

PREDIÇÃO SIMULTÂNEA DOS EFEITOS DE MARCADORES MOLECULARES E SELEÇÃO GENÔMICA AMPLA EM CAJUEIRO¹

JOSÉ JAIME VASCONCELOS CAVALCANTI², MARCOS DEON VILELA DE RESENDE³, FRANCISCO HERBETH COSTA DOS SANTOS⁴, CÁSSIA RENATA PINHEIRO⁵

RESUMO - A seleção genômica ampla (*genome wide selection* - GWS) foi proposta como uma forma de aumentar a eficiência e acelerar o melhoramento genético, enfatizando a predição simultânea dos efeitos genéticos de grande número de marcadores genéticos de DNA dispersos em todo o genoma de um organismo, de forma a capturar os efeitos de todos os locos e explicar a variação genética de um caráter quantitativo. Objetivou-se com o presente trabalho aplicar o princípio da GWS no melhoramento do cajueiro, estimando simultaneamente os efeitos de 238 marcadores avaliados em 74 indivíduos de uma família de irmãos completos, visando a explicar grande porcentagem da variação genotípica total do caráter peso da amêndoa e a aumentar a eficiência do melhoramento do cajueiro. Verificou-se que a capacidade preditiva e a acurácia são praticamente maximizadas na análise com 70 marcadores de maiores efeitos. O aumento do número de marcadores não aumenta linearmente a acurácia da GWS pelo método RR-BLUP. Os 70 marcadores de maiores efeitos capturam 74% da variação genotípica total e propiciam alta acurácia seletiva (86%) da seleção para o peso de amêndoas, enquanto os cinco marcadores de maiores efeitos capturam apenas 19% da variação genotípica total e propiciam acurácia seletiva de apenas 44%. Assim, a seleção assistida (MAS), baseada em poucos (cinco) marcadores de efeitos significativos, propicia eficiência muito inferior à GWS. Os valores genéticos genômicos preditos na população de validação cruzada aproximam-se bem dos valores fenotípicos observados, com correlação de 0,79. A estimação simultânea dos efeitos dos marcadores, segundo o conceito da GWS, é uma alternativa interessante, visando a aumentar a eficiência do melhoramento do cajueiro.

Termos para Indexação: seleção assistida, melhoramento genético, regressão aleatória, BLUP.

SIMULTANEOUS PREDICTION OF THE EFFECTS OF MOLECULAR MARKERS AND GENOME WIDE SELECTION IN CASHEW

ABSTRACT - The genome wide selection (GWS) was proposed as a way to increase efficiency and accelerate the genetic improvement, emphasizing the simultaneous prediction of the genetic effects of a large number of DNA genetic markers scattered throughout the genome of an organism, in order to capture the effects of all loci and explain the genetic variation of a quantitative character. The objective of this study was to apply the principle of GWS in the cashew breeding, simultaneously estimating the effects of 238 markers assessed in 74 individuals from a full-sib family to explain the large percentage of total genotypic variation of the character kernel weight and increasing the efficiency of cashew breeding. It was found that the predictive power and accuracy are nearly maximized in the analysis with 70 markers to the greatest effect. The increase in the number of markers does not increase linearly the accuracy of the method GWS RR-BLUP. The 70 markers with the greatest effects capture 74% of the total genotypic variation and provide a selective high accuracy (86%) of the selection for kernel weight, while the five markers with the greatest effects capture only 19% of the total genotypic variation and provide a selective accuracy of only 44%. Thus assisted selection (MAS) based on a few (five) markers provide significant effects efficiency much lower than GWS. The predicted genomic genetic values in the population of cross-validation had a good approximation of the observed phenotype values, with a correlation of 0.79. The simultaneous estimation of the effects of the markers according to the concept of GWS is an interesting alternative to enhance the efficiency of cashew breeding.

Index terms: assisted selection, breeding, random regression, BLUP.

¹(Trabalho 091-11). Recebido em: 23-03-2011. Aceito para publicação em: 28-08-2012.

