Advances in Forestry Science
Original Article

Repeatability and genetic dissimilarity using biometric traits of black wattle seeds

Paulo César Flôres Junior[1]*   Angela Cristina Ikeda[1]   Guilherme Schnell Schuhli[2]   Luciana Duque Silva[3]   Antonio Rioyei Higa[1]

[1] Universidade Federal do Paraná, Rodovia Régis Bittencourt, 9082-9174 - Jardim Botânico, Curitiba - PR, CEP: 80050-540
[2] EMBRAPA Floresta, Estrada da Ribeira, Km 111 - Parque Monte Castelo, Colombo - PR, 83411-000
[3] ESALQ/USP, Avenida Pádua Dias, 11 - Agronomia, Piracicaba - SP, 13418-900

Abstract
Selecting superior individuals requires a careful genetic evaluation. For this reason, repeated measurements of biometric traits need to be carried out in the same individual. This index is a strategic tool aimed to assist in providing accurate observations and further selection. The objectives of this study were to estimate the repeatability coefficient in many seed traits for black wattle clones; test the best method to estimate the optimal number of measurements required; and to analyze genetic dissimilarity. Seed weight, longitudinal length, transversal length, thickness and hilum size were evaluated in nine clones measuring 50 seeds per clone The following methods were employed in the evaluation: analysis of variance, principal component analysis (PCA) based on the correlation and covariance matrices, and structural analysis. Mahalanobis distance was calculated and a scatter plot was created to evaluate genetic dissimilarity. The results showed that the most effective method to determine the repeatability coefficient was PCA based on the covariance matrix. Twenty measurements are needed in order to achieve a coefficient of determination of 90%. Based on the measures of genetic dissimilarity, it was possible to differentiate clones and to recognize two genetically distinct groups.
Keywords: *Acacia mearnsii*; number of measurements; principal component analysis

Introduction

Repeatability is a measure that expresses the maximum heritability value that can be achieved for a given sample. It is possible to know the component of variance from a random temporary or localized environment solely by estimating the repeatability coefficient of one trait. However, a permanent environment that acts upon a cultivar or progenies shows indistinct genotypic variance (Cruz et al., 2012). Therefore, repeatability, as heritability, is an indispensable tool to guide tree breeding trials (Ferreira et al., 2010; Cruz et al., 2012).

Repeatability coefficient has been successfully applied to assess morphological traits in bacuri (Silva et al., 2009), macaúba (Manfio et al., 2011), elephant grass (Cavalcante et al., 2012), muricizeiro (Lorenço et al, 2013); soybean (Matsuo et al., 2012) and peroba rosa (De Paula et al., 2015). According to Vencovsky (1977), the repeatability coefficient is suitable to studies on perennial plants due to the repetition of certain traits throughout the life of an individual. Therefore, when a particular genotype is considered as being superior, it is expected that this superiority remains constant throughout the whole life of this genotype. Papers addressing the issue of how many plants need to be measured to assure reliable heritability estimates are scarce. Black wattle is no exception and no article was found regarding its repeatability coefficient.

In the south of Brazil, black wattle cultivation has spread throughout the State of Rio Grande do Sul, so that the acacia production chain has become an important economic activity with considerable social and environmental impacts on that region (Grein et al., 2013). This species has been cultivated since 1983 in breeding programs (Mochiutti, 2007). Therefore, black wattle has an important role in the Brazilian forestry sector. Since cultivars are the result of extensive work within breeding programs, it is a priority to legally protect them. In order to accomplish its protection, it is fundamental to distinguish promising genotypes based on variable morphological descriptors from different parts of the plant.

The present study aims to estimate the repeatability coefficient of black wattle seeds as well as to select the best method for its evaluation. It also explores the estimate of the optimal number of measurements and analyzes the genetic dissimilarity between black wattle clones.

Material and methods

Estimate of the repeatability coefficient for the biometric traits of the seeds

The seeds used in this study were collected from a clonal orchard located in Piratini, RS – Brazil. Seeds of two-year-old black wattle trees were collected. The experiment was set up in a randomized block design. Five replicates of nine clones (genotypes) were sampled. Each clone was represented by random samples containing 50 seeds per clone collected in 2014. Five biometric traits were measured for each seed: individual weight (SW); longitudinal length (LL); transversal length (TL); thickness (TH) and hilum size (HS).

