

Nota Científica

Avaliação genética e análise de *deviance* em um teste desbalanceado de procedência e progênie de *Ilex paraguariensis*

José Alfredo Sturion¹, Marcos Deon Vilela de Resende¹

¹Embrapa Florestas, Estrada da Ribeira, Km 111, CP 319, CEP 80411-000, Colombo, PR, sturion@cnpf.embrapa.br; deon@cnpf.embrapa.br

Resumo - Em um teste combinado de procedência e progênie de erva-mate (*Ilex paraguariensis* A. St. Hil.), os efeitos aleatórios do modelo foram testados pelo teste da razão de verossimilhança (LTR). Os dados referem-se ao peso de biomassa foliar obtido aos seis anos de idade, por ocasião da terceira poda, em um teste com sete procedências e 140 progênies de meios-irmãos instalado no delineamento de blocos ao acaso, com dez repetições de parcelas lineares com seis plantas no espaçamento de 3 m x 2 m. A análise de deviance (Anadev) foi realizada e as seguintes conclusões foram obtidas: os componentes de variância associados a todos os efeitos aleatórios do modelo e respectivos coeficientes de determinação (herdabilidades dos efeitos genéticos aditivos, dos efeitos de parcela e dos efeitos de procedência) foram significativamente diferentes de zero, indicando que existe variabilidade genética significativa a ser explorada pelo melhoramento, tanto entre como dentro de populações naturais.

Termos para indexação: Anadev, erva-mate, teste de progênie.

Genetic evaluation and deviance analysis in an unbalanced provenance and progeny test of *Ilex paraguariensis*

Abstract - Likelihood ratio tests (LRT) were applied in an unbalanced provenance/progeny test of mate (*Ilex paraguariensis* A. St. Hil.). The trial was established in a randomized complete block design with seven provenances, 140 half sib families, ten blocks and six plants per linear plot. The spacing was 3 m x 2 m. The following conclusions were obtained from the deviance analysis (Anadev): all random effects and associated variance components and coefficients of determination (heritability and coefficients of determination of plots and provenances) were significant; there is enough genetic variability in both within and between provenance levels which can be adequately explored for breeding purposes.

Index terms: Anadev, maté, progeny test.

Com o propósito de suprir a demanda de sementes de erva-mate (*Ilex paraguariensis* A. St. Hil.), já com determinado grau de melhoramento, a Embrapa Florestas instalou uma rede de testes combinados de procedência e progênie no delineamento de blocos ao acaso, parcelas lineares com seis plantas nos espaçamentos de 3 m x 2 m e 3 m x 1,5 m, variando o número de blocos de cinco a dez, dependendo da disponibilidade de área. Nestes testes, as plantas são avaliadas individualmente com base em seus valores genéticos.

A avaliação genotípica compreende a estimação de componentes de variância (parâmetros genéticos) e a predição de valores genotípicos. As estimativas dos parâmetros genéticos como herdabilidade e correlações genéticas são fundamentais para o

delineamento de estratégias de melhoramento eficientes. A experimentação de campo, via de regra, está associada ao desbalanceamento de dados devido a vários motivos como perdas de plantas e parcelas, quantidades desiguais de sementes e mudas disponíveis por tratamento, rede experimental com diferentes delineamentos experimentais, não avaliação de todas as combinações genótipo-ambiente, dentre outros. Em função disso, o procedimento ótimo de estimação de avaliação genotípica refere-se ao Reml/Blup (máxima verossimilhança residual ou restrita/melhor predição linear não viciada), também denominado genericamente de metodologia de modelos mistos. Estes procedimentos lidam naturalmente com o desbalanceamento, conduzindo a estimações

e predições mais precisas de parâmetros genéticos e valores genéticos, respectivamente (Resende, 2007b).

O procedimento ótimo de seleção é o Blup para efeitos genéticos aditivos (*a*), de dominância (*d*) e genotípicos (*g*), dependendo da situação. O Blup é o procedimento que maximiza a acurácia seletiva e, portanto, é superior a qualquer outro índice de seleção combinada, exceto aquele que usa todos os efeitos aleatórios do modelo estatístico (índice multiefeitos, conforme Resende & Higa, 1994), o qual é o próprio Blup para o caso de dados balanceados (Resende & Fernandes, 1999). O Blup permite também o uso simultâneo de várias fontes de informações como as advindas de vários experimentos instalados em um ou vários locais e avaliados em uma ou várias colheitas. O Blup individual utiliza todos os efeitos do modelo estatístico, contempla o desbalanceamento, utiliza o parentesco genético entre indivíduos em avaliação e considera a coincidência entre unidade de seleção e unidade de recombinação. A predição usando Blup assume que os componentes de variância são conhecidos (Resende, 2002).

