

APLICAÇÃO DA METODOLOGIA DE MODELOS MISTOS (REML/BLUP) NA ESTIMAÇÃO DE COMPONENTES DE VARIÂNCIA E PREDIÇÃO DE VALORES GENÉTICOS EM PUPUNHEIRA (*Bactris gasipaes*)¹

JOÃO TOMÉ DE FARIAS NETO², MARCOS DEON VILELA DE RESENDE³

RESUMO - No Brasil, a pupunheira é uma planta muito útil na alimentação, seja como fonte de frutos ou de palmito. O interesse pela pupunheira, além de ser uma cultura perene, são: crescimento a pleno sol, precocidade, rusticidade, perfilhamento, palatabilidade e não-escurecimento do palmito após o corte. Estimativas de parâmetros genéticos em pupunheira são escassas e se constituem em ferramenta de suma importância para orientar os programas de melhoramento. O objetivo deste trabalho foi estudar a variabilidade genética e estimar o valor genético individual como critério de seleção, usando o procedimento BLUP/REML (Melhor predição linear não viciada/máxima verossimilhança restrita). Adotaram-se duas estratégias de seleção para o caráter produção de palmito: simulando programa de melhoramento a curto prazo (CP- seleção das 9 famílias com 31 indivíduos de maior valor genético) e a longo prazo (LP- seleção das 15 famílias com 53 indivíduos). As progênies foram avaliadas em experimento delineado em blocos ao acaso com três repetições, parcelas lineares de cinco plantas, espaçadas de 2,0 m x 1,0 m e bordadura composta por uma fileira em torno do experimento no Campo Experimental do Matapi, Município de Porto Grande, Estado do Amapá. A avaliação foi realizada aos 26 meses pós- plantio (2ª avaliação), coletando-se dados de altura da planta (AP), diâmetro da planta à altura do colo (DPC), tamanho do palmito (TP), diâmetro do palmito (DP), peso do residuo apical (PRA), basal (PRB) e do palmito líquido (PP) (tipo exportação). Os dados de AP, DPC, TP e DP corresponderam às médias das touceiras que apresentavam mais de uma haste. Já para os caracteres PRA, PRB e PP corresponderam à soma das hastes na touceira. De maneira geral, a população estudada apresenta baixa variabilidade genética. As herdabilidades no sentido restrito em nível de indivíduos foram: AP (18,44%), DPC (3,16%), TP (42,47%), DP (10,54%), PP (5,70%), PRB (6,15%). Os ganhos genéticos preditos em relação à média da população para PP foram de 7,18% na situação de LP e 8,40% para CP, com tamanho efetivo de 30,38 e 19,00, respectivamente.

Termos para indexação: modelos lineares, tamanho efetivo, predição de variáveis aleatórias, espécie perene, genética quantitativa.

APPLICATION OF THE MIXED MODEL METHODOLOGY (REML/BLUP) IN VARIANCE COMPONENTS ESTIMATION AND PREDICTION OF GENETIC VALUES IN PEACH PALM (*Bactris gasipaes*)

ABSTRACT- The peach palm is a very useful plant for feeding Brazilians as fruit or palm heart producer. The interest for the peach palm besides being a perennial culture is: growth in full sun, precocity, rusticity, capacity to shoot, flavor and non-darkening of the palm heart after the cut. Estimates of genetic parameters in peach palm are scarce and constitute the most important tool to guide the improvement programs. The objective of this work was to study the genetic variability and estimate the individual genetic value as selection criterion, using the BLUP/REML procedure (Best linear unbiased prediction/restricted maximum likelihood). Two selection strategies for the palm heart production trait were adopted: a short term (CP - selection of the 9 families with 31 individuals of bigger genetic value) and a long term (LP - selection of the 15 families with 53 individuals). The progenies were evaluated in randomized block design with three replications, the plots were composed by rows of five plants, spaced in 2.0 m x 1.0 m and with a row around the experiment in the Experimental Field of Matapi, Porto Grande municipality, Amapa State, Brazil. The evaluation was accomplished to the 26 months after planting (2nd evaluation) being collected data of plant height (AP), diameter of the plant to the lap height (DPC), palm heart size (TP), palm heart diameter (DP), residual apical weight (PRA), basal weight (PRB) and of the liquid palm heart (PP) (exportation type). The data of AP, DPC, TP and DP corresponded to the clump of roots averages that presented more than a stem. However for the characters PA, PRB and PP corresponded the sum of the stems in the clump of roots. In general, the population presented low genetic variability. The narrow sense heritability at the individuals level was: AP (18.44%), DPC (3.16%), TP (42.47%), DP (10.54%), PP (5.70%), PRB (6.15%). The genetic gain estimated in relation to average of the population for PP were of 7.18% in the LP situation and 8.40% for CP, with effective size of 30.38 and 19.00, respectively.

