Crop Breeding and Applied Biotechnology 4:305-312, 2004
Brazilian Society of Plant Breeding. Printed in Brazil

Estimates of correlations and path analyses for eight high protein and grain yield soybean populations

Odilon Lemos de Mello Filho[1*], Carlos Sigueyuki Sediyama[2], Maurilio Alves Moreira[3], Múcio Silva Reis[2], Rafael Rosalino Dias[2], Guilherme Augusto Massoni de Andrade[2], and Newton Denis Piovesan[4]

Received 21 November 2003

Accepted 21 July 2004

ABSTRACT *- Estimates of phenotypic correlations and the direct and indirect effects of grain yield and protein content components on grain and protein yields per plant were studied as source of information on eight soybean populations. Four* BC_1F_4 *and four* F_4 *populations from crosses of high protein content materials and commercial varieties were studied. Protein yield and grain yield per plant followed an analogous tendency. The CD201-BC population presented positive correlations of protein content with protein and grain yields. All yield components presented positive direct effects for both grain and protein yields. The number of pods per plant, amongst the yield components, presented the largest direct effects on grain and protein yields, except in the OC13-CR population. The direct effect of the protein content on protein yield was positive, yet small. Grain yield per plant was the main determining factor for protein yield per plant.*

Keywords: *Glycine max*, yield components, backcross.

INTRODUCTION

Plant breeding demands continuous phenotypic evaluation of the genetic material that is being generated by the breeding program in measurements of several variables that characterize the individuals. According to Wright (1921), Bravais, in 1846, established and defined the mathematical relationship between two variables, which Galton, in 1888, designated ‘correlation coefficient’.

Knowledge about correlation coefficients is a basic requirement to quantify the magnitude and the direction of the influences of certain traits on others, or when several traits are to be improved simultaneously (Cruz and Regazzi 1997). Despite the great importance of the correlations, the quantification and interpretation of the coefficients may, however, induce a wrong move in the selection strategy, since a high correlation between two traits may be result of the indirect effect of a third trait or a group of other traits (Carvalho et al. 1999).

According to Wright (1921), the ideal scientific method is the investigation of the direct effect of one condition on another in experiments free from any other possible variation causes. Unfortunately, the causes of variation often seem to

[1]Departamento de Fitotecnia, Universidade Federal de Viçosa (UFV), 36570-000, Viçosa, MG, Brasil. *E-mail: odilonlemos@bol.com.br
[2]Departamento de Fitotecnia, UFV
[3]Departamento de Bioquímica e Biologia Molecular, UFV
[4]Bioagro, UFV

be beyond control. In biological sciences, especially, one often has to deal with a group of traits or conditions correlated in a complex of interacting, uncontrollable, and often obscure causes. The correlation degree between two variables can be calculated by the common methods, but it merely presents the result of all connecting paths of influence.

Hence, it is important that the correlation coefficients be split up in components of direct and indirect effects, achieved by the path analysis, developed by Wright (1921). Path analysis consists in the study of the direct and indirect effects of traits on a main variable, usually the yield, which obtains estimates through regression equations on previously standardized variables (Cruz and Regazzi 1997).

The present work had the objective to present information on eight soybean populations through the estimates of phenotypic correlations among different agronomic traits and protein content in the soybean seed and their division into direct and indirect effects of those traits and of the protein content on grain and protein yield per plant.

MATERIAL AND METHODS

The experiment, the development of the eight populations, the agronomic management practices, as well as details of the total protein content analyses can be found in a previous study (Mello Filho et al. 2004).

The phenotypic correlation between the traits and two path coefficient analyses were estimated by the data logarithms. In the first path analysis, the phenotypic correlation coefficients were split into direct and indirect effects of the primary components of the grain yield – weight of one hundred seeds, number of seeds per pod, and number of pods per plant (Figure 1). In the second path analysis, the phenotypic correlation coefficients were partitioned into direct and indirect effects of the protein percentage in the seed and of the primary components of the yield on the protein yield per plant (Figure 2).

