Phenotypic correlation and path analysis for yield in soybean (*Glycine max* (L.) Merril)

Correlación fenotípica y análisis de sendero para el rendimiento de soya (*Glycine max* (L.) Merril)

Rubén Alfredo Valencia-Ramírez[1*] *and Gustavo Adolfo Ligarreto-Moreno*[2†]

[1]Agronomist Engineer, PhD, Colombian Corporation of Agricultural Research, Corpoica, Director Research Center La Libertad, Villavicencio-Meta. [2]Agronomist Engineer, PhD, Faculty of Agronomy, Universidad Nacional de Colombia, Bogotá, Colombia. *Corresponding author: ravalencia@corpoica.org.co; †galigarretom@unal.edu.co

Rec.: 14.12.11 Acept.: 07.12.12

Abstract

This research determinate the correlations between seed yield and agronomic characteristics in soybean (*Glycine max* (L.) Merril) and the direct and indirect effects of certain yield components on seed yield. Through the path coefficients analysis it was identified an indirect selection criteria for hight seed yields. Thirty treatments were evaluated in a strip plot design with four replications in Oxisols soils of the Colombian Orinoquía, in seven environments during 2007 and 2008. The treatments consisted of a combination of six soybean varieties and five nitrogen sources. Combined data over the locations indicated that seed yield had significant positive correlation with number of pods per plant -VT (r = 0.81**), number of nudes per plant -NN (r = 0.67**), number of seed per pod –GV (r = 0.44*), and negative correlation with weight of seed –PS (r = -0.44*). Path coefficient analysis showed that VT gave the greatest direct positive effect (0.628) on seed yield, followed by NN (0.260). However, the number of pods with three seeds (V3) had high association with seed yield (r = 0.84**). The greater attention should be given to these yield components (VT, NN and V3) as indirect selection criteria for genetic improvement of soybean seed yield.

Key words: *Bradyrhizobium japonicum*, characters of yield, Colombia, genetic inheritance, Oxisols, phenothypes.

Resumen

En este trabajo se determinaron correlaciones entre el rendimiento de grano y algunas características agronómicas en soya (*Glycine max* (L.) Merril) y los efectos directos e indirectos de los componentes de rendimiento sobre la producción. A través del análisis de sendero se identificaron criterios de selección indirecta para alto rendimiento. Para el efecto se empleó un diseño de franjas divididas con cuatro repeticiones y 30 tratamientos en Oxisoles de la Orinoquía colombiana, en siete ambientes, durante 2007 y 2008. Los tratamientos consistieron en la combinación de seis variedades de soya y cinco fuentes de nitrógeno. Los datos unidos por ambientes mostraron una correlación fenotípica positiva entre rendimiento de grano con el número de vainas por planta - VT (r = 0.81**), el número de nudos por planta-NN (r = 0.67**), y el número de semilla por vaina-GV (r = 0.44*), mientras que fue negativa con peso de grano -PS (r = -0.44*). El análisis de sendero mostró que VT presentó el mayor efecto positivo directo (0.628) sobre el rendimiento, seguido por NN (0.260). También, el número de vainas con tres granos (V3) tuvo alta asociación fenotípica con rendimiento (r = 0.84**). Es importante prestar atención a estos componentes de rendimiento VT, NN y V3 como criterios de selección indirecta para el mejoramiento genético del rendimiento de grano en soya.

Palabras clave: *Bradyrhizobium japonicum*, caracteres de rendimiento, Colombia, fenotipos, herencia genética, Oxisoles.

Introduction

Soybean (*Glycine max* (L). Merril) is an oleaginous legume of short cycle that grows on tropical, subtropical and temperate regions. Nowadays, there are 10,000 ha cultivated per year in the savanna of the Colombian altillanura, with a potential of in area close to 560,000 ha, this area has the largest competitive and productive advantages for soybean production in Colombia as a source of protein and energy for the poultry and pig production chains (Valencia *et al.*, 2006).

The association between characters of plant breeding interest can be evaluated by phenotypic, genotypic and environmental correlations. Phenotypic correlations are directly estimated with the mean phenotypic values on the field that are a result of the genetic and environmental causes. Genotypic correlation, contrarily, is the genetic part of the phenotypic correlation caused mainly by pleiotropy, although linkage could be a temporal cause (Ceballos, 2003). Pleiotropy is the ability of a gene to affect more than one character on a positive or negative way (Flaconer, 1986). A correlation, indistinctive of its nature, is the ratio of the appropriate covariance over the product of two standard deviations. Thus, the phenotypic correlation between the X and Y characters, according to Falconer (1986) is the following:

$$r_{(x,y)} = COV/\ \delta_{(x)}.\delta_{(y)},$$

where, $r_{(x,y)}$ = correlation of x, y; COV = covariance of x, y; $\delta_{(x)}$ = deviation of x; $\delta_{(y)}$ = deviation of y.

Malik *et al.* (2007) state that the study and knowledge of the direct and indirect effect on yield and its component give the basis for a successful genetic breeding program, taking as reference strongly related characters. The knowledge on the interaction between the characters that are considered important for the breeding program is a fundamental part for the progressive breeding of the species, by making more efficient the parental and progeny selection. According to Oz *et al.* (2009) a breeding program could be planned to increase the seed production if there is a good association between certain agronomical characteristics and production. In several research studies it has been demonstrated that the correlation of a particular character with other characters is of high importance for the indirect selection of desirable genotypes (Sarawgi *et al.*, 1997; Oz *et al.*, 2009).

