

Comparação da Seleção Genômica e Fenotípica Utilizando Índices de Seleção em Milho¹

Gustavo Vitti Môro¹, Mateus Figueiredo Santos² e Cláudio Lopes de Souza Júnior³

¹ Universidade Estadual Paulista – FCAV/UNESP, Jaboticabal, São Paulo, gvmoro@fcav.unesp.br

² Embrapa Hortaliças, Gama, Distrito Federal, mateus@cnph.embrapa.br

³ Universidade de São Paulo – ESALQ/USP, Piracicaba, São Paulo, clsouza@esalq.usp.br

RESUMO - A seleção de genótipos superiores é realizada considerando índices de seleção. Marcadores moleculares aumentam a eficiência da seleção de genótipos. O objetivo deste estudo foi comparar a seleção genômica com a seleção fenotípica, quando esta é realizada utilizando índices de seleção. Foram utilizadas 256 progênies F_{2:3} e um mapa com 177 marcadores microssatélites. As progênies foram avaliadas em experimentos, obtendo-se médias fenotípicas (MF) e médias preditas (MP) para diversos caracteres, com os quais foram estimados diferentes índices de seleção. Obteve-se correlações entre os índices para as MF e para as MP e, após a seleção das melhores progênies, constatou-se o número de progênies superiores coincidentes entre os métodos. As correlações entre as MF e MP foram elevadas, sugerindo que as MP são eficientes para a seleção de genótipos superiores. As estimativas das correlações entre os índices de seleção obtidos para as MF e as MP variaram de medianas até elevadas. A ocorrência de genótipos superiores coincidentes entre a seleção realizada considerando os índices para as MF e MP mostrou que, de modo geral, a coincidência não é elevada. Os resultados indicam que a seleção genômica pode ser implementada nos programas de melhoramento, aumentando a eficiência do processo seletivo.

Palavras-chave: Eficiência, Seleção Genômica, BLUP, Índice de Seleção, Ideótipo.

Introdução

Uma nova abordagem para a utilização de marcadores moleculares na seleção, conhecida como Genome Wide Selection ou Seleção Genômica (SG), foi proposta por Meuwissen et al. (2001). Na SG, consideram-se informações de todos os marcadores do genoma e não apenas daqueles associados ao caráter por meio da análise de QTL, eliminando-se a etapa de construção de mapa genético e mapeamento de QTL.

Nos programas de melhoramento a seleção de genótipos superiores é realizada com base em índices de seleção e não para cada caráter separadamente, e essa estratégia tem sido muito eficiente para a obtenção de genótipos superiores, pois permite obter ganhos para diversos caracteres de importância agrônômica/econômica, simultaneamente. Os índices de seleção funcionam como um caráter adicional, resultante de uma combinação de diversas características sobre as quais se deseja obter respostas à seleção (Santos et al., 2007), independente da existência ou não de correlação entre estes (Smith, 1936; Vilarinho et al., 2003). Embora a resposta para cada caráter possa ser superior àquela baseada no índice, esta

¹ Apoio financeiro: Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP e Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq.

última resulta em genótipos com melhores padrões agronômicos, pois estes alcançam níveis adequados para todos os caracteres considerados no índice, em menor tempo do que os genótipos obtidos a partir da seleção sequencial para cada caráter.

A seleção genômica realizada com base em um índice de seleção pode aumentar a eficiência do processo de seleção em relação à seleção fenotípica, permitindo selecionar genótipos com padrões agronômicos mais próximos aos genótipos ideais (ideótipos), em menor espaço de tempo. O objetivo deste trabalho foi verificar a possibilidade de utilizar a seleção genômica para obtenção de genótipos superiores em milho, quando se pratica seleção considerando índices de seleção e não para cada caráter separadamente, e definir quais índices de seleção oferecem melhor acurácia para a seleção genômica em relação à fenotípica.

Material e Métodos

Foram utilizadas 256 progênies $F_{2:3}$ obtidas do cruzamento entre as linhagens L14-04B e L08-05F e um mapa genético com 177 marcadores microsatélites, com 10 grupos de ligação, comprimento de 2.055,8 cM e intervalo médio entre marcadores de 12,31 cM. Nos anos agrícolas de 2002/2003 até 2006/2007 as progênies $F_{2:3}$ foram avaliadas em 13 ambientes no município de Piracicaba-SP, em experimentos no delineamento látice simples 16 x 16. Com os dados obtidos nesses experimentos foram realizadas análises estatístico-genéticas considerando os seguintes caracteres: produção de grãos (PG), prolificidade (PROL), acamamento e quebramento (ACQ), altura da planta (AP), altura da espiga (AE), posição relativa da espiga (PRE) e intervalo entre florescimentos (IF). Foram obtidas as médias fenotípicas das progênies $F_{2:3}$, que referem-se aos valores fenotípicos médios de ambientes e repetições. Como foram utilizados 13 ambientes nas avaliações, com duas repetições por ambiente, as médias foram obtidas de vinte e seis observações fenotípicas das progênies e, portanto, pode-se inferir que estas médias apresentam excelente precisão experimental, apresentando valores próximos aos valores genéticos das progênies.

