

## Short Communication:

# Preliminary characterization of groundcherry (*Physalis angulata*) from East Java Province, Indonesia based on morpho-agronomic traits

HALIMATUS SADIYAH<sup>1,\*</sup>, ANDY SOEGIAN TO<sup>2</sup>, BUDI WALUYO<sup>2</sup>, SUMERU ASHARI<sup>2</sup>

<sup>1</sup>Department of Agronomy, Faculty of Agriculture, Universitas Jember. Jl. Kalimantan 37, Jember 68121, East Java, Indonesia.

Tel. +62-331-778238, \*email: sadiyah@unej.ac.id

<sup>2</sup>Department of Agronomy, Faculty of Agriculture, Universitas Brawijaya. Jl. Veteran, Malang 65145, East Java, Indonesia

Manuscript received: 21 October 2019. Revision accepted: 25 December 2019.

**Abstract.** Sadiyah H, Soegianto A, Waluyo B, Ashari S. 2020. Short Communication: Preliminary characterization of groundcherry (*Physalis angulata*) from East Java Province, Indonesia based on morpho-agronomic traits. *Biodiversitas* 21: 759-769. Accurate knowledge of the genetic relationships between different germplasms is important for successful crop improvement programs through breeding. This work aimed to i) characterize the groundcherry germplasm collected from the field of eastern Java and Madura Island based on their morpho-agronomic performance, ii) investigate the relationships among the different germplasms, and iii) verify whether simultaneous analysis of quantitative-qualitative data (joint analysis) can reveal new clustering patterns. A total of 26 quantitative and 31 qualitative traits were measured and evaluated for 28 accessions. Principal Component Analysis (PCA) was used to select variables that had significant loading, based on a t-test, for inclusion in cluster analysis. The first clustering method was the Unweighted Pair Group Method with Arithmetic Mean (UPGMA), which only involved quantitative variables using the Manhattan distance. This method was followed by a joint analysis using Gower dissimilarity. Nonparametric multivariate analysis of variance test showed that four clusters were the optimum number for both methods, but the clusters for each method had different members. The coefficient of variation ranged from 3.9 (fruit pH) to 45.5 (productivity per plant), indicating that improvement in the productivity in groundcherry was promising. The mean values of the morpho-agronomic traits for accessions from eastern Java and Madura Island were not significantly different, except for fruit firmness and the mean length of two internodes. In general, geographic isolation did not have much influence on the characteristics of an accession, since there were accessions from these two regions that were members of the same group, identified by the clustering analysis.

**Keywords:** Characterization, genetic distance, groundcherry, *Physalis angulata*

## INTRODUCTION

*Physalis* belongs to the family of *Solanaceae*, and there are approximately 90 known species (Zamora-Tavares et al. 2014). The diverse species of *Physalis* are known by different names, including cutleaf groundcherry, camapu, winter cherry, wild tomato, and mullaca (Osho et al. 2010). In Indonesia, this plant has various regional names, including *cecenet* (Sunda), *ceplukan* (Java), *yor-yoran* (Madura), *lapinonat* (Seram), *angket* (Bali), *dedes* (Sasak), *leletokan* (Minahasa) (Latifah et al. 2014). In botanical terms, Groundcherry (*Physalis angulata* L.) is an annual plant that is branched and upright. The stem is tilted and hollow, and it reaches 1-2 m in height. Most are hairless (glabrous), but some plants have short hairs, especially on younger parts. Single plants have up to 200 leaves, depending on the biotype, and the availability of water and nutrition. In high water treatment, the height can reach 1.7 m (Travlos 2012). The mature fruit is an orange-colored, 1-1.8 cm long and is enclosed by calyx. The flowers are yellow, 5-40 mm in length, borne on stalks. The corolla is yellowish-green, and the anthers are bluish or violet, up to 2.5 mm long (CABI 2019).

*Physalis angulata* is an important species for human consumption and widely used in traditional medicine systems. It contains chemicals with potentially beneficial properties and has a high nutrient content (Aliero and Usman 2016). Recently, the demand for this species has increased, and this presents an opportunity for the commercial cultivation of groundcherry. However, this requires an effort to develop because this species of groundcherry has not been widely cultivated commercially in Indonesia (Chaidir et al. 2015), also grown on a small scale in Jalisco, Mexico (Morales-Saavedra et al. 2019). Among the factors that determine the successful development of groundcherry as a crop are the identification of superior cultivars assembled and selection from available germplasm accessions. This germplasm is a source of genetic diversity in the species and is important as breeding material. Genetic diversity is essential because it is the primary capital for efforts to improve results (Mahbub et al. 2015). Based on genetic sources, breeders can choose populations with large distances, or predetermined characteristics, or they can propagate plants with desirable traits (Trevisani et al. 2016). Accurate knowledge of the genetic relationships between different germplasms is essential for the successful management of

appropriate long-term crop improvement programs through breeding and is also useful in supporting conservation strategies and genetic improvement.

Morphological descriptors can be used to estimate diversity, as the first step in determining the genetic potential of a species in depth. Morpho-agronomic diversity studies are highly recommended for diversity analysis before using other characteristics. Even today, morphology remains useful because the plant breeders and germplasm collectors usually record morphological data during routine germplasm evaluation (Usaizan et al. 2018). These test results increase our understanding of the diversity of germplasm for future collection and improve its use in conservation (Simbaqueba et al. 2011).

Many studies of the chemical components of groundcherry have found compounds with medicinal properties, including anti-diabetic, anti-bacterial, anti-tumor and antioxidant properties (Poojari et al. 2014; Cobaleda-Velasco et al. 2017; Sun et al. 2017; Mahalakshmi and Nidavani 2014). The morphological and agronomical traits of *P. angulata* have been studied by Morales-Saavedra et al. (2019) and Valdivia-Mares et al. (2016) and in different species of *Physalis* (*P. peruviana*: Chacón et al. 2016; Morillo-Coronado and González-Castillo 2018); (*P. minima*: Pagare et al. 2016; Usaizan et al. 2014); (*P. acutifolia*, *P. chenopodifolia*, *P. pubescens*: Valdivia-Mares et al. 2016). To date, there has been no comprehensive study of *Physalis* in Indonesia. The study was carried out only on Mount Kelud in eastern Java (Hadiyanti et al. 2018). This work aimed to i) characterize the groundcherry germplasm collected in the field of eastern Java and Madura Island based on the morphology and agronomic (morpho-agronomic) characteristics, ii) investigate the relationships among populations based on their morpho-agronomic performance, and iii) verify whether the joint analysis of quantitative-qualitative data can reveal new clustering patterns.

## MATERIALS AND METHODS

### Plant material

Experiments were conducted in a greenhouse at PT Perkebunan Nusantara X Ajong Gayasan, located at 76 m a.s.l. at 8°12'50.6"S and 113°40'36.6"E. The study used 28 accessions originating from the districts of eastern Java and Madura Island (see Table 1 and Figure 1). Seeds were germinated in trays using sand as a medium. After 13 days, the seedlings had three leaves and were then planted in a cocopeat medium for 20 days. They were then transplanted into a polybag containing a medium prepared from a 1:2:1 mixture of compost, soil, and sand, which had a pH of 6.2. The experimental work was conducted from August 2018 to January 2019. The temperature during the experimental work was 26-34°C and the humidity was 55-67%. The experimental work used a randomized block design with 28 treatments (accessions) with four replicates, and three to five plants per replicate (plot), totaling 475 plants. The spacing was 1.0 m between rows and 0.7 m between plants.



