# Genetic Variability and Association Analysis of Soybean (*Glycine max* (L.) Merrill) for Yield and Yield Attributing Traits

Soleh Akram[1], B. M. Nahid Hussain[1], Md. Abdullah Al Bari[1], David J. Burritt[2], Mohammad Anwar Hossain[1]✉

1 Department of Genetics and Plant Breeding, Bangladesh Agricultural University, Mymensingh-2202, Bangladesh

2 Department of Botany, University of Otago, New Zealand

✉ Corresponding author: anwargpb@bau.edu.bd

**Abstract** An experiment was conducted using a randomized complete block design to estimate genetic variability, and to analyse correlation and path coefficient between yield and its components of eleven soybean genotypes. Analysis of variance for yield and yield contributing traits showed significant ($p<0.01$) variation among the genotypes. Results of genetic analyses showed a higher phenotypic coefficient of variation compared to their corresponding genotypic coefficient of variation for all the traits measured, which indicates that the traits were influenced by environment to some extent. High heritability values along with high genetic advance and genetic advance as percentage of mean were found for plant height, yield/plant, number of pods/plant and number of seeds/plant, indicated the scope of improvement for these characters. Yield/plant showed a significant positive correlation with the number of branches/plant (0.851), number of pods/plant (0.988), number of seeds/plant (0.988) and 100-seed weight (0.634) whereas it showed a non-significant positive correlation with pod length. In contrast, yield/plant showed a negative correlation with days to first flowering, days to 50% flowering, days to maturity, plant height and number of seeds/pod. Path coefficient analysis showed that the seeds/plant had the highest (2.848) positive direct effect on yield/plant, followed by pods/plant, days to first flowering, number of branches/plant, 100-seed weight, plant height, and pod length. So, the number of pods/plant and seeds/plant could be considered as important selection criteria for yield improvement in soybean.

**Keywords** Genetic Variability; Heritability; Genetic Advance; Correlation; Path coefficient; Yield; Soybean

## Introduction

Soybean (*Glycine max* (L.) Merrill) is the most widely grown leguminous crop in the world and is an important source of protein and oil for food and feed (Sharma et al., 2013; Hossain and Komatsu, 2014). Soybean seeds are rich in proteins, unsaturated fatty acids, minerals (e.g. Ca and P) and vitamins (e.g. A, B, C and D), which meet the nutritional needs of humans and other animals (He and Chen, 2013; Malek et al., 2014; Ghosh et al., 2014). Soybean seed is referred to the "protein hope" of the future, because of its high nutritive value, containing about 42-45% protein (Sathe et al., 2009). As a significant source of protein, soybean seeds are often referred to as "the meat that grows on plant" (Henkel, 2000). Being a legume, soybean plants also fix atmospheric nitrogen, which then becomes available for other plants. In addition, soybeans contain numerous compounds that can act as antioxidants and are beneficial to human health, as they diminish the risk of many diseases (Kumar et al., 2014).

Yield and yield contributing parameters are the most widely targeted traits for soybean improvement programmes worldwide. The improvement of a crop is largely dependent on the nature and magnitude of available genetic variability, heritability and the transfer of desired characters into new varieties. The success of breeding programmes can be enhanced when variability within the existing germplasm is high, which allows the plant breeder to more rapidly produce new varieties or improve existing ones (Meena and Bahadur, 2013, 2014; Ranganatha et al., 2013; Yared and Misteru, 2016). Hence, knowledge of key genetic parameters is crucial for any crop improvement program, providing precise information for selection. Genetic parameters like the genotypic coefficient of variation (GCV), phenotypic coefficient of variation (PCV), heritability and genetic advance (GA) are useful biometric tools for measuring genetic variability (Aditya et al., 2011). Hence, characterising the genetic

background of soybean and determining breeding values should be done before carrying out any improvement programme (Agong et al., 2001).

Yield is a complex polygenic character, resulting from multiple interactions between many yield-contributing traits. Associations between these traits can be evaluated by correlation analysis, which helps in the simultaneous selection for more than one character (Akhtar et al., 2010; Sathya and Jebaraj, 2013; Semahegn and Tesfaye, 2016). The degree of relatedness between important plant traits is an index that can be used to predict yield responses in relation to changes associated with a particular character (Malek et al., 2014). So, for soybean, identification of key characters associated with yield and other contributing components is important for maximizing yields (Aditya et al., 2011; Meitei et al., 2014; Jain et al., 2015). Correlation coefficients alone are inadequate to interpret cause and effect relationships among traits associated with yield, whereas path coefficient analysis allows a better understanding of associations between different characters, by breaking down correlation coefficients associated with the main character into direct and indirect effects (Rahman et al., 2011, 2012; Hossain et al., 2015). Path coefficient analysis provides an effective means to partition correlation coefficients into unidirectional and alternative pathways, thus permitting a critical examination of the specific factors that produce a given correlation; this can then be employed to formulate an effective selection programme (Salahuddin et al., 2010; Lal et al., 2011; Parmar et al., 2013; Jain et al., 2015). Considering the above facts, the present research study was undertaken to estimate genetic variability, heritability and genetic advances, and the associations among yield and yield contributing traits for soybean, using correlation and path analysis. This investigation provides information that could lead to the development of desirable genotypes in future breeding programmes.

# 1 Results and Analysis

## 1.1 Estimation of genetic variability, heritability and genetic advances

The results of the analysis of variance (mean sum of squares) revealed that genotypes differed significantly ($p<0.01$) for all the traits studied, days to first flowering, days to 50% flowering, days to maturity, plant height, number of branches/plant, pod length, number of pods/plant, number of seeds/pod, number of seeds/plant, 100-seed weight and yield/plant. The mean sum of squares, genotypic variance (GV,$\delta^2$g), phenotypic variance (PV,$\delta^2$p), GCV, PCV, heritability ($h^2_b$), GA and genetic advance as percentage of the means (GA %) for yield and yield attributing traits of soybean are presented in Table 1. The values for phenotypic variance for all of the traits studied were higher than those of their corresponding genotypic variance. The highest phenotypic variance was found for number of seeds/plant (421.99) and the lowest for pod length (0.0004). As with the phenotypic variance, the highest genotypic variance was found for the number of seeds/plant (390.85) and the lowest value for pod length (0.0002). The highest PCV and GCV values were recorded for yield/plant (22.62 and 22.07%, respectively), whereas the lowest PCV and GCV values were recorded for pod length (0.54 and 0.38%, respectively). Higher heritability was observed for most of the traits studied however, the highest was found for days to maturity (99.73%) and the lowest for the number of seeds/pod (44.84%). The highest GA was observed for the number of seeds/plant (39.19) and the lowest for pod length (0.02). GA% was the highest for yield/plant (44.38%) and the lowest for pod length (0.56%).

