

# Seleção genômica ampla e novos métodos de melhoramento do milho<sup>1</sup>

Roberto Fritsche-Neto<sup>2</sup>, Marcos Deon Vilela Resende<sup>3</sup>, Glauco Vieira Miranda<sup>4</sup>, Júlio César DoVale<sup>5</sup>

## RESUMO

Os objetivos deste trabalho foram verificar a acurácia do método da Seleção Genômica Ampla (GWS) no melhoramento de milho nas condições de estresse nutricional e propor novos métodos de melhoramento baseados em GWS. Foram estimados os dois componentes da eficiência no uso de nitrogênio e de fósforo (eficiência de absorção e de utilização) em 41 combinações híbridas, em dois experimentos, sob baixa e alta disponibilidades de N e P. Para a genotipagem da população de estimação, foram utilizados 80 marcadores microssatélites. As estimativas dos parâmetros genéticos foram obtidas via REML/BLUP, e a predição dos valores genéticos genômicos, via regressão aleatória (Random Regression - RR) aplicada à seleção genômica ampla (RR-BLUP/GWS). Para os caracteres em que a GWS apresentou altos valores de acurácia, essa foi comparada com os métodos de Seleção Recorrente Intra e Interpopulacional. Com o uso da GWS houve aumento significativo na acurácia seletiva e nos ganhos genéticos por unidade de tempo.

**Palavras-chave:** seleção assistida por marcadores, seleção recorrente, estresses abióticos.

## ABSTRACT

### Genome wide selection and new methods of maize breeding

The objectives of this work were to verify the accuracy of the Genome Wide Selection method (GWS) in the maize breeding for nutritional stress conditions and propose new breeding methods based on GWS. The efficiency of two components of use of nitrogen and phosphorus (absorption and utilization) was estimated in 41 single-cross hybrids assessed in two experiments under low and high availability of N and P. Eighty microsatellite markers were used to genotype the base population. The estimates of genetic parameters were obtained by REML/BLUP and the predictions of genetic genomic values were obtained by random regression (RR), applied to genome wide selection (RR-BLUP/GWS). GWS showed high values of accuracy for the traits; it was comparable to the methods of Recurrent Intra and Interpopulational Selection. It was concluded that there is a significant increase in selective accuracy and in the genetic gains per unit of time with use of GWS.

**Key words:** marker assisted selection, recurrent selection, abiotic stresses.

Recebido para publicação em 12/01/2012 e aprovado em 31/08/2012.

<sup>1</sup> Parte da tese de doutorado do primeiro autor, apresentada ao Programa de Pós-Graduação em Genética e Melhoramento, da Universidade Federal de Viçosa, para obtenção do título de doutor.

<sup>2</sup>Engenheiro-Agrônomo, Doutor. Departamento de Fitotecnia, Universidade Federal de Viçosa, *Campus* Viçosa, Avenida Peter Henry Rolfs, s/n, 36570-000, Viçosa, Minas Gerais, Brasil. roberto.neto@ufv.br (autor para correspondência).

<sup>3</sup>Engenheiro-Agrônomo, Doutor. Embrapa Florestas, Estrada da Ribeira, Km 111, Caixa Postal 319, 83411-000, Colombo, Paraná, Brasil. deon@cnf.embrapa.br

<sup>4</sup>Engenheiro-Agrônomo, Doutor. Departamento de Fitotecnia, Universidade Federal de Viçosa, *Campus* Viçosa, Avenida Peter Henry Rolfs, s/n, 36570-000, Viçosa, Minas Gerais, Brasil. glaucovmiranda@ufv.br

<sup>5</sup>Engenheiro-Agrônomo, Doutor. Departamento de Fitotecnia, Universidade Federal de Viçosa, *Campus* Viçosa, Avenida Peter Henry Rolfs, s/n, 36570-000, Viçosa, Minas Gerais, Brasil. juliocvale@gmail.com

## INTRODUÇÃO

A Seleção Genômica Ampla (Genome Wide Selection – GWS) consiste na predição simultânea (sem o uso de testes de significância para marcas individuais) dos efeitos genéticos de grande número de marcadores genéticos dispersos em todo o genoma de um organismo, de forma a capturar os efeitos de todos os locos, de pequenos e grandes efeitos, e explicar grande parte da variação genética de um caráter quantitativo (Meuwissen *et al.*, 2001). Após a obtenção desses marcadores, seus efeitos são estimados, baseados em dados fenotípicos de uma população conhecida como de estimação. Uma vez estimados os efeitos, eles são testados em uma população de validação e, então, selecionados os marcadores que explicam grande parte da variância genética do caráter em estudo para que sua informação seja efetivamente incorporada à etapa de seleção do programa de melhoramento (Resende *et al.*, 2010).

A avaliação e a seleção baseadas no fenótipo para eficiência no uso de nitrogênio (EUN) ou de fósforo (EUP) utilizam geralmente métodos destrutivos da planta, o que, dependendo da estrutura genética da população, não permite o uso daquelas identificadas como superiores em hibridações. Consequentemente, há redução no potencial dos ganhos a serem obtidos com a seleção. Além disso, alguns métodos convencionais de melhoramento, como o da seleção recorrente, são muito trabalhosos e apresentam baixos ganhos com a seleção por unidade de tempo, tornando-se, muitas vezes, inviáveis economicamente.

Esses fatos podem ser amenizados com a utilização da GWS, que permite identificar precocemente os genótipos superiores, sem utilizar métodos destrutivos, aumentando os ganhos com a seleção e diminuindo o intervalo entre gerações. Por meio desse método, a predição e a seleção podem ser realizadas em fases muito juvenis das plantas, 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. Assim, a GWS propicia uma forma de seleção precoce direta, pois atua precocemente sobre os genes expressos na idade adulta, ao contrário da seleção precoce tradicional, que é indireta e atua (via avaliação fenotípica) sobre os genes ativados na idade precoce, esperando que eles informem parcialmente sobre sua expressão na idade adulta (Resende *et al.*, 2010).

