African Crop Science Journal, Vol. 26, No. 1, pp. ....
Printed in Uganda. All rights reserved

ISSN 1021-9730/2018 $4.00

DOI: http://dx.doi.org/10.4314/acsj.v26i1.1

LINEAR DISCRIMINANT ANALYSIS OF STRUCTURE WITHIN AFRICAN EGGPLANT 'SHUM'

G. SSEREMBA[1,2], P. TONGOONA[1]*, J.S.Y. ELEBLU[1], E.Y. DANQUAH[1] and E.B. KIZITO[2]

[1]West Africa Centre for Crop Improvement, University of Ghana, PMB 30, Accra, Ghana
[2]Department of Agricultural and Biological Sciences, Faculty of Science and Technology, Uganda Christian University, P. O. Box 4, Mukono, Uganda
Corresponding authors: gsseremba@wacci.edu.gh, ptongoona@wacci.edu.gh

(*Received 10 August, 2017; accepted .... 2018*)

ABSTRACT

African eggplant (*Solanum aethiopicum)* lags behind in terms of research attention; leading to, among other gaps, scantiness of information on intra-group genetic diversity. Hierarchical clustering was applied to reveal structure in a pioneer 20-entry set of *S. aethiopicum* Shum group accessions at Africa Solanaceae Research Network secretariat in Uganda, using 61 morpho-agronomic variables collected from a complete randomised design (CRD) with 12 plants per accession in screen-house. A MANOVA preceded linear discriminant analysis, to model each of 61 variables, as predicted by clusters and experiment to filter out non-significant traits. Four distinct clusters emerged, with a cophenetic relation coefficient of 0.87 (P<0.01). Canonical variates that best predicted the observed clusters include petiole length, sepal length (or seed color), fruit calyx length, seeds per fruit, leaf fresh weight (or leaves per plant), fruit fresh yield, seedling vigour, fruits per plant, harvest index and plant growth habit. Results suggest prospects for variety discrimination at various stages; seedling, vegetative and reproductive. The observed diversity will boost chances of effective selection for desired traits; while the canonical discriminant traits are potential morphological markers, within *S. aethiopicum* 'Shum' for a low-cost germplasm characterisation approach.

Key Words: Hierarchical clustering, morphological markers, *Solanum aethiopicum*

RÉSUMÉ

African eggplant (*Solanum aethiopicum)* lags behind in terms of research attention; leading to, among other gaps, scantiness of information on intra-group genetic diversity. Hierarchical clustering was applied to reveal structure in a pioneer 20-entry set of *S. aethiopicum* Shum group accessions at Africa Solanaceae Research Network secretariat in Uganda, using 61 morpho-agronomic variables collected from a complete randomised design (CRD) with 12 plants per accession in screen-house. A MANOVA preceded linear discriminant analysis, to model each of 61 variables, as predicted by clusters and experiment to filter out non-significant traits. Four distinct clusters emerged, with a cophenetic relation coefficient of 0.87 (P<0.01). Canonical variates that best predicted the observed clusters include petiole length, sepal length (or seed color), fruit calyx length, seeds per fruit, leaf fresh weight (or leaves per plant), fruit fresh yield, seedling vigour, fruits per plant, harvest index and plant growth habit. Results suggest prospects for variety discrimination at various stages; seedling, vegetative and reproductive. The observed diversity will boost chances of effective selection for desired traits; while the canonical discriminant

traits are potential morphological markers, within *S. aethiopicum* 'Shum' for a low-cost germplasm characterisation approach.

Key Words: Hierarchical clustering, morphological markers, *Solanum aethiopicum*

INTRODUCTION

African eggplant (*Solanum aethiopicum*) is the third most important solanaceous crop species after tomato and potato, but little is known on the species' germplasm diversity (Cernansky, 2015; Gramazio *et al.*, 2016; Kie³bowicz-Matuk *et al.*, 2014). *Solanum aethiopicum* consists of four recognised morphological groups (Aculeatum, Kumba, Gilo and Shum) each cultivated for a particular plant part and different purposes (Plazas *et a*l., 2014; World Vegetable Center, 2017).

The Shum is a nutrient-rich leafy vegetable; a source of food and income in Uganda (Cernansky, 2015; Pincus, 2015) and other sub-Saharan Africa countries like Cameroon, Nigeria, Burkina-Faso and Côte d'Ivoire (Kouassi *et al*., 2014; Cernansky, 2015). Other groups such as the Gilo are also commercially produced in Latin America, Asia and across Africa (Plazas *et al*., 2014).