Coefficients of correlation despite being useful in the quantification of magnitude and direction of the factor effect in the determination of complex characters, these coefficients do not indicate the exact importance that have the direct and indirect effects of these characters on the variable of interest (Abbott *et al.*, 2007). A high or low coefficient of correlation between two variables could be due to an effect of a third variable or group of variables. Therefore, according to Seker and Serin (2004), the correlation between two attributes does not predict the success of selection.

As a statistical alternative to overcome this difficulty, Wright (1921) developed the pathway analysis or coefficient path to investigate the branching of several causality models on genetic populations and, nowadays, it is used on different knowledge areas. The method assumes the variables are associated within them by lineal relations (Roehrig, 1996) and that their coefficients are standardized, because they are estimated from the correlation coefficients. The analysis allows decomposition of the correlations between two variables in direct and indirect effects (Mitchell, 1992) and, facilitates the identification of possible causal explanations of the observed correlations between a variable of endogenous or dependent response and a series of exogenous predictive or independent variables. This method differs from the multiple regression models, because it establishes causal relationships between variables (Carrascal, 2004; Whittaker *et al.*, 2009).

Wright´s coefficient, or path coefficients, have been used in soybean for diverse type of studies, among them, identify indirect selection criteria in late sowings (Board *et al.*, 1997), effect of the environment over yield (Board *et al.*, 1999), yield components on different sowing types in soybean (Pandey and Torrie, 1973), the effect of population density on yield (Ball *et al.*, 2001), yield components on soybean hybrids (Taware *et al.*, 1997), weed interference with plant growth (Jordan, 1992), yield component between genotypes (Shukla *et al.,* 1998), relationship between pod dehiscence with other agronomical characters (Tiwari and Bhatnagar, 1991), varietal differences, yield components, oil and protein (Malik *et al.*, 2006a), genetic diversity to improve grain yield (Malik *et al.*, 2006b). Villalobos *et al.* (1985) concluded that the characteristics that show the largest direct effect on soybean grain yield are, at the same time, the cause of the largest variability.

In agreement with the above statements, this work had as objective the estimation of the coefficients of correlation between grain yield per hectare with its components and other agronomical characteristics and, the determination of the causal direct and indirect effects of the evaluated variables on soybean grain yield in seven environments of the Colombian Orinoquía.

Materials and methods

Experimental assays of soybean varieties and *Bradyrhizobium japonicum* strains were performed during 2007 and 2008, in Oxisol soils of the Colombian Orinoquía classified as class IV, in environments localized between 04° 03' N and 73° 29' W between 150 and 336 MASL. Accumulated precipitation per crop cycle, mean temperature and mean solar brightness per semester and year are displayed in the Table 1. The experiment was established to determine the genotype-strain x environment interaction of the combination of six soybean varieties and three nitrogen fixation strains (*B. japonicum*), one nitrogen level (150 kg/ha) and a mix of strains, for a total of 30 treatments (Vc) (Table 2) with seven environment used for the path analysis. The design used was a strip plot with four replicates, the main plots were the varieties and subplots were the rhizobium strains or N dosis. Each experimental unit consisted of 16 m^2 composed of eight lines 5m long, a useful plot of four lines, for a total area of 8 m^2, separation between lines was 40 cm and between plants was 5 cm. fertilization was performed based on the soil analysis and the crop requirements. As they were acid soils, dolomite lime was applied on required doses to reach 50% of base saturation according to Raij *et al.* (1985).

Soybean evaluated varieties were: Soyica P-34, Orinoquía-3, Corpoica La Libertad-4, Corpoica Taluma-5, Corpoica Superior-6 and Corpoica Sabana-7, widely described by Valencia *et al.* (2006) and the *B. japonicum* strains: ICA J-01 generated by the ICA (Salamanca and Ramírez, 2000), and J-96 and J-98 from the Bank of Germoplasm of Corpoica that were introduced from Brasil, the mix J-01 + J-96 and, one level of N (150 kg/ha). The origin and denomination of the Brazilian strains is described in Chen *et al.* (2000).

Five plants were randomly selected in each useful plot, to perform observations on the vegetative, phenological, reproductive and

Table 1. Weather characteristics in the Piedemonte and altillanura of the Eastern llanos of Colombia. Meteorological stations La Libertad and Margaritas, 2007 y 008.

Region	Semester	Accumulated rainfall	Mean temperature	Relative humidity	Solar brightness
Piedemonte 2007	A	1370	25.3	87	4.5
	B	1122	25.9	86	5.5
Piedemonte 2008	A	1566	25.0	88	4.3
	B	976	25.7	84	6.2
Altillanura 2008	A	123	26.5	80	4.4
	B	634	26.0	85	6.5

Table 2. Relation of the treatments (Vc) consisted by the combination of soybean varieties and strains of *B. japonicum.*

