BMC Plant Biology

RESEARCH

Open Access

Extent and patterns of morphological and molecular genetic diversity and population structure of Nigerian Taro cultivars

Joy Jesumeda Oladimeji[1,2*], Ayodeji Abe[3], Pullikanti Lava Kumar[2], Paterne A. Agre[2], Oluyinka Johnson Ilesanmi[2], Ramesh Raju Vetukuri[4*] and Ranjana Bhattacharjee[2*]

Abstract

Background Genetic diversity is crucial for conservation efforts as well as breeding programs targeted at the development of improved varieties. Taro, a climate-resilient crop, plays a vital role in the nutritional and economic livelihoods of many households in Nigeria, but its yield is very low due to inadequate genetic improvement efforts. A diversity assessment of Nigerian taro is therefore required to create a premise for its improvement in yield, quality and disease tolerance. In this study, the genetic diversity and population structure of 490 taro cultivars comprising two main gene pools: Dasheen (215) and Eddoe (275), collected from farmers and marketers across seven states in Nigeria was assessed using 3047 Diversity Array Technology single nucleotide polymorphism (DArT-SNP) markers. A subset of 114 taro cultivars, comprising 30 Dasheens and 84 Eddoes were further phenotyped using 24 agro-morphological descriptors.

Results Both phenotypic and molecular characterization revealed higher genetic diversity among the Eddoes than Dasheens. Estimates of gene flow (Nm = 0.353) revealed intermixing of cultivars among the States of collection, with the highest gene flow occurring between cultivars from Anambra and Ondo states and the lowest between Anambra and Kwara states. Population structure and Ward's minimum variance hierarchical cluster based on DArT-SNPs identified four groups, one comprising Dasheen and three comprising Eddoe cultivars. Hierarchical clustering based on phenotypic traits delineated three clusters. Variation between gene pools (49%) was higher than within gene pools (32%). Variation among States of collection was high (41%), while variation among individuals within gene pools (18%) and States of collection (19%) was relatively low. Correlation between phenotypic and genotypic diversity assessments was low ($r = 0.01$), indicating that both approaches were necessary for assessing genetic diversity in taro. However, genotypic assessment provided better information about genetic diversity of the taro cultivars.

*Correspondence:
Joy Jesumeda Oladimeji
joyjesumeda@gmail.com
Ramesh Raju Vetukuri
Ramesh.vetukuri@slu.se
Ranjana Bhattacharjee
r.bhattacharjee@cgiar.org

Full list of author information is available at the end of the article

Conclusion This is the first study that represented germplasm collection across the major taro growing regions of Nigeria. The findings from this study based on agro-morphological characterization and DArT-SNP genotyping are critical for genetic characterization, conservation and breeding of taro in Nigeria, mainly initiating hybridization between the two genepools after careful assessment of ploidy levels of the accessions collected in this study. This will facilitate in developing improved taro varieties with desirable traits, such as higher yield, better disease resistance, and improved nutritional quality.

Keywords Taro gene pools, Dasheens, Eddoes, Agro-morphological descriptors, DArT-SNP markers, Genetic diversity, Population structure

Introduction

Taro [*Colocasia esculenta* (L.) Schott] is a tuber crop from the family Araceae, known for its edible corms and cormels. It is one of the ancient food crops originating from the Indo-Malayan region over 50,000 years ago [1], where many related species and subspecies are found [2–4]. It is believed that the voyagers of East coast of Africa brought taro to West Africa some 2000 years back [3]. Two main gene pools of taro are identified: *Colocasia esculenta* var. *esculenta*, called as Dasheen, and *Colocasia esculenta* var. *antiquorum* called as Eddoe [2, 5], which are differentiated based on origin, ploidy, and corm shape [2, 3, 6]. The Dasheens are of Pacific origin, diploids, and have a large central corm with few or no cormels, while the Eddoes are from Asia, triploids, and have globular corms with several cormels [2, 6]. The diploids have a chromosome number of 2n=2x=28 while triploids have 3n=3x=42 [7, 8]. The entire plant of taro is useful as food, feed, medicine, and fuel [9, 10], and is highly digestible [11–13]. It is an essential component of the socio-cultural life in the Pacific [2, 14], and ethnic groups in south-east and south-south regions of Nigeria [15–17].