The crop's cultivation is dominated by smallholder farmers, who constitute the highest proportion of the sub-Saharan Africa populations (Abukutsa-Onyango, 2014; Von Grebmer *et al*., 2015). Farmers, however, experience low yields and low leaf quality, as a result of a number of constraints such as low soil fertility, drought stress, lack of quality seed, pests, diseases and postharvest deterioration (Abukutsa-Onyango, 2014; Pincus, 2015; Bisamaza and Banadda, 2017). Breeding for improved varieties to address the constraints is deemed the best strategy as it offers a lasting solution (Abukutsa-Onyango, 2014; Cernansky, 2015). However, the existence of selectable What do you mean? within-group variability among the 'Shum' had not been investigated.

Lumped up accessions from the different cultivar groups are reported to exhibit diversity in what? (Plazas *et al*., 2014). Conspicuously, the Shum were notably insufficiently represented when compared with the Gilo, for example (Adeniji *et al*., 2012; Plazas *et al*., 2014). The lumping up and limited representation for the Shum could have masked the exposure of within-group structure (Plazas *et al*., 2014). Because Shum is leafy type, but reproductively propagated crop, adequate inclusion of leaf traits at seedling and harvest-maturity (vegetative) stages; as well as reproductive attributes ought to be ensured in a comprehensive genetic diversity analysis based on morphology (World Vegetable Center, 2017).

Statistical approaches like cluster analysis and linear discriminant analysis (LDA) are popular and adequate for the genetic diversity analysis (Asher *et al*., 2017). Cluster analysis in graphic user interface softwares like R, is effective for large data sets for both quantitative and qualitative variables (Hornik and Böhm, 2017; Murtagh, 2017). The clustering can be based on a number of alogarithms such as complete linkage, Ward's method and the Unweighted Pair Group Method with Arithmetic mean (UPGMA) method (Saraçli *et al.,* 2013).

The UPGMA, also referred to as the "average" method, was applied in this study due to its unraveled Seems incorect popularity and acceptability. The strength of clustering can be evaluated by way of coefficients, such as agglomerative coefficient and cophenetic correlation coefficient (CPCC, Saracli *et al.* (2013). The CPCC is a common and reliable measure of dissimilarity among formed clusters (Saraçli *et al*., 2013; Silva and Dias, 2013; Murtagh, 2017; Nikoliæ *et al*., 2017).

In order to identify and quantify traits' contribution for best prediction of observed clusters in the dendrogram, LDA (also called canonical discriminant analysis) is a suited approach (Harding and Payne, 2012). Discriminant analysis refers to a linear combination of data variates (original variables) as loadings, which maximises between-cluster variance and minimise within-cluster variance (Jombart and Collins, 2015). It relies on allocation of groups (observed or assigned) as dependent variables to be predicted by data (variates).

The LDA has been applied in medicine, animal (Arandas *et al.*, 2017) and plant research. In plants, LDA has been applied in various studies such as taxonomic and germplasm characterisation (Herklotz *et al.*, 2017), phenotypic changes evaluation in plant species over time (Alberti *et al.*, 2017) and crop diseases detection on remote sensing generated data (Bajwa *et al.*, 2017).

In the Shum group of *S. aethiopicum*, LDA was applied in this study to investigate the morphological diversity of available germplasm, in order to ascertain its suitability for conservation and usefulness in genetic improvement of farmes' cultivars for desired traits through breeding.

The objective of this study was, therefore, to investigate the existence and drivers of structure within *S. aethiopicum* Shum group. We hypothesized that there is considerable variation within *S. aethiopicum* Shum group, as a starting point towards selection for superior varieties and affor germplasm characterisation within the subspecies.

MATERIALS AND METHODS

Germplasm and research site. A pioneer set of 20 accessions of *S. aethiopicum* Shum were obtained from the secretariat of Africa Solanaceae Research Network (Afri-SOL) at the Department of Agricultural and Biological Sciences (DABS), in Uganda Christian University (UCU); Mukono Municipality, Uganda. Afri-SOL, affiliated to Sol Genomics Network, is a network of scientists and other stakeholders with interest in solanaceous species in Africa (www.afri-sol.org/). The study was carried out in screen house at DABS-UCU between April 2016 and March 2017. The study accessions are described in Table 1.

Experimental design. The study was carried out using pots in the screen house in 2016, and October 2016 – March 2017, respectively. The potting substrate was composed of thoroughly mixed topsoil (clay loams) and cowdung manure in a ratio of 5topsoil:2 manure. A fortnightly schedule for insecticide (profenofos and cypermethrin at 2 ml L^{-1} of water), fungicide (metalaxyl and mancozeb at 2g L^{-1} of water) and fertiliser (C.A.N at 4 g pot^{-1}) were administered.