Variety	Strain	Treatment (Vc)	Variety	Strain	Treatment (Vc)
Soyica P-34	J-01	1	C. Taluma-5	J-jF	16
Orinoquia 9	J-01	2	C. Superior-6	J-jF	17
C. Libertad-4	J-01	3	C. Sabana-7	J-jF	1F
C. Taluma-5	J-01	4	C. Sabana-7	J-01 + J-96	19
C. Superior-6	J-01	5	C. Superior-6	J-01 + J-96	20
C. Sabana-7	J-01	6	C. Taluma-5	J-01 + J-96	21
C. Sabana-7	J-96	7	C. Libertad-4	J-01 + J-96	22
C. Superior-6	J-96	F	Orinoquia 9	J-01 + J-96	23
C. Taluma-5	J-96	9	Soyica P-34	J-01 + J-96	24
C. Libertad-4	J-96	10	Soyica P-34	150 Kg ha N	25
Orinoquia 9	J-96	11	Orinoquia 9	150 Kg ha N	26
Soyica P-34	J-96	12	C. Libertad-4	150 Kg ha N	27
Soyica P-34	J-98	13	C. Taluma-5	150 Kg ha N	BF
Orinoquia 9	J-98	14	C. Superior-6	150 Kg ha N	29
C. Libertad-4	J-98	15	C. Sabana-7	150 Kg ha N	30

productive characteristics, among them: days to full flowering (DF), days to maturity (DM), plant height in cm (ALM), number of nodes to maturity (NN), filling onset or height to the first pod in cm (IC), number of pods with one (V1), two (V2), three (V3) and four (V4) grains, number of empty pods (VV), number of embryonic abortions (AE), number of grains per pod (GV), number of total pods per plant (VT), 100 grains weight in g (PS), grain yield in t/ha (Rend) calculated based on the experimental plot of 8 m^2, number of rhizobium nodules per plant (NNR), nodule dry weight in mg (PSN), root dry weight in g (PSR) and aerial part dry weight in g (PST). The variable yield was measured from the useful plot and its value was adjusted by the differential between fresh and dry weight at 13% of humidity. Embryonic abortion was established as the number of compartments without grain into one pod.

Statistical analysis

Simple correlations (*r*) were obtained for all the possible combinations of the characteristics associated with grain yield. Genetic correlations were estimated with the software Genes (Cruz, 2007) using the formulas for correlation as described by Falconer (1986) and described by Ceballos (2003).

Phenotypic correlation= $r_{F(xy)} = COV_{F(xy)}/\delta_{F\,(x)}.\delta_{F\,(y)}$

Genetic correlation= $r_{G(xy)} = COV_{G(xy)}/\delta_{G\,(x)}.\delta_{G\,(y)}$

Environmental correlation= $r_{E(xy)} = COV_{E(xy)}/\delta_{E\,(x)}.\delta_{E\,(y)}$

where, $r_{(xy)}$ and $COV_{(xy)}$ are the correlations and phenotypic (F), genetic (G) and environmental (E) covariances between the x, y characters, respectively; $\delta_{(x)}$ and $\delta_{(y)}$ are the phenotypic (F), genetic (G) and environmental (E) standard deviations of x and y respectively.

To establish the consistency of the general mean phenotypic correlations by the effect of the interaction VcA (variety-strain x environment), significance levels ($P < 0.05$ and $P < 0.01$) and the magnitude of each variable were compared with the Pearson correlation coefficients obtained by semester, considering this as highly relevant factor in the interaction. These correlations were the base to evaluate the relative contribution of each component in the grain production using the path analysis, which is a multivariate statistical technique useful in the analysis of the interactions between variables.

Path analysis was applied to grain yield per hectare (X5) as endogenous variable (response or dependent variable) and the variables that comprised yield: number of nodes per plant (X1), number of grains per pod (X2), total pods (X3) and weight of 100 grains (X4) as exogenous variables (predictive or independent variables). The estimation of the effects was performed with model proposed by Ball *et al.* (2001). To fractionate this correlation in direct and indirect effects the software R version 2.7.0 (2008) was used.

Path coefficients were represented by p_{15}, p_{25}, p_{35} and p_{45} that correspond to the direct effects on grain yield from the number of nodes per plant (NN), number of grains per pod (GV), number of total pods per plant (VT) and weight of 100 grains (PS), respectively (Figure 1).

Indirect effects estimated for one component (p.e. X1) through a second component (X2) correspond to the product of the direct coefficient of the second component and to the correlation between both components (Dewey and Lu, 1959). The structural formula of the direct and indirect effects of X1 is defined as:

$$P_{l5} + P_{25}r_{x1x2} + P_{35}r_{x1x3} + P_{45}r_{x1x4} = r_{x1x5}$$

Results and discussion

In Table 3 are included the coefficients of correlation between the characters evaluated in the combination of all the environments and by semester. It is observed a positive and significant relation between grain yield and number of pods with three grains (V3), number of pods per plant or total pods (VT) and number of grains per pod (GV), plant height (ALM), number of nodes per plant (NN), number of rhizobium nodules per plant (NNR), nodule dry weight (PSN), root dry weight (PSR) and aerial part dry weight (PST); and negative and significant with grain weight (PS).

In the first semester the variables V2 and VV, and in the second semester the variable AE had a statistical significance, in contrast to the coefficients of combined analysis for soybean. Similarly, in the second semester the variables GV, PS, NNR and PSN did not show significant differences.

These results show a noticeable differential effect of the semester on the behavior of some agronomical variables of the plant.