Nigeria, with a history of 2000 years of taro cultivation [18], is the largest producer in the world [19], where it is grown both as a sole crop and an intercrop [20]. In 2021, Nigeria produced 26% of the total global taro comprising of 12.4 million tons from 1.8 million hectares [19]. Despite the high contribution of Nigeria to global taro production, taro yield (3.9 t/ha) in Nigeria is low compared to other taro-producing regions of the world, such as Oceania (8.5 t/ha) and Asia (16.1 t/ha). This low yield, which can primarily be attributed to losses due to diseases and lack of improved planting materials [21–23], calls for genetic improvement of taro in Nigeria for resistance to diseases, increased yield, and quality.

Assessment of genetic diversity for crop improvement is a pre-requisite, and different strategies have been used to evaluate the extent of genetic variability among crop cultivars. Phenotypic descriptors are extensively used, both for their low cost, and usefulness in the identification and selection of desirable traits [24–26]. In taro, agro-morphological descriptors such as plant height, plant span, leaf area index, number of suckers per plant, time to maturity, and yield per plant have been used to discriminate among cultivars in Togo [27], Ghana [28], Burkina Faso [29], and Nigeria [21]. However, agro-morphological traits are less efficient due to environment and genotype x environment interactions, making it difficult to distinguish between closely related individuals [30]. It is therefore important to complement with molecular markers, which are not subject to such limitations [31]. Kreike et al. [2] used amplified fragment length polymorphism (AFLP) markers to evaluate the genetic diversity of 255 taro accessions from seven countries in Asia and the Pacific and identified two gene pools, one in Asia and the other in the Pacific. Similarly, Chaïr et al. [3] deployed simple sequence repeat (SSR) markers on 357 taro accessions from 19 countries in Asia, America, Africa, and the Pacific, and found the highest genetic diversity within the Asian accessions. In addition, the authors reported that taro in Africa has a narrow genetic base and is clonally propagated which further limits the genetic diversity of the crop.

Recently, high throughput sequencing technologies, characterized by high reproducibility such as Diversity Array Technology sequencing (DArTseq) [32], provides a large pool of both SNP and SilicoDArT markers [33]. DArTseq reduces the complexity of the genome via use of restriction enzymes and sequencing short reads while replacing the hybridization step, and sequencing is conducted using the Illumina system [34]. DArTseq technology have been successfully used for the assessment of genetic diversity and population structure of different crops including wheat [35, 36], maize [37, 38], cowpea [39], cassava [40], yam [41, 42], and even taro [23]. Fufa et al. [23] assessed the genetic diversity among 282 taro cultivars which included 188 cultivars from Nigeria and 94 from Vanuatu using DArTseq SNP markers, and reported low genetic diversity among the cultivars from Nigeria.

Previous attempts to assess the diversity of taro in Nigeria [21, 23] were based on collections from the south-east region of the country and those from the National Root Crops Research Institute (NRCRI), Umudike. It is well known that most of these collections are lost due to taro leaf blight disease [22]. Additionally, in Nigeria, taro is predominantly grown across the south-east, south-west, and south-south geopolitical zones of the country, which span from the derived savanna to the humid forest

agro-ecological regions [20, 22, 43]. Therefore, to ensure adequate coverage of taro-growing regions, it is expedient to extend the collections of taro cultivars to other regions of the country, particularly within the derived savanna and humid forest, which are the main taro production agroecologies in Nigeria. In this study, therefore, an attempt has been made with the objectives to (i) collect taro germplasm from diverse areas within the major taro production zones of Nigeria; and (ii) assess genetic diversity and population structure based on agro-morphological descriptors and DArT-SNP genotyping.

