OPEN ACCESS Asian Journal of Crop Science

ISSN 1994-7879
DOI: 10.3923/ajcs.2019.32.39

Research Article
Correlations and Path Coefficients for Yield Related Traits in Soybean Progenies

Thi Thuy Hang Vu, Thi Tuyet Cham Le, Dinh Hoa Vu, Thanh Tuan Nguyen and Thi Ngoc Pham

Department of Plant Genetics and Breeding, Faculty of Agronomy, Vietnam National University of Agriculture, TrauQuy-Gia Lam-Hanoi, Vietnam

Abstract

Background and Objective: Soybean breeding is striving to develop high yielding cultivars. Understanding of the association between yield and its components and the contribution of those yield components to yield is important to the breeding and selection process. This study sought to determine the correlations, the direct and indirect effects of yield components on grain yields in soybean progenies. **Materials and Methods:** Two soybean crosses at F_6 and F_7 generations were grown in field condition. The correlations and path coefficents of 8 measured traits, viz growth duration (days), plant height (cm), first pod insertion height (cm), ratio of first pod insertion height to plant height, total number of pods per plant, total number of seeds per plant, 100 seed weight (g) and grain yield per plant (g/plant) were computed. **Results:** There were consistencies of correlations across generations and higher direct and indirect effects in F_6 than in F_7. Most direct effects were in agreement with correlations, indicating true associations. Significant positive correlations (r) and highly positive direct effects on grain yield were observed for total number of pods (r = 0.406-0.928), total number of seeds (r = 0.434-0.939) and 100 seed weight (r = 0.361-0.626) across generations and crosses. Ratio of first pod insertion height to plant height had significant indirect effects on yield via component traits. **Conclusion:** The selection strategy could be applied in early generations for significant yield components. Besides pod and seed related traits, ratio of first pod insertion height to plant height should also be considered for selection.

Key words: Correlations, seed related traits, path coefficients, selection, soybean breeding, yield components, F_6 and F_7 generations

Citation: Thi Thuy Hang Vu, Thi Tuyet Cham Le, Dinh Hoa Vu, Thanh Tuan Nguyen and Thi Ngoc Pham, 2019. Correlations and path coefficients for yield related traits in soybean progenies. Asian J. Crop Sci., 11: 32-39.

Corresponding Author: Thi Thuy Hang Vu, Department of Plant Genetics and Breeding, Faculty of Agronomy, Vietnam National University of Agriculture, TrauQuy-Gia Lam-Hanoi, Vietnam Tel: (84) 915746863

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Soybean (*Glycine max* (L.) Merrill) is an important annual crop because of its well-known quality values for human consumption, livestock and contributions to the environment. High seed yield is the primary goal of soybean breeding in many countries, especially in Asian countries like China, Indonesia, Thailand and Vietnam because the average annual yield remains stagnant at around 1.4-1.8 t ha^{-1} in the last decade[1,2]. The yield of soybean is a complex character and affected greatly by genotype×environment (G×E) interaction. Low heritability estimates for soybean yield have been also reported in various studies[3-5]. In addition, yield is the outcome of association of several yield contributing traits which are themselves interrelated and may influence productivity directly and indirectly[6].

Estimates of association among yield component traits and their effects on yield are useful in designing an effective breeding program. In many studies using various genotypes, soybean grain yields were strongly associated with phenological and morphological traits such as days to maturity and grain filling period[7], plant height and number of branches[6,8] or associated with yield components such as number of pods per plant[9-12], number of seeds per plant, mass of a thousand grains[3,6].

The simple correlation between two traits is commonly expressed by the Pearson's linear correlation coefficient, which consists of the ratio of the two-variable joint variance known as covariance and the product of their respective standard deviations[13]. However, simple correlation may not represent the exact relationship of cause and effect between yield components and yield components with yield and may result in inefficiencies in the selection strategy[3,13]. Thus, this relationship effects of yield contributing traits on yield should be further separated to identify direct and indirect selection criteria for improvement of yield through path analysis[14,15]. Path analysis is particular useful in determining indirect selection criteria, especially in cases where component traits show high heritabilities[16]. Path analysis is widely used in various soybean studies as those reported by Vu *et al.*[17], Oz *et al.*[18], Carvalho *et al.*[19] and Sulistyo *et al.*[20].

The aim of this study was to investigate the consistency in relationship between yield and yield contributing traits using lines derived from cross populations of different genetic backgrounds and generations. Especially, first pod insertion height trait related to mechanical harvest was also included. On the basis of this relationship, prioritized traits and proper generation could be identified and suggested for breeders.

