ARTICLE

Estimation of genetic parameters and verification of early selection efficiency in baru (*Dipteryx alata*)

Alisson Moura Santos[1*], Sebastião Carlos da Silva Rosado[2] and Adelson Nascimento Oliveira[2]

Abstract – *This study aimed at verifying the efficiency of early selection, and estimating the genetic parameters in* Dipteryx alata. *The 66 half-sib families were obtained from three seed provenances at the Brazilian Savannah of the state of Minas Gerais. The experiment was carried out in a randomized block design with 3 replications and 5 plants per plot. Data diameters at breast height (DBH) and total height (TH) were taken at the ages of 64, 125 and 138 months. REML/BLUP methodology was applied in the* D. alata *different provenances and progenies. Analysis indicated significant genetic variability (P < 0.01) between and within provenances. Both traits, DBH and TH, presented high narrow sense heritability for the ages analyzed. High genetic correlation occurred between DBH and TH traits, and between ages (age-age); thus, it allowed indirect selection, as well as early selection with high genetics gains.*

Key words: *Breeding forest, age-age correlation, mixed models, genetic gains.*

INTRODUCTION

Baru (*Dipteryx alata* Vog.) is a species native to Brazil, which grows throughout the cerrado biome (Ratter et al. 2003). It presents high quality wood, and fruits with high protein content (Fernandes et al. 2010, Cruz et al. 2011). It may present 5-10 m tall and 15-40 cm diameter at breast height (DBH), reaching up to 25 m tall and 70 cm DBH at adulthood (Carvalho 2010), being an important species for use in integrated agricultural systems or afforestation of pasture.

The integrated farming has received great attention in recent years, mainly due to the increasingly required conservation farming systems. It is a versatile operating model, which enables to integrate forestry, agriculture and livestock in the same area with economic and environmental benefits (Hendrickson et al. 2008).

However, there is little information concerning forest species, especially the native ones in integrated cultivation conditions. It is a complex system, when compared to conventional forestry system. It requires a holistic understanding of the system as well as synergy between its components. The technological barrier is obtaining appropriate genotypes for such conditions that ensure long-term economic return.

Breeding programs of native plants aiming at intercropped production in an agroforestry system is an eminent alternative to the current agrarian situation, in addition to providing recovery of the species, conservation, and, especially, commercial cultivation.

For obtaining improved genotypes, it is necessary to identify those that meet the requirements in forest cultivation. Selection from provenance tests combined with progeny tests (Sampaio et al. 2000) is a proven efficiency strategy. These tests, which were carried out in the same area and same time, enable to obtain the estimates of genetic parameters, which allow one to infer the genetic variability within populations, and provide subsidies to predict genetic gain and the potential success in the breeding program (Cruz and Carneiro 2006).

Based on the aforementioned, the aim of this study was to estimate genetic parameters between and within half-sib families of *D. alata*, and verify the possibility of early selection.

MATERIAL AND METHODS

The experiment was carried out at *V & M Florestal* farm, in Brasilândia de Minas/MG, and consisted of three

[1] Embrapa Florestas, Estrada da Ribeira, Km 111, C.P. 319, 83.411-000, Colombo, PR, Brazil. *E-mail: alisson.santos@embrapa.br
[2] Departamento de Ciências Florestais, Universidade Federal de Lavras (UFLA), C.P. 3037, 37.200-000, Lavras, MG, Brazil

provenances from the municipalities of Capinópolis/MG, Brasilândia de Minas/MG and Jequitaí/MG. Seeds of 25 plus tree were collected in Brasilândia de Minas, 25 in Capinópolis, and 16 in Jequitaí, totaling 66 open-pollinated families, sampled randomly, spaced at minimum 100 m apart. The experimental design was randomized blocks, with three replications, and plots consisted of five plants, spaced 3 m x 4 m apart, with two border rows. The locations where seeds were collected is characterized by cultivations of *Eucalyptus sp.* (Brasilândia de Minas), soybean and pasture (Capinópolis), and cerrado fragments (Jequitaí).

All experimental plants underwent pruning at 12, 36 and 64 months of age. Diameter at breast height (DBH) and total height (TH) were evaluated in each plant, at the ages of 64, 125 and 138 months.