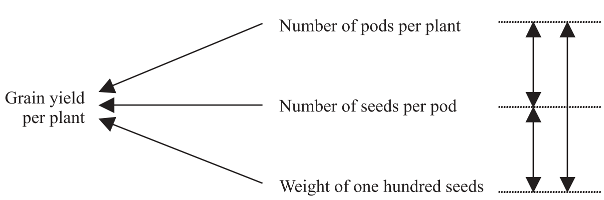

Figure 1. Causal diagram indicating the inter-relationship of the explanatory traits - number of pods per plant, number of seeds per pod and weight of one hundred seeds - on the main trait, grain yield per plant. The single direction arrows indicate the direct effect of the explanatory trait on the main trait, while the bi-directional arrows represent the interdependence between two explanatory variables, quantified by the phenotypic correlation

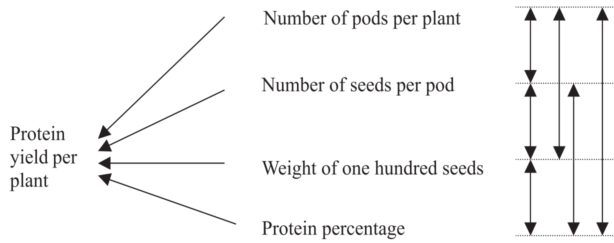

Figure 2. Causal diagram indicating the inter-relationship of the explanatory traits - number of pods per plant, number of seeds per pod, weight of one hundred seeds and protein percentage in the seeds - on the main trait, protein yield per plant. The arrows have the same meaning of the Figure 1

The methodology used for the path coefficient analysis was developed by Wright (1921) and refined by Li (1975) and Cruz and Regazzi (1997). When the relationship between the explanatory variables and the basic variable is structurally multiplicative, Cruz and Regazzi (1997) suggest the logarithmic transformation of the data, because the cause-and-effect analysis bases on the premise of additivity or linearity of the cause-and-effect regression. It was not observed any harmful collinearity. Thus, all presented data refer to the logarithm of the original variables. The correlations and the path analysis were estimated using the software Genes (Cruz 2001).

RESULTS AND DISCUSSION

The estimates of the phenotypic correlation coefficients, in each population, are presented in Table 1. Results of correlations obtained in the present work and the bibliographical citations both refer to the phenotypic correlation coefficients, unless otherwise stated.

The correlation estimates between the grain yield components (number of pods per plant, number of seeds per pod, and weight of one hundred seeds) were variable among the populations, presenting a tendency of the number of pods per plant and of the number of seeds per pod to correlate negatively with the weight of one hundred seeds. Anand and Torrie (1963), Ali et al. (1989), and Santos (1994) found similar results. In agreement with the results obtained by Johnson et al. (1955) and Anand and Torrie (1963), the correlations between the number of pods per plant and weight of one hundred seeds, for CD201-BC and CD206-BC populations, were negative and significant, indicating a strong competition for metabolites between these two traits.

The correlations of each one of the three yield components, with grain yield were, when significant, positive. The OC13-CR population presented positive and significant correlation of all the three mentioned traits to grain yield per plant. Santos et al. (1986), working on beans, found similar

Table 1. Phenotypic correlation coefficients between the logarithms of the traits number of pods per plant, number of seeds per pod, weight of one hundred seeds, grain yield per plant, protein yield per plant, and protein content in the seeds in eight soybean populations

