



**SAMUEL BONFIM FERNANDES**

**ESTRATÉGIAS PARA MELHORAR A  
EFICIÊNCIA DA SELEÇÃO EM  
PROGÊNIES DE MEIOS IRMÃOS POR  
MEIO DE SIMULAÇÃO**

**LAVRAS - MG  
2016**

**SAMUEL BONFIM FERNANDES**

**ESTRATÉGIAS PARA MELHORAR A EFICIÊNCIA DA SELEÇÃO EM  
PROGÊNIES DE MEIOS IRMÃOS POR MEIO DE SIMULAÇÃO**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, área de concentração em Genética e Melhoramento de Plantas, para a obtenção do título de Doutor.

Orientador

Prof. Dr. Daniel Furtado Ferreira

Coorientador

Prof. Dr. Magno Antonio Patto Ramalho

**LAVRAS - MG  
2016**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca  
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Fernandes, Samuel Bonfim.

Estratégias para melhorar a eficiência da seleção em progênies  
de meios irmãos por meio de simulação / Samuel Bonfim  
Fernandes. – Lavras: UFLA, 2016.  
68 p. : il.

Tese(doutorado)–Universidade Federal de Lavras, 2016.

Orientador: Daniel Furtado Ferreira.

Bibliografia.

1. Genética Quantitativa. 2. Seleção entre e dentro. 3. Ganho  
genético. 4. Número de progênies. 5. Parcela de uma planta. I.  
Universidade Federal de Lavras. II. Título.

**SAMUEL BONFIM FERNANDES**

**ESTRATÉGIAS PARA MELHORAR A EFICIÊNCIA DA SELEÇÃO EM  
PROGÊNIES DE MEIOS IRMÃOS POR MEIO DE SIMULAÇÃO  
STRATEGIES TO IMPROVE SELECTION EFFICIENCY IN HALF SIB  
PROGENIES THROUGH SIMULATION**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, área de concentração em Genética e Melhoramento de Plantas, para a obtenção do título de Doutor.

APROVADA em 16 de Setembro de 2016.

Dr. Paulo Evaristo de Oliveira Guimarães	Embrapa Milho e Sorgo
Dr. Fernando Hernique Ribeiro Barrozo Toledo	CIMMYT
Dr. Flávia Maria Avelar Gonçalves	UFLA
Dr. Magno Antonio Patto Ramalho	UFLA

Prof. Dr. Daniel Furtado Ferreira  
Orientador

Prof. Dr. Magno Antonio Patto Ramalho  
Coorientador

**LAVRAS - MG  
2016**

*À minha Família, Pai (Vileste), Mãe (Maria Ilda),  
Lucas (Irmão), Márcia (Irmã), Maria Inácia (Noiva)*

**DEDICO**

## AGRADECIMENTOS

A Deus, por me dar forças nas horas difíceis.

À Universidade Federal de Lavras, pela oportunidade de realizar esse trabalho.

À FAPEMIG e ao CNPq, pela concessão da bolsa de doutorado.

À CAPES, pela bolsa de doutorado sanduíche no exterior.

Ao professor Daniel e ao professor Magno, pela orientação, paciência, disposição para ajudar e, principalmente, por serem um exemplo de profissionalismo e dedicação.

Ao Fernando Toledo, pela ajuda na programação das simulações.

Ao professor Júlio Bueno, pela disponibilização dos computadores, além da disposição para tirar algumas dúvidas que tive ao longo do trabalho.

Aos membros da banca, pela disponibilidade e por contribuírem com a melhoria do trabalho.

A todos os professores do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, pelos ensinamentos transmitidos.

Aos funcionários do Departamento de Biologia, em especial, Lilian, Dona Iron, Zélia e Rafaela, por serem sempre tão atenciosos e prestativos.

A todos os colegas e amigos da genética, pela amizade e companheirismo.

À minha família, pelo apoio e incentivo durante todo o tempo.

À minha noiva Maria Inácia, pela paciência e compreensão em todos os momentos,

**MUITO OBRIGADO!**

## RESUMO

O melhoramento de plantas oferece várias alternativas para a obtenção de cultivares superiores. Quando o objetivo é melhorar o desempenho de uma população, as estratégias utilizadas podem envolver a seleção tanto de indivíduos quanto de progênies. Nesse caso, as progênies de meios-irmãos são, muitas vezes, preferidas, especialmente pela facilidade de condução. Do exposto, realizou-se este trabalho com o objetivo de avaliar, por meio de simulação computacional, estratégias que auxiliem a tomada de decisão com relação aos números de progênies, repetições e plantas por parcela que aumentam a eficiência de programas de melhoramento que utilizam progênies de meios-irmãos. Foram simulados experimentos com progênies de meios-irmãos, considerando um caráter controlado por 100 genes, variando o grau médio de dominância (0 ou 1), a frequência alélica média da população (0,2, 0,5 ou 0,8) e a herdabilidade (0,1, 0,25 ou 0,5). Considerou-se um programa com a capacidade de avaliar 6.000 plantas em experimentos com múltiplas plantas ou com uma planta por parcela. No primeiro caso, utilizou-se  $N_f$  progênies, com  $N_f$  variando entre 60 e 600, em 2, 3, 4 ou 5 repetições, com  $k$  plantas/parcela (5, 10 ou 20). A recombinação aconteceu com e sem o uso de sementes remanescentes, sendo a seleção dentro apenas do lado materno. No caso de uma planta por parcela,  $N_f$  variou entre 12 e 600 progênies, com 500 e 10 repetições, respectivamente. Considerou-se, neste caso, a seleção entre e dentro nos dois sexos. Os resultados obtidos neste estudo mostraram que o aumento do número de repetições em detrimento do aumento do número de plantas por parcela, proporciona maiores ganhos genéticos. Estes ganhos foram, em média, maiores em experimentos com uma planta por parcela. Evidenciou-se também que, é possível obter ganhos genéticos semelhantes com diferentes números de progênies. No entanto, os resultados obtidos com o uso de poucas progênies apresentam grandes oscilações e conseqüentemente maiores riscos. Por outro lado, o uso de poucas plantas por progênie resulta em ganhos, em média, inferiores. Para as situações consideradas neste estudo, o uso de aproximadamente 100 progênies em 60 repetições, resultou na melhor relação entre a magnitude dos ganhos genéticos e a precisão das estimativas obtidas.

Palavras-chave: Genética quantitativa. Seleção entre e dentro. Ganho genético. Número de progênies. Parcela de uma planta.

## ABSTRACT

Plant breeding provides several alternatives for obtaining superior cultivars. When the goal is to improve the performance of a population, the strategies used may involve the selection of either individuals or progenies. In this case, half-sib progenies are often preferred, especially for being easy to work with. In this context, we performed this study to evaluate, through computer simulation, strategies that help in the decision-making regarding to the number of progenies, repetitions and plants per plot to increase effectiveness in breeding programs that uses half-sib progenies. Experiments with half-sib progenies were simulated considering a trait controlled by 100 genes, varying the average degree of dominance (0 or 1), the mean allelic frequency of the population (0,2, 0,5 or 0,8) and heritability (0,1, 0,25 or 0,5). It was considered a breeding program with the ability to assess 6.000 plants in experiments with multiple plants or single plant plot. In the first case, we used  $N_f$  progenies, with  $N_f$  ranging between 60 and 600, in 2, 3, 4 or 5 repetitions with  $k$  plants/plot (5, 10 or 20). Recombination occurred with and without the use of remaining seeds, with selection within progenies only in the maternal side. In the case of one plant per plot,  $N_f$  ranged between 12 and 600 progenies, with 500 and 10 repetitions, respectively. It was considered in this case, the selection among and within progenies in both sexes. The results of this study shows that increasing the number of repetitions over the number of plants per plot, provides larger genetic gain. The gains were, on average, higher in experiments with single plant plot. It also shows that it is possible to obtain similar genetic gain with different number of progenies. However, the results obtained with the use of fewer progeny shows large fluctuations and therefore greater risks. On other hand, the use of fewer plants per progeny, on average, result in lower gain. For the situations considered in this study, the use of about 100 progenies in 60 repetitions resulted in a better ratio between the magnitude of genetic gains and the accuracy of the estimates.

Keywords: Quantitative genetics. Selection among and within. Genetic gain. Number of progenies. Single plant plot.



## LISTA DE FIGURAS

- Figura 1- Estimativa do ganho esperado com a seleção (%) entre (E) e dentro (D) de progênies em experimentos com múltiplas plantas por parcela, considerando-se as intensidades de seleção de 10% E e 20% D (linha vermelha) e 20% E e 10% D (linha azul), com  $k = 10$  plantas/parcela, frequência alélica  $p = 0,2$ ,  $gmd = 0$  e herdabilidades  $h^2 = 0,5$  ( $\blacktriangle$ ),  $h^2 = 0,25$  ( $\bullet$ ) e  $h^2 = 0,1$  ( $\blacksquare$ ). A) Com sementes remanescentes (SR); B) sem SR . . . . . 36
- Figura 2- Estimativa do ganho esperado com a seleção (%) entre (E) e dentro (D) de progênies em experimentos com múltiplas plantas por parcela, considerando-se as intensidades de seleção de 10% E e 20% D,  $gmd = 0$  e  $h^2 = 0,25$ , com o uso de sementes remanescentes (SR). A barra vertical representa o intervalo de confiança de 95%. . . . . 37
- Figura 3- Estimativa do ganho esperado com a seleção (%) entre (E) e dentro (D) de progênies, considerando-se as intensidades de seleção de 10% E e 20% D, com  $h^2 = 0,1$  (linha vermelha),  $h^2 = 0,25$  (linha azul) e  $h^2 = 0,5$  (linha verde) e grau médio de dominância  $gmd = 0$  ( $\blacksquare$ ),  $gmd = 1$  ( $\bullet$ ), com a utilização de sementes remanescentes em parcelas com 5 plantas. . . . . 38

Figura 4-	Ganho esperado com a seleção (%) em função da alteração em $r$ e $k$ quando se utiliza $N_f = 300$ e 10% de intensidade de seleção entre progênies em experimentos com múltiplas plantas por parcela. A barra vertical representa o intervalo de confiança de 95%. . . . .	39
Figura 5-	Estimativa do ganho esperado com a seleção (%) entre (GE), dentro (GD) e ganho total (GT) em experimentos com múltiplas plantas por parcela, considerando-se as intensidades de seleção de 10% E e 20% D, com $gmd = 0$ , $p = 0,2$ , $k$ plantas/parcela ( $k_1 = 5$ , $k_2 = 10$ , $k_3 = 20$ ) e $r$ repetições, com a utilização de 50 sementes remanescentes, selecionando-se número de plantas igual à $N_f$ . A barra vertical representa o intervalo de confiança de 95%. . . . .	40
Figura 6-	Ganho esperado com a seleção entre (GE), com a seleção dentro (GD) e ganho total (GT) em experimentos com múltiplas plantas por parcela, com $gmd = 0$ , $p = 0,2$ , $k$ plantas/parcela ( $k_1 = 5$ , $k_2 = 10$ , $k_3 = 20$ ) e $r$ repetições, obtido com as intensidades de seleção de 10% entre e 20% dentro de progênies, em que o número de sementes remanescentes variou em função do número de progênies selecionadas, sendo selecionadas 120 plantas. A barra vertical representa o intervalo de confiança de 95%. . . . .	42

- Figura 7- Ganho esperado com a seleção entre (GE), com a seleção dentro (GD) e ganho total (GT) em experimentos com múltiplas plantas por parcela, com  $gmd = 0$ ,  $p = 0,2$ ,  $k$  plantas/parcela ( $k1 = 5$ ,  $k2 = 10$ ,  $k3 = 20$ ) e  $r$  repetições, obtido com as intensidades de seleção de 10% entre e intensidades de seleção dentro variando em função do número de progênies selecionadas, totalizando 120 plantas selecionadas. A barra vertical representa o intervalo de confiança de 95%. . . . . 43
- Figura 8- Estimativas do ganho esperado e do ganho realizado com a seleção (%) em experimentos com múltiplas plantas por parcela, considerando-se as intensidades de seleção de 10% entre e 20% dentro de progênies,  $gmd = 0$ ,  $p = 0,2$ ,  $k$  plantas/parcela ( $k1 = 5$ ,  $k2 = 10$ ,  $k3 = 20$ ) e  $r$  repetições, com a utilização de sementes remanescentes. A barra vertical representa o intervalo de confiança de 95%. . . . . 44
- Figura 9- Estimativa do aumento em frequência alélica ( $p$ ) com a seleção entre (E) e dentro (D) de progênies em experimentos com múltiplas plantas por parcela, com as intensidades de seleção de 10% E e 20% D, considerando-se  $gmd = 0$ ,  $k$  plantas/parcela ( $k1 = 5$ ,  $k2 = 10$ ,  $k3 = 20$ ) e  $r$  repetições, com a utilização de sementes remanescentes. A barra vertical representa o intervalo de confiança de 95%. . . . . 45

- Figura 10- Ganho esperado com a seleção entre (GE), com a seleção dentro (GD) e ganho total (GT) em experimentos com uma plantas por parcela, obtido com as intensidades de seleção de 10% entre e 20% dentro das progênieis, com  $gmd = 0$  e  $p = 0,2$ . A barra vertical representa o intervalo de confiança de 95%. . . . . 47
- Figura 11- Razão entre os ganhos médios esperados com a seleção (GES) e a amplitude do intervalo de confiança (IC) de 95% dos ganhos, em experimentos com uma plantas por parcela, com frequências alélicas 0,2 (●), 0,5 (▲) e 0,8 (■). . . . . 48
- Figura 12- Estimativa de ganho esperado com a seleção entre, considerando-se a intensidade de seleção de 10%, com  $N_f = 300$  progênieis, avaliadas em 2, 4 e 20 repetições ( $r$ ) e com 10, 5 e 1 planta por parcela ( $k$ ). A barra vertical representa o intervalo de confiança 95%. . . . . 48

## LISTA DE TABELAS

Tabela 1-	Estimativa do componente da variância de progênie ( $\hat{\sigma}_f^2$ ) e do intervalo de confiança de 95% ( $IC_{95}$ ), com $h^2 = 0,1$ e $gmd = 0$ , em diferentes frequências alélicas ( $p$ ), número de repetições ( $r$ ), número de plantas por parcela ( $k$ ) e número de progênies ( $N_f$ ). . . . .	41
Tabela 2-	Estimativa do componente da variância de progênie ( $\hat{\sigma}_f^2$ ) e do intervalo de confiança de 95% ( $IC_{95}$ ), com $h^2 = 0,1$ e $gmd = 0$ , em diferentes frequências alélicas ( $p$ ), números de repetições ( $r$ ) e número de progênies ( $N_f$ ), em experimentos com uma planta por parcela. . . . .	46

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>14</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>16</b>
<b>2.1</b>	<b>Objetivos dos programas de melhoramento de plantas alógamas</b>	<b>16</b>
<b>2.2</b>	<b>Seleção em progênes de meios-irmãos</b>	<b>18</b>
<b>2.3</b>	<b>Uso da simulação computacional no melhoramento</b>	<b>22</b>
<b>3</b>	<b>MATERIAL E MÉTODOS</b>	<b>26</b>
<b>3.1</b>	<b>Obtenção da população-base</b>	<b>26</b>
<b>3.2</b>	<b>Obtenção das progênes de meios-irmãos</b>	<b>27</b>
<b>3.3</b>	<b>Avaliação das progênes de meios-irmãos</b>	<b>28</b>
<b>3.4</b>	<b>Análise dos experimentos com múltiplas plantas por parcela</b>	<b>30</b>
<b>3.5</b>	<b>Análise dos experimentos com uma planta por parcela</b>	<b>31</b>
<b>3.6</b>	<b>Estimativas de ganho com a seleção e recombinação</b>	<b>32</b>
<b>4</b>	<b>RESULTADOS</b>	<b>35</b>
<b>4.1</b>	<b>Experimentos com múltiplas plantas por parcela</b>	<b>35</b>
<b>4.2</b>	<b>Experimentos com uma planta por parcela</b>	<b>43</b>
<b>5</b>	<b>DISCUSSÃO</b>	<b>49</b>
<b>6</b>	<b>CONCLUSÃO</b>	<b>58</b>
	<b>REFERÊNCIAS</b>	<b>59</b>
	<b>APÊNDICE</b>	<b>66</b>

## 1 INTRODUÇÃO

Em todo programa de melhoramento de populações de plantas, o foco principal é o ganho genético com a seleção, ou seja, o aumento da frequência dos alelos favoráveis para o caráter sob seleção (BIJMA, 2012; HALLAUER; CARENA, 2012). Esse objetivo pode ser alcançado com a seleção de indivíduos ou progênies. Quando se opta pela utilização de progênies, essas podem ser meios-irmãos (MI), irmãos germanos (IG) ou progênies endogâmicas ( $S_1$  e  $S_2$ ) (HALLAUER; CARENA; MIRANDA FILHO, 2010; WEYHRICH; LAMKEY; HALLAUER, 1998). Em muitas situações, as progênies de MI são preferidas, especialmente, pela facilidade de obtenção e condução (BRUMMER; CASLER, 2010).