Statistical analysis was carried out based on REML/BLUP mixed models, according to Resende (2002, 2007a), by the Selegen-REML/BLUP software (Resende 2007b). The linear mixed model used for genetic evaluation at the individual level is given by: $y = Xb + Za + Wc_1 + Qr + e$, where y refers to the data vectors; b, to the fixed block effects; a, to the additive genetic random effects; c_1 to plots random effects; r, to provenances random effects; e, to the residue within the plot X, Z, W and Q were the incidence matrices for b, a, c_1 and r, respectively.

Given that the species is alogamous (Oliveira and Sigrist 2008), it was assumed half-sib relatedness between individuals of each matrix, and thus additive genetic effects were obtained through the use of the additive genetic relatedness matrix, as set by Resende (2007b).

Heritability estimates in level of family means were obtained by the following expression, according to Resende (2002):

$$h_m^2 = \frac{[1 + (n-1)\rho_a]h_a^2}{1 + (n-1)(\rho_a h_a^2)}$$

where ρ_a is the genetic correlation between individuals of the considered type of family ($\rho_a = 1/4$ for half-sibs), h_a^2 is the narrow sense individual heritability, and n is the total number of individuals per family.

Selective accuracies for selecting families were calculated according to Oliveira et al. (2004), and the genetic and phenotypic correlations between the characteristics were obtained using Selegen-REML/BLUP software (Resende 2007b).

The selection of the best families was based on the additive genetic value of progeny, having DBH as a reference due to its greater experimental precision and high correlation

with height. Only those families, which had individuals with positive additive genetic effects were selected. Selection within family was based on individual genotypic value and intensity of 20%, i.e., the individual with greater genotypic value within the family in each block was selected, aiming to avoid high rate of crossings between related individuals. Genotypic values of each selected individual were obtained by summing each genotypic effect to the overall mean. Genetic gain is equal to the mean of the selected predicted genetic values for each trait.

The efficiency of early selection was evaluated using the expression for estimating indirect gain, according to Cotterill and Dean (1990):

$$\Delta A_x = i h_x h_y r_a \sigma_{p(x)}$$

where: ΔA_x: is the expected gain for x characteristic (indirect gain); h_x and h_ys are the narrow sense heritability root of x and y characteristics, respectively; i: is the standardized intensity of selection; r_a: is the additive genetic correlation between both characteristics; $\sigma_{p(x)}$: is the standard deviation of the phenotypic variance of x. The characteristics x and y may be the same at different ages; it may be different characteristics at the same age, or different characteristics at different ages.

RESULTS AND DISCUSSION

Mortality in all experiments was relatively low, being 2.72%, 7.88% and 7.99% for the ages of 64, 125 and 138 months, respectively. There was a higher percentage of death in the seeds of Jequitaí, and this phenomenon was observed from the nursery stage. The deaths may be related to seed quality, considering that they were collected in their natural environment and experienced different degrees of maturation.

Both for DBH and TH, estimates of additive genetic variance between families ($\hat{\sigma}_a^2$) and individual phenotypic variance ($\hat{\sigma}_f^2$) had a clear trend to present higher values in the older ages (Table 1). Regarding the minimum (64 months) and maximum (138 months) ages, a $\hat{\sigma}_a^2$ had increases of 247% and 221% for DBH and TH, respectively. For $\hat{\sigma}_f^2$, these increases were 209% and 153%, respectively. Considering the coefficient of environmental variation (CVe), there was no significant variation in both characteristics; thus, it can be inferred that during this period there was a large release of genetic variability.

