Pre evaluation of cassava (Manihot esculenta Crantz) germplasm for genotypic variation in the identification of K efficient genotypes through different statistical tools

Article *in* Physiology and Molecular Biology of Plants · August 2020

DOI: 10.1007/s12298-020-00867-2

7 authors, including:

Susan JOHN Kuzhivilayil
ICAR-Central Tuber Crops Research Institute
59 PUBLICATIONS **335** CITATIONS
SEE PROFILE

J. Sreekumar
91 PUBLICATIONS **830** CITATIONS
SEE PROFILE

M.N. Sheela
ICAR-Central Tuber Crops Research Institute, Thiruvanathapuram
48 PUBLICATIONS **140** CITATIONS
SEE PROFILE

Sanket J. More
Central Tuber Crops Research Institute
55 PUBLICATIONS **76** CITATIONS
SEE PROFILE

Physiol Mol Biol Plants
https://doi.org/10.1007/s12298-020-00867-2

RESEARCH ARTICLE

Pre evaluation of cassava (*Manihot esculenta* Crantz) germplasm for genotypic variation in the identification of K efficient genotypes through different statistical tools

K. Susan John[1] · J. Sreekumar[2] · M. N. Sheela[3] · S. U. Shanida Beegum[1] · Sanket J. More[1] · G. Suja[1]

Received: 25 May 2020 / Revised: 25 July 2020 / Accepted: 14 August 2020

Abstract Among the tropical tuber crops, cassava (*Manihot esculenta* Crantz) deserves special attention as regards to its higher biological efficiency in terms of dry matter production which incidentally implies to the higher amount of nutrient extraction from the soil resulting in better response to the application of manures and fertilizers. Among the major nutrients, Potassium (K) is considered as the key nutrient for cassava owing to its influence both in tuber yield and tuber quality. The above facts as well as the availability of sufficient cassava genotypes in the germplasm collection of ICAR-CTCRI made us to initiate research work to screen cassava germplasm including the pre breeding lines. The objective being to identify K efficient genotypes which can yield well under limited availability of K so that the external application of K can be reduced. This paper describes the wide variation noticed during the pre evaluation of 83 elite genotypes which was done as a prelude in the screening and identification of K efficient genotypes. The characters studied were tuber yield, tuber characters, plant dry matter percentage, plant K content, tuber quality (starch, cyanogenic glucosides) attributes, physiological efficiency and plant biometric characters. The variation among the genotypes for the above traits was assessed by making some yardstick for classification which in turn helped in determining the percent distribution of genotypes in each category. The variation among the genotypes were further affirmed through principal component analysis, wherein the first five components explained more than 77% of variability and the cluster analysis performed grouped these genotypes into five clusters. The biplot showed the traits which are closely linked to the genotypes. The dendrogram constructed indicated similar genotypes to that of the clusters to the extent of more than 50% revealing the association of members with similar traits in clusters and dendrograms. The study helped in establishing the drastic variation among the genotypes along with identification of six genotypes viz., Aniyoor, 7 Sahya (2), 7 III E3-5, W-19, CR 43-8, 6-6 for further detailed experimentation to identify K efficient genotypes.

Keywords Cassava · Genotypic variation · Physiological efficiency · Principal component analysis · Cluster analysis · Biplot · Dendrogram

✉ K. Susan John
susanctcri@gmail.com

J. Sreekumar
sreekumarctcri@gmail.com

M. N. Sheela
sheelactcri@yahoo.co.in

S. U. Shanida Beegum
siyas31@gmail.com

Sanket J. More
sanketmore1818@gmail.com

G. Suja
sujagin@yahoo.com

[1] Division of Crop Production, Indian Council of Agricultural Research-Central Tuber Crops Research Institute, Sreekariyam, Thiruvananthapuram, Kerala 695 017, India

[2] Section of Social Sciences, Indian Council of Agricultural Research-Central Tuber Crops Research Institute, Thiruvananthapuram, Kerala 695 017, India

[3] Division of Crop Improvement, Indian Council of Agricultural Research-Central Tuber Crops Research Institute, Thiruvananthapuram, Kerala 695 017, India

Published online: 26 August 2020

Introduction

Tropical tuber crops play a significant role in the food and nutritional security of millions of people globally especially in the developing and underdeveloped countries. Among the tropical tuber crops, cassava is the most important with respect to the area under cultivation, productivity, ability to thrive under marginal soil and environmental conditions, tolerance to pests and diseases, the high and quality starch content of the tubers which in turn can be utilized for the preparation of many value-added products of industrial use. Being the National repository of tuber crops germplasm in India, ICAR-Central Tuber Crop Research Institute (CTCRI), Kerala, India conserves more than 2000 genotypes comprising of indigenous, exotic, landraces and breeding lines including wild relatives of cassava. Cassava breeding research was started about 50 years back after the establishment of CTCRI in 1963 and 20 high yielding varieties of cassava were released so far (Edison and Sheela 2008; Malik et al. 2020). Out of the 50,000 edible plant species, cassava is one among the 15 plant species with tuber as the economic produce providing 90% of the world's total food energy intake (More et al. 2019).

The research experience under the long term fertilizer experiment (LTFE) at ICAR-CTCRI since 1977 (Susan John et al. 2005) clearly revealed the strong positive response of cassava to manures and fertilizers. Among the major nutrients, potassium (K) is considered as the 'key nutrient' with respect to its significant role in increasing tuber yield and improving tuber quality through reduction in cyanogenic glucosides responsible for bitterness in cassava tubers and enhancing the tuber starch content through its role in influencing the starch synthetase enzyme responsible for starch production (Susan John et al. 2010). The package of practices (PoP) recommendation of K for cassava is 100 kg K_2O/ha. The source of K usually used is muriate of potash (MOP) having 60% K_2O. MOP is an imported chemical fertilizer, its price for unit of K is high and sometimes, it is not available at times of need. Moreover, as the laterite soils (Ultisols) of Kerala where the study was carried out was under humid tropical climatic situation, these soils are very low in K (< 150 kg ha^{-1}) due to K fixation. K efficient genotypes being worthy in low K soils, attempt was initiated by utilizing the germplasm conserved at ICAR-CTCRI, to identify K efficient genotypes which can mobilize the fixed soil K for plant uptake so that the external application of K can be reduced.

The selection of 83 genotypes for pre evaluation to identify K use efficient genotypes was based on key traits like tuber yield, tuber shape, starch content, plant architecture, harvest index and CMD tolerance of the genotypes available at the field gene bank of cassava comprising of landraces and pre breeding lines (Sheela et al. 2008; Ceballos et al. 2016).