Index terms: linear models, effective size, prediction of random variables, perennial species, quantitative genetics

¹ Trabalho nº 152/2000. Recebido: 21/07/2000. Aceito para publicação: 17/04/2001.

² Eng.º Agr.º, Dr., Pesquisador da Embrapa Amazônia Oriental, Tv. Dr. Eneas Pinheiro s/n. C.P. 48, Belém, PA.

³ Eng.º Agr.º, Dr. Pesquisador da Embrapa Floresta. deon@cnpf.embrapa.br

INTRODUÇÃO

O maior uso da pupunha por habitantes da Amazônia é do fruto, que é consumido diretamente após o cozimento em água com sal ou no preparo de várias iguarias. Os frutos são também utilizados para ração animal e para obtenção de farinha, que é utilizada na fabricação de pães e bolos. Porém, o palmito é o principal motivo do plantio da pupunha em larga escala atualmente (Souza & Silva, 2000). O conhecimento da variabilidade genética associada à principal população de pupunha usada comercialmente no Brasil é de suma importância, tanto para os plantios visando à produção de frutos, quanto à de palmito.

O interesse pela pupunheira, além de ser uma cultura perene, são: crescimento a pleno sol, precocidade, rusticidade, perfilhamento, palatabilidade e não-escurecimento do palmito após o corte (Bovi, 1998; Toney et al., 1999). O aumento da produtividade de palmito em cada ciclo de seleção constitui a meta principal dos programas de melhoramento, necessitando para isso extensivas e demoradas avaliações de progênies em condições de campo. Portanto, estimativas precisas dos componentes de variância são importantes para a predição de valores genéticos e para maximizar a acurácia da seleção em programas de melhoramento da pupunheira. Neste contexto, métodos diferenciados de estimação e predição são necessários, em função das diferentes situações experimentais e do balanceamento associados aos dados experimentais (Resende et al., 1993-1996).

De maneira genérica, o procedimento ótimo de estimação/predição no melhoramento de espécies perenes é o REML/BLUP (máxima verossimilhança restrita/melhor predição linear não viciada). Entretanto, para o caso de dados balanceados, a estimação de componentes de variância pelo método de quadrados mínimos (análise de variância) equivale à estimação por REML (Resende et al., 1996), e a predição de valores genéticos pelo método do índice multi-efeitos (Resende & Higa, 1994) equivale ao BLUP individual, conforme demonstrado por Resende & Fernandes (1999).

Assim, os procedimentos ótimos de estimação/predição podem ser resumidos em dois: (i) análise de variância/índice multi-efeitos para o caso balanceado; (ii) REML/BLUP para os casos desbalanceado e balanceado. No caso de dados balanceados, os dois procedimentos conduzem a resultados idênticos; em casos com pequeno desbalanceamento, os dois procedimentos conduzem a resultados similares e, na presença de grande desbalanceamento, o procedimento (ii) é muito superior ao (i).

Estimativas de parâmetros genéticos em pupunheira são escassas. Farias Neto (1999) relata as seguintes estimativas para a herdabilidade individual no sentido restrito: 0,37; 0,30; 0,35; 0,21 e 0,30 para os caracteres altura da planta, diâmetro do colo, comprimento do palmito, diâmetro do palmito e peso do palmito líquido, respectivamente, aos 15 meses de idade. Clement (1995) e Clement & Bovi (1999) relatam as seguintes estimativas de herdabilidade no sentido restrito: 0,27 a 0,36 para a presença de espinhos no pecíolo/ráquis; 0,49 para a taxa de crescimento relativo da fase de viveiro ao primeiro corte; 0,11 para a razão de área foliar. Estimativas de parâmetros genéticos para outros caracteres ou para os mesmos caracteres em outras idades de avaliação são essenciais para o direcionamento dos programas de melhoramento da espécie.