**Figure 1.** Origin of accessions in any district in East Java Province (eastern Java and Madura Island), Indonesia

**Table 1.** Names and origin of accession evaluated in this study

Accession name	Origin	
	District	Latitude-longitude
BWI-TD	Banyuwangi	8°33'24.9"S 114°20'28.4"E
BLT-SK	Blitar	8°05'52.3"S 112°08'31.7"E
BWS-WS	Bondowoso	7°53'52.6"S 113°52'21.1"E
GRS-SM	Gresik	7°09'10.7"S 112°35'40.1"E
JBR-TG	Jember	8°09'02.5"S 113°43'37.7"E
JBG-MA	Jombang	7°34'01.2"S 112°20'11.1"E
KDR-NL	Kediri	7°47'30.2"S 111°59'48.9"E
KDR-RM(02)	Kediri	7°51'24.4"S 112°01'31.5"E
LMG-KB	Lamongan	7°09'33.6"S 112°21'26.2"E
LMJ-PD	Lumajang	8°07'43.1"S 113°10'29.1"E
MDN-TM	Madiun	7°39'33.8"S 111°31'44.9"E
MGT-BR	Magetan	7°38'38.3"S 111°20'54.3"E
MLG-TP(01)	Malang	8°00'51.6"S 112°45'05.8"E
MJK-GS	Mojokerto	7°28'22.8"S 112°26'30.7"E
NGJ-KS	Nganjuk	7°36'32.4"S 112°05'10.2"E
PMK 1	Pamekasan	7°09'03.2"S 113°28'57.1"E
PBL-PT	Probolinggo	7°42'40.0"S 113°29'54.1"E
PBL-PK	Probolinggo	7°49'02.1"S 113°29'56.3"E
SPG-AS	Sampang	7°12'22.5"S 113°12'32.2"E
SPG-DC	Sampang	7°12'04.9"S 113°19'53.4"E
SPG-GD	Sampang	7°11'29.6"S 113°15'08.3"E
STB-KP	Situbondo	7°41'26.1"S 114°03'28.5"E
SMN-PG(01)	Sumenep	7°00'11.6"S 113°52'13.5"E
SMN-PG(02)	Sumenep	7°00'11.6"S 113°52'13.5"E
SBY-GY	Surabaya	7°20'35.0"S 112°43'36.4"E
TGL-WL	Trenggalek	8°16'51.8"S 111°43'50.4"E
TBN-KR	Tuban	6°53'27.4"S 111°54'38.2"E
TLG-PL(02)	Tulungagung	8°14'19.5"S 112°00'01.6"E

### Evaluation of trait

A total of 57 morpho-agronomic traits were evaluated, consisting of 26 quantitative and 31 qualitative traits. Fifteen quantitative traits and all 31 qualitative traits were measured using the Illustrated Guidelines for the

Description of husk tomato (*Physalis ixocarpa* Brot. ex Horm.) varieties (Peña-Lomelí et al. 2011). Some quantitative traits were transformed to ordinal data, but in this experiment, the analysis was on quantitative variables. There were 11 additional quantitative traits in this study (some of those traits were used in some other *Physalis* study), which were not listed in Peña-Lomelí et al. (2011). The *mean length of internode* (SML) measured as the mean of the first internode of the two main branches. *Flowering time* measured when 5% of the plant had at least one fully open flower. *Total soluble solids* and *pH* calculated from the mean of three measurements per sample made from the ‘juice’ of 40 fruits. *Calyx and fruit length and width* measured with digital vernier calipers using five fruits per plot. *The number of fruit and productivity per plant* estimated from 13 times of fruit harvest every three days. *Fruit weight and volume* measured for three samples with 15 fruit per sample for each plot. *Fruit firmness* measured as the total firmness of the skin and underlying flesh by a puncture test, using five samples per plot, measured twice for each sample at different random positions with a Rheo Tex machine type SD-700 Sun Scientific Co. Ltd. The puncture probe penetrated at 5.5 mm into the flesh.

#### Data analysis

Two methods of analysis were used. In the first method, we considered only quantitative traits, and in the second method, we analyzed qualitative and quantitative data simultaneously (joint analysis). An average value for each quantitative trait and accession was obtained for further analysis.

For each method, first, the collected data were analyzed using Principal Component Analysis (PCA) from a correlation matrix. The Principal Components (PCs) greater than one were selected. The loading factor of PCs is proportional to the correlation coefficient between the PC score and the variable (traits evaluated). Thus, the loading factor can be defined as the correlation coefficient. The data in PCA scaled to unit variance; consequently, the *t*-test for the correlation coefficient (Equation 1) can be applied to test factor loading in PCA. The *t*-statistic has a *t*-distribution with (*n*-2) degrees of freedom (Yamamoto et al. 2014). Variables that had a significant contribution to those selected PCs based on the *t*-test were then used to form clusters.

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad \dots \quad (\text{Equation 1}).$$

Where, *t* is the *t* statistic, *r* is the correlation coefficient, and *n* is the number of variables.

Clustering carried out by the Unweighted Pair Group Method using Arithmetic Averages (UPGMA) used the variables that accounted for a relatively high proportion of the total variability. A dendrogram, including all the accessions, was created using the Manhattan distance for the quantitative traits, and the Gower distance for the joint analysis. The quantitative variables in this research had different scales, so they were transformed before clustering. The transformation involved dividing each

variable by its standard deviation. This transformation has been suggested by some authors as the previous step of calculating the distance in order to eliminate the effect of scale on classification results (Crossa and Franco 2004).

In this research, the Manhattan distance was applied to the transformed data to run the cluster analysis and produce a dendrogram. The Manhattan distance was defined as the sum of the absolute values of the differences of the individuals *j* and *k* for each variable. It was similar to that in the Gower distance formula, but the Gower formula using the number of columns and the data range as denominator. The Gower distance itself has been suggested for analyzing mixed data types of quantitative and qualitative (Gower 1971; Crossa and Franco 2004).

After clustering, the optimal number of the groups was defined by nonparametric multivariate analysis of variance (Permanova). PCA, cluster analysis, and Permanova were calculated using PAST 3.25 (Hammer et al. 2001). The correlation between the dissimilarity matrices described above was also examined. The correlations and their significances were tested by using the Mantel test with 1000 permutations, calculated using Genstat 18.2 software. This procedure allowed us to assess the resemblance between two matrices: quantitative and joint analysis dissimilarity matrices.

## RESULTS AND DISCUSSION

### Morpho-agronomic characterization

Five of the 31 qualitative traits could be used to differentiate the accessions and were included in the analysis. To describe the general behavior of the accessions for each variable, a descriptive analysis for the qualitative (Table 2) and quantitative variables (Table 3) was carried out. The results of the analysis of the qualitative traits (Table 2) showed that only six out of 28 (21.43%) accessions had anthocyanin coloration in their internode (SAC). Only two of those six accessions (PBL-PT and NGJ-KS) had medium intensity coloration, and the other four had weak coloration (KDR-NL, KDR-RM(02), JBG-MA, MJK-GS). Four accessions had no dentation of the margin of their leaves (SMN-PG(01), BLT-SK, PMK1, BWS-WS); the others had medium dentation. The fruit color at physiological maturity was white for three accessions (KDR-NL, BLT-SK, LMJ-PD) and green for the others. The intensity of calyx anthocyanin coloration varied. It was found that 13 out of 28 accessions had very weak anthocyanin coloration, nine accessions had weak coloration, and six accessions had medium anthocyanin coloration.