Analyzing the coefficients of individual genetic variation (CVgi), the increases between the minimum and maximum ages were 30% and 39%, respectively. Similar increases were also observed for the coefficient of genotypic varia-

Table 1. Means estimates, variance components and individual heritability for diameter at breast height (DBH) and total height (TH) in families of *Dipteryx alata* half-sibs, at 64, 125 and 138 months of age[1]

Parameters	Months					
	DBH (cm)			TH (m)		
	64	125	138	64	125	138
$\hat{\sigma}_a^2$	0.9057	2.8661	3.1495	0.3332	0.7025	1.0723
$\hat{\sigma}_{plot}^2$	0.2122	0.1928	0.1602	0.0774	0.1020	0.1781
$\hat{\sigma}_{proc}^2$	0.1952	0.9976	1.3058	0.0244	0.1907	0.2616
$\hat{\sigma}_e^2$	0.4875	0.7434	0.9510	0.2579	0.4755	0.2431
$\hat{\sigma}_f^2$	1.8007	4.7952	5.5666	0.6930	1.4710	1.7552
$\hat{h}_a^2$	0.5029	0.5967	0.5657	0.4808	0.4776	0.6109
$\hat{s}(\hat{h}_a^2)$	0.1293	0.1447	0.1410	0.1264	0.1295	0.1465
h_m^2	0.8199	0.8694	0.8543	0.8064	0.8044	0.8760
$r_{\hat{a}am}^2$	0.9055	0.9324	0.9242	0.8980	0.8969	0.9359
C_{parc}^2	0.1178	0.0402	0.0287	0.1117	0.0694	0.1014
C_{proc}^2	0.1084	0.2080	0.2345	0.0352	0.1296	0.1490
CV_{gi} (%)	18.8286	24.4570	24.4775	14.8460	19.3774	20.7156
CV_{gp} (%)	9.4143	12.2285	12.2387	7.4230	9.6887	10.3578
CV_e (%)	13.2063	12.6930	12.5115	10.8815	12.7167	12.4550
CV_r	1.42	1.92	1.95	1.36	1.52	1.66
PEV	0.0407	0.0936	0.1148	0.0161	0.0343	0.0332
SEP	0.2019	0.3059	0.3388	0.1269	0.1853	0.1823
General mean	5.054	6.9165	7.2503	3.8883	4.3257	4.9987

[1] $\hat{\sigma}_a^2$ = additive genetic variance; $\hat{\sigma}_{parc}^2$ = environmental variance between families; $\hat{\sigma}_{plot}^2$ = genetic variance between provenances; $\hat{\sigma}_e^2$ = residual variance within plots; $\hat{\sigma}_f^2$ = individual phenotypic variance; $\hat{h}_a^2$ = coefficient of narrow sense individual heritability in the block; $\hat{s}(\hat{h}_a^2)$ = standard deviation of heritability estimates; h_m^2 = heritability in level of family means; $r_{\hat{a}am}^2$ = selective accuracy for family selection; C_{parc}^2 = coefficient of determination of family effects; C_{proc}^2 = coefficient of determination of provenance effects; CV_{gi} = coefficient of individual additive genetic variation; CV_{gp} = coefficient of genotypic variation between families; CV_e = coefficient of residual variation; $CV_r = CV_{gi}/CV_e$; PEV = prediction error variance of progeny genotypic values, assuming complete survival; SEP = standard deviation of the progeny predicted genotypic value, assuming complete survival.

tion between families (CVgp). For this reason, estimates of narrow sense individual heritability, obtained at the ages of 64, 125 and 138 months showed significant values ranging from 0.502 to 0.596 for DBH, and 0.480 to 0.610 for TH (Table 1). It can be considered that the increase in heritability as the plant aged was statistically significant only for TH.

Low values (0.04 to 0.11 for DBH, and 0.01 to 0.03 for height) for the prediction error variance (PEV) and its standard deviation (SEP) (Table 1) were obtained. According to Resende and Duarte (2007), PEV is directly related to the experimental precision and accuracy. The lower the PEV value, the greater the accuracy and precision in the inference of predicted genetic values and in the ranking of individuals in evaluation. The low PEV values found indicate that the residual variation of the experiment was of low magnitude. This result shows a favorable situation for effective selection of superior individuals.

The CVgi/CVe ratio, also known as the coefficient of

relative variation (CVr) is another indicator of success in selection. According to Vencovsky (1987), in experimentation with different half-sibs, this ratio indicates positive selection when the situation reaches values greater than or equal to 1.0, which was observed in all ages tested. According to Resende and Duarte (2007), the magnitude of the CVr can also be used to infer the accuracy and precision in genotypic evaluation when associated with the number of repetitions. CVr increased from 1.42 at 64 months to 1.95 at 138 months of age for DBH. Similar behavior was observed for the height; thus, it can be inferred that the selection at 138 months was favorable for obtaining gains.