This study forms a part of the pre evaluation we have conducted in these genotypes as a prelude of the mega activity to identify the K use efficient genotypes. The objective of the preliminary screening was to see the variation among genotypes and to identify genotypes with better plant traits relevant for K use efficiency. The important characters studied were growth (plant height, stem girth) attributes, plant (leaf, stem, tuber) dry matter percentage, tuber yield and tuber characters (number of tubers, length, girth), plant K content, physiological efficiency (PE) of K, tuber quality attributes (starch and cyanogenic glucosides), as they have direct or indirect bearing on K use efficiency. In addition, some general plant attributes like branching, percent sprouting, cassava mosaic disease (CMD) tolerance and flowering which are important in the final selection of K efficient genotypes also were considered. It is known that variation among genotypes with respect to characters related to soil fertility and plant nutrition management is ultimately assessed in terms of nutrient use efficiency (NUE) (Baligar and Fageria 1997). According to Graham (1984), the genetic potential of a crop differ widely among plant species and cultivars within species also differ in their absorption and utilization of nutrients and such differences are attributed to morphological, physiological and biochemical processes in plants and their interaction with climatic, soil, fertilizer, biological and management practices. Moreover, there are several reports indicating nutritional differences among cultivars and strains of plants to inorganic plant nutrition due to genetic makeup (Clark and Duncan 1991; Duncan and Carrow 1999). Here, we aimed to elucidate the distinct genetic variation with respect to the above plant attributes that exists among the genotypes by employing some arbitrary yardsticks for the observations taken on plant attributes. Principal component analysis (PCA) was done to see the extent of variability through the PC's and the contribution of the plant attributes to each PC's. Cluster analysis grouped genotypes with similar traits in one cluster and biplot and dendrogram methods were also done to confirm the close linkage of characters among genotypes belonging to same clusters/groups.

Since the soil under which the present study carried out was low in K, identification of genotypes with better attributes as above can definitely be a criteria in the delineation of K use efficient genotypes. In addition to the above objectives, identification of good performing genotypes can be a source material for further breeding process to evolve K use efficient hybrids.

Table 1 Description of the cassava genotypes used in the pre evaluation study to screen K use efficient genotypes

Sl. No	Genotypes	Origin	Sl. No	Genotypes	Origin
1	Aniyoor	Landrace, Kerala	43	2–18	Breeding line, CTCRI
2	IH3/2	Breeding line, CTCRI	44	Ambakkadan	Landrace, Kerala
3	Sree Rekha	Released variety, CTCRI	45	7/39	Breeding line, CTCRI
4	CR43-11	Breeding line, CIAT	46	CR 9A 125	Breeding line, CIAT
5	Sree Sahya	Released variety, CTCRI	47	7 III C2-5	Breeding line, CTCRI
6	H-97	Released variety, CTCRI	48	I H5-8	Breeding line, CTCRI
7	Sree Vijaya	Released variety, CTCRI	49	Kadakkal	Landrace, Kerala
8	IH5/2	Breeding line, CTCRI	50	Ummanvella	Landrace, Kerala
9	4/31	Breeding line, CTCRI	51	Neelagiri	Landrace, Tamil Nadu
10	4/21	Breeding line, CTCRI	52	CR 54 A-3	Breeding line, CIAT
11	Kalpaka	Released variety, KAU*	53	New 1	Landrace, Kerala
12	7 IVA3-I	Breeding line, CTCRI	54	25/26	Breeding line, CTCRI
13	CR-43-8	Breeding line, CIAT	55	MN7	Breeding line, CTCRI
14	7 III E 1-6	Inbred line, CTCRI	56	Ullichuvala	Landrace, Kerala
15	H-226	Released variety, CTCRI	57	CR 59-8	Breeding line, CIAT
16	Sree Jaya	Released variety, CTCRI	58	7 Sahya (2)	Inbred line, CTCRI
17	Sree Prabha	Released variety, CTCRI	59	New-2	Landrace, Kerala
18	H-165	Released variety, CTCRI	60	99/14(3)	Inbred line, CTCRI
19	C-59/8R	Breeding line, CTCRI	61	43-7	Breeding line, CTCRI
20	Sree Prakash	Released variety, CTCRI	62	7 ulli-2	Breeding line, CTCRI
21	II D 79-6	Inbred line, CTCRI	63	Vellayani Hraswa	Released variety, KAU*
22	I D2 (6-7)	Inbred line, CTCRI	64	H-1687	Released variety, CTCRI
23	I D 7C 1-3	Inbred line, CTCRI	65	7-99 MNA	Breeding line, CTCRI
24	C-1848	Landrace, Kerala	66	6-2 MN4	Breeding line, CTCRI
25	C-21	Landrace, Kerala	67	CR 5/8	Breeding line, CIAT
26	M4 H	Breeding line, CTCRI	68	7MN6	Breeding line, CTCRI
27	7 III E3-5	Inbred line, CTCRI	69	TEMNI	Land race, Kerala
28	7 IV E3-5	Inbred line, CTCRI	70	7/49/MN3	Inbred line, CTCRI
29	7 III C 8-2	Inbred line, CTCRI	71	7MN2	Land race, Kerala
30	4-2	Triploid variety, CTCRI	72	CR43-2	Breeding line, CIAT
31	5-3	Triploid variety, CTCRI	73	4-21	Breeding line, CTCRI
32	7 IVC 4-4	Inbred line, CTCRI	74	CR 43-7	Breeding line, CIAT
33	6-6	Triploid line, CTCRI	75	35/8(2)	Breeding line, CTCRI
34	17/5	Breeding line, CTCRI	76	CR114-0	Breeding line, CIAT
35	16-12	Breeding line, CTCRI	77	43-11	Breeding line, CIAT
36	IH5/15	Breeding line, CTCRI	78	CR-26-1	Breeding line, CIAT
37	35/8	Breeding line, CTCRI	79	T Amba	Landrace, Kerala
38	43-2	Breeding line, CTCRI	80	CR43-6	Breeding line, CIAT
39	W-19	Breeding line, CTCRI	81	4-31	Breeding line, CTCRI
40	4/3	Breeding line, CTCRI	82	25/2	Breeding line, CTCRI
41	MNGA	Released variety, CTCRI	83	CR 43-5	Breeding line, CIAT
42	Sree Harsha	Released variety, CTCRI			

KAU Kerala Agricultural University

Materials and methods

A total of 83 elite cassava genotypes (Table 1) were planted in a row trial with 10 plants each in a row without external application of any fertilizers in a laterite soil at Block V of the experimental farm of ICAR-CTCRI. The soil of the experimental site was sandy clay loam with acidic pH (4.5–5.0), medium in organic matter (0.6–0.75%), low in available N (110–150 kg ha^{-1}), high

in available P (above 50 kg ha^{-1}) and low in exchangeable K (< 110 kg ha^{-1}).

The cut stems called 'setts' were used as planting material. Periodic growth observations, plant dry matter percentage and plant K contents were taken by destructive sampling at 3, 6 and 9 months after planting (MAP). Destructive sampling at these intervals were carried out to see the variation in the above parameters over time and is taken for the purpose of identifying genotypes with better characters. However, in this study, the mean of the three observations for the parameters indicated were included.

In the case of the basic general parameter viz., percent establishment of the setts, number of setts sprouted out of 100 setts planted was taken. As regards to CMD tolerance, it was randomly assessed using virus indexing (Hahn et al. 1980) to see the extent of CMD tolerance in the selected genotypes. In the case of branching, the classes included were branching, top branching and non branching. Observation on flowering was also taken.

Plant height was measured with a metre scale from the base of the lengthiest branch to the topmost new leaf. Stem girth was taken as the diameter of the base of the same branch measured with a twine and the length of the twine was taken in a 30 cm scale. The plants were harvested after 9 months of planting and tuber yield per plant was taken along with number of tubers per plant, length and girth of three tubers selected randomly. Length of the tuber was measured using a 30 cm scale. In the case of tuber girth, the mean of the diameter at three portions of the tuber measured using a twine converted to 30 cm scale was employed. The major determinant used for screening the genotypes for further experiments to identify the K use efficient genotype was the inherent nutrient use efficiency of the crop for K termed as physiological efficiency (PE) of K computed following the formulas suggested by Soon (1992).