O uso de progênies de meios-irmãos no melhoramento se iniciou com o método espiga por fileira (LONNQUIST, 1964). Apesar de ter mais de cinco décadas, esse método é amplamente utilizado em muitas espécies cultivadas (BRUMMER; CASLER, 2010; PALOMINO; RAMALHO; FERREIRA, 2000). Contudo, durante esses mais de cinquenta anos, várias alternativas foram propostas e utilizadas. No que se refere ao número de progênies, de repetições, de tamanho de parcela e intensidade de seleção entre e dentro. São comuns na literatura relatos de experimentos envolvendo, normalmente uma ou duas dessas variáveis, como, por exemplo, o número de repetições e tamanho de parcela (COTTERILL; JAMES, 1984; PERRON; DEBLOIS; DESPONTS, 2013; RESENDE; SOUZA JÚNIOR, 1997). A dúvida que se tem é com relação à melhor opção, quando todas essas variáveis estão envolvidas.

Como esse tipo de informação é de difícil obtenção, sob condições experimentais de campo, a opção que resta é por meio da simulação. O emprego da simulação tem se ampliado com as facilidades computacionais e já vem sendo uti-

lizado há alguns anos e com várias finalidades em termos de melhoramento (BERNARDO, 2003; FERREIRA et al., 2000; FOUILLOUX; BANNEROT, 1988; LEITE et al., 2009; SILVA, 2009). Contudo, não foram encontrados relatos do emprego de simulação visando à identificação das melhores estratégias de seleção utilizando progênies de MI, quando se tem diferentes condições de controle genético do caráter, de frequência alélica, de herdabilidade, bem como variáveis que estão sob o controle do melhorista, tais como número de progênies, de repetições e de plantas/parcela.

É evidente que cada cultura tem suas particularidades, além do mais, diferentes estratégias podem ser mais efetivas em diferentes situações (BRUMMER; CASLER, 2010). Nesse contexto, realizou-se este trabalho com o objetivo de avaliar, por meio de simulação computacional, estratégias que auxiliem a tomada de decisão com relação aos números de progênies, repetições e plantas por parcela que aumentem a eficiência de programas de melhoramento que utilizam progênies de meios-irmãos.



## **2 REFERENCIAL TEÓRICO**

### **2.1 Objetivos dos programas de melhoramento de plantas alógamas**

Os métodos de melhoramentos de plantas alógamas basicamente se concentram em obter populações, híbridos de linhagem ou clones, para o caso de espécies de propagação assexuada. Entretanto, a maioria das espécies alógamas apresenta elevada depressão por endogamia, o que, muitas vezes, limita a produção de híbridos simples, em razão da baixa produtividade das linhagens endogâmicas (SOUZA JÚNIOR, 2001). Uma forma de contornar essa dificuldade é aumentar a frequência dos alelos favoráveis e conseqüentemente a produtividade, o que pode ser conseguido utilizando-se de diferentes métodos de condução de populações segregantes (VENCOVSKY, 1978).

Quando o objetivo é a obtenção de híbridos de linhagens, pode-se citar a utilização dos métodos da descendência de uma única semente, o método genealógico, a seleção gamética e, mais recentemente a obtenção de duplo haplóide (HALLAUER; CARENA; MIRANDA FILHO, 2010). Outro procedimento muito comum no desenvolvimento de cultivares é a utilização de reciclagem de linhagens elite (MIKEL, 2006). Esse método consiste em cruzar linhagens elites e utilizar as populações resultantes desses cruzamentos para a extração de linhagens, que produzirão os novos híbridos. Nesse processo, muitas vezes, a variabilidade genética é reduzida, isso com o decorrer dos anos, aumenta a vulnerabilidade desses genótipos diante de novas pragas e doenças (SOUZA JÚNIOR, 2001).

Os programas de melhoramento que visam a obtenção de híbridos, em muitos casos, não dão atenção ao melhoramento de populações, o que pode não ser uma estratégia adequada. Considerando o caráter teor de óleo no milho, por exemplo, que é controlado por 69 genes (BERNARDO, 2002), a probabilidade de

obtenção de uma linhagem com todos os alelos favoráveis em uma única geração é dada por  $P = (1/4)^{69}$ , ou seja, um em cada  $3,48 \times 10^{41}$  indivíduos terá o genótipo desejado. O Brasil cultiva anualmente quase 60 milhões de hectares com grãos, assim, considerando uma densidade de 70 mil plantas/ha, seria necessário aproximadamente  $8,3 \times 10^{28}$  vezes a área que é cultivada a cada ano, para se obter uma planta com os 69 alelos favoráveis para o teor de óleo. Observa-se que, à medida que o número de genes aumenta, a probabilidade de encontrar o genótipo desejado reduz mais ainda (RAMALHO; ABREU; SANTOS, 2001). O melhoramento, portanto, deve ser conduzido de maneira que se aumente, progressivamente, a frequência média dos alelos favoráveis na população (VENCOVSKY, 1978). Uma vez aumentada essa frequência, maior será a probabilidade de obtenção de cultivares superiores.

Em muitos casos, o melhoramento de linhagens e o melhoramento de populações são tratados como concorrentes, houve inclusive tentativas de comparação entre métodos de melhoramento populacional como a seleção recorrente com o método genealógico. No entanto, o desenvolvimento de linhagens elite depende do melhoramento das fontes de germoplasma, sendo que, ambos precisam ser complementares e só terão sucesso se receberem a mesma atenção (HALLAUER; CARENA; MIRANDA FILHO, 2010; HALLAUER; CARENA, 2012). Embora muitas abordagens sejam utilizadas no melhoramento populacional, frequentemente, as que utilizam progênies de meios irmãos são preferidas, especialmente pela facilidade de condução, seja em culturas anuais como o milho (HALLAUER; CARENA; MIRANDA FILHO, 2010; PALOMINO; RAMALHO; FERREIRA, 2000) ou em culturas perenes como o eucalipto (CASTRO, 1992).

## 2.2 Seleção em progênies de meios-irmãos

A utilização de progênies de meios-irmãos de forma sistemática começou com o método espiga por fileira, com o programa de seleção para alto e baixo teor de óleo e proteína em Illinois, em 1896 (PATERNIANI; FILHO, 1978). O procedimento consistia em plantar espigas de milho selecionadas em fileiras separadas, e avaliar as progênies de cada espiga em linhas separadas, utilizando-se, posteriormente, para recombinação as sementes das linhas selecionadas. Esse método é a forma mais simples de esquema de seleção utilizando progênies de meios-irmãos, aplicável a espécies alógamas, sendo eficiente para caracteres com alta herdabilidade, como é o caso de óleo e proteína no milho (ACQUAAH, 2007). Por não ser tão eficiente para produtividade, não tem sido muito utilizado (HALLAUER; CARENA; MIRANDA FILHO, 2010).

Com o intuito de melhorar a eficiência do método supracitado, Lonquist (1964) e Compton e Comstock (1976) propuseram modificações. O primeiro, em essência, sugeriu avaliar e selecionar progênies de meios irmãos e, posteriormente, selecionar as melhores plantas dentro de cada progênie. Este método ficou conhecido como seleção entre e dentro de progênies de meios-irmãos (PATERNIANI, 1967). Nesse caso, são obtidas espigas (para o caso do milho) de polinização livre da população a ser melhorada, de forma que cada espiga constituirá uma progênie de meio irmãos. As sementes de cada espiga são plantadas em experimentos com repetições, para a seleção entre progênies. Após a seleção entre, colhem-se as sementes dos melhores indivíduos dentro das progênies selecionadas.

Segundo a proposta de Compton e Comstock (1976), podem ser utilizadas sementes remanescentes para serem recombinadas em outra safra. Por ocasião da colheita, são selecionadas as melhores plantas, constituindo a etapa de seleção dentro das progênies (HALLAUER; CARENA; MIRANDA FILHO, 2010; PAN-

DEY; GARDNER, 1992; PATERNIANI; FILHO, 1978). O método de seleção entre e dentro de meios-irmãos tem sido utilizado com sucesso no Brasil, tendo sua eficiência comprovada por diversos autores quanto às magnitudes de parâmetros genéticos, tais como a variância genética aditiva e os ganhos esperados com a seleção para produtividade (CARVALHO et al., 2000; MARTINS et al., 2005; MATTA; VIANNA, 2003).

Embora seja utilizado com sucesso, alguns estudos têm mostrado que é possível melhorar sua eficiência (GRANATE; CRUZ; PACHECO, 2001; RAMALHO, 1977). Entre os fatores que aumentam a eficiência do método, podem ser citadas a utilização de sementes remanescentes para seleção em ambos os sexos, como comentado, a utilização de mais de uma safra por ano, para recombinação das progênies selecionadas e principalmente, o aumento da intensidade de seleção entre e dentro das progênies. A importância da intensidade de seleção pode ser visualizada na expressão do ganho com a seleção (ACQUAAH, 2007):

$$GES = \frac{c_1 i_1 \sigma_A^2}{\sigma_{MI}} + \frac{c_2 i_2 \sigma_A^2}{\sigma_{we}}, \quad (1)$$

em que  $\sigma_A^2$  é a variância genética aditiva,  $\sigma_{MI}$  é o desvio padrão fenotípico entre as progênies,  $\sigma_{we}$  é a raiz quadrada da variância entre plantas dentro das parcelas,  $i_1$  é a intensidade de seleção entre progênies,  $i_2$  é a intensidade de seleção dentro das progênies selecionadas e  $c_1$  e  $c_2$  são as porções da variância aditiva explorada, entre e dentro, que podem ser respectivamente, 1/4 e 3/4 para seleção em ambos os sexos ou 1/8 e 3/8 para seleção apenas do lado materno (VENCOVSKY; BARRIGA, 1992).

Pela expressão apresentada, observa-se que a intensidade de seleção multiplica a variância genética aditiva tanto entre quanto dentro das progênies, de modo que uma intensidade de seleção alta e uma intensidade baixa podem apre-

sentar ganhos genéticos bem distintos. Outro detalhe importante a ser observado é que, grande parte da variância genética aditiva é explorada dentro das progênes de meios irmãos. Essa quantidade de variância genética presente justifica um cuidado maior com o número de plantas a serem selecionadas em cada progênie.

Em geral, quando é feita a seleção entre progênes, a proporção de indivíduos selecionados, normalmente varia entre 10 e 20% (ARAÚJO et al., 2005; CARVALHO et al., 2000; CARVALHO; SANTOS; LEAL, 2005; PATERNIANI; FILHO, 1978). No que se refere à seleção dentro das progênes não há relatos de trabalhos feitos visando à identificação de um número ideal de plantas para representar a progênie selecionada. Em razão da falta dessa informação, diferentes valores tem sido utilizados na literatura, variando entre 2 e 100 plantas (RAMALHO, 1977). O que tem sido sugerido é a utilização de uma intensidade de seleção maior entre do que dentro, pois a seleção entre é feita com base na médias das progênes, e tem maior precisão do que as observações individuais que constituem as unidades de seleção dentro (RAMALHO, 1977).

Estimar o número ideal de indivíduos é de fundamental importância, pois uma vez amostrada corretamente, o aumento no tamanho da amostra não aumenta o ganho com a seleção de forma proporcional (SOUZA JÚNIOR, 2001). Em primeiro momento, pode-se pensar que quanto menos indivíduos selecionados melhor, pois maior seria a intensidade de seleção e, conseqüentemente, o ganho com a seleção. O problema de se utilizar um número muito pequeno é que a amostra poderia não representar bem a progênie selecionada. Nesse caso, as dificuldades em termos amostrais não são devidas à precisão estatística, mas à estrutura genética das populações. Se uma população possui ampla heterogeneidade genética e o tamanho adequado da amostra para essa população for utilizado em uma população mais homogênea, este será maior que o necessário. Porém, se a

população for geneticamente mais homogênea, o tamanho amostral adequado para essa população será insuficiente para representar populações mais heterogêneas. No primeiro caso, haveria desperdício de área, recursos humanos e econômicos, enquanto, no segundo caso, as conclusões seriam afetadas pela insuficiência do tamanho amostral (MARQUEZ-SANCHEZ, 1972). Considerados os tamanhos amostrais mínimos, pode-se aumentar a eficiência do processo seletivo, com uma combinação adequada de seleção entre e de seleção dentro das progênes selecionadas (RAMALHO, 1977).

Além dos fatores já comentados, uma maneira de se aumentar o ganho esperado com a seleção em progênes de meios-irmãos é diminuindo o desvio padrão fenotípico e a raiz quadrada da variância dentro das parcelas, como visto na expressão 1. Esse objetivo pode ser alcançado com o aumento tanto do número de repetições quanto do número de plantas por parcela (CASTRO, 1992; RESENDE, 1995). Alguns trabalhos têm sido feitos com o intuito de identificar tamanhos de parcela, número de repetições e quantidade de progênes a serem utilizadas na avaliação de meios-irmãos (COTTERILL; JAMES, 1984; PERRON; DEBLOIS; DESPONTS, 2013). Em cenoura, por exemplo, utilizando a técnica de Bootstrap, verificou-se que o número de progênes deve ser no mínimo 74, enquanto a utilização de 18 plantas/parcela em duas repetições é suficiente para avaliar a estrutura genética da população (SILVA; VIEIRA; VILLELA, 2009).

Em eucalipto, por meio de análises estatísticas utilizando dados de múltiplos experimentos, Castro (1992) observou que não houve grande vantagem em se utilizar mais de três repetições, quando se utilizam tamanhos de parcela variando entre 3 e 24 plantas/parcela. Nesse caso, a utilização de uma maior quantidade de progênes foi mais vantajosa em relação à utilização de mais plantas por parcela. Em milho testou-se tamanhos entre 5 e 50 plantas por parcela,

avaliando-se 200 progênies de meios irmãos em condições específicas de cerrado e solo fértil. Verificou-se que para solo fértil o ideal seria utilizar 15 plantas por parcela, enquanto para cerrado, o indicado seriam 20 plantas (RESENDE; SOUZA JÚNIOR, 1997). Embora valores entre 15 e 20 possam ser sugeridos, (CHAVES; MIRANDA FILHO, 1992), tamanhos menores que 10 e maiores que 25 plantas por parcela também já foram recomendados (CARGNELUTTI FILHO et al., 2011; COMPTON; BAHADUR, 1977; MARTIN et al., 2005; PALOMINO; RAMALHO; FERREIRA, 2000).

Os trabalhos mencionados abordam fatores importantes para o aumento da eficiência na seleção com progênies de meios-irmãos. No entanto, por se tratarem de condições específicas, podem não ser aplicáveis a situações muito diferentes daquelas estudadas. Como comentado, o ganho esperado com a seleção é influenciado por um grande número de variáveis, de modo que, a realização de experimentos considerando todas essas variáveis pode não ser viável economicamente. Uma alternativa para esse tipo de situação seria a simulação computacional.

### **2.3 Uso da simulação computacional no melhoramento**

Os primeiros trabalhos com simulação surgiram com a utilização do método de Monte Carlo, por Von Neuman, em 1940, com a blindagem de reatores nucleares. Em genética e melhoramento mais especificamente, a utilização da simulação computacional começou com Fraser (1957) e, desde então, tem ocupado um espaço importante, principalmente na solidificação de conceitos de difícil verificação no contexto prático (FERREIRA, 2001).

As contribuições da simulação para a genética e melhoramento de plantas têm sido importantes tanto em áreas clássicas como a genética quantitativa (ABREU et al., 2010), quanto em áreas mais novas como a genômica (DE LOS

CAMPOS; SORENSEN; GIANOLA, 2015). Os modelos, nessas áreas, são inerentemente probabilísticos, por envolverem várias variáveis de natureza aleatória. Evidentemente, tais modelos precisam ser simplificados para que possam ser simulados por algoritmos eficientes e interpretados de forma adequada. A maior restrição se encontra na simplificação exagerada desses modelos, que nem sempre refletem situações reais. Por outro lado, é possível variar os parâmetros de tal forma que uma amplitude realística de fatores possa ser facilmente testada e, mais importante, o efeito relativo destes fatores sobre resultados da seleção pode ser estudado (FERREIRA, 2001).