## 1.2 Estimation of the correlation coefficient

The correlation coefficients for all of the eleven characters measured are presented in Table 2. In the present study 13 associations were significant out of 55. Among them 9 associations were found positive and highly significant (*p<0.01*) and two associations were found to be significant at 5% level of probability. Two associations were found to be negative and significantly associated; one was highly significant (*p<0.01*) and the other was significant at 5% level of probability.

Yield/plant showed significant positive correlations with the number of branches/plant, pod length, number of pods/plant, number of seeds/plant and 100-seed weight. In contrast, it showed non-significant negative correlations with days to first flowering, days to 50% flowering, days to maturity, plant height and the number of seeds/pod. Days to first flowering showed significant (*p<0.01*) positive correlations with days to 50% flowering

and days to maturity, where as it showed a non-significant negative correlation with the number of branches/plant, pod length, number of pods/plant, number of seeds/plant and 100-seed weight. Days to 50% flowering showed a significant ($p<0.01$) positive correlation with days to maturity. Days to maturity showed significant ($p<0.01$) positive correlations with days to first flowering and days to 50% flowering. All of the traits showed negative correlations with plant height, but branches/plant and pod length showed significant negative correlation. The number of branches/plant showed significant positive correlations with pod length, number of pods/plant, number of seeds/plant, whereas it showed a non-significant negative correlation with number of seeds/pod. Pod length showed significant ($p<0.05$) positive correlations with branches/plant whereas it showed a significant ($p<0.05$) negative correlation with plant height. The number of pods/plant showed significant positive correlations with branches/plant and seed yield/plant, but non-significant negative correlations with days to first flowering, days to 50% flowering, days to maturity, plant height. None of the traits showed significant positive or negative correlations with the number of seeds/pod. The number of seeds/plant showed significant positive correlations with yield/plant, number of branches/per plant, number of pods/plant, but non-significant negative correlations with days to first flowering, days to 50% flowering, days to maturity, plant height and number of seeds/pod. 100-seed weight showed significant ($p<0.05$) positive correlations with yield/ plant, but non-significant negative correlations with days to first flowering, days to 50% flowering, plant height and number of seeds/pod.

Table 1Component of variance, coefficient of variation, heritability in broad sense and genetic advance for morphological characters studied

| Traits | MS | PV ($\delta^2$p) | GV ($\delta^2$g) | PCV (%) | GCV (%) | $h^2_b$ (%) | GA | GA (%) |
|---|---|---|---|---|---|---|---|---|
| DFF | 84.394** | 9.44 | 9.38 | 6.67 | 6.65 | 99.42 | 6.29 | 13.67 |
| DFF2 | 90.304** | 10.08 | 10.03 | 6.48 | 6.47 | 99.51 | 6.51 | 13.29 |
| DM | 152.64** | 16.96 | 16.91 | 3.71 | 3.71 | 99.73 | 8.46 | 7.62 |
| PH | 2038.25** | 226.46 | 222.20 | 21.01 | 20.82 | 98.12 | 30.42 | 42.48 |
| BR/Pl | 0.410** | 0.0487 | 0.0403 | 6.21 | 5.65 | 82.88 | 0.38 | 10.59 |
| PDL (cm) | 0.005** | 0.0004 | 0.0002 | 0.54 | 0.38 | 50.00 | 0.02 | 0.56 |
| PD/Pl | 1033.35** | 114.82 | 107.22 | 21.98 | 21.24 | 93.38 | 20.61 | 42.28 |
| SD/Pd | 0.015** | 0.0032 | 0.0014 | 2.93 | 1.96 | 44.84 | 0.05 | 2.71 |
| SD/Pl | 3797.97** | 421.99 | 390.85 | 21.71 | 20.90 | 92.62 | 39.19 | 41.43 |
| Y/Pl (g) | 1.562** | 6.21 | 5.91 | 22.62 | 22.07 | 95.25 | 4.89 | 44.38 |
| 100-W (g) | 55.873** | 0.174 | 0.128 | 3.52 | 3.02 | 73.69 | 0.63 | 5.34 |

Note:** indicates significant differences at $p< 0.01$. MS, Mean sum of square; PCV, Phenotypic coefficient of variation; GCV, Genotypic coefficient of variation; PV ($\delta^2$p), Phenotypic variance; GV ($\delta^2$g), Genotypic variance; GA, Genetic advance; GA (%), Percentage genetic advance; $h^2_b$, Heritability in broad sense. DFF, Days to first flowering; DFF2, Days to 50% flowering; PH, plant height; BR/Pl, number of branches per plant; PDL, Pod length; PD/Pl, number of pods per plant; SD/PD, number of seeds per pod; SD/Pl, number of seeds per plant; 100-W, 100-seed weight; Y/Pl, yield per plant.

## 1.3 Estimation of path coefficients

The path coefficient analysis was performed using correlation coefficients to determine the direct and indirect effects of 11 yield contributing characters. The values are shown in Table 3 and the analysis was performed as correlation coefficient estimates were not adequate to measure the cause and effect of dependent and independent variables.

The analysis showed that the number of seeds/plant (2.848) had a maximum positive and direct effect on yield/plant, followed by the number of pods/plant (1.973). Other important parameters that influenced yield directly and in a positive direction were days to first flowering (0.900), number of branches/plant (0.0728), 100-seed weight (0.0455), and pod length (0.0049). It is likely that these characters dominantly contributed to the yield of soybean plants. In contrast, days to 50% flowering (-0.995), days to maturity (-0.125) and number of seeds/pod (-0.277) had negative direct effects on yield/plant.