Materials and methods

Plant material

Leaf and corm/cormels of 459 taro cultivars which comprised 215 Dasheen and 275 Eddoe cultivars were collected from farmers across seven states: Kwara (110 Dasheens), Oyo [177 (105 Dasheen and 72 Eddoes)], Ekiti (32 Eddoes), Ondo (48 Eddoes), Akwa Ibom (49 Eddoes), Ebonyi (30 Eddoes), and Anambra (13 Eddoes). Additionally, 31 corms (all Eddoes) from five markets across four states (Akwa Ibom: 1 market, 14 corms; Ekiti: 1 market, 4 corms; Ebonyi: 2 markets, 11 corms; Anambra (1 market, 2 corms) were collected and grown at the International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria, and leaf samples were collected. Table 1 shows the collections from each State, the gene pools, and the proportion that was genotyped and phenotyped. Figure 1 shows the location where the collections were made. The source, gene pool, and local name of each cultivar are provided in Supplementary Table S1. The 490 cultivars represent a sum of the collections made in two expeditions: between August and September 2021 representing the growing season (159 cultivars), and between December 2021 and January 2022 representing the harvesting season (331 cultivars). The leaves from all 490 cultivars were collected in properly labeled envelopes and brought to IITA, Ibadan, Nigeria, where these were freeze-dried and stored at the Bioscience Center.

Table 1 Taro cultivars collected from different states in Nigeria

State	Total cultivars collected	Total cultivars genotyped	Total cultivars phenotyped
Kwara	110 Dasheens	110 Dasheens	14 Dasheens
Oyo	177 (105 Dasheens and 72 Eddoes)	177 (105 Dasheens and 72 Eddoes)	25 (16 Dasheens and 9 Eddoes)
Ekiti	36 Eddoes	36 Eddoes	21 Eddoes
Ondo	48 Eddoes	48 Eddoes	11 Eddoes
Akwa Ibom	63 Eddoes	63 Eddoes	25 Eddoes
Ebonyi	41 Eddoes	41 Eddoes	14 Eddoes
Anambra	15 Eddoes	15 Eddoes	4 Eddoes
Total	490 (215 Dasheens; 275 Eddoes)	490 (215 Dasheens; 275 Eddoes)	114 (30 Dasheens; 84 Eddoes)

Agro-morphological evaluation

A subset of 114 taro cultivars was used for agro-morphological characterization, and this selection was based on the availability of corms and cormels, germination rate, and survival of cultivars after one month of planting. The 114 cultivars comprised 14 Dasheens from Kwara, 16 Dasheens and nine Eddoes from Oyo, 21 Eddoes from Ekiti, 11 Eddoes from Ondo, 25 Eddoes from Akwa Ibom, 14 Eddoes from Ebonyi, and four Eddoes from Anambra States (Table 1). These genotypes were planted in two stages, first in a screenhouse at IITA following their collection times, in September 2021 (65 corms), and in March 2022 (49 corms). The seedlings/plants of each cultivar were then transplanted on the field at IITA, Ibadan (N 7° 29' 56.562" and E 3° 54' 27.5868"), Nigeria. The experiment was laid out on the field using an augmented randomized complete block design wherein each cultivar was randomly placed within three blocks. Weeding was done manually as necessary to keep the experimental field clean. Supplemental irrigation was provided to ensure the plants were not moisture-stressed, while mulching was carried out to reduce water loss from the soil.

Data was recorded on 24 agro-morphological traits, which comprised 14 qualitative and 10 quantitative traits following the IPGRI descriptor [44] and Andarini et al. [45] for taro with minor modifications (Supplementary Table S2).

DNA extraction and genotyping

Four leaf discs of approximately 5 mm diameter from freeze-dried leaf samples of each 490 genotypes were punched into labeled 8-strip 1.1 mL propylene tubes with strip caps, up to 12 strips placed on 96-well boxes. A total of six labeled 96-well boxes were then shipped to Intertek, Sweden for DNA extraction and DArTseq genotyping. Total genomic DNA was extracted at Intertek, Sweden using an in-house DNA extraction protocol [46]. The quality and quantity of the DNA were checked on a 0.8% agarose gel and Nanodrop 2000c spectrophotometer (Thermo Scientific, Waltham, MA, USA), respectively. Genotyping was done at Intertek, Sweden using DArT-SNP markers following an in-house protocol [47]. Two blank controls were included in each box during genotyping.

The sequencing was carried out using Hiseq2500 and scoring of DArTseq markers was done using DArTsoft14, an in-house marker scoring software at Intertek. DArTseq markers were scored in binary form as presence/absence (1 and 0, respectively), and aligned to the reference genome of Taro (Taro_V1) [48] to identify chromosome positions.