O modelo ideal não existe, isso fica claro pela famosa expressão de Box e Draper (1987): “Essentially, all models are wrong, but some are useful”. O que se busca é algum modelo que, mesmo com todas as suas simplificações, represente mais fielmente a realidade. Em todos os casos, demanda dos geneticistas o desenvolvimento de modelos biológicos que retratem, da melhor forma possível, os fenômenos de interesse e, dos programadores, as rotinas para o processamento adequado, apesar da imposição de restrições, para que a influência de certos fatores possa ser avaliada (CRUZ, 2001).

Todo processo simulado que envolve um componente aleatório de qualquer distribuição é considerado como pertencente ao método Monte Carlo. A única restrição para o uso desse método é a sua impraticabilidade para distribuições cuja função distribuição seja desconhecida ou cuja inversa não seja possível pela não existência de algoritmos numéricos (MELO, 2010). Inúmeros programas de computador, linguagens de programação e compiladores têm rotinas prontas para a geração de números aleatórios. Esses números são gerados em uma sequência que só se repete após um ciclo superior à ordem dos trilhões de algoritmos. Dessa forma, os números gerados são pseudoaleatórios, mas, para fins práticos, uma vez



que a sequência completa (ciclo) jamais será utilizada, esses números comportam-se como genuinamente aleatórios (FERREIRA, 2001).

Como em todas as técnicas, cuidados devem ser tomados para evitar erros decorrentes de problemas amostrais, escolha inadequada das distribuições de probabilidade dos eventos de natureza aleatória, simplificação inadequada da realidade e erros de implementação do sistema simulado. Observando-se tais cuidados, a utilização dessa técnica no melhoramento é recomendada nos casos em que: a) as soluções analíticas não existem ou o grau de dificuldade e o número de variáveis envolvidas não permitem a realização de inferências adequadas sobre o problema; b) pretende-se comparar a eficiência de um novo procedimento ou técnica em relação a outros já existentes; c) deseja-se verificar a eficiência de métodos que são rotineiramente empregados em algumas etapas de um programa de melhoramento e que podem ser melhorados para propiciarem maior ganho com a seleção (FERREIRA, 2001). Nesses casos, a simulação, por tudo que foi comentado, deverá ser utilizada com frequência, podendo ser mais convincente que qualquer argumento teórico (BAKER, 1995).

Desde o trabalho de Fraser (1957) avaliando o efeito da ligação nas taxas de ganho com a seleção massal e nos avanços dos ganhos genéticos, muitos outros trabalhos têm dado contribuições importantes para o melhoramento. Por exemplo, na avaliação da seleção recorrente (VIANA, 2007), na determinação do número e do tamanho de famílias a serem avaliadas (BERNARDO, 2003; FERREIRA et al., 2000; FOUILLOUX; BANNEROT, 1988; LEITE et al., 2009; SILVA, 2009), na verificação da eficiência de dialelos circulantes na escolha de genitores (VEIGA; FERREIRA; RAMALHO, 2000), na comparação de estratégias de melhoramento (JONES; COMPTON; O.GARDNER, 1971; KERR; DIETERS; TIER, 2004; NUNES; RAMALHO; FERREIRA, 2008; WANG et al., 2003), e em vários outros

trabalhos estudando diferentes aspectos relacionados ao melhoramento genético de plantas (GERLOFF; SMITH, 1988; SILVA; FERREIRA; PACHECO, 2000; SIMMONDS; ROAD, 1991; SOUZA; PETERNELLI; BARBOSA, 2006).

O conhecimento de técnicas computacionais, claramente tem sido útil ao melhoramento de plantas. No entanto, ainda há muitas contribuições a serem feitas. Nesse aspecto, a identificação da melhor combinação entre o número de progênies e de plantas/progênie, bem como do número de repetições e de plantas por parcelas para avaliar progênies de meios irmãos, são exemplos de respostas que podem ser obtidas com o auxílio da simulação computacional.

### 3 MATERIAL E MÉTODOS

Nas simulações realizadas, considerou-se um caráter genérico de herança quantitativa, controlado por  $g = 100$  genes de efeitos iguais e independentes, semelhante ao adotado por Veiga, Ferreira e Ramalho (2000). A frequência média dos alelos favoráveis ( $p$ ), na população base, foi considerada 0,2, 0,5 e 0,8 na média dos  $g$  genes. Foram simulados os modelos aditivo ( $a = 1$  e  $d = 0$ ) e aditivo e dominante ( $a = d = 1$ ), sendo  $a$ , o desvio dos homozigotos em relação à média do valor genotípico para o loco em questão, e  $d$  o desvio dos heterozigotos (WRICKE; WEBER, 1986, p. 48). Interações epistáticas não foram consideradas. Foram realizadas 1.000 simulações Monte Carlo de cada cenário envolvendo as combinações dos fatores estudados. Em cada configuração, foram tomadas estatísticas descritivas relativas ao ganho com a seleção e à precisão experimental.

#### 3.1 Obtenção da população-base

Para cada indivíduo  $i$  da população base ( $i = 1, 2, \dots, N$ , sendo  $N = 10.000$ ), considerando-se que o mesmo possui  $g_1$  locos do tipo  $BB$ ,  $g_2$  do tipo  $Bb$  e  $g_3$  do tipo  $bb$ , sendo o valor genotípico ( $G_i$ ) desse indivíduo dado por:

$$G_i = g_1a + g_2d + g_3(-a), \quad (2)$$

em que  $g = g_1 + g_2 + g_3 = 100$ .

A frequência do alelo favorável na população base variou para cada gene função dos cenários e da distribuição beta. Os valores ao longo dos diferentes genes foram simulados de uma distribuição beta com parâmetros  $\alpha = 10p$  e  $\beta = 10(1 - p)$ . Assim,  $g$  realizações dessa distribuição com parâmetros  $\alpha = 2$  e  $\beta = 8$  para  $p = 0,2$ ,  $\alpha = \beta = 5$  para  $p = 0,5$  e por fim  $\alpha = 8$  e  $\beta = 2$  para  $p = 0,8$

foram obtidas, utilizando-se a função `rbeta` do programa R (R Core Team, 2014). Essas situações ( $p = 0,2$ ,  $p = 0,5$  e  $p = 0,8$ ) foram consideradas separadamente, simulando-se populações-base pouco melhoradas ( $p = 0,2$ ), populações-base com níveis intermediários de melhoramento ( $p = 0,5$ ) e populações-base bastante melhoradas ( $p = 0,8$ ).

Para constituir o genótipo do indivíduo  $i$  no gene  $n$ , foi obtida uma realização  $u$  da distribuição uniforme  $U(0, 1)$ . De modo que, a constituição genotípica do indivíduo  $i$  no gene  $n$  foi  $B_nB_n$  para valores de  $u$  entre 0 e  $p_n^2$ ,  $B_nb_n$  para valores de  $u$  entre  $p_n^2$  e  $p_n^2 + 2p_nq_n$  e  $b_nb_n$  para valores de  $u$  entre  $p_n^2 + 2p_nq_n$  e 1, considerando uma população em equilíbrio de Hardy-Weinberg. O processo foi repetido para os  $g$  genes dos  $N$  indivíduos até que o tamanho final da população base fosse atingido. De posse do genótipo do indivíduo  $i$ , o seu valor genotípico foi calculado utilizando a expressão 2.

### 3.2 Obtenção das progênes de meios-irmãos

As progênes de meios-irmãos foram obtidas de  $N_f$  indivíduos aleatoriamente retirados da população-base, em que  $N_f$  variou em função do método de avaliação das progênes. Cada genitor feminino foi fixado e seus gametas foram sorteados para a formação de cada indivíduo. Os gametas do genitor masculino foram escolhidos aleatoriamente da população. Para cada gene, foi utilizada amostragem com reposição para a formação de cada novo indivíduo. As progênes foram formadas pela junção dos gametas do genitor feminino e dos gametas dos genitores masculinos, nesse caso, todos os indivíduos da população. Se o genótipo do indivíduo no  $n$ -ésimo gene é  $BB$ , seus gametas são do tipo  $B$ , com 100% de probabilidade; se o genótipo do gene é  $bb$ , os gametas são somente do tipo  $b$ ; e se o genótipo é  $Bb$ , os gametas são  $B$  com probabilidade de 50% ou  $b$  também com

probabilidade de 50%. O genótipo resultante foi obtido a partir da junção desses dois gametas, sendo esse processo repetido para todos os genes.

### 3.3 Avaliação das progênies de meios-irmãos

Duas abordagens distintas foram consideradas: 1) experimentos com múltiplas plantas parcela; 2) experimentos com uma planta por parcela. Em ambos os casos considerou-se um programa de melhoramento com recursos para avaliar 6.000 plantas, de modo que o produto do número de progênies ( $N_f$ ), repetições ( $r$ ) e plantas por parcela ( $k$ ) fosse sempre 6.000 ( $N_f \times r \times k = 6.000$ ). Assim, em experimentos com múltiplas plantas por parcela, foram utilizadas 2, 3, 4 e 5 repetições e  $k$  sendo 5, 10 ou 20 plantas/parcela. Foram simuladas combinações de todos os níveis de  $r$  e  $k$ , de modo que o valor de  $N_f$  fosse  $6.000/(k \times r)$ . Logo,  $N_f$  variou entre 60 e 600. Para o caso de uma planta por parcela ( $k = 1$ ), foram utilizadas 10, 20, 40, 60, 80, 100, 150, 200 e 500 repetições, sendo o número de plantas por progênie ( $N_p$ ) igual ao número de repetições. Logo, foram geradas  $N_f$  progênies, com  $N_f$  variando entre 12 e 600.

Para o primeiro caso, os valores fenotípicos dos indivíduos foram obtidos de acordo com o seguinte modelo:

$$F_{ij} = \mu + G_i + e_i + e_{ij}, \quad (3)$$

em que  $\mu$  é a média geral considerada igual a 100, sem perda de generalidade;  $G_i$  é o valor genotípico do indivíduo  $i$  oriundo da expressão 2;  $e_i$  é o erro associado ao  $i$ -ésimo indivíduo, sendo  $e_i \sim NID(0, \sigma_{e_1}^2)$ ;  $e_{ij}$  é o erro associado a  $j$ -ésima parcela contendo o  $i$ -ésimo indivíduo com  $e_{ij} \sim NID(0, \sigma_{e_2}^2)$ .

Para experimentos com uma planta por parcela, os valores fenotípicos dos

indivíduos foram obtidos da seguinte forma:

$$F_i = \mu + G_i + \varepsilon_i, \quad (4)$$

em que  $\varepsilon_i$  é o erro associado ao  $i$ -ésimo indivíduo com  $\varepsilon_i \sim NID(0, \sigma_\varepsilon^2)$ . Para a geração de erros individuais e erros por parcela, considerou-se  $\sigma_{e_1}^2 = 5\sigma_{e_2}^2$  semelhante ao encontrado na literatura (CARDOSO et al., 2003; RAMALHO, 1977). Por serem gerados de forma aleatória e independente, determinou-se que  $\sigma_{e_1}^2 + \sigma_{e_2}^2 = \sigma_\varepsilon^2$ , de modo que ambas as quantidades pudessem ser fixadas em função da herdabilidade.

A magnitude de  $\sigma_\varepsilon^2$  foi definida em função da herdabilidade em nível de indivíduo ( $h^2$ ) que foi 0,1, 0,25 ou 0,50. Para cada situação considerada, o valor de  $\sigma_\varepsilon^2$  foi obtido da relação:

$$h^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_\varepsilon^2}, \quad (5)$$

em que  $\sigma_G^2$  é a variância genética da população base calculada pela seguinte expressão (BERNARDO, 2002, p. 92):

$$\sigma_G^2 = \sum_{f=1}^g 2p_f(1-p_f)[a + d(1-2p_f)]^2 + [2p_f(1-p_f)d]^2. \quad (6)$$

Assim, a variância ambiental é dada por:

$$\sigma_\varepsilon^2 = \frac{\sigma_G^2(1-h^2)}{h^2}. \quad (7)$$

As realizações dos erros normais  $e_i$ ,  $e_{ij}$  e  $\varepsilon_i$  foram simuladas, utilizando-se a função `rnorm` do programa R (R Core Team, 2014).

### 3.4 Análise dos experimentos com múltiplas plantas por parcela

Após a obtenção dos dados, foram realizadas análises de variância de cada um dos experimentos segundo o delineamento utilizado. Para experimentos com múltiplas plantas por parcela utilizou-se o seguinte modelo:

$$Y_{ij\ell} = \mu + f_i + b_j + e_{ij} + d_{(ij)\ell} \quad (8)$$

em que:

$Y_{ij\ell}$  é o valor observado referente à planta  $\ell$  da progênie  $i$  no bloco  $j$ ;

$\mu$  é a constante inerente a cada observação;

$f_i$  é o efeito da progênie  $i$ , ( $i = 1, 2, 3, \dots, Nf$ ),  $f_i \sim NID(0, \sigma_f^2)$ ;

$b_j$  é o efeito do bloco  $j$ , ( $j = 1, 2, 3, \dots, Np$ ),  $b_i \sim NID(0, \sigma_b^2)$ ;

$e_{ij}$  é o efeito do erro experimental associado à observação da parcela  $ij$ ,

$e_{ij} \sim NID(0, \sigma_e^2)$ ;

$d_{(ij)\ell}$  é o efeito da variação entre plantas ( $\ell = 1, 2, 3, \dots, k$ ),  $d_{(ij)\ell} \sim NID(0, \sigma_d^2)$ .

Os componentes de variância foram obtidos segundo as esperanças dos quadrados médios, como apresentando por NGUYEN e SLEPER (1983). A seguir obtiveram-se diferentes estimativas de herdabilidades. A herdabilidade em nível de indivíduo foi estimada pela seguinte relação:

$$\hat{h}_i^2 = \frac{4\hat{\sigma}_f^2}{\hat{\sigma}_f^2 + \hat{\sigma}_e^2 + \hat{\sigma}_d^2} \quad (9)$$

A herdabilidade para a seleção na média de progênes foi obtida da seguinte forma:

$$\hat{h}_m^2 = \frac{\hat{\sigma}_f^2}{\hat{\sigma}_f^2 + \frac{\hat{\sigma}_e^2}{r} + \frac{\hat{\sigma}_d^2}{kr}} \quad (10)$$

Para a seleção dentro das progênes e posterior recombinação, com seleção

apenas do lado materno, a herdabilidade considerada foi:

$$\hat{h}_d^2 = \frac{\frac{1}{2}(3\hat{\sigma}_f^2)}{\hat{\sigma}_d^2}. \quad (11)$$

### 3.5 Análise dos experimentos com uma planta por parcela

Em experimentos de uma planta por parcela utilizou-se o seguinte modelo (VENCOVSKY; BARRIGA, 1992, p. 98):

$$Y_{ij} = \mu + f_i + b_j + e_{ij}, \quad (12)$$

em que:

$Y_{ij}$  é o valor observado referente à progênie  $i$  no bloco  $j$ ;

$\mu$  é o constante inerente a cada observação;

$f_i$  é o efeito da progênie  $i$ , ( $i = 1, 2, 3, \dots, N_f$ ),  $f_i \sim NID(0, \sigma_f^2)$ ;

$b_j$  é o efeito do bloco  $j$ , ( $j = 1, 2, 3, \dots, N_p$ ),  $b_i \sim NID(0, \sigma_b^2)$ ;

$e_{ij}$  é o efeito do erro experimental associado à observação da parcela  $ij$ ,

$e_{ij} \sim NID(0, \sigma_e^2)$ .