Table 3 shows that the FSt mean was 19.398 mm, with a range of 17.25-22.425 mm, and the FNS range of 78-133. These ranges were not significantly different from the results of Nadhifah et al. (2016), who found ranges of 8.5-24 mm for FSt and 129-207 for FNS for *P. angulata* originating from the former Surakarta province.

The quantitative morpho-agronomic characteristics showed considerable variability among the accessions, especially LW, LPL, FNS, FTSW, FNP, and PP, while FL,

FD, FRLD, FDe, FpH, and those characteristics that measured the flower and calyx showed little phenotypic variation (the CV was less than 10%). PP and FNP, two characters related to productivity, had the highest variability (> 40%). This meant that the accession collection showed promise for a breeding process to improve productivity. Morales-Saavedra et al. (2019) observed that the CV ranged 0.9-3.8 %, 20.6-42.6%, and 3.2-20.1% for FD, FNP, and FW, respectively, in the study of *P. angulata* in Mexico, including wild, weedy and cultivated populations.

The FW in this study (1.768) had a higher mean than *P. angulata* in the study of Morales-Saavedra et al. (2019) (1.383). The FpH (4.75), and FTSS (10.44) had a higher mean than those found by Valdivia-Mares et al. (2016), which were 4.22 and 8.5 for FpH and FTSS, respectively. The fruit parameters, including FL, FW, and FV, had a lower variation and lower mean values than those reported by Herrera et al. (2012) for *P. peruviana*, while FF, FD, and FNP had greater variability. *P. angulata* is an annual plant while *P. peruviana* is a perennial plant that can produce a greater number of fruits. The FpH had a lower variation, but a higher mean and the total soluble solid was similar (approximately 10% Brix). The FW, FNS, and the two leaf characteristics (LL and LPL) had higher mean values and greater variation than those found for *P. minima* in the study by Usaizan et al. (2018). Usaizan et al. (2018) found that the mean values of FW, FNS, LL, and LPL for *P. minima* were 1.2 g, 104.5, 8 cm, and 4 cm, and the values of CV were 4%, 4.8%, 2.5%, and 3.5%, respectively. However, CW had lower values for the mean and CV (1.8 cm and 7.1%, respectively, for *P. minima*). The FD and FL had lower mean values but higher variation. These differences could be due to differences between the two species and the influence of environmental effects on quantitative characters.

A two-sample t-test was conducted to investigate whether there was a significant difference between accessions originating from eastern Java and those from

Madura Island. The t-test revealed that only two of the 26 variables were significantly different. Those two traits were SML and FF, with p-values of 0.035 and 0.018, respectively. The accessions from eastern Java had higher values of SML (overall mean = 6.5203) than those from Madura (overall mean = 5.7). The accessions from Madura had higher mean values for FF than those from eastern Java, 387.83 and 359.68, respectively. Morales-saavedra et al. (2019) found the FF of *P. angulata* from Mexico range from 1.1 to 2.4 kg cm<sup>-2</sup> and mean 1.67 kg cm<sup>-2</sup>. Herrera (2012) studied fruit firmness in the fruit of *Physalis peruviana*. The degree of fruit firmness has been used as an important aspect of fruit quality for tomatoes and as a critical handling parameter for peach, pear, apple, and kiwi fruits (Wang et al. 2009).

### Quantitative traits

PCA is commonly used before cluster analysis to determine the relative importance of each variable. The PCA procedure builds new synthetic variables, principal components (PCs), which are the linear combination of the original variables. PCA from a correlation matrix is used when the variables measured have different scales (Jolliffe and Cadima 2016).

PCA for the 26 quantitative traits based on the correlation matrix resulted in seven PCs with a value higher than one and accounted for 79.13% of the total variation (Table 4). These eigenvalues were in the range of 1.11-6.68, and the variation explained by those eigenvalues was in the range 4.28-25.69%. The two-way t-test for the loading factor at  $t_{(24;0.025)}$  was 2.064. Variables were considered to be statistically significant when the t-value was larger than  $t_{(24;0.025)}$ . Based on equation 1, any loading factor greater than 0.38826 or less than -0.38826 was significant. Seven variables (SML, LL, LPL, FF, FDe, FSt, and FpH) were found to be significant and were included in the cluster analysis. The dendrogram from the cluster analysis for quantitative traits is shown in Figure 2.A.

**Table 2.** List of qualitative traits, their code, and frequency

Variable	Category	Code	Frequency (%)
Stem: Anthocyanin coloration of internodes	1. absent	SAC	78.57
	9. present		21.43
Stem: Intensity of anthocyanin coloration of internodes	3. weak	SIAC	92.857
	5. medium		7.143
	7. strong		0
Leaf: Dentation of margin	3. absent/weak	LDM	14.286
	5. medium		85.714
	7. strong		0
Fruit: Main color (at physiological maturity)	1. white	FMC	10.714
	2. green		89.286
	3. yellow		0
	4. orange		0
	5. purple		0
Calyx: Intensity of anthocyanin coloration of the calyx	1. very weak	CIAC	46.43
	2. weak		32.14
	3. medium		21.43
	4. strong		0
	5. very strong		0

**Table 3.** List of quantitative traits and their codes, followed by its summary statistics

Trait	Code	Min	Accession	Max	Accession	Mean	CV (%)
Stem							
<sup>a</sup> Diameter (cm)	SD	1.659	SMN-PG(02)	2.446	MGT-BR	2.031	14.7
Height at first branching (cm)	SH	12.025	PMK-1	17.3	JBG-MA	14.277	18.1
Mean length of internodes (cm)	SML	4.772	SPG-GD	7.91	MDN-TM	6.345	18.1
Leaf							
Length (cm)	LL	8.3	BWS-WS	11.6	SMN-PG(01)	9.648	19.6
Width (cm)	LW	4.538	SPG-DC	6.325	MDN-TM	5.358	22.8
<sup>a</sup> Ratio length/width	LR	1.593	MGT-BR	2.022	SMN-PG(02)	1.802	10.5
Petiole: length (cm)	LPL	4.5	KDR-NL	6.888	LMG-KB	5.676	25.4
Flower							
Diameter (cm)	FID	1.2	*	1.5	*	1.321	8.8
Flowering time (day after sowing)	FIT	48	SPG-GD	61	GRS-SM	54.286	7.1
Fruit							
Length (mm)	FL	12.38	MDN-TM	16.09	PBL-PT	14.766	6.4
Diameter (mm)	FD	11.695	MDN-TM	14.285	PBL-PT	12.967	6.3
Ratio length/diameter	FRLD	1.063	MDN-TM	1.191	BWI-TD	1.14	4.1
Firmness (g/mm)	FF	327	MGT-BR	433.3	SMN-PG(01)	365.708	10.5
<sup>a</sup> Weight per fruit (g)	FW	1.511	SPG-DC	2.145	PBL-PT	1.768	13.5
<sup>a</sup> Volume (cm <sup>3</sup> )	FV	1.525	SPG-DC	2.17222	PBL-PT	1.775	13.7
Density <sup>c</sup> (g/cm <sup>3</sup> )	FDe	0.967	BWS-WS	1.135	PBL-PK	1	5.6
Number of seeds per fruit	FNS	78.242	MDN-TM	133.208	JBG-MA	110.741	21.2
The 1000 seeds weight	FTSW	0.645	LMG-KB	0.911	SBY-GY	0.762	25.3
Stalk (mm)	FSt	17.25	GRS-SM	22.425	MJK-GS	19.398	11.0
<sup>a</sup> total soluble solid (% Brix)	FTSS	9.058	MDN-TM	11.658	MGT-BR	10.443	10.0
<sup>a</sup> pH	FpH	4.51	BWS-WS	4.969	MGT-BR	4.75	3.9
Calyx							
<sup>a</sup> Length (mm)	CL	21.625	MDN-TM	27.925	SPG-AS	25.97	7.4
<sup>a</sup> Width (mm)	CW	11.695	MDN-TM	14.285	PBL-PT	12.967	6.3
<sup>a</sup> Ratio length/width	CR	1.359	KDR-RM(02)	1.839	MJK-GS	1.56	8.8
<sup>a</sup> Number of fruit per plant	FNP	167.688	MDN-TM	571.471	TGL-WL	397.791	41.6
<sup>a</sup> Productivity per plant (g/plant)	PP	288.25	MDN-TM	1161.95	PBL-PT	705.776	45.5