²Pesquisador Embrapa Algodão, CP 174, CEP 58428-095, Campina Grande-PB. E-mail: jaime@cnpa.embrapa.br

³Pesquisador Embrapa Florestas, CP 319, CEP 83411-000, Colombo-PR. E-mail: deon@cnpf.embrapa.br

⁴Doutorando em Agronomia, Universidade Federal do Ceará, Departamento de Fitotecnia, CP 6035, 60.451-970 Fortaleza-CE. E-mail: herbeth.santos@gmail.com

⁵Bolsista nível superior, Embrapa Agroindústria Tropical, CP 3761, CEP 60.511-110 Fortaleza-CE. E-mail: cassiapinheiro@gmail.com

INTRODUÇÃO

O cajueiro é uma das principais espécies frutíferas cultivadas no Brasil, e seu melhoramento constitui-se em importante estratégia no aumento da resistência aos fatores bióticos e abióticos, produtividade e qualidade de seus produtos, contribuindo para o pleno desenvolvimento do sistema produtivo e a sustentabilidade do seu agronegócio. Por ser uma espécie perene, o melhoramento demanda métodos e estratégias eficientes, visando a acelerar o processo seletivo dos caracteres de interesse econômico. Dentre esses, destaca-se o peso da amêndoa, devido à preferência no mercado por tamanhos maiores.

A seleção genômica ampla (*genome wide selection* - GWS) ou seleção genômica (*genome selection* - GS) foi proposta por Meuwissen et al. (2001) como uma forma de aumentar a eficiência e acelerar o melhoramento genético. A GWS enfatiza a predição simultânea (sem o uso de testes de significância para marcas individuais) dos efeitos genéticos de grande número de marcadores genéticos de DNA (SNP, DArT, Microsatélites) dispersos em todo o genoma de um organismo, de forma a capturar os efeitos de todos os locos de um caráter quantitativo (QTL, de pequenos e grandes efeitos) e explicar toda a variação genética. A condição fundamental para isso é que haja desequilíbrio de ligação entre alelos dos marcadores e alelos dos genes que controlam o caráter. A predição é realizada com base em dados genotípicos e fenotípicos de indivíduos pertencentes a uma amostra da população de seleção.

Esses efeitos sobre fenótipos de caracteres quantitativos são somados e usados na predição de valores genéticos de indivíduos apenas genotipados, candidatos à seleção em programas de melhoramento genético. A predição e a seleção podem ser realizadas em fases muito juvenis de plantas e animais, acelerando assim o processo de melhoramento genético. Adicionalmente, a própria predição tende a ser mais acurada por considerar o real parentesco genético dos indivíduos em avaliação, em detrimento do parentesco médio esperado matematicamente (RESENDE, 2007). A GWS propicia uma forma de seleção precoce direta (SPD), pois atua precocemente sobre genes expressos na idade adulta. Ao contrário, a seleção precoce tradicional é indireta, pois atua (via avaliação fenotípica) sobre genes ativados na idade precoce, esperando que esses informem parcialmente sobre genes expressos na idade adulta. Assim, a SPD propiciada pela GWS é especialmente importante para o melhoramento de espécies perenes (RESENDE et al., 2008; GRATTAPAGLIA ; RESENDE, 2011) como o cajueiro.

A GWS, quando aplicada dentro de famílias, equivale à seleção assistida por marcadores (*marker assisted selection* – MAS) tradicional descrita por Lande e Thompson (1990), porém considerando simultaneamente um grande número de marcadores, sem teste prévio de significância estatística de cada um deles. A análise tradicional de QTL para a MAS necessita de pesquisas em backgrounds específicos, baseados em populações de cruzamentos controlados, e, portanto, tem validade apenas para cada família. Assim, uma vantagem adicional da GWS em relação ao estudo e mapeamento de QTLs é que ela pode ser aplicada em toda a população.

Em cajueiro, a construção de mapas genéticos, a detecção de QTL e a MAS têm sido aplicados com sucesso (CAVALCANTI ; WILKINSON, 2007; SANTOS et al., 2010). Santos et al. (2010) detectaram QTLs para sete caracteres em cajueiro, sendo que o número de QTLs detectados por caráter variou de dois a sete, e explicaram entre 3,15% e 21,33% da variação fenotípica do caráter. Esses QTLs marcados serão usados para a MAS, entretanto explicarão pequena fração da variabilidade genética total de cada caráter.

