Correlation and repeatability in progenies of African oil palm

Digner Santiago Ortega Cedillo[1], Willian Silva Barros[2], Fábio Medeiros Ferreira[2], Luiz Antônio dos Santos Dias[3], Rodrigo Barros Rocha[4] and Cosme Damião Cruz[3*]

[1]*Instituto Nacional Autônomo de Investigações Agropecuárias, Santo Domingo de los Colorados, Pichincha, Ecuador.* [2]*Universidade Federal do Amazonas, Campus Itacoatiara, R. Nossa Senhora do Rosário, 3494,69100-00, Itacoatiara, Amazonas, Brasil.* [3]*Universidade Federal de Viçosa,, Viçosa, Minas Gerais, Brasil.* [4]*Centro de Pesquisa Agroflorestal de Rondônia, Empresa Brasileira de Pesquisa Agropecuária, Porto Velho, Rondônia, Brasil. *Author for correspondence. E-mail: cdcruz@ufv.br*

ABSTRACT. The phenotypic, genotypic and environmental correlations between the number of palm fruit bunches and the annual yield and their combinations in African palm were estimated. There were also estimated the coefficients of repeatability and determination, with the objective of indicating the lowest number of successive harvests required to exploit the genotypic potential of the evaluated progenies. The genotypic correlation between the number of bunches per plant and the yield was negative, indicating that when genotypes with a higher yield are selected the number of bunches decreases. In general, the signal of the genotypic correlation differed was opposite to the one from the environmental correlation, evidencing that genetic and environmental factors influence the number of bunches and yield through different physiological mechanisms. Based on the method of principal components applied to the correlation matrix, at least four years of evaluation seem to be necessary, values coefficients of reliability and repeatability were 87.6% e 0.64, respectively.

Key words: *Elaeis guineensis*, genotype stabilization, plant breeding.

RESUMO. Correlações e repetibilidade em progênies de dendê. Estimaram-se as correlações fenotípicas, genotípicas e ambientais entre o número de cachos por planta e a produção anual e suas combinações em dendê (*Elaeis guineensis*). Os coeficientes de repetibilidade e de determinação também foram estimados, com o objetivo de indicar o período mínimo de colheita necessário para se acessar o potencial genotípico das progênies avaliadas. A correlação genotípica entre o número de cachos e a produção foi negativa, indicando que, se forem selecionados os genótipos com maior produção, o número de cachos deverá diminuir. A correlação genotípica geralmente apresentou sinal oposto do que precedia correlação ambiental, evidenciando que as causas genéticas e ambientais influenciam o número de cachos e a produção por meio de mecanismos fisiológicos distintos. Com base no método de componentes principais aplicado à matriz de correlação, o período mínimo de colheitas foi de quatro anos, com confiabilidade de 87,6% e coeficiente de repetibilidade de 0,64.

Palavras-chave: *Elaeis guineensis*, estabilização genotípica, melhoramento genético vegetal.

Introduction

Oil palm (*Elaeis guineensis*), also known as African palm, is native to the west coast of Africa and is cultivated mainly for its high oil production (Quesada, 2000). The crossing of varieties Dura and Pisifera results in the hybrid Tenera, which is commercially used for oil extraction. The oil palm adapts well to the wet tropical climate of the coastline (Pandolfo, 1981). Mean temperatures between 25 and 27°C, rainfall of 1800 to 2000 mm year^{-1}, luminosity over 1,800 hours light^{-1} year^{-1}, and relative air humidity around 80% are development-friendly conditions for this crop. Temperature alters the expression of traits such as number of leaves, number of fruit bunches and fruit oil content, while water availability favors a higher number of bunches (Quesada, 2000).

Oil palm is widely cultivated in Ecuador, especially along the coastline, on the highland and in the eastern part of the country. The largest cultivation area of African palm is the coastal region, where the mean temperature is 25°C, rainfalls (from 1200 to 1500 mm year^{-1}) are unevenly distributed over the year, luminosity varies from 650 to 750 hours light^{-1} year^{-1} and humidity reaches 85%, depending on the locality. In general, an oil palm has 24 to 30 leaves, attains a mean yield of 12 bunches per year and gains 35-75 cm year^{-1} in height, attaining a height of 12 - 15 m after 15 years. The harvesting from the bunches of the adult plant is therefore hampered by the tree height. The weight per bunch varies from 20 to 30 kg, and the oil yield from 5 to 8 tons ha^{-1} year^{-1} (Quesada, 2000).

