Journal of Root Crops, 2011, Vol. 37 No. 2, pp. 150-154
Indian Society for Root Crops
ISSN 0378-2409

Evaluation of Taro (*Colocasia esculenta* (L.) Schott.) Germplasm Using Multivariate Analysis

S.K. Naskar and J. Sreekumar

Central Tuber Crops Research Institute, Sreekariyam, Thiruvananthapuram 695 017, Kerala, India
Corresponding author: J. Sreekumar, e-mail: sreejyothi_in@yahoo.com

Received: 12 July 2011; Accepted: 30 December 2011

Abstract

Principal component and cluster analysis were used to examine diversity among growth traits and yield components of 43 accessions of taro at the Regional Centre of Central Tuber Crops Research Institute, Bhubaneswar, Orissa, India. Principal component analysis identified three vectors with minimal Eigen value criterion, which explained 78% of the variation in the traits. Varimax rotation enabled loading of similar type of variables on a common principal component. The genotypes were further subjected to Hierarchical cluster analysis that resulted in eight clusters containing one to seventeen genotypes. The maximum number of genotypes was included in cluster III. The accessions TL-3 and Megh-1 (Meghalaya) were found to be far apart from other accessions and can be utilized for further breeding programmes.

Key words: Taro, germplasm, diversity, principal component analysis, clusters

Introduction

Taro (*Colocasia esculenta* (L.) Schott.) is a nutritious tuber crop consumed in many parts of India and South East Asia for its edible corms and leaves, while it is an important staple food in the Pacific region (Oceania). It is also called as eddoe and cocoyam in many countries. It is reported that taro, an ancient crop, was in cultivation in India before 5000 B.C. It is grown in many parts of our country and North Eastern region is considered to be one of the centers of origin. The Regional Centre of Central Tuber Crops Research Institute, Bhubaneswar, Orissa, India (RC, CTCRI) has collected a number of germplasm accessions from Chhattisgarh, Orissa, West Bengal and North Eastern region. To choose the right parents for their exploitation in crop improvement programmes, mere collection and maintenance of germplasm of a crop is not alone sufficient but the collected germplasm accessions should be systematically and carefully classified. The germplasm accessions collected from different parts, particularly from North Eastern region seems to be rich in diversity. For estimation of diversity within the germplasm, several workers have postulated a number of classificatory analyses like principal component analysis, factor analysis and clustering of genotypes (Peeters and Martinellis, 1989; Brown, 1991; Ordas et al., 1994 and Smith et al., 1995).

Hence the present investigation was undertaken with 43 accessions of taro to determine the degree of similarity among the accessions and to examine relative importance of the characters.

Materials and Methods

The experimental material consisted of 43 accessions of taro germplasm maintained at RC, CTCRI, Bhubaneswar, Orissa, India. These were collected from different parts of India. The experiment was conducted during 2004-2005 in Randomized Block Design with three replications. Recommended cultural practices were followed to raise good crops. Observations were recorded from five randomly selected plants of each accession on nine quantitative characters *viz.*, plant height (cm), leaf

lamina length and breadth (cm), petiole length (cm), number of cormels, cormel length (cm), cormel width (cm), size of corm (cm^2) and yield (kg per plant).

The mean of the data on different characteristics were used for further analysis. Principal component analysis was carried out using the PROC FACTOR in SAS (SAS Users Guide, 2002) and cluster analysis and plotting of principal component scores of accessions were carried out using GenStat package (GenStat seventh (DE3), 2007). Factor analysis is a generic term for a family of statistical techniques concerned with the reduction of a set of observable variables in terms of a small number of latent factors. It has been developed primarily for analyzing relationships among number of measurable entities. The different factor analytic techniques employ different criteria for extracting factors. The principal component method was used for factor extraction. The PRIORS = SMC option on the PROC FACTOR statement in SAS was used so that the squared multiple correlation could be inserted in the diagonal of the correlation matrix. The number of factors to be extracted was based on the proportion of variance explained. The initial factors obtained were not clearly interpretable. There are different types of rotations that can be done after the initial extraction of factors and the orthogonal rotation Varimax was used. Scatter plots were also drawn using the two main principal component scores of the accessions. Hierarchical cluster analysis with squared city block distance and average linkage algorithm was utilized for cluster analysis of accessions and dendrogram was drawn for showing the clustering pattern of genotypes. The city block distance is similar to Manhattan distance and is measured as $1-|x_i-x_j|/\text{range}$, where range is a measure of variation.