Com base no exposto, objetivou-se com o presente trabalho aplicar o princípio da GWS no cajueiro, estimando simultaneamente os efeitos de 238 marcadores avaliados em 74 indivíduos de uma família de irmãos completos, visando a explicar grande porcentagem da variação genotípica total do caráter peso da amêndoa e a aumentar a eficiência do melhoramento do cajueiro.

MATERIAL E MÉTODOS

Foram avaliados 238 marcadores em 74 indivíduos de uma família de irmãos completos oriundos do cruzamento entre o clone de cajueiro-anão precoce CCP 1001 e uma matriz de cajueiro comum CP 96, plantados em março de 2001. Os marcadores empregados foram do tipo AFLP (224), microsatélites (12) e ISSR (2). Todos esses tipos de marcadores podem ser usados na GWS, desde que a matriz Z seja construída adequadamente.

Nas análises, foi empregado o método da regressão aleatória (RR-BLUP/GWS), que usa preditores do tipo BLUP, com os efeitos de marcadores não ajustados como variáveis classificatórias, e, sim, como variáveis explicativas ou explanatórias. Os estimadores associados à regressão aleatória promovem *shrinkage* ditado por uma função da quantidade λ . Por isso é possível a estimação precisa de um número de parâmetros maior que o próprio número de dados (tamanho da amostra). O parâmetro de regressão é

dado por $\lambda = \sigma_e^2 / \sigma_{gi}^2 = \sigma_e^2 / (\sigma_g^2 / n)$, em que σ_{gi}^2 é a variância genética associada ao loco ou segmento i , e σ_g^2 e σ_e^2 são a variância genética do caráter e a variância residual, respectivamente. A quantidade n é desconhecida *a priori*, mas pode ser inferida conforme descrito adiante.

A predição via RR-BLUP/GWS é descrita a seguir com base em Resende (2007; 2008). O seguinte modelo linear misto geral foi ajustado para estimar os efeitos dos marcadores:

$y = Xb + Zm + e$, em que: y é o vetor de observações fenotípicas; b é o vetor de efeitos fixos; m é o vetor dos efeitos aleatórios de marcadores, e e refere-se ao vetor de resíduos aleatórios. X e Z são as matrizes de incidência para b e m .

A matriz de incidência Z contém os valores 0 e 1 para ausência ou presença do marcador.

As equações de modelo misto genômicas para a predição de m via o método RR-BLUP/GWS equivalem a:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + I \frac{\sigma_e^2}{(\sigma_g^2/n)} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{m} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

O valor genético genômico (VGG) global do indivíduo j é dado por $VGG = \hat{y}_j = \sum Z_j \hat{m}_i$. As equações de predição apresentadas acima assumem *a priori* que todos os locos explicam iguais quantidades da variação genética. Assim, a variação genética explicada por cada loco é dada por σ_g^2/n , em que σ_g^2 é a variação genética total, e n é o número de locos (quando cada loco está perfeitamente marcado por uma só marca). A variação genotípica σ_g^2 foi estimada por REML sobre os dados fenotípicos. Outros métodos, como o BayesA e BayesB, permitem modelar uma variação genética específica para cada loco e podem ser vantajosos em presença de genes de efeitos maiores (MEUWISSEN et al., 2001). Na predição RR-BLUP necessita-se da quantidade $\lambda = \sigma_e^2 / \sigma_{gi}^2 = \sigma_e^2 / (\sigma_g^2 / n)$, em que: n é o número de locos controlando o caráter (assumindo que cada loco está perfeitamente marcado), o qual é desconhecido *a priori*, mas é dado por $[\sum_i p_i(1-p_i)]$. Alternativamente, pode ser expresso como:

$\lambda = \sigma_e^2 / \sigma_{gi}^2 = \sigma_e^2 / (\sigma_g^2 / n) = (1-h^2)/(h^2/n)$, e, portanto, $\lambda = (1-h^2)/(h^2/n) = (1-h^2)/\{(h^2/n) \sum_i p_i(1-p_i)\}$.

