Correlations and path analysis in soybean progenies with resistance source to cyst nematode (race 3)

Daniela Sarti Dvorjak[1], Sandra Helena Unêda-Trevisoli, Wallace de Sousa Leite[3*],
Alysson Jalles da Silva[2], Fabiana Mota da Silva[2], Antonio Orlando Di Mauro[2]

[1]University of Kentucky, Lexington, USA
[2]Paulista State University, São Paulo, Brazil
[3]Federal Institute of Piauí
*Corresponding author, e-mail: leitewallace@hotmail.com

Abstract

Path analysis is an important study that slices the correlation coefficients between two variables to evaluate whether the relationship between them is of cause and effect. This study aimed to estimate the phenotypic and genotypic correlations between agronomic traits and perform a path analysis in order to identify variables for indirect selection aiming at a higher grain yield. Fourteen soybean F_6 lines from the soybean breeding program of FCAV–UNESP, Jaboticabal, São Paulo, Brazil, were analyzed. The experimental design was a randomized block design with three replications. The agronomic traits plant height at maturity (PHM), first pod height (FPH), lodging (Ld), agronomic value (AV), number of pods per plant (NP), number of seeds per plant (NS), and grain yield (GY) were evaluated. Overall, the genotypic correlations were higher than their corresponding phenotypic correlations. The genotypic correlations between grain yield and the traits agronomic value, number of pods per plant, and number of seeds per plant were positive, significant, and of high magnitude. Path analysis showed that the trait number of seeds per plant had the highest direct effect on grain yield, while the trait number of pods per plant had the highest indirect effect through the number of seeds per plant on grain yield

Keywords: Agronomic traits, *Glycine max*, grain yield, indirect selection

Introduction

Soybean (*Glycine max* (L.) Merrill) is considered the most important legume in the world. In Brazil, soybean breeding programs, from an agricultural point of view, aim to obtain more productive and stable cultivars, also better adapted to the various edaphoclimatic regions, in addition to being resistant to diseases of major economic importance. Several pathogens attack soybeans, being *Heterodera glycines* Ichinohe or soybean cyst nematode (SCN) one of the main factors limiting the production of this crop. The genetic variability of *H. glycines* in Brazil has been shown to be higher than that observed in other countries (Asmus et al., 2012). In this context,

several races of this pathogen were found, with emphasis on race 3, one of the most frequent in the several soybean producing regions of Brazil (Santana et al., 2009).

In soybean, grain yield, as in other crops, is a complex trait because it results from the expression and association of different components (El-Mohsen et al., 2013). Therefore, sufficient information about the traits that influence soybean yield is essential.

The knowledge of the correlation between traits is important when it is desired to make a simultaneous selection or when a trait of interest presents a low heritability because this study allows establishing criteria for an efficient

selection (Silva et al., 2015; Hamawaki et al., 2012; Nogueira et al., 2012; Almeida et al., 2010).

To verify the correlation between variables, the correlation coefficient or 'r' is used. The r coefficient evaluates only the degree (intensity) and nature (direction) of the association between traits (El-Mohsen et al., 2013).

Genetic correlations should be more useful than phenotypic correlations in the selection for grain yield because in this case, environmental effects are not considered (Machikowa & Laosuwan, 2011).

The correlation coefficients, although of great use in quantifying the magnitude and direction of factor influences in the determination of complex traits, do not provide the exact relative importance of direct and indirect effects of these factors, which can be measured by means of the path analysis (Nogueira et al., 2012; Alcântara Neto et al., 2011).

Path analysis was developed by Wright (1921) and consists of the study of the decomposition of the correlation coefficient into direct and indirect effects of traits on a main variable, allowing evaluating whether the relationship between two variables is of cause and effect or if it is determined by the influence of another or other variables (Haghi et al., 2012; Akram et al., 2011).

Thus, this study aimed to estimate the phenotypic and genotypic correlations between important agronomic soybean traits, in addition to performing the path analysis in order to identify variables for indirect selection aiming at an increasing grain yield.

