87%, para CE, e de 35%, com variação de 8 a 79%, para a produção de grãos, na geração 0. Na geração 5, a perda foi de 64%, para CE e produção de grãos.

Para que se obtenha maior cobertura de marcadores no genoma, associada a menores custos de genotipagem, o uso de 1 SNP a cada cM parece ser a decisão mais acertada, para a maioria dos cenários aqui simulados. Contudo, ao considerar as populações de baixo LD e com menor variância aditiva, o uso de maiores densidades de SNPs seria necessário, para a obtenção de acurácia de predição satisfatória. No caso de populações biparentais, comuns no melhoramento vegetal, maiores densidades de marcadores serão necessárias, se o LD for baixo (Asoro et al., 2011).

Quando os efeitos dos SNPs foram estimados e usados em populações sem parentesco, com estimação na população de referência (geração 0) e uso para estimação dos GBVs nas demais populações e gerações, não se observou correlação positiva significativa entre os valores estimados e os verdadeiros (Tabela 3). A exceção foi quando a população de seleção e a de referência constituíam a mesma população, ou populações distanciadas apenas por cinco gerações de acasalamento ao acaso. No entanto, a acurácia de predição na geração 5 apresentou redução média de

23%, quando a  $h^2$  foi de 0,3, e de 30%, quando a  $h^2$  foi de 0,7, em comparação à acurácia obtida com genotipagem e fenotipagem do mesmo indivíduo.

Herdabilidade, densidade de SNPs, LD e variância aditiva não tiveram influência sobre a acurácia de predição, para os cenários em que a população de referência e a de seleção não eram derivadas da mesma amostra. Nesse caso, na maioria das situações, a acurácia foi próxima de zero. Portanto, os modelos preditivos não podem ser usados para populações de melhoramento que apresentem parâmetros genéticos similares à população de referência, e há necessidade de que grande parte dos alelos presentes na população de seleção esteja representada na população de referência. O cálculo dos GBVs dificulta a utilização da seleção genômica entre várias populações, uma vez que, na maioria dos casos, os efeitos de marcadores são diferentes entre elas (De Roos et al., 2009). Esses resultados estão de acordo com os obtidos por Toosi et al. (2010), que observaram que a acurácia é reduzida drasticamente quando os alelos específicos de uma amostra, ou de um subgrupo, não estão incluídos na população de referência.

Nakaya & Isobe (2012) argumentam que marcadores em LD significativo com os QTLs de interesse, em geral, não são mantidos em diferentes cruzamentos, o que dificulta o uso da GWS. Como exemplo, se efeitos de marcadores forem estimados a partir de híbridos simples de milho, os valores genético-genômicos estimados em populações heterogêneas não fornecem informações acuradas (Massman et al., 2013). Harris et al. (2009), ao utilizar duas raças de bovinos como populações de referência, verificaram que o uso de uma raça como população de referência e de outra raça para validação resulta em baixa acurácia (-0,10 a 0,3).

Quando os efeitos dos SNPs foram estimados e usados em diferentes amostras da população de referência, observou-se considerável perda de acurácia (Tabela 4). A magnitude dessa perda foi decorrente dos parâmetros contrastantes entre as duas populações.

Houve redução de pelo menos 8% na acurácia, ao se considerar a população de seleção na mesma geração da população de referência, e uma redução de, no mínimo, 22%, ao se considerar o distanciamento de 5 gerações de acasalamentos ao acaso. Esse resultado não é surpreendente, pois alguns marcadores podem estar em elevado LD com um QTL em uma amostra, mas não em outra, especialmente quando as populações

estão distanciadas por muitas gerações ou quando é utilizada menor densidade de marcadores (Andreescu et al., 2007; De Roos et al., 2008).

Verificamos que apesar de sempre haver redução da acurácia de predição, com a utilização de amostras

diferentes para estimação e seleção, a diferença de geração entre as populações apresentou maior relevância do que a herdabilidade e a densidade de marcadores. Um motivo para isso é que alguns marcadores não permanecem em fase de ligação com o

**Tabela 3.** Acurácia de predição dos valores aditivos, em diferentes cenários, e intervalo de confiança "bootstrap" de 95%, com base no método RR-BLUP para seis populações de milho-pipoca, com os efeitos dos marcadores estimados na população de referência (na geração 0) considerados nos cálculos.