The highest positive indirect effects on yield/plant were the number of pods/plant (2.834), followed by the number of branches/plant (2.387), the number of seeds/plant (1.963), 100-seed weight (1.452), pod length (1.225), and days to 50% flowering (0.897). In contrast, days to maturity (-1.581) exhibited the greatest negative indirect effect on yield/plant, followed by days to first flowering (-0.991), pod length (-0.772), and plant height (-0.749). The number of branches/plant (0.851**), the number of pods/plant (0.988**), the number of seeds/plant (0.988**) and 100-seed weight (0.634*), were all positively and significantly correlated with yield/plant (Table 3). Days to first flowering (-0.112) and days to 50% flowering (-0.143), plant height (-0.061), days to maturity (-0.551) and the number of seeds/pod (0.209), were all negatively correlated with yield (Table 3), however, none of them were significant.

The residual effect determines how best the causal factors account for the variability of the dependent variables. The residual effect was 0.00107 indicating that 11 characters contributed 99.99% of the variability in yield per plant, as determined by path analysis. The traits included in this study were adequate and appropriate as they account for almost all yield-associated variation.

Table 2 Phenotypic correlation among yield and yield contributing traits of soybean genotypes

| Traits | DFF | DFF2 | DM | PH (cm) | BR/Pl | PDL (cm) | PD/Pl | SD/PD | SD/Pl | 100-W (g) |
|---|---|---|---|---|---|---|---|---|---|---|
| YPL | -0.112 | -0.143 | -0.061 | -0.551 | 0.851** | 0.438 | 0.988** | -0.209 | 0.988** | 0.634* |
| DFF | | 0.995** | 0.737** | 0.420 | -0.131 | -0.281 | -0.067 | 0.063 | -0.045 | -0.284 |
| DFF2 | | | 0.753** | 0.402 | -0.159 | -0.277 | -0.091 | 0.043 | -0.073 | -0.308 |
| DM | | | | 0.319 | -0.168 | 0.009 | -0.074 | 0.109 | -0.046 | 0.023 |
| PH | | | | | -0.774** | -0.624* | -0.556 | -0.026 | -0.555 | -0.262 |
| BR/Pl | | | | | | 0.592* | 0.831** | -0.079 | 0.838** | 0.548 |
| PDL | | | | | | | 0.391 | 0.340 | 0.429 | 0.339 |
| PD/Pl | | | | | | | | -0.239 | 0.995** | 0.541 |
| SD/PD | | | | | | | | | -0.145 | -0.499 |
| SD/PL | | | | | | | | | | 0.510 |

Note: * and ** indicates significant at 5% and 1% level of probability, respectively. DFF, Days to first flowering; DFF2, Days to 50% flowering; PH, plant height; BR/Pl, number of branches per plant; PDL, Pod length; PD/Pl, number of pods per plant; SD/PD, number of seeds per pod; SD/Pl, number of seeds per plant; 100-W, 100-seed weight; Y/Pl, yield per plant.

Table 3 Partitioning of phenotypic correlation into direct and indirect effects by path coefficient analysis (bold values indicate the direct effect)

| Characters | YPL (g) | DFF | DFF2 | PH(cm) | DM | BR/Pl | PDL (cm) | PD/Pl | SD/PD | SD/Pl | 100-W (g) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| DFF | -0.112 | **0.900** | -0.991 | 0.0495 | -0.052 | -0.0095 | -0.0014 | -0.1322 | -0.0174 | -0.128 | -0.0120 |
| DFF2 | -0.143 | 0.897 | **-0.995** | 0.0505 | -0.0502 | -0.0116 | -0.00135 | -0.1796 | -0.0119 | -0.2079 | -0.0140 |
| PH(cm) | -0.061 | 0.664 | -0.749 | **0.067** | -0.0398 | -0.0122 | 0.00004 | -0.1461 | -0.0302 | -0.1310 | 0.0010 |
| DM | -0.551 | 0.378 | -0.4001 | 0.0214 | **-0.125** | 0.0563 | -0.0031 | 1.097 | 0.0072 | -1.581 | -0.0119 |
| BR/Pl | 0.851** | -0.117 | 0.1592 | -0.0112 | 0.0966 | **0.0728** | 0.0029 | 1.640 | 0.0218 | 2.387 | 0.0249 |
| PDL (cm) | 0.438 | -0.253 | 0.2756 | 0.0006 | 0.0779 | -0.0431 | **0.0049** | -0.772 | -0.094 | 1.225 | 0.0155 |
| PD/Pl | 0.988** | -0.0603 | 0.0905 | -0.0049 | 0.0694 | 0.0605 | 0.0019 | **1.973** | 0.0661 | 2.834 | 0.0246 |
| SD/PD | -0.209 | 0.0567 | -0.0427 | 0.00732 | 0.0032 | 0.00575 | 0.00166 | 0.472 | **-0.277** | -0.4131 | -0.0227 |
| SD/Pl | 0.988** | -0.0405 | 0.0726 | -0.0031 | 0.0692 | 0.0610 | 0.0021 | 1.963 | 0.0401 | **2.848** | 0.0232 |
| 100-W (g) | 0.634* | -0.2377 | 0.3065 | 0.00154 | 0.0327 | -0.0399 | 0.00166 | 1.067 | 0.1381 | 1.452 | **0.0455** |
| | | | | | | | | | | Residual effect = 0.00107 | |

Note: * and ** indicates significant at 5% and 1% probability level, respectively; DFF, Days to first flowering; DFF2, Days to 50% flowering; PH, plant height; BR/Pl, number of branches per plant; PDL, Pod length; PD/Pl, number of pods per plant; SD/PD, number of seeds per pod; SD/Pl, number of seeds per plant; 100-W, 100-seed weight; Y/Pl, yield per plant.

## 2 Discussion

Breeding plants with agronomically and economically superior traits is the ultimate goal of plant breeders. In the present study, an attempt was made to estimate the proportion of variability, heritability, genotypic and phenotypic variances, predicted genetic advance, GCV and PCV, with respect to yield and yield contributing traits

of soybean genotypes. The associations among the traits were also studied by correlation coefficient analysis and their direct and indirect effects on seed yield were estimated by path coefficient analysis.