Estimates of narrow sense individual heritability ($\hat{h}_a^2$) were similar in the different ages tested for both characteristics. The $\hat{h}_a^2$ values suggest the possibility of selection with significant gains for the evaluated characteristics, since, according to Bernardo (2010), the expected gain with selection depends directly on the heritability and intensity of selection. In addition, obtaining high heritability values

is essential to increase the efficiency of early selection (Cotterill and Dean 1990).

Heritability estimates based on family means (h_m^2) were high, being generally higher than estimates at individual level. Estimates obtained using the means have higher accuracy, this is for decreasing the influence of experimental errors when means are used in the selection criteria (Vencovsky and Barriga 1992). This result indicates a greater chance of selecting genotypically superior progenies for the characteristics evaluated. Heritability estimates of high magnitude for *D. alata* were obtained by Siqueira et al. (1993) and Rocha et al. (2009), also indicating a situation favorable for selection.

The coefficients of determination of plot effects (C_{plot}^2) for DBH and height were low, indicating the low environmental variation between plots within the block, and the efficiency of the experimental design. The coefficients of determination of provenance effects (C_{proc}^2) were higher at 125 and 138 months of age. This indicates the influence of provenance effect on total variation at older ages.

The values found for genetic variation between provenances ($\hat{\sigma}_{proc}^2$) showed great variation and trend of higher values at 125 and 138 months of age, as shown in Table 1. These results demonstrate the variability between provenances, which enables selection between provenances for the characteristics evaluated.

Table 2 shows the significance of the additive and provenance genetic effects, at different ages evaluated, by deviance analysis. It is noticed that genetic and provenance effects were highly significant for different ages, except for the plot effects after 138 months of age.

It is confirmed that there was trend of genotypic correlation coefficients, age-age and between the characteristics, to be higher than the phenotypic correlation (Table 3). This behavior is probably related to the magnitude of environmental variances and covariances which acted to reduce the phenotypic correlation.

Estimates of genetic correlations between characteristics at different ages were high. The genetic correlation for diameter at different ages exceeded 0.79. Estimates of high genetic correlation between ages are reported in studies, especially for Eucalyptus (Marques Júnior et al. 1996, Wei and Borralho 1998, Osorio et al. 2003, Lima et al. 2011).

The genetic correlation between the characteristics indicates the direction and magnitude of the correlated response (Falconer and Mackay 1996). Knowing the magnitude of this correlation, which can be caused by pleiotropy and/or genetic linkage (Falconer and Mackay 1996), is critical for the prediction of correlated gain between the characteristics (Cotterill and Dean 1990). Thus, there are favorable prospects for obtaining indirect

Table 2. Deviance analysis for diameter at breast height (DBH) in *Dipteryx alata* evaluated at 64, 125 and 138 months of age, in Brasilândia de Minas

Effect	64 months		125 months		138 months	
	Deviance	LRT Qui-square	Deviance	LRT Qui-square	Deviance	LRT Qui-square
Genotypes	1333.2+	19.7**	2070+	37.8**	2181.82+	38.8**
Provenance	1327.9+	14.3**	2060.2+	27.9**	2175.5+	32.6**
Complete Model	1313.5	-	2032.2	-	2142.9	-

** - *Qui-square significant at 1%. + - Deviance of the model adjusted with none of the effects.*

Table 3. Genotypic (above diagonal) and phenotypic (below the diagonal) correlations between age-age and diameter at breast height (DBH) and total height (TH)

Variables	64 months		125 months		138 months	
	DBH	TH	DBH	TH	DBH	TH
DBH 64	-	0.801	0.804	0.631	0.790	0.580
TH 64	0.758	-	0.694	0.685	0.681	0.632
DBH 125	0.389	0.338	-	0.872	0.987	0.835
TH 125	0.329	0.348	0.827	-	0.881	0.903
DBH 138	0.394	0.339	0.947	0.819	-	0.838
TH 138	0.278	0.283	0.783	0.813	0.804	-

Table 4. Direct and correlated gains for diameter at breast height (DBH) and total height (TH), based on selection for DBH at the ages of 64, 125 and 138 months

Selection	DBH Gain (%)			TH Gain (%)		
	64	125	138	64	125	138
DBH 64	21.8	29.7	23.5	11.3	12.9	14.6
DBH 125		33.5	25.1		13.0	14.3
DBH 138			32.8			39.9

gains in height via DBH direct selection, regardless of the age studied.