Geração	Densidade de SNP	Acurácia do valor fenotípico [(h <sup>2</sup> ) <sup>0,5</sup> ]			
		0,548 (h <sup>2</sup> = 0,3)		0,837 (h <sup>2</sup> = 0,7)	
		Capacidade de expansão	Produção de grãos	Capacidade de expansão	Produção de grãos
População 0					
0	0,1	0,692 (0,641:0,748)	0,666 (0,591:0,725)	0,817 (0,791:0,842)	0,778 (0,750:0,806)
	1,0	0,663 (0,616:0,718)	0,643 (0,575:0,691)	0,761 (0,730:0,786)	0,733 (0,697:0,772)
	10	0,565 (0,509:0,619)	0,548 (0,477:0,614)	0,634 (0,590:0,679)	0,612 (0,566:0,651)
5	0,1	0,495 (0,383:0,555)	0,478 (0,408:0,548)	0,560 (0,502:0,610)	0,546 (0,486:0,606)
	1,0	0,451 (0,355:0,517)	0,436 (0,379:0,501)	0,507 (0,435:0,570)	0,494 (0,444:0,572)
	10	0,351 (0,225:0,407)	0,344 (0,261:0,423)	0,395 (0,296:0,459)	0,389 (0,328:0,444)
População 1					
0	0,1	0,040 (-0,155:0,193)	0,034 (-0,120:0,191)	0,031 (-0,110:0,204)	0,052 (-0,143:0,193)
	1,0	0,023 (-0,212:0,267)	-0,015 (-0,212:0,147)	0,030 (-0,140:0,198)	0,008 (-0,162:0,159)
	10	-0,067 (-0,217:0,096)	-0,080 (-0,237:0,064)	-0,080 (-0,185:0,007)	-0,080 (-0,215:0,002)
5	0,1	0,032 (-0,128:0,139)	0,002 (-0,104:0,119)	0,006 (-0,131:0,163)	0,018 (-0,101:0,110)
	1,0	-0,008 (-0,183:0,103)	-0,012 (-0,115:0,092)	-0,004 (-0,133:0,123)	-0,005 (-0,123:0,113)
	10	-0,058 (-0,182:0,018)	-0,045 (-0,188:0,073)	-0,049 (-0,174:0,044)	-0,053 (-0,170:0,055)
População 2					
0	0,1	0,059 (-0,056:0,162)	0,02 (-0,132:0,142)	0,041 (-0,081:0,193)	0,027 (-0,119:0,114)
	1,0	0,098 (-0,012:0,192)	0,063 (-0,098:0,209)	0,111 (-0,017:0,197)	0,098 (0,000:0,193)
	10	0,002 (-0,159:0,098)	0,003 (-0,175:0,215)	-0,006 (-0,112:0,083)	-0,003 (-0,113:0,127)
5	0,1	0,019 (-0,116:0,131)	-0,014 (-0,188:0,162)	0,007 (-0,121:0,120)	0,021 (-0,094:0,130)
	1,0	0,068 (-0,035:0,193)	0,058 (-0,100:0,179)	0,076 (0,012:0,136)	0,096 (-0,033:0,225)
	10	0,057 (-0,067:0,153)	0,038 (-0,048:0,140)	0,055 (-0,043:0,156)	0,042 (-0,065:0,114)
População 3					
0	0,1	0,045 (-0,158:0,217)	0,064 (-0,204:0,231)	0,090 (-0,189:0,286)	0,080 (-0,110:0,255)
	1,0	0,066 (-0,157:0,273)	0,016 (-0,279:0,299)	0,124 (-0,088:0,330)	0,004 (-0,182:0,213)
	10	-0,029 (-0,264:0,171)	-0,034 (-0,228:0,205)	0,017 (-0,237:0,215)	-0,010 (-0,308:0,133)
5	0,1	0,057 (-0,080:0,174)	0,054 (-0,073:0,198)	0,075 (-0,077:0,229)	0,063 (-0,075:0,235)
	1,0	0,039 (-0,147:0,192)	0,006 (-0,114:0,201)	0,057 (-0,088:0,211)	-0,019 (-0,148:0,135)
	10	0,009 (-0,165:0,132)	-0,011 (-0,146:0,151)	0,038 (-0,122:0,178)	-0,006 (-0,129:0,156)
População 4					
0	0,1	0,000 (-0,132:0,096)	-0,007 (-0,113:0,108)	0,000 (-0,120:0,130)	-0,001 (-0,063:0,086)
	1,0	-0,016 (-0,101:0,127)	-0,011 (-0,139:0,122)	-0,034 (-0,130:0,054)	-0,006 (-0,120:0,104)
	10	-0,015 (-0,146:0,112)	0,005 (-0,143:0,125)	-0,028 (-0,108:0,090)	0,011 (-0,100:0,101)
5	0,1	-0,009 (-0,14:0,096)	0,005 (-0,068:0,079)	0,002 (-0,101:0,116)	-0,001 (-0,073:0,097)
	1,0	-0,011 (-0,125:0,107)	-0,005 (-0,109:0,092)	-0,007 (-0,134:0,085)	0,000 (-0,094:0,098)
	10	-0,003 (-0,084:0,079)	0,024 (-0,078:0,115)	-0,005 (-0,125:0,133)	0,008 (-0,099:0,090)
População 5					
0	0,1	0,069 (-0,269:0,336)	0,013 (-0,330:0,353)	0,067 (-0,282:0,252)	0,024 (-0,309:0,339)
	1,0	0,042 (-0,254:0,293)	-0,014 (-0,276:0,267)	0,111 (-0,247:0,416)	0,025 (-0,271:0,227)
	10	-0,183 (-0,447:0,042)	-0,142 (-0,380:0,124)	-0,185 (-0,367:-0,016)	-0,196 (-0,385:0,064)
5	0,1	0,060 (-0,177:0,325)	0,044 (-0,180:0,255)	0,067 (-0,119:0,239)	0,052 (-0,196:0,229)
	1,0	-0,055 (-0,258:0,127)	-0,082 (-0,249:0,144)	0,008 (-0,234:0,260)	-0,046 (-0,251:0,124)
	10	-0,118 (-0,300:0,067)	-0,110 (-0,250:0,073)	-0,118 (-0,232:-0,005)	-0,138 (-0,299:-0,011)