## 2.1 Estimation of genetic variability, heritability and genetic advance

The success of crop breeding programs largely depends on the presence of genetic variation and the inheritance of traits of interest. Analysis of genetic variation assists the breeder to decide the proper strategy and selection criteria to be used for the improvement of the target traits. Significant differences ($p<0.01$) for different traits among plant genotypes indicated large phenotypic variability and as expected there were inherent genetic differences among the genotypes used in the present study (Table 1). The estimates of GCV reflect the total amount of genotypic variability. However, the study of PCV and GCV is not only useful for comparing the relative amounts of phenotypic and genotypic variation among traits, but is also very useful to estimate the scope for improvement by selection.

Present study showed high $\delta^2p$ and $\delta^2g$ for seeds/plant, pods/plant and plant height. Basavaraj et al. (2015) reported that high $\delta^2p$ indicated a strong influence of environmental factors for these characters. High $\delta^2g$ was also observed for the above characters, indicating a greater contribution of genetic component for the total variation (Basavaraj et al., 2015). The GCV and PCV indicate the presence of considerable variation for these characters, which would allow further improvement by selection of individual traits (Dhanwani et al., 2013; Reni and Rao, 2013; Baraskar et al., 2014). The estimates of phenotypic and genotypic coefficients of variation indicate that the values of PCV were higher than those of GCV, but the difference was not too large. This result could be due to the fact that the traits measured in this study were not greatly influenced by environmental factors. Thus selection based on phenotypic performance of these characters would be an effective way to bring about considerable improvement of these characters. Similar findings have also been reported in other studies (Karnwal and Singh, 2009; Aditya et al., 2011; Dilnesaw et al., 2013; Baraskar et al., 2014; Jain et al., 2015).

Heritability estimates help breeders to select plants based on phenotypic performance and can be used to predict the reliability of a phenotypic value. Therefore, a high heritability helps to effectively select a particular trait (Dhanwani et al., 2013; Islam et al., 2015). The present study showed high heritability for most of the traits measured, which indicates that a large portion of the total variation is under genetic control, and that selection based on phenotypic levels would be useful for the improvement of these traits. High heritability of various yield-contributing traits has been reported in other studies (Karnwal and Singh, 2009; Reni and Rao 2013; Dilnesaw et al., 2013; Kim et al., 2010; Mahbubet al., 2015; Sureshrao et al., 2014; Osekita and Olorunfemi, 2014; Jain et al., 2015; Chandel et al., 2013; Ghodrati, 2013). Heritability estimates appear to be more meaningful when accompanied by estimates of GA and GA% (Shashikanth et al., 2010; Basavaraj et al., 2015). Among the important yield contributing traits for soybean, seeds/plant, pods/plant and plant height showed high heritability values with high GA and GA%, which indicated a high degree of genetic variability for these characters, *i.e.* there is scope to select good genotypes. This result also indicates that mostly additive genetic effects govern these characters and thus selection pressure could profitably be applied for these characters in a soybean breeding programme. Burton et al. (1952) suggested that rapid progress in selection can be achieved when a high genetic advance, which forms the most reliable index for selection, accompanies high heritability. Similar results to those presented in the present study, high heritability with high genetic advance as a percentage of the mean, have been reported for various traits in other studies (Malek et al., 2014; Baruah et al., 2014; Chandel et al., 2013; Dilnesaw et al., 2013; Baraskar et al., 2014).

The genetic parameter analysis also showed high heritability, but very low genetic advance for seed yield/plant, 100-seed weight, number of branches/plant, days to first flowering, days to 50% flowering and days to maturity, which are indicative of non-additive genetic effects (Malek et al., 2014). The above information indicates that there is limited scope for the improvement of these characters through selection, but hybridization followed by progeny selection could be effective (Meitei et al., 2014; Meena and Bahadur, 2014). Iqbal et al. (2003) reported that for characters where low heritability and genetic advance were observed, there is a need to build a broad

germplasm base through collection and acquisition. Baraskar et al. (2014) also reported high heritability and low genetic advance for pod clusters per plant, while Mahbub et al. (2015) observed low genetic advance for branches/plant, plant height, number of seeds/pod, 100 seed weight and seed yield/plant. Heritability and GA together with GCV could provide the best indication of the amount of advancement that could be expected through phenotypic selection (Malek et al., 2014). So, high values of heritability and GA%, along with high GCV for characters like yield/plant, seeds/plant, pods/plant, plant height can be considered to be favourable morphological traits for soybean improvement. In addition, effective phenotypic selection of these traits and high genetic gain from selection for these characters can be achieved. These results agree with previous studies (Ghodrati, 2013; Jain and Ramgiry, 2000; Malik et al., 2006; Karasu et al., 2009; Nassar, 2013, Mahbub et al., 2015). The results of the present study on variability, heritability and genetic advance indicated scope for improvement of soybean grain yield through selection, using parameters like the genetic coefficient of variation, heritability and GA and are of great importance when developing an efficient breeding program for Soybean, because when there is sufficient genetic variation, breeders can exploit additive gene effects, transgressive segregation, and heterosis, to improve yield.

## 2.2 Estimation of the correlation coefficient

The correlation coefficient analysis showed that yield/plant was positively and significantly correlated with the number of branches/plant, pod length, number of pods/plant, number of seeds/plant and 100 seed weight, indicating that as branch number increased, pod number and seed number also increased, which contributed to higher yield/plant. Large pods accommodate more seeds/pod and increased 100-seed weight results in increased yield/plant, which ultimately contributes to higher overall yields. Therefore pod size contributes significantly to final yield and should be considered during selection, as it could lead to a simultaneous improvement in soybean yields (Udensi and Ikpeme, 2012; Islam et al., 2015). Mahbub et al. (2015) and Aditya et al. (2011) reported significant and positive associations between yield related characters and suggested that with an increase in the value of one trait, the value of another trait will also increase. Similar results have also reported in previous studies on soybean (Arshad et al., 2006; Khan et al., 2000; Iqbal et al., 2003, 2010; Arshad et al., 2014; Mohsen et al., 2013; Jain et al., 2015). In addition, significant and positive correlations among yield and other yield contributing traits have been described for soybean, e.g. yield and days to 50% flowering, number of pods/plant, days to maturity, plant height and yield/plant (Sirohi et al., 2007); yield and number of clusters/plant, number of pods/plant, biological yield/plant and harvest index (Chandel et al., 2014), yield and pods/plant and branches/plant (Turkec, 2005).