PE (K) = Biological yield (BY) (kg/plant)/K uptake (kg/plant). Biological yield is the total of the vegetative and tuber yields. Since the nutrient uptake is on dry weight basis, the yield of vegetative parts like leaves and stem and economic part like storage root (tuber) was taken on dry weight basis. For computing the biological yield on dry weight basis, destructive sampling as indicated earlier were done and the fresh weight of plant parts like leaf, stem and tuber were taken and 50 g each of these samples were kept for drying in an oven at 65 ± 5 °C till we got stable results. From the dry weight percentage and fresh weight of each plant part per plant, the dry matter yield of each plant part per plant was arrived. Adding leaf, stem and tuber dry matter yields, the BY per plant was obtained on dry weight basis.

For computing the K uptake, the K content in the leaf, stem, tuber were determined. For this, after collection and processing, the plant K was analysed using triacid digestion with nitric: perchloric: sulfuric acids in the ratio 10:4:1 followed by direct reading in flame photometer (Systronics 128) (Singh et al. 2005). By multiplying the dry matter, yield of each plant part with their respective K content, the K uptake of each plant part was obtained. After adding the K uptake of leaf, stem and tuber, total K uptake of the plant was computed (AOAC 1984).

Tuber quality attributes like cyanogenic glucosides and starch content in cassava tubers were determined following the procedure suggested by Padmaja et al. (2005) and both are expressed on fresh weight basis.

The variation among genotypes with respect to the characters studied was initially attempted through an arbitrary classification by fixing an upper and lower value for each parameter and calculating the percentage of genotypes falling under each class. This was further scientifically tested following different statistical tools like principal component and cluster analysis, biplot and dendrogram methods.

Statistical analyses

The statistical analyses performed included principal component, cluster, biplot and dendrogram analysis. For performing these analyses, as the data set need to be complete, the missing values observed in the data set were imputed using Fully Conditional Specification (FCS) implemented by the Multivariate Imputation by Chained Equations (MICE) algorithm as described in Van Buuren and Groothuis-Oudshoorn (2011). MICE package in R environment was used to perform the statistical computation.

The principal component analysis (PCA) was carried out to extract the variation in terms of the principal components which in turn reflects the importance of the largest contributor to the total variation at each axes of differentiation. Cluster analysis was undertaken to group the genotypes with similar traits with respect to the characters studied. Hierarchical cluster analysis with complete linkage method has been suggested for classifying entries of germplasm collections based on the degree of similarity and dissimilarity (Van Hintum 1995). A combination of cluster and principal component analysis as suggested by Crossa et al. (1995) also had been used to classify the 83 accessions. PCA and Clustering were carried out using the R environment for statistical computing (R version 3.4.1). Biplot and dendrogram analysis also was done to see the associated genotypes with similar traits in each group.

Results and discussion

The variations observed with respect to the different plant characters studied in the preliminary evaluation trial was analysed using different methods and is summarised below. The details as per the arbitrary concept indicating the variation among genotypes as the percentage of genotypes in each group is as follows.

General plant characters It is known that, general traits studied are very important in the selection of genotypes for specific traits for popularization of the identified K efficient genotype or breeding purpose. So, we have taken ample care in finding the percent distribution of 83 genotypes under the various general plant characters. It is seen that, there exists distinct variation among the different genotypes as 54, 34 and 12% are branching, non-branching and top branching respectively. As regards to the sprouting/establishment of the planted cassava stem cuttings, 54% of the genotypes recorded more than 75% establishment. The incidence of CMD at sprouting ranged from 0 to 100% with a mean value of 49%. At this stage, 44% showed more than 75% CMD incidence and 6% had less than 25% of CMD incidence. At harvest, CMD was severe in 43% of the genotypes, less severe in 12% and there was no CMD noticed in 45% of the genotypes as per the procedure undertaken based on Hahn et al. (1980). As regards to the flowering, 31% was flowering and the rest were non-flowering. These type of observations presently made in this evaluation is in conformity with the reports indicating the variation with respect to different general plant morphological and phenotypic attributes as in *Acacia* (Daehler et al. 1999), tomato (Kouam et al. 2018), hop (McAdamet al. 2014) to trace the genetic variability in the selection of superior genotypes for crop improvement.

Biometric characters The plant height ranged from 1.25 to 3.35 m with a mean value of 2.43 m (Fig. 1). However, 4% of the genotypes had less than 1.5 m and 7% had more than 3 m height. The mean stem girth of the genotypes was 9.24 cm with values ranging from 6.0 to 13.58 cm where 70% of the genotypes had 8–11 cm girth and 10% had more than 11 cm stem girth. However, there were 20% genotypes with less than 8 cm girth. This is in agreement with the reports of Jordan et al. (1979) and Dar et al. (2018) in maize, Agong et al. (2001) in tomato, McAdam et al. (2014) in hop, Silva et al. (2019) in castor observing drastic variation in growth parameters during the evaluation of germplasm for the selection of better genotypes for breeding purpose.

Tuber characters and tuber yield The mean tuber number of the genotypes was seen as nine with the number of tubers per plant ranging from 4 to 15. Out of the total, 76% of the genotypes recorded 5–11 tubers per plant, 7% with less than five and 1% with more than 14 tubers per plant. The tuber length ranged from 13 to 52 cm with a mean value of 32 cm. Out of the total genotypes, 79% had 25–40 cm tuber length and there was 1% with less than 20 cm tuber length and 7% having more than 40 cm tuber length. In the case of tuber girth, the mean was 18 cm and it ranged from 11 to 24 cm. A total of 70% genotypes had 15–21 cm tuber girth and there were 4% genotypes with less than 12 cm tuber girth and 16% having more than 21 cm tuber girth.

The tuber yield ranged from 0.01 to 7.52 kg per plant with a mean value of 3.24 kg per plant. It is found that 53% yielded 2–4 kg per plant and 8% of the genotypes had less than 1 kg tuber yield and 7% with more than 6 kg tuber yield per plant. These observations are in conformity with the findings of McAdam et al. (2014) in hop, Chowdhury et al. (2016) in soybean and Otayk (2019) in wheat indicating significant differences in yield and yield attributes in the process of estimating the quantitative genetic parameters for the genetic improvement of these crops.

Plant dry matter percentage The leaf dry matter percentage (LDMP) ranged from 21.92 to 32.44% with a mean value of 27.50%. A total of 77% genotypes had LDMP in the range of 25–30% with 8% of the genotypes with less than 25% and 15% with more than 30% LDMP. Similarly, the stem dry matter percentage (SDMP) ranged from 26.83 to 48.24% with a mean value of 34.92%. Out of the total, 51% of the genotypes had SDMP in the range of 35–40%. There was observed only one genotype with more than 45% and 13 genotypes with less than 30% SDMP. The tuber dry matter percentage (TDMP) ranged from 12.14 to 39.46% with a mean value of 29.04%. Majority of the genotypes (63%) had TDMP in the range of 25–35% (Fig. 2). These conclusions adhere to the observations of Jordan et al. (1979) in sorghum and Iqbal et al. (2019) in cotton revealing significant variation in root and shoot dry weight while studying the genetic variation among genoptypes under phosphorus use efficiency trials.

