Character selection by path and principal analysis for enhanced seed size and yield in local castor bean (*Ricinus communis* L.)

Budi Waluyo¹*, Puji Shandila², Chindy Ulima Zanetta³, Darmawan Saptadi², Noer Rahmi Ardiarini², and Kuswanto³

1Department of Agronomy, Faculty of Agriculture, Universitas Brawijaya, Malang, Indonesia 65145
2Master Program of Agronomy, Faculty of Agriculture, Universitas Brawijaya, Malang, Indonesia 65145
3School of Life Sciences and Technology, Institut Teknologi Bandung, Bandung, Indonesia 40132

budiwaluyo@ub.ac.id

Abstract. Castor bean (*Ricinus communis* L.) is a commercial plant that widely uses in industry and manufacture. The aim of this study is to study the selection of seed size characteristics to increase yields and determine accessions that have a high yield. Twenty-two local castor accessions with two replications were conducted at Malang in 2017. Path analysis was done on 21 characters. The result showed that the seed diameter had a high positive direct effect on seed volume. The character that had a high direct effect on seed weight per plant is weight of inflorescence per plant. Seed diameter had a high positive direct effect on 100-seed weight. In order to identify the patterns of morphological variation, PCA was conducted. In this study, the three first components (PC1, PC2, and PC3) were contributed to 86.42% of the total variation. PC1 was strongly correlated with capsule length, capsule diameter, seed length, seed width, seed thickness, seed diameter, inflorescence length, and seed volume. PC2 varied as a measure of the days to flowering, days to harvesting, plant height, length of main stem, node number in main stem, cluster number. PC3 showed varied as inflorescence weight per plant, seed number per plant, inflorescence number per plant, capsule number per plant, seed weight per plant. The dendrograms obtained using the agglomerative hierarchical clustering generated three main clusters. The accessions that had higher sum of ranking in each cluster in each seed character contributed in first principal component were ASB60, ASB81, ASB22, LMG0216U1, SUKO0316, SUKO0216, LMG0516, TBN0316, TBN0816, LMG0216U2, and TBN1016.
1. Introduction
Castor bean (*Ricinus communis* L.) is one of the oil-producing crop plantations. Castor has not been widely cultivated in Indonesia, because many people do not know the usefulness. Castor bean initially used for traditional medicine, along with the development of technology and science-based research. Castor has been used for biofuel, biopharma, cosmetics, vegetable pesticides, textiles, soaps, paints, varnishes, inks, nylons and plastics (Gupta and Singh, 2015; Rana *et al*., 2012; Salihu *et al*., 2014). Production amount of castor bean in 2015 was reached 1.5 thousand tons (BPS, 2015), but it cannot fulfill domestic needs.

Currently, castor bean was neglected crops. Cultivation of castor bean has not been done intensively. Many people do not know this plant due to a lack of awareness of the importance of castor bean. Another problem faced is low-productivity. Based on data from BPS (2015), the production in Indonesia between 2010 and 2014 tends to decrease 1.8 thousand tons and 1.4 thousand tons, respectively. Production of castor bean is lower than other plantation crops such as sugar, sugar cane, fragrance, patchouli, and tobacco.

One of plant breeding activities that can be done to improve the yield is selection, either direct to yield or indirect through several characters related to yield. In indirect selection, the character chosen for selection criteria should be based on the close relationship with the desired character. The result of study showed that many characters have a mild-to-strong connection from one to another (Kearsey and Pooni, 1996). Although studies of correlation are helpful in determining the components of yield, it does not provide an extent of contributions made by the number of independent traits. Path coefficient analysis is helpful in partitioning the correlation into components due to direct and indirect effects (Rana *et al*., 2015). Principal component analysis (PCA) is a mathematical technique used to categorize a large number of characters into major components and assess their contribution to the total variation. PCA was carried out to determine the relationship between characters and find the characters which varied together, thus segregating the characters (Singh *et al*., 2016, Singh *et al*., 2016). Grouping of local castor bean has been done based on specific morphology characters (Ardiarini *et al*., 2017). The purpose of this research was to study the selection of seed size characteristics to increase yields and determine accessions that have high yield.

2. Materials and methods
The research was carried out in April – October 2017 at the farmer field, Karangploso, Malang Regency. The research was set in a randomized block design. The experimental material comprised of 22 lines of castor bean with two replications. It measured 10 plants in each replication with spacing plant 90 cm x 50 cm.

Twenty-one characters were observed based on Descriptor Draft National Guidelines for the Conduct of Tests for Distinctness, Uniformity, and Stability Castor Bean (*Ricinus communis* L.) (Chakrabarty *et al*., 2006) and Descriptor of Castor Bean (UPOV, 2016). The characters that observed were time of the beginning of flowering, days to harvesting, plant height, length of main stem, number of nodes on main stem, length of internode, capsule length, capsule diameter, seed length, seed width, seed thickness, seed diameter, weight of inflorescence per plant, number of seeds per plant, number of inflorescences per plant, length of inflorescence, number of clusters, number of capsule per plant, weight of seed per plant, 100-seed weight and seed volume.
Table 1. Castor bean accessions and the origins

| No | Accession | Origin |
|----|-----------|--------|
| 1  | TBN0816   | Tuban  |
| 2  | TBN0516   | Tuban  |
| 3  | TBN0116   | Tuban  |
| 4  | TBN0716   | Tuban  |
| 5  | TBN0316   | Tuban  |
| 6  | TBN0416   | Tuban  |
| 7  | TBN0916   | Tuban  |
| 8  | TBN0616   | Tuban  |
| 9  | TBN1016   | Tuban  |
| 10 | TBN0216   | Tuban  |
| 11 | GRT0116   | Garut  |
| 12 | LMG0216U1 | Lamongan|
| 13 | LMG0216U2 | Lamongan|
| 14 | LMG0316   | Lamongan|
| 15 | SMG0316   | Semarang|
| 16 | SUKO0116  | Sukoharjo|
| 17 | SUKO0216  | Sukoharjo|
| 18 | SUKO0316  | Sukoharjo|
| 19 | ASB22     | Asembagus|
| 20 | ASB81     | Asembagus|
| 21 | ASB60     | Asembagus|
| 22 | LMG0516   | Lamongan|