Populations[1]	Traits	Number of pods plant^{-1}	Number of seeds pod^{-1}	Weight of seed	Grain yield plant^{-1}	Protein yield plant^{-1}	% of protein in the seed
OC13-BC	Number of pods plant^{-1}	1	-0.104	0.343	0.796**	0.802**	-0.144
OC13-CR	Number of seeds pod^{-1}	-0.296	1	0.203	0.367	0.338	-0.279
	Weight of seeds	0.334	-0.066	1	0.748**	0.739**	-0.215
	Grain yield plant^{-1}	0.538**	0.453*	0.723**	1	0.990**	-0.312
	Protein yield plant^{-1}	0.499*	0.430	0.748**	0.981**	1	-0.173
	% of protein	-0.289	-0.213	-0.037	-0.310	-0.112	1
CD201-BC	Number of pods plant^{-1}	1	-0.603*	-0.234	0.864**	0.789**	-0.002
CD201-CR	Number of seeds pod^{-1}	0.297	1	0.302	-0.304	-0.289	-0.064
	Weight of seeds	-0.621**	-0.165	1	0.261	0.393	0.731**
	Grain yield plant^{-1}	0.937**	0.515*	-0.377	1	0.983**	0.324
	Protein yield plant^{-1}	0.928**	0.474*	-0.370	0.984**	1	0.497
	% of protein	-0.384	-0.371	0.196	-0.450**	-0.261	1
CD205-BC	Number of pods plant^{-1}	1	0.453*	-0.156	0.903**	0.813**	-0.661**
CD205-CR	Number of seeds pod^{-1}	0.035	1	-0.111	0.635**	0.623**	-0.277
	Weight of seeds	-0.285	0.025	1	0.192	0.331	0.439*
	Grain yield plant^{-1}	0.791**	0.347	0.290	1	0.973**	-0.471*
	Protein yield plant^{-1}	0.740**	0.336	0.352	0.988**	1	-0.246
	% of protein	-0.327	-0.058	0.387	-0.111	0.054	1
CD206-BC	Number of pods plant^{-1}	1	-0.137	-0.213	0.948**	0.929**	-0.200
CD206-CR	Number of seeds pod^{-1}	0.284	1	-0.391	-0.026	-0.061	-0.128
	Weight of seeds	-0.681**	-0.287	1	0.026	0.042	0.083
	Grain yield plant^{-1}	0.941**	0.432	-0.447	1	0.982**	-0.214
	Protein yield plant^{-1}	0.893**	0.433	-0.365	0.976**	1	-0.013
	% of protein	-0.409	-0.096	0.449	-0.319	-0.095	1

[1]BC above the diagonals, and CR below the diagonals. ** and *: Significant at 1 and 5% of probability, respectively, based on the t test

results and also found positive genotypic and environmental correlations between grain yield and each of its components, considering high correlations between grain yield and number of pods per plant.

In studies with soybean, positive - not always significant - correlations of grain yield with number of pods per plant and number of seeds per pod prevailed (Johnson et al. 1955, Anand and Torrie 1963, Simpson and Wilcox 1983). Positive and significant correlations between grain yield and weight of one hundred seeds were obtained by Johnson et al. (1955), Anand and Torrie (1963), Know and Torrie (1964), Byth et al. (1969), and Simpson and Wilcox (1983).

Correlations of grain yield with number of pods per plant and protein yield per plant, as well as between the last two, were positive, high, and significant for all the eight studied populations. This result indicates that the indirect selection for grain yield or protein yield per plant, for number of pods per plant may be a plausible option.

Protein yield per plant followed the same tendency of grain yield per plant, due to the shorter variation of the protein content in relation to the variation of the grain yield, and to the greater magnitude of the grain yield. Similar results were found by Shannon et al. (1972), who conducted F_2 families of six populations originated from the crossing of parents bearing excellent agronomic traits and high grain yield or high protein content in the seeds. These authors evaluated, in the F_4 generation, which types of crosses (high protein (P) x high protein, high protein x high grain yield (G), or high grain yield x high grain yield) were most promising for the increase of grain yield or protein content or both traits. They verified that the correlation coefficients for protein per hectare were similar to the obtained for grain yield, except for one of the six populations. This similarity indicated, according to the authors, that grain yield was the main trait contributing to protein yield per hectare. They concluded that the parent bearers of high protein content used in the experiment were an important factor that contributed to the yield of lines that bear high grain yield and high protein content, since these parents did not present much lower grain yield than the common commercial varieties and, due to the previous selection, they had other desirable agronomic traits.