Note: <sup>a</sup> Additional trait; \*more than one accession**Table 4.** Eigenvalue, accumulated variance and loading of each variable to each of the first 7 PCs

Variables	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7
Eigenvalue	6.68	3.86	3.25	2.49	1.77	1.41	1.11
Cumulative variance	25.69	40.55	53.05	62.62	69.44	74.85	79.13
	<b>PC 1</b>	<b>PC 2</b>	<b>PC 3</b>	<b>PC 4</b>	<b>PC 5</b>	<b>PC 6</b>	<b>PC 7</b>
SD	-0.202	0.241	-0.137	0.113	0.073	-0.086	0.146
SH	-0.16	0.202	-0.118	0.278	0.037	0.316	0.264
SML	-0.235	0.138	0.03	0.115	<b>*0.391</b>	0.076	0.127
LL	-0.153	0.096	<b>*0.452</b>	-0.042	0.071	0.036	0.047
LW	-0.237	0.059	0.309	-0.161	0.204	-0.108	0.021
LR	0.156	0.066	0.29	0.21	-0.267	0.286	0.06
LPL	-0.133	0.122	<b>*0.44</b>	-0.01	-0.122	-0.07	0.127
FID	0.016	-0.055	0.351	-0.056	0.04	0.151	-0.025
FIT	-0.207	0.075	-0.158	0.285	0.222	0.324	0.117
FL	0.311	0.254	-0.054	0.013	0.06	-0.088	0.037
FD	0.332	0.032	-0.011	0.189	0.195	0.007	0.032
FRLD	0.07	0.376	-0.074	-0.224	-0.158	-0.148	0.022
FF	-0.005	-0.07	0.356	0.251	-0.033	0.123	<b>*-0.406</b>
FW	0.22	0.117	0.09	-0.334	0.321	0.117	0.148
FV	0.211	0.145	0.098	-0.3	0.383	0.024	0.104
FDe	0.057	-0.119	-0.091	-0.145	-0.284	<b>*0.462</b>	0.222
FNS	0.185	0.169	0.097	0.34	-0.021	-0.181	0.378
FTSW	0.022	0.148	-0.173	-0.288	0.057	0.258	-0.352
FSt	0.092	0.247	0.083	-0.009	0.078	<b>*0.428</b>	-0.271
FTSS	-0.044	0.357	0.061	-0.048	-0.265	-0.203	0.017
FpH	0.001	-0.041	0.046	-0.33	-0.268	0.216	<b>*0.458</b>
CL	0.21	0.353	0.057	0.071	-0.19	-0.058	-0.13
CW	0.332	0.032	-0.011	0.189	0.195	0.007	0.032
CR	-0.027	0.378	-0.127	0.058	-0.163	0.111	-0.181
FNP	0.3	-0.218	0.048	0.121	-0.075	-0.001	0.005
PP	0.34	-0.155	0.073	0.005	0.036	0.039	0.077

Note: \*= the significant loading factor

According to the Permanova test, the optimal number of the groups required to represent the genetic variability was four. The F-values calculated by the Gower, Manhattan, and Euclidean dissimilarities were 6.361, 6.434, 6.15, successively (Figure 2b), and their probabilities were 0.0001 (not shown). The highest F-values were found for three, six, and seven groups. The dendrogram showed that it was not possible to divide the accessions into five groups, and the Permanova test for five groups could not be conducted.

The number of accessions per group for the four groups ranged from 1 in group 1 to 20 in group 3, and the distance between groups was approximately 8.7 (Figure 2a). The characteristics of each group were explained using a boxplot (Figure 3). The dendrogram showed that group 1 was composed of PBL-PK and was characterized by having the highest mean length of fruit stalk (18.95 mm). Group 2 was composed of five accessions (MDN-TM, JBG-MA, LMG-KB, SMN-PG(01), MJK-GS), characterized by a low value of fruit density, high values for leaf length, petiole length, and mean length of internodes, and a high variation of fruit firmness and fruit pH. Group 3 was composed of 20 accessions characterized by low values for fruit firmness, leaf length, and petiole length, and a medium value for fruit density. There were two outliers for fruit density (SPG-DC with the smallest value, and PBL-PT with the highest value) and one outlier for fruit firmness (TGL-WL). Group 3 also had a high variation in the mean length of the internodes. Group 4 was composed of two accessions (SMN-PG(02), and PMK1). The main characteristics of this group were high values for fruit density and fruit firmness, a low value for the mean length of internodes, and moderate values for leaf length and petiole length.

#### Joint analysis of quantitative and qualitative traits

The joint analysis involved analyzing quantitative and qualitative traits simultaneously. The PCA based on correlation matrices of 26 quantitative and five qualitative traits showed ten PCs with values higher than one, which accounted the 84.20% of the total variation (Table 5). The eigenvalues of the first ten PCs were in the range 1.02-6.87, and the variations explained 3.29-22.17% of the total variation. The variable was considered to be significant if its t-value was higher than  $t_{(29;0.025)} = 2.045$ , or if according to equation 1, it was meant that the loading factor greater than 0.355 or less than -0.355 was significant. From the first ten PCs, 12 variables were found to be significant and were included in cluster analysis. These variables were seven quantitative variables (LL, LPL, FF, FW, FV, FDe, FSt) and all five qualitative variables.

The PCA found seven quantitative traits that were significant and which were included in the cluster analysis. It revealed that five traits were important for both quantitative and joint analysis methods (LL, LPL, FF, FDe, FSt). Four of the seven quantitative variables were different in the two methods. SML and FpH were only included in the first method, and FW and FV were only included in the second method. Interestingly, in the joint analysis, all qualitative variables were assigned as significant for building clusters. This method showed that qualitative traits

were important in clustering because those traits were the genetic expression that relatively not too influenced by the environment. Cluster analysis in the second method used the Gower distance, which weights the variables based on their contributions and can be used for calculating the similarity/dissimilarity of mixed quantitative and qualitative variables (joint analysis). Using 12 variables accounted for a large correlation with the first ten PCs and gave the dendrogram shown in Figure 4a. The optimal number of groups was assessed by Permanova (Figure 4b), based on Gower, Manhattan, and Euclidean similarity/dissimilarity indices. These three indices showed that the steepest decrease in the F-value was from four to five groups. This decrease was larger than from three to four groups. Crossa and Franco (2004) suggested the use of graphs for different numbers of groups and observed the point of maximum increase as a criterion for defining the optimal number of groups. In the present study, we used the criterion in the reverse direction (using the maximum decrease in joint analysis). With seven groups the Gower F value increased, but the other two indices decreased. It was concluded that the optimal number of groups for representing the genetic variability was four. The F-value for six groups could not be investigated because it was not possible to divide the accessions into six groups based on the dendrogram.

