Multivariate analysis unravels genetic diversity and relationship between agronomic traits, protein, and dietary fiber in yardlong bean (*Vigna unguiculata* subsp. *sesquipedalis* Verdc.)

**MUHAMMAD HABIB WIDYAWAN[1,2,♥], AULIA HASANAH[1], TARYONO[1,2], TAUFAN ALAM[1,2],
RAHMI SRI SAYEKTI[2], ABRORY AGUS CAHYA PRAMANA[3], RANI AGUSTINA WULANDARI[1,2,♥♥]**

[1]Department of Agronomy, Faculty of Agriculture, Universitas Gadjah Mada. Jl. Flora No. 1, Bulaksumur, Sleman 55281, Yogyakarta, Indonesia.
Tel./fax.: +62-274-548814, ♥email: habib.widyawan@ugm.ac.id
[2]Center of Agro Technology Innovation, Universitas Gadjah Mada. Jl. Tanjung Tirto, Berbah, Sleman 55573, Yogyakarta, Indonesia.
Tel.: +62-274-497717, ♥♥email: rani.akyun@gmail.com
[3]Department of Biochemistry, Faculty of Medicine, Public Health, and Nursing. Universitas Gadjah Mada. Jl. Farmako, Bulaksumur, Sleman 55281, Yogyakarta, Indonesia

Abstract. *Widyawan MB, Hasanah A, Taryono, Alam T, Sayekti RS, Pramana AAC, Wulandari RA. 2020. Multivariate analysis unravels genetic diversity and relationship between agronomic traits, protein, and dietary fiber in yardlong bean (Vigna unguiculata subsp. sesquipedalis Verdc.). Biodiversitas 21: 5662-5671.* Yardlong bean (*Vigna unguiculata* subsp. *sesquipedalis* Verdc.) is a vegetable legume that possesses the highest amount of protein and total dietary fiber (TDF) content among other vegetables. However, protein and TDF content diversity are found among yardlong bean genotypes. Genetic improvement to produce cultivars with high yield and improved nutritional content is required. This study aimed to analyze genetic diversity and identify the relationship between important agronomic traits from 21 yardlong bean genotypes assessed by multivariate statistical analysis. The experiment was arranged in alpha design with three replicates, each consist of six blocks. Six pod-related important agronomic traits, protein, and TDF content were observed in this study. Total pod weight, pod weight, and pod length were found to be negatively correlated with protein and TDF content based on correlation and principal component analysis (PCA). Although negatively correlated, several variables possess a large and positive direct effect on protein and TDF content, such as the number of pod, pod weight, and days to flower. D^2 analysis and Tocher's method classify genotypes used in this study into five clusters with diverse traits value. Information from this study laid a foundation for the improvement of important agronomic traits as well as nutritional content in yardlong bean.

Keywords: Genetic diversity, multivariate analysis, nutritional content, relationship, yardlong bean

INTRODUCTION

Yardlong bean (*Vigna unguiculata* subsp. *sesquipedalis* Verdc.) (2n = 2x = 22) is one of the most important vegetable legumes in Asia. Yardlong bean consumption per capita in Indonesia was almost three times higher than similar vegetable legumes such as green bean (*Phaseolus vulgaris*) (Ministry of Agriculture 2018). This plant was domesticated from cowpea (*Vigna unguiculata* subsp. *unguiculata*) in Asia. Domestication has results in the change of several traits such as growth habits, longer pods, and larger seed size (Kongjaimun et al. 2012). Fresh pods produced by yardlong bean is rich in vitamins and minerals, especially protein and dietary fiber content. Wills et al. (1994) reported that the protein and dietary fiber content in pods of yardlong bean were higher compared to those of other Chinese vegetables such as bean sprouts, spinach, cabbage, and flowering cabbage. According to Castañeda-Pérez et al. (2019), active peptides isolated from yardlong bean demonstrate an inhibitory effect against α-amylase, α-glucosidase, and dipeptidyl DPP-IV enzymes which suggests yardlong bean as potential therapeutic food for T2DM patients. It makes yardlong bean a potential health-promoting functional food.

Rambabu et al. (2016a) have reported the existence of variations in the nutritional content such as vitamin C and protein content of yardlong bean genotypes. Genetic improvement of yardlong bean is needed to produce cultivars with desired agronomic traits and improved nutritional content. Several authors have been reported the yardlong bean breeding programs that aimed at improving pod quality-related traits such as pod tenderness, total soluble solid (Kongjaimun et al. 2012), and pod color (Kuswanto et al. 2013). Genetic improvement of crops, at least, consists of three different aspects which include genetic diversity, selection of favorable genotypes, and evaluation (Voss-Fels et al. 2019). Evaluation of genetic diversity is the most fundamental step since it plays a major role in the determination of the breeding strategy. Moreover, the selection of targeted traits, especially those that not visible to the eyes such as nutritional content, could be challenging.

Multivariate analysis is a statistical method that involves multiple variables and has been employed to evaluate the genetic diversity of plants as well as to elucidate the relationship between traits (Zanklan et al. 2018). The success of multivariate analysis in the assessment of genetic diversity and the relationship

between traits has been demonstrated in various crop species such as garden pea (Sanwal et al. 2015), maize (Oliveira et al. 2016), and African rice (Lakshmi et al. 2019). Utilization of this technique to analyze the diversity and relationship between traits and nutritional content has been reported in vegetable amaranth (*Amaranthus tricolor*) (Shukla et al. 2010), rapeseed (Yousuf et al. 2011), sorghum (Shegro et al. 2013), and wheat (Hussain et al. 2014). In this present study, multivariate analysis including correlation, path, regression, principal component analysis (PCA), and cluster analysis were performed to reveal the genetic diversity and relationship between important agronomic traits, protein, and total dietary fiber (TDF) content in yardlong bean. This study aimed to reveal the genetic diversity as well as the relationship between important agronomic traits and nutritional content in yardlong bean genotypes using multivariate analysis. Results from this study will lay the foundation for the nutritional content improvement of yardlong bean through a "biofortification" breeding program.

MATERIALS AND METHODS

Plant materials

This study was conducted from September to December 2019 at Kalitirto Field Station, Center of Agro Innovation Technology, Universitas Gadjah Mada, Berbah Subdistrict, Sleman District, Province of Yogyakarta, Indonesia (7°79'58.59"N, 110°46'52.85"E). The soil type in this field station was inceptisol. Twenty-one (21) yardlong bean genotypes obtained from the gene bank of Universitas Gadjah Mada were used in this study. The genotypes consist of 18 germplasm (15 landrace and 3 genotypes courtesy of Indonesian Vegetable Research Center (IVEGRI), Bandung, Indonesia), and 3 commercial cultivars i.e. 'Oriental Seed', 'Parade Tavi', and 'Putih Super' (Table 1).

Procedures

The experiment was arranged in an alpha design with three replicates, each consists of six block-sized four in each block which yield a total of 72 plots of experimental units. Initially, there are 24 genotypes used in this study. However, three genotypes were discarded due to poor seed germination. The plot randomization in this present study was shown in Figure 1. Seeds of each genotype were sown directly in each plot sized 5 m², adopting a spacing of 70 cm between plants and 80 cm between rows. Plants were maintained using the standard cultivation procedure for yardlong bean as described by the Center of Agro Innovation Technology as described by Kathalia (2019). Data were calculated for six important pod characters including the number of pod (NP), pod length (PL), pod diameter (PD), pod weight (PW), total pod weight (TPW), days to flowering (DTF); as well as for two nutritional content namely protein and total dietary fiber (TDF). These traits are important in the yardlong bean breeding program since they were the main target for selection and easy to observed phenotypically. Protein content was quantified

using Kjedahl method while total dietary fiber was quantified using AOAC method (1997).