Path analysis

The path analysis of the direct and indirect effects of the agronomical characters in respect to the grain yield of the Vc treatments (varieties x strains-N), combined by location is displayed in Table 4. The greatest direct effect was obtained with number of nodes (NN), with a coefficient (r) of 0.806 and an input of 53.1%, followed by days to maturity (DM) with -0.378 and 36.2%, respectively. The lowest contribution was obtained with height of plant and root dry weight. Arshad *et al.* (2006) found similar results for days to maturity with

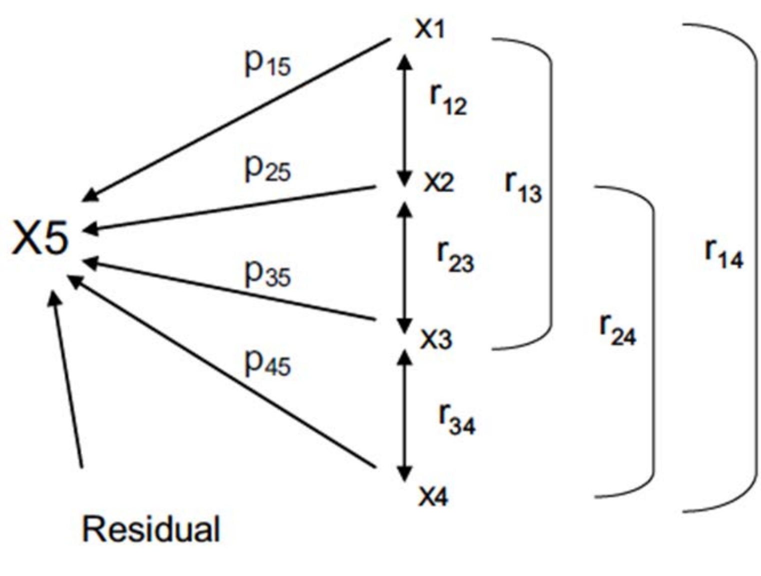

Figure 1. Diagram of the path for soybean with direct and indirect effects of the independent variables on the dependent variables.

Table 3. Coefficients of phenotypic correlation for grain yield/ha with the agronomical variables evaluated, combined per environment and semester, in Oxisols of the Colombian Orinoquía.

Variable	Coef. (r) combinado	Coef. (r) semestre A	Coef. (r) semestre B	Variable	Coef. (r) combinado	Coef. (r) semestre A	Coef. (r) semestre B
DF	-0.07	-0.01	-0.01	GV	0.44*	0.53**	0.18
DM	0.01	0.01	0.02	VV	-0.33	-0.44*	0.23
ALM	0.63**	0.61**	0.56**	AE	0.28	0.28	0.46**
NN	0.67**	0.64**	0.52**	VT	0.81**	0.79**	0.63**
IC	0.11	0.26	0.08	6U	-0.44**	-0.21	-0.18
V1	0.1F	0.18	-0.08	NNR	0.70	0.76	0.27
V2	0.19	0.46**	0.18	6UN	0.63	0.70	0.13
V3	0.84**	0.84**	0.53**	6U[	0.38	0.30	0.51
V4	0.02	0.00	0.07	6UK	0.65	0.68	0.733

La correlación es significativa (*) al nivel P < 0.05 y altamente significativa (**) al nivel P < 0.01.

Table 4. Path and participation coefficients (%) with direct and indirect effects of the agronomical characters on the grain yield of Vc (soybean varieties combined with *B. japonicum* strain), in Oxisols and Colombian Orinoquía.

Caracter	DF	%	DM	%	ALM	%	NN	%	NNR	%	PSN	%	PSR	%	PST	%
Days to flowering (DF)	**-0.131**[†]	**28.1**	0.027	2.6	0.008	0.50	0.020	1.3	-0.026	2.9	-0.010	1.3	-0.020	1.8	0.005	0.5
Days to maturity (DM)	0.079	16.9	**-0.378**	**36.8**	-0.216	14.80	-0.208	13.7	0.023	2.5	0.026	3.3	-0.182	16.7	-0.057	5.6
Plant height (ALM)	0.004	0.9	-0.037	3.6	**-0.066**	**4.50**	-0.063	4.2	-0.014	1.5	-0.009	1.1	-0.045	4.1	-0.036	3.5
Nodes number (NN)	-0.121	25.9	0.443	43.1	0.774	53.00	**0.806**	**53.1**	0.242	26.7	0.161	20.5	0.500	45.8	0.371	36.5
Nodule number (NNR)	0.049	10.5	-0.015	1.5	0.054	3.70	0.074	4.9	**0.246**	**27.1**	0.222	28.2	0.025	2.3	0.074	7.3
Nodule dry weight (PSN)	0.017	3.6	-0.015	1.5	0.028	1.90	0.043	2.8	0.194	21.4	**0.216**	**27.5**	-0.004	0.4	0.054	5.3
Root dry weight (PSR)	-0.004	0.9	-0.012	1.2	-0.018	1.20	-0.016	1.1	-0.003	0.3	0.001	0.1	**-0.026**	**2.4**	-0.017	1.7
Aerial dry weight (PST)	-0.013	2.8	0.049	4.8	0.181	12.40	0.151	9.9	0.099	10.9	0.082	10.4	0.211	19.3	**0.329**	**32.4**
Total correlation	**-0.070**		**0.010**		**0.630**		**0.670**		**0.700**		**0.630**		**0.380**		**0.650**	

[†]Numbers in bold correspond to the direct effects and the rest are indirect effects.

direct negative effects on yield. The selection based on the characteristics with negative input can represent losses in terms of grain yield. For this reason, on a soybean breeding program the selection towards early varieties should be taken cautiously. This effect of the maturity on grain yield is clearly observed in soybean germplasm introduced from other latitudes that when sown in the tropics, as it is a short cycle crop, flowers and matures early.