Diante do exposto, os objetivos deste trabalho foram verificar a acurácia do método da GWS no melhoramento de milho em condições de estresse nutricional e propor novos métodos de melhoramento baseados em GWS.

## MATERIAL E MÉTODOS

### *Material genético e condução experimental*

Para a condução dos experimentos foram utilizados dois grupos de genótipos de milho: o primeiro composto de sete linhagens sem variabilidade genética para EUP, mas contrastantes para EUN (Grupo I), e o segundo de sete linhagens sem variabilidade genética para EUN, mas contrastantes para a EUP (Grupo II). Pelo cruzamento entre as linhagens dos dois grupos, foram obtidas 41 combinações híbridas por meio de cruzamentos em esquema fatorial interpopulacional (desbalanceado). Esses cruzamentos foram realizados de forma recíproca, sendo misturadas as sementes da mesma combinação.

As 41 combinações híbridas foram avaliadas em dois experimentos. O primeiro considerou baixa e alta disponibilidades de nitrogênio (N) e o segundo alta e baixa disponibilidades de fósforo (P). Nesses, foi utilizado o delineamento de blocos ao acaso com duas repetições, em fatorial simples (híbridos x níveis de N ou de P). Os experimentos foram conduzidos em casa de vegetação na Universidade Federal de Viçosa (20° 45' 14" S; 42° 52' 53" W), Brasil, no mês de outubro de 2010.

Primeiramente as sementes foram esterilizadas e germinadas em bandejas de isopor. Após a emergência, foi transplantada uma plântula de cada híbrido para vasos cilíndricos de PVC com 4 dm<sup>3</sup>. No experimento do N, foi utilizado o substrato descrito por Walk *et al.* (2006). Para o experimento do P, o substrato foi composto pela mistura de 50% de areia, 37,5% de vermiculita e 12,5% de solo (horizonte “B” de Latossolo Vermelho-Amarelo Distrófico). O solo foi usado para adsorver o fósforo e não deixá-lo prontamente disponível para a planta.

No experimento do nitrogênio, para os dois níveis, baixo N (BN) e alto N (AN), foi utilizada a solução nutritiva descrita por Chun *et al.* (2005). No segundo experimento, o P foi adicionado na forma de superfosfato triplo, em que, para o nível de baixa disponibilidade de P (BP), misturaram-se 34 mg de P por dm<sup>3</sup> de substrato e no nível de alto P (AP), 192 mg dm<sup>-3</sup>. Os demais nutrientes foram fornecidos via solução nutritiva (Chun *et al.*, 2005).

As plantas foram colhidas no estágio vegetativo de seis folhas completamente expandidas (V6), 25 dias após a semeadura. Para a quantificação da massa da parte aérea seca (MPS, em gramas), a parte aérea da planta foi separada do sistema radicular, acondicionada em sacos de papel e seca em estufa de circulação de ar forçada a 60 °C, por 72 horas.

A quantificação do teor de N foi realizada pelo método descrito por Bremner & Mulvaney (1982) e a do teor de P pela digestão nitroperclórica e leitura em espectrofotômetro a 725 nm.

Por meio desses dados foram estimados os dois componentes da Eficiência de Aquisição de N (EAN) e P (EAP) e Eficiência de Utilização de N (EUtN) e P (EUtP), conforme proposto por Moll *et al.* (1982):

$$EAN = \frac{N(\text{planta})}{N(\text{aplicado})} (\text{mg.mg}^{-1}) \text{ e } EAP = \frac{P(\text{planta})}{P(\text{aplicado})} (\text{mg.mg}^{-1})$$

$$EUtN = \frac{MPS}{N(\text{absorvido})} (\text{g.mg}^{-1}) \text{ e } EUtP = \frac{MPS}{P(\text{absorvido})} (\text{g.mg}^{-1})$$

### Análise de deviance

Para cada nível de disponibilidade de N ou de P, os dados obtidos foram submetidos a análises de deviance, por meio do método da Máxima Verossimilhança Restrita/Melhor Preditor Linear Não Viesado (REML/BLUP), conforme descrito por Resende (2002):

$$\mathbf{y} = \mathbf{Xr} + \mathbf{Zm} + \mathbf{Wf} + \mathbf{Tc} + \mathbf{e}$$

em que  $\mathbf{y}$  é o vetor das médias fenotípicas dos híbridos;  $\mathbf{r}$  é o vetor dos efeitos de repetição (assumidos como fixos) somados à média geral;  $\mathbf{m}$  é o vetor dos efeitos de genitores do Grupo II (assumidos como aleatórios), em que  $m \sim N(0, M)$ , sendo  $M = I\sigma_m^2$ ;  $\mathbf{f}$  é o vetor dos efeitos de genitores do Grupo I (assumidos como aleatórios), em que  $f \sim N(0, F)$ , sendo  $F = I\sigma_f^2$ ;  $\mathbf{c}$  é o vetor dos efeitos da capacidade específica de combinação dos genitores do Grupo I com os genitores do Grupo II (assumidos como aleatórios), em que  $c \sim N(0, C)$ , sendo  $C = I\sigma_c^2$ ; e  $\mathbf{e}$  é o vetor de erros, em que  $e \sim N(0, R)$  e  $R = I\sigma_e^2$ .  $\mathbf{X}$ ,  $\mathbf{Z}$ ,  $\mathbf{W}$  e  $\mathbf{T}$  são matrizes de incidência que relacionam, respectivamente, os efeitos de  $\mathbf{r}$ ,  $\mathbf{m}$ ,  $\mathbf{f}$  e  $\mathbf{c}$  ao vetor  $\mathbf{y}$ .