Data analysis

The mean of all observed traits was calculated using the least square mean using by R software package and preceded by the analysis of variance. Scott-Knott post-hoc analysis with 5% significant level was performed to observe the difference between the mean value of characters observed in this study. Correlation and path analysis was conducted according to Singh and Chaudary (1985). Multiple linear regression preceded by stepwise regression was also performed. The model with the lowest Akaike Information Criterion was selected according to the stepwise regression (Akaike 1974). Correlation, path, and regression analysis were done using agricolae (de Mendiburu 2020) of R software package. PCA and biplot analysis were performed by factoextra (Kassambra and Mundt 2020) and ggbiplot (Vu 2020), respectively, using R software package. Clustering of yardlong bean genotypes was established by Mahalanobis D^2 analysis followed by Tocher's method using biotools (da Silva 2017) of R software package. Heatmap was established by Euclidean distance and average methods using pheatmap (Raivo 2019) of R software package. All statistical analysis was performed using R Studio software version 3.6.2 (R Core Team 2017).
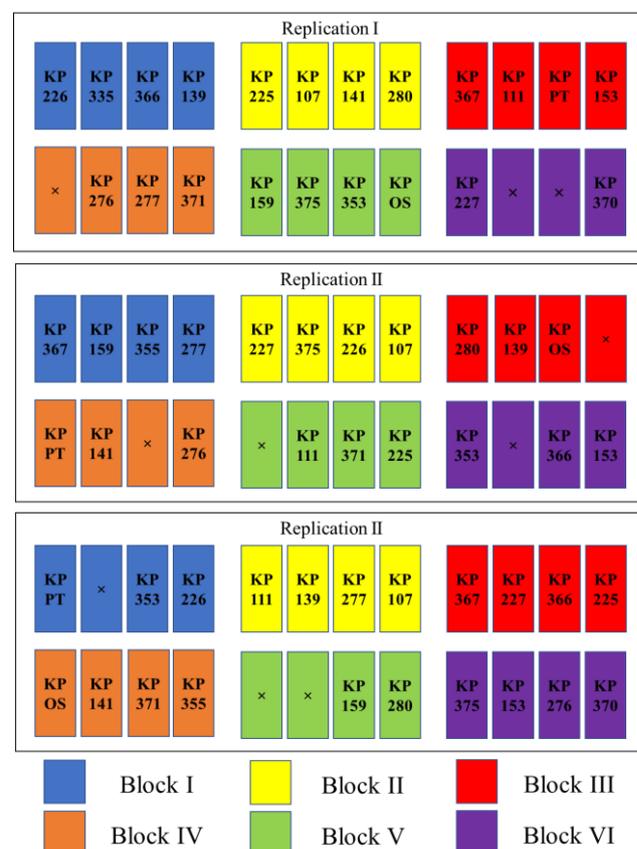

Figure 1. Plot randomization of 21 yardlong bean genotypes

RESULTS AND DISCUSSION

Variability of important agronomic traits

The variability of eight traits in yardlong bean genotypes is shown in Table 1. The average PL was 53.4±2.9 cm, with a minimum and maximum value of 31.7 cm (KP.225) and 88.4 cm (KP.139), respectively. The average NP was 39±3.8, with a minimum and maximum value of 14 (KP.370) and 66.5 (KP.PT), respectively. The average PD was 6.7±0.1 mm, with a minimum and maximum value of 5.6 mm (KP.227) and 7.8 mm (KP.111), respectively. The average PW was 11.7±1.6 g, with a minimum and maximum value of 2.5 g (KP.370) and 29.2 g (KP.PT), respectively. The average TPW was 616.2±89.3 g, with a minimum and maximum value of 254.8 g (KP.153) and 1389.3 (KP.OS), respectively. The average DTF was 42±1.6 days, with a minimum and maximum of 28 days (KP.PS) and 55 days (KP.366), respectively (Table 1).

Genetic diversity of traits among the yardlong bean genotypes were observed in this study. In this regard, commercial cultivars of yardlong bean showed a higher value of NP, PW, TPW, and DTF when compared to those of the germplasm. This was the expected outcome since the important agronomic traits were improved through selection preceded by several evaluations in the cultivar development. Although diversity is observed in morphological traits, Widyawan et al. (2020) reported narrow genetic diversity between yardlong bean at the DNA level as examined by retrotransposon-based molecular marker. The diversity that presents in the important agronomic traits among the genotypes used in this study was beneficial to the genetic improvement. For instance, the genotypes that possess the highest value of TPW but late in the flowering could be crossed with the other genotypes that flowering earlier and followed by selection. Additionally, the genetic relationship between genotypes could be investigated through DNA marker analysis to determine the crosses that yields the largest variability of the progenies.
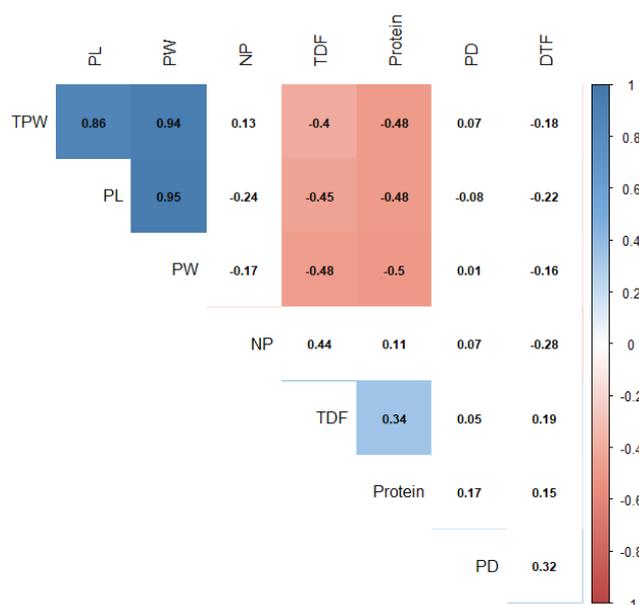
Variability of protein and dietary fiber content

The average protein content in the yardlong bean genotypes was 4.0±0.2 g/100 g, with a minimum and maximum value of 2.8 g/100 g (KP.280) and 5.6 g/100 g (KP.225), respectively. The average TDF in the yardlong bean was 11.6±0.5 g/100 g, with a minimum and maximum value of 8.6 g/100 g (KP.PT) and 17.1 g/100 g (KP.139), respectively (Table 1). In this study, several genotypes showed a high amount of protein and TDF, as indicated by the value of both traits above the average. Those genotypes were KP.139, KP.159, KP.225, KP.227, KP.355, KP.367, KP.371, and KP.375. In contrast, all cultivars i.e. KP.OS, KP.PT, and KP.PS were having protein and TDF value below the average. This finding indicates that nutrient content, such as protein and TDF, was not the priority in the commercial yardlong bean breeding programs. A similar finding was reported by Pidigam et al. (2019), the protein content of check cultivar 'Bhagyalaksmi' was lower compared to most Indian yardlong bean germplasm.

Table 1. Code, origin, source, and mean traits of yardlong bean genotypes and cultivars