The number of accessions per group differed in the four groups. There were four accessions in group 1, 22 accessions in group 2, and only one accession each in groups 3 and 4. The distance between groups was approximately 0.34 (Figure 4a). The main character of each group formed by joint analysis can be seen in the Boxplot (Figure 5).

The dendrogram showed that group 1 was composed of PBL-PT, NGJ-KS, KDR-RM(02), MJK-GS accessions. Group 1 was characterized by having high values of fruit stalk length and fruit volume. This group was distinctive, and all four of its members had anthocyanin present in their internode. Only 6 out of 28 accessions (21.43%) had anthocyanin present in their internode-the two other accessions (KDR NL and JBG-MA) belonged to other groups and had a medium intensity of anthocyanin in their internodes, the only two accessions with such levels of anthocyanin.

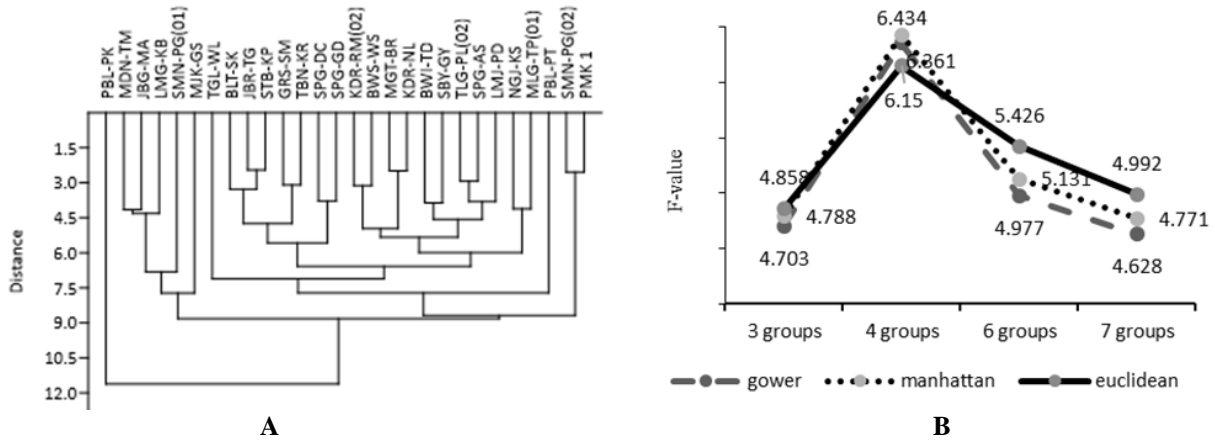
Group 2 was composed of 22 accessions resulting in the highest variability of the four groups. The main characteristics of this group were 1) shorter fruit stalk lengths, smaller fruit volume, and lower fruit weight than accessions in group 1; 2) a medium level of fruit density. There was one outlier, each in fruit density (PBL-PK) and fruit weight (SMN-PG(02)). Both of these outliers had the highest values for these variables.

Group 3 had only one accession (KDR-NL). This accession has a minimum value for petiole length, a relatively low value for fruit firmness, a high value for fruit density, and a medium value for fruit stalk length. KDR-NL was one of three accessions with white as the primary color of the fruit at physiological maturity. There was anthocyanin coloration in its internode but only at a low level. Other qualitative characters were a medium level of

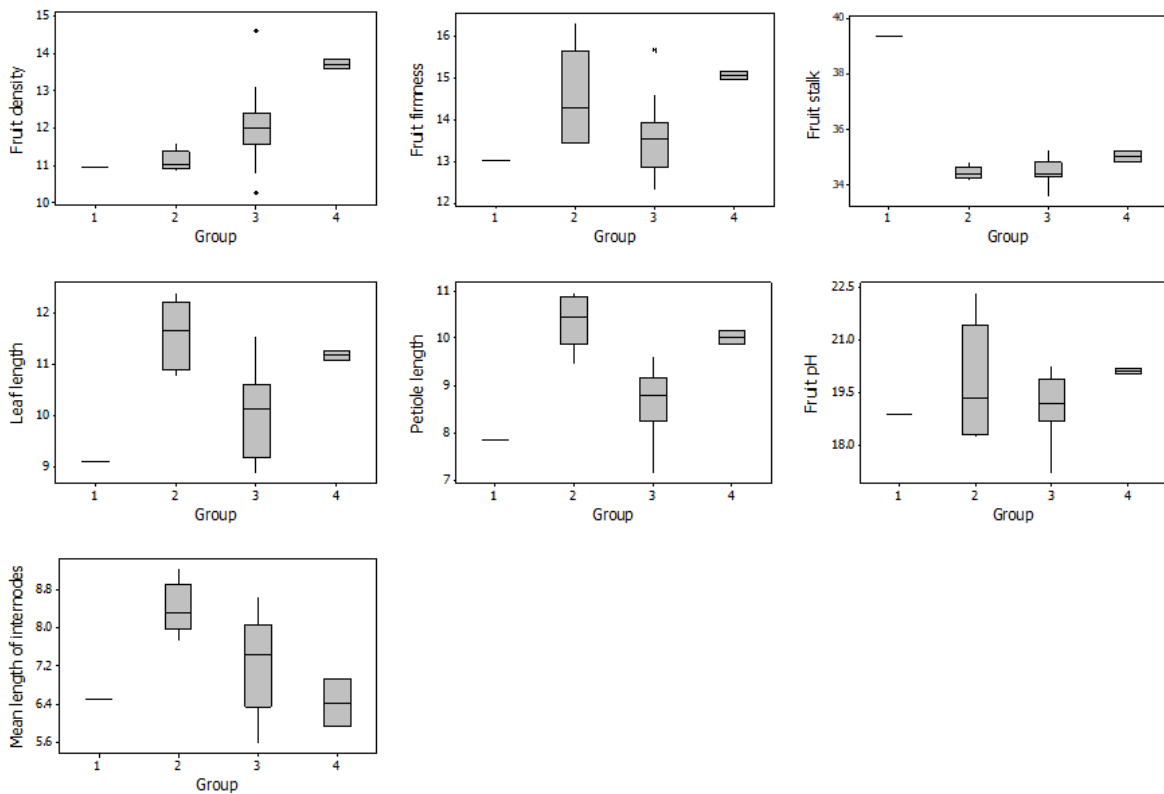
leaf margin dentation, and the absence of calyx anthocyanin coloration.

Group 4 also had only one accession (SMN-PG(01)). This accession was mainly characterized by the highest values of leaf length and fruit firmness, and the second-highest fruit stalk length. SMN-PG(01) also had the third-

lowest of fruit weight and fruit volume. Its qualitative characteristics were no anthocyanin coloration in the internode, weak dentation of leaf margin, green as the primary color of the fruit at physiological maturity, and medium intensity anthocyanin coloration of the calyx.