In order to estimate for each method the repeatability coefficient for all traits of the seeds statistical analyzes included analysis of variance (ANOVA), principal component analysis (based on the correlation and covariance matrices) and structural analysis based on the correlation matrix. In the analysis of variance, the repeatability coefficient was obtained considering two variation factors. The model used was as follows: $Y_{ij} = \mu + G_i + A_j + E_{ij}$, in which Y_{ij} = observation corresponding to the ith subsample; μ = general average; G_i = random effect of the ith sub-sample under influence of the permanent environment (i = 1, 2 ... p); A_i = fixed effect of the temporary environment on the ith measurement; and E_{ij} = effect of the temporary environment associated to the jth measurement on the ith sub-sample (j = 1,2, ..., n).

Repeatability coefficient is given by: $r = \rho = \sigma^2 g/(\sigma^2 + \sigma^2 g)$ Therefore, $\sigma^2 g$ = estimated variance among genotypes; and σ^2 = estimated variance of the experimental error. Subsequently, the results of the analysis of variance for each variable were obtained and the variance components associated to genetic and residual effects were estimated. Repeatability

coefficients were obtained by principal component analysis (Abeywardena, 1972).

Structural analysis based on the correlation matrix between genotypes was used to estimate the repeatability coefficient, in accordance with Mansour et al. (1981). This evaluation is the arithmetic average of correlations between genotypes, considering measurements in pairs (Cruz et al., 2012). The number of measurements (η) to predict the real value of the selected clones in black wattle, considering coefficients of determination of 80, 85, 90, 95 and 99%, was calculated in accordance with Cruz et al. (2012) by the following expression:

$$\eta = r \cdot (R^2 \cdot (1-r))/((1-R^2))$$

The genotypic coefficient of determination (R2) represents the percentage of assurance regarding real value prediction for selected individuals based on the number of measurements. It was obtained by the following expression:

$$R^2 = (\eta \cdot r)/(1+r \cdot (\eta-1))$$

Genetic dissimilarity for the biometric traits of the seeds

Matrices of dissimilarity, residual covariance and genotype averages were produced by employing multivariate analysis. The canonical variables were calculated, as reported by Rao (1952), in order to identify similar genotypes in a two-dimensional scatter plot. The Mahalanobis distance (D2) was used as dissimilarity measure. The criteria described by Singh (1981) were used to quantify the relative contribution of these traits in genetic divergence among clones. All analyzes were performed in GENES software (Cruz, 2013).

Results and Discussion
Estimate of the repeatability coefficient for the biometric traits of seeds

According to the analysis of variance, the differences were statistically significant (p <0.001) for four traits in the sample: longitudinal and transversal length, thickness, and hilum size. Seed weight differences were not statistically significant. Statistically different traits highlight the variability among samples; their average biometry was 4.75 mm length, 3.16 mm width, 2.37 mm thickness and 1.29 mm for hilum size. This significance shows that black wattle clones differ greatly. Therefore, it is possible to obtain selection gains with superior material. It also shows the importance of conducting repeatability studies for this species and the need to estimate the minimum required number of measurements for the statistical usage of this coefficient. Seed weight had a non-statistically significant average (0.00242 g) and demonstrated low-to-middle estimated r amplitude ($0.27 \leq r \leq 0.34$) (Table 1).

Table 1 - Repeatability coefficient (r); coefficient of determination (R²) calculated using the method of analysis of variance (ANOVA), principal component analysis based on the correlation matrix (PCC), principal component analysis based on the covariance matrix (PCCV) and structural analysis based on the correlation matrix (SA) for seed weight (SW), longitudinal length (LL) transversal length (TL), thickness (TH) and hilum size (HS) in 50 measurements using nine black wattle clones.