Plant K content The leaf K content ranged from 0.84 to 2.03% with a mean value of 1.34%. Out of the total, 49% had leaf K in the range of 1.25–1.5% and there were 7% genotypes with less than 1% and 1% genotype with more than 2% leaf K. The mean stem K content was 0.97% but ranged from 0.445 to 1.742%. A total of 65% had the stem K content ranging from 0.75 to 1.25%. However, there were 4% genotypes each with less than 0.5% and more than 1.5% stem K. The tuber K content ranged from 0.490 to 1.839% with a mean of 1.149%. A total of 58% genotypes had tuber K in the range of 1–1.5% and there were 12% genotypes with less than 0.75% and 7% genotypes with more than 1.5% tuber K (Fig. 3). These findings were supported by the views of Lopez et al. (2008) and Wang and Chen (2012) in cotton while studying the K uptake and

Fig. 1 Contrasting canopy architecture of genotypes under different clusters

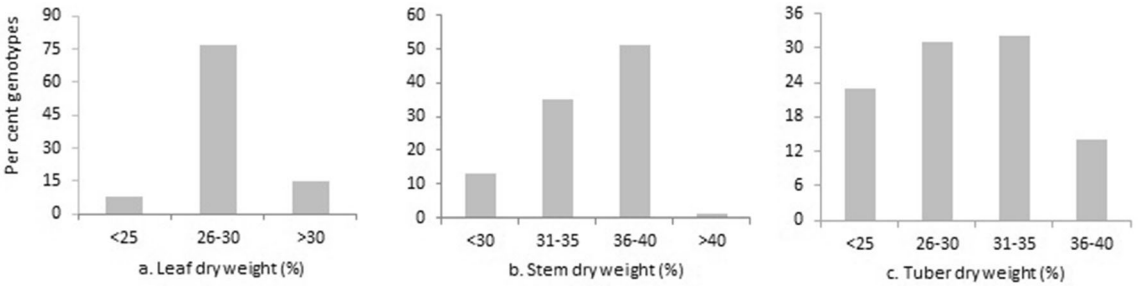

Fig. 2 Variation in leaf, stem and tuber dry matter content (%) of the genotypes

K use efficiency with respect to different genotypes where they observed significant variation in the leaf, stem and boll K content and hence its uptake.

Physiological efficiency The PE of K computed ranged from 68 to 244 kg biological yield per kg K uptake with a mean value of 106 kg/kg. The distribution of PE of the genotypes is shown in Fig. 4. Out of the total, 79% of the genotypes had PE of K ranging from 75 to 125 kg/kg. However, there were 5% genotypes with PE less than 75 kg/kg and 2% with PE more than 175 kg/kg (Fig. 4). Isfan (1990) indicated physiological nutrient efficiency is a genetic trait of the genotype, which can be used in the breeding program to detect high yielding potential genotypes. Sadegh (2017) studied the physiological efficiency of three cultivars of soybean under different levels of potassium and found PE(K) is significant under different levels of K.

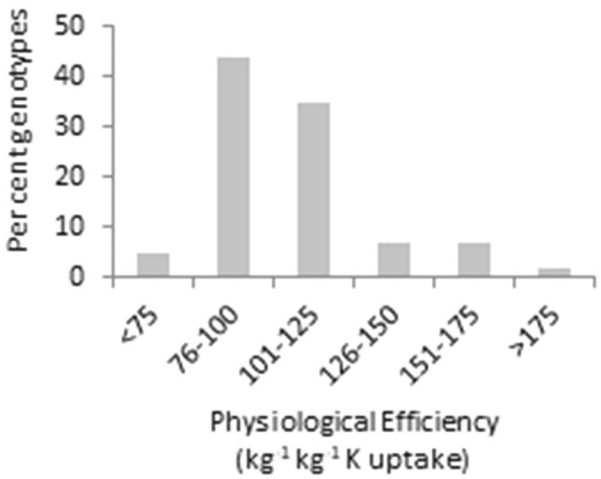

Fig. 4 Variation in physiological efficiency (K) of the genotypes

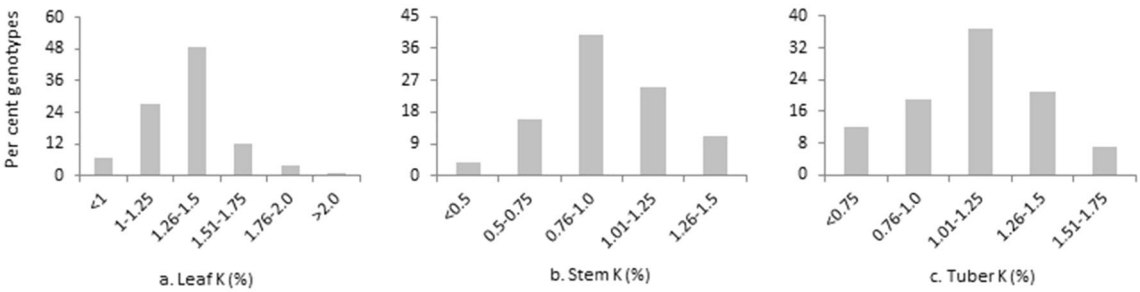

Fig. 3 Variation in K content (%) of the leaf, stem and tubers of the genotypes

Tuber quality attributes The starch content on fresh weight basis varied from 6.07 to 25.80% with a mean value of 16.31%. A total of 52% had tuber starch in the range of 15–20%. However, there were 4% genotypes with less than 10% and 3% genotypes with more than 25% tuber starch (Fig. 5a). As regards to cyanogenic glucoside (HCN) content in cassava tuber, it ranged from 24 to 378 ppm with a mean value of 97 ppm. A total of 53% had HCN in the range of 30–90 ppm. There were 4% genotypes with less than 30 ppm and 7% genotypes with more than 180 ppm tuber cyanogen contents (Fig. 5b). These observations confirm to the studies of McAdam et al. (2014) who observed significant genetic variation among different species of hop of the two families for biochemical traits like colupulone, α-acid and β-acid. Agong et al. (2001) also found significant variation in biochemical quality attributes of different species of tomato.

Statistical tools employed in the genotypic analyses under the pre evaluation trial

The different statistical tools used included PCA, cluster, biplot and dendrogram analysis. Multivariate analysis such as principal component and cluster analysis require complete data set for performing exploratory data analysis and classification of the observations. If there are missing values in the data set, the analysis followed by its interpretation will deprive of some valuable information. If the number of missing values are very small compared to the total size of the data, leaving out few samples with missing features can be adopted. Usually, while dealing with large number of data like this, we come across missing values and the solution is not so trivial. Under such situation, Multivariate Imputation via Chained Equations (MICE) is one of the commonly used packages by R users to impute missing values with plausible data values. These plausible values are drawn from a distribution specifically designed for each missing data point. Here, we used the MICE package in R with *method* = *pmm* for imputing the missing values. Predictive Mean Matching (PMM) is a semi-parametric imputation approach. It is similar to the regression method except that for each missing value, it fills in a value randomly from among the observed donor values from an observation whose regression-predicted values are closest to the regression-predicted value for the missing value from the simulated regression model as suggested by Heitjan and Little (1991) and Schenker and Taylor (1996).