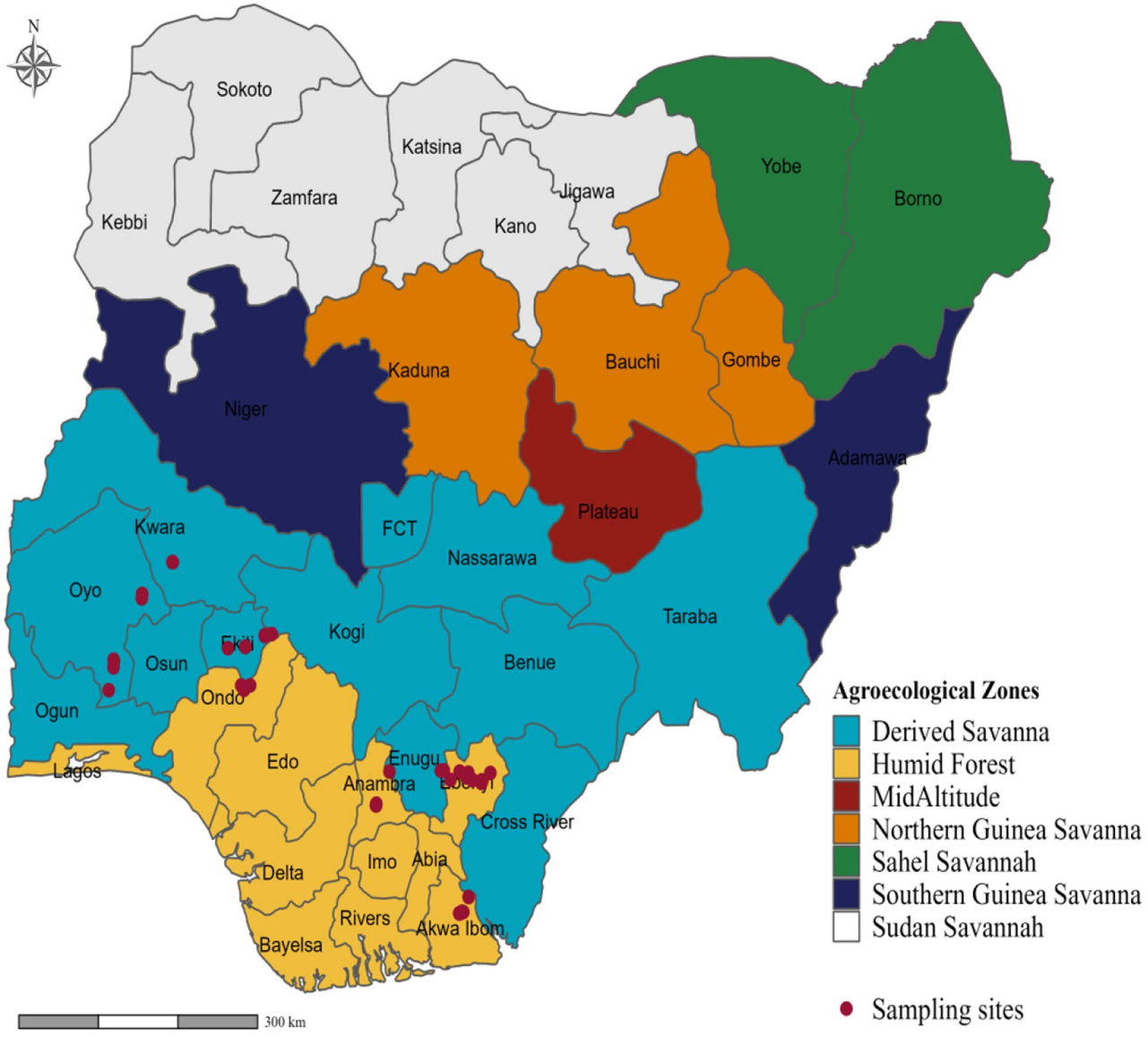

Fig. 1 Map of Nigeria with red dots depicting the sampling sites

Data Analysis

(a) Agro-morphological data

The data obtained from 114 cultivars were pooled together for analysis. Data on 14 qualitative traits were subjected to descriptive statistics in the Excel sheet. Analysis of variance and computation of genetic parameters for the 10 quantitative traits were carried out using the augmented RCBD package in R v 4.1 [49]. The adjusted means of the 10 quantitative traits were used to compute correlation coefficients among the traits and displayed as a choropleth map. The Gower algorithm in Philentropy package in R v 4.1 [50] was used to construct the genetic distance matrix and Ward's minimum variance was used for hierarchical clustering from the adjusted means of the 10 traits.