Trait	Analysis method							
	ANOVA		PCC		PCCV		SA	
	r	R²	r	R²	R	R²	r	R²
SW	0.27	95.02	0.34	96.35	0.34	96.30	0.28	95.16
LL	0.36	96.54	0.43	97.37	0.46	97.73	0.37	97.72
TL	0.19	92.34	0.30	95.93	0.32	95.90	0.20	92.56
TH	0.78	99.40	0.80	99.49	0.80	99.51	0.79	99.48
HS	0.45	97.56	0.50	98.04	0.51	98.08	0.46	97.68

By analyzing the high estimate variability of the repeatability coefficient for seed weight, it is possible to infer that this trait is highly influenced by the environment. Consequently, its use in the selection process for superior genotypes is complex, as the phenotypic value does not represent the genotypic value, since repeatability expresses the maximum value that heritability in the broad sense can achieve. When the variance provided by the permanent effects of the environment is minimized, the repeatability is closer to the estimate of heritability (Cruz et al., 2012). This result indicates that trait is susceptible to annual production variation, so it was not possible to measure the repeatability coefficient by employing analysis of variance.

Resende (2002) proposed the following classification for repeatability coefficient: high repeatability ($r \geq 0.60$); medium repeatability ($0.30 \leq r < 0.60$) and low repeatability ($r < 0.30$). Seed thickness was the trait with the highest *r* values ($0.78 \leq r \leq 0.80$), showing regular results in all measurements. In addition, it demonstrates the efficiency of the methods used to estimate the repeatability coefficient, since there was low variation of ratio among them.

Repeatability coefficient for hilum size ($0.45 \geq r \geq 0.51$) showed medium repeatability (Resende, 2002), which means this trait is subjected to environmental influence. It is important to highlight that besides this result, the coefficient of determination is highly reliable, and it can be applied to all other traits evaluated. Seed length ($0.36 \geq r \geq 0.46$) showed medium magnitude for repeatability, and transversal length ($0.19 \geq r \geq 0.32$) presented low magnitude for *r* values.

According to Ferreira et al. (1999), low estimates of the repeatability coefficient, generally lower than 0.4, make it difficult for the breeder to identify the best genotypic values based on the analyzes of the phenotypic means obtained. Therefore, the longitudinal and transversal length of black wattle seeds are biometric traits hardly used as selection parameter in species breeding programs, as they require higher coefficients for an effective genotype selection. For this reason, hilum size and seed thickness should be neglected in genotype selection.

Low repeatability coefficient values, obtained by means of analysis of variance can be explained by the high mean-square values. Cruz et al. (2012) reported that production fluctuates in many species, what suggests great variability from one year to another. As this effect can have different intensities and forms among genotypes, analysis of variance (used to estimate the repeatability coefficient) is not able to eliminate the experimental error component, and thus the repeatability index will be underestimated. This was noted for all traits using this method. Abeywardena (1972) reported that the most appropriate repeatability coefficient estimate is the one obtained by the PCA method, in which genotypes exhibit cyclical behavior during the evaluations, for a given trait. Among the methods, PCCV was the one with the highest repeatability coefficient values. Based on the results found, this is the best method to estimate the repeatability coefficient for biometric traits of black wattle seeds.

With regards to other methods, low or medium repeatability may be related to the reduced correlations observed in the traits evaluated in each clone. Therefore, the most suitable method to estimate the repeatability coefficient in black wattle seeds is principal component analysis based on the covariance matrix with coefficient of determination of 90%. This result is consistent with those obtained by Negreiros et al. (2008) and Nascimento Filho et al. (2009) in sweet orange and guarana, respectively. According to Neves

et al. (2010), when working with passion fruit families, PCA with a coefficient of determination of 80% was the most suitable for repeatability studies, since when using a coefficient of 90%, fruit length was the only trait to present plausible measurements among six biometric traits evaluated.

As for the number of measurements required for different coefficients of determination (80, 85, 90, 95 and 99%) (Table 2), it was verified that, for all traits evaluated, using a coefficient of determination above 90% resulted in no higher experimental accuracy. Therefore, a coefficient of determination of 90% was used as reference, thus requiring a smaller sample of work. Moreover, it is important to emphasize that using the PCCV method resulted in values higher than those of other methods used for measuring the repeatability coefficient. That is, by using this method, it was possible to reduce in 40% the time needed to perform this work.