The protein percentage presented negative correlation – but of low magnitude – to most of the agronomic traits. The only positive correlation was found for weight of one hundred seeds, significant in the CD201-BC and CD205-BC populations. Johnson et al. (1955) found positive and negative non-significant correlations between protein percentage and weight of one hundred seeds. Simpson and Wilcox (1983) found positive and significant correlation between these traits.

Despite the predominance of the negative and non significant correlation mentioned in literature, the most divergent results have been reported for the correlation between seed protein content and grain yield: Johnson et al. (1955), Byth et al. (1969), and Thorne and Fehr (1970) found negative and significant correlations; Shannon et al. (1972) and Scott and Kephart (1997) found positive and significant correlations. This divergence indicates that when a breeding program aims at a simultaneous selection of materials bearers of high protein seed contents and high grain yield, it is important to use populations in which the selection regarding those two traits are favored by a positive correlation between them.

The correlation between the protein percentage and the protein yield per plant were negative in most of the cases, but of low magnitude. This negative correlation indicates that, in the analyzed populations, a careful selection of plants aiming at high protein content in the seeds is required, since despite the increase in the protein content, the total protein yield per plant can be reduced. However, the CD201-BC population presented positive and relatively high correlations, although not significant, of protein content with protein yield per plant and grain yield, indicating good chances for the simultaneous improvement of these three traits.

It seems that many of the traits are correlated due to a positive or negative mutual association with each other. As more variables are considered in a table of correlations, these indirect associations become more complex and less obvious. In this context, the path coefficient analysis is a way to elucidate the direct and indirect causes of associations. It allows a critical analysis of specific forces producing a given correlation, besides permitting the measurement of the relative importance of each causal factor (Dewey and Lu 1959).

This way, the effects of the yield components on the grain yield and the direct and indirect effects of the grain yield components and of the protein content in the seed on the protein yield per plant were determined.

The path analysis and the partition of the phenotypic correlations into direct and indirect effects of the grain yield explanatory traits on grain yield per plant, based on the causal diagram presented in Figure 1 are shown in Table 2.

The indirect effects of the number of pods per plant on the grain yield - either through number of seeds per pod or weight of one hundred seeds - were low. The direct effects were positive and high, as well as the correlation between them, indicating that this trait can be used as an auxiliary trait in programs of soybean breeding aiming at higher-yielding varieties. Similar results were reported by Duarte and Adams (1972) and Furtado et al. (2002), who worked with beans, and by Ali et al. (1989) and Taware et al. (1997), studying soybeans.

The number of seeds per pod presented, by the weight of one hundred seeds, a very low indirect effect on the grain yield, indicating the low influence the weight of one hundred seeds exerts on the correlation between number of seeds per pod and grain yield. By the number of pods per plant, the indirect effect of the number of seeds per pod on the grain yield was variable among the populations. This indirect effect was mainly responsible for the negative correlation between number of seeds per pod and grain yield for the CD201-BC population. For the CD201-CR, CD205-BC, and CD206-CR populations, this indirect effect was positive and larger when compared to the direct effect on the grain yield. The direct effects were positive, however of low magnitude, except for the OC13-CR population, for which the direct effect was positive and of high magnitude. This indicates that in this population, the trait may be used as auxiliary to increase the efficiency of the indirect selection on the correlated response.