The number of nodules (NNR), nodule dry weight (PSN) and aerial dry weight (PST), had direct positive effects on the yield and notorious indirect effects on yield by number of nodes per plant (NN). In general, the greatest indirect effect of the agronomical variables evaluated was shown with the number of nodes per plant (NN). This characteristic was consistent within semesters, in the combination by locations and because it correlated significantly with grain yield ($r = 0.67$; $P < 0.0001$) and it had a direct input superior to 50% of the total correlation, was included in the path analysis as exogenous variable to value the input of the yield components on grain production.

Grain yield per hectare vs. yield components

For path analysis of the yield components there were selected those exogenous variables directly directed with the endogenus variable (grain yield/ha) such as: number of pods per plant (VT), average number of grains per pod (GV), grain weight (PS) and number of nodes (NN).

The path coefficients of the combined analysis and the averages per semester are displayed in the Table 5, where it can be observed that the total effect or R is the results of the addition of the direct effect with the indirect one. In the path analysis joining the combined effects and by semester, the greatest input on the direct effects was for number of pods per plant (VT) with coefficients between 0.628 and 0.836, and participation on the total correlation between 61.3 and 77.5%. These results agree with the ones from Rajanna *et al.* (2000), who concluded that the number of pods is an important component of yield in the soybean crop. In relation to the number of pods per plant (VT), the research on soybean crop done by Oz *et al.* (2009) concluded that the positive association ($r = 0.458$, $P < 0.01$) with grain yield was stable through the years and locations and that the number of pods per plant happened from the direct effect on a 50%. For Board *et al.* (1997) the number of grains per reproductive node was the best criteria of indirect selection for grain yield.

In second place, it was the number of nodes at maturity (NN), with direct effects from 0.260 to 0.524 and input between 38.8 and 44.5%; the degree of association between

Table 5. Direct and indirect effects of the yield components on grain yield of Vc (variety x strain), in Oxisols of the Colombian Orinoquía.

Effects	**Combined**		**Sem. A**		**Sem. B**	
	Coefic.	**%**	**Coefic.**	**%**	**Coefic.**	**%**
			Node number (NN)			
Direct effects	0.260	38.8	0.524	44.5	0.340	46.0
Indirect effects	—	—	—	—	—	—
-by grain number (GV)	0.001	0.1	-0.248	21.1	-0.033	4.4
-by total pods(VT)	0.345	51.5	0.385	32.7	0.289	39.1
-by grain weight (PS)	0.064	9.6	-0.021	1.7	-0.077	10.4
Total phenotypic correlation	**0.67**		**0.64**		**0.52**	
			Number of grains per pod (GV)			
Direct effects	0.002	0.3	-0.381	28.5	-0.069	21.7
Indirect effects						
-by grain number (GV)	0.156	32.7	0.340	25.4	0.160	50.2
-by total pods(VT)	0.301	63.1	0.594	44.4	0.032	9.9
-by grain weight (PS)	-0.019	3.9	-0.023	1.7	0.058	18.1
Total phenotypic correlation	**0.44**		**0.53**		**0.18**	
			Number of total pods (VT)			
Direct effects	0.628	77.5	0.836	61.3	0.526	66.1
Indirect effects						
-by grain number (GV)	0.143	17.6	0.241	17.7	0.187	23.5
-by total pods(VT)	0.001	0.1	-0.271	19.8	-0.004	0.5
-by grain weight (PS)	0.039	4.8	-0.016	1.2	-0.079	9.9
Total phenotypic correlation	**0.81**		**0.79**		**0.63**	
			Weight of 100 grains (PS)			
Direct effects	-0.133	30.3	0.059	9.4	0.193	34.1
Indirect effects						
-by grain number (GV)	-0.125	28.3	-0.183	29.3	-0.136	24.1
-by total pods(VT)	0.000	0.0	0.149	23.8	-0.021	3.7
-by grain weight (PS)	-0.182	41.3	-0.234	37.5	-0.216	38.2
Total phenotypic correlation	**-0.44**		**-0.21**		**-0.18**	
R^2	**0.75**		**0.79**		**0.47**	

yield and NN was explained by the indirect effect of VT in 51.5% (Table 5). This result is comparable with the ones of Bizeti *et al.* (2004) who found that the number of nodes significantly correlated with grain yield, being the one with the largest direct effect on yield, therefore, for the researchers NN could be helpful to select indirectly genotypes, without forgetting that the total number of pods per plant (VT) give the grains for a high yield, this because there could be plants with low NN but higher number of pods per node.