Expresso de outra forma, sendo:

$$\lambda = n(1-h^2)/h^2 = \{ \sum_i p_i(1-p_i) \} (1-h^2)/h^2$$

Desta forma, de posse de h^2 e das frequências alélicas nos locos marcadores (p_i), obtém-se para uso nas equações de modelo misto. É importante notar que h^2 se refere à herdabilidade ajustada.

Populações de estimação, validação e seleção.

Na prática da seleção genômica ampla, três populações podem ser definidas: população de estimação, validação e seleção. Essas podem exercer duas funções ao mesmo tempo (uma só população usada para estimação e validação). A Figura 1 ilustra a estratégia empregada no presente trabalho.

População de estimação: é também denominada população de descoberta, de treinamento ou de referência. Esse conjunto de dados contempla um grande número de marcadores avaliados em um número moderado de indivíduos, os quais devem ter seus fenótipos avaliados para os vários caracteres de interesse. Equações de predição associam a cada marcador seu efeito (predito por RR-BLUP) no caráter de interesse. Nessa população, são descobertos, via marcadores, os marcadores que explicam os locos que controlam os caracteres, bem como são estimados seus efeitos.

População de validação: as equações de predição de VGG são testadas para verificar suas acurácias em uma amostra independente. Para computar essa acurácia, VGG são preditos (usando os efeitos estimados na população de estimação) e submetidos à análise de correlação com os valores fenotípicos observados. Como a amostra de validação não foi envolvida na predição dos efeitos dos marcadores, os erros dos VGGs e dos valores fenotípicos são independentes, e a correlação entre esses valores é predominantemente de natureza genética e equivale à capacidade preditiva (r_{yy}) da GWS em estimar os fenótipos, sendo dada pela própria acurácia seletiva (r_{gg}), multiplicada pela raiz quadrada da herdabilidade individual (h), ou seja, $r_{yy} = r_{gg}h$. Assim, para estimação da própria acurácia, deve-se obter $r_{yy} = r_{gg}/h$.

População de Seleção: esse conjunto de dados contempla apenas os marcadores avaliados nos candidatos à seleção. Essa população não necessita ter seus fenótipos avaliados. As equações de predição derivadas na população de descoberta são então usadas na predição dos VGGs ou fenótipos futuros dos candidatos à seleção.

População de validação e Jackknife: a metodologia *Jackknife* baseia-se na divisão do conjunto de N dados amostrais em g grupos de tamanho igual a k , de forma que $N = gk$. Em geral, k é tomado como 1, mas pode ser tão grande quanto $N/2$. O estimador do parâmetro de interesse baseia-se em amostras de tamanho $(g-1)k$, em que o i -ésimo grupo de tamanho k foi removido. As validações foram rea-

lizadas com $k = 1$.

RESULTADOS E DISCUSSÃO

Inicialmente, foram computados os valores genéticos genômicos (VGG) usando todos os marcadores. Em seguida, os marcadores foram ordenados por maiores módulos dos efeitos estimados dos marcadores e foram criados arquivos com subconjuntos dos marcadores com maiores módulos dos efeitos estimados (1 a 238). Todos esses arquivos foram analisados para o cômputo dos VGGs nas populações de estimação e de validação. Os resultados são apresentados nas Tabelas 1 e 2.

Verifica-se que a herdabilidade que maximiza a capacidade preditiva na validação cruzada é 0,85 (Tabela 1). O ponto de máximo da capacidade preditiva reflete a coerência interna e intrínseca dos dados em informar sobre o fenótipo. Os pontos de máximo da capacidade preditiva podem então ser usados de forma reversa para informar sobre a herdabilidade capturada pelos marcadores, de forma alternativa ou confirmativa do método REML. Assim, a h^2 e o parâmetro λ são obtidos por sintonia fina ou ajuste na própria validação cruzada via o modelo com maior capacidade preditiva ou maior determinação para prever o fenótipo. Este valor coincide com aquele estimado para o cajueiro por REML sobre os dados fenotípicos (CAVALCANTI et al., 2007). Assim, o valor 0,85 foi empregado na predição.