A partir das estimativas dos componentes de variâncias obtidos, estimou-se a herdabilidade em nível de indivíduo pela relação:

$$\hat{h}_i^2 = \frac{4\hat{\sigma}_f^2}{\hat{\sigma}_f^2 + \hat{\sigma}_e^2}. \quad (13)$$

Em seguida, obteve-se a herdabilidade para a seleção em média de progênies:

$$\hat{h}_m^2 = \frac{\hat{\sigma}_f^2}{\hat{\sigma}_f^2 + \frac{\hat{\sigma}_e^2}{r}}. \quad (14)$$



A herdabilidade para seleção dentro das progênie, considerando-se a possibilidade de recombinação das plantas selecionadas e a seleção em ambos os sexos, foi derivada de Vencovsky e Barriga (1992, p. 98), sendo estimada da seguinte forma:

$$\hat{h}_d^2 = \frac{3\hat{\sigma}_f^2}{\hat{\sigma}_{ps}^2}, \quad (15)$$

nesse caso, como a análise é realizada com totais de parcelas, utilizou-se como denominador da herdabilidade  $\hat{\sigma}_{ps}^2$ , que é a variância fenotípica entre plantas das progênie selecionadas.

### 3.6 Estimativas de ganho com a seleção e recombinação

Inicialmente, foram consideradas duas combinações de intensidade de seleção entre e dentro independente da estratégia de avaliação utilizada. Na primeira combinação, procedeu-se à seleção de 10% das melhores progênie e destas foram selecionados os 20% melhores indivíduos. Na segunda combinação, selecionou-se as 20% melhores progênie e destas os 10% melhores indivíduos, de modo que a intensidade de seleção totalizasse 2% nos dois casos.

O ganho esperado com a seleção entre progênie foi estimado considerando-se a herdabilidade em nível de média de progênie (Eq. 14 e 10) da seguinte forma:

$$GS_e = \hat{h}_m^2 (M'_F - M_F), \quad (16)$$

em que  $M'_F$  é a média fenotípica das progênie selecionadas e  $M_F$  é a média fenotípica da população inicial. Em experimentos com múltiplas plantas por parcela, considerou-se a recombinação dos indivíduos selecionados dentro das melhores

progênies com e sem a adoção de sementes remanescentes (SR). Em caso de uso de SR, o valor de  $GS_e$  é considerado integralmente, caso contrário, o valor de  $GS_e$  é dividido pela metade no ganho genético total, visto que sem o uso de SR a seleção das melhores progênies ocorre em apenas um sexo.

Quando se utilizou SR, considerou-se 50 SR disponíveis, com 5 (10%) e 10 (20%) indivíduos selecionados independente do número de progênies avaliadas. Ao final, o número total de indivíduos selecionados foi igual a  $N_f$ . Utilizou-se também, outra alternativa para estimar o ganho genético total, isto é, mantendo-se constante o número de plantas na população melhorada, e não de progênies como anteriormente. Assim, o número de sementes remanescentes foi igual ao número de plantas por progênie, ou seja,  $N_p$  sementes remanescentes. Logo, para o caso de intensidade de seleção dentro de 20%, com  $N_f = 600$ ,  $N_p = kr = 10$ , foram selecionadas duas plantas. Já com  $N_f = 60$ ,  $N_p = 100$  foram selecionadas 20 plantas. Veja que o número de plantas, ao final do processo foi sempre 120. Para estimar o progresso realizado com a seleção, considerou-se que cada planta produz 50 indivíduos, mantendo-se, novamente, o mesmo número de plantas da população original, 6.000.

Na terceira opção, diferente dos casos comentados até agora, a intensidade de seleção total não foi constante. Isso se deu pois, o número de indivíduos selecionados variou em função de  $N_f$ , já o número de sementes remanescentes foi sempre o mesmo, 50. Como no último caso, o número de indivíduos selecionados foi 120. Assim, para a intensidade de seleção entre de 10%, com  $N_f = 600$ , foram selecionadas 2 plantas dentro do lote de recombinação, intensidade de seleção de  $2 \text{ em } 50 = 0,04$  ou 4%. Para  $N_f = 60$ , foram selecionadas 20 plantas dentro, ou seja,  $20 \text{ em } 50 = 0,4$  ou 40%.

O ganho esperado com a seleção dentro das progênies, por sua vez, foi

estimado considerando  $\hat{h}_d^2$  (Eq. 15 e 11) sendo dado por:

$$GS_d = \hat{h}_d^2(M'_{Fs} - M_F), \quad (17)$$

em que  $M'_{Fs}$  é a média fenotípica das melhores plantas dentro das progênies selecionadas. Como se observa pelos estimadores da herdabilidade para a seleção dentro das melhores progênies, em experimentos com uma planta por parcela, a recombinação ocorreu entre as próprias plantas selecionadas com seleção em ambos os sexos. No caso de múltiplas plantas por parcela, considerou-se a seleção dentro apenas do lado materno.

O ganho genético esperado com a seleção ( $GES$ ), em percentual da média original, foi estimado do seguinte modo:

$$GES = \frac{GS_e + GS_d}{M_F} \times 100. \quad (18)$$

Após a recombinação, o ganho realizado foi calculado comparando-se a média da população melhorada com a média original, da seguinte forma:

$$GR = \frac{(M_M - M_F)}{M_F} \times 100 \quad (19)$$

em que  $M_M$  é a média fenotípica da população melhorada.

A alteração na frequência alélica após a seleção foi calculada da seguinte forma:

$$\Delta p = \frac{(p_m - p)}{p} \times 100, \quad (20)$$

em que  $p_m$  é a frequência alélica média da população melhorada.

Após a obtenção das 1.000 estimativas de cada parâmetro considerado, obteve-se os quantis 2,5% e 97,5% da distribuição Monte Carlo, para a obtenção

dos intervalos de confiança. As simulações e análises dos dados foram feitas no software R 3.2.2 (R Core Team, 2014).

## 4 RESULTADOS

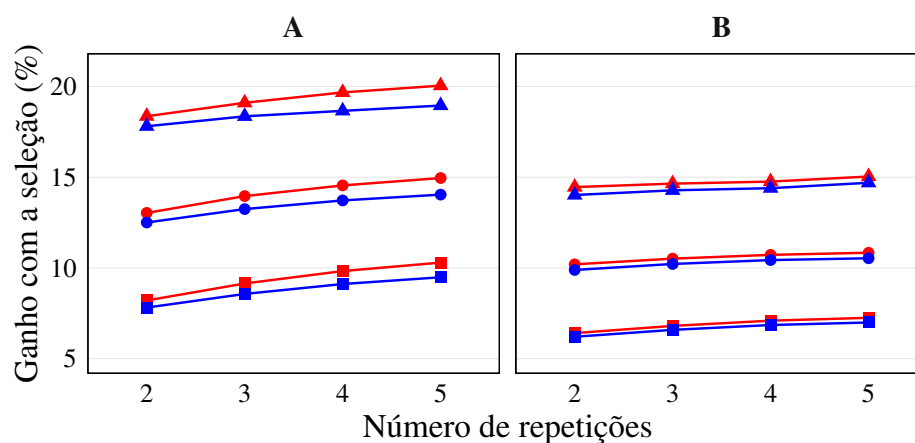
A validade das simulações pode ser constatada pela razão entre a variância aditiva e a variância das progênies de MI de aproximadamente 4 (FIGURAS A.1 e A.2, APÊNDICE). Da mesma forma, verifica-se que a estimativa da herdabilidade em nível de indivíduo no sentido restrito ( $\hat{h}_i^2$ ), na média das 1.000 simulações, aproximou-se do valor pré-definido da herdabilidade em nível de indivíduo ( $h^2$ ) na ausência de dominância e foi inferior a este no modelo com dominância completa (FIGURAS A.3 e A.4, APÊNDICE).

### 4.1 Experimentos com múltiplas plantas por parcela

Inicialmente deve ser salientada a questão da intensidade de seleção entre progênies (E), isto é, 10 ou 20%, independente das alternativas de seleção dentro. Na Figura 1, estão apresentados os resultados considerando  $k$  (10 plantas por parcela) e grau médio de dominância igual a zero. As demais situações não foram incluídas, pois os resultados foram bastante coincidentes. Quando se comparou o efeito da proporção de seleção entre (E) e dentro (D) de progênies, isto é, 10% E ou 20% E, verifica-se que, em todas as situações, a primeira opção propiciou maiores estimativas de ganho. Contudo, é preciso dizer que as diferenças foram muito pequenas sem sementes remanescentes, e mais expressivas quando estas foram utilizadas. Observa-se também que, como era esperado, o ganho foi sempre superior com o incremento da herdabilidade, porém, a tendência não se alterou com uma menor intensidade de seleção entre. Do exposto, todos os co-

mentários serão focados, considerando-se apenas a intensidade de seleção de 10% entre progênies. Em todos os casos, o *GES* foi função da frequência alélica, isto

Figura 1 – Estimativa do ganho esperado com a seleção (%) entre (E) e dentro (D) de progênies em experimentos com múltiplas plantas por parcela, considerando-se as intensidades de seleção de 10% E e 20% D (linha vermelha) e 20% E e 10% D (linha azul), com  $k = 10$  plantas/parcela, frequência alélica  $p = 0,2$ ,  $gmd = 0$  e herdabilidades  $h^2 = 0,5$  ( $\blacktriangle$ ),  $h^2 = 0,25$  ( $\bullet$ ) e  $h^2 = 0,1$  ( $\blacksquare$ ). A) Com sementes remanescentes (SR); B) sem SR

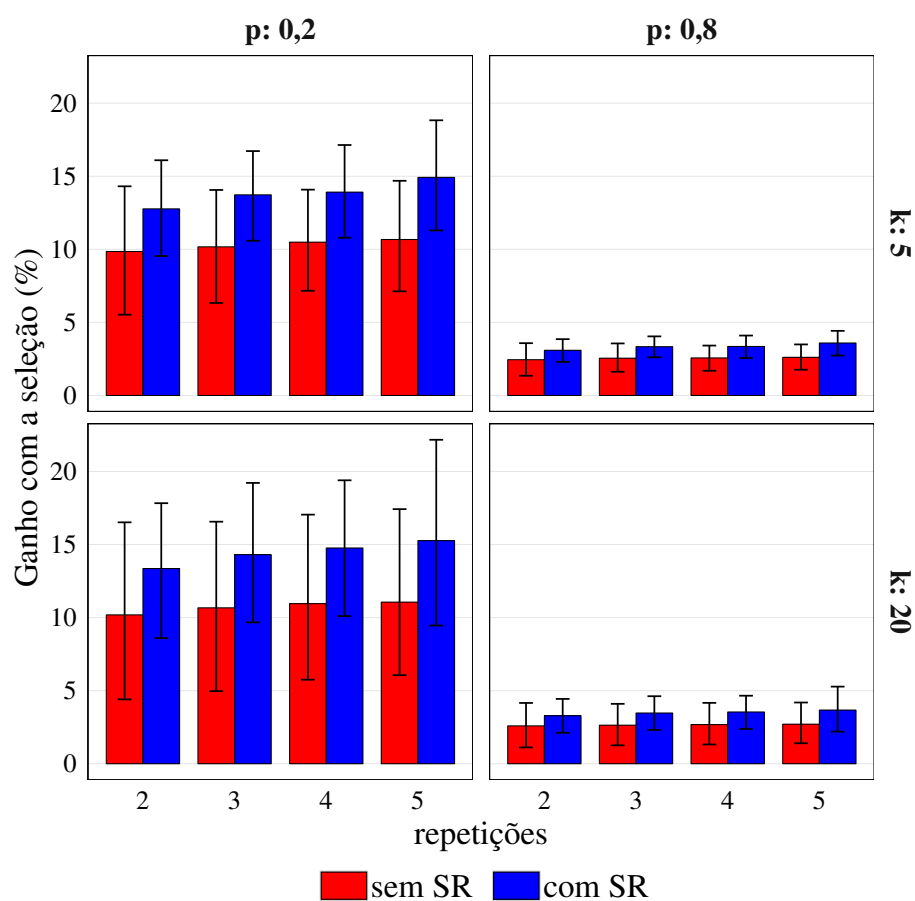


Fonte: Do autor (2016).

é, maior para população pouco melhorada ( $p = 0,2$ ) em relação à mais melhorada ( $p = 0,8$ ) (FIGURAS 2 e 3). Como quando  $p = 0,5$  os valores de *GES* foram intermediários, a maior atenção será dada a  $p = 0,2$  ou  $p = 0,8$ . A questão do emprego de SR ou não, nota-se que quando se utiliza SR, o ganho é sempre superior, como era esperado (FIGURA 1 e 2). Verificou-se também que, para quase totalidade das simulações, as alterações no comportamento com SR em relação ao sem SR, foram pequenas. Desse modo, ênfase será dada à seleção com a utilização de sementes remanescentes.

Uma outra indagação é com relação ao efeito do grau médio de dominância no sucesso com a seleção. Verificou-se que, como esperado, o *GES*

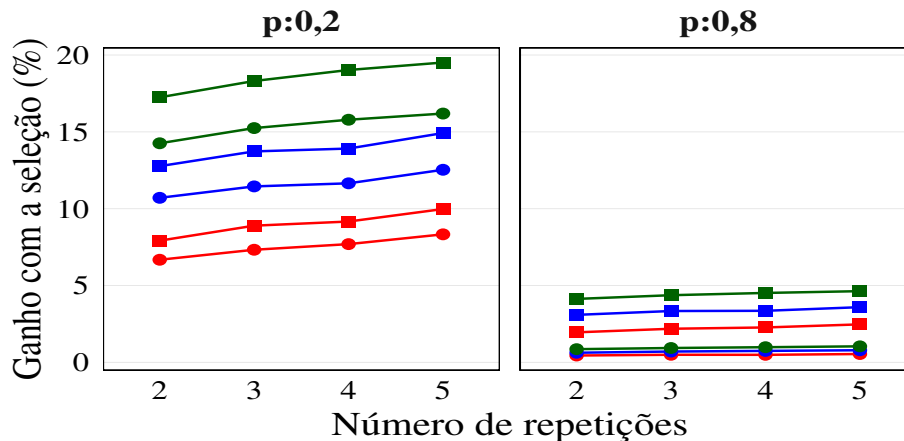
Figura 2 – Estimativa do ganho esperado com a seleção (%) entre (E) e dentro (D) de progênies em experimentos com múltiplas plantas por parcela, considerando-se as intensidades de seleção de 10% E e 20% D,  $gmd = 0$  e  $h^2 = 0,25$ , com o uso de sementes remanescentes (SR). A barra vertical representa o intervalo de confiança de 95%.



Fonte: Do autor (2016).

foi sempre superior com  $gmd = 0$  em relação à  $gmd = 1$  (FIGURA 3). Contudo, independente da frequência alélica, do número de plantas por parcela ou de  $h^2$ , o efeito de  $gmd$  permaneceu praticamente inalterado. Assim, nos demais resultados, uma maior atenção será dada às situações em que não ocorre dominância.

Figura 3 – Estimativa do ganho esperado com a seleção (%) entre (E) e dentro (D) de progênies, considerando-se as intensidades de seleção de 10% E e 20% D, com  $h^2 = 0,1$  (linha vermelha),  $h^2 = 0,25$  (linha azul) e  $h^2 = 0,5$  (linha verde) e grau médio de dominância  $gmd = 0$  (■),  $gmd = 1$  (●), com a utilização de sementes remanescentes em parcelas com 5 plantas.

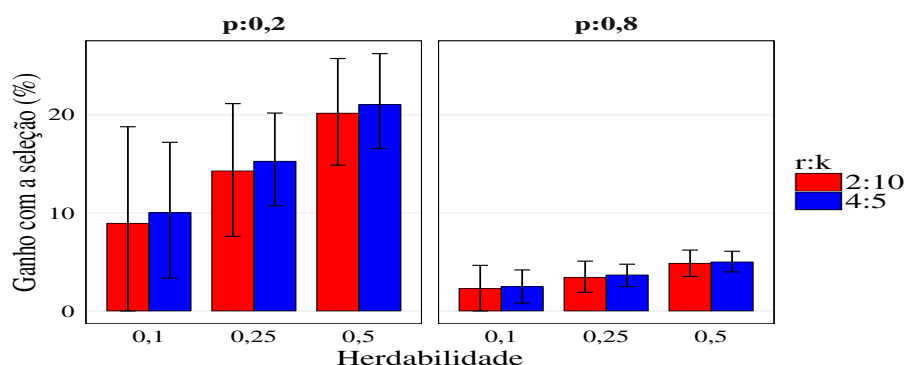


Fonte: Do autor (2016).

Em apenas algumas configurações pode-se comparar o efeito de  $r$  e  $k$  isoladamente. Considerando-se, por exemplo,  $N_f = 300$ , verifica-se que o aumento no número de repetições de 2 para 4 proporcionou maiores ganhos que o aumento do número de plantas por parcela (FIGURA 4). Esse resultado foi mais explícito com  $p = 0,2$ , decorrente da baixa magnitude de ganho observado com frequência alélica 0,8.