**Figure 2.** A. Dendrogram based on 7 quantitative traits significant on first 7 PCs using UPGMA cluster analysis and Manhattan distance on the standardized-variable matrix; B. F-value of Permanova test on quantitative analysis



**Figure 3.** The boxplot of 7 variables included in cluster analysis based on quantitative variables

Table 5. Eigenvalue, accumulated variance and loading of each variable to each of the first 10 PCs

	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC9	PC10
<b>Eigenvalue</b>	6.87	3.97	3.38	2.66	2.44	1.89	1.53	1.23	1.11	1.02
<b>Cumulative variance</b>	22.17	34.97	45.88	54.47	62.35	68.45	73.39	77.34	80.91	84.20
	<b>PC 1</b>	<b>PC 2</b>	<b>PC 3</b>	<b>PC 4</b>	<b>PC 5</b>	<b>PC 6</b>	<b>PC 7</b>	<b>PC 8</b>	<b>PC9</b>	<b>PC10</b>
SD	-0.199	0.216	-0.167	0.041	-0.103	-0.067	-0.039	-0.149	-0.198	0.274
SH	-0.160	0.180	-0.181	0.266	-0.101	0.133	0.292	0.045	0.023	0.177
SML	-0.226	0.151	-0.007	0.208	0.104	-0.171	0.100	-0.265	0.004	0.228
LL	-0.147	0.155	<b>*0.418</b>	0.050	0.106	0.009	0.052	0.049	-0.022	-0.011
LW	-0.230	0.109	0.295	-0.027	0.217	-0.120	-0.084	0.050	-0.089	0.080
LR	0.155	0.084	0.252	0.131	-0.205	0.268	0.268	-0.009	0.145	-0.185
LPL	-0.128	0.166	<b>*0.412</b>	-0.014	-0.020	0.205	-0.020	-0.140	0.046	0.019
FID	0.021	-0.012	0.347	0.010	0.026	-0.093	0.237	-0.020	-0.171	0.182
FIT	-0.199	0.072	-0.184	0.315	-0.026	-0.100	0.224	-0.040	0.255	0.116
FL	0.306	0.240	-0.088	-0.016	-0.014	0.002	-0.113	-0.009	0.050	0.100
FD	0.326	0.040	-0.023	0.196	-0.031	-0.070	-0.087	0.055	0.192	0.205
FRLD	0.070	0.343	-0.116	-0.271	0.029	0.101	-0.065	-0.075	-0.184	-0.126
FF	-0.006	-0.024	<b>*0.368</b>	0.142	-0.267	-0.168	0.031	0.140	0.154	-0.126
FW	0.224	0.130	0.063	-0.127	<b>*0.395</b>	-0.067	0.135	-0.166	0.118	0.063
FV	0.216	0.162	0.066	-0.086	<b>*0.399</b>	-0.119	0.053	-0.167	0.071	0.063
FDe	0.054	-0.142	-0.064	-0.183	-0.016	0.244	<b>*0.381</b>	-0.005	0.320	-0.024
FNS	0.184	0.173	0.039	0.287	-0.158	0.179	-0.118	-0.173	-0.230	0.201
FTSW	0.021	0.123	-0.146	-0.325	0.000	-0.336	0.203	0.237	0.153	0.245
FSt	0.093	0.247	0.028	0.029	-0.006	-0.150	<b>*0.422</b>	0.268	-0.098	-0.157
FTSS	-0.046	0.353	0.027	-0.139	-0.124	0.164	-0.158	0.154	-0.118	0.180
FpH	-0.006	-0.054	0.043	-0.267	0.148	0.346	0.265	0.224	-0.182	0.315
CL	0.209	0.339	0.009	-0.053	-0.189	0.068	-0.035	0.023	-0.091	-0.140
CW	0.326	0.040	-0.023	0.196	-0.031	-0.070	-0.087	0.055	0.192	0.205
CR	-0.025	0.348	-0.170	-0.055	-0.150	0.053	0.081	0.018	0.128	-0.338
FNP	0.291	-0.218	0.063	0.082	-0.122	0.003	0.023	0.127	-0.239	0.101
PP	0.334	-0.151	0.076	0.039	0.020	-0.006	0.069	0.055	-0.177	0.119
SAC	0.034	0.053	-0.103	<b>*0.409</b>	0.270	0.096	0.194	0.089	-0.155	-0.059
SIAC	0.088	0.108	0.009	0.200	<b>*0.446</b>	0.071	-0.157	0.006	0.119	-0.320
LDM	-0.069	-0.123	-0.179	0.076	0.199	<b>*0.479</b>	0.040	0.002	-0.133	-0.017
FMC	-0.088	0.082	0.094	-0.007	0.055	0.313	-0.297	0.275	<b>*0.502</b>	0.302
CIAC	-0.103	0.058	-0.056	0.172	0.163	-0.105	-0.153	<b>*0.676</b>	-0.161	-0.087