Com base no exposto, o presente trabalho teve como objetivos a estimação de componentes de variância, herdabilidade e predição de valores genéticos para vários caracteres em teste de progênies de pupunheira. Visou também à comparação entre os procedimentos REML/BLUP e análise de variância no processo de estimação.

MATERIAL E MÉTODOS

Material experimental

O teste de progênies de meios-irmãos de pupunheiras foi instalado em janeiro de 1998 no Campo Experimental do Matapi, Município de Porto Grande-AP, pertencente ao Centro de Pesquisa Agroflorestal do Amapá - CPAF/AP. As 31 progênies são oriundas do Centro Nacional de Pesquisa da Amazônia Ocidental - CPAA, Manaus. As sementes foram obtidas a partir de matrizes sem espinhos, e a base genética original da população é estreita (Clement & Bovi, 1999). A área apresenta topografia plana, cobertura com vegetação de capoeira, clima tipo Ami, segundo a classificação de Köppen, com temperatura média anual de 27°C, umidade relativa do ar média de 82% e precipitação média anual de 2.700 mm. As características químicas do solo da área experimental são apresentadas na Tabela 1.

As progênies foram avaliadas em experimento delineado em blocos ao acaso, com três repetições, parcelas lineares de cinco plantas, espaçadas de 2,0 m x 1,0 m e bordadura composta por uma fileira em torno do experimento. Avaliadas aos 26 meses (2ª avaliação), foram obtidos dados de altura da planta-AP (medida do solo até o ponto de inserção da folha guia e a primeira folha expandida), diâmetro da planta à altura do colo (DPC), tamanho do palmito- TP (soma do número de toletes de 9 cm), diâmetro do palmito- DP (medido na porção mediana do palmito líquido) e pesos do residuo apical (PRA), basal (PRB) e do palmito líquido- PP (palmito tipo exportação). Os dados obtidos caracterizaram-se como desbalanceados.

Estimativas de componentes de variância e predição de valores genéticos:

Na situação de dados desbalanceados, o modelo linear misto adequado para a descrição dos dados equivale a:

$$y = Xf + Za + Wc + e, \text{ em que:}$$

y , f , a , c , e – vetores de dados, de efeitos fixos (médias de blocos), de efeitos aditivos (aleatório), de efeitos de parcelas (efeitos aleatórios de ambiente comum das parcelas) e de erros aleatórios, respectivamente.

X , Z e W – são matrizes de incidência conhecidas, formadas por valores 0 e 1, as quais associam as incógnitas f , a e c ao vetor de dados y , respectivamente.

A metodologia de modelos mistos permite estimar f pelo procedimento de quadrados mínimos generalizados e predizer a e c pelo procedimento BLUP. Para obtenção destas soluções, basta resolver o seguinte sistema de equações lineares, o qual é denominado equações de modelo misto (MME):

$$\begin{bmatrix} \hat{f} \\ \hat{a} \\ \hat{c} \end{bmatrix} = \begin{bmatrix} X'X & X'Z & X'W \\ Z'X & Z'Z + A^{-1}\lambda_1 & Z'W \\ W'X & W'Z & W'W + I\lambda_2 \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ Z'y \\ W'y \end{bmatrix}$$

em que:

$$\lambda_1 = \frac{1-h^2-c^2}{h^2} = \frac{\sigma_e^2}{\sigma_a^2}; \quad \lambda_2 = \frac{1-h^2-c^2}{c^2} = \frac{\sigma_e^2}{\sigma_c^2}$$

A e I = matrizes de parentesco genético aditivo e matriz identidade de ordem apropriada aos dados, respectivamente.