In Table 5 is observed that VT and NN present both the greatest direct effects and indirect effects, thus they are two useful variables as predictors of grain yield of soybean varieties grown in Oxisols of the Orinoquía. As these variables are highly correlated and as VT is consider an important variable easy to measure in soybean, this should be the first alternative to perform selection processes in yield breeding. In general, the highest average values of VT and NN were obtained with the J-96 and J-98 strains and the lowest with J-01, which is a factor that should be considered for a good interaction between strain, variety and high yield. With the two best strains, the varieties, C. Libertad 4 and Soyica P-34 presented the lowest average values for the cited characteristics and for the number of pods with three grains (V3) (Valencia and Ligarreto, 2010). These results are consistent in the variety Corpoica La Libertad-4, the Soyica P-34 variety has low adaptation to soils with aluminum problems, like Oxisols, and it is sensitive to the green stem syndrome (STW) that negatively and severely affects the quantity and quality of grains and is sown frequently at the piedemonte during the first semester. Corpoica La Libertad-4 despite of having good adaptation, shows susceptibility to *C. sojina*, abd both varieties had low grain quality due to excess of water during the first semester (Valencia *et al.*, 2006; Valencia and Ligarreto, 2010).

The number of grains per pod (GV) is a useful characteristic for yield breeding only because of its indirect effect by total number of pods (VT) since it is correlated with yield at 63.1% being, in this case, the direct effect 0.3%. These results contrast with the ones of Ball *et al.* (2001) and Oz *et al.* (2009) who

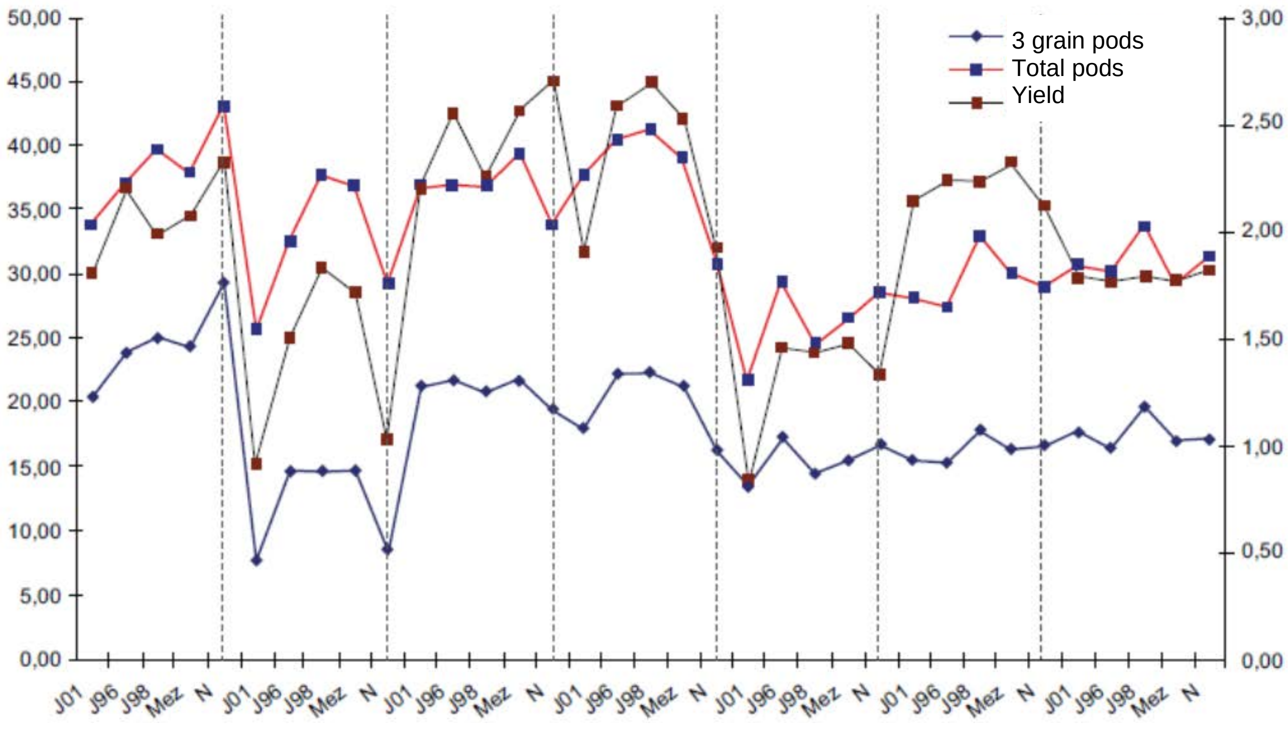

Figure 2. Behavior of the number of pods with three grains (V3), total pods (VT) and soybean grain yield, through environments and strains of *B. japonicum* –N.

found by path analysis that the number of grains per pod had a direct positive effect on grain yield. In this study no correlation was found between grain yield with pods of one (V1), two (V2) of four (V4) grains, but was highly significant with three grains (V3) pods. This variable was consistent through locations and the combination of varieties with strains, with high association with the number of pods per plant (VT) and yield grain/ha (Figure 2), which allows the deduction that V3 and VT are of great importance, highly predictive for grain yield and useful as selection criteria in the breeding program under study.

In relation with the weight of 100 grains (PS), the correlation was negative with grain yield/ha in all the cases; its direct effects, although some were negative, were not relevant to affect the endogenous variable. Direct effects were between -0.133 and 0.193, consider as very low values for interrelation. These results agree with Taware *et al.* (1997), Shuka *et al.* (1998) and Board *et al.* (1999) who, similarly, did not find significant relations between grain weight and yield. In contrast, Malik *et al.* (2007) found in soybean that with the path analysis the grain weight has a direct negative effect on yield. These differences in results, could be attributed to the influence of environmental factors and interactions in this complex variable, and its expression depends on factors like variety, predominant soil and weather conditions and presence of biotic factors like pathogens and plagues.