Note: \*= the significant loading factor

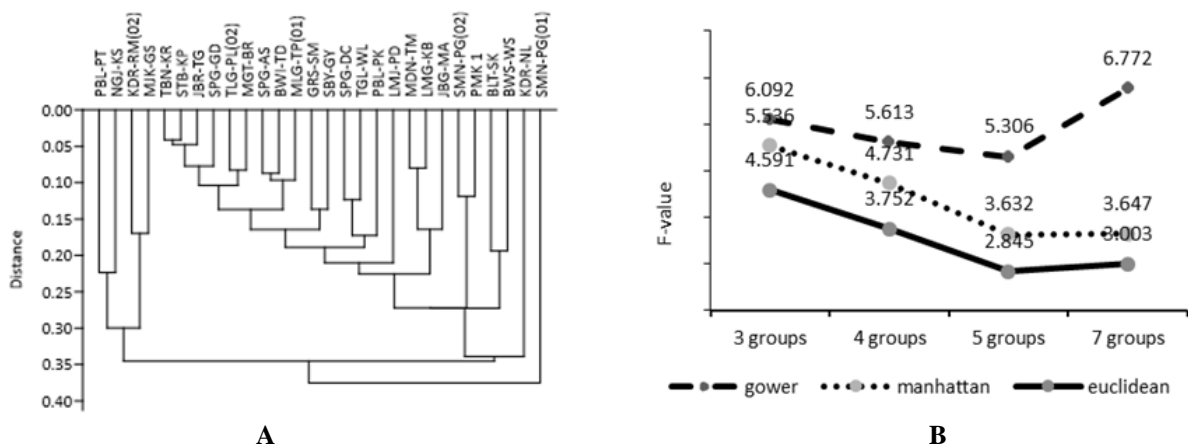
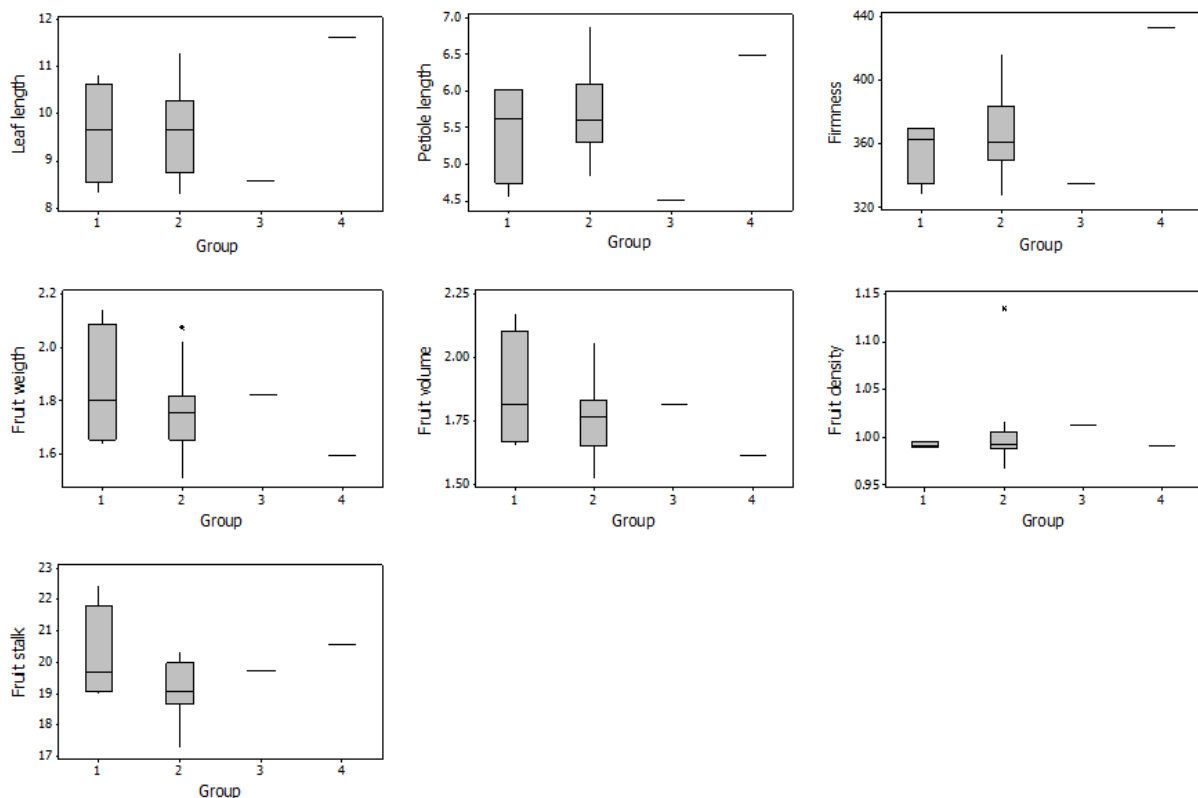


Figure 4. A. Dendrogram based on 10 quantitative and 5 qualitative traits significant on first 10 PCs using UPGMA cluster analysis and Gower distance on the original variable matrix; B. F-value of Permanova test on joint analysis





**Figure 5.** The boxplot of 7 variables included in cluster analysis based on joint analysis

In the present study, four groups were the optimum number for classification of the accessions, based on both quantitative and joint analyses. In a study of cassava, Oliveira et al. (2014) found that the number of groups in the joint analysis was less than the number for the quantitative trait analysis. This difference was possibly caused by there being a higher number of qualitative variables (35) than quantitative variables (13) in classifying the genotypes. It was known that there was less variation in the qualitative variables; therefore fewer groups were needed for classifying the genotypes in the joint analysis. In contrast, the present study had only five qualitative traits and 26 quantitative traits included in the joint analysis.

The dendrograms (Figure 2a and 4a) revealed that there were mixed accessions from eastern Java and Madura Island in one group, even if 15 groups at a Manhattan distance of 4.5 were used for quantitative analysis and 16 groups at a Gower distance of 0.15 were used for the joint analysis. Genetic distance matrices of the accessions ranged from 2.46 to 16.99. The lowest distance values were observed between JBR-TG and STB-KP (2.46), MGT-BR and KDR-NL (2.49), SMN-PG(02) and PMK-1 (2.54), TBN-KR and STB-KP (2.80), and TLG-PL(02) and SP-AS (2.93) (data not shown). Genetic distance matrices of the accessions ranged from 0.041 to 0.636. The lowest distance values were observed between TBN-KR and STB-KP (0.041), JBR-TG and STB-KP (0.046), TLG-PL(02) and JBR-TG (0.075), MDN-TM and LMG-KB (0.079), SPG-GD and STB-KP (0.082) (data not shown). In general,

there was no significant difference between accessions from eastern Java and those from Madura Island, even though there is geographic isolation in the form of the strait. This conclusion was consistent with the results of the t-test on the average of eastern Java versus Madura Island, where only two of 26 quantitative traits had significant differences (fruit firmness and mean length of internodes). This similarity was probably caused by the migration by humans or because morphologically, they have the same type of ecology.

The different clustering patterns were obtained from the joint analysis and quantitative trait analysis. The accessions present in the different groups based on quantitative analysis could be clustered together based on joint analysis, or vice versa. The use of qualitative data in the joint analysis provided additional information and variation that influenced the measurement of dissimilarity among the accessions and affected the grouping. It seems that the parental selection for breeding programs was different when using the results of joint analysis rather than quantitative traits alone.

A Mantel test of 1000 permutations showed significant associations between the two distance matrices (0.5365), with a probability 0.0001. This relation may indicate that the configuration of the distance matrices, formed by using quantitative data and based on joint analysis, had a strong association.

The results presented in this paper are the first data on the morphology and agronomy characteristics of *P.*

*angulata* accessions, especially those from the eastern Java and Madura Island. This study can be used as the basis for the characterization of groundcherry and the development of breeding programs and conservation strategies for groundcherry diversity. In terms of fruit quality characters (fruit volume and fruit weight), the PBL-PT, SMN-PG(02), and PMK 1 accessions were found to be very promising for the cultivation of fresh fruit. The morphological and agronomical traits are essential variables for *P. angulata* research, but the outcomes obtained from this study have been limited due to the very limited reference accessions in *P. angulata* breeding programs. The details record of phenology under cultivation, particularly the time when 50% of the plants present flowers and when 50% develop fruits as a precocity indicator is beneficial to conduct. A molecular-based study is required for a more reliable classification of *P. angulata* germplasms and to overcome these limitations.

To our knowledge, this is the first detailed study that compares the clustering result based on two methods: quantitative and quantitative-qualitative simultaneously; and contributes to the understanding of the characteristics of *P. angulata* germplasms in eastern Java and Madura Island. A nonparametric multivariate analysis of variance (Permanova) test showed that four clusters were optimum for both methods, but the clusters for each method had different members. The present study revealed considerable variation in the morpho-agronomic characteristics of groundcherry accessions. The highest variation was in the productivity per plant (45.5%) indicating that it is a promising focus for improving the productivity of groundcherry. The morpho-agronomic traits for eastern Java and Madura Island were not significantly different, except for fruit firmness and the mean length of two internodes. Based on the two methods, some groups composed of accessions from both eastern Java and Madura Island. In general, geographic isolation did not influence the characteristics of accessions.

