CBAB
CROP BREEDING AND
APPLIED BIOTECHNOLOGY

Correlations and path analysis in peanut

Regina Lucia Ferreira Gomes[1*], and Ângela Celis de Almeida Lopes[1]

ABSTRACT - *Genetic parameters of agronomical traits of peanut were estimated and the coefficients of genotypic correlation between the grain yield and the primary components of the yield apportioned into direct and indirect effects. The cultivars (Tatu, BR 1, L.7 Vermelha, CNPA 75 AM, CNPA 76 AM, CNPA 68 AM, L.8.14.12, L.8.14.01 and L.7 Bege) were evaluated in Teresina, state of Piauí. The highest estimates of the coefficient of genotypic determination and of the relation b were obtained for weight of 100 seeds, number of pods plot^{-1}, grain yield, number of seeds pod^{-1} and pod yield. The spliting of the genotypic correlations into grain yield and the primary components, in direct and indirect effects, shows that the grain yield was positively influenced by the number of pods plot^{-1} and weight of 100 seeds, and negatively by the number of seeds pod^{-1}. In a favorable sense, the number of pods plot^{-1} was the trait of greatest direct influence on the grain yield.*

Key words: *Arachis hypogaea*, coefficient of genotypic determination, correlations, path analysis.

INTRODUCTION

Peanut (*Arachis hypogaea* L.) is a leguminous plant of high nutritional value and is a protein and energy source in developing regions; it produces high quality oil, is rich in essential fatty acids, as well as in the vitamins E, B$_1$ and B$_2$, and has a easy digestibility. Compared with other oleaginous plants, it is relatively drought-resistant and of great socioeconomic importance in semi-arid regions.

In spite of the crop's remarkable potential, the expressivity of peanut is still low, mainly because of the restricted adoption of appropriate technologies and the low genetic quality of the cultivars farmers use (Santos et al. 1993). In the northeastern region, the cultivars planted are generally acquired on open markets or reused from preceding plantings, limiting the crop yield (Santos 1995).

The improvement programs of peanut have achieved considerable progress in the last decades, however, very little of the existing broad genetic variability in the germplasm of the species has been used (Upadhyaya et al. 2002).

The estimation of genetic parameters represents a useful tool in improvement programs, allowing an

[1]Departamento de Biologia, Centro de Ciências da Natureza, Universidade Federal do Piauí, Campos Ministro Petrônio Portela, BL-02, FG-1, Ininga, 64.049-550, Teresina, PI, Brasil. *E-mail: rlgomes@ufpi.br

evaluation of the potential of the variability presented by the study populations.

Traits of agronomical importance can be correlated to each other in different magnitudes. The coefficient of correlation reflects the degree of association and its estimate shows how the selection for a trait influences the expression of other traits. However, despite their great usefulness, the coefficients of correlation do not express the exact relative importance of the direct and indirect effects of these factors (Cruz and Regazzi 1994). The partitioning of the coefficient of correlation by the path analysis, developed initially by Sewall Wright according to Li (1956), consisted in the study of the direct and indirect effects of traits on a basic variable, whose estimates are obtained by regression equations in which the variables are previously standardized. Knowledge on the direct and indirect correlations, especially of the yield, allows breeders to use this additional information to discard or promote genotypes of interest (Pandini et al. 2002).

This study had the aim of estimating genetic parameters of agronomical traits of peanut and to split the coefficients of genotypic correlation between the grain yield and the primary components of the yield in direct and indirect effects.

MATERIAL AND METHODS

Three cultivars and six lines were used: Tatu, BR 1, L.7 Vermelha, CNPA 75 AM, CNPA 76 AM, CNPA 68 AM, L.8.14.12, L.8.14.01 and L.7 Bege, respectively. These genotypes, provided by Embrapa Algodão, presented upright habit type, sequential branching and a short cycle. The lines CNPA 75 AM, CNPA 76 AM and L.7 Bege belong to the Spanish group and have cream-colored seeds, while the other genotypes belong to the group Valence and have red seeds.