Na Tabela 2, são apresentados valores da capacidade preditiva e acurácia da GWS na população de validação, associada aos arquivos com diferentes números de marcadores, em ordem daqueles de maior efeito. Verifica-se que a capacidade preditiva e a acurácia são praticamente maximizadas na análise com 70 marcadores de maiores efeitos. As capacidades preditivas diminuem com o aumento do número de marcadores. O aumento do número de marcadores não aumenta linearmente a acurácia da GWS pelo método RR-BLUP, concordando com os resultados de Fernando et al. (2007). Assim, em torno de 70 locos marcadores são suficientes para maximizar a acurácia na população de validação.

O aumento ou diminuição da acurácia da GWS via RR-BLUP é um compromisso ou balanço entre acréscimo da quantidade N de informação útil via uso de maior número n de locos marcadores e diminuição do tamanho de amostra efetivo para estimar o efeito de cada loco, ou seja, menor número de indivíduos por loco a ser estimado (menor N/n).

Verifica-se, na Tabela 2, que os 70 marcadores de maiores efeitos capturam 74% da variação

genotípica total e propiciam alta acurácia seletiva (86%). O marcador de maior efeito captura apenas 8% da variação genotípica total e propicia acurácia seletiva de apenas 28%. Os cinco marcadores de maiores efeitos capturam apenas 19% da variação genotípica total e propiciam acurácia seletiva de apenas 44%. Assim, a seleção assistida, baseada em poucos marcadores de efeitos significativos, conforme adotado para o cajueiro por Cavalcanti e Wilkinson (2007) e Santos et al. (2010), propicia eficiência muito inferior à GWS.

Neste sentido, mesmo dentro de famílias, a estimação simultânea dos efeitos dos marcadores, segundo o conceito da seleção genômica ampla, é uma alternativa interessante visando a aumentar a eficiência do melhoramento do cajueiro.

O número reduzido de marcadores explicando grande parte da variação genética ou da acurácia máxima possível é muito interessante do ponto de vista prático. Nesse caso, arranjos de DNA com baixa densidade de marcadores previamente selecionados poderiam ser usados nas populações de seleção do cajueiro.

Os efeitos dos 70 marcadores principais e os valores genéticos genômicos preditos dos 74 indivíduos são apresentados na Tabela 3. Verifica-se que os efeitos dos marcadores seguem aproximadamente distribuição normal, obedecendo ao modelo genético poligênico (muitos genes de pequenos efeitos), com efeitos variando entre -0,85 e 0,93 gramas. Os valores genéticos genômicos preditos na população de validação cruzada aproximaram-se bem dos valores fenotípicos observados, com correlação de 0,79.

Os catálogos com os efeitos genéticos estimados dos 70 marcadores, bem como com os valores genéticos genômicos preditos dos 74 indivíduos são prontamente úteis ao melhoramento do cajueiro e podem contribuir significativamente ao aumento da produtividade da cultura.

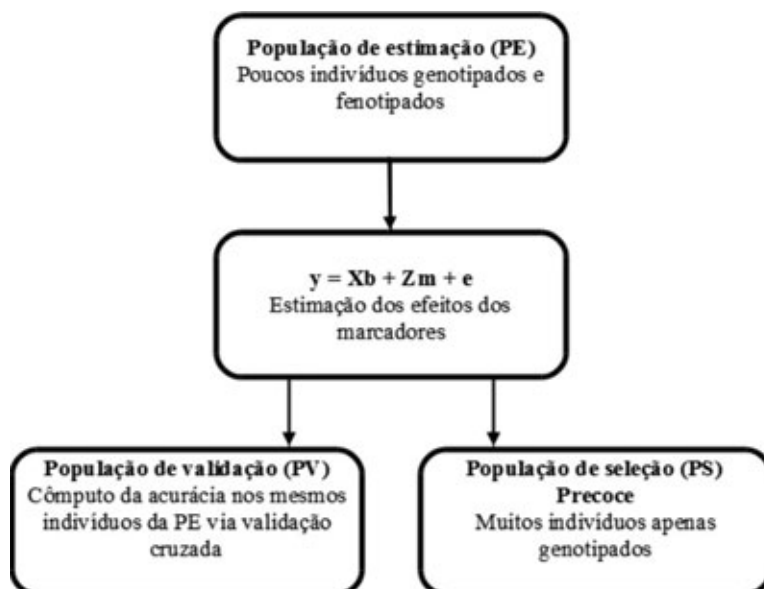


FIGURA 1- Esquema de aplicação da seleção genômica ampla em um programa de melhoramento genético.