Code	Origin	Mean of traits							
		Pod length (cm)	Number of pod	Pod diameter (mm)	Pod weight (g)	Total pod weight (g)	Days to flower	Protein (g/100g)	Total dietary fiber (g/100g)
KP.107[1]	Sleman	43.6[c]	34[e]	6.9[a]	7.5[d]	313.5[e]	44[c]	5.5[a]	9.8[h]
KP.111[1]	Sleman	70.8[b]	49[c]	7.8[a]	16.6[c]	1174.1[b]	41[d]	4.2[d]	8.9[i]
KP.139[1]	Pekalongan	88.4[a]	16[g]	7.3[a]	3.4[e]	330.7[d]	40[d]	4.7[c]	17.1[a]
KP.141[1]	Gresik	42.4[c]	15[g]	7.0[a]	3.9[e]	155.5[e]	44[c]	3.4[e]	9.1[i]
KP.153[1]	Blitar	64.6[b]	15[g]	6.9[a]	4.3[e]	254.8[e]	48[b]	4.1[d]	11.2[f]
KP.159[1]	Unknown	72.0[b]	18[g]	5.9[c]	4.7[e]	366.4[e]	40[d]	5.0[b]	13.2[d]
KP.225[1]	Sukoharjo	31.7[d]	32[e]	7.4[a]	8.8[d]	267.1[e]	52[a]	5.6[a]	12.1[e]
KP.226[1]	Sukoharjo	45.8[c]	25[f]	6.3[b]	4.4[e]	199.2[e]	44[c]	3.0[f]	10.6[g]
KP.227[1]	Sukoharjo	67.6[b]	33[e]	7.0[a]	8.9[d]	593.4[d]	40[d]	4.7[c]	14.1[b]
KP.276[2]	IVEGRI	39.0[d]	59[b]	7.3[a]	17.8[c]	683.2[c]	40[d]	3.8[d]	10.5[g]
KP.277[2]	IVEGRI	55.5[c]	57[b]	5.6[c]	15.4[c]	816.5[c]	40[d]	3.4[e]	11.5[e]
KP.280[2]	IVEGRI	47.6[c]	56[b]	6.7[a]	16.6[c]	783.0[c]	40[d]	2.8[f]	13.0[e]
KP.353[1]	Purbalingga	55.1[c]	58[b]	7.0[a]	18.5[c]	1013.4[c]	41[c]	3.5[e]	11.0[g]
KP.355[1]	Purbalingga	57.6[b]	38[d]	6.5[b]	8.4[d]	478.2[d]	43[c]	4.3[c]	10.5[g]
KP.366[1]	Prambanan	60.2[b]	61[b]	7.1[a]	22.6[b]	1347.7[b]	55[a]	3.2[f]	12.1[e]
KP.367[1]	Prambanan	44.8[c]	28[f]	7.0[a]	7.2[d]	350.4	51[b]	4.2[d]	13.9[c]
KP.371[1]	Prambanan	52.5[c]	36[d]	6.0[c]	7.2[d]	349.8[d]	43[c]	4.7[c]	11.8[e]
KP.375[1]	Prambanan	50.5[c]	30[f]	6.9[a]	6.8[d]	372.2	48[b]	4.2[c]	13.5[c]
KP.OS[2]	Oriental Seed	63.1[b]	58[b]	6.7[a]	21.7[b]	1389.3[a]	40[d]	3.4[e]	8.7[i]
KP.PT[2]	East-West Seed	44.5[c]	66[a]	6.5[b]	29.2[a]	1354.6[a]	42[d]	3.5[e]	8.6[i]
KP.PS[2]	Chia Tai Seed	44.9[c]	62[b]	6.1[b]	20.1[b]	888.5b[c]	38[d]	3.7[e]	10.6[g]
Min		31.7	14	5.6	2.5	254.8	38	2.8	8.6
Max		88.4	66	7.8	29.2	1389.3	55	5.6	17.1
Mean±SE		53.4±2.9	39±3.8	6.7±0.1	11.7±1.6	616.2±89.3	42±1.6	4.0±0.2	11.6±0.5
Stdev		13.8	18.0	0.6	7.6	418.7	7.6	0.8	2.1

Note: Superscript number in the Code column indicates the source of genotypes. 1 = germplasm, 2 = commercial cultivars. The same letters in the mean value of each trait indicate no significant difference between genotypes based on Scott-Knott test

Figure 2. Correlogram of pod traits, protein, and total dietary fiber content in yardlong bean. Note: Colored box indicates significant correlation (p < 0.05) between two variables

Diversity in the important agronomic traits such as pod length and pod number per plants, as well as the protein content using 24 diverse yardlong bean genotypes from Orisha and Andhra Pradesh region of India have been reported by Pidigam et al. (2019). In this present study, KP.225 and KP.139 were identified as genotypes with the highest value of protein content and TDF, respectively. However, their important agronomic traits value such as NP, PW, and TPW are below average. Based on this finding, these genotypes could be utilized as donors in subsequent yardlong bean breeding programs to produce biofortified cultivars with good agronomic performance. A breeding method such as the backcross approach could be used to integrate the high nutritional content traits into the other genotypes as well as commercial cultivars.

Correlation, path, and regression analysis

Correlation, regression, and path analysis of important agronomic traits, protein, and total dietary fiber content were performed to reveal the relationship between the observed traits. The result of the correlation analysis is shown in Figure 2. There was a strong and positive correlation between PL and PW (r = 0.95, p < 0.05), PL and TPW (r = 0.86, p < 0.05), and between PW and TPW (r = 0.94, p < 0.05). This finding suggests that genotypes exhibiting longer pod sizes may possess a higher yield. Similarly, Patel et al. (2016) and Rambabu et al. (2016b) reported a positive correlation between pod length and total pod yield per plant in yardlong bean and cowpea (*Vigna unguiculata* subsp. *unguiculata*), respectively. On the other hand, the present study revealed that most pod related traits were negatively correlated to the protein and TDF content. Significant negative correlation was observed between PW and protein content (r = 0.5, p < 0.05), TPW and protein content (r = 0.48, p < 0.05), and between PL and protein content (r = 0.48, p < 0.05). Moreover, the negative correlation of protein content to yield related component have been observed in green gram (*Vigna radiata*) (Kumar et al. 2013) and chickpea (*Cicer arietinum*) (Gaur et al. 2016). Similarly, this study also revealed a negative correlation between PL and TDF (r = −0.45, p < 0.05), as well as between PW and TDF (r = −0.48). In contrast, NP exhibited a positive correlation with TDF (r = 0.44, p < 0.05).

Correlation analysis is a powerful statistical tool used to measure the degree of association between traits in plants (Silva et al. 2016). The existence of a correlation between traits may also indicate an association between the genes that controlling those traits (Neyhart et al. 2019). The positive correlation will be beneficial in selecting multiple traits based on the assertion that the genetic gain of several traits could be achieved by selecting a single trait that is strongly and positively correlated to many other traits (Anshori et al. 2019). This strategy could be very useful for the selection of traits that are not visible to the naked eyes, need sophisticated tools, or additional analysis such as protein, oil, and nutrient content. In this study, the traits of interest which are protein and TDF content exhibited a negative correlation with the yield-related traits. A negative correlation between protein content as well as dietary fiber to the yield has been reported in various crop species. Thorwarth et al. (2019) and Obala et al. (2020) also reported a negative correlation between yield and protein content in wheat (*Triticum aestivum*) and pigeon pea (*Cajanus cajan*), respectively. Similarly, Londero et al. (2006) reported a negative correlation between yield and dietary fiber content in common bean (*Phaseolus vulgaris*). This finding, therefore, challenges the selection process to develop biofortified cultivars since it will restrict the utilization of indirect selection (Harshman et al. 2016). Thus, it will increase the labor, time, and cost to analyze the nutritional content in the breeding population.

The direct and indirect effect of pod traits on protein and TDF content in the yardlong bean genotypes were revealed by path analysis (Table 2). The result is useful for providing a better interpretation of the relationship between traits that may eventually become the target of improvement. If a trait possesses a positive direct effect on another trait, it could be used directly to perform an indirect selection. On the other hand, if a negative direct effect was present, an examination of traits with positive indirect effects should be conducted to enable indirect selection (Reddy and Jaben 2016). PW, NP, PL, and PD portrayed the largest positive direct effect on the protein (0.363, 0.340, 0.191, and 0.134, respectively). Meanwhile, TPW had a large negative direct effect on the protein content (−0.974). In contrast, Haghi et al. (2012) reported a large and positive direct effect of grain yield on the protein content in soybean (*Glycine max*). Protein content was the major trait that becomes an interest in the genetic improvement program for a long time and is known to have a high heritability in soybean (Duhnen et al. 2017, Stewart-Brown et al. 2019). This might explain the positive direct effect of yield to the protein content in soybean. The model of path analysis for protein content was able to explain a low level of variation of protein content (16.84%). It is

indicated that PW and NP could be used for indirect selection for protein content. Moreover, PL and TDF could be used for indirect selection since they possessed a positive indirect effect through PW (0.345) and TPW (0.390), respectively. Several variables including PW, NP, PL and DTF exerted large direct effects on the TDF content (1.821, 1.400, 0.630, and 0.514, respectively). Similar to the first model, TPW had a large negative indirect effect on the TDF content (−2.741). In contrast to the protein model, the latest model was able to explain 58.76% of TDF content variability. This finding suggests that TDF content could be selected indirectly through PW, NP, PL, and DTF.