The evaluation was realized in Teresina, state of Piaui, lat 5º 5' S, long 42º 29' W and altitude 72 m asl, from February to July 2000, on a plain area of the Neossolo type soil with a sandy franca texture, in an experimental design of randomized complete blocks with five replications. Each plot consisted of a 5 m long row, where the plants were spaced 0.10 m and rows 0.70 m apart.

The evaluated traits were: plant height at flowering, number of pods per plot, percentage of empty pods, number of seeds per pod, weight of 100 seeds, pod yield, grain yield and ratio seed weight to fruit weight.

The statistical analyses were realized using software Genes (Cruz 2001). In all analyses the genotype effects

were considered fixed, implying that the possible inferences to be realized should be limited to the genotypes under study. Analyses of variance were realized and coefficients of variation estimated which provided information regarding the experimental precision. The coefficient of genotypic determination (h^2), the relation b and the phenotypic (r_F), genotypic (r_G) and environmental correlations (r_A) were estimated based on the quadratic components of the analyses, using the expressions proposed by Vencovsky and Barriga (1992).

In the path analysis, the coefficients of genotypic correlation between the grain yield and the primary components of the yield (number of pods per plot, number of seeds per pod and weight of 100 seeds) were partitioned in direct and indirect effects via the others. In the diagram presented in Figure 1, the direct influence of the primary components on the yield is shown by unidirectional arrows and measured by the path coefficients (P), while the indirect influence is measured by the sum of the products of the correlation coefficients by the path coefficients (rP). The bidirectional arrows indicate correlation, which is measured by coefficient r.
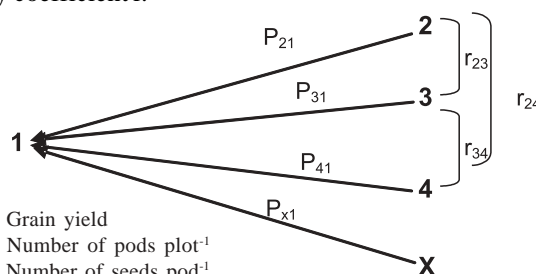
1. Grain yield
2. Number of pods plot^{-1}
3. Number of seeds pod^{-1}
4. Weight of 100 seeds
X. Residual variable

Figure 1. Illustrative causal diagram of the direct and indirect effects of the variables 2, 3 and 4 correlated with 1 and X, non-correlated residual variable

In the referred partition, the following equations were established:

$$r_{12} = P_{12} + r_{23}P_{31} + r_{24}P_{41}$$

$$r_{13} = r_{23}P_{21} + P_{31} + r_{34}P_{41}$$

$$r_{14} = r_{24}P_{21} + r_{34}P_{31} + P_{41}$$

A determination of the variation of the grain yield (GY) by the effects of the primary components, considered jointly, can be obtained by the equation of determination according to Li (1956):

$$GY\,(1) = P^2_{X1} + P^2_{21} + P^2_{31} + P^2_{41} + 2P_{21}r_{23}P_{31} + 2P_{21}r_{24}P_{41} + 2P_{31}r_{34}P_{41} \quad \text{where,}$$

$R^2_1\,(2, 3 \text{ and } 4) = P^2_{21} + P^2_{31} + P^2_{41} + 2P_{21}r_{23}P_{31} + 2P_{21}r_{24}P_{41} + 2P_{31}r_{34}P_{41} \rightarrow$ coefficient of multiple determination of the

grain yield by primary components 2, 3 and 4;

$P^2_{X1} = 1 - R^2_1$ (2, 3 and 4) $\rightarrow$ variation of the grain yield due to the residual variable X;

P^2_{X1} = residual effect.

RESULTS AND DISCUSSION

The genotypes presented significant effects for all traits, demonstrating the existence of variability in the evaluated genotypes, with exception of plant height at flowering (Table 1). The coefficients of variation (CV) evaluated in the experiment showed that the trait weight of 100 seeds (5.26%), relation seed/fruit (5.26%) and plant height at flowering (7.98%) presented the best experimental precision. For the number of seeds pod^{-1}, number of pods plot^{-1} and pod yield, the CVs attained mean values of 15.43, 17.78 and 20.43%, respectively. The high CV values for grain yield (28.12%) and percentage of empty pods (36.61%) can be considered acceptable and are in agreement with literature (Santos et al. 1993).