(b) DArT-SNP marker data

The DArT-SNP genotyping of 490 taro cultivars resulted in a total of 4748 SNPs. The raw data was filtered using VCFtools [51] in R v 4.1 software to eliminate markers with <60% marker information and <0.05 minor allele frequency (MAF). After filtering, 3047 SNP markers were retained for further analysis. PowerMarker v 3.25 [52] was used to compute the diversity parameters including expected heterozygosity (H_e), observed heterozygosity (H_o), minor allele frequency (MAF), and polymorphic information content (PIC). A binary file was generated from the filtered variant call format (VCF) file and was then subjected to cross-validation approaches for population structure analysis. The population structure analysis was conducted based on Admixture [53] using

the Bayesian Information Criterion (BIC). The optimum number of clusters was inferred using k-means analysis after varying the possible number from 2 to 20. A cut-off value of 50% ancestry suggested through the Admixture analysis was used to estimate membership probabilities of the cultivars for the groups identified, while those lower than 50% were considered as admixed. Subsequently, the discriminant analysis of principal component (DAPC) was carried out on the identified clusters using the first 40 principal components in the adegenet package in R v 4.1 [54]. A genetic distance matrix based on Ward's minimum variance was generated for hierarchical clustering of the 490 cultivars using the phylogenetics and evolution (ape) package analysis in R v 4.1 [55, 56].

The genetic differentiation among and within groups (gene pools and States of collection) was estimated using the analysis of molecular variance (AMOVA), and the significance was tested with a non-parametric approach with 99 permutations using GenAlex v. 6.503 [57]. The gene flow (Nm) and fixation index (F_{ST}) among gene pools and States of collection, were computed using GenAlEx 6.503.

(c) Combined analysis of phenotypic and genotypic data

All combined analyses was carried out only on 114 taro cultivars for which both phenotypic and genotypic data was available. The genetic distance matrices of both phenotypic and genotypic data was combined to generate a composite distance matrix. A hierarchical cluster dendrogram (HC) was further generated based on the composite distance matrix to examine the relationships between the composite genetic distance and the individual phenotypic and genotypic genetic distances. The Mantel test was used to estimate the correlation among the phenotypic, genotypic, and composite genetic distance matrices, using the Monte-Carlo method with 9999 permutations for significance estimation. The similarity between the phenotypic and molecular HC was examined using the tanglegram function implemented in the dendextend package in R v 4.1 [58].

Results

Phenotypic diversity and differentiation

(a) Qualitative traits

The frequency distribution of the 14 qualitative traits is presented in Table 2. The Eddoes showed variation for all the qualitative traits except corm fiber color. However, the Dasheens were less variable as they showed no variation across six out of 14 traits, such as Leaf main vein color (all purple), Petiole junction color (all purple), Petiole color of top third (all purple), Petiole color of basal

Table 2 Phenotypic variation among 14 qualitative traits assessed on 114 taro cultivars