Material and methods

The experiment was carried out at the experimental area of the Farm of Teaching, Research and Extension (FEPE) of the São Paulo State University, UNESP–FCAV, campus of Jaboticabal, located in the north of the São Paulo State, at the geographical coordinates 21°15′22″ S and 48°18′58″ W, with an altitude of about 595 m. The soil of the experimental area is classified as an Oxisol (Latossolo Vermelho Eutrófico, Brazilian Soil Classification System), with gently wavy relief. According to Köppen global climate classification the regional climate is classified as Cwa, with a warm and humid summer and dry winter, with an average annual temperature of 22.2 °C and precipitation of 1.451 mm (Vianna et al., 2013). Plowing and harrowing operations were performed before sowing. Planting fertilization was performed according to crop requirements after previous soil analysis.

The evaluated lines belong to the soybean breeding program of FCVA–UNESP and were obtained and selected in previous cycles for agronomic performance and resistance to the soybean cyst nematode.

The experimental design was a randomized block design with three replicates. Fourteen soybean lines resistant to soybean cyst nematodes were evaluated in the F_6 generation of endogamy, as well as four standard cultivars used as controls (Conquista, Vencedora, Coodetec 205, and IAC-Foscarin 31). These genotypes were chosen as controls because they have good agronomic attributes and are recommended for the region.

The genotypes were sown in plots consisting of 4 rows of 5 m, spaced 0.5 m from each other, with an average density of 20 plants per linear meter. The two central rows were considered as useful area, with 0.5 m being discarded from each end. Plants were evaluated at the R8 development stage (Fehr & Caviness, 1977), being collected six plants from the useful area of each plot.

The collected plants were evaluated for the following agronomic traits: plant height at maturity (PHM), expressed in cm; first pod height (FPH), expressed in cm; lodging (Ld), with a visual grading scale varying from 1 (all plants erect) to 5 (all plants lodged); agronomic value (AV), with a visual grading scale based on the observation of the average global aspect of the evaluated plants and taking into account a series of adaptive visual characters, which ranged from 1 (bad plant) to 5 (excellent plant); number of pods per plant (NP); number of seeds per plant (NS); and grain yield (GY), expressed in grams per plant.

Initially, an analysis of variance was performed, followed by estimates of phenotypic and genotypic correlations between the studied

traits (Falconer, 1987), using the t-test to verify if the correlations differed from zero or not.

In the path analysis, the trait grain yield was chosen as the main variable of the multivariate analysis. Therefore, the slice of correlations into direct and indirect effects of the other traits on grain yield was estimated.

To check for multicollinearity in a multivariate data set or correlation matrix, the Montgomery & Peck (1981) classification should be consulted, where NC = number of condition for multicollinearity, which is the division of the largest by the smallest eigenvalue of the data. Therefore, NC < 100 shows weak multicollinearity (it is not a serious problem), 100 < NC <1000 shows a moderate to strong multicollinearity, and NC > 1000 shows severe multicollinearity, leading to higher complications in the study.

Thus, the degree of multicollinearity of the X'X matrix was established based on its number of condition (NC) (Montgomery & Peck, 1981). Multicollinearity problems can make the matrix unique, leading the estimates to be unreliable. To attenuate this adverse effect, the system of normal equations is slightly modified by introducing a constant k (value from 0 to 1) on the diagonal of the X'X matrix (Cruz, 2012). Path analysis can be done from phenotypic, genotypic or environmental correlations. All genetic and statistical analyses were processed using the software Genes (Cruz, 2013).

Results and discussion

The results of the analysis of variance indicated a good experimental precision due to the reduced coefficients of variation (9.39 to 29.58%). In addition, significant differences were observed between genotypes for the studied traits, except for Ld and AV, with estimates of high heritabilities (between 0.6 and 0.88), i.e. with an accuracy of genotype selection (Ac_{gen}), which is the square root of heritability (Resende, 2007), of 77.46 and 93.80%, respectively.