The highest estimates of the coefficient of genotypic determination were observed for weight of 100 seeds (98.63%), number of pods plot^{-1} (95.79%), grain yield (89.94%), number of seeds pod^{-1} (89.55%), and pod yield (84.03%), indicating a greater possibility of achieving superior genotypes in the selection for these traits. Santos and Moreira (1994) found high coefficients of genotypic determination for the initiation of flowering (97.00%), weight of 100 seeds (90.00%) and weight of 100 pods (87.00%). Ntare and Williams (1998) obtained low estimates of heritability, insignificant for pod yield in the initial generations. According to Knauft and Wynne (1995), low estimates of heritability in the narrow sense were found for

grain yield, yield components and traits associates with peanut quality. Regarding relation b there was a favorable situation for the selection of superior genotypes for the traits weight of 100 seeds (3.79), number of pods plot^{-1} (2.13), grain yield (1.34), number of seeds pod^{-1} (1.31), and pod yield (1.03), whose values were above 1.0.

Line CNPA 76 AM presented the highest number of pods per plot and was among the genotypes with lowest percentage of empty pods (Table 2). Regarding the number of seeds per pod, the genotypes Tatu, BR 1, CNPA 75 AM, CNPA 68 AM and L.8.14.01 presented the best results, in other words, three seeds per pod. For the weight of 100 seeds, L.7 Vermelha and L.7 Bege stood out with 59.82 g and 58.46 g, respectively. Considering the yields of pods and of grains, best results were obtained by the genotypes: CNPA 76 AM and L.7 Bege, with a cream skin; and L.7 Vermelha and L.8.14.01, a red skin. These genotypes attained mean yields with hull of over 1.880 kg ha^{-1}, surpassing the mean of yield of the commercial cultivars Tatu (1.294.34 kg ha^{-1}) and Br-1 (1.637.60 kg ha^{-1}). For grain yield, the means varied from 964.11 kg ha^{-1} to 1.437 kg ha^{-1}. The color of the outstanding genotypes showed no difference. Já Rego and Santos (1994) verified that the red-skinned genotypes were more productive than the beige ones. The data percentage referring to the relation seed/fruit, that is, the yield of perfect seeds in relation to the weight of dry pods varied from 54.59 to 67.53%.

In general, the genotypic correlations had the same signal and were superior to the phenotypic correlations (Table 3). This indicates that the phenotypic expression of the association between the traits is diminished by environmental interferences, which hampers the selective process. Estimates of the genotypic correlations of plant height at flowering with weight of 100 seeds, pod yield and grain yield were of moderate to high magnitude

Table 1. Mean squares, estimates of the coefficient of genotypic determination (h^2) and their respective standard deviations and of the relation b, referring to the traits evaluated in nine peanut (*Arachis hypogaea* L.) genotypes

Sources of variation	df	Plant height at flowering cm	Nr of pods plot^{-1}	Empty pods %	Nr of seeds pod^{-1}	Weight of 100 seeds g	Pod yield kg ha^{-1}	Grain yield kg ha^{-1}	Relation seed fruit^{-1} %
						Mean squares			
Blocks	4	21.17	64986.37	21.07	0.06	36.19	524886.31	173945.73	22.48
Genotypes	8	2.33	324991.26**	124.70**	1.49**	415.77**	889692.84**	607833.16**	88.06**
Error	32	1.21	13694.49	36.17	0.16	5.69	142069.27	61156.58	5.69
General mean		13.79 ± 0.49	658.29 ± 52.33	16.41 ± 2.69	2.56 ± 0.18	45.38 ± 1.07	1800.41 ± 168.56	879.52 ± 110.60	45.38 ± 1.07
CV %		7.98	17.78	36.61	15.43	5.26	20.93	28.12	5.26
h^2		48.16 ± 0.07	95.79 ± 0.00	71.05 ± 0.02	89.55 ± 0.00	98.63 ± 0.00	84.03 ± 0.01	89.94 ± 0.00	76.64 ± 0.01
b		0.43	2.13	0.70	1.31	3.79	1.03	1.34	0.81

** - Significant at 1% probability by the F tes

Table 2. Estimates of the means[1] referring to the traits evaluated in nine peanut (*Arachis hypogaea* L.) genotypes

