

**Comunicação**  
**Relationship between fruit traits of custard apple trees**  
**(*Annona squamosa* L.)**

*Keny Henrique Mariguele<sup>1</sup>, Paulo Sérgio Lima e Silva<sup>2</sup>*

**ABSTRACT**

The objective of this study was to estimate simple and partial coefficients of correlation, as well as to divide their effects into direct and indirect using path analysis for custard apple tree traits. Twenty half-sibling progenies were evaluated in a randomized block design with five replicates, and plots consisting of four plants. Six traits were evaluated in the first cropping season (mean number of seeds per fruit and mean weight of the pericarp, pulp, pedicel, seeds per fruit, and the whole fruit), while five traits were evaluated in the first three cropping seasons (mean fruit length and width, total number of fruits ha<sup>-1</sup>, mean fruit weight (in both types of analyses), and fruit yield in kg ha<sup>-1</sup>). The results of this work led to the conclusion that doing selection based on simple correlation estimates may not be convenient, since not always a cause and effect relationship can be verified between two traits. Positive correlations were obtained between number of seeds and seed weight, and between number of fruits and yield. The greatest direct effects were those obtained for pulp weight on fruit weight and for mean number and weight of fruits on fruit yield. The most important indirect effects were obtained for number of seeds and pericarp weight, obtained via pulp weight, on fruit weight, and for fruit length and width, obtained via mean fruit weight, on fruit yield.

**Key words:** Fruit traits , partial correlations, path analysis, simple correlations.

**RESUMO**

**Relação entre caracteres do fruto em pinheira (*Annona squamosa* L.)**

O objetivo deste trabalho foi estimar os coeficientes de correlações simples e parciais, e desdobrar os efeitos diretos e indiretos, pela análise de trilha, entre características de frutos de pinheira. Vinte progênies de meias-irmãs foram avaliadas no delineamento de blocos ao acaso com cinco repetições e parcelas de quatro plantas. Seis características foram avaliadas somente na primeira safra (número médio de sementes por fruto e pesos médios do pericarpo, polpa, pedúnculo, sementes por fruto e do fruto) e cinco características nas três primeiras safras (comprimento e largura médios do fruto, número total de frutos ha<sup>-1</sup>, peso médio do fruto (nos dois tipos de análises) e rendimento de frutos, kg ha<sup>-1</sup>). Concluiu-se que praticar seleção com base nas estimativas de correlação simples pode não ser eficiente, pois nem sempre a relação entre as duas características é de causa e efeito. Correlações positivas foram obtidas entre número médio de sementes - peso médio de sementes e número médio de frutos - rendimento. Os maiores efeitos diretos foram aqueles do peso da polpa sobre peso do fruto e do número e peso médio de frutos sobre o rendimento de frutos. Os principais efeitos indiretos foram de número de sementes e peso do pericarpo, via peso da polpa, sobre peso do fruto, e o do comprimento e largura do fruto, via peso médio do fruto, sobre o rendimento de frutos.

**Palavras-chave:** Análise de trilha, características do fruto, correlações simples, correlações parciais.

Recebido para publicação em dezembro de 2007 e aprovado em maio de 2010

<sup>1</sup> Engenheiro-Agrônomo, Mestre. Doutorando em Genética e Melhoramento. Universidade Federal de Viçosa. Av. Peter Henry Rolfs, s/n, 37500-000 Viçosa-MG, Brasil. kmariguele@hotmail.com

<sup>2</sup> Engenheiro-Agrônomo, Doutor. Universidade Federal Rural do Semi-Árido. C.P. 137, 59625-900 Mossoró-RN, Brasil. paulosergio@ufersa.edu.br

## INTRODUCTION

The custard apple is a typically tropical fruit tree, belonging to the genus *Annona*, of great economic significance in Brazil. Also known as sweetsop or sugar apple, this is an important crop in several states of the Northeastern and Southeastern regions of Brazil. Its commercial cultivation is particularly associated with the trade of fresh fruits in food/produce distribution centers of several of the country's major cities (Dias *et al.*, 2004). There has been increased interest in exploiting this crop under irrigation in the Northeast, since farmers are trying to expand the number of grown species, in order to retain labor and reduce problems with diseases and pests, which occur more frequently in monoculture crops. In addition, with the development of better communication and transportation means, new and extensive markets are emerging for a great variety of tropical and subtropical fruits (Silva *et al.*, 2002).