TABELA 1- Resposta da capacidade preditiva (rgf da população de validação) em função da herdabilidade usada na predição, para a base de dados com 238 marcadores.

Herdabilidade	Capacidade Preditiva
0,50	0,01
0,60	0,03
0,70	0,04
0,80	0,04
0,85	0,05

TABELA 2- Capacidade preditiva (correlação entre fenótipo e valor genético genômico) e acurácia da seleção genômica ampla na população de validação.

Número de Marcas - M	Número de Locos - n	Capacidade Preditiva	Acurácia Seletiva	Determinação ou Confiabilidade da Seleção
238	58,00	0,05	0,05	0,00
200	48,00	0,28	0,30	0,09
150	36,00	0,56	0,61	0,37
100	24,00	0,74	0,80	0,64
85	21,00	0,75	0,82	0,67
80	20,00	0,75	0,82	0,67
75	18,00	0,77	0,84	0,71
70	17,00	0,79	0,86	0,74
67	16,50	0,76	0,83	0,69
63	15,00	0,77	0,83	0,69
50	12,00	0,74	0,80	0,64
40	10,00	0,75	0,81	0,65
30	7,00	0,68	0,74	0,55
20	5,00	0,66	0,71	0,51
15	3,70	0,63	0,68	0,47
10	2,50	0,53	0,58	0,33
5	1,22	0,40	0,44	0,19
4	0,98	0,37	0,40	0,16
3	0,74	0,32	0,34	0,12
2	0,49	0,30	0,33	0,11
1	0,22	0,26	0,28	0,08

TABELA 3- Efeitos dos 70 marcadores principais e valores genéticos genômicos preditos e validados dos indivíduos de cajueiros.

Marcador	Efeito	Indivíduo	Valor Genômico	Valor Fenotípico
8	0,9314	11	26,00	25,60
16	0,9152	129	25,04	24,10
32	0,7628	59	23,18	21,60
4	0,7581	15	23,47	23,80
3	0,7174	107	22,50	22,40
5	0,7051	127	22,03	24,75
7	0,7005	156	22,24	19,50
6	0,6886	109	22,34	21,95
1	0,6673	158	22,16	22,90
59	0,5759	88	22,05	25,10
9	0,5718	13	21,82	23,20
55	0,5718	12	21,72	22,80
64	0,5640	103	21,84	24,15
24	0,5453	55	21,79	25,00
13	0,5442	105	21,36	23,20
37	0,5364	194	21,35	19,80
39	0,5086	18	21,40	20,20
14	0,4967	57	21,37	22,40
70	0,4740	122	21,36	20,65
10	0,4656	54	20,75	19,40
38	0,4396	193	21,18	19,25
49	0,4216	104	20,85	21,00
36	0,3985	89	20,90	19,75
54	0,3887	155	20,53	22,50
11	0,3827	14	20,28	20,20
25	0,3709	51	20,23	20,40
44	0,3696	196	20,23	17,65
15	0,3486	108	19,38	24,15
63	0,3349	121	19,95	18,10
52	0,2877	102	19,88	21,35
58	0,2577	151	19,54	17,05
56	0,0907	58	19,53	19,00
46	0,0232	199	19,55	18,90
43	-0,0057	153	19,33	18,30
67	-0,0695	197	19,13	17,25
34	-0,1154	123	19,33	22,15
66	-0,1574	124	19,38	16,75
45	-0,1902	130	18,68	20,30
42	-0,2536	83	18,83	18,35
65	-0,2623	126	18,61	15,10
41	-0,2872	154	18,88	20,60
47	-0,3292	87	18,32	16,15
57	-0,3517	110	18,54	16,75
69	-0,3544	120	18,43	20,05
27	-0,3742	101	18,39	18,30
33	-0,3814	106	18,10	19,75
35	-0,3992	60	18,11	17,70
48	-0,4027	20	18,12	18,20
60	-0,4091	198	18,05	16,50
53	-0,4490	90	18,31	16,75
62	-0,4598	17	17,81	16,40
30	-0,4666	152	17,78	19,40
17	-0,4698	16	17,78	20,60
61	-0,4717	159	17,63	17,90
50	-0,4839	82	17,62	16,75
21	-0,5094	192	17,36	14,50
31	-0,5232	157	17,51	19,40
19	-0,5453	160	17,07	11,50
23	-0,5699	125	17,40	16,60
51	-0,6519	191	16,91	18,50
26	-0,6687	118	16,38	17,75