Genotypes	Plant height at flowering cm	Nr of pods plot[-1]	Empty pods %	Nr of seeds pod[-1]	Weight of 100 seeds g	Pod yield kg ha[-]	Grain yield kg ha[-1]	Relation seed fruit[-1] %
Tatu	13.86	458.00 c	22.38 the	3 the	37.78 de	1294.34 c	503.68 d	56.02 b
BR 1	13.20	538.60 bc	16.43 ab	3 the	39.02 de	1637.60 bc	785.14 bcd	60.03 ab
L.7 - Vermelha	14.46	739.60 b	11.40 ab	2 b	59.82 the	2257.20 ab	1279.54 ab	64.08 ab
CNPA -75AM	12.56	458.60 c	20.48 ab	3 a	41.94 cd	1432.86 c	531.14 d	56.86 b
CNPA -76AM	13.66	1273.40 a	9.73 b	2 b	34.38 and	2444.06 the	1437.01 the	67.53 a
CNPA -68AM	14.40	519.40 c	23.52 the	3 a	40.68 cd	1475.60 bc	547.26 d	54.59 b
L.8.14.12	13.50	547.20 bc	15.76 ab	2 b	51.08 b	1533.03 bc	702.91 cd	59.65 ab
L.8.14.01	13.72	638.40 bc	15.80 ab	3 a	45.28 c	1882.34 abc	964.11 abcd	61.80 ab
L.7 - Bege	14.76	751.40 bc	11.73 ab	2 b	58.46 the	2246.68 ab[1]	1164.91 abc	63.31 ab

[1] Means followed by the same letters in the same column did not differ in the Tukey test (P < 0.05)

Table 3. Estimates of coefficients of the phenotypic (r_F), genotypic (r_G) and environmental correlation (r_A) between traits evaluated in nine peanut (*Arachis hypogaea* L.) genotypes

Trait	Nr of pods plot[-1]	Empty pods %	Nr of seeds pod[-1]	Weight of 100 seeds g	Pod yield kg ha[-]	Grain yield kg ha[-1]	Relation seed fruit[-1] %
Plant height at flowering (cm)	-0.281*	-0.291*	0.536*	0.452*	0.406*	0.226*	0.226*
	0.332*	-0.337*	-0.416*	0.800*	0.673*	0.581*	0.237*
	0.009	-0.217*	-0.076*	-0.178*	0.083*	0.101*	0.237*
Nr of pods plot[-1]		-0.782*	-0.619*	-0.074*	0.863*	0.875*	0.869*
		-0.906*	-0.646*	-0.082*	0.876*	0.885*	0.998*
		-0.317*	-0.299*	0.223*	0.928*	0.822*	0.101*
Empty pods (%)			0.769*	-0.435*	-0.924*	-0.948*	0.769*
			0.959*	-0.510*	-0.998*	-0.997*	-0.998*
			0.019	-0.121*	-0.399*	-0.538*	-0.844*
Nr of seeds pod[-1]				-0.581*	-0.732*	-0.723*	-0.734*
				-0.609*	-0.819*	-0.779*	-0.890*
				-0.238*	-0.170*	-0.228*	0.022
Weight of 100 seeds (g)					0.374*	0.337*	0.272*
					0.404*	0.353*	0.299*
					0.120*	0.131*	0.224*
Pod yield (kg ha[-1])						0.989*	0.942*
						0.998*	0.997*
						0.997*	0.999*
Grain yield (kg ha[-1])							0.971*
							0.999*
							0.987*

(r_G from 0.581 to 0.800), indicating that within the genotypes, the most productive were the highest. Among the traits number of pods plot[-1], pod yield, grain yield, and relation weight of seed/weight of fruit, the phenotypic and genotypic correlations were positive and high (r_G and r_F > 0.800). These associations can be ascribed to the fact that the traits in question are affected by the same fundamental physiological processes in the plant. Similar results between pod and grain yield were obtained by Moreira et al. (1994). The genotypic correlations of percentage of empty pods and number of seeds pod[-1] with weight of 100 seeds, pod yield, grain yield and relation weight of seed/ weight of fruit were negative and of high

magnitude (r_G > 0.700). Moreira et al. (1994) estimated high but positive genotypic correlations between the percentage of empty pods with pod yield and grain yield. The environmental correlations were generally of low magnitude (r_E < 0.400). However, between pod yield, grain yield and relation weight of seed/weight of fruit, and number of pods plot[-1] with pod yield and grain yield, these correlations were positive and high (r_E > 0.800). These results suggest that the action of the environment influenced the expression of these traits in the same sense, with great intensity. A high coefficient of environmental correlation between pod yield and grain yield was also estimated by Moreira et al. (1994).