Genotypic correlations were higher than their respective phenotypic correlations for most of the traits (Table 1), which indicates that strong intrinsic associations are reduced at the phenotypic level due to residual random effects. Thus, the major contributions in associations between traits are due to genetic causes. When genotypic correlations are higher than their corresponding phenotypic correlations, this indicates that the phenotypic expression is reduced by the influences of the environment (Chandel et al., 2014; Nogueira et al., 2012; Machikowa & Laosuwan, 2011; Almeida et al., 2010; Showkat & Tyagi, 2010). Thus, genotypic correlations are intrinsically more useful than phenotypic correlations in deciding selection strategies.

Table 1. Coefficients of phenotypic (r_F) and genotypic (r_G) correlations among seven agronomic traits evaluated in soybean F_6 progenies

		Trait					
	Correlation	FPH	Ld	AV	NP	NS	GY
PHM	r_F	0.590**	0.249	0.202	−0.348	0.084	0.122
	r_G	0.584*	0.801**	0.981**	−0.348	0.135	0.171
FPH	r_F		0.474*	−0.252	−0.609**	−0.419	−0.289
	r_G		0.964**	−0.180	−0.670**	−0.401	−0.272
Ld	r_F			−0.613**	−0.457	−0.454	−0.341
	r_G			−0.998**	−0.927**	−0.885**	−0.742**
AV	r_F				0.473*	0.737**	0.602**
	r_G				0.430	0.952**	0.957**
NP	r_F					0.763**	0.719**
	r_G					0.849**	0.753**
NS	r_F						0.920**
	r_G						0.939**

PHM: plant height at maturity; FPH: first pod height; Ld: lodging; AV: agronomic value; NP: number of pods per plant; NS: number of seeds per plant; and GY: grain yield. * and ** Significant at 5% and 1% probability, respectively, by the t-test.

All traits, except for first pod height and lodging, were positively correlated with grain yield in both phenotypic and genotypic coefficients.

Positive and significant (P<0.01) phenotypic and genotypic correlations with a high magnitude were observed between the traits grain yield and

agronomic value (0.602** and 0.957**), number of pods per plant (0.719** and 0.753**), and number of seeds per plant (0.920** and 0.939**) (Table 1), indicating that the selection of plants with a higher number of pods and seeds would result in more productive plants. Therefore, these traits can serve as a marker/indicator for the improvement of soybean grain yield.

The most favorable situation for the indirect selection of more productive genotypes was related to the number of pods per plant and number of seeds per plant, which are corroborated by the results observed by Silva et al. (2015), Chandel et al. (2014), Nogueira et al. (2012), and Bárbaro et al. (2007). Different results were obtained by Ghodrati et al. (2013) and Almeida et al. (2010), who verified that the number of seeds per plant did not correlate positively with grain yield.

El-Mohsen et al. (2013) and Akram et al. (2011) concluded that the production components number of pods per plant and number of seeds per plant showed a positive and significant correlation with final grain yield. Chandel et al. (2014) concluded that the trait grain yield was positively and significantly correlated with the number of pods per plant at phenotypic and genotypic levels. The results of our study are consistent with the findings of these authors.

Plant height at maturity showed a high positive and significant phenotypic and genotypic correlation with first pod height, indicating that the higher the plant height is, the higher the first pod height. Almeida et al. (2010) observed a positive and significant correlation of high magnitude between first pod height and grain yield and stated that the selection of plants with a higher first pod height would allow the indirect improvement to the trait grain yield.

Therefore, the negative correlation of first pod height with the number of pods per plant suggests that the higher the first pod height is, there is a trend of decrease the number of pods, besides a higher tendency to lodging. In a study of path analysis of soybean grain yield, Alcântara Neto et al. (2011) observed that the trait first pod height presented a low correlation with all the analyzed traits.

A negative correlation was observed between lodging and agronomic value and with the other traits that are part of the production components, as expected.