Em seguida, foram realizadas as análises de deviance conjunta, utilizando um modelo reduzido, visando apenas identificar a significância da interação híbrido x nível de N ou de P, conforme descrito por Resende (2002):

$$\mathbf{y} = \mathbf{Xr} + \mathbf{Zg} + \mathbf{Wi} + \mathbf{e}$$

em que  $\mathbf{y}$  é o vetor das médias fenotípicas dos híbridos;  $\mathbf{r}$  é o vetor dos efeitos de repetição dentro dos níveis de N ou de P (assumidos como fixos), somados à média geral;  $\mathbf{g}$  é o vetor dos efeitos genotípicos dos híbridos (assumidos como aleatórios), em que  $g \sim N(0, G)$ , sendo  $G = \sigma_g^2$ ;  $\mathbf{i}$  é o vetor da interação híbrido x nível de N ou de P (assumidos como aleatórios), em que  $i \sim N(0, I_{ga})$ , sendo  $I_{ga} = I\sigma_{ga}^2$ ; e  $\mathbf{e}$  é o vetor de erros, em que  $e \sim N(0, R)$  e  $R = I\sigma_e^2$ .  $\mathbf{X}$ ,  $\mathbf{Z}$  e  $\mathbf{W}$  são matrizes de incidência que relacionam, respectivamente, os efeitos de  $\mathbf{r}$ ,  $\mathbf{g}$  e  $\mathbf{i}$  ao vetor  $\mathbf{y}$ .

Análises foram realizadas utilizando o software SELEGEN-REML/BLUP (Resende, 2007a).

### Genotipagem da população de estimação

Na genotipagem foram utilizados 80 marcadores microssatélites (SSR), que foram escolhidos pelo seu conteúdo informativo, pela localização nos cromossomos (pelo menos nove por cromossomo) e, principalmente, por apresentarem associação com caracteres de interesse agrônomo (MaizeGDB - <http://www.maizegdb.org/ssr.php>).

A genotipagem foi realizada apenas nas 14 linhagens genitoras, as quais apresentavam mais de sete gerações de autofecundação. Com isso, pode-se fazer a extrapolação dos resultados dos marcadores para as combinações híbridas. Foram consideradas apenas as marcas que apresentaram frequência superior a 5% na população híbrida.

### Seleção genômica ampla (GWS)

Nessa etapa, foram analisados nos híbridos simples apenas os caracteres identificados como significativos dentro de cada nível de N ou de P, por meio da análise de deviance. Foi empregado o método da regressão aleatória (Random Regression) do tipo BLUP (RR-BLUP) aplicado à seleção genômica ampla (RR-BLUP/GWS), conforme descrito por Resende *et al.* (2010):

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zh} + \mathbf{e}$$

em que  $\mathbf{y}$  é o vetor de médias fenotípicas de cada híbrido;  $\mathbf{b}$  é o vetor de efeitos fixos;  $\mathbf{h}$  é o vetor dos efeitos aleatórios das marcas; e  $\mathbf{e}$  refere-se ao vetor de resíduos aleatórios.  $\mathbf{X}$  e  $\mathbf{Z}$  são as matrizes de incidência para  $\mathbf{b}$  e  $\mathbf{h}$ . A matriz de incidência  $\mathbf{Z}$  contém os valores 0, 1 e 2 para o número de marcas. A estrutura de médias e variâncias é definida como:  $h \sim N(0, G)$ ;  $E(\mathbf{y} = \mathbf{Xb})$ ;  $e \sim N(0, R = I\sigma_e^2)$ ;  $\text{Var}(\mathbf{y}) = \mathbf{V} = \mathbf{ZGZ}' + \mathbf{R}$ ; e  $G = I\sigma_A^2/n$ .

As equações de modelo misto para a predição de  $\mathbf{h}$  por meio do método RR-BLUP/GWS equivalem a:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \frac{\sigma_e^2}{\sigma_A^2/n} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{h}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

em que  $\sigma_A^2$  se refere à variância genética aditiva total do caráter;  $\sigma_e^2$  é a variância residual; e  $n$  o número total de marcadores ponderados, dado aproximadamente por

$$n = \left[ 2 \sum_i p_i (1 - p_i) \right], \text{ em que } p_i \text{ é a frequência da marca } i;$$

$\sigma_e^2 / (\sigma_A^2 / n) = (I - h_{Am}^2) (h_{Am}^2 / n)$ , em que  $h_{Am}^2$  é a herdabilidade no sentido restrito em nível de médias de híbridos em várias repetições, estimada por

$$h_{Am}^2 = \frac{r \cdot h_A^2}{1 + (r - 1)h_A^2}, \text{ sendo } r \text{ o número de repetições. Nes-}$$

se método é considerado que cada loco explica  $\frac{1}{n}\sigma_A^2$ ; ou seja, partes iguais da variância genética são atribuídas a

todos os locos. O valor genético genômico global do indivíduo  $j$  é dado por  $VGG = \hat{y}_j = \sum_i Z_i \hat{h}_i$ .

O procedimento RR-BLUP/GWS permite também identificar os marcadores com maiores efeitos, objetivando processar análises com subgrupos menores de marcadores e determinar quantos e quais maximizam a acurácia seletiva. Desse modo, primeiramente foi analisado todo o conjunto de marcadores em todos os indivíduos fenotipados (população de estimação completa). A partir desses resultados, foram gerados nove subgrupos de marcas, formados por 214 (todas as marcas), 150, 100, 50, 25, 10, 5, 2 e 1. A formação desses subgrupos foi realizada baseando-se no efeito (em módulo) das marcas obtidas na análise inicial; ou seja, considerando todos os 214 marcadores. A partir desses valores, de modo sequencial, foram retirados da análise os de menor efeito, até atingir o extremo de apenas uma marca (a de maior efeito) formando o subgrupo.