QTL ao longo das gerações (Muir, 2007). Além disso, a diminuição das relações de parentesco, causada pelo distanciamento de populações por muitas gerações, tem sido reportada como um empecilho no uso da seleção genômica (Gianola et al., 2009; Habier et al., 2010) e pode ter maior efeito sobre a acurácia dos GBVs do que o valor “per se” de LD (Wientjes et al., 2013).

Para todos os cenários simulados, a estruturação da população em progênies S1 proporcionou os maiores valores de acurácia, seguida das estruturações em FIC e em FMI (Tabela 5). Os menores valores ocorreram na população não estruturada em progênies. Para a capacidade de expansão, na herdabilidade 0,3, a acurácia dos GBVs estimados em população estruturada em FS1 foi superior em, aproximadamente, 15, 30 e 50%, em comparação à acurácia obtida para FIC, FMI e sem parentesco, respectivamente. A estruturação dos indivíduos em progênies de maior relacionamento resultou em maior eficiência da seleção genômica, o que está de acordo com os resultados obtidos por Clark et al. (2012).

Os valores de acurácia estimados foram inversamente proporcionais ao tamanho efetivo das populações. No entanto, o aumento do  $N_e$  dentro de cada população não conduziu a reduções expressivas de acurácia, tendo apresentado, inclusive, maiores valores de acurácia com o aumento do  $N_e$ , em alguns cenários para a produção de grãos. Uma explicação para esse resultado é o fato de se ter utilizado alta densidade de marcadores (10 SNPs por cM). Segundo Grattapaglia & Resende (2011), o tamanho efetivo só teria maior influência, se