Knowledge of the association between traits is of great importance to the selection practice in improvement programs, with the aim of obtaining indirect gains and maintaining genetic variability of non-selected traits. The correlation coefficient is a measure of the linear relation between two variables through the measurement of change in one variable possible, in function of the change in another (Steel *et al*., 1997). This coefficient varies from -1 to +1. It is positive when the two variables vary in the same direction and negative when in opposite direction. The genetic correlation involves associations of inheritable nature, and is consequently of greater importance to improvement programs.

The measure of consistency of the relative position of the trees during successive measurements has traditionally been called repeatability (Turner *et al.*, 1969; Lerner, 1977). In perennial species, it is expected that the initial superiority after selection of a certain genotype would persist throughout the cycle. The truthfulness of this expectation can be verified by the repeatability coefficient of the trait under study (Cruz *et al*., 2004). The repeatability coefficient is used in genetic improvement as the superior limit of heritability and as a criterion to evaluate the effectiveness of the selection process (Lush, 1964). From the statistical point of view, repeatability can be defined as the correlation between the measurements in the same tree whose evaluations were repeated in time or space (Hansche, 1983).

For the estimation of the repeatability coefficient, diverse methodologies are presented, some of which are described in studies presented by Abeywardena (1972), Mansour *et al*. (1981) and Cruz *et al*. (2004). Abeywardena (1972) demonstrated that the most adequate estimate of the repeatability coefficient when the genotypes of perennial species present cyclic performance throughout the evaluations in relation to the studied trait is the one obtained through the method of principal components. In practice, the repeatability coefficient is used to determine the minimum number of multiple measurements that must be taken in each tree in order to be able to perform selection with a certain degree of efficacy, as well as minimal costs and labor (Cruz *et al*., 2004).

In this study the phenotypic, genotypic and environmental correlations and the repeatability coefficient were estimated. The minimum period of successive harvests required to evaluate the yield potential of a genotype using the repeatability estimate obtained by principal components was estimated as well.

Material and methods

The present study was realized at the Santo Domingo Experimental Station at INIAP (Instituto Ecuatoriano de Investigaciones Agropecuarias – Ecuador), with Dura full-sib progenies. The experiment was set up in May 1989 and evaluated five progenies of the crossing Dura x Dura, plus one common control of the crossing Dura x Pisifera (Tenera). The Santo Domingo Experimental Station lies 38 km away from Quinindé (lat 00° 06’ N, long 79° 20’ O, alt 300 m asl). The collected experimental data were the number of bunches and the five-year yield (1992-1996). Estimates of the phenotypic, genotypic and environment correlations between the number of bunches and yield in their different annual combinations (Cruz *et al*., 2004) were obtained. The repeatability coefficient for yield was evaluated based on the method of principal components using Genes software (Cruz, 2001).

Results and discussion

In general, the phenotypic or genotypic and environmental correlations in the evaluated year and in combined years between number of bunches and yield did not agree in magnitude and direction (Table 1), demonstrating that the influence of the environment on the association between the traits is high. According to Cruz *et al*. (2004), if two traits present a favorable genetic correlation, it is possible to obtain gains for one of them by indirect selection in the other associated trait, and that this selection can often result in quicker progress than a direct selection of the desired trait. In this case, the genotypic correlation in most years was negative and the magnitude oscillated from -0.10 to -0.77 and in combined years from -0.16 to -0.61 (Table 1), indicating that selection for number of fruit bunches reduces the yield.

In the case of the environmental correlation, the estimates presented magnitudes over 0.80. Only in 1996 was this correlation a little lower (0.54), but still positive (Table 1). These correlation values indicate that the number of bunches can be favored or impaired as much as the yield by environmental variations, such as rain and luminosity, which are two climatic factors of great importance to the crop.

Rain stimulates the growth of inflorescences of oil palm, and high luminosity causes bunches to ripen within five months. These conditions are given in the rainy period (December to May).