Principal component analysis (PCA)

Here, the PCA was carried out on the standardized dataset (after employing MICE) comprising of 83 genotypes. PCA resulted in extracting six principal components which contributed 77% of the total variability and is presented in Table 2. Sett establishment, plant height, tuber girth, tuber yield and number of tubers per plant contributed significantly to PC1 whereas stem girth, tuber length and stem K% contributed to PC2. Stem dry weight%, leaf K% and leaf dry weight% had significant effect on PC3. PC4 is found influenced by leaf dry weight%, HCN content, sett establishment and PE. The Leaf K%, tuber dry weight% and starch content contributed to PC5. Leaf K%, starch% and HCN content contributed to PC6. Hughes et al. (2015) reported that, the principal component analysis combined with clustering can be a better tool for genetic divergence studies which in turn can be further exploited in crop improvement including breeding programmes.

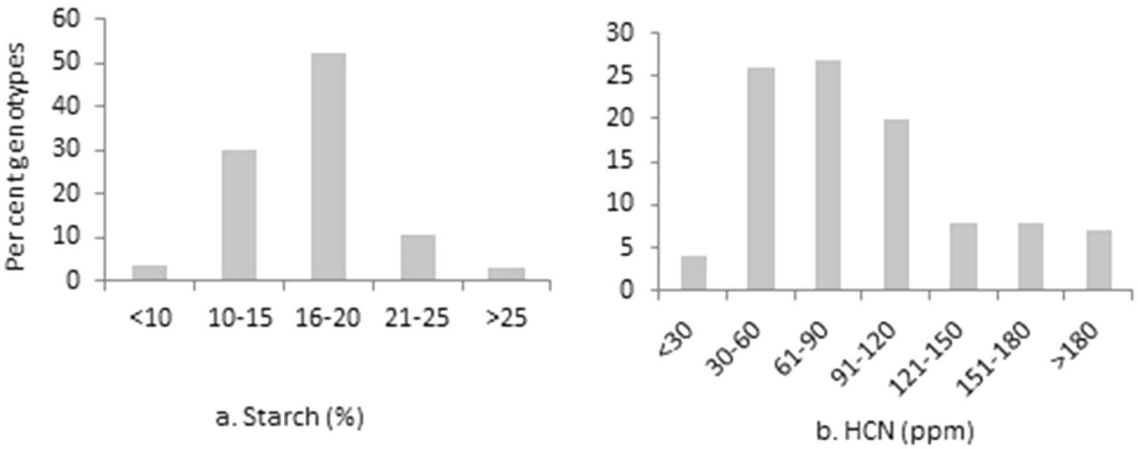

Fig. 5 Variation in tuber quality attributes of the genotypes

Table 2 PCA analysis of the studied traits of cassava genotypes

Variable	PC1	PC2	PC3	PC4	PC5	PC6
Eigen value (Root)	2.040	1.743	1.562	1.177	1.027	1.002
% variation expressed	24.5	17.9	14.4	8.1	6.2	5.9
Cumulative variation expressed (%)	24.5	42.3	56.7	64.5	71.1	77.0
Establishment	− 0.379	0.229	− 0.158	0.373	− 0.042	0.146
Sprouting	− 0.091	0.266	− 0.264	− 0.108	− 0.187	− 0.038
Plant height	− 0.387	− 0.122	− 0.216	− 0.0002	− 0.113	0.285
Stem girth	− 0.232	− 0.372	0.060	− 0.058	− 0.050	− 0.048
Leaf dry weight	− 0.125	0.115	0.344	0.412	− 0.209	0.329
Stem dry weight	− 0.281	0.204	0.128	− 0.143	− 0.230	− 0.226
Tuber dry weight	− 0.173	0.444	0.154	0.112	0.359	− 0.022
Leaf K	− 0.042	− 0.058	− 0.365	0.033	0.526	− 0.409
Stem K	0.169	− 0.251	− 0.304	0.120	0.300	0.444
Tuber K	− 0.131	− 0.076	− 0.472	− 0.147	− 0.285	− 0.099
PE	0.068	− 0.090	0.355	− 0.577	0.033	− 0.169
Starch	− 0.235	0.411	− 0.011	0.010	0.364	0.092
HCN	− 0.056	− 0.177	0.102	0.611	− 0.012	− 0.418
Tuber number	− 0.326	− 0.237	0.033	− 0.205	0.129	0.253
Tuber length	− 0.241	− 0.260	0.219	− 0.021	0.307	0.014
Tuber girth	− 0.370	− 0.076	− 0.114	0.086	− 0.138	− 0.290
Tuber yield	− 0.330	− 0.246	0.222	− 0.004	0.107	− 0.024

Cluster analysis

The hierarchical cluster analysis performed with complete linkage methods on the Euclidean distance matrix of the 83 genotypes using R package resulted in five clusters. The results of the cluster membership is presented in Table 3. The number of genotypes in clusters 1,2,3,4 and 5 respectively were 9, 48, 17, 5, 4. The members of the clusters had almost similar traits. Among all the clusters, the Cluster 2 had 48 members followed by Cluster 3 with 17 genotypes. The mean values of the different parameters of the five clusters with the grant centroid values are presented in Table 4.

Mean of the plant characters of the clusters over the grant centroid along with the important members in each cluster as well as the characters important for each cluster are described as follows. Among the five clusters, cluster 1 comprised of some important genotypes viz., Aniyoor, 7 III E3-5 and 7 Sahya 2 which were selected later as K efficient based on detailed K level experimentation. These genotypes were selected from pre evaluation trial as candidates for further detailed K experimentation due to their high PE(K) coupled with high values on characters like % sett establishment, plant height, stem girth, leaf, stem, tuber dry weight percentage, starch%, tuber characters viz., number, length, girth and yield. After strict K level experimentation

Table 3 Cluster composition of the evaluated genotypes of cassava for K use efficiency

Cluster	Number of genotypes	Cluster members
1	9	Aniyoor, C59/8R, ID2(6-7), 7 III E3-5, CR54 A-3, MN7, 7 Sahya (2), 43-7, TEMNI
2	48	IH3/2, Sree Rekha, CR 43-11, Sree Sahya, H97, Sree Vijaya, IH5/2, 4/31, 4/21, Kalpaka, 7 IVA3-I, CR 43-8, 7III E-1-6, H 226, Sree Jaya, Sree Prabha, H 165, Sree Prakash, II D 79-6, I D 7C 1-3, C-1848, C-21, M4 H, 7 IV E3-5, 7 III C 8-2, 4-2, 5-3, 6-6, 17/5, 16-12, IH5/15, 43-2, W-19, 4/3, Sree Harsha, 2-18, CR 9A 125, Kadakkal, Ummanvella, New 1, 25/26, Ullichuvala, CR 59-8, New 2, 99/14(3), 7 Ulli-2, Vellayani Hraswa, 6-2MN4
3	17	7 IVC 4-4, 35/8, 7 III C2-5, Neelagiri, CR5/8, 7/49/MN3, 7MN2, CR43-2, 4-21, CR43-7, 35/8(2), CR114-0, CR 26-1, 25/2, CR43-5, 43-11, 4-31
4	5	MNGA, Ambakkadan, I H5-8, 7-99 MNA, T Amba
5	4	7/39, H1687, 7MN6, CR43-6