a densidade de marcadores utilizada fosse insuficiente para amostrar o genoma. Isso pode ser explicado pelas expressões  $r_{mq}^2 = [n_m / (n_m + M_e)]$  e  $M_e = 2N_eL$ , em que:  $N_e$  é o tamanho efetivo populacional e  $L$  é o tamanho do genoma da espécie (Goddard et al., 2011). Assim, a proporção da variação aditiva do caráter explicada pelos marcadores ( $r_{mq}^2$ ) é inversamente proporcional ao  $N_e$  e diretamente proporcional ao número  $n_m$  de marcadores. Outro aspecto importante é que, quando o tamanho efetivo populacional é pequeno, e os valores associados aos genótipos são grandes, pode ocorrer uma superestimação do valor aditivo (Nakaya & Isobe, 2012).

Em geral, menores tamanhos da população de referência são requeridos em populações de autopolinização, pois, populações com maior diversidade genética requerem tamanhos populacionais maiores, para que obtenha alta acurácia na estimação dos GBVs (Mujibi et al., 2011). Nesse contexto, é de se esperar que, para condições similares de seleção, maiores acurácias sejam obtidas para FS1, FIC, FMI e sem parentesco. No entanto, a extrapolação desses resultados para qualquer cenário de análise exige cautela, pois, o valor de LD e a magnitude da variância aditiva das populações aqui simuladas podem não refletir a magnitude desses parâmetros nas diferentes populações de melhoramento. De acordo com Wientjes et al. (2013), quanto menor for o tamanho efetivo da população de referência, maior é a importância das relações de parentesco para a eficiência da GWS, em relação à magnitude do LD.

**Tabela 4.** Acurácia de predição dos valores aditivos, em diferentes cenários, intervalo de confiança "bootstrap" de 95%, com base no método RR-BLUP, e efeitos dos marcadores em diferentes amostras da população de referência (0).

Densidade de SNP	Acurácia do valor fenotípico $[(h^2)^{0.5}]$			
	0,548 ( $h^2 = 0,3$ )		0,837 ( $h^2 = 0,7$ )	
	Capacidade de expansão	Produção de grãos	Capacidade de expansão	Produção de grãos
	Geração 0			
0,1	0,613 (0,573:0,652)	0,612 (0,570:0,666)	0,646 (0,603:0,686)	0,641 (0,597:0,699)
1,0	0,581 (0,544:0,628)	0,588 (0,537:0,652)	0,612 (0,571:0,646)	0,606 (0,559:0,664)
10	0,483 (0,438:0,548)	0,489 (0,427:0,560)	0,544 (0,483:0,609)	0,513 (0,467:0,580)
	Geração 5			
0,1	0,483 (0,424:0,526)	0,460 (0,381:0,512)	0,529 (0,455:0,583)	0,530 (0,468:0,580)
1,0	0,436 (0,373:0,503)	0,428 (0,373:0,467)	0,473 (0,414:0,528)	0,465 (0,406:0,506)
10	0,310 (0,236:0,407)	0,334 (0,292:0,405)	0,383 (0,305:0,463)	0,341 (0,277:0,426)

**Tabela 5.** Acurácia de predição dos valores aditivos, em diferentes estruturas populacionais e tamanhos efetivos ( $N_e$ ), com intervalo de confiança "bootstrap" de 95%, de acordo com o método RR-BLUP.

$N_e$	Acurácia do valor fenotípico [ $(h^2)^{0.5}$ ]			
	0,548 ( $h^2 = 0,3$ )		0,837 ( $h^2 = 0,7$ )	
	Capacidade de expansão	Produção de grãos	Capacidade de expansão	Produção de grãos
	População estruturada em progênies endogâmicas (FS <sub>i</sub> )			
10	0,820 (0,783:0,853)	0,765 (0,730:0,805)	0,883 (0,856:0,913)	0,840 (0,820:0,869)
20	0,778 (0,763:0,816)	0,757 (0,745:0,794)	0,874 (0,849:0,894)	0,857 (0,838:0,882)
	População estruturada em progênies de irmãos completos (FIC)			
20	0,695 (0,619:0,773)	0,659 (0,589:0,694)	0,789 (0,761:0,857)	0,760 (0,694:0,813)
38	0,664 (0,620:0,739)	0,708 (0,649:0,757)	0,781 (0,727:0,830)	0,791 (0,746:0,834)
	População estruturada em progênies de meios-irmãos (FMI)			
38	0,570 (0,477:0,660)	0,596 (0,521:0,662)	0,696 (0,633:0,766)	0,703 (0,683:0,733)
72	0,559 (0,501:0,608)	0,528 (0,493:0,569)	0,682 (0,644:0,702)	0,672 (0,642:0,696)
	População estruturada - sem parentesco entre os indivíduos (OP)			
1.000	0,405 (0,355:0,450)	0,472 (0,431:0,520)	0,556 (0,511:0,607)	0,667 (0,639:0,686)