MATERIALS AND METHODS

The present investigation was carried out in an open field of Vietnam National University of Agriculture, Hanoi (Vietnam). About 36 lines of two crosses at F_6 and F_7 generations were evaluated in spring and summer cropping seasons, respectively, in the period of February-September, 2017. Crosses were created in 2014 in an attempt of breeding for yield and mechanical harvest and advanced to F_5 by single-seed descent from 2014-2016.

Plant materials: The F_6 and F_7 generations of two crosses, GBVN004898×DT26 (referred as LSB5 hereafter) and VI045032×GBVN004904 (referred as LSB10 hereafter) were used in this study. GBVN004898 and GBVN004904 were two soybean accessions provided by Plant Resources Center, Vietnam while DT26 is a commercial variety. VI045032 is an accession provided by the The World Vegetable Center.

Experimental design and cultural details: The experimental design was randomized blocks with two replications. Each line was planted in double rows in a plot of 3 m^2 (1×3 m). The seeds were sown at 10×40 cm spacing in both growing seasons. Standard agronomic practices, designed to facilitate growth with minimum environmental stress were followed to ensure similar expression of the genetic potential in both seasons. Irrigation was provided to ensure optimal growth. Pest and disease control was also applied when necessary.

Measured traits: The evaluated traits were:

- **Growth duration (GD, days):** Recorded as the number of days from sowing day to the day that 95-100% of pods in the plot turned brown
- **Plant height (PH, cm):** Measured at maturity as the length in centimeter from ground level to the tip of the main stem
- **First pod insertion height (FPIH, cm):** Measured at maturity the length in centimeter from the ground level to up to the insertion node of the first pod on the main stem
- **Total number of pods per plant (TNP):** Obtained by counting the total number of filled pods per plant at maturity

- **Total number of seeds per plant (TNS):** Obtained by counting threshed, air-dried seeds per plant
- **100 seed weight (100 SW, g):** Recorded in grams per 100 random whole seeds after air-drying to uniform moisture
- **Grain yield (GY, g/plant):** Measured from mass of threshed, air-dried grains per plant

Correlations and path analysis: Correlations and path analysis were carried out for each cross at both F_6 and F_7 generations. Measurements of individual plants were averaged across replications to obtain the means for each line. Correlations (r) among pairs of traits based on line mean basis of individual plants were obtained by using Pearson's equation:

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{(n-1)S_X S_Y}$$

where, $\overline{X}$ and $\overline{Y}$ are the sample means of X and Y, respectively, while S_X and S_Y are the standard deviations of X and Y, respectively.

Direct and indirect effects of yield components on grain yield were computed using the Excel program as described by Akintunde[21]. Grain yield was the response variable while other traits, including GD, PH, FPIH, FPIH/PH, TNP, TNS and 100 SW were the predictor variables. Briefly, means of traits were standardized, applying the equation STANDARDIZE ($X^* = (X-\overline{X})/S_X$) in Excel. Regression analysis on a set of standardized variables resulted in direct path coefficients.

Correlation coefficient r of the predictor variables with grain yield was split in to the direct effect (P_{ij}) and indirect effects which were determined by multiplying appropriate r and direct path coefficient P_{ij} as per the following equation[21]:

$$r_{ij} = P_{ij} + \sum r_{ik}P_{kj}$$

where, r_{ij} is the correlation coefficient between the predictor variable (i) and response variable (j); P_{ij} is the path coefficient measuring the direct effect of the predictor variable (i) on the response variable (j); $\sum r_{ik}P_{kj}$ is summation of components of indirect effects of a given predictor variable (i) on the response variable (j) via all other predictor variables.

Bootstrapping with 5000 times by PROCESS[22] ver 3.1 were conducted to compute confidence intervals of direct and indirect effects for the significant test. If zero is not in the interval, then the researcher can be confident that the direct and indirect effect is different from zero.

Data obtained from 72 lines, i.e., 36 lines for each of the two crosses were joined for correlation and path analysis as conducted for each cross.

RESULTS

Correlations between grain yield and yield components: The intergeneration correlations between the F_6 and the F_7 in cross GBVN004898×DT26 (LSB5), VI045032×GBVN004904 (LSB10) and joint crosses were positive for most traits but not statistically significant, such as for GD (r = 0.090 and 0.035 for LSB5 and LSB10 crosses, respectively) and for FPIH (r = 0.151 and 0.055 for LSB5 and LSB10 crosses, respectively (Table 1). This revealed the random selection of plants for evaluation.