Breeding programs aim at improving cultivars with respect to a number of traits. For this reason, the prior knowledge on the nature and magnitude of the correlations between traits of interest is extremely important. If two traits are correlated, it is possible to obtain gains in one of them by the indirect selection of the other. This is advantageous especially when a trait of high economic value has low heritability or is difficult to evaluate, when compared with another trait associated with the first. As a result, selection can be accomplished based on the trait that has higher heritability or is easier to evaluate, but maintaining the breeding goal focused on the wanted trait (Ferreira *et al.*, 2003).

Simple correlation allows the magnitude and meaning of an association between two traits to be evaluated, without providing the required information about the direct and indirect effects of a group of traits in relation to a dependent variable of greater importance (Cruz, 2001). In certain cases, simple correlation coefficients may produce significant misunderstandings about the relationship that exists between two variables, and may not represent an actual measurement of cause and effect. Consequently, a high or low coefficient of correlation between two variables may result from the effect produced by a third variable or a by group of variables on those two variables. A more informative measure of the relationship between variables is the coefficient of partial correlation, which is estimated by removing the effects of other variables on the association under study. (Cruz *et al.*, 2004). Path analysis is a technique that allows the effects of several independent variables on a basic variable to be studied; estimates for those variables are obtained by means of regression equations, in which the variables are initially standardized (Vencovsky & Barriga, 1992; Cruz, 2001).

The objective of this study was to estimate simple and partial coefficients of correlation, as well as to divide their effects into direct and indirect using path analysis for custard apple traits.

## MATERIALS AND METHODS

Twenty half-sibling progenies were evaluated in a randomized block design with five replicates and plots consisting of four plants each. Six traits were evaluated in the first cropping season (mean number of seeds per fruit and mean weight of the pericarp, pulp, pedicel, seeds per fruit, and the whole fruit), while five traits were evaluated in the first three cropping seasons (mean fruit length and width, total number of fruits ha<sup>-1</sup>, mean fruit weight (in both types of analyses), and fruit yield in kg ha<sup>-1</sup>). These traits were evaluated for all fruits of each plant. Then, plot means were calculated in order to estimate the coefficients of correlation. Simple and partial correlations were estimated and a path analysis was conducted. Statistical and genetic analyses were performed using the Genes software (Cruz, 2007). Prior to conducting the path analysis, the multicollinearity between the independent variables (Cruz, 2001) must be diagnosed; according to Montgomery & Peck (1981), this multicollinearity is represented by the condition number (CN), calculated by the ratio of the largest to the smallest eigenvalue. After performing the diagnosis, the Pearson's coefficient of correlation for each independent variable was divided into direct and indirect effects on the dependent variable, as proposed by Wright (1921 and 1923) and detailed by Li (1975). Among the traits evaluated during the first cropping season, fruit weight was considered the dependent variable, whereas yield was considered the dependent variable among the traits evaluated over a three year period.

## RESULTS AND DISCUSSION

The eigenvalues of the correlation matrix were 3.37, 0.76, 0.55, 0.22, and 0.09 for traits evaluated in the first year. Eigenvalues for traits evaluated over three years were 2.65, 1.01, 0.26, and 0.07. Based on these eigenvalues, we obtained condition numbers of 36.3 for traits evaluated during one year, and 37.2 for traits evaluated over three years. Since these values were below 100, they indicate a weak multicollinearity, and a path analysis could be run, as the estimates were not biased.