A validação cruzada foi realizada pela reamostragem de grupos de indivíduos via procedimento Jackknife (Helshe & Forrester, 1983), conforme descrito por Resende *et al.* (2010). Nesse processo, em cada subgrupo analisado a GWS foi avaliada ao calcular a correlação entre o valor genético predito por esse método e o valor fenotípico observado nos indivíduos. Essa correlação é conhecida como capacidade preditiva ( $r_{\hat{y},y}$ ), sendo dada teoricamente pela acurácia de seleção ( $r_{gg}$ ) multiplicada pela raiz quadrada da herdabilidade individual ( $h$ ) ( $r_{\hat{y},y} = r_{gg}h$ ) (Resende *et al.*, 2010).

Para os caracteres que apresentaram altos valores de  $r_{gg}$  e de  $r_{\hat{y},y}$ , a GWS foi simulada fazendo uma comparação com a fenotípica quanto ao ganho de seleção por unidade de tempo. Como padrão da seleção fenotípica foram considerados os métodos da Seleção Recorrente Intrapopulacional (SRI) e Recíproca (SRR). Para isso, no caso da SRI foi considerada (simulada) a formação de uma população-base, originada do livre cruzamento entre diversas linhagens. Dessa população, foram obtidas 200 progênies de meios-irmãos maternos para avaliação e 200 progênies  $S_j$  para recombinação. Essas progênies foram avaliadas em um experimento com duas repetições, sendo a parcela constituída de 25 plantas, totalizando 10.000 plantas avaliadas. Além disso, foi considerada a seleção de 10% das progênies superiores. Para a SRR, foi considerada a formação de duas populações-base, originadas do livre cruzamento entre diversas linhagens dentro de cada grupo heterótico (GH). Entre as populações, foram obtidas 200 progênies de meios-irmãos maternos interpopulacionais para avaliação e dentro de cada população 200 progênies  $S_j$  para recombinação. Essas progênies foram avaliadas em um experimento com duas repetições, sendo a parcela constituída de 25 plantas, totalizando 10.000 plantas ava-

liadas. Além disso, foi considerada a seleção de 10% dos híbridos interpopulacionais superiores.

No caso da GWS intrapopulacional, foi considerada, na mesma população de 10.000 plantas, a seleção dos 200 melhores indivíduos antes do florescimento e a sua recombinação; ou seja, uma seleção a 2% de intensidade, em nível de indivíduo, com coincidência entre as unidades e gerações de avaliação e recombinação. No caso da GWS recíproca, a seleção foi considerada na mesma intensidade e no mesmo tamanho efetivo da intrapopulacional; entretanto, divididas dentro de cada GH. Assim, considerando variâncias genéticas e fenotípicas igualmente divididas nos dois GHs, os ganhos com a seleção obtidos por meio das SRI e SRR seriam teoricamente iguais.

Os ganhos com a seleção em desvios-padrão genéticos aditivos foram estimados por  $GS = K \cdot r_{gg}$ , em que  $k = 1,76$  para os métodos da SRI e SRR e de 2,42 para o da GWS; e  $r_{gg}$  é a acurácia seletiva (calculada via validação cruzada para o método GWS), obtida por  $\sqrt{1/4 h_{Am}^{2*}}$  para os métodos da SRI e SRR, em que  $h_{Am}^{2*} = \frac{r(1/2)h_A^2}{1+(r-1)(1/4)h_A^2}$ . A

quantidade  $r$  refere-se ao número de repetições. Esse valor de  $1/4$  ponderando a herdabilidade refere-se aos tipos de progênie considerados para a avaliação e recombinação. A partir dessas estimativas foi obtida, para cada caráter, a eficiência relativa dos métodos baseados na GWS

em relação à SRI e SRR:  $EF_{GWS(\%)} = \frac{GS_{GWS} \cdot t_{SRI, SRR}}{GS_{SRI, SRR} \cdot t_{GWS}} \cdot 100$ , em

que  $GS$  é o ganho de seleção estimado em desvios-padrão genético aditivos para cada um dos métodos (GWS, SRI e SRR); e  $t$  é o tempo médio (em anos) do ciclo de cada um dos métodos, sendo 1,5 para o da SRI e da SRR e de 1,0 ou 0,5 para o da GWS, considerando a possibilidade de um ou dois ciclos por ano, respectivamente.

As análises foram realizadas utilizando o software SELEGEN Genômica RR-BLUP/GWS (Resende, 2007b).

## RESULTADOS E DISCUSSÃO

### Análises de deviance

Nas análises de deviance individuais para cada nível de N ou de P, pelo menos uma das fontes de variação genética foi significativa em quase todos os caracteres avaliados ( $p \leq 0,05$ ), sendo em sua maioria para os genitores do Grupo I no experimento do N e do Grupo II no experimento do P. As exceções foram EAP e EUTp em AP e EUTn em BN. As diferenças genéticas significativas observadas indicam existência de variabilidade genética, o que possibilita seleção e ganhos genéticos dentro de cada nível de N ou de P. Os coeficientes de variação apresentaram valores dentro do aceitável para esse tipo de estudo (Fritsche-Neto *et al.*, 2010).

Nas análises de deviance conjuntas, a variância de híbridos e da sua interação com os níveis de N ou de P foi significativa para a maioria dos caracteres avaliados, sendo as exceções EUtN e EUtP. Isso indica que a seleção deve ser realizada para ambientes específicos e não em razão do comportamento médio, pois, provavelmente, os genes são expressos diferentemente de acordo com a disponibilidade do nutriente para as plantas (Fritsche-Neto *et al.*, 2010).

### Acurácia seletiva da GWS

Na população híbrida foram identificados 214 alelos com frequência superior a 5%, com média de 2,68 marcas por loco, variando de dois a cinco.

Houve grande variação entre os caracteres para o número de marcas que maximizaram a capacidade preditiva e a acurácia da GWS, sendo esse de cinco para MPS em AN e BN e EAN em AN a 25 marcas para EUtN em AN (Tabela 1). No caso do P, essa foi de um para MPS em AP a 50 marcas para EUtP em BP (Tabela 2).