2.1 Data analysis
2.1.1 Path analysis. To determine correlation between observed characters is used as an approach of (Singh and Chaudhary, 1979):

\[ r_{xy} = \frac{\text{Cov}_{g}(XY)}{\sqrt{\text{V}_{g}X \cdot \text{V}_{g}Y}} \]

Where:
- \( r \) = correlation
- \( \text{Cov} \) = covariance X and Y
- \( X \) = variance X
- \( Y \) = variance Y
The direct effect of yield components to yield calculated by path analysis using a matrix method (Singh and Chaudhary, 1979):

\[
\begin{bmatrix}
 r_1 \\
r_2 \\
r_3
\end{bmatrix} =
\begin{bmatrix}
 r_{11} & r_{12} & r_{13} \\
r_{21} & r_{22} & r_{23} \\
r_{31} & r_{32} & r_{33}
\end{bmatrix}
\begin{bmatrix}
 p_1 \\
p_2 \\
p_3
\end{bmatrix}
\text{ or } R_y = Rx.P_i
\]

Based on above equation, the value of Pi (direct influence) can be calculated using following formula:

\[
P_i = \frac{p_i}{g_{1844}/g_{1876}} - 1.
\]

Where :
- \( R_x \) = Correlation matrix between independent variable
- \( R_x^{-1} \) = Inverse matrix \( R_x \)
- \( P_i \) = Path coefficient showing the direct effect of each independent variable on the dependent variable
- \( R_y \) = Correlation coefficient between independent variable \( X_1 \) and the dependent variable

While to look for residual value or influence of unobserved properties (P residual) are as follows:

\[
P_{\text{residual}} = \sqrt{1 - (r_{1y}p_{1y} + r_{2y}p_{2y} + r_{3y}p_{3y} + \cdots + r_{ny}p_{ny})}
\]

Data processing and data analysis using OPSTAT (Online Statistical Analysis Tools) (Sheoran et al., 1998).

### 2.1.2 Principal component analysis (PCA).

Variability analysis was performed on quantitative characters determined by principal component analysis (PCA) with Pearson correlation type. The PCA method is usually used to identify significantly different variables in the data. Quantitative characters are expressed in mean values. The character that contributes to the variability is determined based on factor loading value \( >\pm 0.6 \). Number of principal components that affect the total variability is determined by eigenvalue \( >1 \) (Woolford, 2015).

Genetic distance based on cluster analysis is used to determine the proximity of genetic distance between plant taxon. Genetic distances can be grouped by the un-weighted pair-group method average (UPGMA) agglomeration method and the Pearson correlation coefficient. Genetic distance is displayed in the form of dendrogram that shows the value of similarity. Dendrogram in the form of branched lines, branches located at the same point 1 called a cluster. One cluster formed shows the similarity of plant character. Based on an estimation of character similarity, the accession that belongs to one cluster has close genetic distance. The more clusters formed, the more varied characters are analyzed. It was performed analysis of data using Microsoft® Excel 2007/XLSTAT Version 2009.3.02.

### 3. Results

The extent of correlation coefficient are shown in Table 1, path analysis is presented in Table 2, Table 3, Table 4. While principal component analysis is shown in Table 5 and dendrogram showed in Figure 1.
3.1 Path Analysis

Based on Table 1, there were a significant correlation among seed volume, 100-seeds weight, and seed weight per plant, to other characters. A positive correlation indicates that the increase of one character values causes increasing to another character, whereas a negative correlation indicates that an increase in the value of one character leads to the decrease of another character (Freeman and Julious, 2010). While the determination of characters that can be used as effective selection criteria can be seen from the magnitude of the direct effect on the result (Pi), the correlation between character and result (rx;y) and the difference between characters and results with direct effect of the character on the yield value is less than 0.05 (Syukur et al., 2010).

### Table 2. Correlation coefficient between characters and weight of seed per plant, 100 seed weight and seed volume

| Characters                           | Weight of seed per plant (g) | 100 seed weight (g) | Seed volume (mm$^3$) |
|--------------------------------------|------------------------------|---------------------|----------------------|
| Time of beginning of flowering (dap) | -0.11                        | -0.05               | 0.06                 |
| Days to harvesting (dap)             | -0.06                        | 0.10                | 0.36                 |
| Plant height (cm)                    | 0.21                         | -0.17               | 0.06                 |
| Length of main stem (cm)             | 0.11                         | -0.08               | 0.12                 |
| Number of nodes in main stem         | 0.15                         | 0.02                | 0.17                 |
| Length of internode (cm)             | -0.03                        | -0.37               | -0.32                |
| Capsule length (mm)                  | 0.23                         | 0.51*               | 0.96*                |
| Capsule diameter (mm)                | 0.38                         | 0.53*               | 0.94*                |
| Seed length (mm)                     | 0.26                         | 0.51*               | 0.98*                |
| Seed width (mm)                      | 0.34                         | 0.49*               | 0.94*                |
| Seed thickness (mm)                  | 0.32                         | 0.58*               | 0.99*                |
| Seed diameter (geometric)(mm)        | 0.31                         | 0.54*               | 0.99*                |
| Weight of inflorescence per plant (g)| 0.96*                        | 0.01                | 0.16                 |
| Number of seeds per plant            | 0.89*                        | -0.14               | -0.13                |
| Number of inflorescence per plant    | 0.80*                        | -0.04               | 0.02                 |
| Length of inflorescence (cm)         | 0.55*                        | 0.10                | 0.41                 |
| Number of clusters                   | 0.10                         | -0.17               | -0.02                |
| Number of capsule per plant          | 0.85*                        | -0.12               | -0.17                |
| Weight of seed per plant (g)         | 1.00                         | 0.21                | 0.30                 |
| 100-seed weight (g)                  | 0.21                         | 1.00                | 0.54*                |
| Seed volume (mm$^3$)                 | 0.30                         | 0.54*               | 1.00                 |