Results and Discussion

Taro breeding involves three main steps (i) creation of genetic variation (ii) evaluation of progenies and selection of superior individuals (iii) procedures associated with the release of a new variety. The strategy can be simple when existing varieties are highly variable and breeding is aimed at improving a few genetically simple traits through selection. The process will become complex when the breeder has to improve several characters simultaneously. The main task in evaluating the accessions for future breeding programmes is to include several characters simultaneously and the use of multivariate technique plays a crucial role in selecting parents for future generation.

The first four Principal Components (PC) showed eigen values more than one and cumulatively contributed 78.36% of variability. The first PC explained 36.43% of the total variability and the second, third and fourth principal components explained 17.18, 13.28 and 11.46% variation respectively. However, factor analysis was carried out with reduced correlation matrix to identify most important variables and for which three factors were extracted by the PROPRTION criterion. The rotated factor patterns after Varimax rotation are presented in Table 1. At present there is no information on the study of principal component or factor analysis of taro. Plant height, leaf lamina length and petiole length had significantly higher first principal factor loadings after rotation as 0.7609, 0.7582 and 0.7438 respectively. Number of cormels and yield had significantly higher second principal factor loadings of 0.7963 and 0.7871 respectively.

Table 1. Factor loadings of different characters with respect to different principal factors (Varimax rotation) in taro

Charactes	Principal Factor		
	PF1	PF2	PF3
Plant height	0.7609	0.2530	-0.1706
Leaf lamina length	0.7582	0.1219	-0.0650
Petiole length	0.7438	0.2097	0.1043
Leaf lamina breadth	0.4132	0.1711	0.4621
Cormel length	-0.0727	0.1798	0.5643
Cormel width	0.2910	0.2217	0.0271
Number of cormels	0.2539	0.7963	0.0969
Size of corm	-0.0407	-0.1404	0.5000
Yield (kg per plant)	0.2598	0.7871	-0.0096

The principal component scores obtained in the analysis were used to plot all the genotypes using PF1 and PF2 (Fig.1), which cumulatively explained about 54% of the variability. The genotypes, Panchmukhi, Megh-1 and TL-3 were clearly separated in the plot and the genotypes *viz.*, TL-3, Tellibari-3, Bastar-9 and BL-3, which were separated towards the positive side of the PF1 is supposed to