The correlation estimates varied with trait combinations, generations and crosses. In the LSB5 cross, TNP, TNS and 100SW were positively associated with GY (r was in the range of 0.2-0.8) (Table 2). The TNP and TNS were positively correlated while TNS was non-significantly negatively associated with 100 SW. The GD was negatively associated with TNP and TNS regardless of generations. All other associations, including GY with GD, PH, FPIH and FPIH/PH were variable and negligible.

As was the case in cross LSB5, in the cross LSB10, TNP and TNS were consistently and significantly positively associated with GY (r>0.850) (Table 3). The TNP was significantly and positively correlated with TNS but these two components were negatively correlated with 100 SW in the F_7 generation (r = -0.366 and r = -0.363, respectively). Generally, morphological traits (PH, FPIH and FPIH/PH) were non-significantly and negatively correlated with GY, except for significant and positive correlation of FPIH/PH and GY in the F_7.

Table 1: Analysis of intergeneration correlation measured traits in F_6 and F_7 for each cross and joint analysis of the crosses GBVN004898×DT26 (LSB5) and VI045032×GBVN004904 (LSB10)

Crosses	GD	PH	FPIH	FPIH/PH	TNP	TNS	100 SW	GY
LSB5	0.090	0.368*	0.151	0.291	-0.017	0.017	0.178	0.243
LSB10	0.035	0.284	0.055	-0.025	-0.038	-0.139	-0.199	0.002
Joint analyses	0.569**	0.401**	-0.001	0.016	-0.061	-0.122	0.611**	0.106

*'**r-value is statistically significant at 0.05 and 0.01 probability levels, respectively, FPIH: First pod insertion height, FPIH/PH: Ratio of first pod insertion height to plant height, GD: Growth duration, GY: Soybean grain yield, PH: Plant height, TNP: Total number of pods per plant, TNS: Total number of seeds per plant, 100 SW: 100 seed weight

Table 2: Correlations between seed yield and yield components in F_6 and F_7 generations of the cross GBVN004898×DT26

Traits	Generation	GY	GD	PH	FPIH	FPIH/PH	TNP	TNS
GD	F_6	0.121						
	F_7	-0.072						
PH	F_6	-0.280	-0.149					
	F_7	-0.076	-0.371*					
FPIH	F_6	-0.160	0.137	0.025				
	F_7	0.135	0.209	-0.068				
FPIH/PH	F_6	0.210	0.158	-0.594**	0.728**			
	F_7	0.161	0.338*	-0.441**	0.920**			
TNP	F_6	0.406*	-0.297	0.080	-0.011	0.101		
	F_7	0.805**	-0.121	0.208	0.120	0.041		
TNS	F_6	0.434**	-0.267	0.059	-0.049	0.092	0.952**	
	F_7	0.822**	-0.114	0.114	0.264	0.212	0.948**	
100 SW	F_6	0.297	0.419**	-0.209	0.029	0.105	0.026	-0.064
	F_7	0.510**	0.021	-0.221	-0.093	-0.012	0.038	-0.006

*,**r-value is statistically significant at 0.05 and 0.01 probability levels, respectively, FPIH: First pod insertion height, FPIH/PH: Ratio of first pod insertion height to plant height, GD: Growth duration, GY: Soybean grain yield, PH: Plant height, TNP: Total number of pods per plant, TNS: Total number of seeds per plant, 100 SW: 100 seed weight

Table 3: Correlations between seed yield and yield components in F_6 and F_7 generations of the cross VI045032×GBVN004904

Traits	Generation	GY	GD	PH	FPIH	FPIH/PH	TNP	TNS
GD	F_6	-0.007						
	F_7	0.056						
PH	F_6	0.242	0.188					
	F_7	-0.376*	0.422**					
FPIH	F_6	-0.129	0.226	0.196				
	F_7	-0.008	0.366*	0.505**				
FPIH/PH	F_6	-0.179	0.184	-0.003	0.980**			
	F_7	0.387*	0.008	-0.442**	0.542**			
TNP	F_6	0.852**	0.002	0.242	-0.006	-0.054		
	F_7	0.928**	0.029	-0.412*	-0.032	0.402*		
TNS	F_6	0.870**	-0.036	0.256	0.024	-0.025	0.957**	
	F_7	0.939**	0.071	-0.393*	0.002	0.417*	0.986**	
100 SW	F_6	0.361*	0.179	-0.099	0.000	0.016	0.126	0.144
	F_7	0.063	0.038	0.187	0.092	-0.087	-0.366*	-0.363*