Note: * is significant different at 5% level.

Seed weight per plant had a significant positive correlation with weight of inflorescence per plant, number of seed per plant, number of inflorescence per plant, length of inflorescence and number of
capsule per plant. The highest correlation coefficient was between weight of seed per plant and weight of inflorescence per plant ($r = 0.96$). The high direct effect found on weight of inflorescence per plant. Results of the analysis showed that the residue obtained is 0.00012. This value defines the number of characters that represent observed to determine the effect of direct and indirect between the characters of the character of the results. It shows a path analysis model using matrix above can explain the relationship of components affecting the total weight of the fruit of 99.98%, so there are 0.012% other characters that have not been described in the model used.

Weight of 100 seed had a significant positive correlation with capsule length, capsule diameter, seed length, seed width, seed thickness, and seed volume. The highest correlation coefficient was between 100-seed weight and seed thickness ($r = 0.58$). Path analysis that is shown in Table 3, the high direct effect found on seed diameter. Results of the analysis showed that the residue obtained is 0.00189.

Seed volume had a significant positive correlation with capsule length, capsule diameter, seed length, seed width, seed thickness, seed diameter, and 100 seed weight. The highest correlation coefficient was between seed volume and seed diameter ($r = 0.99$). Based on Table 2 seed thickness have high direct effect value. The analysis showed that the residue is 0.00002.

Seed character so important. So it is important to know characters that have a relation with seed character. The advantage of physical characteristics of seeds can be used as the basis for selection of raw materials industry and biodiesel (Waluyo et al., 2017).
|   | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
|---|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|
| 1 | -0.27 | -0.05 | 0.02 | 0.04 | 0.01 | -0.01 | 0.01 | 0.00 | 0.00 | -0.02 | 0.00 | 0.09 | -0.14 | 0.10 | 0.01 | -0.03 | -0.05 | 0.10 | -0.06 | -0.01 | 0.12 |
| 2 | -0.23 | -0.06 | 0.02 | 0.04 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 | -0.06 | -0.03 | 0.03 | 0.03 | -0.21 | -0.51 | 0.10 | 0.01 | -0.03 | -0.07 | 0.09 | 0.08 |
| 3 | -0.16 | -0.03 | 0.04 | 0.03 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 | -0.01 | 0.00 | 0.00 | 0.00 | -0.01 | 0.10 | 0.11 | 0.00 | -0.02 | 0.08 | 0.11 |
| 4 | 0.04 | -0.05 | 0.02 | 0.04 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 | -0.01 | -0.22 | 0.03 | 0.00 | -0.02 | -0.22 | 0.02 | 0.00 | -0.02 | -0.02 | 1.11 |
| 5 | 0.11 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.01 | 0.01 | 0.02 | 0.07 | 0.03 | 0.00 | 0.08 | 0.03 | 0.01 | 0.01 | 0.03 | 0.01 | 0.36 |
| 6 | -0.09 | -0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.36 |
| 7 | 0.01 | -0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.05 | -0.66 | -0.19 | 0.21 | 0.00 | -0.06 | -0.01 | 0.12 |
| 8 | 0.01 | 0.01 | 0.00 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.00 | -0.65 | -0.17 | 0.21 | 0.00 | -0.06 | -0.01 | 0.00 | 0.12 |
| 9 | -0.02 | -0.03 | 0.00 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.00 | -0.67 | -1.23 | 0.30 | 0.01 | 0.00 | -0.06 | 0.00 | 0.05 |
| 10 | -0.07 | -0.03 | 0.00 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.00 | -0.64 | -1.20 | 0.16 | 0.00 | 0.00 | -0.06 | 0.00 | 0.03 |
| 11 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.00 | -0.64 | -1.20 | 0.16 | 0.00 | 0.00 | -0.06 | 0.00 | 0.03 |
| 12 | 0.03 | -0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.68 | -1.26 | 0.12 | 0.01 | 0.00 | -0.05 | 0.07 | 0.01 |
| 13 | 0.04 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.71 | -0.22 | 0.71 | 0.05 | 0.07 | 0.07 | 0.01 | 0.25 |
| 14 | 0.04 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.71 | -0.22 | 0.71 | 0.05 | 0.07 | 0.07 | 0.01 | 0.25 |
| 15 | 0.11 | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.71 | -0.22 | 0.71 | 0.05 | 0.07 | 0.07 | 0.01 | 0.25 |
| 16 | -0.10 | -0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.71 | -0.22 | 0.71 | 0.05 | 0.07 | 0.07 | 0.01 | 0.25 |
| 17 | -0.23 | -0.05 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.71 | -0.22 | 0.71 | 0.05 | 0.07 | 0.07 | 0.01 | 0.25 |
| 18 | 0.06 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.71 | -0.22 | 0.71 | 0.05 | 0.07 | 0.07 | 0.01 | 0.25 |
| 19 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.71 | -0.22 | 0.71 | 0.05 | 0.07 | 0.07 | 0.01 | 0.25 |
| 20 | 0.01 | -0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.71 | -0.22 | 0.71 | 0.05 | 0.07 | 0.07 | 0.01 | 0.25 |
| 21 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.71 | -0.22 | 0.71 | 0.05 | 0.07 | 0.07 | 0.01 | 0.25 |