Table 2 - Number of measurements needed for different coefficients of determination calculated using the methods of analysis of variance (ANOVA), principal component analysis based on correlation matrix (PCC), principal component analysis based on the covariance matrix (PCCV) and structural analysis based on the correlation matrix (SA) for seed weight (SW), longitudinal length (LL), transversal length (TL), thickness (TH) and hilum size (HS)

Traits	Analysis method				
	R(%)	ANOVA	PCC	PCCV	AS
SW	80	11	8	8	11
	85	15	11	11	15
	90	24	17	18	23
	95	50	36	36	49
	99	250	187	188	252
LL	80	8	6	8	7
	85	11	8	11	10
	90	17	13	17	16
	95	34	26	34	33
	99	177	134	177	168
TL	80	17	10	9	17
	85	24	14	13	23
	90	36	22	20	37
	95	79	45	41	77
	99	411	232	212	398
TH	80	2	2	1	2
	85	2	2	2	2
	90	3	3	3	3
	95	6	5	5	5
	99	28	26	25	27
HS	80	5	4	4	5
	85	8	6	6	7
	90	12	9	9	11
	95	24	19	19	23
	99	124	99	97	118

Genetic dissimilarity for the biometric traits of seeds

According to the method proposed by Singh (1981) regarding the relative contribution of the traits used to study genetic divergence among genotypes, it was verified that seed weight contributes less to the analysis of diversity than other traits (Table 3). These results corroborate those found by Nick et al. (2010) in a study of genetic divergence in manioc samples. According to Oliveira et al. (2004), traits such as seed weight, which scarcely contribute to diversity studies should be kept in future works, since no previous information about them in black wattle has been found. Therefore, over time, it must be verified if this trait can remain in or must be excluded from repeatability studies in production cycles.

Table 3 - Relative contribution of five biometric traits of seeds evaluated in 50 seeds of nine black wattle clones.

Biometric traits	Relative contribution %
Seeds weight	0.4868
Longitudinal length	8.7276
Transversal length	5.0383
Thickness	71.5806
Hilum size	14.1667

Unweighted Pair Group Method with Arithmetic Mean (UPGMA) cluster analysis divided the clones into two groups based on their similarity. Results infer that clones seven and eight are the most different as they form a group far from the other clones (Figure 1). UPGMA is indicated for quantitative data, considering standard deviation and an invariant scale (Moreira et al., 2014). Taking into account that quantitative data is assessed, this measure is highly recommended in order to explain the results. Genetic divergence analysis has been successfully used in studies aimed to check similarity among genotypes, such as *Persica prunnus* (Silva et al, 2015); *Capsicum chinense* (Vasconcelos et al., 2012); and *Euterpe edulis* (Yokomizo et al., 2012).

Graphic dispersion of the principal component analysis (Figure 2), using thickness and hilum size (which correspond to more than 85.74% of the total variation among genotypes), was consistent with the genetic similarity clustering based on Mahalanobis distance. It confirmed that clones seven and eight are genetically different from each other and constitute a distinct group from the others. When two canonical variables explain over 80% of the total variation, they can be used in genetic divergence studies (Cruz et al., 2012). Both methods demonstrated the same dissimilarity among the studied mother trees and highlighted the genetic divergence among clones.
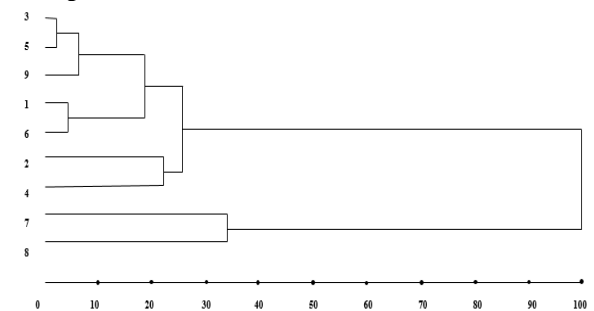

Figure 1 - Dissimilarity dendrogram constructed by means of UPGMA cluster analysis using five biometric traits of black wattle seeds. Vertical numbers are clones identification and horizontal numbers are percentage of dissimilarity.

In tree breeding programs, it is desirable to prioritize materials with high average and wider genetic diversity.