*,**r-value is statistically significant at 0.05 and 0.01 probability levels, respectively, FPIH: First pod insertion height, FPIH/PH: Ratio of first pod insertion height to plant height, GD : Growth duration, GY: Soybean grain yield, PH: Plant height, TNP: Total number of pods per plant, TNS: Total number of seeds per plant, 100 SW: 100 seed weight

More significant correlations were estimated for the joint analyses than for separate populations. Correlation estimates for the joint analyses in the F_6 and F_7 generations also indicated significant and positive associations between TNP, TNS and 100 SW with GY (Table 4). Quite similar to LSB5 and LSB10, TNP and TNS were also negatively correlated with 100 SW in F_7 of the joint analysis (r = -0.608 and r = -0.594, respectively).

Path coefficient analysis: In the LSB5 cross, TNS and 100 SW had positive and significant direct effects on GY in both generations, of which the direct effects of TNS were more important (0.491 and 0.768 for TNS in F_6 and F_7, respectively) (Table 5). Positive direct effects on GY were also observed for GD, PH and FPIH/PH but only significant for PH (0.670) and FPIH/PH (1.428) in the F_6. Path analysis also revealed that absolute values of direct effects in F_6 were greater than those in F_7, except for TNS and 100 SW and greater than those of total indirect effects for most cases, except for PH and TNP. In the F_7 of LSB5, only indirect effects of TNP on GY through TNS were significantly positive while most indirect effects of other yield components were negligible.

In the LSB10 cross, TNP, TNS and 100 SW also had positive direct effects on GY in both generations (Table 5). In particular, TNS had highest and significant positive direct effects of 0.581 and 0.949 in F_6 and F_7, respectively, followed by 100 SW (0.264 and 0.330) and TNP (0.238 and 0.127). Among indirect effects, a few indirect effects on GY in F_7 were significant such as indirect effects of PH via TNS (-0.373), of FPIH/PH via TNS (0.396) and of TNP via TNS (0.936).

Similarly, estimates of direct and indirect effects for the joint analysis were in agreement with those estimates for

Table 4: Correlations between seed yield and yield components in F_6 and F_7 generations for the joint analyses of two crosses GBVN004898×DT26 and VI045032×GBVN004904

Traits	Generation	GY	GD	PH	FPIH	FPIH/PH	TNP	TNS
GD	F_6	0.446**						
	F_7	0.080						
PH	F_6	-0.155	-0.229					
	F_7	-0.229	-0.281*					
FPIH	F_6	-0.016	0.176	0.089				
	F_7	0.035	-0.146	0.149				
FPIH/PH	F_6	0.021	0.213	-0.144	0.969**			
	F_7	0.168	-0.015	-0.340**	0.873**			
TNP	F_6	0.551**	-0.098	0.130	0.003	-0.009		
	F_7	0.651**	-0.502**	0.147	0.240*	0.191		
TNS	F_6	0.610**	-0.056	0.116	0.025	0.016	0.953**	
	F_7	0.683**	-0.467**	0.103	0.303**	0.274*	0.981**	
100 SW	F_6	0.626**	0.655**	-0.328**	0.094	0.161	0.095	0.103
	F_7	0.127	0.714**	-0.293*	-0.287*	-0.151	-0.608**	-0.594**

*,**r-value is statistically significant at 0.05 and 0.01 probability levels, respectively, FPIH: First pod insertion height, FPIH/PH: Ratio of first pod insertion height to plant height, GD: Growth duration, GY: Soybean grain yield, PH: Plant height, TNP: Total number of pods per plant, TNS: Total number of seeds per plant; 100 SW: 100 seed weight

each cross at the F_6 and F_7 generations. Direct effects of TNS and 100 SW were significantly positive and higher than total indirect effects (e.g., 0.809 compared with -0.199; 0.441 compared with 0.184 in F_6, respectively) (Table 5). Among indirect effects, the effects of TNP via TNS were consistently positive and highest (0.771 and 1.084 in F_6 and F_7 respectively). Besides, GD, PH, FPIH and FPIH/PH also seemed to have significant indirect effects on GY via either TNS or 100 SW depending on generations (e.g., indirect effects of GD via 100 SW were significant of 0.289 and 0.512 in the F_6 and F_7).