Normalmente, as capacidades preditivas diminuem com o aumento do número de marcadores além de um número ótimo (Resende *et al.*, 2010). Neste trabalho, pode-se observar que, quando foram consideradas as 214 marcas, as capacidades preditivas foram muito baixas, ou seja, o aumento do número de marcadores não aumenta linearmente a acurácia da GWS pelo método RR-BLUP (Tabelas 1 e

2). Assim, é fundamental, para cada caráter a ser selecionado, estabelecer o número ótimo de marcadores que maximizem a acurácia na população de validação.

Para os caracteres EAN, EUtN e MPS em AN e para EAN em BN, nos pontos máximos de acurácia para cada caráter, os marcadores explicaram 55%, 49%, 49% e 30% da variação genética total, respectivamente, e propiciaram alta acurácia seletiva (74%, 70%, 70% e 55%, respectivamente) (Tabela 1). No experimento do P, para os caracteres MPS em AP, EAP e EUtP em BP nos pontos máximos de acurácia para cada caráter os marcadores explicaram 88%, 66% e 87% da variação genética total, respectivamente, e também propiciaram alta acurácia seletiva (94%, 82% e 93%, respectivamente) (Tabela 2). Não foram obtidos resultados consistentes para o caráter MPS em BN e em BP. Isto pode ser devido, principalmente, ao número reduzido de marcadores utilizados, quando se considerar o tamanho do genoma do milho, de modo que não foram detectadas marcas em desequilíbrio de ligação com os supostos QTL que controlam esses caracteres.

A diminuição no número de marcas usadas na GWS a partir do ponto de máxima acurácia também faz com que ocorra redução na capacidade preditiva, acurácia e proporção da variância genética explicada (Resende *et al.*, 2010). Exemplo disso é que quando foi considerada apenas a marca de maior efeito, explicou-se apenas 3%, 6%,

**Tabela 1.** Estimativas da herdabilidade no sentido restrito individual ( $h^2_A$ ), da capacidade preditiva ( $r^2_{y\hat{y}}$ ) e da acurácia ( $r^2_{gg}$ ) da seleção genômica, proporção da variância genética (% Vg) explicada pelas marcas considerando diferentes quantidades desses subgrupos nas análises de validação da Seleção Genômica Ampla e acurácia fenotípica ( $r_{fen}$ ) para as eficiências de absorção (EAN) e de utilização (EUtN) de N e das massas de parte aérea seca (MPS) em alto (AN) e baixo (BN) nitrogênio em milho, Viçosa-MG, Brasil

Caráter	$h^2_A$	Estimativa	Número de marcas consideradas								
			214	150	100	50	25	10	5	2	1
EAN (AN)	0,62	$r^2_{y\hat{y}}$	0,61	0,62	0,62	0,63	0,63	0,66	0,68	0,33	0,22
		$r^2_{gg}$	0,67	0,68	0,68	0,69	0,69	0,72	0,74	0,36	0,24
		$r_{fen}$	0,39	0,39	0,39	0,39	0,39	0,39	0,39	0,39	0,39
		%Vg	0,45	0,46	0,46	0,47	0,47	0,52	0,55	0,13	0,06
EAN (BN)	0,72	$r^2_{y\hat{y}}$	0,43	0,44	0,46	0,47	0,47	0,48	0,34	0,25	0,16
		$r^2_{gg}$	0,50	0,51	0,52	0,54	0,54	0,55	0,39	0,28	0,18
		$r_{fen}$	0,37	0,37	0,37	0,37	0,37	0,37	0,37	0,37	0,37
		%Vg	0,25	0,26	0,27	0,29	0,29	0,30	0,15	0,08	0,03
EUtN (AN)	0,13	$r^2_{y\hat{y}}$	0,19	0,26	0,30	0,33	0,33	0,32	0,30	0,24	0,25
		$r^2_{gg}$	0,40	0,55	0,63	0,70	0,70	0,69	0,64	0,51	0,53
		$r_{fen}$	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17
		%Vg	0,16	0,30	0,40	0,49	0,49	0,48	0,41	0,26	0,28
MPS (AN)	0,75	$r^2_{y\hat{y}}$	0,59	0,59	0,60	0,60	0,61	0,65	0,65	0,28	0,21
		$r^2_{gg}$	0,63	0,64	0,64	0,65	0,65	0,70	0,70	0,30	0,22
		$r_{fen}$	0,40	0,40	0,40	0,40	0,40	0,40	0,40	0,40	0,40
		%Vg	0,40	0,41	0,41	0,42	0,43	0,49	0,49	0,09	0,05
MPS (BN)	0,66	$r^2_{y\hat{y}}$	0,20	0,22	0,23	0,25	0,27	0,34	0,35	0,16	-0,07
		$r^2_{gg}$	0,23	0,24	0,25	0,28	0,30	0,38	0,39	0,18	-0,08
		$r_{fen}$	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38
		%Vg	0,05	0,06	0,06	0,08	0,09	0,14	0,15	0,03	0,01

28% e 5% da variação genética para EAN em BN, EAN, EUtN e MPS em AN, respectivamente (Tabela 1). Já no P apenas 23%, 2% e 1% da variação genética para EAP, EUtP e MPS em BP, respectivamente (Tabela 2). Assim, é possível observar que a Seleção Assistida por Marcadores (SAM) baseada em um ou poucos marcadores de efeitos significativos propicia eficiência inferior à GWS. A exceção foi para o caráter MPS em AP, em que, considerando apenas a marca de maior efeito, foi possível explicar 88% da variação genética. Nesse sentido, a GWS apresenta outra grande vantagem em relação aos métodos SAM. Isso se deve à GWS estimar o efeito de todos os marcadores de modo simultâneo, o que reduz significativamente viés na estimativa dos seus efeitos. Com isso, é possível encontrar marcas que estão em desequilíbrio de ligação com o caráter de interesse e não eram consideradas ou identificadas por meio das técnicas tradicionais.