The indirect effects of weight of one hundred seeds on the grain yield by number of seeds per pod were very low, and were diverse by the number of pods per plant, though, predominantly, negative. For the CD201-CR and CD206-CR populations, the high and negative indirect effect of the weight of one hundred seeds by the number of pods per plant was the chief responsible for the negative correlation between weight of one hundred seeds and grain yield, indicating a compensation effect between weight of one hundred seeds

Table 2. Direct (underlined) and indirect effects of the primary components of the grain yield (number of pods per plant, number of seeds per pod, and weight of one hundred seeds) on the grain yield per plant, in eight soybean populations[1]

Population	Trait	Effect through			Total effect (r)
		Number of pods $plant^{-1}$	Number of seeds pod^{-1}	Weight of one hundred seeds	
OC13-BC	Number of pods $plant^{-1}$	<u>0.679</u>	-0.036	0.153	0.796**
	Number of seeds pod^{-1}	-0.071	<u>0.347</u>	0.090	0.367
	Weight of seeds	0.233	0.070	<u>0.445</u>	0.748**
OC13-CR	Number of pods $plant^{-1}$	<u>0.535</u>	-0.193	0.196	0.538**
	Number of seeds pod^{-1}	-0.159	<u>0.650</u>	-0.039	0.456**
	Weight of seeds	0.178	-0.043	<u>0.587</u>	0.723**
CD201-BC	Number of pods $plant^{-1}$	<u>1.105</u>	-0.136	-0.106	0.864**
	Number of seeds pod^{-1}	-0.666	<u>0.226</u>	0.136	-0.304
	Weight of seeds	-0.259	0.068	<u>0.452</u>	0.261
CD201-CR	Number of pods $plant^{-1}$	<u>1.064</u>	0.075	-0.202	0.937**
	Number of seeds pod^{-1}	0.316	<u>0.253</u>	-0.054	0.515*
	Weight of seeds	-0.661	-0.042	<u>0.326</u>	-0.377
CD205-BC	Number of pods $plant^{-1}$	<u>0.822</u>	0.137	-0.055	0.903**
	Number of seeds pod^{-1}	0.372	<u>0.302</u>	-0.039	0.635**
	Weight of seeds	-0.128	-0.034	<u>0.354</u>	0.192
CD205-CR	Number of pods $plant^{-1}$	<u>0.937</u>	0.010	-0.156	0.791**
	Number of seeds pod^{-1}	0.033	<u>0.301</u>	0.014	0.347
	Weight of seeds	-0.267	0.007	<u>0.549</u>	0.290
CD206-BC	Number of pods $plant^{-1}$	<u>1.058</u>	-0.035	-0.075	0.948**
	Number of seeds pod^{-1}	-0.145	<u>0.256</u>	-0.137	-0.026
	Weight of seeds	-0.226	-0.100	<u>0.352</u>	0.026
CD206-CR	Number of pods $plant^{-1}$	<u>1.151</u>	0.062	-0.272	0.941**
	Number of seeds pod^{-1}	0.327	<u>0.220</u>	-0.115	0.432
	Weight of seeds	-0.784	-0.063	<u>0.400</u>	-0.447

[1]R^2 = 1.00 for all populations. **and *: Significant at 1 and 5%, respectively, based on the t test

and number of pods per plant in those populations. That compensation effect was also verified by Gravois and Helms (1992), working on rice. The authors verified that as the planting density was increased, the density of the panicles increased significantly, while the number of filled grains per panicle decreased. Board et al. (1999) found a small indirect effect of the number of seeds per plant by the weight of one hundred seeds (size) on grain yield, indicating the absence of a compensation effect among these traits. The direct effects of the weight of one hundred seeds on the grain yield were positive for all eight populations, indicating that this trait can be used to assist the selection of high grain yield material. Ali et al. (1989) found positive direct effects of both number of pods and weight of one hundred seeds on the grain yield.

The direct effect of the number of pods per plant on the grain yield, obtained in the present study was positive and high. According to Board et al. (1997), the grain yield was highly correlated to number of seeds, size of seeds (weight of one hundred seeds), and number of pods. Based on the correlations, just one of the three referred traits would be enough to represent the components of the grain yield, since the three traits were almost equally important for the increase of the grain yield. However, the path analysis revealed that among the three traits, the number of seeds was the most important. The direct effect of the seed size on grain yield was negligible and the number of pods was negative.