DISCUSSION

Correlations between seed yield and yield components can vary across different generations and crossing populations and even cultivation regimes. These were observed not only in soybean[6,23] but also other crops like sesame[24], mungbean[25], cowpea[26], pigeonpea[27], cotton in the F_3 and F_4 generations[28]. In this study, although variations were observed, negative and positive correlations were consistent for most pair-wise traits at the F_6 and F_7 generations in both crosses. Majority of significant estimates of correlations between yield and yield components were positive in nature, viz over generations, 11 positive out of 14 significant correlations in LSB5, 14 positive out of 21 significant correlations in LSB10 and 15 positive out of 24 significant correlations in jointed population analysis (Table 2-4). This suggested improvement of those individually traits would simultaneously improve other pair-wise traits due to correlated responses. Especially, there were consistencies in significantly positive correlations between TNP and TNS, which were in agreement with reported data[8,10,18,20]. In other studies, TNP and TNS were also significantly positive correlated with yield regardless of environment[6]or growth habit[11]. These indicated the priority of TNP and TNS in yield improvement. In addition, the correlation between grain yield and GD was not significant, except for the joint analysis in F_6. As a result, this trait cannot be prioritized for selecting superior genotypes and selecting for early maturing genotypes without effects on yield would be possible.

In most cases, TNP and TNS were negatively correlated with 100 SW, indicating that increasing seed size was accompanied with decrease in seed number. Balla and Ibrahim[29] also reported negative associations between seed weight and both number of pods and of seeds per plant, although the degree of association varied. The FPIH is an important agronomic trait in soybean breeding for mechanical harvest[30,31]. The elevation of plant height often tends to increase the height of first pod insertion, because tall plants have a denser canopy and restrict the interception of photosynthetically active radiation in the lower parts of the plant[32]. However, FPIH was often negatively correlated with yield components[18,31]. In this study, FPIH showed non-significant associations with GY, TNP and TNS. Alternatively, FPIH/PH revealed higher and positive correlations with GY, TNP, TNS and 100 SW, indicating possibility to select plants of adequate FPIH/PH for mechanical harvest without reducing TNP, TNS and GY. Therefore, FPIH/PH could be a useful criterion in soybean breeding for

Table 5: Estimates of direct and indirect effects of yield components in F_6 and F_7 generations for each cross and for the joint analyses of two crosses GBVN004898×DT26 (LSB5) and VI045032×GBVN004904 (LSB10)