Watering was carefully carried out using tap water delivered through a hosepipe on a daily basis. A completely randomised design

TABLE 1. List of *S. aethiopicum* Shum accessions used for the screen house study

Serial no.	Accession	Pedigree
1	168G	SAS168/G/2015
2	183G	SAS183/G/2015
3	163	SAS163/2015
4	163P	SAS163/P/2015
5	157P	SAS157/P/2015
6	160	SAS160/2015
7	163G	SAS163/G/2015
8	183P	SAS183/P/2015
9	108	SAS108/2015
10	157G	SAS157/G/2015
11	148	SAS148/2015
12	145	SAS145/2015
13	168P	SAS168/P/2015
14	184G	SAS184/G/2015
15	137	SAS137/2015
16	184P	SAS184/P/2015
17	141	SAS141/2015
18	108P	SAS108/P/2015
19	185G	SAS185/G/2017
20	185P	SAS185/P/2015

was used in which 24 plants (12 for data collection at seedling/vegetative stages, and remaining 12 for reproductive/fruiting stage) per accession were evaluated. Each pot was planted with two seedlings at the four-leaf stage (four weeks after sowing in a nursery); followed by thinning to one plant per pot. The individual plant in a pot was the observational unit.

Data collection and analysis

Hierarchical clustering. Data were collected on a total of 61 morpho-agronomic traits, at different stages of crop growth, namely cotyledonous and seedling, vegetative and reproductive/ fruiting. This number of variables excludes those that did not significantly differ for at least two accessions ($P<0.05$). The vegetative stage, at 4 weeks after transplanting (WAT), is the harvest stage (since *S. aethiopicum* Shum is a leafy vegetable), at which leaf, shoot and root weights were taken. A summary of the scale or units of measurement for the different traits is included in Table 2, and most of these traits have been described earlier by Plazas *et al.* (2014), Adeniji *et al.* (2012, 2013) and Kubie (2013). The variables were summarised in Excel to obtain means. A table of means per trait for each accession was then imported into R statistical software for UPGMA reordered hierarchical cluster analysis. A mantel test (cophenetic correlation coefficient, CPCC) for significance of clustering was also performed.

Discriminant analysis of clusters. After clustering, the raw data file (in Excel) was updated with a column named ‘cluster’, to assign a group number to each accession (member) as determined from dendrogram. A multivariate analysis of variance (MANOVA) was then carried out (for the 61 variables) to identify traits that significantly ($P<0.01$) contributed to observed clusters in the dendrogram. The following two-way MANOVA model was analysed in GenStat 12th Edition (VSN International Ltd):

$$y_{ij} = \mu + C_i + R_j + \varepsilon_{ij}$$

Where:

y_{ij} is the observed trait measurement for a group of accessions in the i^{th} cluster (C) and j^{th} experiment (R); μ and ε_{ij} refer to overall cluster mean for a trait and random error, respectively.

The traits (variables) which did not significantly differ for at least two clusters ($\alpha = 1\%$) were excluded from discriminant analysis (DA). The DA in GenStat generates c-1 number of discriminant functions (DFs, where c = number of clusters), correlations between data variates and DFs, discriminant scores for cluster means and Mahalanobis’ (D^2) inter-cluster distances. Finally, a Pearson’s correlation between canonical variates was carried out to identify variate (morphological marker) alternatives for strongly correlated traits.

RESULTS

Hierarchical clustering. Observed clustering in the dendrogram (Fig. 1) showed four distinct clusters, with a CPCC of 0.87 ($P<0.05$). Cluster 3 was the richest; followed by cluster 2. Cluster 4 and cluster1 had only one and two members, respectively (Table 3).

The mean performance of accessions also showed differences for different variables. Accessions 160 and 145, which constituted cluster 4 had higher petiole length (PEL), sepal length (SEL), fruit length (FRL), fruit breadth (FRB), fruit calyx length (FCL), fruit cross section diameter (FCD), fruit fresh weight (FFW) and seeds per fruit (SPF) than the rest of accessions. Accession 185P Cluster 1 uniquely had the highest number of fruits per plant (FPP) and fruit fresh yield (FFY).

Discriminant analysis of clusters. The MANOVA eliminated 11 variables on the basis of non-significance in cluster discrimination (Table 4). The eliminated variables include FPL,