O emprego da simulação possibilita, visualizar o efeito do número de progênies no sucesso com a seleção, o que quase nunca é possível nas condições

Figura 4 – Ganho esperado com a seleção (%) em função da alteração em  $r$  e  $k$  quando se utiliza  $N_f = 300$  e 10% de intensidade de seleção entre progênieis em experimentos com múltiplas plantas por parcela. A barra vertical representa o intervalo de confiança de 95%.



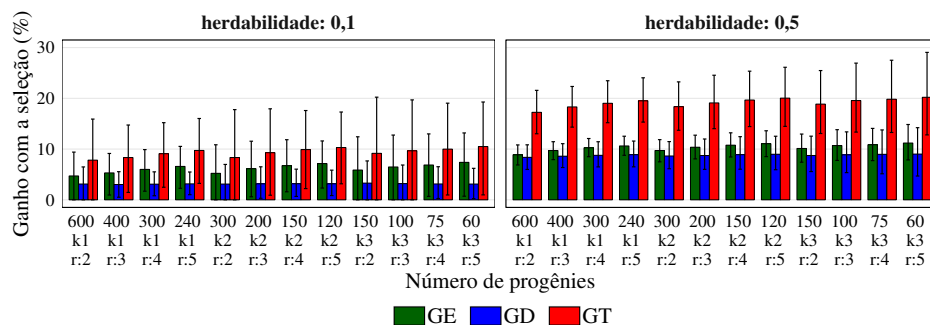
Fonte: Do autor (2016).

de campo. Essa informação não é consensual entre os melhoristas. O que ocorre com o *GES*, quando  $N_f$  é variável, pode ser observado na Figura 5. Apenas os resultados para as herdabilidades 0,1 e 0,5 foram apresentados, pois o caso  $h^2 = 0,25$  foi intermediário a essas situações. Considerando-se o produto  $N_f$ ,  $k$  e  $r$  constante, percentualmente, o *GES* foi sempre maior, como já mencionado, para  $p = 0,2$ .

Pelo que foi mostrado na Figura 5, ao que tudo indica, a variabilidade genética entre as progênieis de meios irmãos ( $\hat{\sigma}_f^2$ ), em média, não varia muito em função da amostra utilizada. Esse resultado pode ser verificado na Tabela 1, onde estão apresentadas as estimativas de  $\hat{\sigma}_f^2$ , que possibilitaram estimar o *GES* apresentados na Figura 5. Veja que na média das 1.000 amostras,  $\hat{\sigma}_f^2$  variou muito pouco em função do número de progênieis, inclusive, se for considerado o intervalo de confiança, os valores são todos iguais, a diferença sendo apenas amostral. Depreende-se que, mesmo com a avaliação de um número menor de progênieis, a representatividade da variância genética, na média, não altera muito. Assim, as variações na amplitude do intervalo de confiança e as maiores estimativas médias



Figura 5 – Estimativa do ganho esperado com a seleção (%) entre (GE), dentro (GD) e ganho total (GT) em experimentos com múltiplas plantas por parcela, considerando-se as intensidades de seleção de 10% E e 20% D, com  $gmd = 0$ ,  $p = 0,2$ ,  $k$  plantas/parcela ( $k_1 = 5$ ,  $k_2 = 10$ ,  $k_3 = 20$ ) e  $r$  repetições, com a utilização de 50 sementes remanescentes, selecionando-se número de plantas igual à  $N_f$ . A barra vertical representa o intervalo de confiança de 95%.



Fonte: Do autor (2016).

do  $GES$ , com menor  $N_f$ , são função do número de repetições.

Sob a ótica do melhoramento, existem duas vertentes a serem consideradas: o ganho com a seleção em um único ciclo seletivo, ou em programas de seleção recorrente. Nos casos até agora apresentados, a ênfase foi no sentido de mostrar os resultados no contexto de um programas de seleção recorrente, já que ao final, gerou-se o mesmo número de progênies da etapa anterior, para a continuidade do processo de seleção. Nesse caso, o que se observa é que não há alteração em relação ao que foi apresentado antes. Verifica-se que, mesmo com a seleção de  $N_f$  plantas, quando se considerou o contexto de um programa de seleção recorrente, e 120 na última situação, como a intensidade de seleção foi a mesma, 2%, os ganhos foram bastante similares (FIGURA 6).

Alternativamente, pode-se considerar que cada planta produza 50 sementes no ciclo, terceira opção da intensidade de seleção dentro, seriam obtidas ao final da seleção E e D, 120 indivíduos, para se ter 6.000 plantas novamente. Como

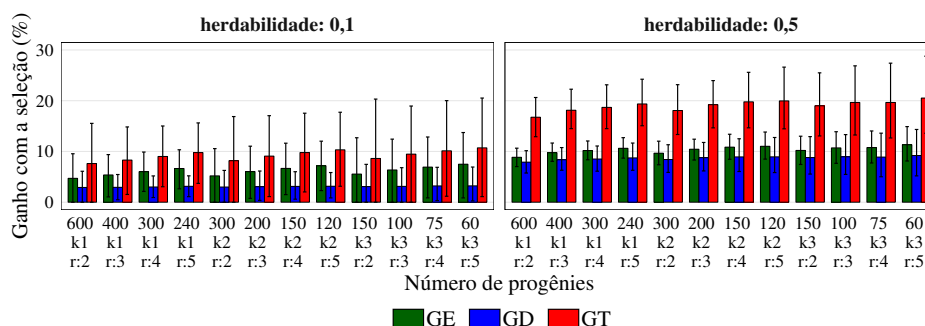
Tabela 1 – Estimativa do componente da variância de progênie ( $\hat{\sigma}_f^2$ ) e do intervalo de confiança de 95% ( $IC_{95}$ ), com  $h^2 = 0,1$  e  $gmd = 0$ , em diferentes frequências alélicas ( $p$ ), número de repetições ( $r$ ), número de plantas por parcela ( $k$ ) e número de progênies ( $N_f$ ).

$h^2$	$p$	$k$	$r$	$N_f$	$\hat{\sigma}_f^2$	$IC_{95}$
0,1	0,2	5	2	600	8,21	(0,00 – 17,04)
0,1	0,2	5	3	400	7,84	(1,38 – 14,40)
0,1	0,2	5	4	300	7,94	(2,20 – 13,96)
0,1	0,2	5	5	240	8,01	(2,63 – 13,67)
0,1	0,2	10	2	300	8,10	(0,00 – 17,65)
0,1	0,2	10	3	200	8,12	(0,82 – 16,22)
0,1	0,2	10	4	150	8,07	(1,79 – 15,08)
0,1	0,2	10	5	120	7,93	(2,09 – 13,89)
0,1	0,2	20	2	150	8,54	(0,00 – 19,09)
0,1	0,2	20	3	100	8,15	(0,00 – 17,39)
0,1	0,2	20	4	75	7,90	(0,69 – 16,02)
0,1	0,2	20	5	60	7,81	(0,64 – 14,79)

Fonte: Do autor (2016)

no caso anterior, a seleção entre não é alterada. Contudo, para se ter um número final de 120 plantas, a intensidade de seleção dentro é variável. Nota-se que, nessa situação, o ganho esperado com a seleção é mais variável que nas situações anteriores. Ocorre ligeiro benefício nos casos em que  $N_f$  foi maior (FIGURA 7).

Figura 6 – Ganho esperado com a seleção entre (GE), com a seleção dentro (GD) e ganho total (GT) em experimentos com múltiplas plantas por parcela, com  $gmd = 0$ ,  $p = 0,2$ ,  $k$  plantas/parcela ( $k1 = 5$ ,  $k2 = 10$ ,  $k3 = 20$ ) e  $r$  repetições, obtido com as intensidades de seleção de 10% entre e 20% dentro de progênies, em que o número de sementes remanescentes variou em função do número de progênies selecionadas, sendo selecionadas 120 plantas. A barra vertical representa o intervalo de confiança de 95%.

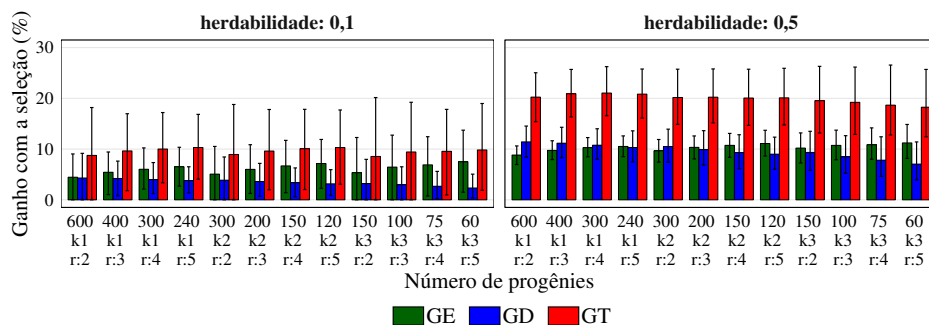


Fonte: Do autor (2016).

Finalmente, no presente estudo, foi possível ter tanto o ganho esperado quanto o ganho realizado com a seleção. Com essa estratégia pode-se também, mostrar a consistência da simulação realizada. Observa-se, na Figura 8, que o ganho realizado foi sempre ligeiramente inferior ao esperado, contudo, a diferença se mantém constante nas diferentes alternativas simuladas. Deve ser salientado que, os intervalos de confiança estimados a partir das 1.000 amostras, no geral, sobrepõem-se.

Por meio da simulação, é possível estimar também, as alterações nas frequências alélicas com a seleção, o que não é possível em condições experimen-

Figura 7 – Ganho esperado com a seleção entre (GE), com a seleção dentro (GD) e ganho total (GT) em experimentos com múltiplas plantas por parcela, com  $gmd = 0$ ,  $p = 0,2$ ,  $k$  plantas/parcela ( $k1 = 5$ ,  $k2 = 10$ ,  $k3 = 20$ ) e  $r$  repetições, obtido com as intensidades de seleção de 10% entre e intensidades de seleção dentro variando em função do número de progênies selecionadas, totalizando 120 plantas selecionadas. A barra vertical representa o intervalo de confiança de 95%.



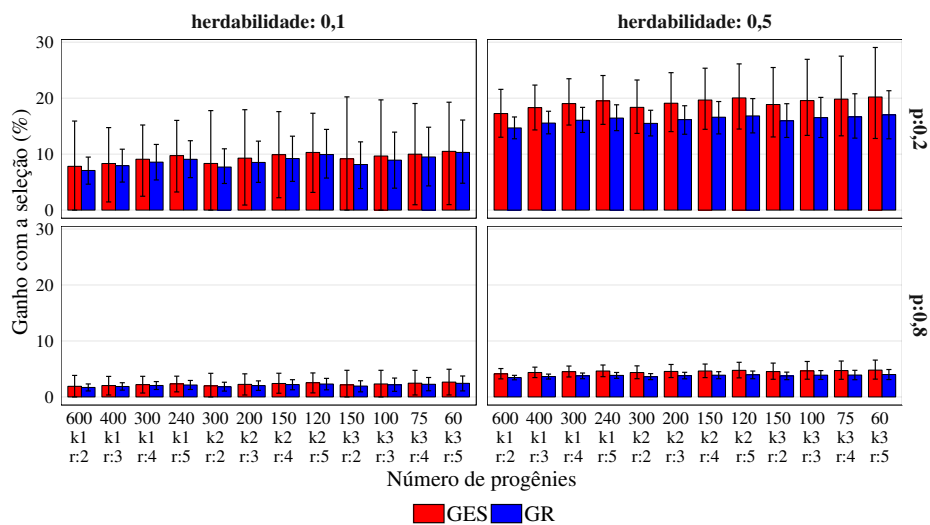
Fonte: Do autor (2016).

tais. Evidencia-se, como era esperado, que as alterações nas frequências alélicas foram mais expressivas nas populações com menor frequência alélica. Observa-se novamente que, a herdabilidade tem grande efeito nas alterações das frequências alélicas com a seleção (FIGURA 9). O mesmo comentário realizado para *GES*, pode ser feito com relação ao efeito do número de progênies nas alterações em frequência alélica, isto é, em média as alterações foram maiores com menor  $N_f$  e maiores números de repetições. Contudo, ocorre sobreposição nos intervalos de confiança.

#### 4.2 Experimentos com uma planta por parcela

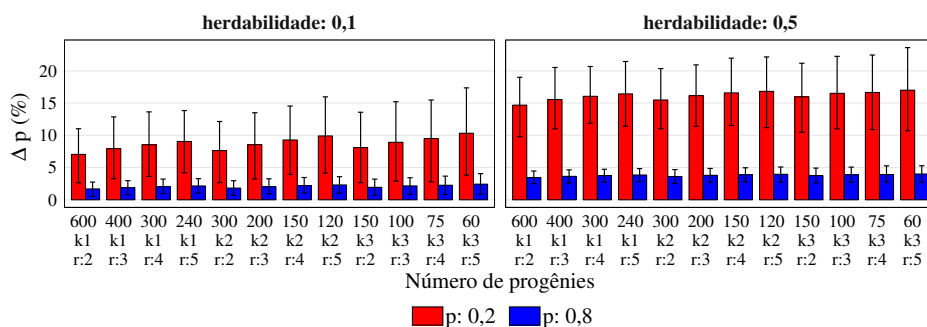
Com o evidente efeito do número de repetições, principalmente nos intervalos de confiança, considerou-se a utilização de experimentos com uma planta por parcela. Nesse caso, as mesmas observações anteriores, com relação ao *gmd*, à intensidade de seleção E e D e às frequências alélicas, são válidas para o caso de

Figura 8 – Estimativas do ganho esperado e do ganho realizado com a seleção (%) em experimentos com múltiplas plantas por parcela, considerando-se as intensidades de seleção de 10% entre e 20% dentro de progênies,  $gmd = 0$ ,  $p = 0,2$ ,  $k$  plantas/parcela ( $k_1 = 5$ ,  $k_2 = 10$ ,  $k_3 = 20$ ) e  $r$  repetições, com a utilização de sementes remanescentes. A barra vertical representa o intervalo de confiança de 95%.



Fonte: Do autor (2016).

Figura 9 – Estimativa do aumento em frequência alélica ( $p$ ) com a seleção entre (E) e dentro (D) de progênes em experimentos com múltiplas plantas por parcela, com as intensidades de seleção de 10% E e 20% D, considerando-se  $gmd = 0$ ,  $k$  plantas/parcela ( $k_1 = 5$ ,  $k_2 = 10$ ,  $k_3 = 20$ ) e  $r$  repetições, com a utilização de sementes remanescentes. A barra vertical representa o intervalo de confiança de 95%.



Fonte: Do autor (2016).

uma planta por parcela. Por essa razão, o foco será dado às alterações em GES, considerando  $gmd = 0$ ,  $p = 0,2$  e  $p = 0,8$ ,  $h^2 = 0,1$  e  $h^2 = 0,5$ , com  $N_f$  e  $r$  variável.

Na Figura 10, é mostrado o GES em função da variação em  $N_f$ , considerando-se os casos de  $h^2 = 0,1$  e  $h^2 = 0,5$ , para a frequência alélica de 0,2. Observe que, em média, o GES total aumenta com a diminuição do número de progênes avaliadas e o consequente aumento de  $r$  para  $h^2 = 0,1$ . Quando  $h^2$  é de 0,5, o incremento também ocorre, porém até  $N_f = 40$ , depois ele decresce. Como se observa, o IC é grande e há sobreposição nos intervalos em todos os casos.

Apesar da variância de progênie se manter constante, como no caso de múltiplas plantas por parcela (TABELA 2), o que se observa é que há um ponto em que a amplitude dos intervalos de confiança de  $\hat{\sigma}_f^2$  e de  $\hat{\sigma}_e^2$  é minimizada (FIGURA A.5, APÊNDICE). A razão entre os ganhos médios obtidos nas diferentes configurações e a amplitude dos intervalos de confiança dos mesmos, mostra que existem combinações de número de progênes e de repetições que propiciam

maximizar o ganho médio e minimizar a amplitude dos intervalos de confiança (FIGURA 11). Verifica-se que, no geral, os valores dessa razão são máximos em torno de  $N_f = 100$ , conseqüentemente  $r = 60$ . Como esperado, em casos de herdabilidades mais elevadas o aumento do número de progênies em detrimento do número de repetições continuou sendo vantajoso para  $N_f > 100$ . Já com  $h^2 = 0,1$ , a redução do número de repetições gerou intervalos de confiança proporcionalmente maiores.

