

## NÚMERO MÍNIMO DE FAMÍLIAS DE MEIOS-IRMÃOS PARA REPRESENTAR UMA POPULAÇÃO DE MILHO-PIPOCA<sup>1</sup>

Maria José Granate<sup>2</sup>  
Cosme Damião Cruz<sup>3</sup>  
Cleso Antônio Patto Pacheco<sup>4</sup>

### RESUMO

O número mínimo de famílias de meios-irmãos, no melhoramento genético, é aquele que representa uma população, permitindo obter estimativas consistentes e estáveis dos parâmetros genéticos. A média e a variância genética foram escolhidas para caracterizar a população de 166 famílias de meios-irmãos de milho-pipoca CMS 43, avaliada em Sete Lagoas e em Coimbra, MG, no delineamento em blocos casualizados, com duas repetições, em 1997-1998. A técnica de simulação *bootstrap* foi utilizada para se obterem as informações necessárias. A determinação do número mínimo é feita pela análise visual dos gráficos da estabilização da média e da variância de subamostras. Na característica prolificidade foram necessárias 20 famílias; na característica produção de grãos/ha, 161; e na capacidade de expansão, 141. Este é o número mínimo de famílias considerado adequado para representar esta população. O método de simulação de subamostras proposto contribui para reduzir os trabalhos e os custos da pesquisa agrônômica, mantendo a precisão e a exatidão desejadas.

Palavras-chaves: *Zea mays*, método de simulação, melhoramento.

<sup>1</sup> Aceito para publicação em 11.02.2001.

<sup>2</sup> EPAMIG/CTZM, Vila Gianetti, 46, 36571-000 Viçosa, MG. granate@correio.ufv.br

<sup>3</sup> Departamento de Biologia Geral, Universidade Federal de Viçosa, 36571-000 Viçosa, MG. cdacruz@mail.ufv.br

<sup>4</sup> Embrapa/CNPMS. Caixa Postal 151, 35701-970 Sete Lagoas, MG. cleso@cnpmis.embrapa.br

**ABSTRACT****MINIMUM HALF-SIB FAMILY NUMBER REPRESENTING A  
POPCORN POPULATION**

The minimum number of half-sib families to be used in breeding is a number that represents a population, allowing consistent and stable estimates of the genetic parameters. Mean and genetic variance were chosen to characterize a population of 166 half-sib families of popcorn CMS 43, evaluated at Sete Lagoas and Coimbra, Minas Gerais. A randomized complete - block design was used, with two replications, in 1997-1998. The bootstrap simulation technique was used to generate information to study the minimum number of half-sib families. The minimum number of families is determined by visually analyzing the graphics with the estimates of the sub samples mean and variance. For the trait prolificity 20 families were needed for stable mean and variance estimates; for the trait grain yield/ha, 161 families; and for the trait expansion volume, 141. This is the minimum number acceptable to represent this population. The proposed sub sampling simulation method helps reduce costs and agricultural research, while maintaining desired precision and accuracy.

Key words: *Zea mays*, simulation method, breeding.

**INTRODUÇÃO**

O milho (*Zea mays* L.) é um dos cereais mais importantes na alimentação humana. Embora sendo a planta submetida à maior intensidade de domesticação e seleção, ainda não se esgotou a sua variabilidade, sendo possível continuar o melhoramento. Existem centenas de variedades de milho que incluem o milho comum e os vários tipos especiais de milho, entre eles o milho-pipoca. As suas sementes são pequenas, com pericarpo duro envolvendo um endosperma com teores de óleo e umidade relativamente elevados, os quais, sob a ação do calor (177 – 180 °C), rebentam a película do pericarpo, formando a pipoca, muito apreciada pelo seu sabor característico (9, 17). Considerado primitivo, o milho-pipoca teria sido o primeiro a ser consumido, estourado em fogueiras antes da invenção das panelas (14).

As famílias de meios-irmãos (FMI) são muito utilizadas na seleção recorrente do milho. É indispensável que elas sejam avaliadas com a máxima exatidão e precisão para que as estimativas sejam confiáveis e também que representem adequadamente a população estudada para que as inferências sejam válidas, porque nas várias repetições não figuram os mesmos indivíduos. A exatidão será tanto maior quanto mais elevado o número de famílias, e a precisão aumenta com o número de repetições; entretanto, os recursos humanos e financeiros necessários para a condução dos trabalhos serão também proporcionais ao número de famílias e de repetições presentes. Assim, é de interesse para os melhoristas a

determinação do número mínimo de famílias a serem avaliadas e que represente as características genéticas de uma população.

Para fins de melhoramento genético, o tamanho mínimo da amostra foi definido por Wu et al. (16) como o número mínimo de elementos necessários para estimar a média e a variância de um dos caracteres da população, com precisão razoável. Uma amostra de 250 plantas com duas repetições foi considerada confiável para a estimação dos componentes da variância no milho (16), mas apenas 30 plantas de milho foram tidas como suficientes para representar uma população segregante, também na cultura do milho (16). Estreitamente relacionado com a dimensão da parcela, o número mínimo de plantas de milho por parcela estabelecido por Camacho Palomino (1) foi de 9,6, considerando o caráter de menor precisão.