## Conclusões

1. Para populações de reduzido desequilíbrio de ligação (LD) e menor variância aditiva, o uso de maiores densidades de marcadores é necessário, para que se obtenha acurácia de predição satisfatória.

2. É necessário que o candidato à seleção seja relacionado à população da qual os efeitos de marcadores foram estimados, em caso contrário, qualquer semelhança dos parâmetros genéticos entre candidato e população não se traduzirá em alta acurácia na seleção genômica.

3. Há redução de pelo menos 8% da acurácia, quando se consideram diferentes amostras como população de seleção e de estimação.

4. A estruturação dos indivíduos em progênies de maior relacionamento genético resulta em maior eficiência da seleção genômica.

5. O menor tamanho efetivo da população tem impacto positivo sobre os valores de acurácia.

## Agradecimentos

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig), pelo apoio financeiro.

## Referências

- ALBRECHT, T.; AUINGER, H.J.; WIMMER, V.; OGUTU, J.O.; KNAACK, C.; OUZUNOVA, M.; PIEPHO, H.P.; SCHÖN, C.C. Genome-based prediction of maize hybrid performance across genetic groups, testers, locations, and years. **Theoretical and Applied Genetics**, v.127, p.1375-1386, 2014. DOI: 10.1007/s00122-014-2305-z.
- ANDREESCU, C.; AVENDANO, S.; BROWN, S.R.; HASSEN, A.; LAMONT, S.J.; DEKKERS, J.C.M. Linkage disequilibrium in related breeding lines of chickens. **Genetics**, v.177, p.2161-2169, 2007. DOI: 10.1534/genetics.107.082206.
- ASORO, F.G.; NEWELL, M.A.; BEAVIS, W.D.; SCOTT, M.P.; JANNINK, J.-L. Accuracy and training population design for genomic selection on quantitative traits in elite North American oats. **Plant Genome**, v.4, p.132-144, 2011. DOI: 10.3835/plantgenome2011.02.0007.
- CLARK, S.A.; HICKEY, J.M.; DAETWYLER, H.D.; WERF, H.J. van der. The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. **Genetics Selection Evolution**, v.44, p.4-12, 2012. DOI: 10.1186/1297-9686-44-4.
- CROSSA, J.; BEYENE, Y.; KASSA, S.; PÉREZ, P.; HICKEY, J.M.; CHEN, C.; DE LOS CAMPOS, G.; BURGUEÑO, J.; WINDHAUSEN, V.S.; BUCKLER, E.; JANNINK, J.L.; CRUZ, M.A.; BABU, R. Genomic prediction in maize breeding populations with genotyping-by-sequencing. **G3**, v.3, p.1903-1926, 2013. DOI: 10.1534/g3.113.008227.
- CROSSA, J.; PÉREZ, P.; HICKEY, J.; BURGUEÑO, J.; ORNELLA, L.; CERÓN-ROJAS, J.; ZHANG, X.; DREISIGACKER, S.; BABU, R.; LI, Y.; BONNETT, D.;