Trait	Descriptor	Proportion of genotypes	Proportion of each gene pool	Trait	Descriptor	Frequency of genotypes	Frequency of each gene pool
LC	Green	31%	Dasheen: 26%; Eddoes: 74%	PCU	Green	41%	Eddoes: 100%
	Dark green	69%	Dasheen: 26%; Eddoes: 74%		Purple	59%	Dasheens: 45%; Eddoes: 55%
LO	Drooping	40%	Dasheen: 39%; Eddoes: 61%	PCL	Green	75%	Dasheens: 35%; Eddoes: 65%
	Horizontal	5%	Dasheen: 17%; Eddoes: 83%		Purple	25%	Eddoes: 100%
	Cup-shaped	15%	Dasheen: 63%; Eddoes: 37%	CFC	Light yellow	74%	Eddoes: 100%
	Erect-apex down	40%	Dasheen: 2%; Eddoes: 98%		Purple	26%	Dasheens: 100%
LMC	Green	24%	Eddoes: 100%	CSS	Smooth	1%	Eddoe: 100%
	Purple	76%	Dasheens: 34%; Eddoes: 66%		Scales present	31%	Dasheens: 48%; Eddoes: 52%
LM	Entire	33%	Dasheens: 42%; Eddoes: 58%		Fibrous and scales present	68%	Dasheens: 17%; Eddoes: 83%
	Undulate	67%	Dasheens: 18%; Eddoes: 82%	CS	Conical	3%	Dasheens: 67%; Eddoe: 33%
LVP	Y pattern	40%	Dasheens: 24%; Eddoes: 76%		Round	34%	Dasheens: 8%; Eddoes: 92%
	Y pattern extending to secondary veins	60%	Dasheens: 28%; Eddoes: 72%		Elliptical	1%	Dasheen: 100%
LMVC	Green	60%	Eddoes: 100%		Dumbbell	38%	Eddoes: 100%
	Whitish	6%	Eddoes: 100%		Cylindrical	2%	Dasheen: 50%; Eddoe: 50%
	Purple	34%	Dasheens: 77%; Eddoes: 23%		Elongated	22%	Dasheens: 92%; Eddoes: 8%
PJP	Absent	19%	Eddoes: 100%	CBC	White	34%	Eddoes: 100%
	Small	42%	Dasheens: 14%; Eddoes: 86%		Purple	34%	Dasheen: 77%; Eddoe: 23%
	Medium	39%	Dasheens: 52%; Eddoes: 48%		Yellow/ green	25%	Eddoes: 100%
PJC	Green	51%	Eddoes: 100%		Pink/ red	7%	Eddoes: 100%
	Purple	49%	Dasheens: 54%; Eddoes: 46%				

LC: Leaf lamina color, LO: Leaf lamina orientation, LMC: Leaf lamina margin color, LM: Leaf lamina margin, LVP: Leaf vein pattern, LMVC: Leaf main vein color, PJP: Petiole junction pattern, PJC: Petiole junction color, PCU: Petiole color of top third, PCL: Petiole color of basal third, CFC: Corm fiber color, CSS: Corm skin surface, CS: Corm shape, CBC: Corm bud color

Table 3 Variance component, heritability, and genetic advance for ten quantitative traits on 114 taro cultivars

Traits	Mean	PV	GV	Heritability (%)
EM	34.20	732.96	644.67	87.95
PL	43.87	288.43	196.20	68.02
NST	4.88	16.27	11.77	72.34
LLWR	1.13	0.01	0.01	98.38
LA	544.41	144924.69	118400.46	81.70
PH	54.21	430.93	307.51	71.36
NLPP	18.25	146.15	119.61	81.84
PS	63.66	667.75	285.33	42.73
NCMPH	21.39	318.90	303.74	95.24
TY	0.86	0.45	0.35	78.27

PV: Phenotypic variation, GV: Genetic variation, EM: Days to emergence, PL: Petiole length, NST: Number of stolons per plant, LLWR: Leaf length to width ratio, LA: Leaf area, PH: Plant height, NLPP: Number of leaves per plant, PS: Plant span, NCMPH: Number of cormels per hill, TY: Total yield

third (all green), Corm fiber color (all purple) and Corm bud color (all purple).

(b) Quantitative traits

Significant differences among the cultivars was observed for days to emergence (EM), leaf length to width ratio (LLWR), and number of cormels per hill (NCMPH)) (Supplementary Table S3). There was a high and positive correlation (79%) between total yield (TY) and NCMPH. However, there was a moderate (>45%) correlation between TY and Petiole length (PL), Plant height (PH), and Leaf area (LA) (Supplementary Figure S1). The genetic parameters of the measured traits showed moderate [42.73% for Plant span (PS)] to high heritability (95.24% for NCMPH). A low (0.96 for LLWR) to moderate [43.46 for Number of stolons per plant (NST)] environmental coefficient of variation (Table 3) was also observed.

The hierarchical clustering based on 10 quantitative traits grouped the cultivars into four clusters (Fig. 2). Cluster I had the least number of cultivars (7) with highest agronomic performance. The cluster included one Dasheen (from Kwara) and six Eddoes (one from Ekiti, two from Akwa Ibom, and three from Ebonyi). This cluster recorded highest average Total yield (TY) (1.93 Kg per plant), Number of cormels per hill (NCMPH) (54 cormels), Leaf area (LA) (1346.39 m^2) and number of Days to emergence (EM) (73 days). In contrast, Cluster II represented cultivars with least agronomic performance, and showed lowest average EM (31 days). This cluster had 28 cultivars, which included 8 Dasheens and 20 Eddoes from all seven States: Akwa Ibom (11 Eddoes), Anambra (1 Eddoe), Ebonyi (3 Eddoes), Ekiti (2 Eddoes), Kwara (4 Dasheens), Ondo (1 Eddoe), and Oyo [6 (4 Dasheens and 2 Eddoes)]. The average TY was 0.43 Kg/plant, NCMPH was 11 cormels, and the LA was 140.36 m^2.