O pesquisador poderá ter de escolher entre avaliar progênes numerosas, resultantes de um pequeno número de cruzamentos, ou avaliar muitos cruzamentos, cada um com poucos descendentes. Se poucos indivíduos permitirem avaliar satisfatoriamente determinado cruzamento, muitos cruzamentos poderão ser avaliados ao mesmo tempo (16). Um número menor de famílias que represente adequadamente uma população, preservando as suas propriedades genéticas, permitirá avaliar muitas populações simultaneamente. Quanto ao caráter rendimento da forragem fresca de *Agropyrum desertorum*, a estimativa da herdabilidade de 80%, usando 10 grupos de 30 FMI, com três repetições, foi considerada equivalente à de 82% obtida com 300 FMI (12). Também com relação ao rendimento de forragem fresca, mas de *Pascopyrum smithii*, nove grupos de 30 FMI, em dois ambientes, com duas repetições no primeiro e três no segundo, forneceram a estimativa da herdabilidade de 61,9%, quase idêntica à de 61,5% obtida com 270 FMI (13). Em conseqüência, nos dois casos, 30 FMI foram consideradas representativas da população. Em cana-de-açúcar consideraram-se necessários 40 indivíduos para representar cada população (16).

Consultaram-se 140 das 173 teses sobre melhoramento de milho existentes na Biblioteca Central da UFV, das quais 41 referem-se a FMI (7). Verificou-se que o número de famílias utilizadas variou de 50 a 1.200, consideradas *a priori* representativas das populações de inferência, sendo o número médio de 368 famílias. Embora as diferentes finalidades de cada tese e a diversidade dos recursos disponíveis possam explicar os diferentes números de famílias utilizados, constatou-se a ausência de uma metodologia que definisse o número mínimo de FMI a utilizar no melhoramento do milho.

Vários procedimentos estatísticos, desenvolvidos com outras finalidades, podem ser adaptados para o estudo do número mínimo de FMI que melhor represente uma população, para fins de melhoramento genético. Neste trabalho é apresentado um método empírico que permite estimar o número mínimo de FMI para representar uma população, no melhoramento do milho-pipoca, adaptando conceitos dos métodos de simulação, da técnica *bootstrap* e de metodologias desenvolvidas para o estudo do tamanho mínimo da parcela.

## MATERIAL E MÉTODOS

Utilizaram-se famílias de meios-irmãos do composto de milho CMS 43-Composto Pipoca Branco Tipo Americano, obtido pela recombinação de materiais considerados superiores e tolerantes às principais doenças, existentes no banco de germoplasma do Centro Nacional de Pesquisa de Milho e Sorgo. Desde 1979, a população derivada do composto foi submetida a mais de cinco ciclos de seleção massal estratificada e, a partir de 1988, a vários ciclos de seleção entre e dentro de famílias de meios-irmãos. A seleção foi a favor de boas características agrônômicas, em especial resistência ao acamamento, a doenças e pragas, cor do grão e da pipoca, dimensões dos grãos, tipo de grão e forma da pipoca (3, 10, 11). Os ensaios foram conduzidos no ano agrícola de 1997-1998 nos campos experimentais do Centro Nacional de Pesquisa de Milho e Sorgo da Empresa Brasileira de Pesquisa Agropecuária, em Sete Lagoas, e da Universidade Federal de Viçosa, em Coimbra.

O delineamento foi em blocos casualizados, com 179 tratamentos, dos quais 166 famílias de meios-irmãos do composto CMS 43 que se pretendiam avaliar e 13 tratamentos com testemunhas adicionais, em duas repetições. Entre as testemunhas estavam representadas a variedade Zélia e a população de milho CMS 43 do ciclo anterior. Cada parcela constou de uma fileira de 5 m, com 50 sementes por parcela, em 25 covas espaçadas de 0,20 m. O espaçamento entre fileiras foi de 0,90 m. O plantio foi efetuado manualmente, e o estande ideal após o desbaste seria de 25 plantas por fileira. Os tratos culturais e as adubações foram os habituais na cultura do milho. As medições dos seguintes caracteres avaliados referem-se a cada parcela e foram feitas na época da colheita: CE – capacidade de expansão, razão entre o volume de pipoca expandida e o volume de grãos antes do pipocamento: uma amostra de 30 ml de grãos foi tomada de cada parcela, medida em proveta graduada, estourada em pipoqueira elétrica, regulada automaticamente em 237 °C, o volume da pipoca expandida foi medido em proveta graduada de 1.000 ml; PRODU – produtividade, peso dos grãos debulhados por unidade de área, expresso em kg/ha; ALTP – altura de planta, média das medições feitas do nível do solo à inserção da folha bandeira, em seis plantas competitivas tomadas ao acaso, em cm; ALTES – altura da espiga, média das distâncias do nível do solo até à inserção da espiga, nas mesmas seis plantas, em cm; NESP – número de espigas colhidas; e PROLI – prolificidade, razão entre o número de espigas colhidas e o número de plantas existentes.