TABLE 2. Summary of 61 variables recorded

Variable	Unit/scale	Variable	Unit/scale	Variable	Unit/scale
Days to emergence (DG)	Days	Shoot flesh biomass (SBF)	g	Flowers per inflorescence (FLI)	Counts
Cotyledonous leaf length (CLBL)	mm	Shoot dry biomass (SBD)	g	Stamen length (mm) (N=5)(STL)	mm
Cotyledonous leaf width (CLBW)	mm	Harvest fresh index (HIF)	Ratio	Petal length (mm) (N=5) (PEL)	mm
Visual seedling vigour (VSV)	Scale (1-5)	Harvest dry index (HID)	Ratio	Sepal length (N=5) (SEL)	mm
Seedling leaf length (SLBL)	mm	Plant growth habit (PGH)	Scale (1-9)	Relative style length (mm) (RSL)	mm
Seedling leaf width (SLBW	mm	Spines on stem (SOS)	Scale (1-9)	Pollen production (POL)	Scale (0-7)
Seedling fresh weight (SDFW)	mg	Stem pubescence (SPU)	Scale (1-4)	Fruit length (mm) (FRL)	mm
Seedling dry weight (SDDW)	mg	Stem colour (SC)	Scale (1-3)	Fruit breadth (mm) (FRB)	mm
Leaf blade length (LBL)	cm	Petiole colour (PC)	Scale (1-3)	Fruit pedicel length (mm) (FPL)	mm
Leaf blade width (LBW)	cm	Petiole prickles (PP)	Scale (1-9)	Fruit pedicel thickness (mm) (FPT)	mm
Leaf area	cm^2	Leaf blade lobbing (LL)	Scale (1-9)	Fruit color at physiological ripeness (FCP)	Scale (1-9)
Leaves per plant (LPP)	Counts	Leaf blade tip angle (LTA)	Scale (1-9)	Fruit position (FRP)	Scale (1-9)
Plant branching (PB)	Counts	Leaf blade color (LBC)	Scale (1-3)	Fruit calyx length (N=10) (FCL)	mm
Plant height (PH)	cm	Leaf vein pigmentation (LVP)	Scale (1-3)	Fruit cross section diameter (N=10) (FCD)	mm
Plant width (PW)	cm	Leaf mid rib color (LMC)	Scale (1-3)	Locules per fruit (N=10) (LPF)	Counts
Leaf relative water content (LRWC)	%	Leaf spines on upper surface (LSU)	Scale (1-9)	Fruit fresh weight (N=10) (FFW)	g/fruit
Leaf fresh weight (LFW)	mg	Leaf spines on lower surface (LSL)	Scale (1-9)	Fruits per inflorescence (N=5) (FPI)	Counts
Leaf dry weight (LDW)	mg	Leaf pubescence on upper surface (LHU)	Scale (1-9)	Fruits per plant (FPP)	Counts
Leaf fresh yield (LFY)	g	Leaf pubescence on lower surface (LHL)	Scale (1-9)	Fruit fresh yield per plant (FFY)	g $plant^{-1}$
Leaf dry yield (LDY)	g	Date to first flower appearance (FLW)	Days	Seed color (SEC)	Scale (1-9)
				Seeds per fruit (SPF)	Counts

The LRWC (measured at vegetative/harvest maturity stage) = [(FW-DW)/(TW-DW)]%100; where FW = fresh weight of leaf sample, TW = Turgid weight of leaf sample, and DW = Dry weight of leaf sample

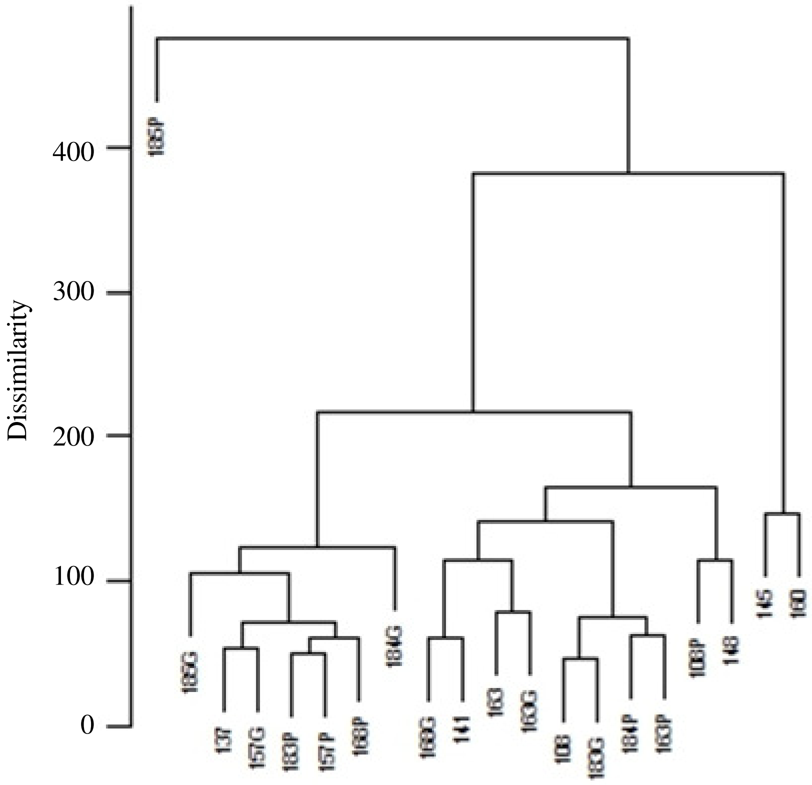

Figure 1. Reordered hierarchical clustering of study accessions using UPGMA method.