The partitioning of the genotypic correlations between grain yield and the primary components: number of pods plot[-1], number of seeds pod[-1] and weight of 100 seeds into direct and indirect effects (Table 4) shows that the grain yield was positively influenced by the number of pods plot[-1] and weight of 100 seeds and negatively by the number of seeds pod[-1].

The positive genotypic correlation of elevated magnitude between the number of pods plot[-1] and grain yield ($r_G = 0.885$) was caused by the direct effect (1.302) of the number of pods plot[-1] in the grain yield that compensated the indirect negative effects by the number of seeds/pod (- 0.352) and of the weight of 100 seeds (- 0.065) (Table 4).

The indirect and negative effects of the number of seeds pod[-1] through the number of pods plot[-1] (- 0.842) and the weight of 100 seeds (- 0.482) helped that the estimate of the coefficient of genotypic correlation between number of seeds pod[-1] and grain yield was negative and of high magnitude (- 0.779), although the direct effect of the number of seeds/pod in the grain yield had been positive (0.544).

The estimate of the coefficient of positive genotypic correlation and of low magnitude between weight of 100 seeds and grain yield ($r_G=0.353$) is result of the direct positive effect (0.791) counterbalanced by the indirect negative effects through the number of pods plot[-1] (- 0.107) and the number of seeds pod[-1] (- 0.331). Moreira et al. (1994) obtained an estimate of positive genotypic correlation of high magnitude ($r_G=0.877$). Nevertheless, the direct effect of the weight of 100 seeds on the grain yield was negative (- 2.928).

For improvement programs it is important to identify traits of greater direct effect in a favorable sense for selection among those of high correlation with the basic variable, so that the response correlated through the indirect selection is efficient (Cruz and Regazi, 1994). In our study, the number of pods plot[-1] was the trait of strongest direct influence (1.179) in a favorable sense on the grain yield; this trait should therefore be considered when aiming at the selection of productive genotypes. In soybean, the number of pods plant[-1] also presented a potential for indirect selection of genotypes with higher yield (Pandini et al. 2002). According to Wynne and Coffelt (1982); cited by Godoy et al. (1999), the highest reproductive efficiency can be expressed by quantitative traits of moderate heritability, such as the number of pods plant[-1] or the percentage of fully developed pods. The weight of 100 seeds also presented a direct effect in the favorable sense on the grain yield and can be used in the indirect selection since it is a trait of higher heritability, as suggested in the study of Coffelt et al.(1989).

The degree of genotypic determination of the grain yield by the primary components of the yield attained a value of $R^2=1.00$. A hundred percent of the variation of the grain yield were thus genotypically explained by the trait in consideration.

Table 4. Path analysis: partitioning of the genotypic correlations into grain yield (1) and the trait number of pods plot[-1] (2), number of seeds pod[-1] (3) and weight of 100 seeds (4) in direct and indirect effects

Mode of action	Genotypic correlation
Number of pods plot[-1] x grain yield (kg ha[-1])	
Direct effect	1.302
Indirect effect through the number of seeds pod[-1]	-0.352
Indirect effect through the Weight of 100 seeds	-0.065
Total effect	0.885
Number of seeds pod[-1] x grain yield (kg ha[-1])	
Direct effect	0.544
Indirect effect through the number of pods plot[-1]	-0.842
Indirect effect through the weight of 100 seeds	-0.482
Total effect	-0.779
Weight of 100 seeds x grain yield (kg ha[-1])	
Direct effect	0.791
Indirect effect through the number of pods plot[-1]	-0.107
Indirect effect through the number of seeds pod[-1]	-0.331
Total effect	0.353
R^2_1 (2, 3 and 4)	1.00
R^2_x	0.00