Para a determinação do número mínimo de famílias, pelo método *bootstrap* (4) de simulação de subamostras, iniciou-se a primeira rodada com subamostras de 20 famílias, por ser inferior ao menor número de famílias encontrado na literatura para representar uma população (7). O aumento do tamanho das subamostras ou incremento, de uma rodada para a outra, foi de uma família, isto é, o menor possível. O número de famílias

foi sucessivamente aumentado até se atingir o total de 166. Para cada tamanho de subamostra fizeram-se 10 rodadas, de forma aleatória, com reposição. Em cada subamostra foi feita a análise de variância e estimados os parâmetros genéticos, obtendo-se em seguida a média das 10 subamostras do mesmo tamanho. Assim, foram realizadas 8.754 análises de variância. Camacho Palomino (1), no seu trabalho de simulação de subamostras para estudo do tamanho da parcela, apresenta os resultados de 270 análises de variância.

Neste trabalho, por analogia com o Método da Curvatura Máxima (5, 15), da Curvatura Máxima Modificado (6) e o de Wu (16), representaram-se graficamente as estimativas dos parâmetros obtidos a partir das análises de variância das subamostras, em função do tamanho das respectivas subamostras. A média e a variância genética, porque melhor expressam as características genéticas de uma população de interesse para o melhorista, foram utilizadas no eixo das ordenadas, e o número de famílias, no eixo das abscissas. Verificou-se que, a partir de determinado número de famílias, os valores dos referidos parâmetros se estabilizavam, considerando que, a partir desse ponto, o aumento do número de famílias presentes nas subamostras não provocava mudanças importantes nas estimativas dos parâmetros. Deste modo, o número mínimo de famílias para representar uma população foi determinado visualmente como o ponto a partir do qual a estimativa do parâmetro escolhido tornava-se estável. A determinação visual do ponto de estabilização das médias dos parâmetros é uma avaliação subjetiva e, como tal, sujeita a variar com o observador. Neste trabalho, quando era necessário escolher entre dois pontos de estabilização, optou-se pelos valores mais elevados, porque conferiam maior precisão e exatidão.

Foram determinadas as retas dos limites superiores e inferiores da média mínima não significativa (LMS), pelo teste t, tomando-se como referência a média obtida com as 166 famílias, por ser considerado um critério auxiliar útil.

Aplicaram-se as equações empíricas, obtidas por Lessman e Atkins (6) e por Meier e Lessman (8), para determinar os pontos que correspondem ao tamanho ideal de amostra, segundo esses autores.

O aplicativo computacional GENES (2), desenvolvido na UFV, foi utilizado para realizar todas as computações.

## RESULTADOS E DISCUSSÃO

Os valores obtidos em relação ao número mínimo de indivíduos, utilizando as expressões de Lessman e Atkins (6) e de Meier e Lessman (8), costumam ser muito semelhantes (8). No entanto, os obtidos neste

trabalho, quanto ao número mínimo de famílias, pela aplicação das referidas expressões, diferem muito entre si. Além disso, quanto aos caracteres CE, ALTP, NESP e PROLI, são fornecidos valores inferiores ao menor tamanho da subamostra utilizada e até inferior à unidade. Em relação ao caráter PRODU, os valores estimados pelas referidas expressões são exageradamente altos, muito superiores ao total de famílias avaliadas (7).

As médias da característica CE em 1.459 subamostras de vários tamanhos são apresentadas na Figura 1. Onze pontos, representando 110 subamostras, estão fora dos LMS em relação à média obtida com as 166 famílias. Após o último desses 11, o primeiro tamanho de amostra que fica entre os dois LMS tem 45 famílias. O valor da média da característica CE parece estabilizado a partir de subamostras com 121 famílias.

O gráfico que representa as médias da variância genética,  $V_g$ , do caráter CE nas subamostras, é apresentado na Figura 2. Verifica-se que 12 pontos, correspondentes a 120 subamostras, estão fora dos LMS e que o primeiro ponto dentro dos LMS, depois desses 12, representa 10 subamostras com o tamanho de 40 famílias. A média da variância genética pode considerar-se visualmente estabilizada a partir das subamostras com 141 famílias de tamanho.

Assim, em relação à característica CE, este método de simulação de subamostras fornece os valores 40, 45, 121 e 141. Considerando que um número maior de famílias confere maior precisão, escolhe-se o tamanho de 141 famílias para representar a população. Quando comparado com os números de famílias encontrados na literatura, verifica-se que é inferior ao tamanho médio de 368 famílias.