Para a seleção genômica, os valores genéticos de cada marcador molecular foram estimados utilizando-se o método BLUP (*Best linear unbiased predictor*) (Meuwissen et al., 2001; Bernardo; Yu, 2007; Meuwissen, 2007), de acordo o modelo: $\mathbf{y} = \bullet \mathbf{1} + \mathbf{X}\mathbf{g} + \mathbf{e}$, em que \mathbf{y} é o vetor de dimensão $N_{F_2} \times 1$ com as médias fenotípicas das progênie $F_{2:3}$; \bullet corresponde à média fenotípica geral das progênies $F_{2:3}$; $\mathbf{1}$ é o vetor de uns de dimensão $N_{F_2} \times 1$ que relaciona a média ao vetor \mathbf{y} ; \mathbf{X} é a matriz de dimensões $N_{F_2} \times N_M$, com elementos 1 se a planta F_2 é homozigota para o marcador originado da linhagem L14-04B, e -1 se a planta F_2 é

homozigota para o marcador originado da linhagem L08-05F e 0 se for heterozigoto; \mathbf{g} é o vetor $N_M \times 1$ dos BLUPs dos valores genéticos dos marcadores, considerado aleatório; \mathbf{e} é o vetor $N_M \times 1$ dos resíduos. A variância genética e a variância do erro foram obtidas das análises de variância conjuntas realizadas, assumindo-se a variância de cada marcador como sendo igual a V_G/N_M (Meuwissen et al. 2001). Foi obtido um vetor (\mathbf{g}) contendo as estimativas dos valores genéticos de cada marcador molecular do genoma (\hat{g}_{is}), utilizando-se uma equação de modelo-misto, considerando μ como fixo e o efeito dos g_{is} como aleatório. Com a genotipagem das progênies $F_{2:3}$ e as estimativas dos valores genéticos dos marcadores moleculares foram obtidas as médias preditas das progênies com base em todos os marcadores do genoma, por: $\hat{\mathbf{y}} = \hat{\mu} \mathbf{1} + \mathbf{X}\hat{\mathbf{g}}$, em que $\hat{\mathbf{y}}$ é o vetor com as estimativas das médias preditas das progênies $F_{2:3}$; $\hat{\mu}$ corresponde à estimativa da média fenotípica geral das progênies $F_{2:3}$; $\mathbf{1}$ é o vetor de uns de dimensão $N_{F_2} \times 1$ que relaciona a média ao vetor \mathbf{y} ; \mathbf{X} é a mesma matriz descrita anteriormente e $\hat{\mathbf{g}}$ é o vetor $N_M \times 1$ das estimativas dos BLUPs dos valores genéticos dos marcadores.

Com as médias fenotípicas e preditas com base nos efeitos de todos os marcadores do genoma para cada um dos caracteres analisados, foram construídos diferentes índices de seleção para as progênies, envolvendo os caracteres considerados, com exceção da altura da espiga. A expressão geral dos índices de seleção (I), que foram obtidos para cada genótipo e para cada metodologia de obtenção das médias, é: $\mathbf{I} = \sum_i^i \mathbf{b}_i \mathbf{X}_i$, em que: \mathbf{b}_i é o peso para o caráter i e \mathbf{X}_i é o valor observado ou predito para o caráter i . Considerando cada estratégia de obtenção das médias foram calculados, para cada progênie, os seguintes índices de seleção: Smith-Hazel (Smith, 1936; Hazel, 1943), Williams (Williams, 1962), Pesek e Baker (Pesek; Baker, 1969), Dickerson, Allaire e Henderson (Dickerson et al., 1954, Allaire; Henderson, 1966), Smith (Smith et al., 1981), Mulamba e Mock (Mulamba; Mock, 1978) e, também foi aplicada uma seleção “subjetiva”, onde os genótipos foram ranqueados pela produtividade, selecionando-se aqueles com produtividade acima da média e com valores adequados para os demais caracteres, com maior rigor para o acamamento e quebramento de plantas.