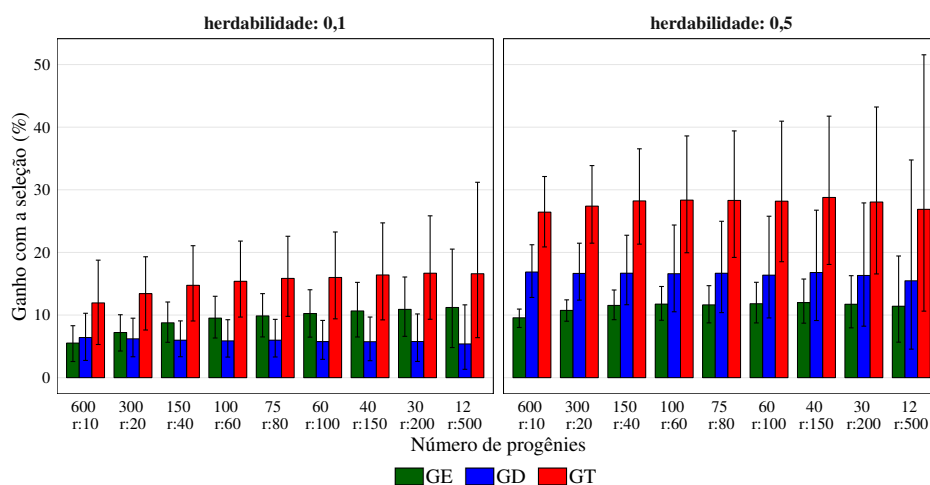
Tabela 2 – Estimativa do componente da variância de progênie ( $\hat{\sigma}_f^2$ ) e do intervalo de confiança de 95% ( $IC_{95}$ ), com  $h^2 = 0,1$  e  $gmd = 0$ , em diferentes frequências alélicas ( $p$ ), números de repetições ( $r$ ) e número de progênies ( $N_f$ ), em experimentos com uma planta por parcela.

$h^2$	$p$	$r$	$N_f$	$\hat{\sigma}_f^2$	$IC_{95}$
0,1	0,2	10	600	7,88	(3,40 – 12,16)
0,1	0,2	20	300	7,99	(4,21 – 11,66)
0,1	0,2	40	150	7,96	(4,52 – 11,75)
0,1	0,2	60	100	7,89	(4,64 – 11,79)
0,1	0,2	80	75	7,98	(4,58 – 12,03)
0,1	0,2	100	60	7,83	(4,08 – 12,06)
0,1	0,2	150	40	7,87	(3,86 – 12,71)
0,1	0,2	200	30	7,94	(3,78 – 13,32)
0,1	0,2	500	12	7,83	(2,23 – 15,65)
0,1	0,8	10	600	7,94	(3,43 – 12,64)
0,1	0,8	20	300	7,89	(4,18 – 11,80)
0,1	0,8	40	150	7,88	(4,54 – 11,96)
0,1	0,8	60	100	7,91	(4,46 – 11,94)
0,1	0,8	80	75	7,83	(4,52 – 11,91)
0,1	0,8	100	60	7,85	(4,36 – 12,00)
0,1	0,8	150	40	7,92	(3,87 – 12,95)
0,1	0,8	200	30	7,86	(3,78 – 13,39)
0,1	0,8	500	12	7,69	(2,63 – 15,66)

Fonte: Do autor (2016)

O ganho genético total, obtido com a utilização de múltiplas plantas por

Figura 10 – Ganho esperado com a seleção entre (GE), com a seleção dentro (GD) e ganho total (GT) em experimentos com uma plantas por parcela, obtido com as intensidades de seleção de 10% entre e 20% dentro das progênies, com  $gmd = 0$  e  $p = 0,2$ . A barra vertical representa o intervalo de confiança de 95%.

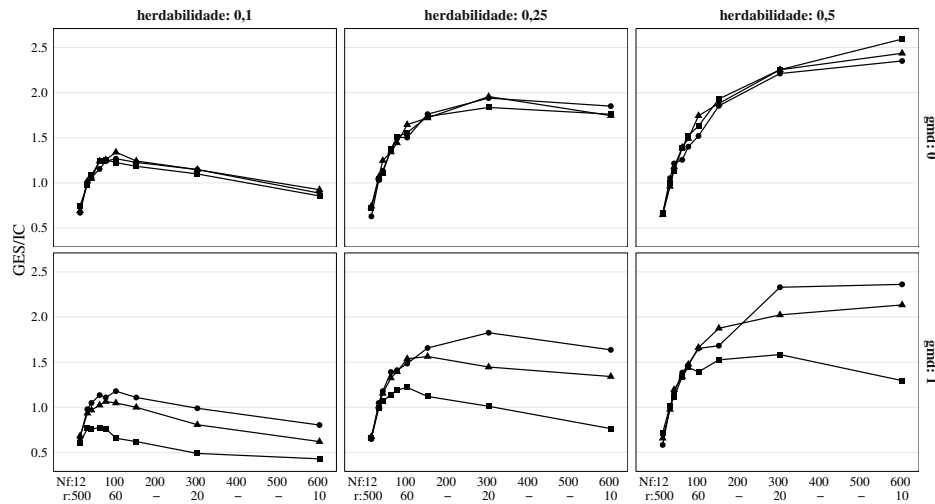


Fonte: Do autor (2016).

parcela ou com parcelas de uma planta não são diretamente comparáveis, já que no primeiro caso, considera-se a seleção dentro apenas do lado materno, enquanto no segundo, a mesma é feita em ambos os sexos. A seleção entre progênie, por outro lado, é semelhante nos dois casos. Nesse sentido, considerando o aumento do número de repetições em detrimento do aumento no número de plantas por parcela, semelhante ao apresentado na Figura 4, o uso de  $r = 20$  e  $k = 1$  foi vantajoso (FIGURA 12). Como esperado, embora ocorra sobreposição nos intervalos de confiança, para um número de progênies fixo, maiores números de repetições proporcionaram um ganho médio superior e um menor  $IC$ .

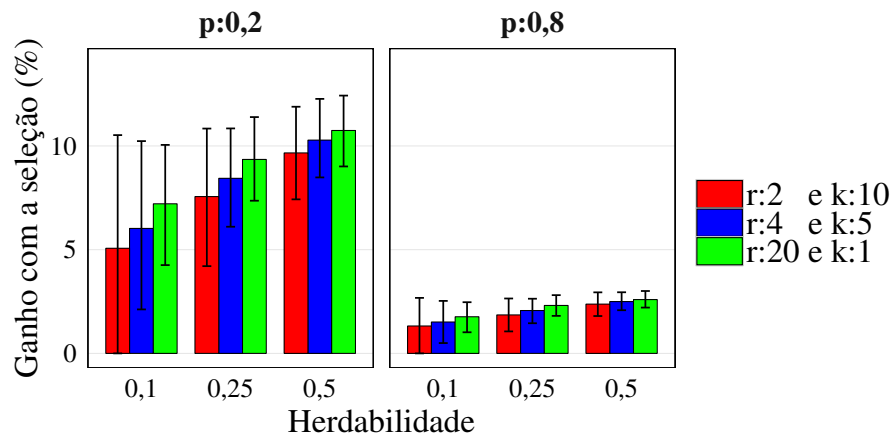


Figura 11 – Razão entre os ganhos médios esperados com a seleção (*GES*) e a amplitude do intervalo de confiança (*IC*) de 95% dos ganhos, em experimentos com uma plantas por parcela, com frequências alélicas 0,2 (●), 0,5 (▲) e 0,8 (■).



Fonte: Do autor (2016).

Figura 12 – Estimativa de ganho esperado com a seleção entre, considerando-se a intensidade de seleção de 10%, com  $N_f = 300$  progênies, avaliadas em 2, 4 e 20 repetições (*r*) e com 10, 5 e 1 planta por parcela (*k*). A barra vertical representa o intervalo de confiança 95%.



Fonte: Do autor (2016).

## 5 DISCUSSÃO

A existência de questionamentos a respeito de estratégias que possam otimizar os programas de melhoramento na seleção de plantas ou progênies é frequente entre os melhoristas. Para responder a esses questionamentos, em qualquer situação estarão envolvidas inúmeras variáveis, o que sob condições reais, é impossível de se avaliar. Uma alternativa possível é a utilização de facilidades computacionais disponíveis, como a simulação. No presente trabalho, para a avaliação da resposta à seleção em progênies de meios-irmãos, foram gerados 360.000 experimentos, envolvendo a variação de propriedades genéticas da população/caráter, tais como frequência alélica e grau médio de dominância, bem como experimentais, sob o controle do melhorista, isto é, número de progênies, de plantas/parcela e de repetições, além de propriedades relacionadas à qualidade dos experimentos, no caso a herdabilidade.

Quando se utiliza simulação, a primeira pergunta a ser feita é se as pressuposições estabelecidas foram atendidas. Um dos pressupostos estabelecidos no presente trabalho foi a herdabilidade em nível de indivíduo. Dessa forma, pode-se comprovar se as estimativas de  $h^2$ , obtidas a partir das simulações, se ajustam às herdabilidades propostas. Isso foi possível de ser realizado, quando se considerou o  $gmd = 0$ . Em todas as situações a  $\hat{h}^2$  das 1.000 simulações foi, em média, similar à  $h^2$  pré-estabelecida, comprovando a eficiência da estratégia utilizada. Esse mesmo procedimento tem sido utilizado em inúmeros trabalhos com simulação (ABREU et al., 2013; NUNES; RAMALHO; FERREIRA, 2008).

Uma outra estratégia de validação da simulação utilizada, foi a verificação da coincidência entre a variância aditiva paramétrica e as estimativas de  $\sigma_A^2$ , obtidas nos experimentos de avaliação das progênies. Constatou-se que a coincidência

foi muito grande, a despeito do erro amostral que está sempre presente em amostras de tamanho finito (WRAY et al., 2013). A comparação dos ganhos esperados com os realizados é outra alternativa que ratifica a validade do processo de seleção utilizado. Nesse caso, a coincidência também foi muito alta (FIGURA 8). Isso ocorre porque, o presente estudo não considerou o efeito da interação genótipos x ambientes, de modo que, uma semelhança entre os mesmos era esperada. Ressalta-se que grande parte das discussões estão assentadas no que ocorreu na média das 1.000 simulações. Contudo, em todas as simulações foi colocado o intervalo de confiança (*IC*), para mostrar que dependendo da situação, diferenças nas estimativas da média do *GES* podem ser atribuídas ao efeito da amostragem genética, uma vez que os *IC*'s se sobrepõem.

Quando se realiza a seleção utilizando progênies, duas situações são possíveis, utilizando sementes remanescentes ou não. A primeira opção é normalmente mais demorada, no entanto, a seleção entre progênies capitaliza o ganho genético nos dois sexos (HALLAUER; CARENA; MIRANDA FILHO, 2010, p. 265). Constatou-se, em todas as simulações, que as estimativas de *GES* mantiveram a mesma proporcionalidade e que nas condições simuladas, o processo com sementes remanescentes propiciou maior *GES* (FIGURA 1). Além do mais, ao que parece, é o processo mais utilizado atualmente.

Com relação ao grau médio de dominância, foram adotadas duas situações,  $gmd = 0$  ou  $gmd = 1$ . Embora existam algumas controvérsias, ao que tudo indica, não ocorre sobredominância no controle genético dos caracteres. Um grande número de estimativas a respeito da importância da dominância, foi obtido para a produtividade de grãos na cultura do milho, os resultados de 99 estimativas mostraram que, na média a relação  $\hat{\sigma}_D^2/\hat{\sigma}_A^2$  foi de 0,61 (HALLAUER; CARENA; MIRANDA FILHO, 2010, p. 171). Para outros caracteres, essa relação foi ainda

menor. Desse modo, pode-se inferir que o *gmd* utilizado neste trabalho, reflete o intervalo em que o mesmo ocorre com maior predominância, nos diferentes caracteres das espécies cultivadas.

Considerando-se um loco, a  $\sigma_A^2$  é obtida pela expressão  $\sigma_A^2 = 2p(1-p)[a + d(1-2p)]^2$  e  $\sigma_D^2 = [2p(1-p)d]^2$ . Depreende-se que, quando  $p = 0,5$ , a contribuição dos locos em heterozigose, ou o efeito da dominância ( $d$ ), não afeta a estimativa de  $\sigma_A^2$ . Portanto, a discussão do efeito de *gmd*, só tem significado para as frequências alélicas 0,2 e 0,8. Bernardo (2002, p. 94) mostra que  $\sigma_A^2$  é máxima com  $p = 0,25$  e  $\sigma_D^2$  é máxima com  $p = 0,5$ . Assim, o comportamento de  $\sigma_A^2$  e  $\sigma_D^2$  fica realçado nas simulações realizadas, já que as estimativas de *GES* foram de maior magnitude quando  $p = 0,2$  em relação ao  $p = 0,8$ . Do exposto, a discussão das simulações será focada nas variáveis que podem ser controladas pelos melhoristas, implicando em possíveis sugestões/recomendações. No entanto, é oportuno salientar que, embora considere-se apenas os casos de  $p = 0,2$  ou  $p = 0,8$ ,  $gmd = 0$  e com sementes remanescentes, as observações podem ser generalizadas para a maioria dos caracteres e populações de plantas que os melhoristas trabalham. Há de se ressaltar também que, quanto menor a frequência alélica média do alelo desejado na população, maior é o ganho esperado (FALCONER; MACKAY, 1996, p. 28), o que está de acordo com o que foi comentado anteriormente.

Quando se seleciona as melhores progênies, essa seleção é feita com base em experimentos com repetições. A seleção das melhores plantas no lote de recombinação por outro lado, baseia-se no fenótipo de um único indivíduo, ou seja, seleção massal. Esse tipo de seleção é de menor eficiência, especialmente em situações de baixas herdabilidades (ABREU et al., 2010). Nesse sentido, um maior ganho com o aumento da intensidade de seleção (menor proporção de indivíduos selecionados) entre progênies, em detrimento da intensidade de seleção

dentro é esperado, quando o caráter tem  $h^2$  baixa, inferior a 0,5 por exemplo, o que é comum para a maioria dos caracteres, independente da espécie. Esse resultado foi observado nas simulações realizadas, em que a maior  $h^2$  foi de 0,5, pois, embora ocorra um aumento da contribuição da seleção dentro para o *GES* com maiores valores de  $h^2$ , esse não foi suficiente para justificar a adoção de uma intensidade de seleção entre de 20%, em relação ao uso de 10%. Essa maior vantagem do aumento da intensidade de seleção, principalmente com baixa  $h^2$ , está de acordo com as discussões apresentadas por Pereira e Vencovsky (1988), na qual os autores afirmam que, em alguns casos, uma das formas de se compensar baixas herdabilidades é aumentar a intensidade de seleção.

Para a obtenção de maiores herdabilidades, além do controle ambiental, a principal alternativa é aumentar o número de repetições e/ou plantas por parcela. Nas simulações realizadas, partiu-se do pressuposto de que a área experimental era constante, 6.000 plantas. Assim, o número de plantas por parcela, por exemplo, é dependente do número de repetições. As estimativas do *GES* variaram com  $h^2$ ,  $r$ ,  $k$  ou  $N_f$  (FIGURAS 5, 6, 7 e 10), mas, de modo geral, os valores são comparáveis aos que são normalmente citados na literatura. Novamente, tomando a cultura do milho como referência, constatou-se em um levantamento dos *GES* de trabalhos na literatura que em média os ganhos variaram de 3 a 7%, embora existam valores inferiores ou superiores a esse (HALLAUER; CARENA; MIRANDA FILHO, 2010, p. 632). Depreende-se, mais uma vez que, as simulações realizadas são coerentes com o que os melhoristas obtém de resultado prático com a seleção. Além disso, nota-se que, quanto maior a  $h^2$ , maior é o ganho com a seleção, independente das outras variáveis (FIGURA 4). Novamente, pode-se inferir que uma das principais estratégias para otimizar o progresso com a seleção é ter experimentos mais precisos, durante a avaliação das progênies.