Table 1. Estimates of the phenotypic ($\hat{r}_p$), genotypic ($\hat{r}_g$) and environmental ($\hat{r}_e$) correlation coefficients between number of bunches (B) and yield (Y), evaluated in full-sib of oil palm families, 1992-96.

Traits	$\hat{r}_p$	$\hat{r}_g$	$\hat{r}_e$
B1992-Y92	0.40	*	0.87
B1993-Y93	0.02	-0.10	0.88
B1994-Y94	0.26	0.21	0.80
B1995-Y95	-0.40	-0.77	0.82
B1996-Y96	-0.49	-0.72	0.54
B-Y1992-93	-0.01	-0.16	0.87
B-Y1992-94	-0.19	-0.30	0.89
B-Y1992-95	-0.42	-0.61	0.89
B-Y1992-96	-0.44	-0.61	0.83

* undetermined correlation, where the genetic variance was zero or negative.

The equally important genetic factors that determine the number and size of bunches must also be considered. In general, in the first years, the bunches are small and abundant. The number of fruit bunches decreases, but bunches are larger and heavier as the plant develops (Lim and Toh, 1985; Henson, 1993). The genotypic correlation, in general, presented a different direction from the environmental correlation, which was positive. According to Falconer (1989), it is evident in such cases that the causes of the genetic and environmental variations that influence the traits are determined by different physiological mechanisms.

Table 2 displays the phenotypic and genotypic correlation estimates for number of bunches in different years and their combinations. The phenotypic correlation for all years was positive and strong, with exception of 1992. In the case of the genotypic correlation in 1992 the estimates were zero, due to the estimate of negative genetic variance in this year. For the other years, this correlation was strong and positive, evidencing that the genotype performances in the different years of evaluation are consistent in relation to the number of bunches.

Phenotypic correlation estimates of high magnitude were found in 1992, 1993 and 1996 for yield (Table 3). The highest correlation in relation to individual years was found in 1993. In 1994, negative correlation estimates of low magnitude were detected. This situation is possibly due to the fact that from 1993 on, the yield increased (12 to 18 tons ha^{-1} $year^{-1}$) compared to the first harvest, which was low (5 to 8 tons ha^{-1} $year^{-1}$). In the combinations of years, this correlation was positive and below 0.68. A similar result to the phenotypic was found for the genotypic correlation, expressing a certain inconsistency in the estimates from one year to the other. The inconsistency in yield from one year to the other is mainly result of the two climatic periods that occur at the site where the experiments were undertaken.

Table 2. Estimates of the phenotypic (above the diagonal) and genotypic (below the diagonal) correlation coefficients, for number of bunches in individual and combined years, evaluated in full-sib oil palm families, during 1992-96.

	1992	1993	1994	1995	1996	1992-93	1992-94	1992-95	1992-96
1992	-	0.56	0.15	0.18	0.34	0.70	0.43	0.37	0.36
1993	-	-	0.78	0.82	0.89	0.98	0.94	0.93	0.92
1994	-	0.83	-	0.88	0.94	0.71	0.94	0.94	0.95
1995	-	0.89	0.92	-	0.97	0.75	0.89	0.94	0.95
1996	-	0.96	0.96	1.00	-	0.84	0.97	0.99	0.99
1992-93	-	1.00	0.78	0.83	0.93	-	0.91	0.88	0.88
1992-94	-	0.96	0.96	0.93	1.00	0.93	-	0.99	0.99
1992-95	-	0.95	0.96	0.97	1.00	0.91	0.99	-	1.00
1992-96	-	0.95	0.96	0.97	1.00	0.91	0.99	1.00	-

* * undetermined correlation, where the genetic variance was zero or negative.

Table 3. Estimates of the phenotypic (above the diagonal) and genotypic correlation coefficients (below the diagonal), for yield in individual years and combined years, evaluated in full-sib oil palm families from 1992-96.