As estimativas para cada índice de seleção considerando as médias preditas das progênies com base em todos os marcadores foram comparadas com aquelas obtidas considerando as médias fenotípicas observadas dessas progênies, a partir da obtenção das estimativas de correlação e dos coeficientes de determinação entre essas estimativas.

Também, as progênes superiores foram selecionadas considerando os índices de seleção e a seleção subjetiva, para os dois procedimentos de obtenção de médias, com intensidades de seleção de 10% e 20%, selecionando-se os 26 e os 52 melhores genótipos, respectivamente. Para os genótipos selecionados foram contados o número de coincidências de progênes $F_{2:3}$ selecionadas entre o método envolvendo marcadores moleculares e o fenotípico.

Resultados e Discussão

Para todos os caracteres avaliados, as fontes de variação ambiente, progênie e a interação entre esses dois fatores foram altamente significativas indicando, respectivamente, que os ambientes utilizados eram contrastantes, que existe variabilidade genética entre as progênes $F_{2:3}$ e que estas apresentaram performance diferencial de acordo com o ambiente. Os coeficientes de variação experimental obtidos na análise conjunta estão de acordo com os reportados na literatura para esses mesmos caracteres, indicando que os experimentos foram realizados com boa precisão. Todos os componentes da variância foram diferentes de zero e as estimativas foram obtidas com boa precisão, o que foi verificado pelos intervalos de confiança dessas estimativas (dados não apresentados).

Os coeficientes de herdabilidade ao nível de médias foram elevados para todos os caracteres, variando de 0,80 para acamamento e quebramento de plantas até 0,94 para produção de grãos. Este ponto é de fundamental importância, pois os valores fenotípicos observados nas progênes são muito próximos dos seus valores genéticos, o que afeta diretamente a predição de médias utilizando marcadores moleculares, tornando essas estimativas muito precisas (Bernardo; Yu, 2007) (Tabela 1). Com relação às médias fenotípicas, verifica-se que para todos os caracteres ocorreu uma grande amplitude de variação, confirmando a existência de elevada variabilidade nas progênes $F_{2:3}$ utilizadas neste trabalho. Os intervalos de variação para as médias preditas considerando informações de todos os marcadores moleculares diferiram daqueles considerando as médias fenotípicas, indicando que a amplitude dos valores fenotípicos difere da amplitude dos valores genotípicos das progênes. Assim como era esperado, devido à grande quantidade de observações utilizadas para estimar as médias fenotípicas (26 observações), os BLUPs dos marcadores e, conseqüentemente, as médias preditas das progênes foram estimadas com boa precisão, o que é confirmado pela correlação e coeficientes de determinação obtidos entre as médias fenotípicas e preditas (Bernardo; Yu, 2007; Meuwissen, 2007). Para todos os caracteres as correlações foram elevadas, variando de 0,72 para acamamento e quebramento de plantas até 0,84 para produção de grãos, indicando que as médias preditas se aproximam das médias

fenotípicas para todos os caracteres, podendo ser utilizada para a seleção de genótipos superiores (Meuwissen et al., 2001). Os coeficientes de determinação, que indicam quanto os marcadores conseguem explicar do fenótipo também foram elevados para todos os caracteres, variando de 0,52 até 0,71 para acamamento e quebramento de plantas e para produção, respectivamente. Esses valores elevados indicam, portanto, que os marcadores moleculares, por meio da seleção genômica, são eficientes em prever o valor fenotípico de uma progênie utilizando informações de seu genótipo, explicando grande parte da variação fenotípica observada (Bernardo; Yu, 2007; Meuwissen, 2007).

As estimativas dos coeficientes de correlação entre cada índice de seleção considerando as médias fenotípicas e as médias preditas com base nos marcadores moleculares variaram de medianas considerando os índices de seleção propostos por Smith-Hazel (IS 1), Williams (IS 2) e Smith (IS 4), até elevada para os índices propostos por Pesek e Baker (IS 3), Mulamba e Mock (IS 5) e Dickerson - Allaire e Henderson (IS 6) (Tabela 2). Com base nos valores dessas estimativas, pode-se inferir que o índice de seleção que mais se aproxima quando se utilizam médias fenotípicas e médias preditas com base em marcadores moleculares é o índice proposto por Mulamba e Mock, no qual os pesos dos caracteres envolvem o ranqueamento e a soma dos ranques dos genótipos para os caracteres considerados no índice.