Note: Diagonal and bold values are the direct effect coefficient values. (1) Time to flowering (2) Days to harvesting (3) Plant height, (4) Length of main stem, (5) Number of nodes in main stem, (6) Length of internode, (7) Capsule length, (8) Capsule diameter, (9) Seed length, (10) Seed width, (11) Seed thickness, (12) Seed diameter, (13) Weight of inflorescence per plant, (14) Number of seeds per plant, (15) Number of capsules per plant, (16) Length of capsule, (17) Number of clusters, (18) Number of seedlings per plant, (19) Weight of seed per plant, (20) 100 seed weight, (21) Seed volume.
Table 4. Path analysis of characters to 100 seed weight

|   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  | 14  | 15  | 16  | 17  | 18  | 19  | 20  | 21  |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1 | 2.68| -0.59| -0.20| -0.95| -1.52| -0.02| 0.07| -0.02| -0.11| -0.64| 0.01| 1.29| 0.92| -0.23| -0.37| 0.48| -0.06| -0.08| -0.40| -0.31|
| 2 | 2.24| -0.70| -0.18| -0.87| -1.39| -0.01| -0.37| 0.07| -0.80| -1.12| 0.13| 4.62| 0.92| -0.36| -0.36| 0.71| -0.06| -0.11| -0.22| -2.04|
| 3 | 1.55| -0.36| -0.35| -0.75| -1.07| -0.02| -0.07| 0.00| -0.11| -0.39| 0.00| 0.93| -1.04| 0.34| -0.05| 0.80| -0.06| 0.03| 0.74| -0.31|
| 4 | 2.41| -0.58| -0.25| -1.05| -1.62| -0.02| -0.04| 0.00| -0.26| -0.67| 0.04| 1.96| -0.31| 0.14| -0.24| 0.77| -0.07| -0.01| 0.38| -0.67|
| 5 | 2.43| -0.58| -0.22| -1.02| -1.67| -0.01| -0.06| 0.01| -0.35| -0.82| 0.06| 2.55| -0.46| 0.17| -0.24| 0.72| -0.07| 0.00| 0.54| -0.94|
| 6 | 0.86| -0.13| -0.14| -0.36| -0.45| -0.05| 0.41| -0.10| 0.61| 0.60| -0.14| -3.42| -0.29| 0.26| -0.08| 0.33| -0.04| 0.03| -0.09| 1.82|
| 7 | -0.14| -0.20| -0.02| -0.03| -0.08| 0.02| -1.31| 0.29| -1.84| -2.09| 0.40| 10.65| -0.65| -0.33| 0.00| 0.53| 0.01| -0.08| 0.81| -5.43|
| 8 | -0.13| -0.16| 0.00| -0.01| -0.07| 0.01| -1.25| 0.31| -1.75| -2.20| 0.40| 10.54| -1.95| -0.02| 0.19| 0.56| 0.01| -0.01| 1.37| -5.32|
| 9 | 0.16| -0.29| -0.02| -0.15| -0.31| 0.02| -1.26| 0.28| -1.92| -2.22| 0.41| 11.06| -0.81| -0.29| -0.02| 0.58| 0.00| -0.08| 0.93| -5.57|
| 10| 0.68| -0.31| -0.05| -0.28| -0.55| 0.01| -1.10| 0.27| -1.70| -2.50| 0.39| 10.80| -1.52| -0.08| 0.03| 0.58| -0.01| -0.04| 1.21| -5.33|
| 11| 0.04| -0.21| 0.00| -0.09| -0.24| 0.02| -1.24| 0.29| -1.85| -2.31| 0.42| 11.13| -1.14| -0.21| 0.04| 0.47| 0.01| -0.06| 1.14| -5.63|
| 12| 0.31| -0.29| -0.03| -0.18| -0.38| 0.01| -1.24| 0.29| -1.88| -2.39| 0.42| 11.29| -1.14| -0.21| 0.01| 0.57| 0.00| -0.06| 1.10| -5.66|
| 13| -0.37| 0.10| -0.06| -0.05| -0.12| 0.00| -0.13| 0.09| -0.23| -0.57| 0.07| 1.95| -6.64| 1.64| 0.79| 0.67| 0.00| 0.36| 3.42| -0.91|
| 14| -0.35| 0.15| -0.07| -0.08| -0.16| -0.01| 0.25| 0.00| 0.32| 0.12| -0.05| -1.35| -6.24| 1.75| 0.75| 0.53| -0.01| 0.39| 3.17| 0.76|
| 15| -1.12| 0.29| 0.02| 0.28| 0.45| 0.00| 0.00| 0.07| 0.05| -0.09| 0.02| 0.12| -5.91| 1.47| 0.89| 0.29| 0.02| 0.35| 2.87| -0.09|
| 16| 0.99| -0.38| -0.21| -0.63| -0.92| -0.01| -0.54| 0.13| -0.85| -1.11| 0.15| 4.92| -3.40| 0.71| 0.20| 1.30| -0.05| 0.13| 1.96| -2.30|
| 17| 2.22| -0.53| -0.25| -0.95| -1.44| -0.02| 0.08| -0.03| -0.04| -0.29| -0.03| 0.34| -0.39| 0.22| -0.27| 0.82| -0.08| 0.01| 0.37| 0.10|
| 18| -0.54| 0.20| -0.03| 0.02| -0.02| 0.00| 0.27| 0.00| 0.37| 0.27| -0.06| -1.80| -6.11| 1.71| 0.78| 0.41| 0.00| 0.39| 3.04| 0.99|
| 19| -0.30| 0.04| -0.07| -0.11| -0.25| 0.00| 0.12| 0.50| -0.85| 0.14| 3.49| -6.36| 1.55| 0.72| 0.72| -0.01| 0.34| 3.57| -1.72|
| 20| 0.15| -0.25| -0.02| -0.12| -0.28| 0.02| -1.25| 0.29| -1.88| -2.35| 0.42| 11.25| -1.06| -0.23| 0.02| 0.53| 0.00| -0.07| 1.08| -5.68|