Significant correlations were found for days to first flowering with days to 50% flowering; days to maturity with days to first flowering and 50% flowering, which showed that early flowering genotypes matured early; number of branches/plant with number of pods/plant and number of seeds/plant, which indicated that more branched genotypes accommodated more pods/plant and significantly increased seeds/plant. In contrast, a significant negative correlation was found for plant height, with the number of branches/plant and pod length, which showed that increased plant height could reduce both branch number and pod length. This result are consistent with those of Ramteke et al. (2010) who found a significant negative correlation of yield with days to flowering and maturity, and negative correlations for seed weight with days to flowering, maturity, plant height and number of nodes per plant. Arshad et al. (2006) also reported significant and negative correlation between yield and days to maturity.

Apart from the significant positive correlations, the present study also showed non-significant positive correlations (Table 3) for 16 associations. The observed positive correlations among various traits have been observed in other studies e.g., pods/plant and branches/plant by Turkec (2005); number of pods/plant and number of seeds/pod by Malik et al. (2006); pod height, days to flowering, days to maturity, plant height and number of branches/plant by Malik et al. (2007); plant height, pod yield, pods/plant by Sarutayophat (2012); plant height, number of pods/plant and number of seeds/ pod by Ali et al. (2013); pod height, plant height, branches/plant, pods/plant and 100 seed weight by Anwar and Malik (2013); 100 seed weight by Arshad et al. (2014); number of

branches/plant, number of pods/plant and 1000-seed weight by Mohsen et al. (2013). In contrast, yield/plant was negatively correlated with days to first flowering, days to 50% flowering, days to maturity, plant height and number of seeds/pod. Negative correlations among the characters like days to first flowering, days to 50% flowering and days to maturity indicate that any delay with respect to these characters could result in decreased yields (Malik et al., 2007). Negative correlations with yield and yield attributing traits for soybean have also been reported in other studies (Iqbal et al., 2010; Ali et al., 2013). Henry and Krishna (1990) and Akinyele and Osekita (2006) reported that for characters having negative correlations with one another it would be difficult to select for desirable traits. In addition, 26 associations were found to be negative and non-significant, with these associations involving complex linked relationships between various combinations of character pairs. The negative associations of these characters could be problematic with respect to combining these traits to produce a single genotype with increased yields. Suitable recombination might be obtained through bi-parental mating, mutation breeding or diallel selective mating, by breaking undesirable linkages, as suggested by Ghafoor et al. (1990). Alternatively, characters with negative associations and non-significant correlations could be disregarded when selecting for crop improvement (Henry and Krishna, 1990; Akinyele and Osekita, 2006).

Overall, the results of the present study indicate that selection of high yielding soybean genotypes would be possible by carefully balancing the number of branches/plant, pod length, number of pods/plant, number of seeds/plant and 100-seed weight, with moderate plant heights, and days to flowering and maturity. These results should be considered when determining the selection criteria for future varietal improvement of soybean. Any traits, which do not show any significant association or very negligible association, can be discarded to reduce the number of traits considered.

### 2.3 Estimation of path coefficients by partitioning phenotypic correlations into direct and indirect effects

Partitioning path correlation coefficients into direct and indirect effects provides information on the actual contribution of independent variables with respect to a dependent variable. The present study showed a maximum direct effect on yield for the number of seeds/plant, which influences yield/plant directly in a positive direction, followed by number of pods/plant, number of branches/plant, 100-seed weight and pod length (Table 3). These characters have direct positive effects on yield/plant, indicating that these are the main contributors to yield for soybean plants. Therefore, during phenotypic selection the main emphasis should be given to these traits for producing high yielding soybean genotypes. These results are in agreement with other studies (Arshad et al., 2006; Malik et al., 2006; Jain et al., 2015).

Days to first flowering, days to 50% flowering, plant height, days to maturity and number of seeds/pod were negatively correlated with yield/plant. Again, days to 50% flowering, days to maturity and number of seeds/ pod had direct negative effects on yield. This result demonstrates that for soybean selection on the basis of these traits might lead to a yield compromise (Malik et al., 2007). Sharma et al. (1983) reported that days to maturity and days to flowering contributed mostly to the seed yield. Differences between studies might be due to the influence of environmental factors (Malik et al., 2007). Plant height and days to first flowering were negatively correlated with yield/plant, but the direct effect on yield/plant was positive. This positive effect could be due to the fact that any positive indirect effects nullified any direct negative effects that plant height and days to first flowering might have on yield. The residual effect on seed yield/plant was low and this indicates that the traits under study could be used to determine any effects on seed yield (Chandel et al., 2014). The above information suggests that a highly significant positive correlation, with the highest positive direct effect were observed for the number of seeds/plant followed by the number of pods/plant. Therefore, the number of seeds/plant and pods/plant can be considered as critical criteria for improving yield. Similar conclusions were found in other studies (Iqbal et al., 2003; Chettri et al., 2003; Malik et al., 2006; Jain et al., 2015).

In conclusion, yield in soybeans is a complex polygenic character, so direct selection based on yield might not give positive results without giving due consideration to genetic background. Importantly, correlation analysis can help when examining selection criteria for improving yield through indirect selection of its component traits,

which are highly correlated. Considering the inter-relationships studied and path analysis of the various component characters with seed yield and among themselves, it is clear that seeds/plant, pods/plant, 100-seed weight, pod length and number of branches/plant are important traits to be considered when breeding to improve the yield of soybean.

## 3 Materials and Methods

### 3.1 Plant materials and experimental design

The experiment was conducted in the net house of the Department of Genetics and Plant Breeding, Bangladesh Agricultural University, Mymensingh during the period from June to November 2015. Eleven soybean genotypes including seven cultivars *viz.,* BINA Soybean-1, BINA Soybean-2, BINA Soybean-3, BINA Soybean-4, BARI Soybean-5, BARI Soybean-6, SBM-9, SBM-15, SBM-18, SBM-22 and Sohag were used as plant material. The seedlings were grown in earthen pot filled with field soil and five seedlings were finally kept in each pot. The experiment was carried out in a Randomized Complete Block Design (RCBD) with three replicates. Fertilizer and irrigation was applied following standard methods and techniques maintaining recommended doses. Other intercultural operations were done whenever necessary.