continuação...

continuação...

Marcador	Efeito	Indivíduo	Valor Genômico	Valor Fenotípico
29	-0,6726	19	16,33	18,40
40	-0,6744	128	16,05	16,95
68	-0,7040	195	15,19	15,50
12	-0,7071	53	14,55	11,00
28	-0,7086	52	15,05	16,40
18	-0,7436	86	14,68	16,95
20	-0,8226	81	14,00	13,75
22	-0,8231	-	-	-
2	-0,8537	-	-	-

CONCLUSÕES

1-A capacidade preditiva e a acurácia (86%) da seleção para o peso de amêndoas são praticamente maximizadas na análise com 70 marcadores de maiores efeitos, os quais capturam 74 % da variação genotípica total do caráter.

2-A seleção assistida, baseada em poucos (cinco) marcadores de efeitos significativos, propicia eficiência muito inferior à seleção genômica ampla.

3-Os valores genéticos genômicos preditos na população de validação cruzada apresentam alta correlação (79%) com os valores fenotípicos observados e, portanto, a seleção genômica ampla pode ser usada com eficiência no melhoramento do cajueiro.

AGRADECIMENTOS

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, pelo apoio financeiro.

REFERÊNCIAS

- CAVALCANTI, J.J.V.; RESENDE, M.D.V.; CRISÓSTOMO, J.R.; BARROS, L.M.; PAIVA, J.R. Genetic control of quantitative traits and hybrid breeding strategies for cashew improvement. **Crop Breeding and Applied Biotechnology**, Londrina, v.7, p.186-195, 2007.
- CAVALCANTI, J.J.V.; WILKINSON, M.J. The first genetic maps of cashew (*Anacardium occidentale* L.). **Euphytica**, Wageningen, v. 157, p. 131-143, 2007.
- FERNANDO, R.L.; HABIER, D.; STRICKER, C.; DEKKERS, J.C.M.; TOTTIR, L.R. Genomic selection. **Acta Agriculturae Scandinavica, Section A - Animal Science**, Copenhagen, v.57, p.192-195, 2007.
- GRATTAPAGLIA, D.; RESENDE, M.D.V. Genomic selection in forest tree breeding. **Tree Genetics and Genomes**, Davis, v.7, p.241-255, 2011.
- LANDE, R.; THOMPSON, R. Efficiency of marker-assisted selection in the improvement of quantitative traits. **Genetics**, Austin, v.124, p.743-756, 1990.
- MEUWISSEN, T.H.E.; HAYES, B.J.; GODDARD, M.E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, Austin, v.157, p.1.819-1.829, 2001.
- RESENDE, M.D.V. **Genômica quantitativa e seleção no melhoramento de plantas perenes e animais**. Colombo: Embrapa Florestas, 2008. 330 p.
- RESENDE, M.D.V. **Matemática e estatística na análise de experimentos e no melhoramento genético**. Colombo: Embrapa Florestas, 2007. 561 p.
- 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**, Colombo, v.56, p.63-78, 2008.
- SANTOS, F.H.C; CAVALCANTI, J.J.V.; SILVA, F.P. Detection of quantitative trait loci for physical traits of cashew apple. **Crop Breeding and Applied Biotechnology**, Londrina, v.10, p.101-109, 2010.