## ACKNOWLEDGEMENTS

Thanks to The Indonesian Education Scholarship (LPDP) for its financial support to the research reported in this article

## REFERENCES

- Aliero AA, Usman H. 2016. Leaves of groundcherry (*Physalis angulata* L.) may be suitable in alleviating micronutrient deficiency. *Food Sci Technol* 4: 89-94.
- CABI. 2019. *Physalis Angulata*. In: *Invasive Species Compendium*. CAB International, Wallingford, UK.
- Chacón MI, Sánchez Y del P, Barrero LS. 2016. Genetic structure of a Colombian Cape Gooseberry (*Physalis peruviana* L.) collection by means of microsatellite markers. *Agron Colomb* 34: 5-16.
- Chaidir L, Epi, Taofik A. 2015. Eksplorasi, identifikasi, dan perbanyakan tanaman Ciplukan (*Physalis angulata* L.) dengan menggunakan metode generatif dan vegetatif. *Jurnal ISTEK* 9: 82-103. [Indonesian]
- Cobaleda-Velasco M, Alanis-Bañuelos RE, Almaraz-Abarca N, Rojas-López M, González-Valdez LS, Ávila-Reyes JA, Rodrigo S. 2017. Phenolic profiles and antioxidant properties of *Physalis angulata* L. as quality indicators. *J Pharm Pharmacogn Res* 5: 114-128.
- Crossa J, Franco J. 2004. Statistical methods for classifying genotypes. *Euphytica* 137: 19-37.
- Gower JC. 1971. A general coefficient of similarity and some of its properties. *Biometrics* 27: 857-871.
- Hadiyanti N, Supriadi, Pardono. 2018. Diversity of ciplukan (*Physalis* spp) on the gradient of Mt. Kelud, East Java. *Berita Biologi* 16: 135-146. [Indonesian]
- Hammer Ø, Harper DAT, Ryan PD. 2001. PAST: Paleontological statistics software package. *Palaeontol Electron* 4: 1-9.
- Herrera A, Fischer G, Chacón M. 2012. Agronomical evaluation of cape gooseberries (*Physalis peruviana* L.) from central and north-eastern Colombia. *Agron Colomb* 30: 15-24.
- Jolliffe IT, Cadima J. 2016. Principal component analysis: a review and recent developments subject areas. *Phil Trans A* 374:1-16
- Latifah N, Hidayati AA, Yunas SR, Sulistyorini E. 2014. Ciplukan (*Physalis angulata* L.). *Encyclopedia of Anticancer Plant*. www.cccr.farmasi.ugm.ac.id.
- Mahalakshmi A, Nidavani R. 2014. *Physalis angulata* L.: an ethnopharmacological review. *Indo Am J Pharm Res* 4: 1479-1486.
- Mahbub M, Rahman M, Hossain M, Mahmud F, Kabir M. 2015. Genetic variability, correlation and path analysis for yield and yield components in soybean. *American-Eurasian J Agric Environ Sci* 15: 231-236.
- Morales-Saavedra J del C, Zaragoza FAR, Toledo DC, Hernández CVS, Vargas-Ponce O. 2019. Agromorphological characterization of wild and weedy populations of *Physalis angulata* in Mexico. *Sci Hortic (Amsterdam)* 246: 86-94.
- Morillo-Coronado AC and González-Castillo JA. 2018. Characterization of genetic diversity uchuva (*Physalis peruviana* L.) In *Boyacá. Biotecnol. en el Sect. Agropecu y Agroindustrial* 16: 26-33.
- Nadhifah A, Suratman, Pitoyo A. 2016. Keekerabatan genetik ciplukan (*Physalis angulata* L.) di wilayah eks-Karesidenan Surakarta berdasarkan karakter morfologis, palinologis dan pola pita isozim. *J Tumbuhan Obat Indonesia* 9: 1-10. [Indonesian]
- Osho A, Adetunji T, Fayemi SO, Moronkola DO. 2010. Antimicrobial activity of essential oils of *Physalis angulata*. *L. Afr J Tradit Complement Altern.Med* 7: 303-306.
- Oliveira EJ, Ferreira CF, Santos VS, Oliveira GA. 2014. Development of a cassava core collection based on single nucleotide polymorphism markers. *Genet Mol Res* 13(3): 6472-6485.
- Pagare S, Mishra RP, Tripathi N, Rathore M, Kumar B. 2016. Assessment of genetic diversity among different biotypes of *Physalis minima*. *Indian J Weed Sci* 48: 417-420.
- Peña-Lomelí A, Magaña-Lira N, Montes-Hernández S, Sánchez-Martínez J, Santiaguillo-Hernández JF, Grimaldo-Juárez O. 2011. Illustrated Guidelines for the Description of Husk Tomato (*Physalis ixocarpa* Brot. ex Horm.) Varieties. SNICS-SAGARPA, Universidad Autónoma Chapingo, Chapingo, Estado de México, México.
- Poojari S, Porika R, Mamidala E. 2014. Phytochemical analysis and in vitro antidiabetic activities of *Physalis angulata* fruit extracts. *Nat J Integr Res Med* 5: 34-38.
- Simbaqueba J, Sánchez P, Sanchez E, Núñez Zarantes VM, Chacon MI, Barrero LS, Mariño-Ramírez L. 2011. Development and characterization of microsatellite markers for the cape gooseberry *Physalis peruviana*. *PLoS ONE* 6: e26719. DOI: 10.1371/journal.pone.0026719
- Sun C-P, Qiu C-Y, Zhao F, Kang N, Chen L-X, Qiu F. 2017. Physalins V-IX, 16,24-cyclo-13,14-seco withanolides from *Physalis angulata* and their antiproliferative and anti-inflammatory activities. *Sci Rep* 7: 1-10.
- Travlos IS. 2012. Invasiveness of cut-leaf ground-cherry (*Physalis angulata* L.) populations and impact of soil water and nutrient availability. *Chil J Agric Res* 72: 358-363.
- Travisani N, Schmit R, Beck M, Guidolin AF, Coimbra JLM. 2016. Selection of fisális populations for hybridizations, based on fruit traits. *Rev Bras Frutic* 38: e-568. DOI: 10.1590/0100-29452016568.
- Usaizan N, Abdullah NAP, Ahmad SH, Saleh G. 2014. Preliminary phytochemical screening and GC-MS analysis of ethanol extract of *Physalis minima* L. (*Solanaceae*). *J Adv Agric Technol* 1: 100-103.
- Usaizan N, Ashikin N, Abdullah P, Saleh G, Kashiani P. 2018. Genetic diversity of *Physalis minima* L. accessions based on morphological traits. *Adv Plants Agric Res* 8: 151-157.
- Valdivia-Mares LE, Zaragoza FAR, González JJS, Vargas-Ponce O. 2016. Phenology, agronomic and nutritional potential of three wild husk

- tomato species (*Physalis*, *Solanaceae*) from Mexico. *Sci Horti* (Amsterdam) 200: 83-94.
- Wang YW, Wang J, Yao C, Lu QJ. 2009. Firmness measurement of peach by impact force response. *J Zhejiang Univ Sci B* 10: 883-889.
- Yamamoto H, Fujimori T, Sato H, Ishikawa G, Kami K, Ohashi Y. 2014. Statistical hypothesis testing of factor loading in principal component analysis and its application to metabolite set enrichment analysis. *BMC Bioinformatics* 15: 51. DOI: 10.1186/1471-2105-15-51.
- Zamora-Tavares P, Vargas-Ponce O, Sánchez-Martínez J, Cabrera-Toledo D. 2014. Diversity and genetic structure of the husk tomato (*Physalis philadelphica* Lam.) in Western Mexico. *Genet Resour Crop Evol* 62:141-153