LBW, LL, LMC, LSL, LVP, PC, PP, SBD, SBF, and SOS. Linear discriminant analysis (LDA) on the remaining 50 variables (now factor variates) produced 3 discriminant functions (DFs) consisting of vector loadings or DF scores for each variate. The canonical Pearson's product-moment correlation coefficients for DF 1, DF 2 and DF 3 were 0.9999, 0.9973 and 0.9727, respectively. A biplot of the first two DF scores showed a clear separation among 4 clusters. The first DF separated Cluster 4 from the rest; while the second DF distinguished Cluster 1, Cluster 2 and cluster 3 (Fig. 2). Further, discriminant scores for cluster means under DF 1, DF 2 and DF 3 were highest for Cluster 1 and 4, Cluster 1; followed by Cluster 3 and Cluster 1; which was in turn followed by Cluster 2. The D^2 (Mahalanobis' inter-cluster distances) scores were highest between Cluster 4 and any of the remaining clusters followed by Cluster 1 and Cluster 3 ($D^2 = 2694$), and Cluster 1 and Cluster 2 ($D^2 = 1352$).

Cluster 2 and Cluster 3 had the lowest score for inter-cluster distance at 438 (Table 5). Correlations between 50 data variates and 3 DFs produced SEC (-0.05), PEL (-0.03), SEL (-0.02), SPF (-0.02) and FCL (-0.02) as the canonical variates in the first DF. In the second DF, LFY and FFY had the highest scores, both at 0.06 and -0.06, respectively; followed by VSV and FPP both at 0.04 and -0.04, respectively. In the last DF, the following variates had high scores in the order of HID (0.13), followed by HIF (0.11), PGH (0.08), and LPP and LFY both at 0.07. Strong correlations between some variates were also observed. For instance, HIF and HID (r = 0.52), SEC and SEL (r = 0.62), and LFY and LPP (r = 0.76) are correlated traits.

TABLE 3. Cluster membership of study accessions

Serial no.	Accession	Cluster
1	168G	3
2	183G	3
3	163	3
4	163P	3
5	157P	2
6	160	4
7	163G	3
8	183P	2
9	108	3
10	157G	2
11	148	3
12	145	4
13	168P	2
14	184G	2
15	137	2
16	184P	3
17	141	3
18	108P	3
19	185G	2
20	185P	1

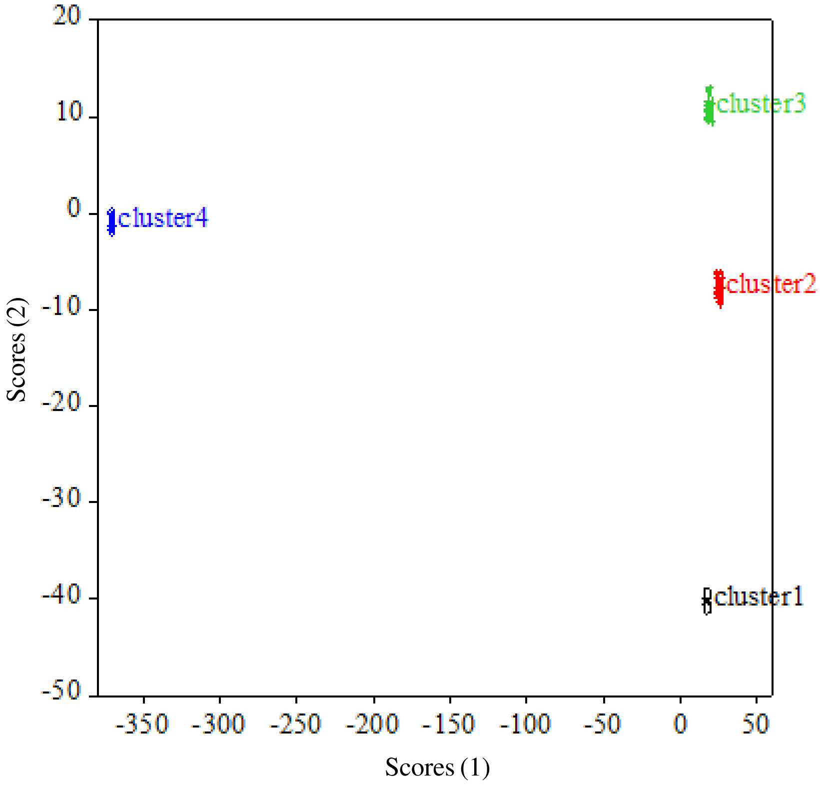

Figure 2. Biplot of first two discriminant functions showing cluster dissimilarities.