Since the selection was based on families with higher additive genetic value, the number of selected individuals ranged at different ages. 50% of the families of the experiment were selected at 138 months of age, based on DBH. It was observed that the indirect gain in height by DBH selection at 64 months of age was approximately 11.3% (Table 4), with selection intensity of 9.8%, following the selection criteria based on the additive genetic effect. At the ages of 125 and 138 months, gains were 12.9% and 14.6%, respectively. This evolution of indirect gain in height is important with the advance of the breeding program, given that the heights become more difficult to be measured and the errors increase significantly.

The efficiency of a forest breeding program is directly related to the time required to obtain gains. Thus, the coincidence of the performance of the characteristics of interest in early and adult ages is essential for greater program effectiveness.

Confronting individuals selected for DBH at 64 and 138 months of age, it is verified that the coincidence is high, with correlated gain of 23.5% (Table 4). Taking into account that the gain for DBH at the age of 138 months was approximately 32.8%, it verified that the early selection at 64 months had an efficiency of 71.7%, evidencing the possibility of using the early selection.

Early selection is a particular case of indirect selection of fundamental importance for the species since it favors the selection of individuals at early age. Based on these results, the feasibility of selecting the best individuals at an early age envisions the possibility of the use of intra-population recurrent selection more effectively. Moreover, the efficiency of early selection is also evidenced by the high values of narrow sense heritability. This enables greater dynamism in the breeding program, allowing recombination of the best individuals in a shorter period of time. Taking into account that mostly native plants have moderate growth and long interval between generations, the use of this strategy is key to obtaining gains in a shorter period of time.

The selection for vegetative and clonal propagation is a strategy that leads to faster gains than those based on sexual propagation. However, for this species, there are few studies that demonstrate the viability of asexual propagation.

According to the results, 38, 29 and 47 families from Capinópolis showed the highest additive genetic values. Most of the selected families come from this provenance. This may be associated to the high means found in this population, and to the fact that most individuals in Capinópolis are found in pasture and soybean plantations, probably having been pre-selected at the time these cultures were deployed.

ACKNOWLEDGEMENTS

Estimação de parâmetros genéticos e verificação da eficiência de seleção precoce em Baru (*Dipteryx alata*)

Resumo – *Este trabalho objetivou avaliar a eficiência de seleção precoce e estimar os parâmetros genéticos em* Dipteryx alata. *Foram avaliadas 66 famílias de meio-irmãos, de três procedências no delineamento de blocos ao acaso com três repetições e cinco plantas por parcela. Foram obtidos os dados de diâmetro à altura do peito (DAP) e altura total aos 64, 125 e 138 meses de idade. Obtiveram-se, para ambas as características, as estimativas dos parâmetros genéticos via metodologia de REML/BLUP. Verificou-se a existência de variabilidade genética significativa entre e dentro das procedências. Observou-se alta correlação genética entre DAP e altura total e entre as idades (idade-idade), possibilitando assim a seleção indireta e seleção precoce com significativos ganhos.*

Palavras-chave: *Melhoramento florestal, correlação idade-idade, modelos mistos, ganhos de seleção.*

REFERENCES

Bernardo R (2010) **Breeding for quantitative traits in plants**. Stemma, Woodbury, 400p.

Carvalho PER (2010) **Espécies florestais brasileiras: recomendações silviculturais, potencialidades e uso da madeira**. Embrapa Informação Tecnológica, Brasília, 644p.

Cotterill PP and Dean A (1990) **Successful tree breeding with index selection**. CSIRO Publications, Melbourne, 80p.

Cruz CD and Carneiro PSC (2006) **Modelos biométricos aplicados ao Melhoramento genético**. Editora UFV, Viçosa, 585p.

Cruz KS, Silva MA, Freitas O and Neves VA (2011) Partial characterization of proteins from baru (*Dipteryx alata* Vog) seeds. **Journal of the Science of Food and Agriculture 91**: 2006-2012.