The direct and indirect effects of the grain yield primary components (number of pods per plant, number of seeds per pod, and weight of one hundred seeds) and of the protein percentage in the seed on the protein yield per plant are presented in Table 3 for the eight assessed populations. The indirect effects of the number of pods on the protein yield were all low. The direct effects were high and, among the studied traits, the number of pods was the main determinant of the protein yield. High direct effects linked to high correlations with the basic variable, favorable to selection, indicate that the correlated answer through indirect selection might be efficient. It is worth make an observation regarding the protein yield per plant - that is the result of the multiplication of protein content in the seeds by grain yield - when the protein content barely varies and the grain yield (highly correlated to the number of pods) varies in larger magnitude, the expression of the trait grain yield (or number of pods) becomes prevailing in the determination of the protein yield per plant; and grain yield per plant becomes the main determining trait of the protein yield per plant.

The direct effects of the other two yield components (number of seeds per pod and weight of one hundred seeds) on the protein yield were positive for all populations presenting, however, a smaller magnitude than the direct effects of the number of pods. The indirect effects of the number of seeds per pod and weight of one hundred seeds on

Table 3. Direct (underlined) and indirect effects of the primary components of the grain yield (number of pods per plant, number of seeds per pod, and weight of one hundred seeds) and of the seed protein percentage on the protein yield per plant[1]

Population	Trait	Effect through				Total effect (r)
		Number of pods plant^{-1}	Number of seeds pod^{-1}	Weight of one hundred seeds	% of protein	
OC13-BC	Number of pods plant^{-1}	<u>0.701</u>	-0.038	0.156	-0.018	0.802**
	Number of seeds pod^{-1}	-0.073	<u>0.353</u>	0.092	-0.035	0.338
	Weight of seeds	0.240	0.072	<u>0.454</u>	-0.03	0.739**
	% of protein	-0.100	-0.098	-0.098	<u>0.124</u>	-0.173
OC13-CR	Number of pods plant^{-1}	<u>0.559</u>	-0.202	0.205	-0.063	0.499*
	Number of seeds pod^{-1}	-0.166	<u>0.682</u>	-0.041	-0.046	0.430
	Weight of seeds	0.186	-0.045	<u>0.614</u>	-0.008	0.748**
	% of protein	-0.161	-0.145	-0.023	<u>0.218</u>	-0.112
CD201-BC	Number of pods plant^{-1}	<u>1.012</u>	-0.127	-0.097	0.000	0.789**
	Number of seeds pod^{-1}	-0.610	<u>0.210</u>	0.125	-0.014	-0.289
	Weight of seeds	-0.237	0.063	<u>0.412</u>	0.154	0.393
	% of protein	-0.002	-0.013	0.301	<u>0.211</u>	0.497
CD201-CR	Number of pods plant^{-1}	<u>1.140</u>	0.080	-0.212	-0.081	0.928**
	Number of seeds pod^{-1}	0.338	<u>0.270</u>	-0.056	-0.078	0.474*
	Weight of seeds	-0.708	-0.044	<u>0.341</u>	0.041	-0.370
	% of protein	-0.438	-0.100	0.067	<u>0.210</u>	-0.261
CD205-BC	Number of pods plant^{-1}	<u>0.903</u>	0.151	-0.061	-0.180	0.813**
	Number of seeds pod^{-1}	0.409	<u>0.332</u>	-0.043	-0.075	0.623**
	Weight of seeds	-0.141	-0.037	<u>0.389</u>	0.120	0.331
	% of protein	-0.597	-0.092	0.171	<u>0.272</u>	-0.246
CD205-CR	Number of pods plant^{-1}	<u>0.940</u>	0.104	-0.156	-0.054	0.740**
	Number of seeds pod^{-1}	0.033	<u>0.300</u>	0.014	-0.010	0.336
	Weight of seeds	-0.268	0.007	<u>0.548</u>	0.064	0.352
	% of protein	-0.307	-0.017	0.212	<u>0.166</u>	0.054
CD206-BC	Number of pods plant^{-1}	<u>1.079</u>	-0.034	-0.075	-0.041	0.929**
	Number of seeds pod^{-1}	-0.147	<u>0.251</u>	-0.138	-0.026	-0.061
	Weight of seeds	-0.230	-0.098	<u>0.353</u>	0.017	0.042
	% of protein	-0.215	-0.032	0.029	<u>0.205</u>	-0.013
CD206-CR	Number of pods plant^{-1}	<u>1.207</u>	0.066	-0.285	-0.095	0.893**
	Number of seeds pod^{-1}	0.343	<u>0.232</u>	-0.120	-0.022	0.433
	Weight of seeds	-0.822	-0.066	<u>0.419</u>	0.105	-0.365
	% of protein	-0.494	-0.022	0.188	<u>0.233</u>	-0.095