R^2_1 (2, 3 and 4): coefficient of genotypic determination
R^2_x: effect of the residual variable

Correlações e análise de trilha em amendoim

RESUMO - *Parâmetros genéticos de caracteres agronômicos de amendoim foram estimados e os coeficientes de correlação genética entre a produção e os componentes primários da produção desdobrados em efeitos diretos e indiretos. Os genótipos Tatu, BR 1, L.7 Vermelha, CNPA 75 AM, CNPA 76 AM, CNPA 68 AM, L.8.14.12, L.8.14.01 e L.7 Bege foram avaliados em Teresina, PI. As mais altas estimativas do coeficiente de determinação fenotípica e da relação b, foram obtidas para peso de 100 sementes, número de vagem parcela^{-1}, produção, número de sementes vagem $^{-1}$ e produção de vagem. O desdobramento das correlações genotípicas dentro da produção e os componentes primários, em efeitos diretos e indiretos, mostra que a produção foi positivamente influênciada pelo número de vagem parcela^{-1} e peso de 100 sementes, e negativamente pelo número de sementes vagem^{-1}. No sentido favorável, o número de vagem parcela^{-1} foi ao caráter de maior influência direta sobre a produção.*

Palavras-chave: *Arachis hypogaea*, coeficiente de determinação genotípica, correlações, análise de trilha.

REFERENCES

Coffelt TA, Seaton ML and Van Scoyoc SN (1989) Reproductive efficiency of 14 virginia - type peanut cultivars. **Crop Science 29**: 1217-1220.

Cruz CD (2001) **Programa Genes: aplicativo computacional em genética e estatística**. Editora UFV, Viçosa, 648p.

Cruz CD and Regazzi AJ (1994) **Modelos Biométricos aplicados as melhoramento genético**. Editora UFV, Viçosa, 390p.

Godoy IJ, Moraes AS, Zanotto MD, and Santos FC (1999) Melhoramento do amendoim. In: Borém A (ed.). **Melhoramento de espécies cultivadas**. Editora UFV, Viçosa, p. 51-94.

Knauft DA and Wynne JC (1995) Peanut breeding and genetics. **Advances in Agronomy 55**: 393-445.

Li CC (1956) The concept of path coefficient and its impact on population genetics. **Biometrics 12**: 190-210.

Moreira JAN, Santos RC, Carvalho LP and Santos JW (1994) Correlações genotípicas, fenotípicas e ambientais em linhagens F$_5$ de amendoim. In: **Relatório técnico anual do Centro de Pesquisa de Algodão - 1992-93**. Embrapa CNPA, Campina Grande, p. 376-378.

Ntare BR and Williams JH (1998) Heritability of components of a simple physiological model for yield in groundnut under semiarid rainfed conditions. **Field Crops Research 58**: 25-33.

Pandini F, Vello NA and Lopes ACA (2002) Heterosis in soybeans for seed yield components and associated traits. **Brazilian Archives of Biology and Technology 45**: 401-412.

Rego GM and Santos RC (1994) Competição entre genótipos avançados de amendoim de película vermelha e bege sob condições irrigadas em Itabaiana - SE. In: **Relatório técnico anual do Centro de Pesquisa de Algodão - 1992-93**. Embrapa CNPA, Campina Grande, p. 360-362.

Santos RC (1995) Peanut crop: a viable alternative to Brazilian Northeast growers. **Ciência e Cultura 47**: 9-10.

Santos RC and Moreira JAN (1994) Estimação dos parâmetros genéticos em populações F$_5$ de amendoim. In: **Relatório técnico anual do Centro de Pesquisa de Algodão - 1992-93**. Embrapa CNPA, Campina Grande, p. 405-407.

Santos RC, Guimarães MB, Moraes JS and Brito SFM (1993) **Fenologia, reprodução e crescimento de genótipos de amendoim no nordeste brasileiro**. Embrapa CNPA, Campina Grande, 8p. (Embrapa CNPA. Pesquisa em Andamento, 16).

Upadhyaya HD, Bramel PJ, Ortiz R and Singh S (2002) Developing a mini core of peanut for utilization of genetic resources. **Crop Science 42**: 2150-2156.

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