Teoricamente, com populações maiores e maior número de marcadores aumenta-se a probabilidade de encontrar marcas em desequilíbrio de ligação com características de interesse (Flint-Garcia *et al.*, 2003). Entretanto, neste trabalho, mesmo com tamanho efetivo e número de marcadores relativamente pequeno, foi possível serem obtidos bons resultados, pois em todas as características avaliadas, nos pontos de máxima acurácia da GWS, esta foi superior à acurácia fenotípica (Tabelas 1 e 2). O acréscimo na acurácia é devido ao uso da matriz de parentesco real; ou seja, baseada nas marcas que os indivíduos apresentam e não como a tradicional, com base em esperança matemática (Resende *et al.*, 2010).

O número reduzido de marcadores explicando grande parte da variação genética ou maximizando a acurácia é muito interessante do ponto de vista prático. Nesses casos, um pequeno grupo de marcadores, previamente selecionados, poderia ser usado nas populações de seleção. Com isso, seria possível reduzir o intervalo entre ciclos de seleção, ou, ainda, usar métodos não destrutivos de avaliação, possibilitando aumentar os ganhos genéticos.

### Ganhos com a seleção e novos métodos de melhoramento

Considerando as estimativas dos ganhos preditos, em desvios-padrão genéticos aditivos, com um ciclo de GWS por ano, é possível obter aumento na eficiência de seleção em relação aos métodos de seleção recorrente para os caracteres EAP e EUtP em BP e para MPS em AP de 1.028%, 960% e 1.241%, respectivamente (Tabela 3). Quando foi considerada a possibilidade de dois ciclos de GWS por ano, esses valores passaram para 3.317%, 3.098% e 4.004%, respectivamente. No experimento do N, com um ciclo de GWS por ano, para os caracteres EAN em baixo N (BN), EAN, EUtN e MPS em AN, o aumento na eficiência foi de 307%, 394%, 816,99% e 364%, respectivamente. Na possibilidade de dois ciclos de GWS por ano, esses valores passaram para 991%, 1.271%, 2.636% e 1.174%, respectivamente.

Existem poucos métodos propostos para a implementação da GWS no melhoramento do milho. Neste trabalho foi proposto, primeiramente, um esquema de Seleção Recorrente Genômica Recíproca - SRGR (Figura 1A). Como

**Tabela 2.** Estimativas da herdabilidade no sentido restrito individual ( $h^2_A$ ), da capacidade preditiva ( $r_{y\hat{y}}$ ) e da acurácia ( $r_{gg}$ ) da seleção genômica, proporção da variância genética (%Vg) explicada pelas marcas considerando diferentes quantidades desses subgrupos nas análises de validação da Seleção Genômica Ampla e acurácia fenotípica ( $r_{fen}$ ) para as eficiências de absorção (EAP) e de utilização (EUtP) de P e massas de parte aérea seca (MPS) em alto (AP) e baixo (BP) fósforo, em milho, Viçosa-MG, Brasil

Caráter	$h^2_A$	Estimativa	Número de marcas consideradas								
			214	150	100	50	25	10	5	2	1
EAP (BP)	0,11	$r_{y\hat{y}}$	-0,04	0,05	0,14	0,21	0,22	0,31	0,34	0,36	0,21
		$r_{gg}$	-0,08	0,12	0,32	0,46	0,50	0,70	0,75	0,82	0,48
		$r_{fen}$	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,16
		%Vg	0,01	0,01	0,10	0,21	0,25	0,49	0,57	0,66	0,23
EUtP (BP)	0,17	$r_{y\hat{y}}$	0,43	0,47	0,47	0,50	0,42	0,44	0,09	0,20	-0,08
		$r_{gg}$	0,79	0,87	0,88	0,93	0,78	0,82	0,16	0,38	-0,15
		$r_{fen}$	0,20	0,20	0,20	0,20	0,20	0,20	0,20	0,20	0,20
		%Vg	0,63	0,75	0,77	0,87	0,60	0,67	0,03	0,14	0,02
MPS (AP)	0,06	$r_{y\hat{y}}$	-0,08	0,03	0,14	0,22	0,26	0,26	0,26	0,30	0,32
		$r_{gg}$	-0,24	0,07	0,41	0,63	0,75	0,75	0,76	0,88	0,94
		$r_{fen}$	0,13	0,13	0,13	0,13	0,13	0,13	0,13	0,13	0,13
		%Vg	0,06	0,01	0,17	0,40	0,56	0,57	0,58	0,77	0,88
MPS (BP)	0,11	$r_{y\hat{y}}$	-0,23	-0,15	-0,06	0,05	0,11	0,13	0,07	0,06	0,04
		$r_{gg}$	-0,54	-0,33	-0,13	0,11	0,26	0,29	0,16	0,14	0,09
		$r_{fen}$	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,16
		%Vg	0,29	0,11	0,02	0,01	0,07	0,09	0,03	0,02	0,01

em todo método baseado no princípio GWS, é necessário primeiramente gerar a população de estimação, na qual serão realizadas a genotipagem e a fenotipagem dos genótipos avaliados. No caso da SRGR, a seleção deve ser realizada dentro de cada grupo heterótico, mas baseada no desempenho dos híbridos interpopulacionais. Nessa fase, podem ser genotipadas as linhagens genitoras e fenotipados os híbridos obtidos do cruzamento entre elas. A partir desses dados, são estimados os valores genéticos dos marcadores e o número desses que maximizam a acurácia seletiva para cada caráter de interesse. A validação é realizada nos mesmos híbridos, mas de forma independente, por meio do procedimento Jackknife (Heltshe & Forrester, 1983). Finalmente, faz-se a seleção genômica dentro de cada grupo heterótico, a partir dos resultados obtidos na população de descoberta e de validação.