Atribuindo-se valores iniciais para os componentes de variância nas MME, obtêm-se as predições para os efeitos a e c. Calculando-se as variâncias desses efeitos preditos, obtêm-se as estimativas de variâncias $\hat{\sigma}_a^2$ (variância genética aditiva) e $\hat{\sigma}_c^2$ (variância entre parcelas), as quais, provavelmente, serão diferentes dos valores iniciais utilizados nas MME, significando que os valores iniciais não foram plausíveis ou verossímeis. Desta forma, deve-se resolver novamente as MME, usando estes componentes de variância calculados. Procedendo-se sucessivamente desta maneira, atinge-se a convergência para os componentes de variância, ou seja, têm-se que os valores utilizados nas MME equivalem às próprias variâncias dos efeitos preditos, significando que os valores utilizados nas MME passaram a ser plausíveis ou verossímeis com o conjunto de dados. Este é o princípio da verossimilhança.

Os estimadores REML para obtenção das variâncias dos efeitos aleatórios, empregando-se o algoritmo EM (Expectation-Maximization) são:

$$\hat{\sigma}_e^2 = [y'y - \hat{f}'X'y - \hat{a}'Z'y - \hat{d}'Z'y - \hat{c}'W'y] / [N - r(x)]$$

$$\hat{\sigma}_a^2 = [\hat{a}'A^{-1}\hat{a} + \hat{\sigma}_e^2 \text{tr} A^{-1}C^{22}] / q$$

$$\hat{\sigma}_c^2 = [\hat{c}'c + \hat{\sigma}_e^2 \text{tr} C^{33}] / s,$$

tr = operador traço matricial;

r(x) = posto da matriz X;

N-r(x) = número de graus de liberdade do erro;

q = número de indivíduos;

s = número de parcelas;

N = número total de dados.

C^{22} e C^{33} provêm de:

$$C = \begin{bmatrix} C^{11} & C^{12} & C^{13} \\ C^{21} & C^{22} & C^{23} \\ C^{31} & C^{32} & C^{33} \end{bmatrix} = \text{inversa generalizada da matriz dos coeficientes das MME.}$$

As estimações e predições pelo procedimento REML/BLUP foram realizadas empregando-se o software DFREML (Meyer, 1998).

Para o método tradicional, baseado na análise de variância, empregou-se o software SELEGEN (Resende & Oliveira, 1997), o qual utiliza os estimadores da variância apresentados por Vencovsky & Barriga (1992).

Predição de ganhos genéticos

Os valores genéticos foram preditos pelo vetor \hat{a} , como desvios da média geral do experimento. Assim, o ganho genético predito foi estimado fazendo-se a média dos valores genéticos dos indivíduos selecionados. Com base nos valores genéticos individuais preditos, os ganhos genéticos preditos com a prática da seleção para peso de palmito foram estimados admitindo-se duas estratégias de seleção: 31 indivíduos (pertencentes a 9 progênies), representando a população de produção de sementes (simulando programa em curto prazo) e com 53 indivíduos (pertencentes a 15 progênies), para compor a população de melhoramento (simulando programa a médio e longo prazos). O tamanho efetivo (N_e) foi estimado pela expressão fornecida por Resende & Bertolucci (1995).

RESULTADOS E DISCUSSÃO

O número de famílias e indivíduos selecionados, intensidade de seleção, ganho genético esperado, média esperada e tamanho efetivo com o emprego das estratégias de seleção a curto e longo prazos são apresentados na Tabela 2. Verificam-se baixos ganhos genéticos esperados com valores semelhantes considerando as duas modalidades de seleção. A prática da seleção elevaria a média da população para o caráter peso de palmito após um ciclo de seleção de 227 g para 243 e 246 g para curto e longo prazos, respectivamente.