- MATHEWS, K. Genomic prediction in CIMMYT maize and wheat breeding programs. **Heredity**, v.112, p.48-60, 2014. DOI: 10.1038/hdy.2013.16.
- DAETWYLER, H.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**, v.193, p.347-365, 2013. DOI: 10.1534/genetics.112.147983.
- DAETWYLER, H.D.; VILLANUEVA, B.; WOOLLIAMS, J.A. Accuracy of predicting the genetic risk of disease using a genome-wide approach. **PLoS One**, v.3, p.e3395, 2008. DOI: 10.1371/journal.pone.0003395.
- DE LOS CAMPOS, G.; HICKEY, J.M.; PONG-WONG, R.; DAETWYLER, H.D.; CALUS, M.P.L. Whole-genome regression and prediction methods applied to plant and animal breeding. **Genetics**, v.193, p.327-345, 2013. DOI: 10.1534/genetics.112.143313.
- DE ROOS, A.P.W.; HAYES, B.J.; GODDARD, M.E. Reliability of genomic predictions across multiple populations. **Genetics**, v.183, p.1545-1553, 2009. DOI: 10.1534/genetics.109.104935.
- DE ROOS, A.P.W.; HAYES, B.J.; SPELMAN, R.J.; GODDARD, M.E. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. **Genetics**, v.179, p.1503-1512, 2008. DOI: 10.1534/genetics.107.084301.
- ENDELMAN, J.B. Ridge regression and other kernels for genomic selection with R Package rrBLUP. **The Plant Genome**, v.4, p.250-255, 2011. DOI: 10.3835/plantgenome2011.08.0024.
- GIANOLA, D.; DE LOS CAMPOS, G.; HILL, W.G.; MANFREDI, E.; FERNANDO, R. Additive genetic variability and the Bayesian alphabet. **Genetics**, v.183, p.347-363, 2009. DOI: 10.1534/genetics.109.103952.
- GODDARD, M.E.; HAYES, B.J. Genomic selection. **Journal of Animal Breeding and Genetics**, v.124, p.323-330, 2007. DOI: 10.1111/j.1439-0388.2007.00702.x.
- GODDARD, M.E.; HAYES, B.J.; MEUWISSEN, T.H.E. Using the genomic relationship matrix to predict the accuracy of genomic selection. **Journal of Animal Breeding and Genetics**, v.128, p.409-421, 2011. DOI: 10.1111/j.1439-0388.2011.00964.x.
- GRATTAPAGLIA, D.; RESENDE, M.D.V. Genomic selection in forest tree breeding. **Tree Genetics & Genomes**, v.7, p.241-255, 2011. DOI: 10.1007/s11295-010-0328-4.
- HABIER, D.; FERNANDO, R.L.; DEKKERS, J.C. The impact of genetic relationship information on genome-assisted breeding values. **Genetics**, v.177, p.2389-2397, 2007. DOI: 10.1534/genetics.107.081190.
- 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. **Genetics Selection Evolution**, v.42, p.5-16, 2010. DOI: 10.1186/1297-9686-42-5.
- HARRIS, B.L.; JOHNSON, D.L.; SPELMAN, R.J. Genomic selection in New Zealand and the implications for national genetic evaluation. In: ICAR BIENNIAL SESSION, 36., 2008, Niagara. **Identification, breeding, production, health and recording of farm animals: proceedings**. Rome: ICAR, 2009. p.325-330. (ICAR Technical Series, 13). Editors: J. D. Sattler.
- HESLOT, N.; JANNINK, J.-L.; SORRELLS, M.E. Perspectives for genomic selection applications and research in plants. **Crop Science**, v.55, p.1-12, 2015. DOI: 10.2135/cropsci2014.03.0249.
- JONAS, E.; DE KONING, D.-J. de. Does genomic selection have a future in plant breeding? **Trends in Biotechnology**, v.31, p.497-504, 2013. DOI: 10.1016/j.tibtech.2013.06.003.
- MASSMAN, J.M.; GORDILLO, A.; LORENZANA, R.E.; BERNARDO, R. Genomewide predictions from maize single-cross data. **Theoretical and Applied Genetics**, v.126, p.13-22, 2013. DOI: 10.1007/s00122-012-1955-y.
- MEUWISSEN, T.H.; HAYES, B.J.; GODDARD, M.E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v.157, p.1819-1829, 2001.
- MUIR, W.M. Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. **Journal of Animal Breeding and Genetics**, v.124, p.342-355, 2007. DOI: 10.1111/j.1439-0388.2007.00700.x.
- MUJIBI, F.D.N.; NKURUMAH, J.D.; DURUNNA, O.N.; STOTHARD, P.; MAH, J.; WANG, Z.; BASARAB, J.; PLASTOW, G.; CREWS JR., D.H.; MOORE, S.S. Accuracy of genomic breeding values for residual feed intake in crossbred beef cattle. **Journal of Animal Science**, v.89, p.3353-3361, 2011. DOI: 10.2527/jas.2010-3361.
- NAKAYA, A.; ISOBE, S.N. Will genomic selection be a practical method for plant breeding? **Annals of Botany**, v.110, p.1303-1316, 2012. DOI: 10.1093/aob/mcs109.
- POLAND, J. Breeding-assisted genomics. **Current Opinion in Plant Biology**, v.24, p.119-124, 2015. DOI: 10.1016/j.pbi.2015.02.009.
- R CORE TEAM. **R: A language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2013. Disponível em: <<http://www.R-project.org>>. Acesso em: 16 jul. 2013.
- RESENDE JR, M.F.; MUNOZ, P.; ACOSTA, J.J.; PETER, G.F.; DAVIS, J.M.; GRATTAPAGLIA, D.; RESENDE, M.D.V.; KIRST, M. Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. **New Phytologist**, v.193, p.617-624, 2012. DOI: 10.1111/j.1469-8137.2011.03895.x.
- RESENDE, M.D.V. de; BARBOSA, M.H.P. **Melhoramento genético de plantas de propagação assexuada**. Colombo: Embrapa Florestas, 2005. 130p.
- RESENDE, M.D.V.; RESENDE JR., M.F.R.; SANSALONI, C.P.; PETROLI, C.D.; MISSIAGGIA, A.A.; AGUIAR, A.M.; ABAD, J.M.; TAKAHASHI, E.K.; ROSADO, A.M.; FARIA, D.A.; PAPPAS JR., 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. DOI: 10.1111/j.1469-8137.2011.04038.x.