Therefore, breeding contrasting materials enables the identification of the ones that best express the non-additive fraction of genetic variance (Nascimento et al., 2014). When different groups of clones are crossed, it is possible to obtain superior individuals in the following generation, ensuring maximum genetic gains in the next breeding cycle. With regards to black wattle, diallel crossing is the strategy used to ensure that desirable traits can be obtained and thus, reliable decisions can be made in breeding programs of this species.
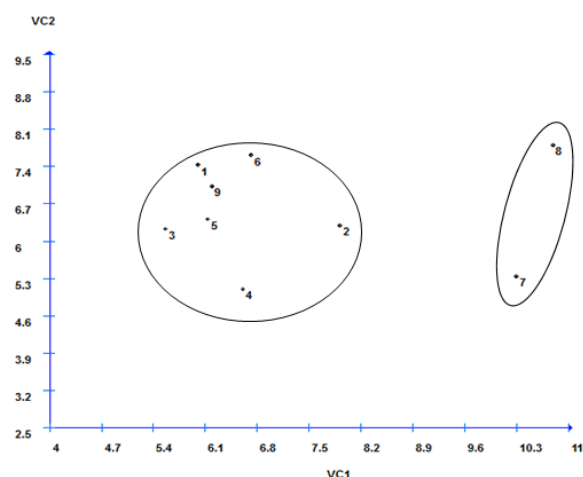

Figure 2 - Graphical dispersion (2D) of the score of nine clones, using two canonical variables, set by linear combination of five biometric traits of black wattle seeds.

Conclusion

Principal component analysis based on the covariance matrix is the most effective method to estimate the repeatability coefficient for black wattle seeds since it results in the highest values. Twenty measurements are needed in order to achieve a coefficient of determination of 90%.

Based on the measures of genetic dissimilarity using Mahalanobis distance and graphic dispersion methods, it was possible to differentiate clones and to recognize two genetically distinct groups.

Acknowledgements

References

Abeywardena, V (1972) An application of principal component analysis in genetics. *Journal of genetics* 61: 27-51. doi: 10.1007/BF02984099

Cavalcante M, Lira MA, Santos MVF, Pita EBAF, Ferreira RLC, Tabosa JN (2012) Coeficiente de repetibilidade e parâmetros genéticos em capim-elefante. *Pesquisa Agropecuária Brasileira* 47 (4): 569-575.

Cruz CD, Regazzi AJ, Carneiro PCS (2012) *Modelos biométricos aplicados ao melhoramento genético* 4th Edition Volume 1 Viçosa - MG: Editora UFV

Cruz CD (2013) GENES – a software package for analysis in experimental statistics and quantitative genetics. *Acta Scientiarum Agronomy* 35 (3):271-276. doi: 10.4025/actasciagron.v35i3.21251.

Ferreira RP, Botrel MA, Pereira AV, Cruz CD (1999) Avaliação de cultivares de alfafa e estimativas de repetibilidade de caracteres forrageiros. *Pesquisa Agropecuária Brasileira* 34 (6): 994-1002. doi: 10.1590/S0100-204X1999000600010

Ferreira RP, Vasconcelos ES, Cruz CD, Barioni Júnior N, Rassini JB, Freitas AR, Vilela D, Moreira A (2010) Determinação do coeficiente de repetibilidade e estabilização genotípica das características agronômicas avaliadas em genótipos de alfafa no ano de estabelecimento. *Revista Ceres* 57 (5): 642-647.

Grein A, Silva BC, Wendel CF, Tischer CA, Sierakowski MR, Moura ABD, Iacomini M, Gorin PAJ Somas-Tosin FF, Riegel-Vidotti IC (2013) Structural characterization and emulsifying proporties of polysaccharides of *Acacia mearnsii* De Wild gum. *Carbohydrate Polymers* 93: 313-320. doi: 10.1016/j.carbpol.2012.09.041

Lourenço IP, Figueiredo RW, Alves RE, Aragão FAZ, Moura CFH (2013) Caracterização de frutos de genótipos de muricizeiros cultivados no litoral cearense. *Revista Ciência Agronômica* 44 (3): 499-504. doi: 10.1590/S1806-66902013000300011

Manfio CE, Motoike SY, Santos CEM, Pimentel LD, Queiroz V, Sato AY (2011) Repetibilidade em características biométricas do fruto de macaúba. *Ciência Rural* 41 (1):70-76. doi: 10.1590/S0103-84782011000100012

Mansour H, Nordheim EV, Rutledge JJ (1981) Estimators of repeatability. *Theorical Applied Genetics* 60:151-156. doi: 10.1007/BF00264520

Matsuo E, Sediyama T, Cruz C, Oliveira RCT (2012) Análise de repetibilidade em alguns descritores morfológicos para soja. *Ciência Rural* 2 (2):189-196. doi: 10.1590/S0103-84782012000200001

Mochiutti S (2007) *Produtividade e sustentabilidade de plantações de acácia-negra (Acacia mearnsii De Wild.) no Rio Grande do Sul*. Tese. Universidade Federal do Paraná, Curitiba. 286 p.