Path analysis was used to establish a partition of the correlation coefficients into direct and indirect effects. It provides an effective way to unravel the relationship between traits hereby accounting for the interaction between them as the basis of selection strategy in a breeding program (Kumar et al. 2014). The direct effect could be useful to provide a guide to select traits that are likely to yield an overall improvement of the trait of interest (Lakshmi et al. 2019). On the other hand, the direct value of some traits acted oppositely with their correlation coefficient (e.g., positive direct effect follows negative correlation and *vice versa*). Singh and Chaudhary (1985) suggest that this might be caused by the indirect effect of other traits that masked the direct effect. Adjustment of the model, either by removing or adding variables, could be done to provide more insight into the relationship between the traits.

Multiple regression is the appropriate method of analysis employed when the research problem involves a single metric dependent variable that is presumed to be related to two or more metric independent variables (Hair et al. 2014). Regression analysis results for the two models revealed the relationship between pod traits, protein, and TDF content. To provide a better model for the prediction of protein and TDF, a stepwise regression analysis was performed. Only variables that exhibited significant

regression i.e. TPW to protein content and NP, TPW, and DTF to TDF were included to establish the model. The regression equation for protein and TDF content based on stepwise regression were as follows:

$$\text{Protein content} = 4.637 - 0.0009238 \text{ TPW} \quad \text{…………….…..} \ (1)$$

$$\text{TDF} = -11.516 + 0.211 \text{ NP} - 0.0124 \text{ TPW} + 0.245 \text{ DTF} \ \text{…….} \ (2)$$

The first model showed that TPW was the only variable that had a significant negative relationship with the protein content. In addition to the path analysis result for protein content, it could be concluded that the first model could be neglected due to the low variation that could be explained by the model. Therefore, improvement to the model could be done in future studies through the addition of variables in the model. Protein content in the plant has been identified to be controlled by multiple genes or Quantitative Trait Loci (QTL) with various gene actions i.e. additive or epistasis (Sarvamangala et al. 2011; Santos et al. 2012; Purnamasari et al. 2019). It could explain the reason for the low amount of variation that could be explained by the first model. The second model revealed that TDF had a positive relationship with NP (0.211) and DTF (0.245); and a negative relationship with TPW (−0.0124). The first and second models were able to explain the protein content and TDF by 19.03% and 56.27%, respectively, and this was indicated by the coefficient of determination (R^2). Based on the regression analysis, it can be inferred that NP and DTF could be considered for selection to increase the TDF content the in yardlong bean. In the case of protein content, adjustment of the model by addition of more variables could be done to improve the reliability of the model. Yield-related traits are likely difficult to predict protein content in plants, as reported by Salimi and Moradi (2012) and El-Mohsen et al. (2013). There is no significant regression obtained between protein content and yield-related traits.

Table 2. The direct and indirect effect of pod related traits on protein and dietary fiber content in yardlong bean

Protein	Number of pod	Pod length	Pod diameter	Pod weight	Total pod weight	Days to flower	Total dietary fiber
Number of Pod	**0.340**	−0.032	0.013	−0.062	−0.127	−0.027	0.003
Pod Length	−0.082	**0.134**	−0.015	0.345	−0.837	−0.021	−0.003
Pod Diameter	0.024	−0.011	**0.191**	0.004	−0.068	0.030	0.000
Pod Weight	−0.058	0.127	0.002	**0.363**	−0.915	−0.015	−0.004
Total Pod Weight	0.044	0.115	0.013	0.341	**−0.974**	−0.017	−0.003
Days to Flower	−0.095	−0.029	0.061	−0.058	0.175	**0.095**	0.001
Total Dietary Fiber	0.150	−0.060	0.010	−0.174	0.390	0.018	**0.008**
Variation explained: 16.84%.							
Total dietary fiber	Number of pod	Pod length	Pod diameter	Pod weight	Total pod weight	Days to flower	Protein
Number of Pod	**1.400**	−0.151	0.001	−0.310	−0.356	−0.144	0.000
Pod Length	−0.336	**0.630**	−0.001	1.730	−2.358	−0.113	−0.002
Pod Diameter	0.098	−0.050	**0.011**	0.018	−0.192	0.165	0.001
Pod Weight	−0.238	0.598	0.000	**1.821**	−2.577	−0.082	−0.002
Total Pod Weight	0.182	0.541	0.001	1.712	**−2.741**	−0.093	−0.002
Days to Flower	−0.392	−0.139	0.003	−0.291	0.493	**0.514**	0.001
Protein	0.154	−0.302	0.002	−0.910	1.316	0.077	**0.004**
Variation explained: 58.76%.							

Note: numbers in bold indicate the direct effect of each variable on protein or total dietary fiber content.

Table 3. Eigenvalue, contribution to variability, and factor loading for the principal component (PC) analysis of the agronomic traits and nutritional content in yardlong bean genotypes.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Number of Pod	0.2170	0.7963	−0.4785	−0.0261	0.1258	−0.2655	0.0454	−0.0305
Pod Length	−0.9392	−0.0154	−0.0723	−0.0408	−0.2590	0.1463	0.1487	−0.0093
Pod Diameter	0.0954	−0.3679	−0.7806	0.3491	0.2838	0.2091	0.0077	−0.0014
Pod Weight	−0.9541	−0.0270	−0.1953	−0.0329	−0.1844	0.0213	−0.1066	−0.0629
Total Pod Weight	−0.8959	0.1571	−0.3668	−0.0359	−0.0877	−0.1467	−0.0448	0.0756
Days to Flower	0.2733	−0.7284	−0.3331	−0.4486	−0.0696	−0.2775	0.0254	−0.0092
Protein	0.6500	−0.0593	−0.1546	0.4418	−0.5857	−0.1092	−0.0039	0.0020
Total Dietary Fiber	0.6329	0.3169	−0.3366	−0.4646	−0.2409	0.3331	−0.0270	0.0109
Eigenvalue	3.549	1.430	1.264	0.739	0.611	0.358	0.039	0.011
Variability (%)	44.360	17.869	15.804	9.236	7.639	4.469	0.488	0.136
Cumulative (%)	44.360	62.229[*]	78.033	87.269	94.907	99.376	99.864	100.000

Note: PC: Principal Component; *: the percentage of cumulative variation that could be explained by PC1 and PC2.

Principle Component Analysis (PCA)

The results of the PCA revealed that the first two principal components were able to explain 62.29% of the total variation (Table 3). The eigenvalue of the first and second principal component were 3.549 and 1.430, respectively. These values were classified as satisfied according to Kaiser's criterion (eigenvalue >1) (Kaiser 1958). In the first principal component, PW (0.9541), PL (0.9392), TPW (0.8959), protein (0.6500) and TDF (0.6329) had greater contributions to the total morphological variability (44.36%). The second principal component was able to explain 17.87% of variability, with NP (0.7963), DTF (0.7284), PD (0.3679), and TDF (0.3169) identified as the largest contributors in the component. Apart from the contribution of each trait to total variability, PCA was also able to dissect the genotypes contribution to total diversity. Genotypes are said to have meaningful contributions if they exhibit principal component scores of more than 1 (Mellidou et al. 2020). A list of genotypes with a significant contribution based on the principal component score in PC1 and PC2 are tabulated in Table 4. Based on the first PC, there were 10 genotypes selected with the principal component score value ranging from 1.0327 (KP.371) to 3.0599 (KP.139). On the other hand, four genotypes were selected from the second PC including KP.227 (1.3011), KP.277 (1.3011), KP.159 (1.9586), and KP.139 (2.4215). Several genotypes i.e. KP.227, KP.159, and KP.139 were presented in both PC1 and PC2. These genotypes tended to possess large variations of the important agronomic traits and nutritional content. Moreover, utilization of these genotypes as parents could be beneficial in the yardlong bean breeding program. Since they could potentially generate a larger magnitude of genetic diversity (Lakshmi et al. 2019).