Genotypic and phenotypic correlations of the yield components are displayed in Table 6. In general, the genotypic correlations were superior to their respective phenotypic correlations, which could indicate a larger contribution of the genetic factors on the association; however, several phenotypic coefficients with less magnitude resulted statistically significant, in contrast to the genotypic coefficients, suggesting a high environmental effect on the studied characters. When comparing the genetic and phenotypic correlations between strains for the yield components and having into account the variable V3, it is evident that the best phenotypic correlations for

Table 6. Genotypic and phenotypic correlations (superior and inferior diagonal, respectively), of the soybean components for *B. japonicum* strains J-96, J-98 and J-01.

	NN	V3	GV	VT	PS	Yield
Strain J-96						
NN	1	0.893**	0.626	0.918**	-0.255	0.909**
V3	0.847*	1	0.631	1.005**	-0.606	0.923**
GV	0.602	0.599	1	0.848*	0.619	0.634
VT	0.823*	0.980**	0.728*	1	-0.447	0.954**
PS	-0.240	-0.440	0.478	-0.311	1	-0.616
Yield	0.853*	0.868*	0.583	0.870*	-0.395	1
Strain J-98						
NN	1	1.028**	0.546	1.000**	-0.432	1.146**
V3	0.906*	1	0.241	1.000**	-1.129	1.092**
GV	0.504	0.258	1	1.000**	0.611	0.114
VT	0.567	0.835*	0.116	1	1.000**	1.000**
PS	-0.369	-0.747	0.471	-0.658	1	1.000**
Yield	0.802*	0.743	0.017	0.314	-0.691	1
Strain J-01						
NN	1	0.832*	0.694	0.777	-0.730	1.338*
V3	0.782*	1	0.589	0.962**	-1.075	1.851*
GV	0.628	0.582	1	0.870	-0.120	0.627
VT	0.711*	0.938**	0.779*	1	-1.114	1.777*
PS	-0.643	-0.938**	-0.113	-0.868**	1	1.000**
Yield	0.610	0.870**	0.323	0.797**	-0.522	1

grain yield between strains were achieved with VT, V3 and NN, and the lower associations with GV and PS. Genetic correlations were lower for GV, while the other variables of interest were similar. In general, PS had a negative association with the variables of interest V3, NN and VT, which confirms the low importance of grain yield in selection processes. This characteristic does not contribute to the breeding of grain yield in soybean in the Oxisols of the Orinoquía. The correlation values higher than 1.0 can be attributed to sampling errors and interaction factors GA.

For the above stated, it is important that the plant breeder knows deeply the interactions between characters and their interactions with the environment to develop a successful progressive breeding program of the species. Ariyo (1995) states that a better selection criterion is considering in a simultaneous way the variables significantly associated with yield, if possible combining vegetative and reproductive characters.

Conclusions

- The path analysis is a useful statistical tool to identified interrelated characters as indirect selection alternative of grain yield for soybean in the Colombian Orinoquia.
- The number of pods per plant (VT) and pods with three grains (V3) are useful phenotypic characteristics as selection criteria to maximize grain yield in new soybean varieties in Oxisol in the Colombian Orinoquía.
- The strains of *B. japonicum* J-96 and J-98 are new options for biological nitrogen fixation, due to its great direct input on grain yield, and indirect by number and weight of rhizobium nodules.

References

Abbott, L.; Pistorale, S.; and Filippini, O. 2007.Análisis de coeficientes de sendero para el rendimiento de semillas en Bromus catharticus. Cien. Inv. Agr. 34(2):141 - 149.

Ariyo, O. 1995. Correlations and path-coefficients analysis of components of seed yield in soybeans. African crop Sci. J. 3(1):29 - 33.

Arshad, M.; Ali, N.; and Ghafoor, A. 2006.Character correlation and path coefficient in soybean [*Glycine max* (L.) Merrill]. Pak. J. Bot. 38(1):121 - 130.

Ball, R.;McNew, R.;Vories, E.;Keisling, T.; and Purcell, L. 2001. Path analyses of population density effects on short-season soybean yield. Agron. J. 93:187 - 195.

Bizeti, H. S.; Carvalho, C. G.; Souza, J. R.; and Destro, D. 2004. Path analysis under multicollinearity in soybean. Brazilian Archives of Biology and Technology 47(5):669 - 676.

Board, J.; Kang, M.; and Harville, B. 1997. Path analyses identify indirect selection criteria for yield of late-planted soybean. Crop Sci. 37:879 - 884.

Board, J.; Kang, M.; and Harville, B.1999. Path analyses of the yield formation process for late-planted soybean. Agron J. 91:128 - 135.

Carrascal, L. M. 2004. Path analysis. Depto. Biodiversidad y Biología Evolutiva. Museo Nacional De Ciencias Naturales. Madrid, España. Available in: http://www.161.111.161.171/estadistica2004/pathanalysis.pdf [Revision date: Febrero de 2010]

Ceballos, H. 2003. Genética cuantitativa y fitomejoramiento. Universidad Nacional de Colombia. Palmira. 524 p.

Chen, L. S.; Figueredo, A.; Pedrosa, F. O.; and Hungria, M. 2000.Genetic characterization of soybean rhizobia in Paraguay. Appl. Environ. Microb. 66(11):5099 - 5103.