As médias de 1.459 subamostras, obtidas com a subdivisão das 166 famílias, em relação ao caráter PRODU, são apresentadas na Figura 3. Verifica-se que 10 círculos, representando 100 subamostras, estão fora dos LMS. Após o último desses 10, o primeiro tamanho de amostra que fica dentro dos LMS corresponde a 10 amostras com 86 famílias de tamanho. Visualmente, o valor da média da característica PRODU nas subamostras parece estabilizado a partir das 10 amostras cujo tamanho é de 148 famílias.

Na Figura 4, apresenta-se o gráfico que representa as médias da variância genética,  $V_g$ , do caráter PRODU nas subamostras. Verifica-se que apresenta 14 círculos, que correspondem a 140 subamostras, fora dos LMS. O primeiro círculo entre as retas dos LMS, depois do último desses 14, corresponde a 10 subamostras com 64 famílias cada uma. Este gráfico mostra grande oscilação entre os pontos representados após o valor 64, pelo que o número de 64 famílias não é considerado suficiente para representar esta população. O ponto de estabilização visual das médias do parâmetro  $V_g$  nas subamostras parece ser a partir de 161 famílias. Este valor é considerado preferível para representar a população.

Por estes processos obtiveram-se 64, 86, 148 e 161 famílias para representar esta população de milho, em relação ao caráter PRODU.