PCA was used to simplify the relationship between sets of variables that were inter-correlated. This analysis produced new variables, also referred to as the principal components, which cover the essential information from the original dataset (Lever et al. 2017). Each principal component could be further dissected to reveal the traits that largely contributed to the total variation. This information is very useful in the study of the genetic diversity of crop plants. To increase efficiency, observation could be done on the characters that are known to have largely contributed to the total morphological variability. Lakshmi et al. (2019) reported that yield, days to 50% flowering, and a number of spikelets per panicle were contributed larger than the other traits to the total morphological diversity in the African wild rice (*Oryza glaberrima*). Biplot analysis preceding PCA was used to determine the pattern of classification of the studied genotypes (Figure 3). The biplot analysis revealed that commercial cultivars such as KP.OS, KP.PS, and KP.PT were clustered together along with the other genotypes such as KP.277, KP.111, KP.280, and KP.353. These three commercial cultivars as well as the other genotypes mentioned previously are likely to exhibit a similar PW, TPW, and PL. These findings corroborate the result of the descriptive statistics. Biplot analysis could be used to reveal the relationship between characters as well as confirmed the results of correlation analysis (Cai et al. 2018). PL, PW, and TPW were identified to exhibit strong correlation, as indicated by the narrow-angle between them in the biplot. The opposite direction of the characters arrow indicated a negative correlation between them, as shown by the protein content and TDF to the other important agronomic traits such as PL, PW, and TPW.

Table 4. The score of each genotype in the first and second PC that has a significant contribution to diversity

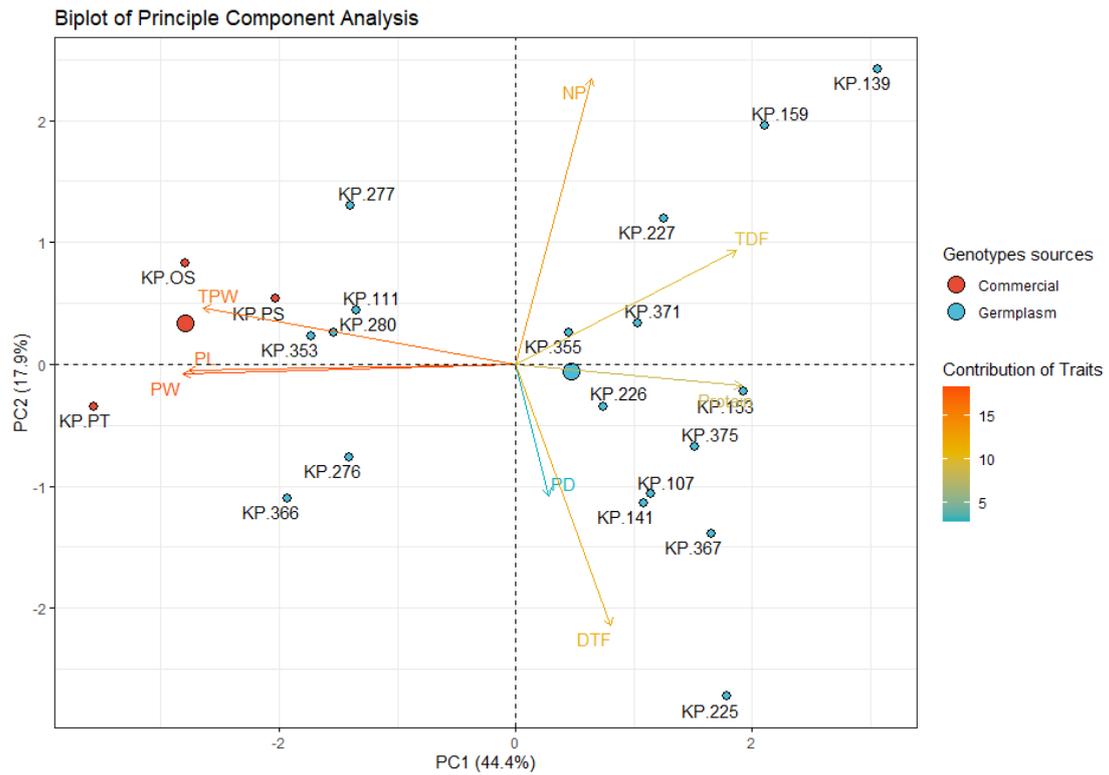| PC1		PC2	
Genotypes	Score	Genotypes	Score
KP.107	1.1440	KP.139	2.4215
KP.139	3.0599	KP.159	1.9586
KP.141	1.0778	KP.227	1.1939
KP.153	1.9220	KP.277	1.3011
KP.159	2.1068		
KP.225	1.7857		
KP.227	1.2552		
KP.367	1.6498		
KP.371	1.0337		
KP.375	1.5114		

Note: PC: Principal Component

Figure 3. Biplot of Principal Component Analysis for pod traits, protein, and total dietary fiber content
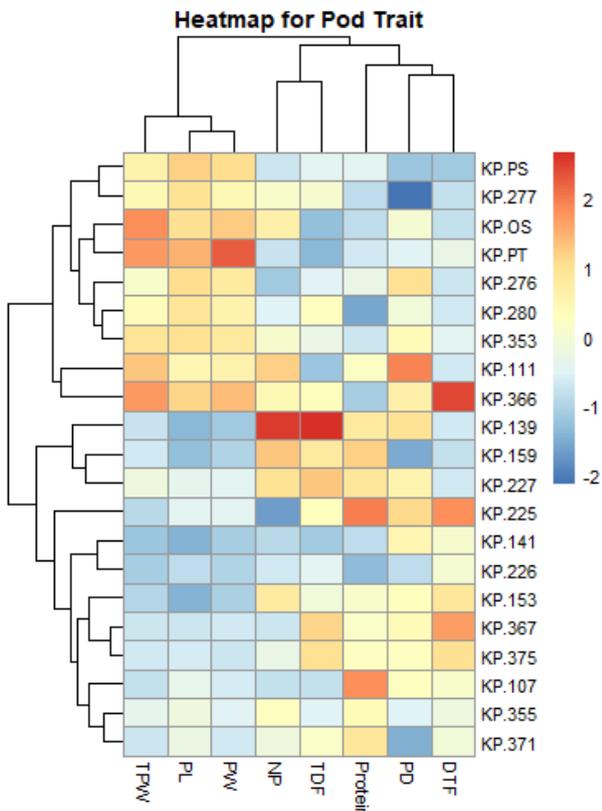

Figure 4. Heatmap representing the genetic variation of the agronomic traits and nutritional content in the yardlong bean

Cluster analysis

A heatmap was constructed to unravel the variability of each trait among the genotypes used in this study (Figure 4). The same technique has been applied by Anshori et al. (2020) to select doubled-haploid rice lines under hydroponic salinity screening. Based on the heatmap, it can be inferred that there is a broad genetic diversity in NP, TDF, Protein, PD, and DTF. The variation in PW, TPW, and PL was smaller compared to the aforementioned traits. Cluster analysis was performed based on the values of D^2 using Tocher's method. Tocher's is a distance-based optimization method that is commonly used accompanying the hierarchical clustering method and is expected to provide more accuracy (da Silva and Dias 2013; Meena and Bahadur 2015). It will provide more information to the other clustering methods i.e. heatmap or dendrogram since it could identify the optimum number of clusters (Pires et al. 2013).