### 3.2 Data collection on different growth parameters and yield attributes

Data on days to first flowering, days to 50% flowering, days to maturity, plant height (cm), number of branches/plant, pod length (cm), number of pods/plant, number of seeds/pod, number of seeds/plant, 100 seed weight (g) and yield/plant (g) were recorded. Five plants from each replication were picked for collecting data.

### 3.3 Statistical analysis and estimation of genetic parameters

The recorded data for different parameters were assembled and organized properly for statistical analysis using SAS software version 9.3 (SAS Institute Inc. 2010) following RCBD design in three replications. Genotypic and phenotypic variances were estimated according to the formula given by Johnson et al. (1955). Heritability in broad sense ($h^2_b$) was estimated according to the formula suggested by Johnson et al. (1955) and Hanson (1961). Genotypic and phenotypic coefficient of variations was estimated according to Burton (1952) and Singh et al. (1997). Estimation of GA was done following formula given by Johnson et al. (1955) and Allard (1975). GA% was calculated by the formula of Comstock et al. (1952). The phenotypic correlations were estimated by the formula suggested by Miller et al. (1958). Correlation coefficients were further partitioned into components of direct and indirect effects by path coefficient analysis, as developed by Wright (1921) and later described by Dewey and Lu (1959).

### Authors' contribution

Soleh Akram and Mohammad Anwar Hossain conceived the idea. Soleh Akram, Mohammad Anwar Hossain and Md. Abdullah Al Bari participated in designing and executing the experiment. Soleh Akram and B. M. Nahid Hussain managed the trial and recorded the data. Soleh Akram and Md. Abdullah Al Bari analysed the data. Soleh Akram prepared the manuscript and Mohammad Anwar Hossain, Md. Abdullah Al Bari and David J. Burritt critically assessed and evaluated the results, constructed the manuscript, and improved each version as required. All authors read and approved the final manuscript.

### References

Aditya J.P., Pushpendra B.P., and Anuradha B., 2011, Genetic variability, heritability and character association for yield and component characters in soybean (*G. max* (L.) Merrill), Journal of Central European Agriculture, 12(1): 27-34
https://doi.org/10.5513/JCEA01/12.1.877

Agong S.G., Schittenhelm S., and Fried W., 2001, Genotypic variation of Kenyan tomato (*Lycopersicon esculentum* Mill.) germplasm, Journal of Food Technology in Africa, 6(1): 13-17
https://doi.org/10.4314/jfta.v6i1.19277

Akinyele B.O., and Osekita O.S., 2006, Correlation and path coefficient analyses of seed yield attributes in okra [*Abelmoschus esculentus* (L.) Moench], African Journal of Biotechnology, 5(14): 1330-1338

Akhter T., Ivy N.A., Rasul M.G., and Mian M.A.K., 2010, Variability and character association of reproductive traits in exotic rice germplasm, Bangladesh Journal of Plant Breeding and Genetics, 23(1): 39-43

Ali A., Iqbal Z., Safdar M.E., Ashraf M., Aziz M., Asif M., Mubeen M., Noorka I.R., and Rehman A., 2013, Comparison of yield performance of soybean varieties under semi-arid condition, Journal of Animal and Plant Science, 23: 828-832

Allard R.W., ed., 1975, Principles of plant breeding, John Wiley and Sons Incorporation, New York

Anwar M.F., and Malik., 2013, Evaluation of genetic diversity in soybean (*Glycine max* L.) genotypes based on agronomic and biochemical traits, Pakistan Research Repository, 188

Arshad M., Ali N., and Ghafoor A., 2006, Correlation and path analysis of soybean (*Glycine max* L.), Pakistan Journal of Botany, 38(1): 121-130

Arshad W., Zeeshan M., Khan M.I., Ali S., Hussain M., and Rahman S.U., 2014, Character association and causal effect analysis for yield and yield components among early maturing genotypes of soybean (*Glycine max* (L.) Merrill), Journal of Renewable Agriculture, 2: 1-4
https://doi.org/10.12966/jra.03.01.2014

Baraskar V.V., Kachhadia V.H., Vachhanl J.H., Barad H.R., Patel M.B., and Darwankar M.S., 2014, Genetic variability, heritability and genetic advance in soybean [*Glycine max* (L.) Merrill], Electronic Journal of Plant Breeding, 5: 802-806

Baruah S., Sarma M.K., Baishya D., Sharma A.A., Borah R., and Bhuyan J., 2014, Genetic variation for seed yield and yellow mosaic virus resistance in soybean [*Glycine max* (L.) Mill.], International Journal of Scientific and Research Publications, 4: 1-10

Basavaraj B.L., Vilas G.D., and Vijayakumar R., 2015, Study on genetic variability and character inter-relation of quality and yield components in tomato (*Solanum lycopersicum* L.), Hort. Flora Research Spectrum, 4(2): 108-115

Burton G.W., 1952, Quantitative inheritance in grasses, Proc. 6th International Grassland Cong., 1: 277-283

Chandel K.K., Patel N.B., and Patel J.B., 2013, Genetic variability analysis of soybean (*Glycine max* (L.) Merrill), AGRES-An International e-Journal, 2: 318-325

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

Chettri M., Mondol S., and Nath R., 2003, Studies on correlation science bulletin and path analysis in soybean [*Glycine max* (L.) Merrill] in the Darjeeling hills, Journal of Hill Research, 16: 101-103

Comstock R., Robinson H., and Gowen J., 1952, Estimation of average dominance of genes, Heterosis, 494-516

Dewey D.R., and Lu K.H., 1959, A correlation and path-coefficient analysis of components of crested wheat grass seed production, Agronomy Journal, 51: 515-518
https://doi.org/10.2134/agronj1959.00021962005100090002x

Dhanwani R.K.,Sarawgi A.K., Solanki A., and Tiwari J.K., 2013, Genetic variability analysis for various yield attributing and quality traits in rice (*O. sativa* L.), The Bioscan: An International Quarterly Journal of Life Science, 8(4): 1403-1407

Dilnesaw Z., Abady S., and Getahun A., 2013, Genetic variability and heritability of soybean [*Glycine max* (L.) Mill.] genotypes in Pawe district, Metekel zone, BenishanguleGumuz regional state, north-western Ethiopia, Wudpecker Journal of Agricultural Research, 2: 240-245