A significant phenotypic correlation was observed between agronomic value and the number of pods, number of seeds, and grain yield and, while a genotypic correlation was observed between agronomic value and the number of seeds per plant and grain yield, indicating that plants with a higher agronomic value tend to be more productive since this trait considers the overall phenotypic aspect of the plant. The trait agronomic value can be used in the process for indirect selection since it presented a high correlation in the phenotypic and genotypic coefficient with grain yield.

For path analysis from the phenotypic data, the number of condition was below 100, showing weak multicollinearity. However, the genotypic correlation matrix of traits presented values well above 1000, showing severe multicollinearity among traits. Thus, a path analysis with collinearity was performed and, to minimize the adverse effect, the constant k was introduced with an adequate value to estimate path coefficients. The values of k were considered within the range $0<k<1$ since the X'X matrix is in the format of correlations. Therefore, the value of k used was equal to 0.1006826, being the lowest value found to allow the stabilization of all coefficients.

However, to better understand the inter-relationships between traits, a path analysis was performed and its results are shown in Table 2. The analyzed variables have a great effect on the main variable grain yield since high determination coefficients (88% and 82% for phenotypic and genotypic correlations, respectively) and low effects of the residual variable (0.349 and 0.427 for phenotypic and genotypic correlations, respectively) were observed.

Table 2 shows that the number of seeds per plant had the highest influence on soybean yield since it had a higher direct effect on the phenotypic and genotypic coefficients (0.999 and 0.628, respectively). In turn, the number of pods per plant has a high indirect effect via number of seeds per plant on soybean yield, in

Table 2. Estimates of direct and indirect phenotypic and genotypic effects of the evaluated traits relative to GY (main variable) by means of path analysis in soybean F_6 progenies

Variable ==> PHM	Phenotypic estimate	Genotypic estimate
Direct effect on GY	0.03300889	0.0903076
Indirect effect via FPH	0.9276974	0.06215741
Indirect effect via Ld	−0.00339918	0.00830679
Indirect effect via AV	0.08415308	0.01956752
Indirect effect via NP	−0.04887024	−0.10335175
Indirect effect via NS	−0.0355623	0.08494837
TOTAL (direct and indirect)	0.1221	0.1710
Variable ==> FPH		
Direct effect on GY	0.15715694	0.10645215
Indirect effect via PHM	0.01948515	0.05273061
Indirect effect via Ld	−0.0064706	0.01207393
Indirect effect via AV	0.04439131	−0.0029826
Indirect effect via NP	−0.08549483	−0.19894989
Indirect effect via NS	−0.41826678	−0.25201769
TOTAL (direct and indirect)	−0.2892	−0.272
Variable ==> Ld		
Direct effect on GY	−0.01365682	0.01036923
Indirect effect via PHM	0.00821591	0.07234542
Indirect effect via FPH	0.07446096	0.12395288
Indirect effect via AV	0.10777732	−0.05906711
Indirect effect via NP	−0.06412112	−0.33483114
Indirect effect via NS	−0.45364705	−0.55618563
TOTAL (direct and indirect)	−0.3410	−0.7424
Variable ==> AV		
Direct effect on GY	−0.17587683	0.01657001
Indirect effect via PHM	0.0066744	0.10664424
Indirect effect via FPH	−0.03966641	−0.01916139
Indirect effect via Ld	0.0083689	−0.03696319
Indirect effect via NP	0.73628945	0.12765233
Indirect effect via NS	0.06648037	0.70520111
TOTAL (direct and indirect)	0.6023	0.9570
Variable ==> NP		
Direct effect on GY	0.14043171	0.29707315
Indirect effect via PHM	−0.01148709	−0.03141801
Indirect effect via FPH	−0.09567714	−0.07129101
Indirect effect via Ld	0.0062357	−0.01168716
Indirect effect via AV	−0.08326009	0.00712013
Indirect effect via NS	0.76237491	0.53312642
TOTAL (direct and indirect)	0.7186	0.7529
Variable ==> NS		
Direct effect on GY	0.99944272	0.62831635
Indirect effect via PHM	0.00277935	0.01220959
Indirect effect via FPH	−0.06577018	−0.04269796
Indirect effect via Ld	0.00619883	−0.00917884
Indirect effect via AV	−0.12956846	0.03565535
Indirect effect via NP	0.10712131	0.25206656
TOTAL (direct and indirect)	0.9202	0.939
Determination coefficient (R^2)	0.87790838	0.81808536
Effect of the residual variable	0.34941611	0.42651453