A análise de variância (Anova) e análise de regressão foram, durante muito tempo, o principal esteio da análise e modelagem estatística. Entretanto, estas técnicas têm limitação para lidar com dados desbalanceados e com parentesco entre tratamentos. O método Reml permite lidar com dados desbalanceados e com parentesco entre tratamentos. Tal procedimento constitui-se no procedimento padrão para análise estatística em uma grande gama de aplicações. Em experimentos agrônomicos e florestais, o Reml tem substituído com vantagens o método Anova. Na verdade, o Reml é uma generalização do Anova para situações mais complexas. Para situações simples, os dois procedimentos são equivalentes, mas para as situações mais complexas encontradas na prática, o Anova é um procedimento apenas aproximado. O Reml é um método eficiente no estudo das várias fontes de variação associadas à avaliação de experimentos de campo, permitindo desdobrar a variação fenotípica em seus vários componentes genéticos, ambientais e de interação genótipo x ambiente (Resende, 2007b).

Como principais vantagens práticas, o Reml/Blup permite: comparar indivíduos ou variedades através do tempo (gerações, anos) e espaço (locais, blocos); simultânea correção para os efeitos ambientais, estimação de componentes de variância e predição de valores genéticos; lidar com estruturas complexas

de dados (medidas repetidas, diferentes anos, locais e delineamentos); aplicar a dados desbalanceados e a delineamentos não ortogonais; utilizar simultaneamente um grande número de informações, provenientes de diferentes gerações, locais e idades, gerando estimativas e predições mais precisas; permite o ajuste de vários modelos alternativos, podendo-se escolher o que se ajusta melhor aos dados e, ao mesmo tempo, é parcimonioso, isto é, apresenta menor número de parâmetros (Resende, 2002).

No caso de dados desbalanceados, o Anova conduz a estimativas imprecisas de componentes de variância e, conseqüentemente, a predições inacuradas de valores genéticos.

Na análise de modelos mistos com dados desbalanceados, os efeitos do modelo não são testados via testes F, tal como se faz no método de análise de variância. Nesse caso, para os efeitos aleatórios, o teste cientificamente recomendado é o teste da razão de verossimilhança (LTR). Para os efeitos fixos, um teste F aproximado pode ser usado. Um quadro similar ao quadro de análise de variância pode ser elaborado. Tal quadro pode ser denominado análise de deviance (Anadev) e é estabelecido segundo os seguintes passos, conforme Resende (2007a):

Obtenção do logaritmo do ponto de máximo da função de verossimilhança residual (*L*) para modelos com e sem o efeito a ser testado;

Obtenção da deviance $D = -2\log L$ para modelos com e sem o efeito a ser testado;

Fazer a diferença entre as deviances para os modelos sem e com o efeito a ser testado, obtendo a razão de verossimilhança (*LR*);

Testar, via LTR, a significância dessa diferença usando o teste qui-quadrado com 1 grau de liberdade.

Essa análise foi aplicada a um teste combinado de procedência e progênie integrante do programa de melhoramento genético da Embrapa Florestas, com o objetivo de verificar a significância dos efeitos do modelo matemático e também divulgar a metodologia para aplicações práticas rotineiras em programas de melhoramento genético. A presente aplicação refere-se a um modelo diferente daquele apresentado por Resende (2007b), o qual é dado por:

$$y = u + b + a + parc + proc + e$$

em que *b* refere-se ao efeito fixo de blocos, *a* ao efeito aleatório genético aditivo dentro de procedências, *parc* é ao efeito aleatório de parcela, *proc*, efeito aleatório de

procedências e *e* refere-se ao resíduo dentro da parcela, respectivamente.

O material genético consiste de sete procedências de erva-mate plantadas, respectivamente, em março e julho de 1997, em dois testes combinados de procedência e progênie em Ivaí, PR, e Chapecó, SC (Tabela 1).

Ivaí está localizada na latitude de 25° 01' Sul, longitude de 50° 48' oeste, altitude de 600 m e está sob a influência do tipo climático Cfb – clima subtropical úmido mesotérmico, com verões quentes e geadas pouco frequentes, com tendência de concentração das chuvas nos meses de verão, sem estação seca definida. A média das temperaturas dos meses mais quentes é superior a 22 °C e a dos meses mais frios é inferior a 18 °C, temperatura média anual entre 17 °C e 18 °C, precipitação média anual em torno dos 1.500 mm e excedente hídrico variando de 500 mm a 800 mm. O experimento foi instalado em solo da classe Latossolo Bruno distrófico típico a moderado, textura muito argilosa, fase relevo suave ondulado segundo Sistema... (1999). Estes solos caracterizam-se por serem profundos acentuadamente drenados, porosos, muito argilosos (72% de argila) e de coloração bruno avermelhada-

escura. Quimicamente, são ácidos com saturação de bases baixa e saturação com alumínio elevada. Ocorrem em relevo suave ondulado com declives em torno de 4% e em altitudes variando de 650 m a 750 m e originados de rochas sedimentares (argilito).