	LSB5		LSB10		Joint analysis	
Crosses	F_6	F_7	F_6	F_7	F_6	F_7
Direct effect of GD	0.079	-0.009	-0.003	-0.006	0.189*	0.084
Ind. via PH	-0.100	0.002	0.081	-0.114	0.021	0.065
Ind. via FPIH	-0.167	-0.041	-0.422	0.096	0.035	-0.023
Ind. via FPIH/PH	0.226	0.062	0.309	-0.002	-0.066	0.004
Ind. via TNP	0.077	-0.009	0.001	0.004	0.023	-0.046
Ind. via TNS	-0.131	-0.087	-0.021	0.067	-0.045	-0.516
Ind. via 100 SW	0.137	0.010	0.047	0.012	0.289*	0.512*
Total indirect effects	0.042	-0.063	-0.005	0.062	0.257*	-0.003
Direct effect of PH	0.670*	-0.006	0.432	-0.271	-0.093	-0.232
Ind. via GD	-0.012	0.003	-0.001	-0.003	-0.043	-0.023
Ind. via FPIH	-0.031	0.013	-0.365	0.132	0.018	0.023
Ind. via FPIH/PH	-0.848	-0.080	-0.005	0.129	0.044	0.087
Ind. via TNP	-0.021	0.015	0.058	-0.052	-0.031	0.014
Ind. via TNS	0.029	0.088	0.149	-0.373*	0.094	0.114
Ind. via 100 SW	-0.068	-0.109	-0.026	0.062	-0.145*	-0.210*
Total indirect effects	-0.950	-0.070	-0.191	-0.105	-0.062	0.004
Direct effect of FPIH	-1.217*	-0.196	-1.867	0.261	0.196	0.154
Ind. via GD	0.011	-0.002	-0.001	-0.002	0.033	-0.012
Ind. via PH	0.017	0.000	0.085	-0.137	-0.008	-0.035
Ind. via FPIH/PH	1.040	0.168	1.642	-0.158	-0.299	-0.224
Ind. via TNP	0.003	0.009	-0.001	-0.004	-0.001	0.022
Ind. via TNS	-0.024	0.203	0.014	0.002	0.021	0.335*
Ind. via 100 SW	0.009	-0.046	0.000	0.030	0.042	-0.206
Total indirect effects	1.056	0.331	1.738	-0.269	-0.212	-0.119
Direct effect of FPIH/PH	1.428*	0.182	1.676*	-0.292	-0.309	-0.256
Ind. via GD	0.012	-0.003	0.000	0.000	0.040*	-0.001
Ind. via PH	-0.398	0.002	-0.001	0.120	0.013	0.079
Ind. via FPIH	-0.886	-0.181	-1.830	0.142	0.190	0.135
Ind. via TNP	-0.026	0.003	-0.013	0.051	0.002	0.018
Ind. via TNS	0.045	0.163	-0.015	0.396*	0.013	0.303*
Ind. via 100 SW	0.034	-0.006	0.004	-0.029	0.071*	-0.108
Total indirect effects	-1.218	-0.021	-1.855	0.679	0.330	0.425
Direct effect of TNP	-0.258	0.073	0.238	0.127	-0.235	0.092
Ind. via GD	-0.023	0.001	0.000	0.000	-0.019	-0.042
Ind. via PH	0.054	-0.001	0.105	0.112	-0.012	-0.034
Ind. via FPIH	0.014	-0.024	0.011	-0.008	0.001	0.037
Ind. via FPIH/PH	0.145	0.008	-0.091	-0.117	*0.003*	*-0.049*
Ind. via TNS	0.468	0.729*	0.556	0.936*	0.771*	1.084*
Ind. via 100 SW	0.009	0.019	0.033	-0.121	0.042	-0.437*
Total indirect effects	0.665	0.731*	0.615	0.801*	0.786*	0.559*
Direct effect of TNS	0.491*	0.768*	0.581*	0.949*	0.809*	1.105*
Ind. via GD	-0.021	0.001	0.000	0.000	-0.011	-0.039
Ind. via PH	0.040	-0.001	0.111	0.106	-0.011	-0.024
Ind. via FPIH	0.059	-0.052	-0.045	0.001	0.005	0.047
Ind. via FPIH/PH	0.131	0.039	-0.043	-0.122	-0.005	-0.070
Ind. via TNP	-0.246	0.070	0.227	0.125	-0.223	0.091
Ind. via 100 SW	-0.021	-0.003	0.038	-0.120	0.045	-0.427*
Total indirect effects	-0.057	0.054	0.288	-0.010	-0.199	-0.422*
Direct effect of 100 SW	0.327*	0.495*	0.264*	0.330*	0.441*	0.718*
Ind. via GD	0.033	0.000	0.000	0.000	0.124*	0.060
Ind. via PH	-0.140	0.001	-0.043	-0.051	0.030	0.068
Ind. via FPIH	-0.035	0.018	0.001	0.024	0.019	-0.044
Ind. via FPIH/PH	0.150	-0.002	0.026	0.026	-0.050	0.039
Ind. via TNP	-0.007	0.003	0.030	-0.046	-0.022	-0.056
Ind. via TNS	-0.032	-0.004	0.083	-0.345	0.083	-0.656*
Total indirect effects	-0.030	0.016	0.097	-0.392	0.184	-0.591*

*Statistically significant by bootstrapping analysis, FPIH: First pod insertion height, FPIH/PH: Ratio of first pod insertion height to plant height, GD: Growth duration, GY: Soybean grain yield, PH: Plant height, TNP: Total number of pods per plant, TNS: Total number of seeds per plant, 100 SW: 100 seed weight mechanized harvesting. Kang *et al.*[31] also suggested adequate ratio of first pod height to plant height to be 0.16-0.25.

Path coefficients analysis provides a means to dissect correlation, direct and indirect effects of more than one character on yield in a particular population. Thus, path analysis provides clearer pictures of trait associations which help breeders to determine efficient selection strategy. In the present study, the direct effects were in agreement with correlations for most traits across generations, crosses and joint analyses, indicating that correlations explained the true relationships between yield and yield components (Table 2-5). While correlations priority suggested TNP and TNS, the path coefficients identified TNS and 100 SW with consistent and significant positive direct effects on grain yield across the F_6 and F_7 generations and populations (Table 5). In addition, different from previous studies, through this path analysis, TNP should only be emphasized as an indirect contributor toward grain yield. The very high positive direct effects in the F_6 and significant indirect effects of FPIH/PH on grain yield via either TNS and 100 SW could imply FPIH/PH to be considered in selection, particularly selection for mechanical harvest as in this study. The direct and indirect effects of remaining traits, such as GD, PH, FIPH were too low and non-significant to be considered important. Those traits are well-known to often suffer greater genotype×environment interaction effects so that analysis could not allow accurate inferences about their effects[6]. As a result, those traits cannot be prioritized for selecting superior genotypes.