TABLE 4. Mean squares of some of the variables and their significance in clustering of study accessions

Source	d.f	CLBL	CLBW	DG	FCD	FCL	FCP	FFW	FFY	FLW	FPI	FPL	FPP	FPT
Experiment	1	2515.8	539.1	1.008	113.2	4.123	4.453	345.37	168538.0	27319.4	45.72	19.97	30030.5	191.5
Cluster	3	418.8***	86.6***	102.5***	3154.5***	2040.5***	129.0***	3144.1***	1159353.0***	272.3***	32.0***	39.3**	17918.5***	50.0***
Residual	48	7.418	1.348	2.259	11.28	3.276	2.902	15.16	19902.0	24.11	0.809	6.471	321.2	3.862

*, ** and *** stand for significance of cluster distinctiveness based on particular trait at á = 5, 1 and 0.1%, respectively

CLBL = ..., CLBW = ..., DG = ..., FCD = ..., FCL = ..., FCP = ..., FFW = ..., FFY = ..., FLW = ..., FPI = ..., FPL FPP FPT

TABLE 5. Discriminant scores for cluster means and inter-cluster distances

Discriminant scores for cluster means			
	Scores (1)	Scores (2)	Scores (3)
Cluster 1	17.35	-40.17	-11.31
Cluster 2	25.42	-7.8	4.16
Cluster 3	18.58	10.92	-2.2
Cluster 4	-370.95	-0.88	0.57

Inter-cluster distances - Mahalanobis (D-squared)				
	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	0			
Cluster 2	1352	0		
Cluster 3	2694	438	0	
Cluster 4	152466	157172	151879	0

DISCUSSION

Please section this section based on those in the Results section

The results on hierarchical clustering indicated existence of strong population structures within the *S. aethiopicum* 'Shum' group, contrary to observations made by Adeniji *et al*. (2012, 2013) and Plazas *et al*. (2014). However, it is notable that the duo did not assess the structure based on Mantel test, which gives a cophenetic correlation coefficient (CPCC), and they compared all the four recognised morpho-types of *S. aethiopicum* (Shum, Gilo, Kumba and Aculeatum) in 'one' analysis (Adeniji *et al*., 2012, 2013; Plazas *et al*., 2014; Gramazio *et al*., 2016). The CPCC is one of the reliable measures of genetic diversity applicable to both molecular and morpho-agronomic data variables (Murtagh, 2017; Nikoliæ *et al*., 2017). The higher the CPCC (closer to 1.0) at high confidence level, is the stronger the distinctiveness of observed dendrogram clusters and reliability of hierarchical clustering (Nikoliæ *et al*., 2017). After clustering, it was imperative to carryout MANOVA to eliminate data redundancy, which could have occurred as a result of introducing 'cluster' as a predictor variable for the response of different traits measured. After MANOVA, linear discriminant analysis (LDA) considers 'cluster' as a dependent (response) variable as predicted by measured traits (Arandas *et al*., 2017).

From LDA results, very high canonical correlation coefficients indicate that all three discriminant functions (DFs) generated were generally very important in dissecting variatal contribution to distinct clusters. This is in agreement with guidelines by Harding and Payne (2012) and Jombart and Collins (2015), which specify that .... Specifically, the order of importance of the DFs in distinguishing among clusters was such that: DF 1 was slightly superior to DF 2, and the DF 2 is slightly superior to DF 3.

The biplot of DF 2 scores and DF 1 scores showed a visual display of four distinct clusters. The distinctness of clusters was supported by generally high values of both inter-cluster distances and discriminant scores for cluster means. Members of cluster4 (accessions 160 and 145) possessed uniquely higher mean values for petiole length (PEL), sepal length (SEL) or seed color (SEC), fruit calyx length (FCL) and seeds per fruit (SPF) than the rest of accessions. The SEL and SEC were moderate-strongly correlated. Specifically, accessions 160 and 145 had orange-yellow fruits. Further, the biplot showed that DF 1 largely served to separate cluster4 from the rest of clusters. The observation suggests that other than leaf area, reproductive traits are putatively good morphological markers in *S. aethiopicum* 'Shum' group. Similar views were made by Adeniji *et al.* (2012) and Plazas *et al.* (2014) when they studied ....

Judging from highest discriminant score for Cluster 1 mean under DF 2, supported by correlations between data variates and DF 2, it implies that Cluster 1 (only one member, accession 185P, represented this cluster) was considered unique. This was due to extremely high fruit fresh yield per plant (FFY), the least leaf fresh weight (LFW) and relatively high number of fruits per plant. The FFY and LFW were the most important predictors of cluster distinctiveness in DF 2; the traits had the highest scores from the correlation between all data variates and the DF 2. Reduced leaf size and high reproductive efficiency are known adaptive traits for survival under harsh environments (Fita *et al*., 2015; Basu *et al*., 2016; Mwale *et al*., 2017).

Cluster 1 members showed the highest departure from Cluster 2 and Cluster 3 (basing on discriminant scores for cluster means and inter-cluster distances) under DF 2. This implies that accession 185P is different from

Cluster 2 and Cluster 3 members for LFW and FFY. Although Cluster 2 and Cluster 3 had the shortest inter-cluster distance, their discriminant scores for cluster means were most extreme under DF 2. Thus, the traits that made Cluster 1 unique are also important in discriminating between Cluster 2 and Cluster 3.