Genotypes used in this study were classified into 5 clusters (Table 5). Cluster I was the largest with 7 genotypes (KP.367, KP. 371, KP.375, KP.159, KP.139, KP.107, and KP. 225), followed by cluster III and II with 4 (KP.277, KP.280, KP.PS, and KP.276) and 3 genotypes (KP.366, KP.PT, and KP.OS,) respectively. Cluster IV and V had two members each i.e. KP.227 and KP.355 in cluster IV; KP.111 and KP.355 in cluster V. Genotypes This result shows an agreement in terms of genotype classification compared to heatmap analysis, although it seems to be divided the genotypes into two major clusters. Most of the genotypes that belong to the same clusters based on were also grouped in the heatmap. Similar results also have been reported in soybean (*Glycine max*) (Cantelli et al. 2016).

Table 5. Member of each cluster and their corresponding average value for each trait based on Tocher's method

Cluster	Genotypes	Mean							
		NP	PL	PD	PW	TPW	DTF	Protein	TDF
I	KP.367, KP. 371, KP.375, KP.159, KP.139, KP.107, KP. 225	27.77	54.79	6.74	6.53	335.73	45.36	4.84	13.03
II	KP.366, KP.PT, KP. OS	61.73	55.93	6.77	24.51	1363.87	45.47	3.36	9.80
III	KP.277, KP.280, KP.PS, KP.276	58.55	46.75	6.43	17.49	792.80	39.60	3.43	11.15
IV	KP.227, KP.355	35.85	62.60	6.75	8.63	535.80	41.50	4.54	12.32
V	KP.111, KP.353	53.20	62.95	7.38	17.55	1093.75	40.95	3.84	9.91

Note: NP: number of pod; PL: pod length; PD: pod diameter; PW: pod weight; TPW: total pod weight; DTF: days to flowering; TDF: total dietary fiber

Table 6. Average of inter and intracluster distance of yardlong bean genotypes by D2 analysis

	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V
Cluster I	**93.06**	1068.93	498.64	241.17	798.72
Cluster II		**36.26**	571.38	828.78	271.11
Cluster III			**108.90**	259.47	301.57
Cluster IV				**115.82**	558.45
Cluster V					**161.75**

Note: Intra cluster distance are typed in bold

Each cluster possesses different characteristics of important agronomic traits and nutritional content, as shown by its average value (Table 5). The highest value of TPW was belong to cluster II and V, respectively. Cluster II exhibited the largest average value of most important agronomic traits i.e. TPW, PW, NP, and DTF. Two commercial cultivars i.e. KP.PT and KP.OS and one genotype; KP.366 belonged to this cluster. In contrast, the value of protein and TDF content in cluster II were the lowest in this study. This finding supports the results of descriptive statistics analysis: commercial cultivars are superior in the important agronomic traits but low in the nutritional content. The highest value of protein and TDF content were obtained in cluster I and IV, respectively. These clusters possess the lowest value of important agronomic traits i.e. NP, PW, and TPW. Interestingly, clusters that exhibited the highest agronomic traits were lowest in the nutritional content value and *vice versa*. Cluster III, on the other hand, placed in the middle rank on almost all of the traits observed in this study.

The average of inter and intracluster distance of yardlong bean genotypes based on the important agronomic traits, protein, and TDF content are shown in Table 6. The average value of the intracluster distance was 103.16, with cluster II being the smallest (36.26) and cluster V being the largest (161.75). The inter-cluster distance exhibited a larger variation than the intracluster distance, with an average of 539.82. The largest inter-cluster distance was found between clusters I and II (1068.93) while the smallest was found between clusters I and IV (241.17). Knowledge of inter and intracluster distance would be useful in a breeding program, since it provides an insight to select parentals candidate for hybridization. This was crucial to obtain maximum heterosis or genetic variability in their offspring (Kumar et al. 2017).

In this present study, traits such as PW, TPW, and PL showed less diversity among the genotypes. The aforementioned traits could be considered as the most important traits in yardlong bean since they are the major determinant of yield. Moreover, these traits also results of domestication from yardlong bean known ancestor, cowpea (*Vigna unguiculata* subsp. *unguiculata*) (Kongjaimun et al. 2012). Therefore, this could be the reason behind narrow diversity among PW, TPW, and PL in yardlong bean. On the other hand, nutritional content such as that protein and TDF content were found to be diverse among genotypes. Protein content in yardlong bean is also found to be diverse among genotypes. Rambabu et al. (2016c) and Rambabu et al. (2016d) reported broad genetic diversity in the yardlong bean germplasm from India. On the other hand, a report on breeding for fiber content in yardlong bean is still limited to date. Genes and quantitative trait loci (QTL) for fiber content in yardlong bean have been identified (Suanum et al. 2016) and known to exhibited dominant effect (Watcharatpong et al. 2020).

The D^2 statistics is a powerful tool used to measure divergence among genotypes and it had been successfully applied for parental selection in the cross-hybridization program (Kumar et al. 2008), as well as in the classification of rice genotypes (Lakshmi et al. 2019). In this study, genotypes that belong to different clusters were assumed to possess distinct genetic background as revealed by their respective mean value. This is beneficial for the development of biofortified yardlong bean cultivars with good agronomic traits. The crossing of genotypes from different clusters could yield progenies with high genetic diversity and desirable morphological features (Li et al. 2014).

Breeding of crops to improve their nutritional content is somewhat challenging as it involves a complex background of genetic interaction (Mahender et al. 2016). To ensure an effective and efficient improvement, it is important to elucidate the relationship between nutritional content and other traits. In this case, multivariate analysis was used to elucidate this relationship. A negative relationship between the nutritional content (protein and total dietary content) and the majority of the agronomic traits were observed in this study. As a strategy, the construction of a selection

index could be applied to overcome this problem. The construction of selection index for multiple traits with negative genetic correlation has been demonstrated in common bean (*Phaseolus vulgaris*) (Lima et al. 2015), soybean (Woyann et al. 2019), rice (Anshori et al. 2019), as well as in the tree breeding scenario as exemplified in scots pine (*Pinus sylvestris*) (Hong et al. 2014) and Khasi pine (*Pinus kesiya*) (Missanjo and Matsumura 2017). Therefore, utilization of the selection index could simultaneously improve agronomic traits and nutrient content in yardlong bean. The other approach utilizing genomic tools such as marker-assisted selection, genome-wide association studies (GWAS), and genomic selection (GS) could be considered to assist the genetic improvement of important agronomic traits as well as nutrient content in yardlong bean.

Multivariate analysis revealed the genetic diversity and relationship between important agronomic traits and nutritional content (protein and TDF) in yardlong bean. It was also revealed that there were genetically diverse important agronomic traits and nutritional content among the yardlong bean genotypes. Commercial cultivars used in this study exhibited low nutritional content despite having good agronomic traits. On the other hand, the nutritional content was higher in the genotypes that possessed the less favorable agronomic traits. Thus, there is a possibility to develop yardlong bean cultivars with improved agronomic traits and nutritional content from the genotypes used in this study. Multivariate analysis revealed that most of the important agronomic traits were negatively correlated to the nutritional content. Therefore, simultaneous improvement of agronomic traits and nutritional content should be conducted using appropriate strategies. This study provides fundamental information to support the genetic improvement of yardlong bean.