Moreover, most direct and indirect effects of yield components were higher in the F_6 than in the F_7. Significant effects of TNS and 100 SW were also observed in the F_6 and F_7. Together with the steady pattern of correlations and agreement between correlations and direct effects across generations, crosses and joint analyses, this suggested that selection could be applied at the earlier generation for significant traits.

CONCLUSION

Present correlation and path analysis of soybean seed yield and yield components revealed consistency across generations, crosses and joint analysis for most cases. Total number of seeds and 100 seed weight should be prioritized for selecting superior soybean genotypes. Total number of pods and ratio of first pod insertion height to plant height should be considered in specific selection strategies such as breeding for mechanical harvest. Selection for yield improvement through significant yield components can also be applied in early generations.

SIGNIFICANCE STATEMENT

This study discovered the consistencies in associations of total number of pods and seeds, 100 seed weight and ratio of fist pod insertion height to plant height with yield across different cross populations or generations, suggesting breeders to prioritize those traits and decide appropriate generation for selection in soybean breeding. This study will help the researcher and breeders to uncover the use of lines developed from crosses rather than an array of established cultivars. Thus a new theory on the genetic gains of different selection strategies may be arrived at.

ACKNOWLEDGMENTS

The first author acknowledges funding by the Vietnam National University of Agriculture, Vietnam during 2016-2018.

REFERENCES

1. Vinh, M.Q., D.K. Thinh, D.T. Bang, D.H. At and L.H. Ham, 2009. Current status and research directions of induced mutation application to seed crops improvement in Vietnam. Proceedings of the International Joint FAO/IAEA Symposium on Induced Plant Mutations in the Genomics Era, August 12-15, 2008, Rome, pp: 341-345.
2. Nguyen, B.A.T., Q.T. Nguyen, X.L.T. Hoang, P.T. Nguyen and L.S.T. 2014. Evaluation of drought tolerance of the Vietnamese soybean cultivars provides potential resources for soybean production and genetic engineering. BioMed. Res. Int., Vol. 9. 10.1155/2014/809736
3. De Carvalho, C.G.P., C.A.A. Arias, J.F.F. de Toledo, M.F. de Oliveira and N.A. Vello, 2002. Correlation and path analyses in soybean lines sowed at different sowing dates. Pesq. Agropec. Bras., 37: 311-320.
4. Karasu, A., M. Oz, A.T. Goksoy and Z.M. Turan, 2009. Genotype by environment interactions, stability and heritability of seed yield and certain agronomical traits in soybean [*Glycine max* (L.) Merr.]. Afr. J. Biotechnol., 8: 580-590.
5. Hakim, L. and S. Suyamto, 2017. Gene action and heritability estimates of quantitative characters among lines derived from varietal crosses of soybean. Indones. J. Agric. Sci., 18: 25-32.
6. Ferrari, M., I.R. Carvalho, A.J. de Pelegrin, M. Nardino and V.J. Szareski *et al*., 2018. Path analysis and phenotypic correlation among yield components of soybean using environmental stratification methods. Aust. J. Crop Sci., 12: 193-202.
7. Bekele, A. and G. Alemahu, 2011. Desirable traits influencing grain yield in soybean (*Glycine max* (L.) Merrill). Innovative Syst. Des. Eng., 2: 14-23.