Ghafoor A., Zubair M., and Malik B.A., 1990, Path analysis in mash (*Vigna mungo* L.), Pakistan Journal of Botany, 22(2): 160-167

Ghodrati G., 2013, Study of genetic variation and broad sense heritability for some qualitative and quantitative traits in soybean (*Glycine max* L.) genotypes, Journal of Current Opinion in Agriculture, 2: 31-35

Ghosh J., Ghosh, P.D., and Choudhury P.R., 2014, An assessment of genetic relatedness between soybeans [*Glycine max* (L.) Merrill] cultivars using SSR markers, American Journal of Plant Sciences, 05: 3089-3096
https://doi.org/10.4236/ajps.2014.520325

Hanson W.D., 1961, Heritability, statistical genetics and plant breeding, National Academy of Science, National Research Council, Washington, pp.125-140

He F.J., and Chen J.Q., 2013, Consumption of soybean, soy foods, soy isoflavones and breast cancer incidence: Differences between Chinese women and women in Western countries and possible mechanisms, Food Science and Human Wellness, 2: 35-38
https://doi.org/10.1016/j.fshw.2013.08.002

Henkel J., 2000, Soy: health claims for soy protein, question about other components, FDA consumer (Food and Drug Administration), 34(3): 18-20

Henry A., and Krishna G.V., 1990, Correlation and path coefficient analysis in pigeon pea, The Madras Agricultural Journal, 77(9-12): 443-446

Hossain S.,Haque M.M., and Rahman J., 2015, Genetic variability, correlation and path coefficient analysis of morphological traits in some extinct local aman rice (*Oryza sativa* L.), Journal of Rice Research, 3: 158

Hossain Z., and Komatsu S., 2014, Potentiality of soybean proteomics in untying the mechanism of flood and drought stress tolerance, Proteomes, 2(1): 107-127
https://doi.org/10.3390/proteomes2010107

Iqbal S., Mahmood T., Tahira M.A., Anwar M., and Sarwar M., 2003, Path-coefficient analysis in different genotypes of soybean [*Glycine max* (L.) Merrill], Pakistan Journal of Biological Science, 6(12): 1085-1087
https://doi.org/10.3923/pjbs.2003.1085.1087

Iqbal Z., Arshad M., Ashraf M., Naeem R., Malik M.F., and Waheed A., 2010, Genetic divergence and correlation studies of soybean [(*Glycine max* (L.) Merrill] genotypes, Pakistan Journal of Botany, 42: 971-976

Islam M.A., Raffi S.A., Hossain M.A., and Hasan A.K., 2015, Analysis of genetic variability, heritability and genetic advance for yield and yield associated traits in some promising advanced lines of rice, Progressive Agriculture, 26: 26-31
https://doi.org/10.3329/pa.v26i1.24511

Jain P.K., and Ramgiry S.R., 2000, Genetic variability of metric traits in Indian germplasm of soybean [*Glycine max* (L.) Merrill], Advances in Plant Science, 13: 127-131

Jain S., Srivastava S.C., Singh K.S., Indapurkar Y.M., and Singh B.K., 2015, Studies on genetic variability, character association and path analysis for yield and its contributing traits in soybean [*Glycine max* (L.) Merrill], Agricultural Research Communication Center, Legume Research, 38(2): 182-184
https://doi.org/10.5958/0976-0571.2015.00031.4

Johnson H.W., Robinson F.P.F., and Comstock R.E., 1955, Estimation of genetic and environmental variability in soybean, Journal of Agronomy, 47: 314-318
https://doi.org/10.2134/agronj1955.00021962004700070009x

Karasu A., Oz M., Goksoy A.T., and Turan Z.M., 2009, Genotype by environment interactions, stability, and heritability of seed yield and certain agronomical traits in soybean [*Glycine max* (L.) Merrill], African Journal of Biotechnology, 8: 580-590

Karnwal M.K., and Singh K., 2009, Studies on genetic variability, character association and path coefficient for seed yield and its contributing traits in soybean [*Glycine max* (L.) Merrill], Legume Research, 32: 70-73

Khan A., Hatam M., and Khan A., 2000, Heritability and interrelationship among yield determining components of soybean varieties, Pakistan Journal of Agricultural Research, 16: 5-8

Kim H.K., Kim Y.C., Kim S.T., Son G.B., Choi Y.W., Kang J.S., Park Y.H., Cho Y.S., and Choi I.S., 2010, Analysis of quantitative trait loci (QTLs) for seed size and fatty acid composition using recombinant inbred lines in soybean, Journal of Life Science, 20: 1186-1192
https://doi.org/10.5352/JLS.2010.20.8.1186

Kumar A.V., Kumar S., Lal K., Jolly M., and Sachdev A., 2014, Influence of gamma rays and ethylmethanesulphonate (EMS) on the levels of phytic acid, raffinose family oligosaccharides and antioxidants in soybean seeds of different genotypes, Journal of Plant Biochemistry and Biotechnology, 24: 204-209

Lal H., Rai M, Verma A., and Vishwanath, 2011, Analysis of genetic divergence of Dolichos Bean (*Lablab purpureus* L.) genotypes, Vegetable Science, 32(2): 129-132

Mahbub M.M., Rahman M.M., Hossain M.S., Mahmud F., and Kobir M.M.M., 2015, Genetic variability, correlation and path analysis for yield and yield components in soybean, American-Eurasian Journal of Agriculture & Environmental Science, 15: 231-236

Malek M.A., Raffi M.Y., Afroj M.S.S., Nath U.K., and Mondol M.M.A., 2014, Morphological characterization and assessment of genetic variability, character association, and divergence in soybean mutants, The Scientific World Journal, 1-12
https://doi.org/10.1155/2014/968796 PMid:25197722 PMCid:PMC4146355

Malik M.F.A., Muhammad A., Qureshi A.S., and Abdul G., 2007, Assessment of genetic variability, correlation and path analyses for yield and its components in soybean, Pakistan Journal of Botany, 39: 405-413

Malik M.F.A., Qureshi A.S., Ashraf M., and Ghafoor A., 2006, Genetic variability of the main yield related characters in soybean, International Journal of Agriculture and Biology, 8: 815-819

Meena O.P, and Bahadur V., 2013, Assessment of breeding potential of tomato (*Lycopersicon esculentum* Mill.) germplasm using D$^2$ analysis, The Bioscan, 8: 1145-1148