Note: Diagonal and bold are the direct effect coefficient values. (1) Time of beginning of flowering, (2) Days to harvesting, (3) Plant height, (4) Length of main stem, (5) Number of nodes in main stem, (6) Length of internode, (7) Capsule length, (8) Capsule diameter, (9) Seed length, (10) Seed width, (11) Seed thickness, (12) Seed diameter, (13) Weight of inflorescence per plant, (14) Number of seeds per plant, (15) Number of inflorescences per plant, (16) Length of inflorescence, (17) Number of clusters, (18) Number of capsule per plant, (19) Weight of seed per plant, (21) Seed volume.
Table 5. Path analysis characters to seed volume

|    | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1  | -0.11 | 0.11 | 0.01 | 0.08 | -0.05 | 0.00 | 0.00 | -0.01 | 0.01 | 0.06 | 0.01 | -0.03 | -0.03 | 0.02 | 0.07 | -0.02 | -0.07 | 0.05 | -0.04 | 0.00 |
| 2  | -0.09 | 0.13 | 0.01 | 0.07 | -0.05 | 0.00 | 0.01 | 0.04 | 0.10 | 0.14 | -0.11 | -0.03 | 0.03 | 0.07 | -0.03 | -0.06 | 0.07 | -0.02 | -0.01 |
| 3  | -0.06 | 0.07 | 0.01 | 0.06 | -0.04 | 0.00 | 0.00 | 0.00 | 0.01 | 0.03 | 0.00 | -0.02 | 0.04 | -0.03 | 0.01 | -0.03 | -0.06 | -0.02 | 0.07 | 0.01 |
| 4  | -0.10 | 0.11 | 0.01 | 0.09 | -0.05 | 0.00 | 0.00 | 0.00 | 0.03 | 0.06 | 0.04 | -0.05 | 0.01 | -0.01 | 0.04 | -0.03 | -0.07 | 0.00 | 0.04 | 0.00 |
| 5  | -0.10 | 0.11 | 0.01 | 0.08 | -0.06 | 0.00 | 0.00 | 0.01 | 0.05 | 0.07 | 0.07 | -0.06 | 0.02 | -0.02 | 0.04 | -0.03 | -0.07 | 0.00 | 0.05 | 0.00 |
| 6  | -0.03 | 0.02 | 0.01 | 0.03 | -0.02 | -0.01 | -0.01 | -0.05 | -0.08 | -0.05 | -0.16 | 0.08 | 0.01 | -0.02 | 0.01 | -0.01 | -0.04 | -0.02 | -0.01 | 0.02 |
| 7  | 0.01 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 | 0.15 | 0.23 | 0.21 | 0.44 | -0.26 | 0.07 | 0.00 | -0.03 | -0.02 | 0.01 | 0.00 | 0.13 | -0.03 |
| 8  | 0.01 | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 | 0.16 | 0.22 | 0.20 | 0.44 | -0.26 | 0.07 | 0.00 | -0.03 | -0.02 | 0.01 | 0.00 | 0.13 | -0.03 |
| 9  | -0.01 | 0.05 | 0.00 | 0.01 | -0.01 | 0.00 | 0.03 | 0.14 | 0.24 | 0.20 | 0.45 | -0.27 | 0.03 | 0.03 | 0.00 | -0.02 | 0.00 | 0.05 | 0.09 | -0.03 |
| 10 | -0.03 | 0.06 | 0.00 | 0.02 | -0.02 | 0.00 | 0.02 | 0.14 | 0.22 | 0.22 | 0.43 | -0.26 | 0.05 | 0.01 | -0.01 | -0.02 | -0.01 | 0.03 | 0.11 | -0.03 |
| 11 | 0.00 | 0.04 | 0.00 | 0.01 | -0.01 | 0.00 | 0.03 | 0.15 | 0.23 | 0.20 | 0.47 | -0.27 | 0.04 | 0.02 | -0.01 | -0.02 | 0.01 | 0.04 | 0.11 | -0.03 |
| 12 | -0.01 | 0.05 | 0.00 | 0.02 | -0.01 | 0.00 | 0.03 | 0.15 | 0.24 | 0.21 | 0.46 | -0.28 | 0.04 | 0.02 | 0.00 | -0.02 | 0.00 | 0.04 | 0.10 | -0.03 |
| 13 | 0.02 | -0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.05 | 0.03 | 0.05 | 0.08 | -0.05 | 0.24 | -0.15 | -0.14 | -0.03 | -0.01 | -0.23 | 0.32 | 0.00 |
| 14 | 0.01 | -0.03 | 0.00 | 0.01 | -0.01 | 0.00 | -0.01 | 0.00 | -0.04 | -0.01 | -0.06 | 0.03 | 0.22 | -0.16 | -0.13 | -0.02 | -0.01 | -0.25 | 0.30 | 0.01 |
| 15 | 0.05 | -0.05 | 0.00 | 0.02 | 0.02 | 0.00 | 0.00 | 0.03 | -0.01 | 0.01 | 0.02 | 0.00 | 0.21 | -0.13 | -0.16 | -0.01 | 0.02 | -0.22 | 0.27 | 0.00 |
| 16 | -0.04 | 0.07 | 0.01 | 0.05 | -0.03 | 0.00 | 0.01 | 0.07 | 0.11 | 0.10 | 0.17 | -0.12 | 0.12 | -0.07 | -0.04 | -0.05 | -0.05 | -0.08 | 0.18 | -0.01 |
| 17 | -0.09 | 0.10 | 0.01 | 0.08 | -0.05 | -0.01 | 0.00 | -0.02 | 0.00 | 0.03 | -0.03 | -0.01 | 0.01 | -0.02 | 0.05 | -0.03 | -0.08 | 0.00 | 0.04 | 0.01 |
| 18 | 0.02 | -0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | -0.05 | -0.02 | -0.07 | 0.04 | 0.22 | -0.16 | -0.14 | -0.02 | 0.00 | -0.25 | 0.28 | 0.01 |
| 19 | 0.01 | -0.01 | 0.00 | 0.01 | -0.01 | 0.00 | 0.01 | 0.06 | 0.06 | 0.08 | 0.15 | -0.09 | 0.23 | -0.14 | -0.13 | -0.03 | -0.01 | -0.22 | 0.33 | -0.01 |
| 20 | 0.01 | 0.01 | 0.00 | -0.01 | 0.00 | 0.00 | 0.01 | 0.08 | 0.12 | 0.11 | 0.27 | -0.15 | 0.00 | 0.02 | 0.01 | -0.01 | 0.01 | 0.03 | 0.07 | -0.06 |