	1992	1993	1994	1995	1996	1992-93	1992-94	1992-95	1992-96
1992	-	0.88	0.24	0.47	0.74	0.95	0.84	0.82	0.82
1993	0.95	-	0.37	0.69	0.90	0.98	0.92	0.97	0.98
1994	0.18	0.42	-	-0.08	-0.01	0.33	0.68	0.48	0.36
1995	0.54	0.77	-0.32	-	0.91	0.62	0.45	0.73	0.80
1996	0.88	1.00	-0.13	1.00	-	0.86	0.67	0.86	0.93
1992-93	0.98	0.99	0.33	0.69	1.00	-	0.92	0.94	0.95
1992-94	0.85	0.95	0.67	0.41	0.77	0.92	-	0.94	0.89
1992-95	0.87	1.00	0.42	0.69	1.00	0.98	0.95	-	0.99
1992-96	0.87	1.00	0.28	0.80	1.00	0.99	0.90	1.00	-

By the method of principal components obtained from the correlation and co-variance matrices, the respective estimates of the coefficients of repeatability were 0.69 and 0.74, with corresponding determination coefficients of 91.8 and 93.6%. According to Abeywardena (1972), the principal components method is most appropriate for the estimation of repeatability when the genotypes present cyclic performance during the evaluations, in relation to the studied trait. In the case of oil palm, it is known that the plant produces a lot in one year, mainly in the rainy period, which begins in December. In this period, a large number of inflorescences emerge, which result in a higher yield. The bunches attain maturity after five months, so the first bunches would appear in April or May and the others in the following months, depending on the growth of the other inflorescences. In this first year, the yield is therefore normally high. However, new inflorescences appear in lower number in the months of drought season (June to December), increasing the probability of a low yield in the following year. Based on the method of principal components obtained from the correlation matrix, and considering a value of R^2 of 90%, 4 years of successive harvests to assess the phenotypic value of the progenies would be necessary. This period could be reduced to 3 years if the repeatability is estimated by principal components based on the covariance matrix.

Table 4. Estimates of the coefficients of repeatability (r), of determination (R^2) and number of years of successive harvests (n) required to obtain different R^2, in relation to the yield of oil palm full-sib families, during 1992-96.

			n and arbitrary R^2 values				
Method	r	R^2	80	85	90	95	99
PCA[1]	0.69	91.8	2	3	4	9	44
PCA[2]	0.74	93.6	1	2	3	7	34

[1] Principal components obtained through the correlation matrix; [2] Principal components obtained through the covariance matrix.

Table 5 presents results of the study of genotypic stabilization, based on combinations of two, three, four and all five evaluated years. This analysis aimed to determine the minimal number of successive harvests, taking the repeatability and determination coefficient into consideration. According to Cedillo (2003) who used different methods to estimate the repeatability and determination coefficients in oil palm, the method of the principal components presented the highest coefficients compared to the other methods, corroborating the cyclic performance of the oil palm yield.

Table 5. Minimal period of years of successive harvests (n), repeatability (r) and determination (R^2), in relation to the yield of full-sib families of oil palm.

Years	n	r[1]	R^2 (%)
1992-93	2	0.89	94.4
1993-94	2	0.60	74.8
1994-95	2	0.37	53.5
1995-96	2	0.92	95.7
1992-94	3	0.65	85
1993-95	3	0.61	82.3
1994-96	3	0.56	79.2
1992-95	4	0.64	87.6
1993-96	4	0.68	89.6
1992-96	5	0.69	91.8

[1] Estimated by principal components based on the correlation matrix.

For the first two years of evaluation, 1992/93, the repeatability and determination coefficients, established at 0.89 and 94.4%, respectively, indicated good consistency in the family performance evaluated in these two years. For the following biennials 1993/94 and 1994/95 on the other hand, the estimates of repeatability, considered relatively low, were 0.60 and 0.37, respectively, indicating a lower consistence of family performance in these years of evaluation. It was also observed that for the biennial 1995/96 the repeatability estimate was 0.92 (R^2 = 95.7%), which could indicate the beginning of yield stabilization, since in these harvests yield values still attained around 20 tons/ha. This expresses that the evaluation in 1996 was well represented by the one of 1995. It is therefore safe to conclude that the minimal necessary period of successive harvests to assess the phenotypic value of the families is 4 years from the third year on after planting.

Similar results were found by Bastidas (1989), who mentions 4 years as the time required for the selection of parent plants. His study was based on the estimates of the yield correlation coefficient between plants and years.