8. Aditya, J.P., P. Bhartiya and A. Bhartiya, 2011. Genetic variability, heritability and character association for yield and component characters in soybean (*G. max* (L.) Merrill). J. Cent. Eur. Agric., 12: 27-34.
9. Board, J.E., M.S. Kang and M.L. Bodrero, 2003. Yield components as indirect selection criteria for late-planted soybean cultivars. Agron. J., 95: 420-429.
10. Nagarajan, D., T. Kalaimagal and E. Murugan, 2015. Correlation and path coefficient analysis for yield and yield attributes in soybean, *Glycine max* L. (Merr). Int. J. Farm Sci., 5: 28-34.
11. Teodoro, P.E., L.P. Ribeiro, C.C.G. Correa, R.A.A. da Luz Jr., A.D.S. Zanuncio, D.P. Capristo and F.E. Torres, 2015. Path analysis in soybean genotypes as function of growth habit. Biosci. J., 31: 794-799.
12. Machado, B.Q.V., A.P.O. Nogueira, O.T. Hamawaki, G.F. Rezende and G.L. Jorge *et al*., 2017. Phenotypic and genotypic correlations between soybean agronomic traits and path analysis. Genet. Mol. Res., Vol. 16. 10.4238/gmr16029696.
13. Cruz, C.D., A.J. Regazzi and P.C.S. Carneiro, 2012. Biometric Models Applied to Genetic Improvement. 4th Edn., UFV., Vicosa, Brazil, (In Portuguese)..
14. Wright, S., 1921. Correlation and causation. J. Agric. Res., 20: 557-585.
15. Dewey, D.R. and K.H. Lu, 1959. A correlation and path-coefficient analysis of components of crested wheatgrass seed production. Agron. J., 51: 515-518.
16. Falconer, D.S. and T.F.C. Mackay, 1996. Introduction to Quantitative Genetics. 4th Edn., Pearson, England, ISBN-13: 978-0582243026, Pages: 480.
17. Vu, T.T.H., T.H. Nguyen and D.H. Vu, 2007. The influences of seeding seasons on the traits and their impacts on yield of soybean individuals. Vietnam J. Agric. Rural Dev., 12-13: 47-51.
18. Oz, M., A. Karasu, A.T. Goksoy and Z.M. Turan, 2009. Interrelationships of agronomical characteristics in soybean (*Glycine max*) grown in different environments. Int. J. Agric. Biol., 11: 85-88.
19. Carvalho, I.V., M. Nardino, G.H. Demari, V.J. Szareski and D.N. Follmann *et al*., 2017. Relations among phenotypic traits of soybean pods and growth habit. Afr. J. Agric. Res., 12: 450-458.
20. Sulistyo, A., Purwantoro and K.P. Sari, 2018. Correlation, path analysis and heritability estimation for agronomic traits contribute to yield on soybean. IOP Conf. Ser.: Earth Environ. Sci., Vol. 102. 10.1088/1755-1315/102/1/012034.
21. Akintunde, A., 2012. Path analysis step by step using excel. J. Tech. Sci. Technol., 1: 9-15.
22. Hayes, A.F., 2018. Introduction to Mediation, Moderation and Conditional Process Analysis: A Regression-Based Approach. 2nd Edn., The Guilford Press, New York, ISBN: 9781462534654, Pages: 692.
23. Abd El-Mohsen, A.A., G.O. Mahmoud and S.A. Safina, 2013. Agronomical evaluation of six soybean cultivars using correlation and regression analysis under different irrigation regime conditions. J. Plant Breed. Crop Sci., 5: 91-102.
24. Subramanian, S. and M. Subramanian, 1994. Correlation studies and path coefficient analysis in sesame (*Sesamum indicum* L.). J. Agron. Crop Sci., 173: 241-248.
25. Nguyen, D.T., H.T.T. Vu, L.M. Bielig and R.J. Lawn, 2016. Expression and heritability of late flowering and other quantitative traits in cultivated x Australian wild mungbean hybrids. Crop Pasture Sci., 67: 1235-1251.
26. Ribeiro, H.L.C., C.A.F. Santos, L. da Silva Diniz, L.A. do Nascimento and E.D. Nunes, 2016. Phenotypic correlations and path analysis for plant architecture traits and grain production in three generations of cowpea. Rev. Ceres, Vicosa, 63: 033-038.
27. Yerimani, A.S., S. Mehetre and M.N. Kharde, 2013. Association analysis for yield and its component traits in F_3 and F_4 populations of pigeonpea (*Cajanus cajan* (L.) Mill sp.). Mol. Plant Breed., 4: 189-195.
28. Tonk, F.A., D. Istipliler, M. Tosun, E. Ilker and C. Gizem, 2018. Correlation and path analysis in yield and quality traits in F_3 and F_4 generation of Carmen × Devetuyu-176. J. Anim. Plant Sci., 28: 107-112.
29. Balla, M.Y. and S.E. Ibrahim, 2017. Genotypic correlation and path coefficient analysis of soybean [*Glycine max* (L.) Merr.] for yield and its components. Agric. Res. Technol., Vol. 7.
30. Ramteke, R., D.H. Singh and P. Murlidharan, 2012. Selecting soybean (*Glycine max*) genotypes for insertion height of the lowest pod, the useful trait for combine harvester. Indian J. Agric. Sci., 82: 511-515.
31. Kang, B.K., H.T. Kim, M.S. Choi, S.C. Koo and J.H. Seo *et al*., 2017. Genetic and environmental variation of first pod height in soybean [*Glycine max* (L.) Merr.]. Plant Breed. Biotechnol., 5: 36-44.
32. Muller, M., M. Rakocevic, A. Caverzan and G. Chavarria, 2017. Grain yield differences of soybean cultivars due to solar radiation interception. Am. J. Plant Sci., 8: 2795-2810.