O delineamento experimental empregado foi o de blocos ao acaso, com dez repetições de seis plantas por parcela, no espaçamento de 3 m x 2 m. As progênies foram aleatorizadas independentemente das procedências, ou seja não se adotou o arranjo hierárquico de progênies dentro de procedência. A poda de formação foi efetuada em 1999, aos dois anos de idade. As podas de produção foram efetuadas, respectivamente, em julho de 2001 e agosto de 2003. Os resultados apresentados nesse trabalho referem-se ao peso de massa foliar obtido por ocasião da segunda poda de produção, aos seis anos de idade.

As estimativas dos parâmetros genéticos e fenotípicos foram efetuadas através do programa genético-estatístico Selegen-Reml/Blup, desenvolvido por Resende (2007b) e a análise de deviance (Anadev) realizada (Tabela 2).

Tabela 1. Procedências e progênies de erva-mate coletadas em sete regiões do sul do Brasil e plantadas em Ivaí, PR.

Procedências	Lat. (Sul)	Long. (Oeste)	Alt. (m)	Progênies
Ivaí, PR (IV)	25° 01'	50° 48'	600	25
Colombo, PR (CO)	25° 20'	49° 14'	920	18
Barão de Cotegipe, RS (BR)	27° 38'	52° 23'	530	21
Quedas do Iguaçu, PR (QUE)	25° 25'	52° 55'	590	24
Pinhão, PR (PI)	25° 41'	51° 40'	1.050	22
Cascavel, PR (CA)	24° 57'	53° 27'	750	25
Antonio Olinto, PR (AO)	25° 59'	50° 12'	790	5
Total				140

Tabela 2. Análise de deviance (Anadev) a partir do programa Selegen-Reml/Blup.

Efeito	Deviance	LTR (Qui-quadrado)	Comp. Variância	Coef. Determinação
Genético dentro de Procedências	21.857,10 ⁺	133,80**	2,682396	h ² a = 0,2219
Parcela	21.875,90 ⁺	152,60**	1,021126	c ² parc = 0,0845
Procedência	21.855,91 ⁺	132,61**	2,395794	c ² proc = 0,1982
Resíduo			8,000118	c ² residuo = 0,4954
Modelo completo	21.723,30			c ² total = 1,0000
Bloco		F = 2,50**		

Qui-quadrado tabelado: 3,84 e 6,63 para os níveis de significância de 5% e 1%, respectivamente, + deviance do modelo ajustado sem os referidos efeitos.

O programa Selegen-Reml/Blup fornece as deviances quando se rodam os modelos com ou sem os efeitos a serem testados (basta zerar os coeficientes de determinação de c^2 correspondentes, na tela do Selegen). De posse dessas deviances, torna-se fácil construir a tabela de análise de deviance. No presente teste, verifica-se que os efeitos genéticos aditivos de indivíduos, parcelas e procedências são significativos. Consequentemente, os respectivos componentes de variância são significativamente diferentes de zero, assim como os respectivos coeficientes de determinação (herdabilidade – h^2a , o coeficiente de determinação dos efeitos de parcela – c^2_{parc} e o coeficiente de determinação dos efeitos de procedência c^2_{proc}), conforme obtidos pelo modelo 5 do Selegen.

Esses resultados indicam que existe variabilidade genética significativa a ser explorada pelo melhoramento, tanto entre como dentro de populações naturais. A seleção pode então ser realizada em vários níveis: entre populações, dentro das melhores populações e simultaneamente entre e dentro de populações, combinando os valores genéticos de populações e de indivíduos dentro de populações. A melhor estratégia depende do valor de cada população para os caracteres

de qualidade dos produtos da erva-mate, por exemplo, o teor de cafeína.

Referências

- SISTEMA Brasileiro de Classificação de Solos. Brasília, DF: Embrapa Produção de Informação; Rio de Janeiro: Embrapa Solos, 1999. 412 p.
- RESENDE, M. D. V. de. **Genética biométrica e estatística no melhoramento de plantas perenes**. Brasília: Embrapa Informação Tecnológica, 2002. 975 p.
- RESENDE, M. D. V. de. **SELEGEN-REML/BLUP: sistema estatístico e seleção genética computadorizada via modelos lineares mistos**. Colombo: Embrapa Florestas, 2007a. 359 p.
- RESENDE, M. D. V. de. **Matemática e estatística na análise de experimentos e no melhoramento genético**. Colombo: Embrapa Florestas, 2007b. 362 p.
- 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, n. 28/29, p. 37-55, jan./dez. 1994

Recebido em 09 de abril de 2010 e aprovado em 27 de outubro de 2010