Cruz, C. D. 2007. Programa GENES verso windows - Análises estatística, biométrica, multivariada e diversidade genética de dados quantitativos e moleculares. Editora UFV, Universidade Federal de Viosa (Brasil). Available in: www.ufv.br/dbg/genes/genes.htm [Revised date: Febrero de 2009]

Dewey, D.R.; and Lu, K.H. 1959. A correlation and path coefficient analysis of components of crested wheatgrass seed production. Agron. J. 51:515 - 518.

Falconer, D. 1986. Introducción a la genética cuantitativa. Edt. Continental. México. 383 p.

Jordan, N. 1992.Differential interference between soybean (*Glycine max*) varieties and common cocklebur (*Xanthium strumarium*): A path analysis. Weed Sci. 40(4):614 - 620.

Malik, M.; Qureshi, A.; Ashraf, M.; and Ghafoor, A. 2006a. Genetic variability of the main yield related characters in soybean. Int. J. Agric. Biol. 8(6):815 - 819.

Malik, M.; Qureshi, A.; and Ghafoor, A. 2006b.Utilization of diverse germplasm for soybean yield improvement. Asian J. Plant Sci. 5(4):663 - 667.

Malik, M.; Ashraf, M.; Qureshi, A.; and Ghafoor, A. 2007. Assesment of genetic variability, correlation and path analysis for yield and its components in soybean. Pakistan J. Bot. 39:405 - 413.

Mitchell, R. 1992. Testing evolutionary and ecological hypotheses using path analysis and structural equation modeling. Funct. Ecol. (6):123 - 129.

Oz, M.; Karasu, A.; Goksoy, A.; and Turan, Z. 2009. Interrelationships of agronomical characteristics in soybean (*Glycine max*) grown in different environments. Int. J. Agric. Biol.11:85 - 88.

Pandey, J. and Torrie, J. 1973. Path coefficient analysis of seed yield components in soybeans (*Glycine max* (L.) Merr.).Crop Sci. 13:505 - 507.

R programming for Bioinformatics.version2.7.0. 2008. The R Foundation for Statistical Computing ISBN 3-900051-07-0.

Raij, B.; da Silva, N.; Bataglia, O.;Quaggio, J.; Hiroce, R.; Cantarella, H.; Belinazzi, R. Jr.; Dechen, A.; and Trani, P. 1985. Recomendaes de adubao e calagem para o Estado de Sao Paulo. Bol. Tec. 100. Campinas, SP, Brasil. 241 p.

Rajanna, M.; Viswanatha, S.; Kulkarni, R.; and Ramesh, S. 2000. Correlation and path analysis in soybean (*Glycine max* (L.) Merrill). Crop Res. Hisar. 20(2):244 - 247.

Roehrig, S. 1996. Probabilistic inference and path analysis. Decision Support Systems 16:55 - 66.

Salamanca, C. and Ramírez, M. 2000. ICA J 01 biofertilizante para soya. Inoculante de rizobio para los Llanos Orientales. Plegable divulgativo N° 16 Corpoica-Ministerio de agricultura. FOCC. Villavicencio, Colombia.

Sarawgi, A.; Rastogi, N.; and Soni, D. 1997. Correlation and path analysis in rice accessions from Madhya Pradesh. Field Crops Res. 52:161 - 167.

Seker, H. and Serin, Y. 2004. Explanation of the relationships between seed yield and some morphological traits in smooth bromegrass (*Bromusinermis*- Leyss.) by path analysis. Europ. J. Agron. 21:1 - 6.

Shukla, S.; Singh, K.; and Pushpendra. 1998. Correlation and path coefficient analysis of yield and its components in soybean (*Glycine max* L. Merrill). Soybean Gen. Newsl. 25:67 - 70.

Taware, S.; Halvankar, G.; Raut, V.; and Patil, V. 1997. Variability, correlation and path analysis in soybean hybrids. Soybean Gen. Newsl. 24:96 - 98.

Tiwari, S. P. and Bhatnagar, P. S. 1991. Pod shattering as related to other agronomic attributes in soybean. Trop. Agric. (Trinidad) 68(1):102 - 103.

Valencia, R.; Carmen, H.; Vargas, H.; and Arrieta, G. 2006. Variedades mejoradas de soya para zonas productoras actuales y potenciales de Colombia. Innovación y cambio tecnológico-Corporación Colombiana de Investigación Agropecuaria (Corpoica) 4(2-3):7 - 15.

Valencia, R. and Ligarreto-M, G. 2010. Análisis de la interación soya-cepa (*Bradyrhyzobium japonicum*) x ambiente, en oxisoles de la Orinoquía colombiana. Agron. Col. 28(3):361 - 372.

Villalobos, E.; Barrantes, R.; and Echandi, R. 1985. Asociación de algunas características cuantitativas con el rendimiento de la soya (*Glycine max*) en el trópico mediante dos técnicas de regresión. Agron. Costarr. 9(1):65 - 70.

Whittaker, A.; Vazzana, C.;Vecchio, V.; and Benedettelli, S. 2009. Evaluation of direct and indirect effects of flavonoids, mineral elements and dry weight on antiradical scavenging activity in leaf material of field-grown Trifoliumpratense cultivars using Path Analysis. Field Crops Res. 113(1):1 - 11.

Wright, S. 1921. Correlation and causation. J. Agric. Res. 20:557 - 585.