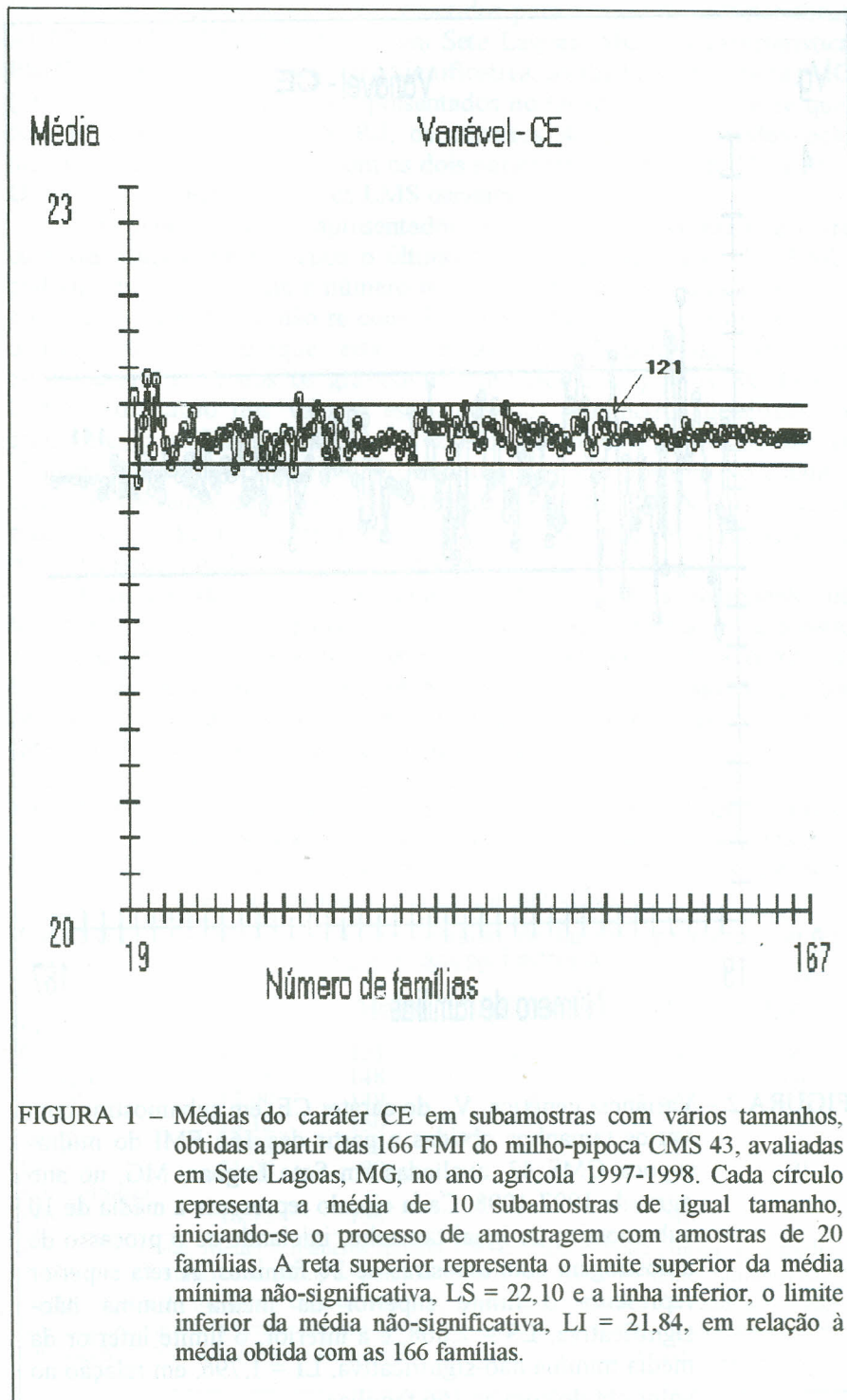


FIGURA 1 – Médias do caráter CE em subamostras com vários tamanhos, obtidas a partir das 166 FMI do milho-pipoca CMS 43, avaliadas em Sete Lagoas, MG, no ano agrícola 1997-1998. Cada círculo representa a média de 10 subamostras de igual tamanho, iniciando-se o processo de amostragem com amostras de 20 famílias. A reta superior representa o limite superior da média mínima não-significativa, LS = 22,10 e a linha inferior, o limite inferior da média não-significativa, LI = 21,84, em relação à média obtida com as 166 famílias.

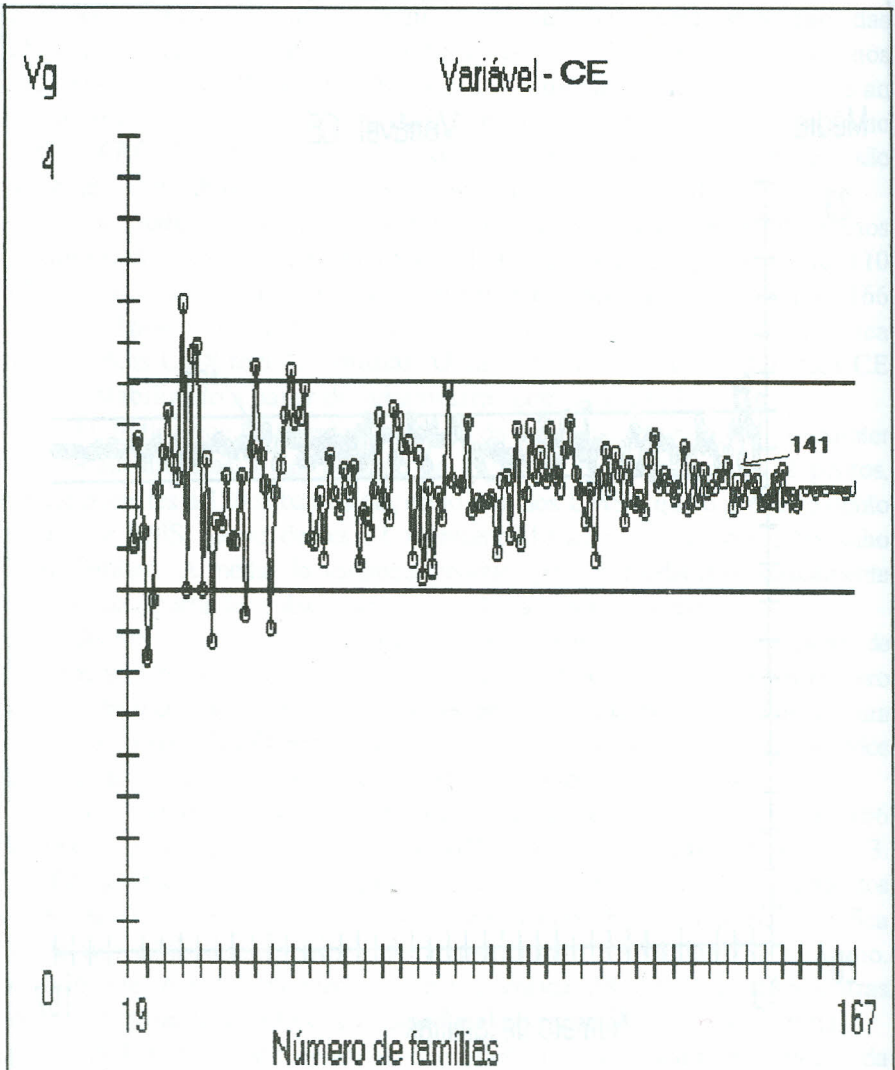


FIGURA 2 – Variância genética,  $V_g$ , do caráter CE, em subamostras com vários tamanhos obtidas a partir das 166 FMI do milho-pipoca CMS 43, avaliadas em Sete Lagoas, MG, no ano agrícola 1997-1998. Cada círculo representa a média de 10 subamostras de igual tamanho, iniciando-se o processo de amostragem com amostras de 20 famílias. A reta superior representa o limite superior da média mínima não-significativa,  $LS = 2,808$ , e a inferior, o limite inferior da média mínima não-significativa,  $LI = 1,796$ , em relação ao valor obtido com as 166 famílias.



Estes procedimentos foram repetidos para avaliar as características ALTP, ALTES, NESP e PROLI, em Sete Lagoas, MG, e a característica PRODU, a única estatisticamente significativa, avaliada em Coimbra, MG (7). Os resultados obtidos são apresentados no Quadro 1. Verifica-se que, com exceção do caráter PROLI, os números de famílias obtidos pelo método visual de estimação, com os dois parâmetros, variam de 121 a 161. Os valores encontrados com os LMS oscilam entre 32 e 77.