Note: Diagonal and bold are the direct effect coefficient values. (1) Time of beginning of flowering, (2) Days to harvesting, (3) Plant height, (4) Length of main stem, (5) Number of nodes in main stem, (6) Length of internode, (7) Capsule length, (8) Capsule diameter, (9) Seed length, (10) Seed width, (11) Seed thickness, (12) Seed diameter, (13) Weight of inflorescence per plant, (14) Number of seeds per plant, (15) Number of inflorescences per plant, (16) Length of inflorescence, (17) Number of clusters, (18) Number of capsule per plant, (19) Weight of seed per plant, (20) 100-seed weight.
3.2 Principle component analysis

The principal component analysis (PCA) was performed to assess the variation in 21 characters. Based on factor loadings of these characters, first three components of PCA explained 86.420% of the total variation as shown in Table 5. The projection of characters on three principal components is shown in Figure 1. The first principal component (PC1) that have eigenvalue 7.899 was contributing 37.616% of the total variation with the characters on capsule length, capsule diameter, seed length, seed width, seed thickness, seed diameter, length of inflorescence and seed volume. The principal component second (PC2) had eigenvalue 5.382 and accounted for 25.629% of total variation with the contributed characters on the time of beginning of flowering, days to harvesting, plant height, length of main stem, number of nodes on main stem and number of clusters. The principal component third (PC3) had eigenvalue 4.867 and accounted for 23.175% of total variation with the contributed characters weight of inflorescence per plant, number of seeds per plant, number of inflorescence per plant, number of capsule per plant and weight of seed per plant.

Table 6. Minimum, maximum, mean, standard deviation, and principal component analysis of castor bean characters

| Characters and Components | Min    | Max    | Mean   | S. Dev | PC1    | PC2    | PC3    |
|---------------------------|--------|--------|--------|--------|--------|--------|--------|
| Eigenvalue                | -      | -      | -      | -      | 7.90   | 5.38   | 4.87   |
| Variability (%)           | -      | -      | -      | -      | 37.62  | 25.63  | 23.18  |
| Cumulative %              | -      | -      | -      | -      | 37.62  | 63.25  | 86.42  |
| Time of beginning of flowering | 59.03 | 111.04 | 74.64  | 15.90  | 0.31   | 0.87*  | -0.09  |
| Days to harvesting        | 107.75 | 161.88 | 127.46 | 18.41  | 0.57   | 0.71*  | -0.17  |
| Plant height              | 80.75  | 229.80 | 128.89 | 43.28  | 0.29   | 0.69*  | 0.27   |
| Length of main stem       | 54.13  | 186.73 | 91.44  | 35.71  | 0.40   | 0.88*  | 0.12   |
| Number of nodes in main stem | 11.20 | 24.13  | 15.51  | 4.01   | 0.44   | 0.83*  | 0.12   |
| Length of internode       | 6.47   | 12.95  | 8.77   | 1.69   | -0.18  | 0.56   | 0.21   |
| Capsule length            | 15.44  | 21.59  | 16.94  | 1.65   | 0.89*  | -0.32  | -0.20  |
| Capsule diameter          | 14.32  | 18.29  | 15.41  | 1.03   | 0.89*  | -0.37  | -0.02  |
| Seed length               | 11.90  | 16.46  | 13.20  | 1.31   | 0.94*  | -0.23  | -0.18  |
| Seed width                | 6.41   | 9.06   | 7.58   | 0.62   | 0.94*  | -0.11  | -0.08  |
| Seed thickness            | 4.81   | 6.81   | 5.37   | 0.53   | 0.93*  | -0.29  | -0.14  |
| Seed diameter             | 7.26   | 10.03  | 8.11   | 0.74   | 0.96*  | -0.22  | -0.13  |
| Weight of inflorescence per plant | 161.80 | 689.25 | 244.14 | 112.98 | 0.26   | -0.19  | 0.94*  |
| Number of seeds per plant | 168.33 | 1206.50| 404.06 | 209.79 | 0.00   | -0.06  | 0.99   |
| Number of inflorescence per plant | 1.63  | 21.50  | 4.74   | 4.01   | 0.03   | -0.45  | 0.85*  |
| Length of inflorescence   | 20.56  | 48.10  | 32.50  | 7.46   | 0.63*  | 0.38   | 0.47   |
| Number of clusters        | 9.00   | 32.60  | 16.79  | 6.30   | 0.27   | 0.91*  | 0.17   |
| Number of capsule per plant | 59.40 | 550.50 | 159.31 | 95.81  | -0.06  | -0.15  | 0.97*  |
| Weight of seed per plant  | 44.04  | 287.36 | 98.28  | 52.86  | 0.41   | -0.19  | 0.87*  |
| 100 seed weight           | 19.06  | 68.43  | 28.77  | 11.82  | 0.52   | -0.33  | -0.18  |
| Seed volume               | 200.06 | 528.20 | 286.16 | 87.21  | 0.94*  | -0.27  | -0.15  |
Note: *) character that has a contribution to principal component