Among the traits analyzed during the first year, significant simple correlations were found between all pairs of traits, except between pericarp weight/pedicel weight and pericarp weight/seed weight (Table 1). However, when the effects of other traits were removed from the model to estimate partial correlations, all correlations were non-significant, except for the correlation between seed number and seed weight (0.78),

**Table 1.** Simple and partial correlations between pairs of traits of twenty custard apple progenies

Pairs of traits	Simple correlation	Partial correlation
Number of seeds and pericarp weight	0.52*	0.18 <sup>ns</sup>
Number of seeds and pulp weight	0.67**	0.04 <sup>ns</sup>
Number of seeds and pedicel weight	0.65**	0.42 <sup>ns</sup>
Number of seeds and seed weight	0.88**	0.78*
Number of seeds and fruit weight	0.66**	0.01 <sup>ns</sup>
Pericarp weight and pulp weight	0.72**	0.17 <sup>ns</sup>
Pericarp weight and pedicel weight	0.43 <sup>ns</sup>	0.09 <sup>ns</sup>
Pericarp weight and seed weight	0.41 <sup>ns</sup>	-0.19 <sup>ns</sup>
Pericarp weight and fruit weight	0.76**	0.40 <sup>ns</sup>
Pulp weight and pedicel weight	0.45*	0.01 <sup>ns</sup>
Pulp weight and seed weight	0.64**	0.11 <sup>ns</sup>
Pulp weight and fruit weight	0.89**	0.65 <sup>ns</sup>
Pedicel weight and seed weight	0.52*	-0.12 <sup>ns</sup>
Pedicel weight and fruit weight	0.44*	-0.03 <sup>ns</sup>
Seed weight and fruit weight	0.63**	0.14 <sup>ns</sup>

\*\* , \* and <sup>ns</sup>: significant at 1 and 5% probability, and non-significant by the t test, respectively.

with a determination coefficient of 60.8%. For the traits evaluated over three years, only the correlation between number of fruits and fruit yield (0.99) was significant, showing, by means of the coefficient of determination, that 99.3% of the variation in yield is explained by the variation in number of fruits. These data illustrate the importance of estimating partial correlations to know the actual relationship between traits.

If the coefficient of correlation between a causal factor and the dependent variable is equal or similar to its direct effect in magnitude and signal, this correlation may explain the true association that exists between them. When a positive coefficient of correlation is obtained, but the direct effect is negative or negligible, the correlation is produced by the indirect effects, which are then taken into consideration in the analysis. When Pearson's coefficient of correlation is negligible and the direct effect is high and positive, then the indirect effects are responsible for the lack of correlation, equally deserving attention in the analysis. Finally, when a negative correlation is obtained, with a high and positive direct effect, the indirect effects should be removed from the

analysis, and only the direct effects should be used (Vencovsky & Barriga, 1992).

Despite the high magnitude of the totals between the explanatory variables over the main variable (fruit weight), the direct effect was high (0.61) only for pulp weight (Table 2). The most important indirect effects were for number of seeds via pulp weight (0.41) and rind weight via pulp weight (0.44). The coefficient of determination showed that 83.0% of the variation in fruit weight can be explained by the variation of number of seed, pericarp weight, pulp weight, pedicel weight and seed weight (Table 2). With respect to the influence of a variable on pulp weight in umbu trees (*Spondias tuberosa*), the best combination of a direct effect with a positive correlation was found for the variable fruit width (Santos & Nascimento, 1998). According to these authors, pulp weight can be increased via a positive selection for this trait or via a negative selection for plant height. The most important variable for rind weight and stone weight was fruit width.

The greatest direct effects on fruit yield and explanatory variables were number of fruits (0.99) and average fruit

**Table 3.** Estimates of direct and indirect effects of explanatory traits on the main trait fruit yield in 20 custard apple progenies

Explanatory traits	r	Effects on fruit yield (kg ha <sup>-1</sup> )				
		Direct	Indirect			
			FL	FW	NF	MFW
Fruit length (FL)	0.097	-0.022	-	0.018	-0.048	0.148
Fruit width (FW)	0.22	0.025	-0.016	-	0.064	0.145
Number of fruits (NF)	0.98	0.993	0.001	0.002	-	-0.013
Average fruit weight (MFW)	0.09	0.168	-0.019	0.022	-0.080	-
Coefficient of determination: 0.994						
Residual effect: 0.080						

**Table 2.** Estimates of direct and indirect effects of explanatory traits on the main trait (fruit weight) in 20 custard apple progenies