Nos quatro gráficos apresentados, o primeiro círculo que fica entre os LMS, imediatamente após o último círculo que fica além dos LMS, poderia ser tomado como o número mínimo de famílias para representar a população. No entanto, não se considera indicado usar o valor obtido com o uso dos LMS porque está relativamente afastado do início da estabilização, em todos os gráficos. Entre esses dois pontos, verifica-se bastante oscilação dos valores estimados dos parâmetros genéticos. A estabilização indica que as estimativas dos parâmetros genéticos não se alteram com o aumento do número de famílias nas amostras. No entanto, como a diferença entre o valor obtido utilizando os LMS e o valor estabilizado não é significativa, esse valor constitui um número de referência que o melhorista poderá utilizar.

Considerando que a característica CE é a mais importante no melhoramento do milho-pipoca e que a variância genética é parâmetro indispensável no melhoramento genético, o uso de 141 famílias para representar a população desse milho é considerado o menor número adequado, porque permite a estimação estável dos dois principais parâmetros genéticos dessa característica, mantendo a precisão e a exatidão desejadas.

QUADRO 1 – Valores estimados do número mínimo de famílias de meios-irmãos capaz de representar a população de milho-pipoca CMS 43, obtidos a partir das avaliações de 166 famílias de meios-irmãos, em Sete Lagoas e Coimbra, no ano agrícola 1997-1998.

Caracteres	$h^2$	Número mínimo de famílias para cada caráter, segundo dois parâmetros de avaliação			
		Média	LMS-média	$V_g$	LMS-variância
Em Sete Lagoas					
CE	46,10	121	45	141	40
PRODU	42,51	148	86	161	64
ALTP	17,90	149	32	150	47
ALTES	18,12	150	51	137	72
NESP	54,90	124	52	126	33
PROLI	51,08	20	24	20	20
Em Coimbra					
PRODU	18,72	158	77	148	52

$h^2$  – herdabilidade, Média – números mínimos de famílias obtidos utilizando a média, LMS-média – números mínimos de famílias obtidos, utilizando os limites mínimos não-significativos da média da característica,  $V_g$  – números mínimos de famílias obtidos, utilizando a variância genética, LMS-variância – números mínimos de famílias obtidos, utilizando os limites mínimos não-significativos da variância genética da característica.