The results supports that the period of 4 of years for harvesting is most favorable to the selection of these genotypes in the environmental conditions evaluated in this work. In *Anacardium orcidentale* clones per example, the most favorable period to the selection for cashew nuts production is 3 year-old trees (Cavalcanti *et al.*, 2000).

Conclusion

The genotypic correlation between number of bunches and production is negative, soon the selection of the genotypes of larger production results in individuals with smaller number of bunches. In a general way, the sign different of the genotypic correlation in relation the environmental correlation, evidences that the genetic and environmental causes have influences in the number of bunches and in the production through different physiologic mechanisms.

The genotypic correlations of the variable number of bunches among individual years are high and of positive sign. However, for the character production these correlations present inconsistency in certain years.

For the method of principal components obtained of correlation matrix the number of minimum measurements to evaluate the genotypes is four years.

Acknowledgments

We are also indebted to CNPq, FAPEMIG, FAPEAM and INIAP for financial support.

References

ABEYWARDENA, V. An application of principal component analysis in genetics. *J. Genet.*, Sadashivanagar, v. 61, n. 1, p. 27-51, 1972.

BASTIDAS, S.P. Sistema de melhoramento em palma africana em el MIRA-TUMACO. *In*: CURSO CORTO, 11., 1989. Nariño. *Anais...* Nariño: Prociandino, 1989. p. 19-47.

CAVALCANTI, J.J.V. *et al.* Repetibilidade de caracteres de produção e porte da planta em clones de cajueiro-anão precoce. *Pesq. Agropecu. Bras.*, Brasília, v. 35, n. 4, p. 773-777, 2000.

CEDILLO, D.S.O. *Análises biométricas aplicadas ao melhoramento de dendê* (*Elaeis guineensis*, Jacq). 2003. Dissertação (Mestrado em Genética e Melhoramento)-Universidade Federal de Viçosa, Viçosa, 2003.

CRUZ, C.D. *Programa GENES*: versão Windows - aplicativo computacional em genética e estatística. Viçosa: UFV, 2001.

CRUZ, C.D. *et al. Modelos biométricos aplicados ao melhoramento genético*. 3. ed. Viçosa: UFV, 2004.

FALCONER, D.S. *Introduction to quantitative genetics*. London: Longman, 1989.

HANSCHE, P.E. Response to selection. *In*: MOORE, J.N.; JANICK, J. (Ed.). *Methods in fruit breeding*. West Lafayette: Purdue University Press, 1983. p. 154-171.

HENSON, Y.E. Factors determining mesocarp oil bunch ratio in the oil palm: A physiological perspective. *In*: NATIONAL SEMINAR ON PALM OIL EXTRACTION RATE: PROBLEMS AND ISSUES, 1993, Porim. *Proceedings*… Porim: Palm Oil Research Institute of Malaysia, Ministry of Primary Industries, 1993. p. 27-35.

LERNER, I.M. *The genetics basis of selection*. New York: John Wiley & Sons, 1977.

LIM, K.H.; TOH, P.Y. The accuracy and precision of bunch analysis. *In*: SYMPOSIUM ON IMPACT OF THE POLLINATING WEEVIL ON THE MALAYSIAN OIL PALM INDUSTRY, 1985, Kuala Lumpur. *Proceedings*... Kuala Lumpur: Kementerian Perusahaan Utama, 1985. p. 91-109.

LUSH, J.L. *Melhoramento dos animais domésticos*. Rio de Janeiro: Seagra, 1964.

MANSOUR, H. *et al*. Estimators of repeatability. *Theor. Appl. Genet*., Berlin, v. 60, n. 3, p. 151-156, 1981.

PANDOLFO, C. *A cultura do dendê na Amazônia*. Belém: SUDAM, 1981.

QUESADA, G. *Cultivo e industria de la palma africana*: documento preliminar para guía del productor. Viernes: Programa Nacional de Palma Africana, 2000.

STEEL, R.G.D. *et al*. *Principles and procedures of statistics*: a biometrical approach. 3rd ed. New York: McGraw-Hill, 1997.

TURNER, H.N. *et al*. *Quantitative genetics in sheep breeding*. New York: Cornell University, 1969.

Received on July 31, 2006.
Accepted on July 24, 2007.