Table 4 Mean values of the different parameters of the five cluster

Parameters	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Grand centroid
Establishment (%)	78.70	89.10	15.80	28.80	29.50	48.38
Sprouting (%)	26.0	64.10	12.80	47.20	50.00	40.02
Plant height (m)	2.59	2.56	1.84	1.62	2.59	2.24
Stem girth (cm)	9.63	9.18	8.73	7.58	12.54	7.73
Leaf dry weight (%)	28.67	27.51	26.73	27.49	24.87	27.05
Stem dry weight (%)	35.80	35.51	29.83	31.87	30.87	32.78
Tuber dry weight	31.61	30.63	22.62	32.09	17.92	26.97
Leaf K (%)	1.25	1.37	1.32	1.17	1.40	1.30
Stem K (%)	0.68	0.97	1.22	0.88	1.19	0.99
Tuber K (%)	0.95	1.23	1.02	0.84	1.41	1.09
PE (kg/kg)	135.70	102.80	113.90	122.0	161.0	127.08
Starch (%)	16.76	17.64	11.51	16.60	9.66	14.43
HCN (ppm)	106.82	93.83	113.92	66.84	76.45	91.57
Tuber number	11.22	8.81	7.35	4.60	9.75	8.35
Tuber length (cm)	37.33	31.67	29.82	20.00	38.25	31.41
Tuber girth (cm)	21.33	19.04	15.41	12.00	19.25	17.41
Tuber yield (kg/plant)	6.50	3.07	2.47	1.00	3.43	3.29

in a split plot design for 3 years and testing under farmer participatory trials, Aniyoor and 7 III E3-5 were identified as K efficient which could yield well without K and at half the recommended dose of K as 50 kg/ha K_2O. The genotypes under this cluster possessed the maximum plant height (2.59 cm), leaf (28.67%) and stem (35.80%) dry weight percentage, tuber number (11.22), tuber girth (21.33 cm) and tuber yield (6.50 kg/plant). Some released hybrids viz., H 97, H 226, H 165, Sree Rekha, Sree Sahya, SreeVijaya, Kalpaka, Sree Jaya, Sree Prabha, Sree Prakash, Sree Harsha, Vellayani Hraswa and some very popular local cultivars were seen grouped in cluster 2. The genotypes under this cluster had the highest sprouting (64.1%) and starch percentage (17.64%). In addition, they had high plant height, stem girth, leaf, stem, tuber DW%, leaf and tuber K % and tuber attributes like number, length, girth and yield. Some CMD resistant accessions like CR 5/8, CR 43/2, CR43/7, CR114-0 including a local popular cultivar 'Neelagiri' was found in cluster 3 which in turn possess highest stem K% (1.22%) in addition to high stem girth and leaf K %. Among the five genotypes in Cluster 4, MNGA and Ambakkadan are popularly known for their CMD resistance and high yield respectively. They have high sprouting %, leaf and tuber DW%, high starch and lowest cyanogen content in the tubers in addition to the highest tuber dry weight percentage (32.09%). As regards to cluster 5, it contained the most popular hybrid variety H-1687 associated with characters like high sprouting%, plant height, stem girth, leaf, stem and tuber K content, PE (K), low cyanogen and tuber atributes like number, length, girth and yield. This cluster also had maximum plant height (2.59 m), stem girth (12.54 cm), highest leaf (1.40%) and tuber (1.41%) K%, highest PE (K) (161 kg/kg) and highest tuber length (38.25 cm). These type of clustering was undertaken by Pahadi et al. (2017) in maize with the aim of identifying better performing genotypes based on important traits. Suryanarayana et al. (2017) grouped 30 genotypes of maize into six clusters based on non-hierarchical Euclidean cluster analysis in the process of studying their genetic divergence. Contrasting genotypes under different clusters are illustrated in Fig. 6.

Biplot analysis

The biplot of the studied characters and the associated genotypes is presented in Fig. 7. The biplot shows the characters which are closely linked with the genotypes or genotypes which behave similarly with respect to different plant characters. If we see the genotypes under different clusters (Table 3) along with the biplot diagram, it can be seen that, for instance, Nilgiri belonging to cluster 3 is associated with high stem K%. Similarly, 6-2 MN4 and Ummanvella in cluster 2 and TEMNI in cluster 1 had high HCN. CR 43-2 under cluster 2 have high stem girth and tuber length. 7 Sahya (2) under cluster 1 is with high tuber yield and tuber length. New 1 under cluster 2 is with high starch too. This in turn indicated the linkage of the plant characters with genotypes under biplot to the important attribute of the cluster under which the genotypes belong.

Dendrogram analysis

The cluster dendrogram (Fig. 8) divided the genotypes into 5 groups with the number of genotypes as 11, 42, 16, 12

Fig. 6 Contrasting tuber characters of genotypes under different clusters

and 2 in each group. An analysis of the linkage of the genotypes in the five clusters (Table 3) with the genotypes in the five dendrogram sections revealed the percentage analogy of the members in clusters 1, 2, 3, 4 and 5 to the genotypes in sections 3, 2, 4, 1, and 5 of the dendrogram are 44.44, 68.75, 47.06, 80 and 50% respectively.

As per cluster (Table 3) and dendrogram (Fig. 8), some important genotypes which was grouped as common in both the above analyses are Aniyoor and 7 Sahya (2) (cluster 1, dendrogram section 3), the released varieties from CTCRI viz., Sree Sahya, H 97, Sree Vijaya, H 226, Sree Prakash, Sree Jaya, Sree Harsha (Cluster 2, dendrogram section 2), CMD resistant genotypes viz., CR43-2, CR 43-7 and CR 43-5 (Cluster 3, dendrogram section 4), the CTCRI released CMD tolerant variety, MNGA, the very popular local lanrace Ambakkadan (Cluster 4, dendrogram section 1) and the most popular hybrid released from CTCRI, H 1687 (cluster 5, dendrogram section 5). The six genotypes screened from the pre evaluation trial for further field experimentation at different levels of K to identify the K efficient genotypes were Aniyoor, 7 Sahya (2), CR 43-8, W-19, 7 III E3-5 and 6-6. Among these, Aniyoor, 7 Sahya (2) and 7 III E3-5 belonged to cluster 1 which in turn had high PE(K), % sett establishment, plant height, stem girth, leaf, stem, tuber dry weight percentage, starch, tuber characters viz., number, length, girth and yield. Similarly genotypes viz., W-19, CR 43-8 and 6-6 belonging to cluster 3 possessed high sprouting and starch percentage, plant height, stem girth, leaf, stem, tuber DW%, leaf and tuber K%, and tuber attributes like number, length, girth and yield. As regards to the characters linked to these genotypes as per biplot, it was seen that, the above six genotypes are linked to most of the characters evaluated and specifically associated with characters like percent