The third discriminant function (DF 3) and DF 2 were important in separating Cluster 1 from Cluster 2, where LFY (or LPP; these are strongly correlated), FFY, VSV, FPP, HID (or HIF; HID and HIF (Fig. ....) are correlated traits), and PGH are responsible traits. Contrary to Adeniji *et al.* (2012) and Plazas *et al.* (2014) whose findings emphasized reproductive traits as major drivers of variability in *S. aethiopicum*, results from this study indicated equal importance of both vegetative and reproductive traits in distinguishing among the Shum accessions. Additionally, the visual seedling vigor offers opportunity for early-stage discrimination among Shum germplasm.

CONCLUSION

There is significant genetic diversity within *S. aethiopicum* Shum, which can support breeding interventions to develop improved varieties, as a way of addressing the crop's productivity constraints. Thirteen canonical variates, which best explained the structure by way of high loadings in each discriminant function, are potential morphological markers in the Shum. The canonical variates include seedling vigour for early-stage separation among accessions or individuals of a breeding population. At vegetative (harvest maturity of 8 weeks after planting); petiole length, leaf fresh weight, number of leaves per plant, harvest fresh index, harvest dry index and plant growth habit are the major discriminating variables. Varieties with favorable scores of the said traits (harvest index and leaf fresh weight, in particular) at vegetative stage are of direct importance to consumers and farmers. To breeders and seed companies, the possession of reproductive fitness traits is key to sustainability. As such; sepal length, seed color, fruit calyx length, number of seeds per fruit, fruit fresh yield and number of fruits per plant which were among the best predictors of diversity; offer opportunities for improving crop productivity through cross-breeding research and seed supply to farmers.

ACKNOWLEDGEMENT

This work was supported by WACCI/German Academic Exchange Programme, Makerere University/EU-Intra-ACP mobility project and Forum for Agricultural Research in Africa (FARA/PAEPARD-CRFII). We are also grateful to Ms. Julian Mildred Nakanwagi for assistance in routine screen house activities and data entry, Ms. Nahamya Pamela Kabod for providing the germplam and Dr. Michael for allowing the use of facilities at the Faculty of Science and Technology (UCU) during the study.

REFERENCES

Abukutsa-Onyango, M. 2014. Strategic repositioning African indigenous vegetables and fruits with nutrition, economic and climate change resilience potential. In: *Novel Plant Bioresources*. Gurib-Fakim, A. (Eds.). London: John Wiley and Sons. pp. 361-369.

Adeniji, O.T., Kusolwa, P.M. and Reuben, S.O.W.M. 2012. Genetic diversity among accessions of *Solanum aethiopicum* L. groups based on morpho-agronomic traits. *Plant Genetic Resources* 10(03):177-185. https://doi.org/10.1017/S1479262112000226

Adeniji, O.T., Kusolwa, P. and Reuben, S. 2013. Morphological descriptors and micro satellite diversity among scarlet eggplant groups. *African Crop Science Journal* 21(1):37-49.

Alberti, M., Correa, C., Marzluff, J.M., Hendry, A.P., Palkovacs, E.P., Gotanda, K.M. and Zhou, Y. 2017. Global urban signatures of phenotypic change in animal and plant populations. *Proceedings of the National Academy of Sciences* 201606034.

Arandas, J.K.G., da Silva, N.M.V., Nascimento, R. de B., Pimenta Filho, E.C., Albuquerque Brasil, L.H. de and Ribeiro, M.N. 2017. Multivariate analysis as a tool for phenotypic characterization of an endangered breed. *Journal of Applied Animal Research* 45(1):152-158. https://doi.org/10.1080/09712119.2015.1125353

Asher, J., Williams, I.D. and Harvey, E.S. 2017. Mesophotic depth gradients impact reef fish assemblage composition and functional group partitioning in the main Hawaiian Islands. *Frontiers in Marine Science* 4. https://doi.org/10.3389/fmars.2017.00098

Bajwa, S., Rupe, J. and Mason, J. 2017. Soybean disease monitoring with leaf reflectance. *Remote Sensing* 9(2):127 One page?. https://doi.org/10.3390/rs9020127

Basu, S., Ramegowda, V., Kumar, A. and Pereira, J.G. 2016. Plant adaptation to drought stress. *F1000Research* 5:1554. https://doi.org/10.12688/f1000research.7678.1

Bisamaza, M. and Banadda, N. 2017. Solar drying and sun drying as processing techniques to enhance the availability of selected African indigenous vegetables, *Solanum aethiopicum* and *Amaranthus lividus* for nutrition and food security in Uganda. *African Journal of Food Science and Technology* 8(1):001–006.

Cernansky, R. 2015. Super vegetables *Nature* 522(7555):146.