Comparando-se o aumento do número de repetições em relação ao aumento do plantas por parcela, verifica-se pelos ganhos com a seleção entre com  $N_f = 300$ , que há maior vantagem em se aumentar  $r$ . O que se observa com  $h^2 = 0,1$ , é a possibilidade de obtenção de ganhos nulos, se forem utilizadas apenas 2 repetições e  $k = 10$ . Embora os  $IC$ 's se sobreponham, nessa mesma situação o limite inferior para  $r = 20$  e  $k = 1$  foi em torno de 5 (FIGURA 12). Esse resultado é esperado, pois considerando-se a variância fenotípica entre média de progênies (HALLAUER; CARENA; MIRANDA FILHO, 2010, p. 273), pode-se visualizar que, na maioria dos casos, é preferível aumentar  $r$  do que  $k$ , para que se tenha maior  $h_m^2$ , e conseqüentemente, maior GES. Muitos resultados encontrados na literatura apontam para uma maior vantagem de se utilizar poucas ou mesmo uma planta por parcela, principalmente no contexto de plantas perenes (HAAPENEN, 1992; JANSSON; DANELL; STENER, 1998; SCARPINATI et al., 2009). No entanto, o uso de uma planta por parcela em culturas anuais pode ser inviável, do ponto de vista prático, especialmente se forem consideradas as operações de manejo todas mecanizadas. Assim, o número de plantas por parcela varia em função da cultura, em cenoura, por exemplo, verificou-se que uma amostra de 18 plantas por parcela seriam o suficiente (SILVA; VIEIRA; VILLELA, 2009). Já em milho, podem ser recomendadas de 10 a 20 plantas por parcela, dependendo da condição de solo considerada (RESENDE; SOUZA JÚNIOR, 1997). Outros trabalhos também chegaram a resultados semelhantes, indicando que a quantidade de plantas ideal para avaliação de produtividade em milho devesse ser entre 15 e 20, mas semelhante ao que foi observado no presente estudo, dada a necessidade de se escolher entre  $k$  e  $r$ , a segunda opção deveria ser privilegiada (CHAVES; MIRANDA FILHO, 1992).

Como visto, embora ocorra sobreposição dos intervalos de confiança, o aumento do número de repetições proporcionou GES maiores, principalmente com

baixos valores de  $h^2$ . Nos casos de  $h^2 = 0,5$ , esse aumento também é verificado, porém com menor intensidade. Com base em estudos de re-amostragem, Cotterill e James (1984) mostram que com uma herdabilidade de 0,4, o uso de 50 plantas por progênie é suficiente para que se tenha 91% de probabilidade de detectar diferença acima de 0,25 desvios-padrão fenotípico. Os mesmos autores chegam à conclusão de que 10 a 20 indivíduos são suficientes para se avaliar uma progênie. Os resultados obtidos neste estudo concordam com o trabalho mencionado, no sentido de que, com valores de  $N_p$  entre 40 e 60, no geral, a razão entre o  $GES$  e a amplitude do  $IC$  esteve próxima dos valores máximos. Já a utilização de  $N_p$  entre 10 e 20, considerando as Figuras 7 e 11, foi a melhor opção apenas em altas herdabilidades.

O aumento do número de plantas por progênie possibilita uma avaliação mais acurada de cada progênie, no entanto, maior  $N_f$  permite maior intensidade de seleção entre. Como a intensidade de seleção foi mantida constante, o principal efeito observado, como comentado, é em relação à amplitude do intervalo de confiança. Verifica-se que números muito pequenos de  $N_f$  resultam em uma grande variação da estimativa da variância genética, o que representa um risco de obtenção de ganhos consideravelmente inferiores (FIGURA A.5, APÊNDICE). Com relação ao efeito de  $N_f$  e  $N_p$  na precisão das estimativas dos parâmetros genéticos, em um trabalho realizado em trevo vermelho, Xie e Mosjidis (1997) utilizando re-amostragem em situações experimentais com duas e com quatro repetições, verificaram que o aumento na precisão das estimativas dos parâmetros genéticos, tais como a herdabilidade e o ganho, nunca foi alcançado aumentando-se somente o número de plantas por progênies. Para que isso ocorresse, foi necessário também, o aumento do número de progênies. Os autores mostram que, o decréscimo nas estimativas do erro padrão foi pequeno com o aumento do número

de plantas por progênie de três para seis, mas com o aumento do número de progênies, esse decréscimo foi proporcionalmente maior. Nesse sentido, Cottrell e James (1984) observaram que o ordenamento das progênies com a utilização de apenas cinco indivíduos, apresentou baixa correlação com o ordenamento utilizando todo o banco de dados, de modo que, mesmo a utilização de seis indivíduos mencionada no trabalho anterior, pode estar abaixo do recomendado.

Se, por um lado, o uso de poucas progênie implica em maiores intervalos de confiança para as estimativas de ganhos genéticos, por outro lado, o uso de valores elevados de  $N_f$  e, por consequência  $N_p$  pequenos, resultaram em ganhos médios inferiores, não sendo, portanto, recomendáveis, principalmente em situações de baixa  $h^2$ . Em experimentos avaliando duas espécies florestais com 225 e 275 progênies em 5 e 12 repetições por local, Perron, DeBlois e Despons (2013) também observaram que a amplitude do *IC* diminuiu com o aumento de  $N_f$ . No entanto, os autores constataram que a variação em  $N_p$  teve maior impacto na magnitude dos parâmetros testados e assim o aumento do número de plantas por progênie teria maior relevância. Nesse caso, a conclusão a que se chegou foi que, o uso de 75 progênies e 12 plantas por progênie por local parecem ser valores mínimos para se estimar a herdabilidade em meios irmãos, tanto para altura quanto para densidade de madeira.

Como no caso anterior, em meios-irmãos de cenoura, o número adequado para a avaliação de progênies está em torno 74 (SILVA; VIEIRA; VILLELA, 2009). No entanto, os resultados apresentados, na Figura 11, mostram que, embora  $N_f = 75$  apresente uma boa relação entre a magnitude e precisão das estimativas de ganho, os melhores resultados podem ser obtidos com  $N_f$  em torno de 100 ou ainda maiores com herdabilidades mais altas. Como pode ser observado nos resultados deste trabalho, o número ótimo de progênies, no geral, depende da herdabilidade.



Desse modo, com baixa  $h^2$  o ganho entre é mais importante, o que sugere a necessidade maiores valores de  $N_p$ , para que se tenha uma avaliação acurada do efeito das progênies (LINDGREN; MULLIN, 1997).

A maior preocupação com o número de progênies em um programa de seleção recorrente, normalmente está relacionada à manutenção do tamanho efetivo populacional ( $N_e$ ) e, conseqüentemente, da variabilidade, para que se possibilite a obtenção de ganhos futuros. Alguns trabalhos têm mostrado que, quando a principal preocupação é o ganho a curto prazo, há pouca vantagem em se utilizar grandes  $N_e$ 's. Baker e Curnow (1969), por exemplo, sugerem que os limites seletivos não seriam alcançados em muito menos de 30 gerações, utilizando um tamanho efetivo de 16. Da mesma forma, os resultados obtidos por Weyhrich, Lamkey e Hallauer (1998) para progênies  $S_1$ , indicam não haver vantagem, a curto prazo, em se utilizar mais de 10 linhagens para intercrossamentos. O fato de a variância de progênies ter sido em média, constante, favorece a argumentação em relação ao uso de poucas progênies. No entanto, a grande variação nas estimativas, com  $N_f$  pequeno, evidenciam uma maior exposição à deriva genética. Nesse sentido, existem casos onde o recomendado seriam maiores valores de  $N_f$ , por exemplo, quando a população é não adaptada, para que se possa compensar a perda de variabilidade, em decorrência da seleção natural (PEÑA-ASIN; ÁLVAREZ; ORDAS, 2013). Segundo Pereira e Vencovsky (1988), um maior tamanho efetivo tende a compensar algumas dificuldades que impedem o progresso máximo possível com a seleção, no entanto, os autores deixam claro que essa estratégia, só é recomendável se for conseguida sem sacrificar a intensidade de seleção, pois dessa última depende o progresso imediato.

O aumento da intensidade de seleção obviamente proporciona maiores ganhos e, até certo ponto, a maneira mais adequada de se obter esse aumento é com

a avaliação de um maior número de progênies (CHAVES; MIRANDA FILHO, 1992). A dificuldade encontrada é que a condução das atividades de melhoramento com número excessivo de progênies, demanda mais recursos humanos e financeiros (VIEIRA; MARCOS; BOSCO, 2006), além da maior dificuldade de se ter experimentos precisos, restringindo, muitas vezes, o uso de um maior  $N_f$ . Os resultados apresentados, na Figura 7, mostram que, em se utilizando maiores intensidades de seleção dentro o ganho total aumenta. No entanto, o ponto máximo do *GES* em torno de  $N_f = 300$  e  $N_p = 20$  com  $h^2 = 0,5$ , confirmam os comentários anteriores a respeito do número mínimo de plantas por progênie. Obviamente, há que se considerar o esquema de recombinação utilizado, mas, fica claro que os ganhos, mesmo com intensidade total de seleção variando, foram menores que os obtidos com  $k = 1$ .

## 6 CONCLUSÃO

Os resultados obtidos, neste estudo, mostram que o aumento do número de repetições em detrimento do aumento do número de plantas por parcela, proporciona maiores ganhos genéticos. Esses ganhos foram, em média, maiores em experimentos com uma planta por parcela. Evidenciou-se também que é possível obter ganhos semelhantes com diferentes números de progênies. No entanto, os resultados obtidos com o uso de poucas progênies apresentam grandes oscilações e conseqüentemente maiores riscos. Por outro lado, o uso de poucas plantas por progênie resulta em ganhos, em média, inferiores. Para as situações consideradas, neste estudo, o uso de aproximadamente 100 progênies em 60 repetições, resultou na melhor relação entre a magnitude dos ganhos genéticos e a precisão das estimativas obtidas.

## REFERÊNCIAS

- ABREU, G. B. et al. Computer Simulation for the Evaluation of Recombination Strategies in Intrapopulation Recurrent Selection in Eucalyptus. **Silvae genetica**, v. 62, n. 1/2, p. 68–79, 2013.
- ABREU, G. B. et al. Strategies to improve mass selection in maize. **Maydica**, v. 55, n. 3/4, p. 219–255, 2010.
- ACQUAAH, G. **Principles of Plant Genetics and Breeding**. 2. ed. Oxford: Blackwell, 2007. 569 p.
- ARAÚJO, P. A. D. et al. Avaliação de famílias de meios-irmãos de milho em diferentes espaçamentos entre linhas. **Revista Brasileira de Milho e Sorgo**, v. 4, n. 3, p. 428–435, 2005.
- BAKER, L. H.; CURNOW, R. N. Choice of Population Size and Use of Variation Between Replicate Populations in Plant Breeding Selection Programs. **Crop Science**, v. 9, p. 555–560, 1969.
- BAKER, R. J. **Selection indices in plant breeding**. Florida: Boca Raton, 1995. 218 p.
- BERNARDO, R. **Breeding for quantitative traits in plants**. Woodbury, MN: Stemma Press, 2002. 369 p.
- BERNARDO, R. Parental selection number of breeding populations and size of each population in inbred development. **Theoretical and Applied Genetics**, v. 107, n. 7, p. 1252–1256, 2003.
- BIJMA, P. Long-term genomic improvement - new challenges for population genetics. **Journal of Animal Breeding and Genetics**, v. 129, n. 1, p. 1–2, 2012.
- BOX, G. E. P.; DRAPER, N. R. **Empirical model building and response surfaces**. New York: Wiley, 1987. 669 p.
- BRUMMER, E. C.; CASLER, M. D. Improving selection in forage, turf, and biomass crops using molecular markers. In: YAMADA, T.; SPANGENBERG, G. (Ed.). **Molecular Breeding of Forage and Turf: The Proceedings of the 5th International Symposium on the Molecular Breeding of Forage and Turf**. [S.l.]: Springer New York, 2010.

CARDOSO, M. J. et al. Melhoramento genético das cultivares de milho CMS 47 e BR 5039 (São Vicente) Na Região meio-norte do Brasil. **Revista Brasileira de Milho e Sorgo**, v. 2, n. 3, p. 88–96, 2003.

CARGNELUTTI FILLHO, A. et al. Tamanho ótimo de parcela em milho com comparação de dois métodos. **Ciência Rural**, v. 41, n. 11, p. 1890–1898, 2011.

CARVALHO, H. et al. Avaliação De Progênes De Meios-Irmãos. **Pesquisa Agropecuária Brasileira**, v. 35, n. 8, p. 1577–1584, 2000.

CARVALHO, H. W. L. d. et al. Três ciclos de seleção entre e dentro de progênes de meios-irmãos na população de milho CMS-52. **Pesquisa Agropecuária Brasileira**, v. 35, n. 8, p. 1621–1628, 2000.

CARVALHO, H. W. L. de; SANTOS, M. X. dos; LEAL, M. d. L. d. S. Estimativas de parâmetros genéticos na população de milho CPATC-3 no nordeste brasileiro. **Agrotrópica**, v. 17, p. 47–52, 2005.

CASTRO, N. **Número de repetições e eficiência da seleção em progênes de meio irmãos em Eucalyptus camaldulensis**. 121 p. Tese (Doutorado) — Escola Superior de Agricultura de Lavras, Lavras, 1992.

CHAVES, L.; MIRANDA FILHO, J. Plot size for progeny selection in maize (*Zea mays* L.). **Theoretical and Applied Genetics**, v. 84-84, p. 963–970, sep 1992.

COMPTON, W. A.; BAHADUR, K. Ten Cycles of Progress from Modified Ear-to-Row Selection in Corn. **Crop Science**, v. 17, p. 378–380, 1977.

COMPTON, W. A.; COMSTOCK, R. E. More on modified ear-to-row selection in corn. **Crop Science**, v. 16, n. 74, p. 122, 1976.

COTTERILL, P.; JAMES, J. Number of offspring and plot sizes required for progeny testing. **Silvae Genetica**, v. 33, n. 6, p. 203–209, 1984.

CRUZ, C. D. A informática no melhoramento genético. In: NASS, L. et al. (Ed.). **Recursos genéticos e melhoramento de plantas**. Rondonópolis-MT: Fundação-MT, 2001. p. 1086–1118.

DE LOS CAMPOS, G.; SORENSEN, D.; GIANOLA, D. Genomic heritability: what is it? **PLoS genetics**, v. 11, n. 5, p. e1005048, 2015.

FALCONER, D. S.; MACKAY, T. F. **Introduction to Quantitative Genetics**. 4. ed. [S.l.]: Pearson, 1996. 480 p.

- FERREIRA, D. F. Uso da simulação no melhoramento. In: NASS, L. L. et al. (Ed.). **Recursos genéticos e melhoramento de plantas**. Rondonópolis/MT: Fundação-MT, 2001. p. 1119–1141.
- FERREIRA, W. D. et al. Family number in common bean selection. **Genetics and Molecular Biology**, v. 23, n. 2, p. 403–409, jun 2000.
- FOUILLOUX, G.; BANNEROT, H. Selection methods in the common bean (*Phaseolus vulgaris*). In: **Genetic resources of Phaseolus Beans: their maintenance, domestication, evolution and utilization**. Netherlands: Kluwer Academic Publishers, 1988. p. 503–542.
- FRASER, A. Simulation of genetics systems by automatic digital computers ii: Effect of linkage on rates of advance under selection. **Australian Journal of Biological Science**, v. 10, p. 492–499, 1957.
- GERLOFF, J. E.; SMITH, O. S. Choice of method for identifying germplasm with superior alleles. **Theoretical and Applied Genetics**, v. 76, p. 209–216, 1988.
- GRANATE, M.; CRUZ, C. D.; PACHECO, C. A. P. Número mínimo de famílias de meios-irmãos para representar uma população de milho-pipoca. **Revista Ceres**, v. 48, n. 276, p. 209–221, 2001.
- HAAPENEN, M. Effect of plot size and shape on the efficiency of progeny tests. **silva Fennica**, v. 26, n. 4, p. 201–209, 1992.
- HALLAUER, A. R.; CARENA, M. J. Recurrent selection methods to improve germplasm in maize. **Maydica**, v. 57, p. 266–283, 2012.
- HALLAUER, A. R.; CARENA, M. J.; MIRANDA FILHO, J. **Quantitative Genetics in Maize Breeding**. 3. ed. New York: Springer, 2010. v. 6. 663 p.
- JANSSON, G.; DANELL, Ö.; STENER, L.-G. Correspondence between single-tree and multiple-tree plot genetic tests for production traits in *Pinus sylvestris*. **Canadian Journal of Forest Research**, v. 28, n. 3, p. 450–458, 1998.
- JONES, L. P.; COMPTON, W. A.; O.GARDNER, C. Comparison of Full and Half-Sib Reciprocal Recurrent Selection. **Theoretical and Applied Genetics**, v. 41, p. 36–39, 1971.
- KERR, R.; DIETERS, M.; TIER, B. Simulation of the comparative gains from four different hybrid tree breeding strategies. **Canadian Journal of Forest Research**, v. 34, n. 1, p. 209–220, 2004.