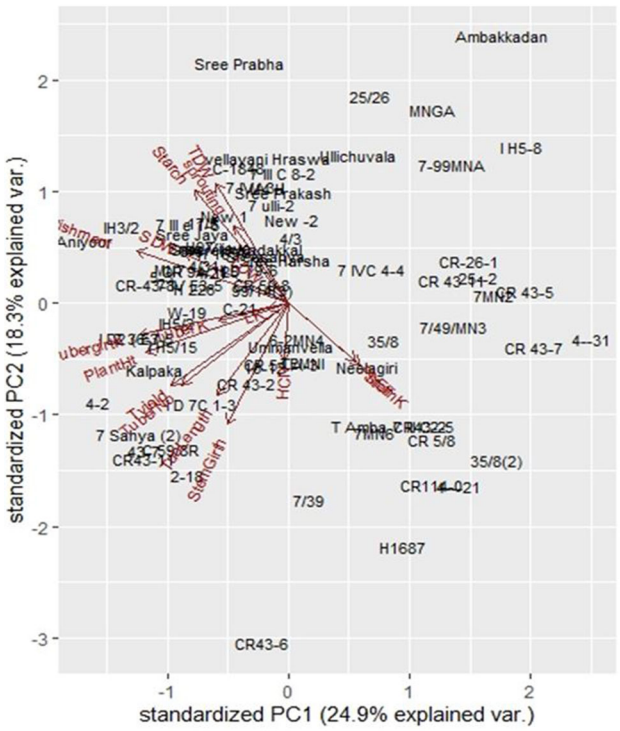

Fig. 7 Biplot comparison of 83 cassava genotypes

establishment of the setts, tuber number, tuber girth, tuber yield, starch, stem dry weigh percentage, tuber and leaf K% which in turn have a bearing on selection of genotypes for further detailed experimentation to arrive at the best K use efficient genotypes. The ultimate aim of this pre evaluation trial being to trace out the dissimilarity among genotypes and the similarity or linkage among genotypes of the same cluster/dendrogram units resulted in selecting valid genotypes for further detailed K experimentation to evolve better genotypes with specific traits as K use efficiency. Rahim et al. (2010) already carried out similar studies and showed that, if genotypes are with maximum dissimilarity/high variation, it can help in the evolution of better hybrids with the required traits like good yield and nutrient use efficiency.

Conclusion

It is important to have wide genetic variability along with better genotype traits to evolve good varieties either through selection or through breeding targeting on some specific traits. In cassava, as a prelude to screen and identify K use efficient genotypes, pre evaluation of 83 elite genotypes was done to establish the distinct variation among genotypes through different statistical tools like principal component, cluster, biplot as well dendrogram methods in addition to an arbitrary analysis. These analyses revealed the wide variation among genotypes with respect to the characters evaluated like plant dry matter percentage, plant K content, tuber yield and other tuber attributes like number, length, girth, physiological efficiency of K, growth characters like plant height and stem girth, tuber quality traits like starch and cyanogenic glucoside content of cassava tubers. The six principal components extracted could account for 77% of the variability in the genotypes along with information on the most important traits responsible for each PC's. The five clusters generated through cluster analysis grouped the genotypes based on the close linkage/

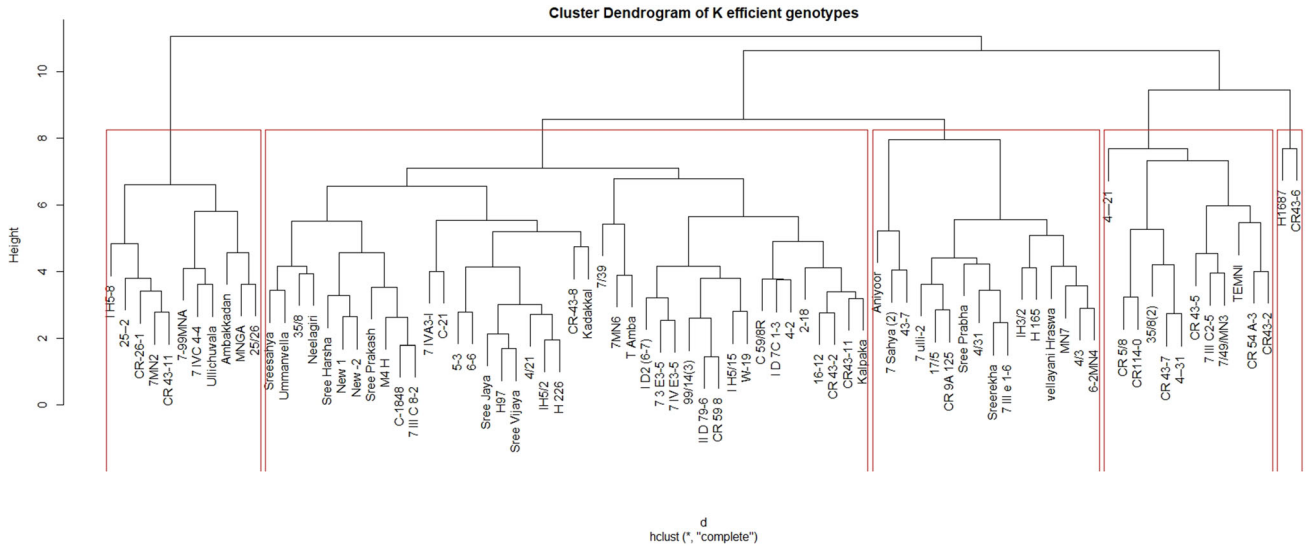

Fig. 8 Cluster analysis of the 83 cassava genotypes based on physiological efficiency

similarity of certain traits in addition to giving details of the important genotypes in each cluster. The major characters linked to different genotypes were understood through biplot. The dendrogram also separated the genotypes into five groups. A close analysis of the different statistical tools together showed that, genotypes viz., Aniyoor, 7 Sahya (2), 7 III E3-5, W-19, 6-6 and CR 43-8 selected as candidates for further detailed testing through field experiments at different levels of K to evolve K use genotypes were the common members in groups/clusters generated under the different statistical analysis. These genotypes in turn posessed important traits like percent establishment of the setts, tuber number, tuber girth, tuber yield, starch, stem dry weigh percentage, tuber, leaf K% and physiological efficiency (K) which in turn have a bearing on K use efficiency. Among these genotypes, further experiments resulted in evolving Aniyoor and 7 III E3-5 as K efficient, W-19 and CR 43-8 as N efficient. Hence, the present pre evaluation helped in confirming the wide genetic dissimilarity among genotypes as well as helped in identifying valuable genotypes with valid characters for further trials in the process of evolving K use efficient genotypes.

References

Agong SG, Schittenhelm S, Friedt W (2001) Genotypic variation of Kenyan tomato (*Lycopersiconesculentum* L.) germplasm. J Food Tech Africa 6(1):13–17. https://doi.org/10.4314/jfta.v6i1.19277

AOAC (1984) Official methods of analysis, 14th edn. Association of Analytical Chemists, Washington DC

Baligar VC, Fageria NK (1997) Nutrient use efficiency in acid soils: nutrient management and plant use efficiency. In: Monitz AC, Furlani AMC, Fageria NK, Rosolem CA, Cantarells H (eds) Plant-soil interactions at low pH: sustainable agriculture and forestry production. Brazilian Soil Science Society Campinas, Brazil, pp 75–93

Ceballos H, Perez JC, Orlando JB, Lenis JI, Morante N, Calle F, Pino L, Hershey H (2016) Cassava breeding I: the value of breeding value. Front Plant Sci. https://doi.org/10.3389/fpls.2016.01227

Chowdhury J, Karim M, Khaliq Q, Solaiman A, Ahmed J (2016) Genotypic variations in growth, yield and yield components of soybean genotypes under drought stress conditions. Bangladesh J Agric Res 40(4):537–550. https://doi.org/10.3329/bjar.v40i4.26929

Clark RB, Duncan RR (1991) Improvement of plant mineral nutrition through breeding. Field Crops Res 27:219–240