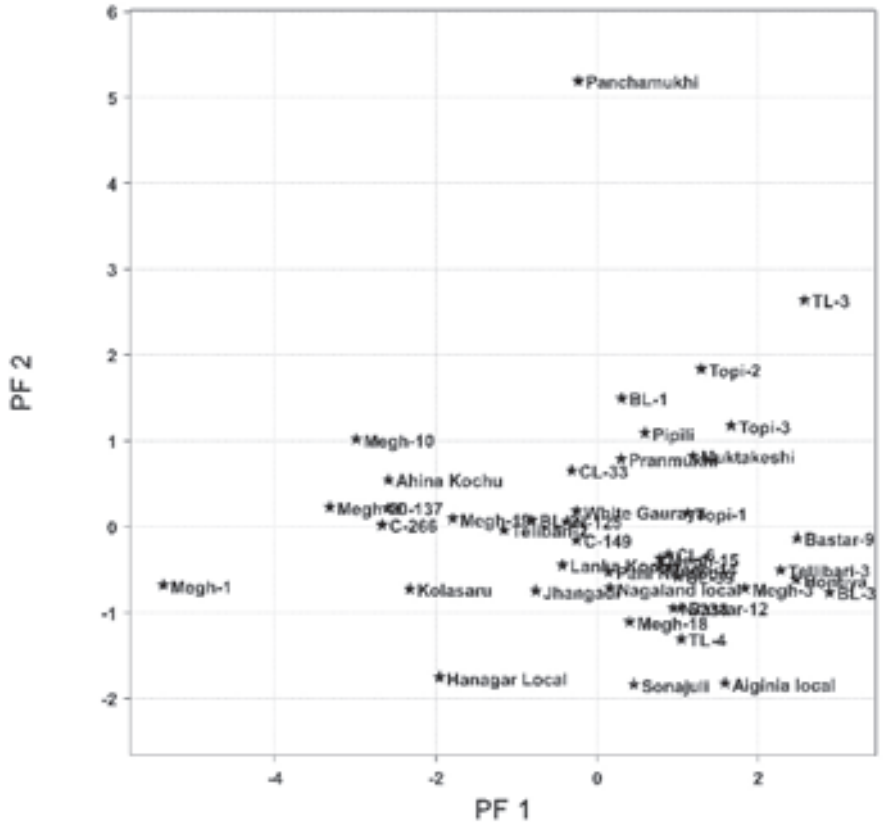

Fig.1. Distribution of taro germplasm based on principal components 1 and 2

be superior in growth characteristics. The genotypes, Panchamukhi, TL-3, Topi-2 and Topi-3, which were separated positive towards the second factor is supposed to have high yield characteristics. The results obtained in this study indicated significant differences among accessions in different agronomic characters of taro, whether it is growth or yield characteristics. The first principal factor (PF) indicated the contribution of plant height, leaf lamina length and petiole length to total variation, which were growth characters and can be designated as growth factor. The second PF was found to have significantly high loadings for number of cormels and the final yield. Also the third PF had high loadings for cormel length. Thus the second and third PF showed high loadings in yield characters and hence can be called as yield factors. Perusal of the analysis revealed that plant height, leaf lamina length, petiole length, number of cormels and yield were the main variability contributing characters accounting 54% of the total variation (78%). The plot of PF1 vs PF2 indicated the superior genotypes in the first and second principal factor loadings and the genotypes which were superior for high yield characters along with clearly separated genotypes that can be used as parents in hybridization programmes for the improvement of yield. The superiority of genotypes, Panchamukhi, TL-3, Topi-2, Topi-3, Megh-1 and Bastar-9 is established in Fig.1.

The hierarchical method of cluster analysis with city block distance measure and average linkage algorithm divided the genotypes into eight clusters. The cluster membership of different accessions is presented in Table 2. Maximum number of

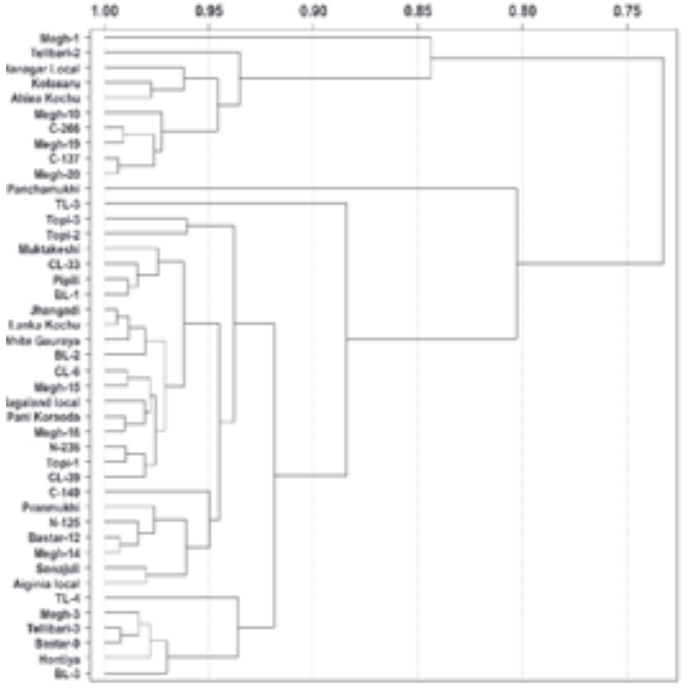

Fig. 2. Dendrogram showing the clustering pattern of different taro genotypes

Table 2. Cluster membership and number of genotypes in each cluster of taro germplasm