LEITE, M. et al. Sample size for full sib evaluation in sugarcane. **Pesq. agropec. bras.**, v. 44, n. 12, p. 1562–1574, 2009.

LINDGREN, D.; MULLIN, T. J. Balancing gain and relatedness in selection. **Silvae Genetica**, v. 46, n. 2-3, p. 124–129, 1997.

LONNQUIST, J. H. A modification of the ear-to-row procedure for the improvement of maize populations. **Crop Science**, v. 4, p. 227–228, 1964.

MARQUEZ-SANCHEZ, F. Tamaño de muestra para representar poblaciones de maíz. **Agrociencia**, v. 8, p. 163–177, 1972.

MARTIN, T. N. et al. Bases genéticas de milho e alterações no plano experimental. **Pesq. agropec. bras.**, v. 40, n. 1, p. 35–40, 2005.

MARTINS, I. S. et al. Comparação entre os processos de seleção entre e dentro e o de seleção combinada, em progênies de *Eucalyptus grandis*. **Cerne**, v. 11, n. 1, p. 16–24, 2005.

MATTA, F.; VIANNA, J. Eficiência relativa dos processos de seleção entre e dentro de meios-irmãos em população de milho-pipoca. **Rev. Ciênc. Agron.**, v. 27, n. 3, p. 548–556, 2003.

MELO, D. **Número ideal de genitores para a geração de populações sintéticas em eucalipto via simulação**. 81 p. Tese (Doutorado) — Universidade Federal de Lavras, Lavras, 2010.

MIKEL, M. a. Availability and Analysis of Proprietary Dent Corn Inbred Lines with Expired U.S. Plant Variety Protection. **Crop Science**, v. 46, n. 6, p. 2555, 2006.

NGUYEN, H. T.; SLEPER, D. a. Theory and application of half-sib matings in forage grass breeding. **Theoretical and Applied Genetics**, v. 64, n. 3, p. 187–196, feb 1983.

NUNES, J.; RAMALHO, M. A. P.; FERREIRA, D. F. Inclusion of genetic relationship information in the pedigree selection method using mixed models. **Genetics and Molecular Biology**, v. 31, n. 1, p. 73–78, 2008.

PALOMINO, E.; RAMALHO, M.; FERREIRA, D. F. Tamanho de amostra para avaliação de famílias de meios-irmãos de milho. **Pesq. agropec. bras.**, v. 35, n. 7, p. 1433–1439, 2000.

PANDEY, S.; GARDNER, C. O. Recurrent selection for population, variety and hybrid improvement in tropical maize. **Advances in Agronomy**, Cali, v. 48, p. 87, 1992.

PATERNIANI, E. Selection Among and Within Half-Sib Families in a Brazilian Population of Maize (*Zea mays* L.). **Crop Science**, v. 7, p. 212–216, 1967.

PATERNIANI, E.; FILHO, J. M. Melhoramento de populações. In: PATERNIANI, E. (Ed.). **Melhoramento e produção de milho no Brasil**. Piracicaba: Fundação Cargill, 1978. p. 202–256.

PEÑA-ASIN, J.; ÁLVAREZ, A.; ORDAS, B. Molecular changes during intra and inter recurrent selection of two populations of maize: One adapted and one non adapted to the selection environment. **Euphytica**, v. 193, n. 3, p. 359–367, 2013.

PEREIRA, M. B.; VENCOVSKY, R. Limites da seleção recorrente. **Pesq. agropec. bras.**, v. 23, n. 7, p. 769–780, 1988.

PERRON, M.; DEBLOIS, J.; DESPONTS, M. Use of resampling to assess optimal subgroup composition for estimating genetic parameters from progeny trials. **Tree Genetics & Genomes**, v. 9, n. 1, p. 129–143, 2013.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2014. Disponível em: (<http://www.R-project.org/>).

RAMALHO, M. A. P. **Eficiência relativa de alguns processos de seleção intrapopulacional no milho baseados em famílias não endógamas**. 122 p. Tese (Doutorado) — ESALQ-USP, 1977.

RAMALHO, M. A. P.; ABREU, A. d. F. B.; SANTOS, J. B. Melhoramento de espécies autógamas. In: NASS, L. et al. (Ed.). **Recursos genéticos e melhoramento de plantas**. Rondonópolis-MT: Fundação-MT, 2001. p. 201–230.

RESENDE, M. D. V. Delineamentos de experimentos de seleção para a maximização da acurácia seletiva e do progresso genético. **Revista Árvore**, v. 19, n. 4, p. 479–500, 1995.

RESENDE, M. D. V. d.; SOUZA JÚNIOR, C. L. D. Número de repetições e tamanho de parcela para seleção de progênies de milho em solos sob cerrado e fértil. **Pesq. agropec. bras.**, v. 32, p. 781–788, 1997.

SCARPINATI, E. A. et al. Influência do modelo de análise estatística e da forma das parcelas experimentais na seleção de clones de *Eucalyptus spp.* **Revista Árvore**, v. 33, n. 4, p. 769–776, 2009.



- SILVA, G.; VIEIRA, J.; VILLELA, M. Tamanho de amostra para avaliação de caracteres de cenoura em sistemas de cultivo agroecológico. **Horticultura Brasileira**, v. 27, n. 2, p. 166–170, 2009.
- SILVA, H.; FERREIRA, D. F.; PACHECO, C. A. P. Avaliação de quatro alternativas de análise de experimentos em látice quadrado, quanto à estimação de componentes de variância. **Bragantia**, v. 59, n. 1, p. 117–123, 2000.
- SILVA, J. W. **Emprego da simulação na escolha do número de indivíduos no método “bulk” no melhoramento de plantas autógamas**. Tese (Doutorado) — Universidade Federal de Lavras, Lavras, 2009.
- SIMMONDS, N. W.; ROAD, W. M. Selection for local adaptation in a plant breeding programme. **Theoretical and Applied Genetics**, v. 82, p. 363–367, 1991.
- SOUZA, E. F. M.; PETERNELLI, L.; BARBOSA, M. Design and model effects definitions in the initial stage of a plant breeding program. **Pesq. agropec. bras.**, v. 41, n. 3, p. 369–375, 2006.
- SOUZA JÚNIOR, C. L. D. Melhoramento de espécies alógamas. In: NASS, L. et al. (Ed.). **Recursos genéticos e melhoramento de plantas**. Rondonópolis-MT: Fundação-MT, 2001. p. 159–199.
- VEIGA, R. D.; FERREIRA, D. F.; RAMALHO, M. Eficiência dos dialelos circulantes na escolha de genitores. **Pesq. agropec. bras.**, v. 35, n. 7, p. 1395–1406, 2000.
- VENCOVSKY, R. Herança quantitativa. In: PATERNIANI, E. (Ed.). **Melhoramento e produção de milho no Brasil**. Piracicaba: Fundação Cargill, 1978. p. 122–201.
- VENCOVSKY, R.; BARRIGA, P. **Genética biométrica no fitomelhoramento**. Ribeirão Preto: Sociedade Brasileira de Genética, 1992. 496 p.
- VIANA, J. M. S. Breeding strategies for recurrent selection of maize. **Pesq. agropec. bras.**, v. 42, n. 10, p. 1383–1391, 2007.
- VIEIRA, J. V.; MARCOS, W.; BOSCO, J. Número mínimo de famílias de meios-irmãos para avaliação de uma população de cenoura. **Pesq. agropec. bras.**, v. 41, n. 2, p. 365–367, 2006.
- WANG, J. et al. Comparison of Two Breeding Strategies by Computer Simulation. **Crop Science**, v. 43, p. 1764–1773, 2003.

WEYHRICH, R. a.; LAMKEY, K. R.; HALLAUER, A. R. Responses to Seven Methods of Recurrent Selection in the BS11 Maize Population. **Crop Science**, v. 38, n. 2, p. 308–321, 1998.

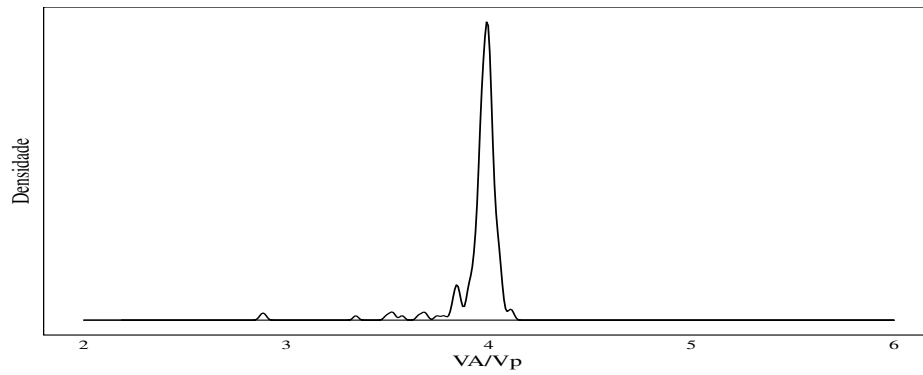
WRAY, N. R. et al. Pitfalls of predicting complex traits from SNPs. **Nature reviews**, Nature Publishing Group, v. 14, n. 7, p. 507–15, 2013. ISSN 1471-0064.

WRICKE, G.; WEBER, W. E. **Quantitative genetics and selection in plant breeding**. Berlin: de Gruyter, 1986. 406 p.

XIE, C.; MOSJIDIS, J. A. Influence of sample size on precision of heritability and expected selection response in red clover. **Plant Breeding**, v. 116, p. 83–88, 1997.

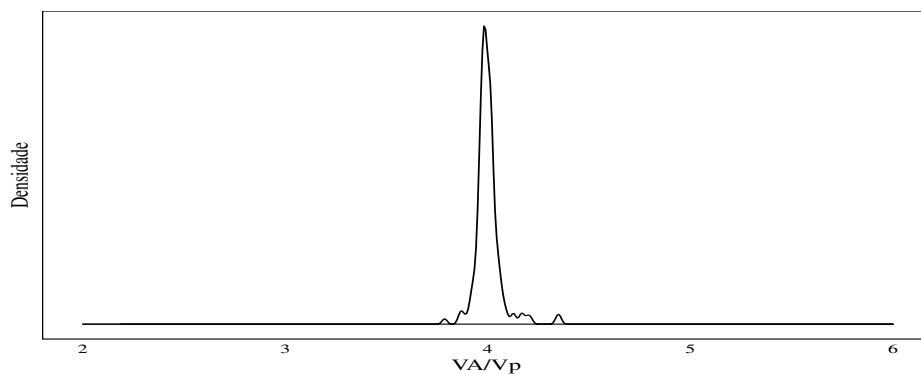
**APÊNDICE**

Figura A.1 – Distribuição da razão entre a variância aditiva da população e a variância de progênie, estimada pela análise de variância ao longo das simulações, com múltiplas plantas por parcela.



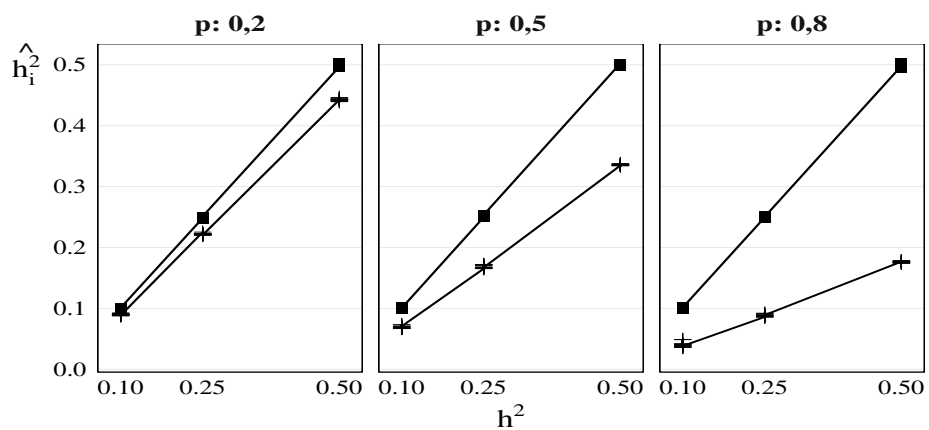
Fonte: Do autor (2016).

Figura A.2 – Distribuição da razão entre a variância aditiva da população e a variância de progênie, estimada pela análise de variância ao longo das simulações, com uma planta por parcela.



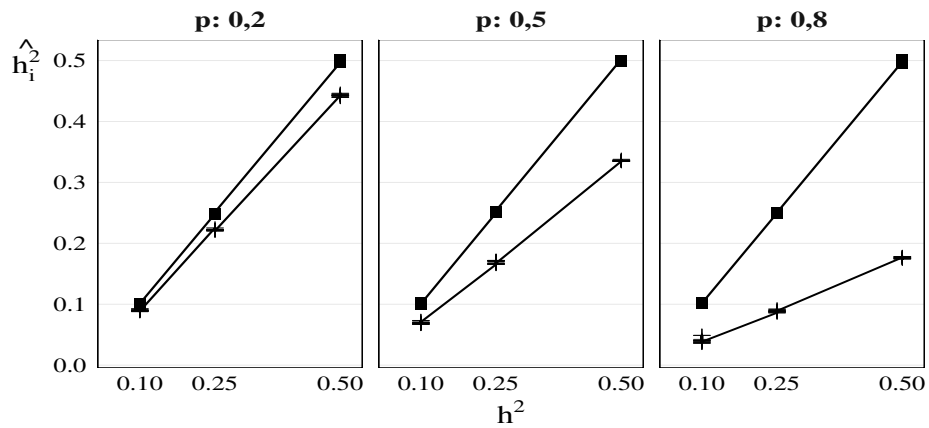
Fonte: Do autor (2016).

Figura A.3 – Herdabilidade em nível de indivíduo no sentido restrito ( $\hat{h}_i^2$ ), estimada considerando-se diferentes frequências alélicas ( $p$ ), para valores de  $h^2$  pré-fixados, com  $gmd = 0$  (■) e  $gmd = 1$  (+), com múltiplas plantas por parcela.



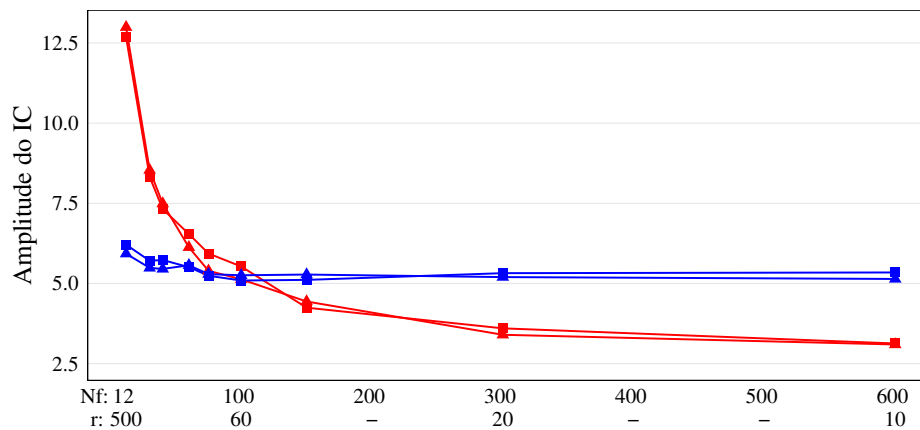
Fonte: Do autor (2016).

Figura A.4 – Herdabilidade em nível de indivíduo no sentido restrito ( $\hat{h}_i^2$ ), estimada considerando-se diferentes frequências alélicas ( $p$ ), para valores de  $h^2$  pré-fixados, com  $gmd = 0$  (■) e  $gmd = 1$  (+), com uma planta por parcela.



Fonte: Do autor (2016).

Figura A.5 – Amplitude do intervalo de confiança (IC) de 95% da variância de progênie (linha vermelha) e da variância do erro (linha azul), em experimentos com uma plantas por parcela, com  $h^2 = 0,5$ , sem dominância e frequências alélicas 0,2 (■) e 0,8 (▲).



Fonte: Do autor (2016).