ACKNOWLEDGEMENTS

REFERENCES

Akaike H. 1974. A new look at the statistical model identification. IEEE Trans Auto Cont 19: 716-723.
Anshori MF, Purwoko BS, Dewi IS, Ardie SW, Suwarno WB. 2019. Selection index based on multivariate analysis for selecting doubled-haploid rice lines in lowland saline prone area. SABRAO J Breed Genet 51: 161-174.
Anshori MF, Purwoko BS, Dewi IS, Suwarno WB, Ardie SW. 2020. Cluster heatmap for detection of good tolerance trait on doubled-haploid rice lines under hydroponic salinity screening. Proceeding of International Conference on Sustainable Cereals and Crops Production Systems in the Tropics. Indonesian Agency for Agricultural Research and Development (IAARD)-Ministry of Agriculture of Indonesia, Makassar, 23-25 September 2019. [Indonesian].
AOAC. 1997. Official Methods of Analysis 16th ed. AOAC, Washington DC.
Cai T, Meng X, Liu X, Liu T, Wang H, Jia Z, Yang D, Ren X. 2018. Exogenous hormonal application regulates the occurrence of wheat

tillers by changing endogenous hormones. Front Plant Sci 9: 1-17. DOI: 10.3389/fpls.2018.01886
Cantelli DAV, Hamawaki OT, Rocha MR, Nogueira APO, Hamawaki RL, Sousa LB, Hamawaki CDI. 2016. Analysis of the genetic divergence of soybean lines through hierarchical and Tocher optimization methods. Genet Mol Res 15(4): 1-13. DOI: 10.4238/gmr.15048836.
Castañeda-Pérez E, Jiménez-Morales K, Quintal-Novelo C, Moo-Puc R, Chel-Guerrero L, Betancur-Ancona D. 2019. Enzymatic protein hydrolysates and ultrafiltered peptide fractions from Cowpea *Vigna unguiculata* L. bean with in vitro antidiabetic potential. J Iran Chem Soc 16: 1773-1781. DOI: 10.1007/s13738-019-01651-0
da Silva AR, Dias CTS. 2013. A cophenetic correlation coefficient for Tocher's method. Pesq Agropec Bras 48(6): 589-596.
da Silva AR. 2017. Biotools: Tools for biometry and applied statistics in agricultural science. Retrieved on April 13, 2020 from https://cran.r-project.org/web/packages/biotools/index.html.
de Mendiburu F. 2020. Agricolae: Statistical procedures for agricultural research.https://cran.r-project.org/web/packages/agricolae/index.html.
Duhnen A, Gras A, Teyssedre S, Romestant M, Claustres B, Mangin JDB. 2017. Genomic selection for yield and seed protein content in soybean: A study of breeding program data and assessment of prediction accuracy. Crop Sci 57(3): 1325-1337.
El-Mohsen AAA, Mahmoud GO, Safina SA. 2013. Agronomical evaluation of six soybean cultivars using correlation and regression analysis under different irrigation regime conditions. J Plant Breed Crop Sci 5(5): 91-102.
Gaur PM, Singh MK, Samineni S, Sajja SB, Jukanti AJ, Kamatam S, Varshney RK. 2016. Inheritance of protein content and its relationships with seed size, grain yield, and other traits in chickpea. Euphytica 209: 253-260. DOI: 10.1007/s10681-016-1678-2
Haghi Y, Boroomandan P, Moradin M, Hassankhali M. 2012. Correlation and path analysis for yield, oil and protein content of Soybean (*Glycine max* L.) genotypes under different levels of nitrogen starter and plant density.
Biharean Biol. 6: 32-37
Harshman JM, Evans KM, Hardner CG. 2016. Cost and accuracy of advanced breeding trial designs in apple. Hortic Res 3: 1-8. DOI: 10.1038/hortres.2016.8
Hair JF, Black WC, Babin BJ, Anderson RE. 2014. Multivariate Data Analysis. 7th Edition. Pearson Education Limited, London.
Hussain SB, Wahid MA, Zubair M, Babar M, Wahid K. 2014. Assessment of germplasm using multivariate analysis for grain yield and quality traits in spring wheat. Pak J Bot 46: 989-994.
Hong Z, Fries A, Wu HX. 2014. High negative genetic correlations between growth traits and wood properties suggest incorporating multiple traits selection including economic weights for the future Scots pine breeding program. Ann For Sci 71: 463-472. DOI: 10.1007/s13595-014-0359-3
Ministry of Agriculture. 2018. Agricultural Statistics 2018. Center for Agricultural Data and Information System, Jakarta.
Kaiser HF. 1958. The varimax criterion for analytic rotation in factor analysis. Psychometrika 23: 187-200.
Kassambra A, Mundt F. 2020. Factoextra: Extract and visualize the results of multivariate data analyses. https://cran.r-project.org/web/packages/factoextra/index.html.
Kathalia M. 2019. Response of yardlong bean (*Vigna unguiculata* subsp. *sesquipedalis*) to *Bradyrhizobium* inoculation. [Bachelor Thesis]. Universitas Gadjah Mada, Yogyakarta. [Indonesian].
Kumar B, Verma AK, Singh HP, Misra HO, Lal RK. 2008. Genetic divergence and correlations study in *Chlorophytum borivilianum*. J New Seed 9(4): 321-329. DOI: 10.1080/15228860802492182
Kumar K, Prasad Y, Mishra SB, Pandey SS, Kumar R. 2013. Study on genetic variability and correlation and path analysis with grain yield and yield attributing traits in green gram (*Vigna radiate* L. Wilczek). Bioscan 8: 1551-1555.
Kumar B, Mali H, Gupta E. 2014. Genetic variability, character association, and path analysis for economic traits in menthofuran rich half-sib seed progeny of *Mentha piperita* L. Biomed Res Int: 1-7.
Kumar S, Singh P, Sharma M. 2017. Genetic divergence and stability (AMMI) study in chickpea (*Cicer aeritinum* L.) under northwestern Himalayas of Jammu and Kashmir, India. Leg Res 40(5): 1-5.
Kuswanto, Waluyo B, Hardaningsih P. 2013. Segregation and selection of observed yardlong bean (*Vigna sesquipedalis* L. *fruwirth*) to get expected lines of purple pod. Int Res J Agric Sci Soil Sci 3: 88-92.

Kongjaimun A, Kaga A, Tomooka N, Somta P, Vaughan DA, Srinives P. 2012. The genetics of domestication of yardlong bean, *Vigna unguiculata* (L.) Walp. ssp. *unguiculata* cv.-gr. *sesquipedalis*. Ann Bot 109: 1185-1200. DOI: 10.1093/aob/mcs048

Lakshmi VGI, Sreedhar M, Vanisri S, Anantha MS, Rao LVS, Gireesh C. 2019. Multivariate analysis and selection criteria for identification of African rice (*Oryza glaberrima*) for genetic improvement of *indica* rice cultivars. Plant Genet Res 17(6): 499-505. DOI: 10.1017/S1479262119000327

Lever M, Krzywinski M, Altman N. 2017. Principal coordinate analysis. Nat Methods 14(7): 641-642.

Li Y, Wei YC, Li ZQ, Wang SH, Chung L. 2014. Relationship between progeny growth performance and molecular marker-based genetic distances in *Eucommia ulmoides* parental genotypes. Genet Mol Res 13: 4736-4746. DOI: 10.4238/2014.July.2.3.

Lima DC, Abreu AB, Ferreira CADC, Ramalho MAP. 2015. Breeding common bean populations for traits using selection index. Sci Agri 72: 132-137.

Londero PMG, Ribeiro ND, Filho AC, Rodrigues JA, Antunes IF. 2006. Heritability of dietary fiber content and grain yield in common bean populations. Pesq Agropec Bras 41: 51-58.

Mahender A, Anandan A, Pradan SK, Pandit E. 2016. Rice grain nutritional traits and their enhancement using relevant genes and QTLs through advanced approaches. SpringerPlus 5: 1-18. DOI: 10.1186/s40064-016-3744-6.

Meena OP, Bahadur V. 2015. Breeding potential of indeterminate tomato (*Solanum lycopersicum* L.) accessions using D² analysis. SABRAO J Bred Genet 47: 49-59.