Cluster	Genotype	Number of genotypes
I	Tellibari-3 (NE), Megh-3 (Meghalaya), Hontiya (NE), Aiginia local (Orissa), Baster-9 (Chhattisgarh), BL-3 (Orissa), Sonajuli (NE)	7
II	TL-4 (NE)	1
III	Megh-15 (Meghalaya), Lanka Kochu, Topi-1, Pipili, BL-1 (Bhubaneswar), N-238, Nagaland local (NE) , CL-39, Pani Koraoda, Jhangdi (Orissa), BL-2, CL-33, CL-6 (Orissa), Megh-18 (NE), Muktakeshi, Topi-3, Topi-2 (Orissa)	17
IV	N-125 (NE), White Gauraya (Assam), Megh-14 (Meghalaya), Baster-12 (Chhattisgarh), Pranmukhi, C-149, Tellibari-2(NE)	7
V	TL-3 (NE)	1
VI	Panchamukhi (Assam)	1
VII	C-266 , C-137 (Orissa), Megh –19 (Meghalaya), Ahina Kochu (Assam), Megh -20, Megh-10, Hanagar Local (NE) , Kolasaru (Orissa)	8
VIII	Megh-1(Meghalaya)	1
	Total	43

genotypes of 17 is in cluster III and four other clusters have only one member each. The relative association among the different genotypes is presented in the form of a dendrogram (Fig. 2). The resemblance coefficient between the two genotypes is at which their branches join. The dendrogram elaborates the relative magnitude of resemblance among the genotypes as well as the cluster. The perusal of clustering pattern indicated that there was no parallelism between the geographical and genetic diversity in the present investigation. For example the collections from Meghalaya were distributed in all the clusters. The cluster VII had eight genotypes, which was followed by cluster I and cluster IV with seven genotypes.

The mean performance of genotypes in different clusters and overall mean with regard to different traits revealed wide range of variation (Table 3). The result indicated that the cluster VIII with Megh-1 exhibited highest means for plant height (47.71 cm), number of cormels (18.65) and yield (2.38 kg per plant), cluster VII were superior for leaf lamina length (25.12 cm), petiole length (31.14 cm) and cormel width (6.72 cm), cluster VI for leaf lamina breadth (37.30 cm) and cormel length (8.08 cm), which includes Panchamukhi and cluster V, which contains TL-3 exhibited superior size of corm (236.48 cm^2). The other clusters like I, II, III and IV contained genotypes, which had poor to average performance for most of the characters. Cluster I contained one genotype, which showed high yield due to its robust growth and number of cormels. Cluster VII that clustered 8 genotypes had superior values for plant height, leaf lamina length, petiole length, cormel length and medium values for leaf lamina breadth, cormel length, low values

Table 3. Cluster mean and general mean for different characters in taro

Cluster/ Trait	Plant height (cm)	Leaf lamina length (cm)	Petiole length (cm)	Leaf lamina breath (cm)	Cormel length (cm)	Cormel width (cm)	Number of cormels	Size of corm (cm^2)	Total yield (kg per plant)
I	27.47	17.86	18.01	12.69	4.61	6.09	4.48	76.18	0.47
II	27.3	20.99	22.20	15.70	3.89	4.89	5.10	73.73	1.04
III	29.93	19.13	21.74	13.33	5.78	6.58	6.55	81.09	0.76
IV	33.19	22.96	24.71	14.98	5.55	5.86	5.84	93.42	0.74
V	22.48	17.91	21.35	12.26	5.35	5.36	5.75	236.48	0.29
VI	21.63	15.68	17.38	37.30	8.08	6.23	8.77	129.51	1.02
VII	41.89	25.12	31.14	20.84	5.39	6.72	8.02	95.70	1.00
VIII	47.71	24.68	29.98	16.54	4.81	6.52	18.65	70.46	2.38
Mean	31.45	20.54	23.32	17.96	5.43	6.03	7.89	107.07	0.96

for number of cormels and corm size, which got reflected in total yield. Other clusters of interest were cluster II and cluster VI, both had one genotype each. Cluster I exhibited medium values for all the characters, whereas cluster VI showed superior performance for leaf lamina breadth and cormel length with short plants.