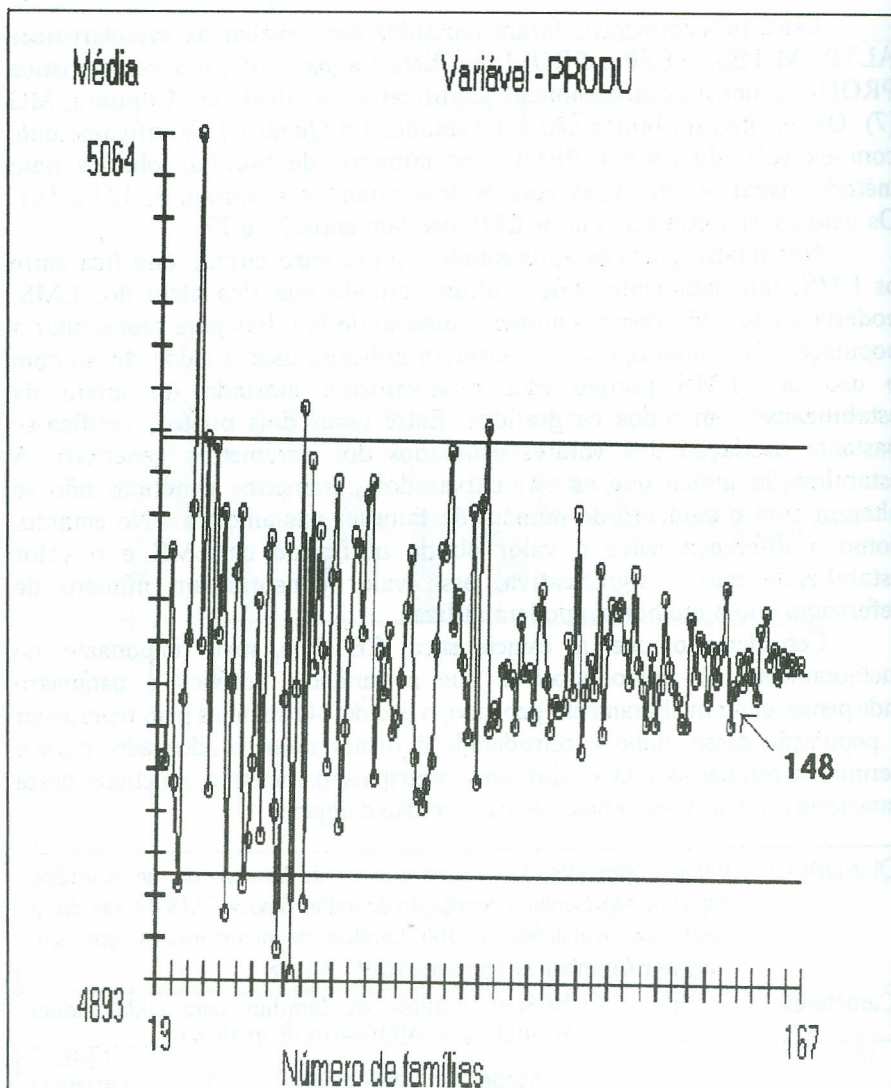
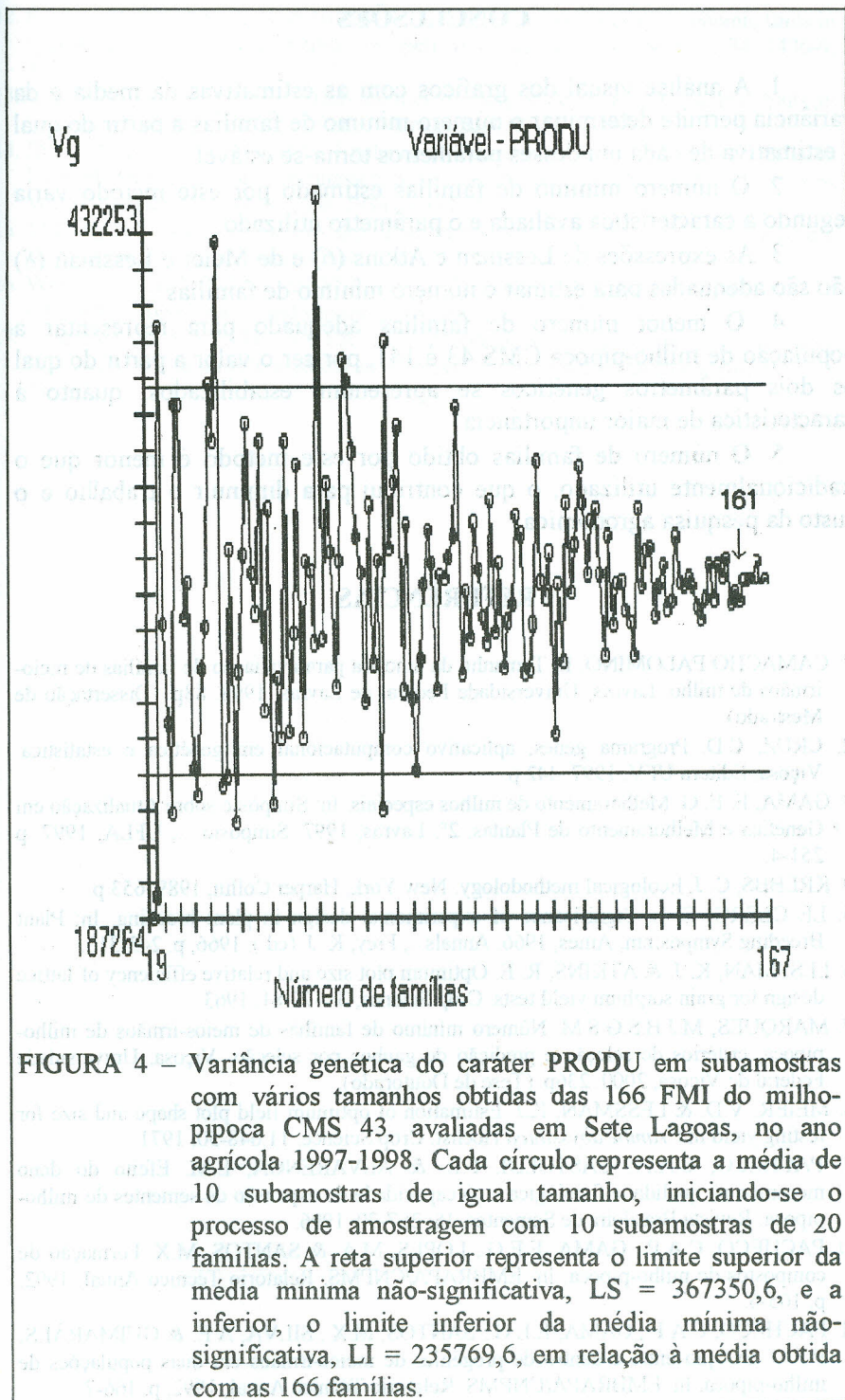


FIGURA 3 – Médias do caráter PRODU em subamostras com vários tamanhos obtidas a partir das 166 FMI do milho-pipoca CMS 43, avaliadas em Sete Lagoas, MG, no ano agrícola 1997-1998. Cada círculo representa a média de 10 subamostras com igual número de famílias, iniciando-se o processo de amostragem com amostras de 20 famílias. A reta superior representa o limite superior da média mínima não-significativa,  $LS = 5002,70$ , e a inferior, o limite inferior da média mínima não-significativa,  $LI = 4914,43$ , em relação à média obtida com as 166 famílias.