Fita, A., Rodríguez-Burruezo, A., Boscaiu, M., Prohens, J. and Vicente, O. 2015. Breeding and domesticating crops adapted to drought and salinity: A new paradigm for increasing food production. *Frontiers in Plant Science* 6. https://doi.org/10.3389/fpls.2015.00978

Gramazio, P., Blanca, J., Ziarsolo, P., Herraiz, F. J., Plazas, M., Prohens, J. and Vilanova, S. 2016. Transcriptome analysis and molecular marker discovery in *Solanum incanum* and *S. aethiopicum*, two close relatives of the common eggplant (*Solanum melongena*) with interest for breeding. *BMC Genomics* 17(1). https://doi.org/10.1186/s12864-016-2631-4

Harding, S. and Payne, R. 2012. A guide to multivariate analysis in GenStat (15th Edition). 5 The Waterhouse, Waterhouse Street, Hemel Hempstead, Hertfordshire HP1 1ES, UK: VSN International.

Herklotz, V., Mieder, N. and Ritz, C.M. 2017. Cytological, genetic and morphological variation in mixed stands of dogroses (Rosa section Caninae; Rosaceae) in Germany with a focus on the hybridogenic *R. micrantha*. *Botanical Journal of the Linnean Society* 184(2):254–271.

Hornik, K. and Böhm, W. 2017. *Cluster Ensembles: Package "clue."* CRAN. Not complete!

Jombart, T. and Collins, C. 2015. A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.0. *Imperial College London-MRC Cent Outbreak Anal Model* 43. Retrieved from http://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf

Kielbowicz-Matuk, A., Rey, P. and Rorat, P. 2014. Interplay between circadian rhythm, time of the day and osmotic stress constraints in the regulation of the expression of a Solanum Double B-box gene. *Annals of Botany* 113:831–842. https://doi.org/doi:10.1093/aob/mct303

Kouassi, A., Béli-Sika, E., Tian-Bi, T., Alla-N'Nan, O., Kouassi, A., N'Zi, J.C. and Tio-Touré, B. 2014. Identification of Three Distinct Eggplant Subgroups within the Solanum aethiopicum Gilo Group from Côte d'Ivoire by Morpho-Agronomic Characterization. *Agriculture* 4(4):260-273.

https://doi.org/10.3390/agriculture 4040260

Kubie, L. D. 2013. Evaluation of genetic diversity in garden egg (*Solanum aethiopicum*) germplasm in Ghana. University of Ghana, Ghana. Not complete!

Murtagh, F. 2017. Massive data clustering in moderate dimensions from the dual spaces of observation and attribute data clouds. *arXiv Preprint arXiv:1704.01871*. Retrieved from https://arxiv.org/abs/1704.01871

Mwale, S.E., Ssemakula, M.O., Sadik, K., Alladassi, B., Rubaihayo, P., Gibson, P. and Edema, R. 2017. Estimates of combining ability and heritability in cowpea genotypes under drought stress and non-stress conditions in Uganda. *Journal of Plant Breeding and Crop Science* 9(2):10-18.

Nikoliæ, A., Kostadinoviæ, M., Vanèetoviæ, J., Stankoviæ, G. and Ignjatoviæ-Miciæ, D. 2017. Comparison of statistical methods for genetic similarity evaluation of maize inbred lines. *Field & Vegetable Crops Research/Ratarstvo I Povrtarstvo* 54(1).

Pincus, L.M. 2015. Increasing indigenous vegetable yield and nutritional quality through traditionally-and scientifically-informed soil fertility management. University of California, Davis, USA. Not complete!

Plazas, M., Andujar, I., Vilanova, S., Gramazio, P., Herraiz, F.J. and Prohens, J. 2014. Conventional and phenomics characterization provides insight into the diversity and relationships of hypervariable scarlet (*Solanum aethiopicum* L.) and gboma (*S. macrocarpon* L.) eggplant complexes. *Frontiers in Plant Science* 5: pages missing? https://doi.org/10.3389/fpls.2014.00318

Saraçli, S., Doðan, N. and Doðan, Ý. 2013. Comparison of hierarchical cluster analysis methods by cophenetic correlation. *Journal of Inequalities and Applications* 1:203. Looks not ok!

Silva, A.R. and Dias, C.T.S. 2013. A cophenetic correlation coefficient for Tocher's method. *Pesquisa Agropecu?ria Brasileira* 48(6):589-596. https://doi.org/10.1590/S0100-204X2013000600003

Von Grebmer, K., Bernstein, J., de Waal, A., Prasai, N., Yin, S. and Yohannes, Y. 2015. 2015 Global hunger index: Armed conflict and the challenge of hunger. International Food Policy Research Institute. Not complete!

World Vegetable Center. 2017. Annual Report 2016 (No. Publication 17-814) Shanhua, Taiwan: World Vegetable Center. p. 67.