Na SRGR as populações seriam geradas pela livre recombinação das melhores linhagens genitoras dentro de cada grupo heterótico. Após a sua obtenção, elas são semeadas a campo e de cada planta retira-se, no estado juvenil, uma amostra foliar para extração de DNA e a genotipagem com os marcadores identificados no processo de estimação e validação. As plantas são então classificadas em razão dos seus valores genéticos genômicos dados pelas marcas que apresentam em seu genótipo. As consideradas superiores para a composição de híbridos interpopulacionais são identificadas antes do florescimento e as demais, eliminadas. Com isso, há coincidência entre as unidades e gerações de seleção e de recombinação em apenas um ciclo a campo. Além disso, a seleção é realizada no indivíduo, em ambos os sexos, e não precisa ser feita no ambiente ideal de avaliação ou cultivo. Desse modo, há aumento na acurácia seletiva, nos ganhos por unidade de tempo e, conseqüentemente, na eficiência do processo de melhoramento.

Outro ponto positivo da SRGR é a possibilidade de monitorar a frequência dos alelos nos dois grupos heteróticos. Assim, podem ser selecionados os que mais contribuem para os desvios de dominância e para maximizar a divergência genética entre as populações; ou seja, trabalhar nos dois principais componentes da heterose. Além

disso, a SRGR pode ser facilmente adaptada para a Seleção Recorrente Genômica Intrapopulacional - SRGI (Figura 1B).

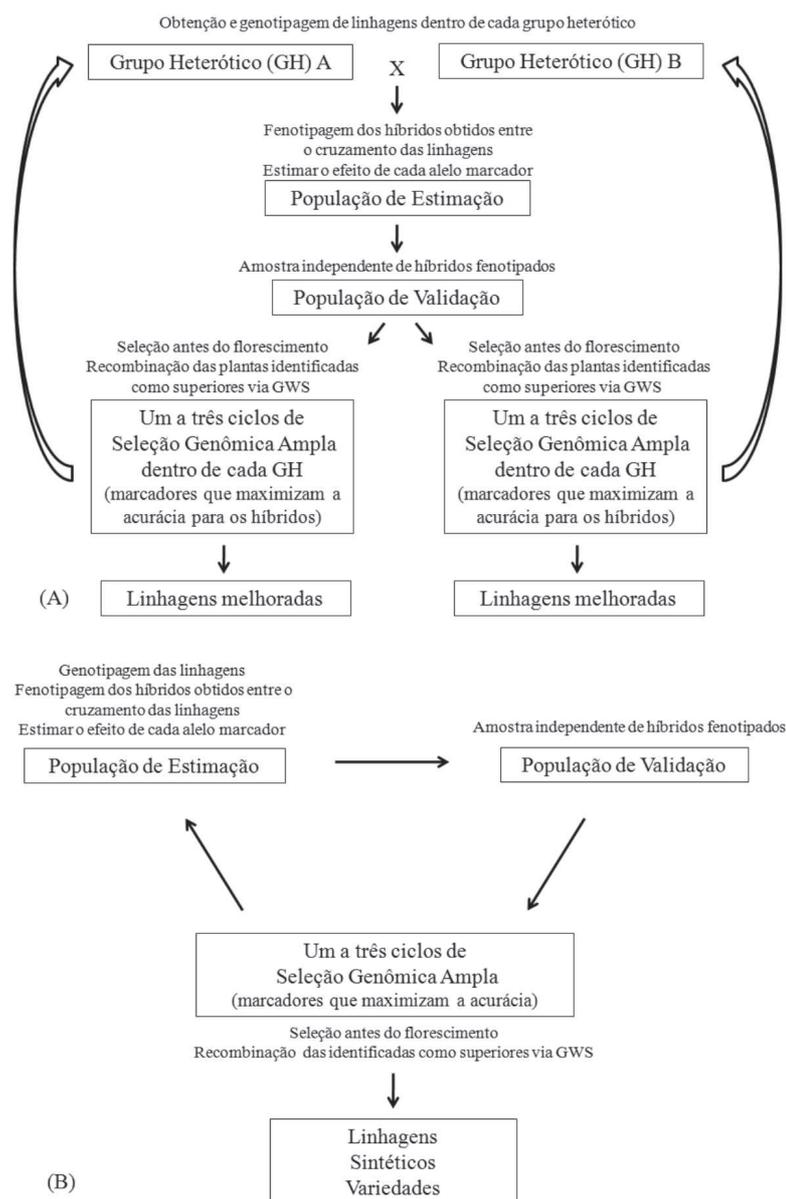
Pode-se usar a GWS várias vezes sem a avaliação fenotípica. Entretanto, é recomendável reestimar os efeitos e o número dos marcadores que maximizam a capacidade preditiva para cada caráter a cada duas ou três gerações de seleção. Isso se deve ao fato de que com a seleção pode haver mudanças significativas no desequilíbrio de ligação da população, principalmente pelo aumento na frequência ou na fixação de alguns alelos (Flint-Garcia *et al.*, 2003).

Nas fases de reestimação dos efeitos das marcas, poderiam ser utilizadas ferramentas como a de duplos haploides, buscando acelerar o processo de obtenção de linhagens puras. Nesse sentido, também poderiam ser usadas testadores na obtenção dos híbridos interpopulacionais, o que diminuiria significativamente os custos.