Explanatory traits	r	Effects on fruit weight (kg ha <sup>-1</sup> )					
		Direct	Indirect				
			NS	PW	PlpW	PedW	SW
Number of seeds (NS)	0.66	0.01	-	0.141	0.410	-0.009	0.116
Pericarp weight (PW)	0.76	0.27	0.003	-	0.443	-0.006	0.053
Pulp weight (PlpW)	0.89	0.61	0.004	0.195	-	-0.006	0.084
Pediceal weight (PedW)	0.44	-0.01	0.004	0.112	0.274	-	0.068
Seed weight (SW)	0.63	0.13	0.004	0.110	0.392	-0.008	-
Coefficient of determination:	0.83						
Residual effect:	0.41						

weight (0.17) (Table 3), whereas the most important indirect effects were fruit length (0.15) and fruit width (0.145), via fruit weight. Despite the small magnitude of the three last values it is important to point out that the estimate for the residual effect (0.08) suggests a good precision of the estimates for direct and indirect effects. The coefficient of determination indicated that 99.4% of the variation in fruit yield can be explained by the variation in this set of variables. It can be observed that partial correlation between fruit number and fruit yield and direct effect estimates obtained by the path analysis are in agreement.

In a similar study with grape vine, Leão & Costa (2003) found that the greatest direct effects on yield per plant were fruit diameter (0.5200) and number of bunches per plant (0.9434), while the most important indirect effects were mean bunch weight (-0.4012) and bunch length (-0.5700) via number of bunches per plant.

## CONCLUSION

The results of this work led to the conclusion that practicing selection based on simple correlation estimates may not be convenient, since not always a cause and effect relationship can be verified between both traits. Positive correlations were obtained between number of seeds and seed weight, and between number of fruits and fruit yield. The greatest direct effects were obtained for pulp weight on fruit weight and for mean number and weight of fruits on fruit yield. The most important indirect effects were observed for number of seeds and pericarp weight, obtained via pulp weight, on fruit weight, and for fruit length and width, obtained via mean fruit weight, on fruit yield.

## REFERENCES

- Cruz CD (2007) Programa Genes - Aplicativo computacional em genética e estatística.
- Cruz CD (2001) Programa Genes: versão Windows: aplicativo computacional em genética e estatística. Viçosa, UFV, 648p.
- Cruz CD, Regazzi AJ, Carneiro PCS (2004) Modelos biométricos aplicados ao melhoramento genético Volume 1. 3a. ed., Viçosa, UFV, 480p.

Dias NO, Souza IVB, Silva JCG da, SILVA KS, Bomfim MP, Alves JFT, Rebouças TNH, Viana AES, São José AR (2004) Desempenho vegetativo e reprodutivo da pinheira (*Annona squamosa* L.) em função de diferentes comprimentos de ramos podados. Revista Brasileira de Fruticultura, 26:389-391.

Ferreira MAJF, Queiroz MA, Braz LT, Vencovsky R (2003) Correlações genotípicas, fenotípicas e de ambiente entre dez caracteres de melancia e suas implicações para o melhoramento genético. Horticultura Brasileira, 21:438-442.

Leão PC de S, Costa JG da. (2003) Estimates of repeatability and path coefficients on grapes. Crop Breeding and Applied Biotechnology, 3:231-236.

Li CC (1975) Path analysis - a primer. The Boxwood Press, Pacific Grove, 346p.

Montgomery DC, Peck EA (1981) Introduction to linear regression analysis. New York, John Wiley & Sons, 504p.

Santos CAF, Nascimento CE de S (1998) Relação entre caracteres quantitativos do umbuzeiro (*Spondias Tuberosa* A. Câmara). Pesquisa Agropecuária Brasileira, 33:449-456.

Silva J da, Silva ES da, Silva PSL e. (2002) Determinação da qualidade e do teor de sólidos solúveis nas diferentes partes do fruto da pinheira (*Annona squamosa* L.). Revista Brasileira de Fruticultura, 24:562-564.

Vencovsky R, Barriga P (1992) Genética biométrica no fitomelhoramento. São Paulo, Sociedade Brasileira de Genética, 466p.

Wright S (1921) Correlation and causation. Journal Agriculture, v. 20:557-585.

Wright S (1923) The theory of path coefficient - a replay to Niles' criticism. Genetics, 8:239-255.