From the clustering pattern it was observed that the genotypes from different geographical regions were not grouped together in a cluster and vice versa suggesting that geographical diversity does not necessarily represent genetic diversity. This may be due to free exchange of genetic material among different regions (Singh et al., 2007). It is suggested that hybridization among the diverse parents is likely to produce progressive segregates in further generation. The intra and inter cluster similarity matrix are presented in Table 4. The maximum intra cluster similarity of 86.0 was observed in cluster IV followed by cluster-I (85.6), cluster III (85.5) and cluster VII (84.4). The maximum inter cluster similarity was found between cluster VII and cluster V followed by cluster VIII and cluster II, cluster VI and cluster V, cluster VII and cluster II and cluster IX and cluster VIII. The clusters which showed low similarity, that means, clusters that were far apart, were cluster V and cluster VIII, cluster VI and cluster VIII and cluster I and cluster VIII with values 44.5, 43.5 and 57.5 respectively. The low similarity between TL-3 (cluster V) and Megh-1 (cluster VIII) is discernible in Fig.1 also. The accessions belonging to this far apart groups can be used as parents for further breeding programmes. The results of hierarchical cluster analysis and principal factor analysis confirmed the findings of each other. The plots PF1 and PF2 both accounting for 54% cumulative variation (Fig.1) revealed clear differentiation of genotypes according to their cluster membership.

Table 4. Between and within cluster similarity matrix of different taro genotypes

Cluster	I	II	III	IV	V	VI	VII	VIII
I	85.6							
II	79.5	----						
III	81.3	78.7	85.5					
IV	79.5	81.1	82.8	86				
V	77.3	74.3	74.2	74.7	0			
VI	65.9	60.5	69.0	64.9	63.0	0		
VII	68.2	69.2	75.5	79.1	59.4	60.0	84.4	
VIII	57.4	60.1	62.8	65.2	44.5	43.5	73.6	0

Conclusion

Forty three accessions of taro were evaluated for their growth and yield characters using multivariate statistical techniques such as principal factor analysis and cluster analysis. Principal factor analysis helped in identifying characters which contributed to total variation explained by each factor. The growth characters such as plant height, leaf lamina length and petiole length contributed much of the variation in factor one, whereas the yield characters influenced more in the second principal factor. Clustering the accessions helped in identifying parents with diverse characters for future breeding programmmes.

References

Brown, J. S. 1991. Principal component and cluster analysis of cotton cultivation variability across U.S.Cotton Belt. *Crop Sci.*, **31**: 915-922.

GenStat Seventh Edition (DE3). 2007. Service Pack 1, Version 7.2.0.220, Lawes Agricultural Trust, supplied by VSN International Ltd, UK.

Ordas, A., Malvar, R. A. and DeRon, A.M. 1994. Relationship among American and Spanish population of maize. *Euphytica*, **79**: 149-161.

Peeters, J. P. and Martinellis, J.A. 1989. Hierarchical cluster analysis as a tool to manage variation in germplasm collections. *Theoretical appl. Genet.*, **78**: 42-48.

SAS Users Guide. 2002. SAS Institute Inc., Cary, NC, USA.

Singh, R.A., Yadav, R. and Pahuja, S.K. 2007. Evaluation of Egyptian clover germplasm using multivariate analysis. *Indian J. Genet.*, **67**(2): 135-139.

Smith, S.E., Guarino, L., Al-Dos, A. and Conta, D. M. 1995. Morphological and agronomical affinities among middle eastern alfalfa accessions from Oman and Yemen. *Crop Sci.*, **35**: 1188-1194.