[1]R^2 = 1.00 for all populations. ** and *: Significant at 1 and 5%, respectively, based on the t test

the protein yield followed the same tendencies of the effects presented on grain yield, so the statements for grain yield are also valid for protein yield. The indirect effects of the grain yield components through protein content on the protein yield per plant were low, indicating that the protein percentage, relatively, is not a factor that indirectly alters the correlation of the components of the grain yield to the protein yield per plant.

The indirect effects of the protein content in the seeds on the protein yield per plant, through the number of pods per plant were negative for all populations. This effect was mainly responsible for the low correlation between protein content and protein yield per plant in the populations CD201-CR, CD205-BC, CD205-CR, and CD206-CR. In these populations, this effect was of larger magnitude than the direct effect of the protein content in the seed on the protein yield per plant.

The direct effects of the protein content in the seeds on the protein yield per plant were positive, but presented a low magnitude and, except for the CD201-CR and CD205-CR populations, their signs were opposite to the correlations. This low magnitude confirms that in these populations grain yield was the most important among the traits that determine the protein yield per plant – protein content and grain yield.

CONCLUSIONS

The main conclusions were: *i*) Protein yield per plant followed the same tendency of grain yield per plant. *ii*) The CD201-BC population presented positive correlations of protein content with protein and grain yield. *iii*) All grain yield components presented positive direct effects for both grain and protein yields. *iv*) The number of pods per plant, among the grain yield components, presented the largest direct effects on grain and protein yields, except for the OC13-CR population; *v*) The direct effect of the protein content on protein yield was positive, though small; and *vi*) The grain yield per plant was the main factor contributing to protein yield per plant.

ACKNOWLEDGEMENTS

To the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the financial support.

Estimativas de correlações e análises de trilha para oito populações de soja de alto teor protéico e alta produtividade de grãos

RESUMO - *Estimaram-se as correlações fenotípicas e efeitos diretos e indiretos dos componentes da produção e do teor protéico sobre a produção de grãos e de proteína por planta em oito populações de soja. Para isso utilizaram-se quatro populações RC_1F_4 e quatro F_4 derivadas de cruzamentos entre materiais de alto teor protéico e variedades comerciais. A produção de proteína apresentou tendência semelhante à da produção de grãos por planta. A população CD201-RC apresentou correlações positivas de teor protéico com produção de proteína e de grãos. Os componentes da produção apresentaram efeitos diretos positivos sobre produção de grãos e proteína. Entre os componentes da produção o número de vagens apresentou maiores efeitos diretos sobre a produção de grãos e de proteína, exceto na população OC13-CR. O efeito direto do teor protéico sobre a produção de proteína foi positivo, todavia baixo. A produção de grãos por planta foi o principal determinante da produção de proteína.*

Palavras chave: *Glycine max*, componentes da produção, retrocruzamento.