Meena O.P., and Bahadur V., 2014, Assessment of genetic variability, heritability and genetic advance among tomato (*Solanum lycopersicum* L.) germplasm, Agricultural Science Digest, 27: 185-192

Meitei K.M., Bora G.C., Singh J.S., and Sinha K.A., 2014, Morphology based genetic variability analysis and identification of important characters for tomato (*Solanum lycopersicum* L.) crop improvement, American-Eurasian Journal of Agricultural & Environmental Sciences, 14(10): 1105-1111

Miller P., Williams J., Robinson H., and Comstock R., 1958, Estimates of genotypic and environmental variances and co-variances in upland cotton and their implications in selection, Agronomy Journal, 50: 126-131
https://doi.org/10.2134/agronj1958.00021962005000030004x

Mohsen A.E., Ashraf A., Mahmoud G.O., and 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
https://doi.org/10.5897/JPBCS2013.0389

Nassar M.A.A., 2013, Heterosis and combining ability for yield and its components in some crosses of soybean, Australian Journal of Basic and Applied Sciences, 7: 566-572

Osekita O.S., and Olorunfemi O., 2014, Quantitative genetic variation, heritability and genetic advance in the segregating F$_3$ populations in soybean (*Glycine max* (L.) Merrill), International Journal of Advanced Research, 2: 82-89

Parmar A.M., Singh A.P., Dhillon N.P.S., and Jamwal M., 2013, Genetic variability studies for morphological and yield traits in Dolichos bean (*Lablab purpureus* L.), World Journal of Agricultural Sciences, 9(1): 24-28

Rahman M.M., Bashar M.K., and Rasul M.G., 2012, Molecular characterization and genetic variation in rice, LAP: Lambert Academic Publishing GmbH and Co. KG, Saarbrucken, Germany, l-45

Rahman M.M., Rasul M.G., Bashar M.K., Syed M.A., and Islam M.R., 2011, Parent selection for transplanted aman rice breeding by morphological, physiological and molecular diversity analysis, Libyan Agriculture Research Center Journal International, 2(l): 29-35

Ramteke R., Kumar V., Murlidharan P., and Dinesh K., 2010, Study of genetic variability and traits interrelationship among released soybean varieties [*Glycine max* (L.) Merrill] of India, Electronic Journal of Plant Breeding, 1: 1483-1487

Ranganatha H.M., Patil S.S., Manjula S.M., and ArvindkumarB.N., 2013, Genetic variability studies in segregating generation of upland cotton (*Gossypium hirsutum*L.), Molecular Plant Breeding, 4: 84-88

Reni Y.P., and Rao Y.K., 2013, Genetic variability in soybean [*Glycine max* (L.) Merrill], International Journal of Plant, Animal and Environmental Science, 3: 35-38

Salahuddin S., Abro M., Kandhro M., Salahuddin L., and Laghari S., 2010, Correlation and path coefficient analysis of yield components of upland cotton [*Gossypium hirsutum* (L.)] sympodial, World Applied Sciences Journal, 8: 71-75

Sarutayophat T., 2012, Correlation and path coefficient analysis for yield and its components in vegetable soybean, Songklanakarin Journal of Science and Technology, 34: 273-277

SAS Institute Inc., 2010, SAS/STAT user's guide, 3rd edition, SAS Institute Inc., Cary

Sathe S.K., Sharma G.M., Kshirsagar H.H., Su M., and Roux K.H., 2009, Effects of long-term frozen storage on electrophoretic patterns, immune reactivity, and pepsin in vitro digestibility of soybean (*Glycine max* L.) proteins, Journal of Agricultural and Food Chemistry, 57: 1312-1318
https://doi.org/10.1021/jf802905n PMid:19187024

Sathya, R., and Jebaraj S., 2013, Inter-Relationship and cause effect analysis among drought and physiological traits in three line aerobic rice hybrids, Plant Gene and Trait, 4: 70-73
https://doi.org/10.5376/pgt.2013.04.0013

Semahegn Y., and Tesfaye M., 2016, Characters associations and path analysis in safflower (*Carthamus tinctorious*) accessions, Molecular Plant Breeding, 7(31): 1-5

Sharma S., Kaur M., Goyal R., and Gill B.S., 2013, Physical characteristics and nutritional composition of some new soybean (*Glycine max* (L.) Merrill) genotypes, Journal of Food Science and Technology, 51: 551-557.
https://doi.org/10.1007/s13197-011-0517-7 PMid:24587531 PMCid:PMC3931876

Sharma S.M., Rao S.K., and Goswami U., 1983, Genetic variation, correlation and regression analysis and their implications in selection of exotic soybean, Mysore Journal of Agricultural Sciences, 17(1): 26-30

Shashikanth, Basavaraj N., Hosamani R.M., and Patil B.C., 2010, Genetic variability in tomato (*Solanum lycopersicum* L.), Karnataka Journal of Agricultural Science, 23(3): 536-537

Singh D.N., Sahu A., and Parida A.K., 1997, Genetic variability and correlation studies in tomato (*Lycopersicon esculentum* Mill.), Journal of Environment and Ecology, 15: 135-141

Sirohi S.P.S., Malik S., Singh S.P., Yadav R., and Meenakshi, 2007, Genetic variability, correlations and path coefficient analysis for seed yield and its components in soybean [*Glycine max* (L.) Merrill], Progressive Agriculture, 7: 119-123

Sureshrao S.S., Singh V.J., Gampala S., and Rangare N.R., 2014, Assessment of genetic variability of the main yield related characters in soybean, International Journal of Food, Agriculture and Veterinary Sciences, 4: 69-74

Turkec A., 2005, Correlation and path analysis of yield component in soybean varieties, Turkish Journal of Field Crops, 10: 43-48

Udensi O., and Ikpeme E.V., 2011, Correlation and path coefficient analyses of seed yield and its contributing traits in *Cajanus cajan* (L.) Mill. sp., American Journal of Experimental Agriculture, 2(3): 351-358
https://doi.org/10.9734/AJEA/2012/952

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

Yared S., and Misteru T., 2016, Variability, heritability and genetic advance analysis for some morphological characters in oilseed brassica breeding lines, Molecular Plant Breeding, 7(20): 1-8