A questão experimental do uso de duas repetições com uma planta por parcela merece ser comentada. Essa prática é usual em experimentos em condição de estresse nutricional em casa de vegetação e com materiais geneticamente uniformes. Em muitos artigos, de diversos grupos de pesquisa de todo mundo, é utilizada essa metodologia (Trachsel *et al.*, 2009; Chun *et al.*, 2005; Liu *et al.*, 2004; Zhu *et al.*, 2005). Isso se deve ao fato de que os ganhos em precisão com o aumento da parcela ou no número de repetições nesse tipo de estudo são pequenos. Assim é mais vantajoso aumentar o número de genótipos avaliados. Tecnicamente, isso deve ser avaliado pelos próprios resultados experimentais usando estatísticas adequadas. Estimativas de acurácia seletiva são a melhor escolha nesse caso. Neste trabalho, as acurácias foram da ordem de 70% para a maioria dos caracteres, as quais são suficientes para maximizar o ganho genético por unidade de tempo, por meio do uso do novo método de seleção recorrente proposto, o qual explora grande parte da variação genética aditiva presente em uma população e utiliza uma só estação de plantio, havendo coincidência entre unidade de seleção e de recombinação. Os métodos tradicionais de seleção recorrente utilizam apenas 1/4 ou 1/2 da variação genética aditiva, quando se usam progênies de meios-irmãos ou irmãos germanos, respectivamente.

**Tabela 3.** Estimativas dos ganhos com a seleção (GS), em desvios-padrão genético aditivos, em cada método de seleção considerado e a eficiência relativa (ER) da Seleção Genômica Ampla (GWS) com um (GWS 1) ou dois (GWS 2) ciclos por ano, em relação aos métodos de Seleção Recorrente (SR), para as eficiências de absorção (EAN e EAP) e de utilização (EUtN e EUtP) e massa de parte aérea seca (MPS) em baixo (BN) e alto nitrogênio (AN) ou em baixo (BP) e alto fósforo (AP), em milho, Viçosa-MG, Brasil

Estimativa	Método	Caráter						
		EAN (AN)	EAN (BN)	EUtN (AN)	MPS (AN)	EAP (BP)	EUtP (BP)	MPS (AP)
GS	SR	0,69	0,65	0,31	0,70	0,29	0,35	0,22
	GWS	1,80	1,32	1,67	1,70	1,97	2,25	1,83
ER (%)	GWS 1	394	307	817	364	1028	960	1241
	GWS 2	1271	991	2636	1174	3317	3098	4004



**Figura 1.** Esquemas propostos de Seleção Recorrente Genômica Recíproca (A) e de Seleção Recorrente Genômica Intrapopulacional (B).

## CONCLUSÕES

Com o uso da GWS houve aumento significativo na acurácia seletiva e nos ganhos genéticos por unidade de tempo. Com base nos resultados, foram propostos os métodos da Seleção Recorrente Genômica Recíproca e Seleção Recorrente Genômica Intrapopulacional para a implementação da GWS no melhoramento do milho.

Devido à alta acurácia observada para a GWS, é possível concluir que os métodos propostos de seleção recorrente genômica neste trabalho podem modificar significativamente a forma como é feita a seleção nos programas de melhoramento do milho, tanto para condições de estresses abióticos como para outros caracteres de interesse agrônômico.

## AGRADECIMENTOS

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico, à Coordenação de Pessoal de Nível Superior e à Fundação de Amparo à Pesquisa do Estado de Minas Gerais, pelo suporte financeiro.

## REFERÊNCIAS

Bremner JM & Mulvaney CS (1982) Nitrogen total. In: Page AL (Ed.) Methods of soil analysis. Madison, Soil Science Society of America Press. p.595-624.

Chun L, Mi G, Li J, Chen F and Zhang F (2005) Genetic analysis of maize root characteristics in response to low nitrogen stress. Plant and Soil, 276:369-382.

Flint-Garcia SA, Thornsberry JM & Buckler ES (2003) Structure of linkage disequilibrium in plants. Annual Review of Plant Biology, 54:357-374.

- Fritsche-Neto R, Miranda GV, DeLima RO, Souza LV & Silva J (2010) Herança de caracteres associados à eficiência de utilização do fósforo em milho. *Pesquisa Agropecuária Brasileira*, 45:465-471.
- Heltshe JF & Forrester NE (1983) Estimating species richness using the jackknife procedure. *Biometrics*, 39:1-11.
- Liu Y, Mi GH, Chen FJ, Zhang JH & Zhang FS (2004) Rhizosphere effect and root growth of two maize (*Zea mays* L.) genotypes with contrasting P efficiency at low P availability. *Plant Science*, 167:217-223.
- Meuwissen T, Hayes B & Goddard M (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157:1819-1829.
- Moll RH, Kamprath EJ & Jackson WA (1982) Analysis and interpretation of factors which contribute to efficiency of nitrogen utilization. *Agronomy Journal*, 74:562-564.
- Resende MDV (2002) *Genética biométrica e estatística no melhoramento de plantas perenes*. Brasília, Embrapa Informação Tecnológica. 975p.
- Resende MDV (2007a) SELEGEN-REML/BLUP: Sistema estatístico e seleção genética computadorizada via modelos lineares mistos. Colombo, Embrapa Florestas. 361p.
- Resende MDV (2007b) SELEGEN Genômica RR-BLUP. Sistema de Seleção Genômica Ampla (GWS) computadorizada via modelos lineares mistos. CD-ROM.
- Resende MDV, Resende Jr. MFRR, Aguiar AM, Abad JIM, Missiaggia AA, Sansaloni C, Petroli C & Grattapaglia D (2010) Computação da seleção genômica ampla (GWS). Colombo, Embrapa Florestas. 79p. (Série Documentos, 210).
- Trachsel S, Messmer R, Stamp P & Hund A (2009) Mapping of QTLs for lateral and axile root growth of tropical maize. *Theoretical and Applied Genetics*, 119:1413-1424.
- Walk TC, Jaramillo R & Lynch JP (2006) Architectural tradeoffs between adventitious and basal roots for phosphorus acquisition. *Plant Soil*, 279:347-366.
- Zhu JM, Kaeppler SM & Lynch JP (2005) Mapping of QTLs for lateral root branching and length in maize (*Zea mays* L.) under differential phosphorus supply. *Theoretical and Applied Genetics*, 111:688-695.