PHM: plant height at maturity; FPH: first pod height; Ld: lodging; AV: agronomic value; NP: number of pods per plant; NS: number of seeds per plant; and GY: grain yield.

the phenotypic and genotypic coefficient (0.762 and 0.533, respectively), which are higher than the residual effect (0.349 and 0.426, respectively). Consequently, a greater emphasis should be given to these components during selection to obtain high yields. The residual effect on seed yield per plant was low, showing that the traits under study were sufficient to obtain the effects on grain yield.

The results, therefore, showed that the number of seeds per plant and the number of pods per plant were the most important attributes since they contributed with direct and indirect effects, respectively, to determine the highest grain yield, being in accordance with the results of other researches (Silva et al., 2015; Li et al., 2013). Chandel et al. (2014) observed that path analysis in 70 soybean genotypes showed that the number of seeds per pod had a positive direct effect, as well as an indirect effect via other traits on seed yield per plant.

The trait number of pods per plant presented a high favorable phenotypic and genotypic correlation with grain yield, with values of 0.7186 and 0.7529, respectively. The direct effect of the number of pods on grain yield is small for both correlations (0.1404 and 0.297), which is masked by the indirect effect via number of seeds per plant of a high magnitude, with values of 0.762 and 0.533 for the phenotypic and genotypic coefficient, respectively. Therefore, in order to obtain satisfactory gains in grain yield due to a low direct effect of the number of pods, the simultaneous selection of both traits must be performed.

Haghi et al. (2012), Alcântara Neto et al. (2011), and Sedghi & Amanpour-Balaneji (2010) observed that the most influential trait on soybean yield was the number of pods per plant because it was the production component that had the highest direct effect on the total grain yield per plant. These authors stated that the number of pods can be considered a good option in soybean improvement in the case of an indirect selection for yield.

The indirect selection for grain yield, when combined with the resistance to soybean cyst nematode, may represent an efficient strategy to be adopted for the optimization of

the genetic gain in soybean. Nogueira et al. (2012) and Akram et al. (2011) concluded that phenotypic and genotypic correlations, as well as path analysis, indicated the total number of pods per plant as a trait with the highest favorable effect on soybean grain yield.

Plant height at maturity presented a low correlation with grain yield, exerting a direct influence on the phenotypic correlation of only 0.033 and genotypic correlation of 0.09. Its contribution was mostly determined by the indirect effect via first pod height for the phenotypic correlation (0.928) and via number of pods (−0.103) for the genotypic correlation. Differing from this result, Alcântara Neto et al. (2011) verified that plant height at maturity had an influence on the total grain yield per plant, with a contribution of indirect effect via number of pods per plant.

The correlation between the first pod height and grain yield was negative. It had a direct effect on grain yield of 0.157 and 0.106, respectively, for the phenotypic and genotypic coefficient. The trait first pod height was strongly influenced by other traits, especially for the number of seeds per plant, with values of −0.4183 and −0.252, respectively, in the phenotypic and genotypic estimates. Bárbaros et al. (2007) observed low values of correlation coefficients between first pod height and the number of seeds per plant, besides a negative direct effect of it on the number of pods per plant.

The trait lodging had a little direct effect on grain yield, with a value of −0.013 for the phenotypic estimate and 0.010 for the genotypic estimate, not exceeding the magnitude of the residual effect and being fundamentally determined by the indirect phenotypic and genotypic effect via number of seeds per plant (−0.4537 and −0.5562, respectively).