Analysis of genetic distance based on cluster analysis is used to determine the proximity of genetic distance between taxon. Genetic distances can be grouped by the Un-weighted Pair-Group Method Average (UPGMA) agglomeration method and the Pearson correlation coefficient. Pearson correlation is a simple correlation that is useful to standardize the data to be analyzed, so it can determine the group of each pair of characters that connects between traits. Analysis of genetic distance for 21 quantitative characters at 22 local castor accessions is shown on the dendrogram (Figure 1). The results showed that Pearson correlation similarity coefficient divide genotypes into 3 cluster at coefficient 0.92, Cluster 1 consist of 6 accessions, that is ASB60, ASB81, ASB22, SMG0316, SUKO0116, and GRT0116. Cluster 2 consist of 1 accession is LMG 0216U1. Cluster 3 consists of 15 accessions, namely SUKO0316, SUKO0216, LMG0516, TBN0316, TBN0816, TBN0516, TBN0216, TBN0416, TBN0716, TBN0916, LMG0316, LMG0216U2, TBN0616, TBN0116, and TBN1016.

Figure 1. Dendrogram of 22 accessions local castor bean based on 21 morphological characters

Based on PCA and each character mean, information is obtained that the physical characteristics of seeds have high variability (Table 6). Thus, based on this character selection can be made to potential accessions. The accession position which is divided into clusters shows the uniqueness of each accession. Therefore, in each cluster, the best performance can be chosen based on the number of accession rankings arranged based on the length of the seeds, seed width, seed thickness, seed diameter (geometric), seed volume, and weight of the planting seeds. Selected accessions based on the number of the best ranking for each character was ASB60, ASB81, ASB22, LMG0216U1, SUKO0316, SUKO0216, LMG0516, TBN0316, TBN0816, LMG0216U2, and TBN1016.
4. Discussion

This research showed that volume of seed, weight of 100 seeds, and weight of seeds per plant is supported by other characters that affect. It can be seen from the correlation values. Correlation coefficient describes a simple relationship between the characters. Explained that correlation becomes the direct effect and indirect effect. Path analysis could explain how far each character is assigned a value to the main character either directly or indirectly. The direct effect is the influence of a character to another character directly without character intermediaries. While the indirect effect gives influence to a character through other characters.

Characters with a high value of direct effect can be used as an identifier character selection. The residual value is the total value of the direct influence of the rest of the characters measured yet unidentified. The residual value nearly zero means that the path analysis is used more effectively explain the cause and effect of the correlation values and character were observed more complete to explain the values of its direct or indirect influence.

The principal component analysis (PCA) is a multivariate analysis used to identify the character of the plant that contributes to the variation in the genotype. Based on the result of PCA, the character that contributes to the variability is determined based on factor loading value >±0.6. The result of the main group is used to identify a character that characterizes a variety (Afuape, 2011). The principal component analysis revealed that the first principal component strongly correlated with capsule length, capsule diameter, seed length, width seed, thickness seed, and seed volume. The second principal component varied as a measure of weight of inflorescence per plant, number of seeds per plant, seed weight per plant, number of inflorescence per plant, number of capsules per plant and yield per plot. The third principal component showed varied morphological characters (plant height, number of nodes on main stem), physiological (time of beginning of flowering, days to harvesting) and yield (number of clusters). In our study, the three components contributed 86.747% of the total variation. Characters that contribute to diversity are characters that have the largest and positive vector values whereas negative values have a negative contribution (Arif et al., 2015). Associations of these traits may lead to take effect seed size and yield. This correlation also can be explained in a way that total fluctuations in yield are governed principally by changes in one or more component. This condition which indicated these characters were potential for a simultaneous selection scheme in more than one trait for increasing castor bean seed size dan yield.

Based on the quantitative characters, the Pearson correlation coefficient of similarity between devide accessions to 3 cluster at 0.92. The similarity value obtained is higher then the genetic distance between accessions is closer (Dewi et al., 2013). Characteristic similarity analysis among accessions is useful for accurately interpreting the genetic distance so as to design new genotypes efficiently (Kantiet al., 2015).