Com relação ao número e a porcentagem de genótipos superiores coincidentes entre a seleção realizada considerando informações de marcadores moleculares e àquela realizada com base em informações fenotípicas para os diferentes índices de seleção utilizados verificou-se que, para a intensidade de seleção de 10% (26 genótipos selecionados) ocorreram 46,15% de genótipos superiores coincidentes para os índices de seleção de Smith-Hazel (IS 1), de Williams (IS 2), de Smith (IS 4) e de Dickerson - Allaire e Henderson (IS 6). Para o índice de Pesek e Baker (IS 3) e para a seleção subjetiva (SJ) houve 50% de coincidência de genótipos superiores selecionados e, para o índice proposto por Mulamba e Mock (IS 5), ocorreu a maior coincidência entre genótipos superiores selecionados considerando informações fenotípicas e de marcadores para a seleção, com 53,85% de coincidência. Considerando a intensidade de seleção de 20%, as menores coincidências foram obtidas para os índices de Smith (48,08%), de Smith-Hazel (51,92%) e de Williams (53,85%). Para o índice de Mulamba e Mock e para a seleção subjetiva foi obtida uma coincidência de 55,77%, para o índice de Pesek e Baker a coincidência foi de 59,62% e, para o índice proposto por Dickerson - Allaire e Henderson houve 63,46% de coincidência sendo, portanto, a maior coincidência de genótipos superiores selecionados para esta intensidade de seleção. Esses

resultados sugerem que, dependendo da intensidade de seleção utilizada, o índice de seleção com maior coincidência de genótipos superiores considerando informações fenotípicas e de marcadores moleculares pode variar, mas de modo geral a coincidência não é elevada.

Conclusões

A predição de médias com base em informações de todos os marcadores do genoma é eficiente, sugerindo que a seleção genômica pode ser implementada nos programas de melhoramento. A obtenção dos diferentes índices de seleção para essas duas estratégias de obtenção de médias confirma a possibilidade de utilização da seleção genômica no melhoramento de plantas, já que as correlações entre os índices considerando as médias fenotípicas e preditas variaram de medianas até elevadas. Pode-se sugerir que o índice de seleção mais adequado para ser utilizado conjuntamente com informações de marcadores moleculares é aquele proposto por Mulamba e Mock.

Um exemplo para aplicação da SG nos programas de melhoramento seria a realização de uma avaliação inicial para obtenção das informações fenotípicas, a partir das quais seriam obtidas as predições dos valores genéticos dos marcadores moleculares do genoma. Posteriormente, realiza-se o primeiro ciclo de seleção utilizando um índice de seleção com informações fenotípicas e, em seguida, a seleção é realizada utilizando-se um índice de seleção com informações dos marcadores moleculares, por um ou dois ciclos. Na época ideal de cultivo realiza-se novamente uma avaliação fenotípica, obtendo-se um índice e praticando-se a seleção com base nessas médias e, paralelamente, estimam-se os BLUPs dos marcadores novamente. Com os valores dos BLUPs obtêm-se um índice de seleção e realiza-se novamente um ou dois ciclos de seleção considerando essas informações. Continuando esse processo, aumenta-se a eficiência do programa pois, com a utilização de marcadores moleculares, a seleção pode ser realizada na fase de plântula e fora do ambiente ideal de cultivo, o que possibilita a realização de mais de um ciclo de seleção por ano.

Literatura Citada

- ALLAIRE, F.R., HENDERSON, C.R. Selection practiced among dairy cows: II. Total production over a sequence of lactations. **J. Dairy Sci.**, v.49, p. 1435-1440, 1966.
- BERNARDO, R; YU, J. Prospects for genome wide selection for quantitative traits in maize. . **Crop Science**, Madison, v. 47, p.1082-1090, 2007.
- DICKERSON, G.E., BLUNN, C.T., CHAPMAN, A.G., KOTTMAN, R.M., KRIDDER, J.L., WARWICK, E.J., WHATLEY JR., J.A., BAKER, M.L., WINTERS, L.M. Evaluation of developing inbred lines of swine. **Res. Bull.** 551, Missouri Agric. Exp. Stn., Columbia, 1954.
- HAZEL, L.N. The genetic basis for constructing selection indexes. **Genetics**, Austin, v.28, p.476-490, 1943.