For the agronomic value index, the direct effect on grain yield was negative and the total effect was positive and of a high magnitude, with a value of 0.6023 for the phenotypic analysis because it had a high indirect effect via number of pods per plant on grain yield in the phenotypic coefficient (0.736). The high correlation verified between agronomic value and grain yield in the genotypic coefficient is determined to a large

extent by the indirect effect via number of seeds per plant on yield (0.705). Thus, traits with a high favorable correlation with the main variable, but with a low direct or unfavorable effect, indicates the absence of cause and effect, i.e. this auxiliary trait does not cause a determining direct effect in the changes of the main variable, with other traits that may have a higher impact in terms of gains.

It is evident from the path analysis that the maximum direct effects, as well as indirect influences through other traits, were exerted by the number of seeds per plant and number of pods per plant. Thus, they can be considered as the traits of yield due to their important contribution. Therefore, emphasis should be given to these traits during phenotypic selection for the development of soybean cultivars with high grain yield.

The results indicated that, for a greater success in the selection process aiming to increase soybean grain yield, the genotypes with the highest number of pods per plant and mainly with the highest number of seeds per plant were the most favorable, presenting the highest potential for an efficient indirect selection. Indirect selection with the aim at increasing grain yield, together with the resistance to the soybean cyst nematode, may represent an efficient strategy to be adopted for the optimization of the genetic gain in soybean.

Conclusions

Genotypic correlations showed higher coefficients when compared to their corresponding phenotypic correlations for most of the traits, indicating that most of the contributions in the associations between traits are due to genetic causes.

Phenotypic and genotypic correlations, as well as path analysis, indicated that the agronomic value, number of pods per plant, and number of seeds per plant had favorable effects on soybean grain yield per plant.

The number of seeds per plant had the highest direct effect, while agronomic value and the number of pods had the highest indirect effect via number of seeds per plant.

The traits number of seeds per plant, number of pods per plant, and agronomic value can be used as auxiliary in the indirect selection process aiming to increase grain yield in soybean breeding programs.

References

Akram, R.M., Fares, W.M., Fateh, H.S.A., Rizk, A.M.A. 2011. Genetic variability, correlation and path analysis in soybean. *Egyptian Journal Plant Breed* 15: 89 – 102.

Alcântara Neto, F., Gravina, G.A., Monteiro, M.M.S., Morais, F.B., Petter, F.A., Albuquerque, J.A.A. 2011. Análise de trilha do rendimento de grãos de soja na microrregião do Alto Médio Gurguéia. *Comunicata Scientiae* 2: 07-112.

Almeida, R.D., Peluzio, J.M., Afferri, F.S. 2010. Correlações fenotípicas, genotípicas e ambientais em soja cultivada sob condições várzea irrigada, sul do Tocantins. *Bioscience Journal* 26: 95-99.

Asmus, G.L., Teles, T.S., Anselmo, J., Rosso, G.T. 2012. Races of Heterodera glycines in the Northeast of Mato Grosso do Sul, Brazil. *Tropical Plant Pathology*, 37: 146-148.

Bárbaro, I.M., Centurion, M.A.P., MAURO, A.O.D., Unêda-Trevisoli, S.H., Costa, M.M., Muniz, F.R.S.; Silveira, G.D., Sarti, D.G.P. 2007. Variabilidade e correlações entre produtividade de grãos e caracteres agronômicos de soja com aptidão para cultivo em áreas para reforma de canavial. *Científica* 35: 136-145.

Chandel, K.K., Patel, N.B., Patel, J.B. 2014. Correlation coefficients and path analysis in soybean (*Glycine max* L. Merrill). *AGRES – An International e-Journal* 3: 25-31.

Cruz, C.D. 2013. GENES - a software package for analysis in experimental statistics and quantitative genetics. *Acta Scientiarum* 35: 271-276.