Moreira JP, Shimizu JY, Sousa VA, Moraes MLT, Moura NF, Aguiar AV (2014) Ganho esperado na seleção de progenies de *Pinus elliottii var. elliottii* em idade precoce para produção de madeira. *Pesquisa Florestal Brasileira* 34 (78): 99-109. doi: 10.4336/2014.pfb.34.78.488.

Nascimento WMO, Gurgel FL, Bhering LL, Ribeiro OD (2014) Pré-melhoramento de camucamuzeiro: estudo de parâmetros genéticos e dissimilaridade. *Revista Ceres* 61(4): 538-543. doi: 10.1590/0034-737X201461040013

Nascimento Filho FJ, Atroch AL, Cruz CD, Carneiro PCS (2009) Repetibilidade da produção de sementes em clones de guaraná. *Pesquisa Agropecuária Brasileira* 44(6): 605-612. doi: 10.1590/S0100-204X2009000600009

Negreiros JRS, Saraiva LL, Álvares TK, Roncatto G (2008) Estimativas de repetibilidade de caracteres de produção em laranjeiras-doces no Acre. *Pesquisa Agropecuária Brasileira* 43: 1763-1768. doi: 10.1590/S0100-204X2008001200017

Neves LG, Bruckner CH, Cruz CD, Barelli MAA (2010) Avaliação da repetibilidade no melhoramento de famílias de maracujazeiro. *Revista Ceres* 57(4): 480-485. doi: 10.1590/S0034-737X2010000400007

Nick C, Carvalho SP, Jesus MAS, Custódio TN, Marim BG, Assis LHB (2010) Divergência genética entre subamostras de mandioca. *Bragantia* 69 (2): 289-298. doi: 10.1590/S0006-87052010000200005

Oliveira ACB, Sediyama MAN, Pedrosa MW, Garcia NCP, Garcia SLR (2004) Divergência genética e descarte de variáveis em alface cultivada sob sistema hidropônico. *Acta Scientiarum* 26: 211-217. doi: 10.4025/actasciagron.v26i2.1894

De Paula SRP, Nogueira AC, Bastos COM (2015) Physiological quality and repeatability in biometric characters of seeds from different trees of *Aspidosperma Polyneuron* Müll.Arg. *Australian Journal of Basic and Applied Sciences* 9(31): 147-154.

Rao RC (1952) *Advanced statical methods in biometric research*. New York: J. Wiley 330p.

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

Silva JOC, Cremasco JPG, Matias RGP, Silva DFP, Salazar AH, Bruckner CH (2014) Divergência genética entre populações de pessegueiro baseada em características da planta e do fruto. *Ciência Rural* 44 (10): 1770-1775. doi: 10.1590/0103-8478cr20131010

Silva, R.G.; Chaves, M.C.L.; Arnhold, E.; Cruz, C.D. Repetibilidade e correlações fenotípicas de caracteres do fruto de bacuri no Estado do Maranhão. *Acta Scientiarum Agronomy*, Maringá, v.31, n.4, p.587-591, 2009. doi: 10.4025/actasciagron.v31i4.723

Singh, D (1981) The relative importance of characters affecting genetic divergence. *The Indian Journal of Genetics e Plant Breeding 4*1: 237-245.

Vasconcelos CS, Barbieri RL, Neitzke RS, Priori D, Fischer SZ, Mistura CC (2012) Determinação da dissimilaridade genética entre acessos de *Capsicum chinense* com base em características de flores. *Revista Ceres* 59 (4):493-498. doi: 10.1590/S0034-737X2012000400009

Vencovsky R (1977) *Princípios de genética quantitativa*. Piracicaba: ESALQ. 97p.

Yokomizo GK, Queiroz JAL, Cavalcante ES, Pinheiro IN, Silva PAR, Mochiutti S (2012) Caracterização fenotípica e genotípica de progênies de *Euterpe oleracea* coletadas no Afuá-Pará nas condições do Amapá. *Cerne* 18 (2): 205-213. doi: 10.1590/S0104-77602012000200004