A determinação do número adequado de indivíduos a serem selecionados deve considerar a endogamia na geração de plantio, para a composição da população de produção de sementes e a manutenção de um tamanho populacional efetivo compatível com a obtenção do limite seletivo para a composição da população de melhoramento (Resende & Bertolucci, 1995). No presente trabalho, maximizou-se o ganho, preestabelecendo um tamanho efetivo mínimo de 19 e 30,38 para a composição das populações de produção de sementes e de melhoramento, respectivamente. Como as duas estratégias conduziram praticamente ao mesmo ganho genético e tendo em vista a baixa

TABELA 1- Características químicas do solo da área experimental

Profundidade de coleta	PH	K	Ca + Mg	Al	H	SB	CTC	MO	P
	H ₂ O								
					mmolc / dm ³			g / dm ³	mg / dm ³
00 - 20 cm	4,8	0,89	6,0	6,0	33,6	6,89	46,49	22	10

TABELA 2 - Número de famílias e indivíduos selecionados, intensidade de seleção, ganho genético (g), ganho genético (%) e tamanho efetivo (N_e) com a seleção em progênies de pupunheira para peso de palmito. Macapá - AP. 2000

Estratégia de seleção	Nº de famílias selecionadas	Nº de indivíduos selecionados	Intensidade de seleção	Ganho genético	Ganho genético	Média	N_e
				predito (g)	predito (%)	Predita (g)	
Curto prazo	09	31	7,92	19,06	8,40	246	19,00
Longo prazo	15	53	13,55	16,30	7,18	243	30,38
Média da população	227g						

TABELA 3 - Estimativas de componentes de variância genética aditiva (σ^2_a), residual (σ^2_e), dentro (σ^2_d) e herdabilidade (h^2) para sete caracteres em uma população de *Bactris gasipaes* por meio dos procedimentos da análise de variância (ANOVA) e REML. Macapá - AP. 2000

Caracteres ¹	σ^2_a		σ^2_e		σ^2_d		h^2	
	ANOVA	REML	ANOVA	REML	ANOVA	REML	ANOVA	REML
AP	0,0065	0,0189	0,00921	0,00321	0,084	0,095	6,85	18,36
DPC	-0,1276	0,0517	0,07333	0,00164	1,3078	1,6233	-9,45	3,16
PRB	32,6446	28,8425	12,8401	27,3811	435,1319	434,1202	7,16	6,15
PRA	-85,3843	0,0000	45,2432	42,6374	1193,537	1105,060	-7,01	0,00
DP	0,01671	0,0178	0,0091	0,00273	0,1536	0,1617	10,00	10,54
TP	22,2041	22,8642	3,5307	2,15009	44,7915	45,9681	41,21	42,47
PP	1017,11	1015,008	72,4507	69,4897	17800,222	17478,413	5,61	5,70

1 AP: altura da planta; DPC: diâmetro da planta à altura do colo; PRB: peso do resíduo basal; PRA: peso do resíduo apical; DP: diâmetro do palmito; TP: tamanho do palmito; PP: peso do palmito.

variação genética na população, recomenda-se apenas a utilização da estratégia de curto prazo e posterior cruzamento dos indivíduos selecionados com indivíduos selecionados em outras populações, visando a gerar uma nova população com maior variabilidade genética aditiva.

As estimativas de herdabilidade no sentido restrito (Tabela 3) são consideradas baixas, revelando baixa variação genética aditiva na população e, conseqüentemente, a estimação de ganhos genéticos de pequena magnitude. Uma possível causa para as baixas estimativas de herdabilidade pode ser a estreita base genética da população, haja vista que as progênies são provenientes de número reduzido de plantas.

As estimativas dos componentes de variância e herdabilidade obtidas pelos métodos da análise de variância (ANOVA) e REML encontram-se na Tabela 3. As estimativas

foram divergentes para os dois procedimentos. Observam-se estimativas negativas de componentes de variância quando se utiliza a ANOVA, resultado normalmente observado quando o valor paramétrico do componente de variância é próximo de zero. Neste caso, o procedimento REML, que apresenta restrição de não-negatividade, conduz a melhores resultados. Exceto para os caracteres DP, TP e PP, os dois procedimentos conduziram a herdabilidades muito diferentes, revelando que o procedimento REML deve ser utilizado em casos de dados desbalanceados.

CONCLUSÕES

1 - A população de pupunheira estudada apresenta baixa variabilidade genética aditiva, o que a torna inadequada para

futuros trabalhos de melhoramento a longo prazo.

2 - Os ganhos genéticos esperados para produção de palmito foram de 8,40 e 7,18% para as condições de curto e longo prazos.

3 - Os procedimentos REML e ANOVA conduziram a resultados divergentes para a maioria dos caracteres.