Crossa J, Delacy IH, Taba S (1995) The use of multivariate methods in developing a core collection. In: Hodgkin T, Brown AHD, Van Hintum TJL, Morales EAV (eds) Core collections of plant genetic resources. Wiley, Chichester, pp 77–92

Daehler CC, Yorkston M, DudleyN SW (1999) Genetic variation in morphology and growth characters of *Acacia koa* in the Hawaiian Islands. Int J Plant Sci 160(4):767–773

Dar TH, Shakeel R, Verma S (2018) Comparative germplasm characterization of maize (*Zea mays* L.) in Rajouri Region of PirPanjal Himalaya J & K (India), based on morphological and ISSR markers. J Crop Sci Biotechnol 21:43–55. https://doi.org/10.1007/s12892-017-0128-0

Duncan RR, Carrow RN (1999) Turfgrass molecular genetic improvement for abiotic/edaphic stress environment. Adv Agron 67:233–306

Edison S, Sheela MN (2008) Diversity of tuber crops and wild relatives. In: Kannaiyan S (ed) Agrobiodiversity hotspots and access and benefit sharing. Narosa Publishing House, New Delhi, pp 241–248

Graham RD (1984) Breeding for nutritional characteristics in cereals. In: Tinker PB, Lauchi A (eds) Advances in plant nutrition. Praiger Publisher, New York, pp 57–102

Hahn SK, Terry ER, Leuschner K (1980) Breeding cassava for resistance to cassava mosaic disease. Euphytica 29:673–683

Heitjan DF, Roderick JAL (1991) Multiple imputation for the fatal accident reporting system. J R Stat Soc C Appl 40(1):13–29. https://doi.org/10.2307/2347902

Hughes AN, Louise A, Arsene IZB, Konan KC (2015) Phenotypic diversity of farmers's traditional maize (*Zea mays* L) varieties in Cote d' Ivoire. Maydica 60(2):112–118

Iqbal A, Gui H, Zhang H, Wang X, Pang N, Dong Q, Song M (2019) Genotypic variation in cotton genotypes for phosphorus use efficiency. Agron 9(11):689

Isfan D (1990) Nitrogen physiological efficiency index in some selected spring barley cultivars. J Plant Nutr 13(8):907–914

Jordan WR, Miller FR, Morris DE (1979) Genetic variation in root and shoot growth of sorghum in hydroponics. Crop Sci 19:468–472

Kouam EB, Dongmo JR, Djeugap JF (2018) Exploring morphological variation in tomato (*Solanumlycopersicum*): acombined study of disease resistance, genetic divergence and association of characters. Agricultura Tropica Et Subtropica 51(2):71–82

Lopez M, El-Dahan MAA, Leidi EO (2008) Genotypic variation in potassium uptake in dryland cotton. J Plant Nutr 31(11):1947–1962

Malik AI, Kongsil P, Nguyen VA, Sholihin WO, Srean P, Sheela MN, López-Lavalle LAB, Utsumi Y, Lu C, Kittipadakul P, Nguyen HH, Ceballos H, Nguyen TH, Selvaraj M, Aiemnaka GP, Labarta R, Chen S, Amawan S, Sok S, Youabee L, Seki M, Tokunaga H, Wang W, Li K, Nguyen HA, Van Dong N, Ham LH, Ishitani M (2020) Cassava breeding and agronomy in Asia: 50 years of history and future directions. Breed Sci 70(2):145–166

McAdam EL, Vaillancourt RE, Koutoulis A, Whittock SP (2014) Quantitative genetic parameters for yield, plant growth and cone chemical traits in hop (*Humuluslupulus* L.). BMC Genet 15:22. https://doi.org/10.1186/1471-2156-15-22

More SJ, Ravi V, Saravanan R (2019) Tropical tuber crops. In: de Freitas ST, Pareek S (eds) Postharvest physiological disorders in fruits and vegetables. CRC Press, Boca Raton, pp 719–757. https://doi.org/10.1201/b22001

Otayk SM (2019) Evaluation of agronomic traits and assessment of genetic variability in some popular wheat genotypes cultivated in Saudi Arabia. Aust J Crop Sci 13(6):847–856

Padmaja G, Moorthy SN, Bala N, Lila B, Sundaresan S, Sajeev MS, Nanda SK, Susan John K, Rajalekshmy L, Sudha Devi KS, Manikantan NM (2005) Analytical methodologies for tropical tuber crops, laboratory manual series-2. ICAR-Central Tuber Crops Research Institute, Thiruvananthapuram, p 119

Pahadi P, Sapkota M, Thapa DB, Pradhan S (2017) Cluster and principal component analysis for the selection of maize (*Zea mays* L.) genotypes. Int J Exp Res Rev 9:5–10

Rahim MA, Mia AA, Mahmud F, Zeba N, Afrin K (2010) Genetic variability, character association and genetic divergence in mungbean (*Vignaradiata* L. Wilczek). Plant Omics 3:1–6

Sadegh B (2017) Evaluation of agronomic, physiological, agro-physiological, recovery and utilization efficiencies in three cultivars of soyabean. Biosci Biotechnol Res Commun 10(1):230–235

Schenker N, Taylor JMG (1996) Partially parametric techniques for multiple imputation. Comput Stat Data Ann 22(4):425–446

Sheela MN, Radhika VS, Susan John K, Abraham K (2008) Variation in crude protein, dry matter and starch in inbred and backcross lines of cassava. J Root Crops 34(2):115–119

Silva SA, Souza DR, Araujo GM, Almeida VO, Neves CG (2019) Genetic parameters in agronomic characters for selection of superior lineages of *Ricinuscommunis* L. Cienc Rural 49(1):e20180547. https://doi.org/10.1590/0103-8478cr20180547

Singh D, Ckonkar PK, Dwivedi BS (2005) A manual on soil, plant and water analysis. Centre for Soil-Plant-Water Analysis, Division of Soil Science and Agricultural Chemistry, Indian Agricultural Research Institute, New Delhi, Westville Publishing House, New Delhi, p 200

Soon YK (1992) Differential response of wheat genotypes to phosphorus in acid soils. J Plant Nutr 15:513–526

Suryanarayana L, Reddi Sekhar D, RatnaBabu D, Ramana AV, Srinivasa Rao V (2017) Cluster and principal component analysis in maize. Int J Curr Microbiol Appl Sci 6(7):354–359

Susan John K, Ravindran CS, James G (2005) Long term fertilizer experiments: 3 decades experience in cassava, technical bulletin series No. 45, ICAR- Central Tuber Crops Research Institute, Thiruvananthapuram, Kerala, India, p 83

Susan John K, Suja G, Sheela MN, Ravindran CS (2010) Potassium: the key nutrient for cassava production, tuber quality and soil productivity: an overview. J Root Crops 36:44–52

Van Buuren S, GroothuisOudshoorn CG (2011) MICE: multivariate imputation by chained equations in R. J Stat Softw 45(3):1–67. https://doi.org/10.18637/jss.v045.i03

Van Hintum TJL (1995) Hierarchical approaches to the analysis of genetic diversity in crop plants. In: Hodgkin T, Brown AHD, Van Hintum TJL, Morales EAV (eds) Core collection of plant genetic resources. Wiley Sayce Publication, New York, pp 23–34

Wang L, Chen F (2012) Genotypic variation of potassium uptake and use efficiency in cotton (*Gossypiumhirsutum*). J Plant Nutr Soil Sci 175:303–308