The accessions in this research were taken from different regions. Accessions from the same region are not always in the same cluster. The cluster grouping is not based on the origin region but the similarity of the characters, so it is likely that accessions from other regions at same cluster. Dendrogram of the genetic distance gives information that the lower value of similarity, the more differences that will cause diversity between accessions, so the value of genetic distance is high and vice versa.

Acknowledgments

The authors gratefully acknowledge to Fefi Ramayanti, SP., and Ainur Rofiq Edy Kurniawan, SP., that helped on observation, measurement, and collect of the data during the research.
References
1. Afuape, 2011. Multivariate assessment of the agromorphological variability and yield components among sweet potato (*Ipomoea batatas* (L.) Lam) landraces. African J. Plant Sci. 5, 123–132.
2. Ardiarini, N.R., Waluyo, B., Saptadi, D., Kuswanto, 2017. Pengelompokan galur-galur jarak kepyar (*Ricinus communis* L.) lokal berdasarkan karakter morfologi spesifik, in: Prosiding Seminar Nasional Pembangunan Pertanian II: Arah Dan Tantangan Pembangunan Pertanian Dalam Era SDG’s. Malang 25 November 2017. Badan Penerbitan Fakultas Pertanian UB, Malang, pp. 333–335.
3. Arif, M., Khurshid, H., Siddiqui, S.U., Jatoi, S.A., Jan, S.A., Ilyas, M., Ghafoor, A., 2015. Estimating spatial population structure through quantification of oil content and phenotypic diversity in Pakistani Castor Bean (*Ricinus communis* L.) germplasm. Sci. Technol. Dev. 34, 147–154.
4. BPS, 2015. Produksi Tanaman Perkebunan 2011-2015.
5. Chakrabarty, S. K., C.L., Mukta., N., 2006. Draft National Guidelines for the Conduct of Tests for Distinctness, Uniformity, and Stability Castor (*Ricinus communis* L.). Indian Council of Agricultural Research., India.
6. Dewi, I.S., Arisanti, Y., Purwoko, B.S., Hariyadi, M.S., 2013. Keragaman genetic beberapa genotype jarak pagar (*Jatropha curcas* L.) berdaya hasil tinggi berdasarkan karakter morfologi, agronomi, dan isozim. J. AgroBiogen. 9, 28–38.
7. Freeman, J. V, Julious, S.A., 2010. Scope Papers (Collection of Statistical Papers first published in Scope), IESE Research Papers. https://doi.org/10.1177/0149206311406265
8. Gupta, N., Singh, A., 2015. A review on Ricinus communis Linn. Int. Ayurvedic Med. J. 3, 491–495.
9. Kanti, M., Anjani, K., Usha Kiran, B., Vivekananda, K., 2015. Agro-morphological and molecular diversity in castor (*Ricinus communis* L.) germplasm collected from Andaman and Nicobar Islands, India. Czech J. Genet. Plant Breed. 51, 96–109. https://doi.org/10.17221/205/2014-CJGPB
10. Kearsey, M.J., Pooni, H.S., 1996. The Genetical Analysis of Quantitative Traits. Chapman and Hall, Birmingham.
11. Rana, M., Dhamija, H., Prashar, B., Sharma, S., 2012. *Ricinus communis* L. - A Review 4, 1706–1711.
12. Rana, M., Sharma, R., Sharma, P., Kumar, S., Kumar, D., Dogra, R.K., 2015. Correlation and path-coefficient analysis for yield and its contributing traits in capsicum, *Capsicum annuum* L. Int. J. Farm Sci. 5, 66–73.
13. Salihu, B.Z., Gana, A.K., Apuyor, B.O., 2014. Castor oil plant (*Ricinus communis* L.): botany, ecology, and uses. Int. J. Sci. Res. 3, 1333–1341.
14. Sheoran, O., Tonk, D., Kaushik, L., Hasija, R., Pannu, R., 1998. Statistical software package for agricultural research workers. Recent advances in information theory, statistics & computer applications by D.S. Hooda& R.C. Hasija [WWW Document]. Dep. Math. Stat. CCS HAU, Hisar.
15. Singh, R.K., Chaudhary, B.D., 1979. Biometrical Methods in Quantitative Genetics Analysis. Kalyani Publishers, New Delhi.
16. Singh, S., Prakash, A., Chakraborty, N.R., Wheeler, C., Agarwal, P.K., Ghosh, A., 2016. Trait selection by path and principal component analysis in *Jatropha curcas* for enhanced oil yield. Ind. Crops Prod. 86, 173–179. https://doi.org/10.1016/j.indcrop.2016.03.047
17. Syukur, M., Sujiprihati, S., Yunianti, R., Nida, K., 2010. Pendugaan komponen ragam, heritabilitas dan korelasi untuk menentukan kriteria seleksi cabai (*Capsicum annuum* L.) populasi...
F5. J. Hortik. Indones. 1, 74–80.

18. UPOV, 2016. DRAFT Castor Bean (*Ricinus communis* L.) Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability. International Union for the Protection of New Varieties of Plants.

19. Waluyo, B., Ramayanti, F., Saptadi, D., Ardiarini, N.R., Kuswanto, 2017. Keragaman genetik karakteristik fisik biji jarak kepyar (*Ricinus communis* L.) local untuk bahan baku industri, in: Prosiding Seminar Nasional Pembangunan Pertanian II: Arah Dan Tantangan Pembangunan Pertanian Dalam Era SDG’s. Malang 25 November 2017. Badan Penerbitan Fakultas Pertanian UB, Malang, pp. 317–321.

20. Woolford, S., 2015. (Factor) Analyze This PCA or EFA. Genetic Counseling Training Program, July 31, 2015. National Human Genome Research Institute, The Johns Hopkins University.