REFERENCES

Ali AAA, Fraj BH and Ibraheen SA (1989) Correlation and path-coefficient analysis of yield and certain traits of soybean (*Glycine max* (L) Merrill) in Iraq. In: Pascale AJ (ed.) **IV Conferencia mundial de investigación en soja – actas**. SRL, Buenos Aires, p. 975-981.

Anand SC and Torrie JH (1963) Heritability of yield and other traits and interrelationships among traits in F_3 and F_4 generations of three soybean crosses. **Crop Science 3**: 508-511.

Board JE, Kang MS and Harville BG (1997) Path analyses identify indirect selection criteria for yield of late-planted soybean. **Agronomy Journal 37**: 879-884.

Board JE, Kang MS and Harville BG (1999) Path analyses of the yield formation process for late-planted soybean. **Agronomy Journal 91**: 128-135.

Byth DE, Weber CR and Caldwell BE (1969) Correlated truncation selection for yield in soybeans. **Crop Science 9**: 699-702.

Carvalho CGP, Oliveira VR, Cruz CD and Casali VWD (1999) Análise de trilha sob multicolinearidade em pimentão. **Pesquisa Agropecuária Brasileira 34**: 603-613.

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

Cruz C D and Regazzi AJ (1997) **Modelos biométricos aplicados ao melhoramento genético**. Editora UFV, Viçosa, 390p.

Dewey DR and Lu KH (1959) A correlation and path-coefficient analysis of components of crested wheatgrass seed production. **Agronomy Journal 51**: 515-518.

Duarte RA and Adams MW (1972) A path coefficient analysis of some yield component interrelation in field beans (*Phaseolus vulgaris* L). **Crop Science 12**: 579-582.

Furtado MR, Cruz CD, Cardoso AA, Coelho ADF and Peternelli LA (2002) Análise de trilha do rendimento do feijoeiro e seus componentes primários em monocultivo e em consórcio com a cultura do milho. **Ciência Rural 23**: 217-220.

Gravois KA and Helms RS (1992) Path analysis of rice yield and yield components as affected by seeding rate. **Agronomy Journal 84**: 1-4.

Johnson HW, Robinson HF and Comstock RE (1955) Genotypic and phenotypic correlations in soybeans and their implications in selection. **Agronomy Journal 47**: 477-483.

Kwon SH and Torrie JH (1964) Heritability of and interrelationships among traits of two soybean populations. **Crop Science 4**: 196-198.

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

Mello Filho OL, Sediyama CS, Moreira MA, Reis MS, Dias RR, Andrade GAM, and Piovesan ND (2004) Selection of high protein content and high yield soybean families. **Crop Breeding and Applied Biotechnology 4**: 171-177.

Santos PC, Cardoso AA, Vieira C and Silva JC (1986) Herdabilidade e correlações do rendimento com seus componentes em dois cruzamentos de feijão. **Revista Ceres 33**: 404-412.

Scott RA and Kephart KD (1997) Selection for yield protein and oil in soybean crosses between adapted and introduced parents. **Field Crops Research 49**: 177-185.

Shannon JG, Wilcox JR and Probst AH (1972) Estimated gains from selection for protein and yield in the F_4 generation of six soybean populations. **Crop Science 12**: 824-830.

Simpson AC and Wilcox JR (1983) Genetic and phenotypic associations of agronomic characteristics in four high protein soybean populations. **Crop Science 23**: 1077-1081.

Taware SP, Halvankar GB and Raut VM (1997) Variability correlation and path analysis in some germplasm lines of soybean (*Glycine max*) with high oil content. **Indian Journal of Agricultural Sciences 67**: 476-477.

Thorne JC and Fehr WR (1970) Incorporation of high-protein exotic germplasm into soybean populations by 2- and 3-way crosses. **Crop Science 10**: 652-655.

Wright S (1921) Correlation and causation. **Journal of Agricultural Research 20**: 557-585.