Mellidou I, Krommydas K, Nianiou-Obediat I, Ouzounidou G, Kalivas A, Ganopoulos I. 2020. Exploring morpho-physiological profiles of a collection of tomato (*Solanum lycopersicum*) germplasm using multivariate statistics. Plant Genet Res 18: 88-97. DOI: 10.1017/S1479262120000088

Missanjo E, Matsumura J. 2017. Multiple trait selection index for simultaneous improvement of wood properties and growth traits in *Pinus kesiya* Royle ex Gordon in Malawi. Forests 8: 1-7. DOI: 10.3390/f8040096

Neyhart JL, Lorenz AJ, Smith KP. 2019. Multi-trait improvement by predicting genetic correlations in breeding crosses. G3 (Bethesda) 9 (10): 3153-3165. DOI: 10.1534/g3.119.400406

Obala J, Saxena RK, Singh VK, Kale SM, Garg V, Kumar CVS, Saxena KB, Tongoona P, Sibiya J, Varshney RK. 2020. Seed protein content and its relationships with agronomic traits in pigeon pea is controlled by both main and epistatic effects QTLs. Sci Rep 10: 1-17. DOI: 10.1038/s41598-019-56903-z

Oliveira GHF, Amaral CB, Silva FAM, Dutra SMF, Marconato MB, Moro GV. 2016. Mixed models and multivariate analysis for selection of superior maize genotypes. Chilean J Agric Res 76: 417-431.

Patel UV, Parmar VK, Patel PB, Malviya AV. 2016. Correlation and path analysis study in cowpea (*Vigna unguiculata* (L.) Walp). Int J Sci Env Tech 5: 3897-3904.

Pidigam S, Munnam SB, Nimmarajula S, Gonela N, Adimulam SS, Yadla H, Bandari L, Amarapalli G. 2019. Assessment of genetic diversity in yardlong bean (*Vigna unguiculata* (L.) Walp subsp. *sesquipedalis* Verdc.) germplasm from India using RAPD markers. Genet Resour Crop Evol 66: 1231-1242.

Pires LC, Machado TMM, De Araujo AM, Da Silva JBL, Euclydes RF, Costa MDS, Olson TA. 2013. Cluster evaluation of Brazilian and Moroccan goat populations using physical measurements. R Bras Zootec 42: 713-720.

Purnamasari I, Sobir, Syukur M. 2019. Diversity and inheritance in cowpea (*Vigna unguiculata)* on protein and yield components characters. Biodiversitas 20(5): 1294-1298. DOI: 10.13057/biodiv/d200507

R Core Team. 2017. R: A language and environment for statistical computing. https://www.R-project.org/.

Raivo K. 2019. Pheatmap: Pretty Heatmaps. https://cran.r-project.org/web/packages/pheatmap/index.html.

Rambabu E, Reddy KR, Kamala V, Saidaiah P, Pandravada SR. 2016a. Morphological characterization of yardlong bean under-exploited vegetable. Int J Sci Nat 7: 344-348.

Rambabu E, Reddy KR, Kamala V, Saidaiah P, and Pandravada SR. 2016b. Correlation and path analysis for quality, yield, and yield components in yardlong bean (*Vigna unguiculata* (L.) Walp. ssp. *sesquipedalis* Verdc.). Env Ecol 34: 1655-1661.

Rambabu E, Reddy KR, Kamala V, Saidaiah P, and Pandravada SR. 2016c. Genetic variability and heritability for quality, yield and yield components in yardlong bean (*Vigna unguiculata* (L.) Walp. ssp. sesquipedalis Verdc.). Green Farming 7: 63-67.

Rambabu E, Reddy KR, Kamala V, Saidaiah P, and Pandravada SR. 2016d. Genetic divergence for quality, yield, and yield components in yardlong bean [*Vigna unguiculata* (L.) Walp. ssp. *sesquipedalis* Verdc.]. Leg Res 3470: 1-5.

Reddy VR, Jabeen F. 2016. Narrow sense heritability, correlation, and path analysis in maize (*Zea mays* L.). SABRAO J Bred Genet 48: 120-126.

Silva TN, Moro GV, Moro FV, Dos Santos DMM, Buzinaro R. 2016. Correlation and path analysis of agronomic and morphological traits in maize. Cienc Agron 47: 351-357.

Salimi S, Moradi S. 2012. Effect the correlation, regression, and path analysis in soybean genotypes (*Glycine max* L.) under moisture and normal condition. Int J Agron Prod 3: 447-454.

Santos CAF, Da Costa DCC, Da Silva WR, Boiteux L. 2012. Genetic analysis of total seed protein content in two cowpea crosses. Crop Sci 55: 2501-2506.

Sanwal SK, Singh B, Singh B, Mann A. 2015. Multivariate analysis and its implication in breeding of desired plant type in garden pea (*Pisum sativum*). Indian J Agric Sci 85: 1298-1302.

Sarvamangala C, Gowda MVC, Varshney RK. 2011. Identification of quantitative trait loci for protein content, oil content, and oil quality for groundnut (*Arachis hypogaea* L.). Field Crop Res 122(1): 49-59. DOI: 10.1016/j.fcr.2011.02.010

Shegro A, Labuschagne MT, Shargie NG, Biljon A. 2013. Multivariate analysis of nutritional diversity in sorghum landrace accessions from western Ethiopia. J Biol Sci 13(2): 67-74.

Shukla S, Bhargava A, Chatterjee A, Pandey AC, Mishra BK. 2010. Diversity in phenotypic and nutritional traits in vegetable amaranth (*Amaranthus tricolor*), a nutritionally underutilized crop. J Sci Food Agric 90(1): 139-144. DOI: 10.1002/jsfa.3797

Singh RK, Chaudary BD. 1985. Biometrical Methods in Quantitative Genetics Analysis. Kalyani Publishers, New Delhi.

Stewart-Brown BB, Song Q, Vaughn JN, Li Z. 2019. Genomic selection for yield and seed composition traits within an applied soybean breeding program. G3 (Bethesda) 9(7): 2253-2265. DOI: 10.1534/g3.118.200917

Suanum Q, Somta P, Kongjaimun A, Yimram A, Kaga A, Tomooka N, Takahashi Y, Srinives P. 2016. Co-localization of QTLs for pod fiber content and pod shattering in F₂ and backcross populations between yardlong bean and wild cowpea. Mol Breed 36: 1-11. DOI: 10.1007/s11032-016-0505-8

Thorwarth P, Liu G, Ebmeyer E, Schacht J, Schachschneider R, Kazman E, Reif JC, Wurschum T, Longin CFH. 2019. Dissecting the genetics underlying the relationship between protein content and grain yield in a large hybrid wheat population. Theor App Genet 132(2): 489-500. DOI: 10.1007/s00122-018-3236-x

Voss-Fels KP, Stahl A, Hickey LT. 2019. Q & A: Modern crop breeding for future food security. BMC Biol 17: 18. DOI: 10.1186/s12915-019-0638-4

Vu VQ. 2020. A biplot based on ggplot2. https://github.com/vqv/ggbiplot.

Watcharatpong P, Kaga A, Chen X, Somta P. 2020. Narrowing down a major QTL region conferring pod fiber contents in yardlong bean (*Vigna unguiculata*), a vegetable cowpea. Genes (Basel) 11(4): 363. DOI: 10.3390/genes11040363.

Widyawan MH, Wulandary S, Taryono. 2020. Genetic diversity analysis of yardlong bean genotypes (*Vigna unguiculata* subsp. *sesquipedalis*) based on IRAP marker. Biodiversitas 21(3): 1101-1107. DOI: 10.13057/biodiv/d210333

Wills RBH, Wong AWK, Scriven FM, Greenfield H. 1994. Nutrient composition of Chinese vegetables. J Agric Food Chem 32: 413-416.

Woyann LG, Meira D, Zdiarski AD, Matei G, Milioli AS, Rosa AC, Madella LA, Benin G. 2019. Multiple trait selection of soybean for biodiesel production in Brazil. Ind Crops Prod 140: 1-7. DOI: 10.1016/j.indcrop.2019.111721

Yousuf M, Ajmal SU, Munir M, Ghaffor A. 2011. Genetic diversity analysis for agro-morphological and seed quality traits in rapeseed (*Brassica campestris* L.). Pak J Bot 43: 1195-1203.

Zanklan AS, Becker HC, Sorensen M, Pawelzik E, Gruneberg WJ. 2018. Genetic diversity in cultivated yam bean (*Pachyrizus* spp.) evaluated through multivariate analysis of morphological and agronomical traits. Genet Resour Crop Evol 65: 811-843. DOI: 10.1007/s10722-017-0582-5.