- TOOSI, A.; FERNANDO, R.L.; DEKKERS, J.C. Genomic selection in admixed and crossbred populations. **Journal of Animal Science**, v.88, p.32-46, 2010. DOI: 10.2527/jas.2009-1975.
- VIANA, J.M.S.; PIEPHO, H.-P.; SILVA, F.F. e. Quantitative genetics theory for genomic selection and efficiency of breeding value prediction in open-pollinated populations. **Scientia Agricola**, v.73, p.243-251, 2016. DOI: 10.1590/0103-9016-2014-0383.
- VIANA, J.M.S.; VALENTE, M.S.F.; SILVA, F.F. e; MUNDIM, G.B.; PAES, G.P. Efficacy of population structure analysis with breeding populations and inbred lines. **Genetica**, v.141, p.389-399, 2013. DOI: 10.1007/s10709-013-9738-1.
- WIJNTJES, Y.C.J.; VEERKAMP, R.F.; CALUS, M.P.L. The effect of linkage disequilibrium and family relationships on the reliability of genomic prediction. **Genetics**, v.193, p.621-631, 2013. DOI: 10.1534/genetics.112.146290.
- XU, S.; ZHU, D.; ZHANG, Q. Predicting hybrid performance in rice using genomic best linear unbiased prediction. **Proceedings of the National Academy of Sciences of the United States of America**, v.111, p.12456-12461, 2014. DOI: 10.1073/pnas.1413750111.
- YABE, S.; OHSAWA, R.; IWATA, H. Potential of genomic selection for mass selection breeding in annual allogamous crops. **Crop Science**, v.53, p.95-105, 2013. DOI: 10.2135/cropsci2012.03.0167.
- ZHAO, Y.; METTE, M.F.; REIF, J.C. Genomic selection in hybrid breeding. **Plant Breeding**, v. 134, p.1-10, 2015. DOI: 10.1111/pbr.12231.
- ZHAO, Y.; ZENG, J.; FERNANDO, R.; REIF, J.C. Genomic prediction of hybrid wheat performance. **Crop Science**, v.53, p.802-810, 2013. DOI: 10.2135/cropsci2012.08.0463.

---

Recebido em 3 de julho de 2015 e aprovado em 26 de agosto de 2016