## CONCLUSÕES

1. A análise visual dos gráficos com as estimativas da média e da variância permite determinar o número mínimo de famílias a partir do qual a estimativa de cada um desses parâmetros torna-se estável.
2. O número mínimo de famílias estimado por este método varia segundo a característica avaliada e o parâmetro utilizado.
3. As expressões de Lessman e Atkins (6) e de Meier e Lessman (8) não são adequadas para estimar o número mínimo de famílias.
4. O menor número de famílias adequado para representar a população de milho-pipoca CMS 43 é 141, por ser o valor a partir do qual os dois parâmetros genéticos se apresentam estabilizados, quanto à característica de maior importância.
5. O número de famílias obtido por este método é menor que o tradicionalmente utilizado, o que contribui para diminuir o trabalho e o custo da pesquisa agrônômica.

## REFERÊNCIAS

1. CAMACHO PALOMINO, E. Tamanho da amostra para avaliação de famílias de meio-irmãos de milho. Lavras, Universidade Federal de Lavras, 1998. 88p. (Dissertação de Mestrado).
2. CRUZ, C.D. Programa genes, aplicativo computacional em genética e estatística. Viçosa, Editora UFV, 1997. 442 p.
3. GAMA, E. E. G. Melhoramento de milhos especiais. In: Simpósio sobre Atualização em Genética e Melhoramento de Plantas, 2º, Lavras, 1997. Simpósio ..., UFLA, 1997. p. 251-4.
4. KREBBS, C. J. Ecological methodology. New York, Harper Collin, 1989. 653 p.
5. LE CLERG, E. L. Significance of experimental design in plant breeding. In: Plant Breeding Symposium, Ames, 1966. Annals..., Frey, K. J. (ed.). 1966, p. 243-313.
6. LESSMAN, K. J. & ATKINS, R. E. Optimum plot size and relative efficiency of lattice design for grain sorghum yield tests. Crop Science, 3: 477-81. 1963.
7. MARQUES, M.J.B.S.G.S.M. Número mínimo de famílias de meios-irmãos de milho-pipoca, critérios de seleção e predição de ganhos por seleção. Viçosa, Universidade Federal de Viçosa, 2000. 236p. (Tese de Doutorado).
8. MEIER, V.D. & LESSMAN, K.J. Estimation of optimum field plot shape and size for testing yield in *Crambe abyssinica* Hochst. Crop Science, 11:648-50, 1971.
9. PACHECO, C.A.P.; CASTOLDI, F.L. & ALVARENGA, E.M. Efeito do dano mecânico na qualidade fisiológica e na capacidade de expansão de sementes de milho-pipoca. Revista Brasileira de Sementes, 18: 267-70, 1996.
10. PACHECO, C.A.P.; GAMA, E.E.G.; LOPES, M.A. & SANTOS, M.X. Formação de compostos de milho-pipoca. In: EMBRAPA/CNPMS. Relatório Técnico Anual, 1992, p. 165-6.
11. PACHECO, C.A.P.; GAMA, E.E.G.; SANTOS, M.X., SILVA, A.F. & GUIMARÃES, P.E.O. Seleção entre e dentro de progênies de meios-irmãos em duas populações de milho-pipoca. In: EMBRAPA/CNPMS. Relatório Técnico Anual, 1992, p. 166-7.

12. RAY, I. M.; FRANK, A. B. & BERDHL, J. D. Genetic variances of agronomic traits in tetraploid crested wheatgrass under competitive conditions. *Crop Science*, 34: 1436-9, 1994.
13. RAY, I. M. & HARMS, J. P. Heritabilities of morphological and agronomic traits in western wheatgrass. *Journal of Range Management*, 47: 60-3, 1994.
14. RUST, W.F. & LEYDEN, B.W. Evidence of maize use at early and middle preclassic La Venta Olmec sites. In: Johannessen, S.(ed.). *Corn and culture in the Prehistoric New World*. Boulder, WestView Press, 1994. p. 181-201.
15. SMITH, H.F. An empirical law describing heterogeneity in the yields of agricultural crops. *Journal of Agricultural Science*. 28: 1-23, 1938.
16. WU, K. W.; HEINZ, D. J.; MEYER, H. K. & LADD, S. L. Minimum sample size for estimating progeny mean and variance. *Crop Science*, 18: 57-62, 1978.
17. ZINSLY, J.R. & MACHADO, J.A. 1987. Milho-pipoca. In: Paterniani, E. & Viégas, G.P. (eds). *Melhoramento e produção do milho*. Campinas, Fundação Cargill, 1987. p. 413-21.