Falconer DS and Mackay TFC (1996) **Introduction to quantitative genetics**. Longman, London, 463p.

Fernandes DC, Freitas JB, Czeder LP and Naves MMV (2010) Nutritional composition and protein value of the baru (*Dipteryx alata* Vog.) almond from the Brazilian Savanna. **Journal of the Science of Food and Agriculture 90**: 1650-1655.

Hendrickson JR, Hanson JD, Tanaka DL and Sassenrath G (2008) Principles of integrated agricultural systems: Introduction to processes and definition. **Renewable Agriculture and Food System 23**: 265-271.

Lima JL, Souza JC, Ramalho MAP, Andrade HB and Sousa LC (2011) Early selection of parents and trees in *Eucalyptus* full-sib progeny tests. **Crop Breeding and Applied Biotechnology 11**: 10-16.

Marques Júnior OG, Andrade HB and Ramalho MAP (1996) Assessment of the early selection efficiency in *Eucalyptus cloeziana* F. Muell. In North West of Minas Gerais State. **Silvae Genetica 45**: 359-361.

Oliveira VR, Resende MDV, Nascimento CES, Drumond MA and Santos CAF (2004) Variabilidade genética de procedências e progênies de umbuzeiro via metodologia de modelos lineares mistos (REML/BLUP). **Revista Brasileira de Fruticultura 26**: 53-56.

Oliveira MIB and Sigrist MR (2008) Fenologia reprodutiva, polinização e reprodução de *Dipteryx alata* Vogel. (Leguminosae-Papilionoideae) em Mato Grosso do Sul, Brasil. **Revista Brasileira de Botânica 31**: 195-207.

Osorio LF, White TL and Huber DA (2003) Age-age and trait-trait correlation for *Eucalyptus grandis* Hill ex Maiden and their implications for optimal selection age and design of clonal trials. **Theoretical and Applied Genetics 106**: 735-743.

Ratter JA, Bridgewater S and Ribeiro JF (2003) Analysis of the floristic composition of the Brazilian cerrado vegetation III: Comparison of the woody vegetation of 376 areas. **Edinburgh Journal of Botany 60**: 57-109.

Resende MDV (2002) **Genética biométrica e estatística no melhoramento de plantas perenes**. Embrapa Informação Tecnológica, Brasília, 975p.

Resende MDV (2007a) **Matemática e estatística na análise de experimentos e no melhoramento genético.** Embrapa Florestas, Colombo, 561p.

Resende MDV (2007b) **Software SELEGEN – REML/BLUP: Sistema estatístico e seleção computadorizada via modelos lineares mistos**. Embrapa Florestas, Colombo, 359p.

Resende MDV and Duarte JB (2007) Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical 37**: 182-194.

Rocha RB, Rocha MGB, Santana RC and Vieira AH (2009) Estimação de parâmetros genéticos e seleção de procedências e famílias de *Dipteryx alata* Vogel (Baru) utilizando metodologia de Reml/Blup e E(QM). **Cerne 3**: 331-338.

Sampaio PTB, Resende MDV and Araújo AJ (2000) Estimativas de parâmetros genéticos e métodos de seleção para o melhoramento genético de *Pinus caribaea* var. *Hondurensis*. **Pesquisa Agropecuária Brasileira 11**: 2243-2253.

Siqueira ACMF, Nogueira JCB and Kageyama PY (1993) Conservação dos recursos genéticos ex situ do cumbaru (*Dipteryx alata* Vog.) – Leguminosae. **Revista Instituto Florestal de São Paulo 2**: 231-243.

Vencovsky R (1987) Herança quantitativa. In Paterniani E and Viegas GP (eds.) **Melhoramento e produção do milho**. Fundação Cargill, Campinas, p.137-214.

Vencovsky R and Barriga P (1992) **Genética biométrica no fitomelhoramento**. Revista Brasileira de Genética, Ribeirão Preto, 486p.

Wei X and Borralho NMG (1998) Genetic control of growth traits of *Eucalyptus urophylla* S T Blake in South East China. **Silvae Genetica 47**: 158-165.