Cruz, C. D., Regazzi, A. J., Carneiro, P. C. S. 2012. *Modelos Biométricos Aplicados ao Melhoramento Genético*. Editora UFV, Viçosa, Brasil. 514 p.

El-Mohsen, A.A.A., Mahmoud, G.O., Safina, S.A. 2013. Agronomical evaluation of six soybean cultivars using correlation and regression analysis under different irrigation regime conditions. *Journal of plant breeding and crop science* 5: 91-102.

Falconer, D.S. 1987. *Introdução à genética quantitativa*. Tradução de Silva MA. & Silva JC. Universidade Federal de Viçosa. Imprensa Universitária, Viçosa, Brasil. 279 p.

Fehr, W.R., Caviness, J.A. 1977. *Stages of soybean development.* Aimes: Iowa State University, 11 p. (Special Report, 80).

Ghodrati, G.R., Sekhavat, R., Mahmoodinezhadedezfully, S.H., Gholami, A. 2013. Evaluation of correlations and path analysis of components seed yield in soybean. *International Journal of Agriculture* 3: 795-800.

Haghi, Y., Boroomandan, P., Moradin, M., Hassankhali, M., Farhadi, P., Farsaei, F., Dabiri, S. 2012. Correlation and path analysis for yield, oil and protein content of Soybean (Glycine max L.) genotypes under different levels of nitrogen starter and plant density. *Biharean Biologist* 6: 32-37.

Hamawaki, O.T., Sousa, L.B., Romanato, F.N., Nogueira, A.P.O., Santos Júnior, C.D., Polizel, A.C. 2012. Genetic parameters and variability in soybean genotypes. *Comunicata Scientiae* 3: 76-83.

Li Y.S., Du M., Zhang Q.Y., Hashemi M., Liu X.B., Hebert S.J. 2013. Correlation and path coefficient analysis for yield components of vegetable soybean in Northeast China. *Legume Research – An International Journal* 36: 284-288.

Machikowa, T., Laosuwan, P. 2011. Path coefficient analysis for yield of early maturing soybean. *Sonklanakarin Journal of Science and Technology* 33: 365-368.

Montgomery, D.C., Peck E.A. 1981. *Introduction to linear regression analysis.* John Wiley e Sons, New York, 504 p.

Nogueira, A.P.O., Sediyama, T., Sousa, L.B., Hamawaki, O.T., Cruz, C.D., Pereira, D.G., Matsuo, E. 2012. Análise de trilha e correlações entre caracteres em soja cultivada em duas épocas de semeadura. *Bioscience Journal* 28: 877-888.

Resende, M. D. V. 2007. *SELEGEN-REML/BLUP: sistema estatístico e seleção genética computadorizada via modelos lineares mistos.* Embrapa Florestas, Colombo, Brasil. 359 p.

Santana, H., Pires, E., Comerlato, A.P., Nasu, E.G., Furlanetto, C. 2009. Genetic variability in field populations of the soybean cyst nematode from the States of Paraná and Rio Grande do Sul, Brazil. *Tropical Plant Pathology*, 34: 261-264.

Sedghi, M., Amanpour-Balaneji, B. 2010. Sequential path model for grain yield in soybean. *Notulae Scientia Biologicae*, 2: 104-109.

Showkat, M., Tyagi, D. 2010. Correlation and path analysis of some quantitative traits in soybean (Glycine max L. Merrill). *Research Journal of Agricultural Sciences* 1: 102-106.

Silva, A.F., Sediyama, T., Silva, F.C.S., Bezerra, A.R.G., Ferreira, L. V. 2015. Correlation and path analysis of soybean yield components. *International Journal of Plant, Animal and Environmental Sciences* 5: 177-179.

Vianna, V.F., Unêda-Trevisoli, S.H., Desidério, J.A., Santiago, S., Charnai, K., Ferreira Júnior, J.A., Ferraudo, A.S., Mauro, A.O. 2013. The multivariate approach and influence of characters in selecting superior soybean genotypes. *African Journal of Agricultural Research* 8: 4162-4169.

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