4 - O procedimento REML/BLUP deve ser o preferido para a estimação/predição no melhoramento de plantas perenes, cujos experimentos, em geral, geram dados desbalanceados.

REFERÊNCIAS BIBLIOGRÁFICAS

BOVI, M.L.A. **Palmito de pupunha: informações básicas para o cultivo**. Campinas. Instituto Agrônomo, 1998. 50p. (Boletim Informativo, 173).

CLEMENT, C.R. **Growth and genetic analysis of pejibaye palm (*Bactris gasipaes*) in Hawaii**. 1995. 221f. PhD. (Thesis) – University of Hawaii, Honolulu, 1995.

CLEMENT, C.R.; BOVI, M.L.A. Melhoramento genético de pupunheira: conhecimentos atuais e necessidades. In: SEMINÁRIO DO AGRONEGÓCIO PALMITO DE PUPUNHA NA AMAZÔNIA, 1., 1999, Porto Velho, **Anais...** p.57-70.

FARIAS NETO, J.T. Estimativas de parâmetros genéticos em progênies de meios-irmãos de pupunheira. **Boletim de Pesquisa Florestal**, Colombo, v.39, p.109-117, 1999.

MEYER, K. DFREML – Version 3.0b User Notes. Armidale: Institute of Animal Genetics of Edinburgh-Scotland/Animal Genetics and Breeding Unit of the University of New England, 1998. 31p.

RESENDE, M.D.; OLIVEIRA, E.B. de. Sistema Selegen - Seleção genética computadorizada para o melhoramento de espécies perenes. **Pesquisa Agropecuária Brasileira**, Brasília, v. 32, n.9, p.931-939, 1997.

RESENDE, M.D.V. de; FERNANDES, J.S.C. Procedimento BLUP

individual para delineamentos experimentais aplicados ao melhoramento florestal. **Revista de Matemática e Estatística**, São Paulo, v.17, p.89-107, 1999.

RESENDE, M.D.V. de; HIGA, A.R. Maximização da eficiência da seleção em testes de progênies de *Eucalyptus* através da utilização de todos os efeitos do modelo matemático. **Boletim de Pesquisa Florestal**, Colombo, v.28/29, p.37-55, 1994.

RESENDE, M.D.V. de; HIGA, A.R.; LAVORANTI, O.J. Predição de valores genéticos no melhoramento de *Eucalyptus* – Melhor Predição Linear. In: CONGRESSO FLORESTAL BRASILEIRO, 7., 1993, Curitiba. **Anais...** p.144-147.

RESENDE, M.D.V. de; PRATES, D.F.; YAMADA, C.K.; JESUS, A. de. Estimação de componentes de variância e predição de valores genéticos pelo método de máxima verossimilhança restrita (REML) e melhor predição linear não viciada (BLUP) em Pinus. **Boletim de Pesquisa Florestal**, Colombo, v. 32/33, p.23-42, 1996.

RESENDE, M.D.V.; BERTOLUCCI, F.L.G. Maximization of genetic gain with restriction on effective population size and inbreeding in *Eucalyptus grandis*. In: IUFRO CONFERENCE ON *Eucalyptus* PLANTATIONS: IMPROVING FIBRE YIELD AND QUALITY, 1995. **Proceedings...** Hobart: IUFRO, 1995.p.167-170.

SEARLE, S.R.; CASELLA, G.; McCULLOCH, C.E. **Variance components**. New York. J. Wiley, 1992. 528p.

SOUZA, A.G.C.; SILVA, S.E.L. Melhoramento de fruteiras da Amazônia. In: SIMPÓSIO BRASILEIRO DE MELHORAMENTO DE FRUTEIRAS, 2., 2000, Viçosa: **Anais...** p.49-55.

TONET, R.M.; FERREIRA, L.G.S.; OTOBONI, J.L.M.A. **A cultura da pupunheira**. Campinas: Coordenadoria de Assistência Técnica Integral - CATI, 1999. 44p. (Boletim Técnico, 237).

VENCOVSKY, R.; BARRIGA, P. **Genética Biométrica no Fitomelhoramento**. Ribeirão Preto:SBG, 1992. 496p.