- 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. 1819-1829, 2001.
- MEUWISSEN, T. H. E. Genomic selection: marker assisted selection on genome-wide scale. **Journal of Animal Breeding and Genetics**, v. 124, p. 321-322, 2007.
- MULAMBA, N.N., J.J. MOCK. Improvement of yield potential of the Eto Blanco maize (*Zea mays* L.) population by breeding for plant traits. **Egypt. J. Genet. Cytol.** v.7, p. 40–51, 1978.
- PESEK, J., BAKER, R.J. Desired improvement in relation to selection indices. **Can. J. Plant Sci.** v.49, p. 803-804, 1969.
- SANTOS, F.S.; AMARAL JÚNIOR, A.T.; FREITAS JÚNIOR, S.P.; RANGEL, R.M.; PEREIRA, M.G. Predição de ganhos genéticos por índice de seleção na população de Milho-Pipoca UNB-2U sob Seleção Recorrente. **Brangantia**, Campinas, v.66, n.3, p.389-396, 2007.
- SMITH, H.F. A discriminant function for plant selection. **Annals of Eugenics**, London, v.7, p.240-250, 1936.
- SMITH, O.S., A.R. HALLAUER, and W.A. RUSSELL. Use of index selection in recurrent selection programs in maize. **Euphytica** v.30, p. 611–618, 1981.
- VILARINHO, A.A.; VIANA, J.M.S.; CÂMARA, T.M.M.; SANTOS, J.F. Seleção de progênies endogâmicas S₁ e S₂ em um programa de melhoramento intrapopulacional de milho pipoca. **Acta Scientiarum**, Maringá, v.24, n.5, p.1419-1425, 2003.
- WILLIAMS, J.S. The evaluation of a selection index. **Biometrics**, North Carolina, v.18, p.375-393, 1962.

Tabela 1. Herdabilidade ao nível de médias ($h_{\bar{X}}^2$) e seus intervalos de confiança (IC), médias fenotípicas (\bar{X}_F), médias preditas considerando informações dos marcadores (\bar{X}_{SG}), intervalos de variação (IV), correlações entre as médias fenotípicas e as preditas (r) e coeficientes de determinação (R^2) considerando os caracteres avaliados.

Caráter ^a	PG (t/ha)	ACQ (%)	PROL (esp/pl)	AP (m)	AE (m)	PRE	IF (dias)
$h_{\bar{X}}^2$	0,94	0,80	0,91	0,92	0,91	0,93	0,86
IC	[0,93;0,95]	[0,76;0,83]	[0,89;0,93]	[0,91;0,94]	[0,89;0,93]	[0,89;0,92]	[0,84;0,89]
\bar{X}_F	4,04	3,15	0,88	1,84	0,97	0,53	0,97
IV	(1,78;9,43)	(1,41;6,84)	(0,53;1,32)	(1,60;2,17)	(0,79;1,17)	(0,47;0,58)	(-1,03;3,70)
\bar{X}_{SG}	4,04	3,15	0,88	1,84	0,97	0,53	0,97
IV	(1,86;6,75)	(2,20;4,50)	(0,64;1,13)	(1,68;2,02)	(0,86;1,09)	(0,48;0,57)	(-0,35;2,68)
r	0,84	0,72	0,82	0,78	0,79	0,83	0,82
R^2	0,71	0,52	0,67	0,61	0,62	0,69	0,67

^a - PG (produção de grãos), ACQ (acamamento e quebramento de plantas), PROL (prolificidade), AP (altura da planta), AE (altura da espiga), PRE (posição relativa da espiga), IF (intervalo entre florescimentos).

Tabela 2. Correlações entre os índices de seleção e a seleção subjetiva considerando as médias fenotípicas e as médias preditas com base nos marcadores do genoma e coincidência de genótipos superiores selecionados, com 10% (26 genótipos selecionados) e 20% (52 genótipos selecionados) de intensidade de seleção.

Parâmetro		Índice de Seleção ^a						
		IS 1	IS 2	IS 3	IS 4	IS 5	IS 6	SJ
	r	0,46	0,49	0,70	0,47	0,79	0,71	-
Intensidade 10%	N ° coincidentes	12	12	13	12	14	12	13
	% coincidentes	46,15	46,15	50	46,15	53,85	46,15	50,00
Intensidade 20%	N ° coincidentes	27	28	31	25	29	33	29
	% coincidentes	51,92	53,85	59,62	48,08	55,77	63,46	55,77

^a - IS 1 refere-se ao índice de seleção de Smith-Hazel (Smith, 1936; Hazel, 1943), IS 2 ao de Williams (Williams, 1962), IS 3 ao de Pesek e Baker (Pesek; Baker, 1969), IS 4 ao de Smith (Smith et al., 1981), IS 5 ao de Mulamba e Mock (Mulamba; Mock, 1978), IS 6 ao índice proposto por Dickerson – Allaire e Henderson (Dickerson et al., 1954; Allaire; Henderson, 1966) e SJ à seleção genômica.