Cluster III had 14 cultivars comprising five Dasheens and nive Eddoes from six States, namely Kwara (4 Dasheens), Oyo (1 Dasheen; 1 Eddoe), Ondo (3 Eddoes), Ekiti (2 Eddoes), Akwa Ibom (2 Eddoes) and Anambra (1 Eddoes). The average TY was 1.24 Kg/plant, NCMPH was 26, LA was 908.44 cm2 while the EM was 38 days. Cluster IV had the highest number of cultivars (65), comprising 16 Dasheens and 49 Eddoes from the seven States, namely Akwa Ibom (11 Eddoes), Anambra (2

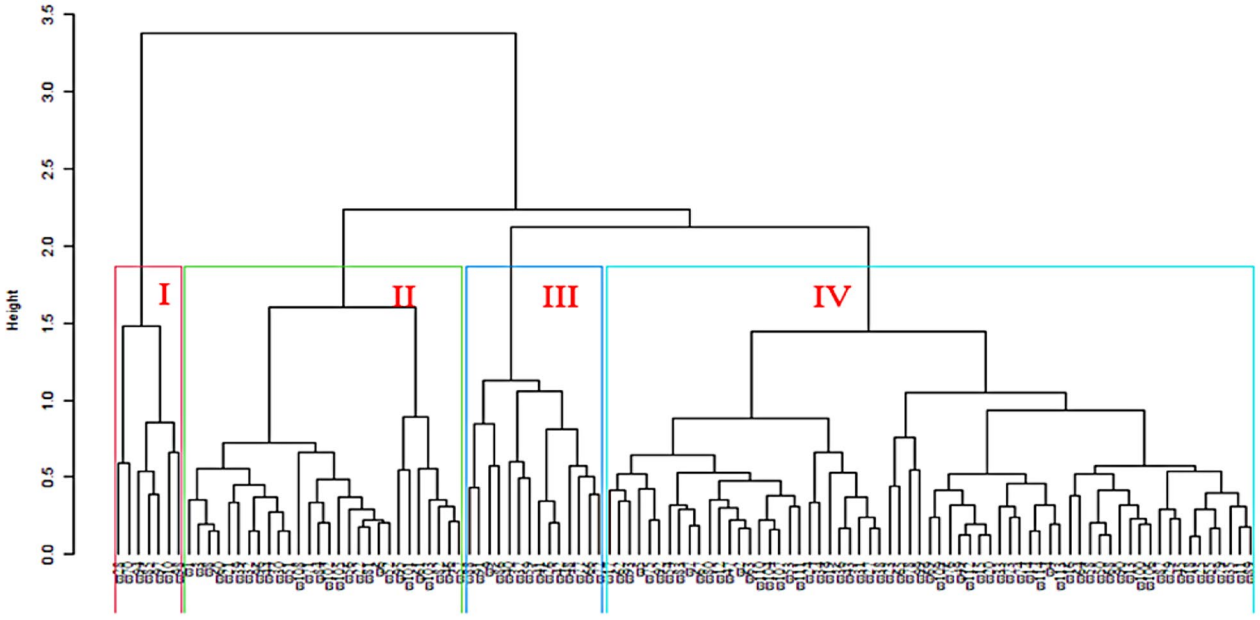

Fig. 2 Ward's minimum variance hierarchical cluster dendrogram using ten quantitative traits on 114 taro cultivars

Table 4 Average diversity metrics for 490 taro cultivars using 3047 DArT-SNPs

Taro cultivars	Number of cultivars	Minor allele frequency	Expected heterozygosity	Observed heterozygosity	Polymorphic information content
Eddoes	275	0.304	0.398	0.244	0.314
Dasheen	215	0.017	0.025	0.03	0.021
Entire cultivars	490	0.249	0.347	0.15	0.281

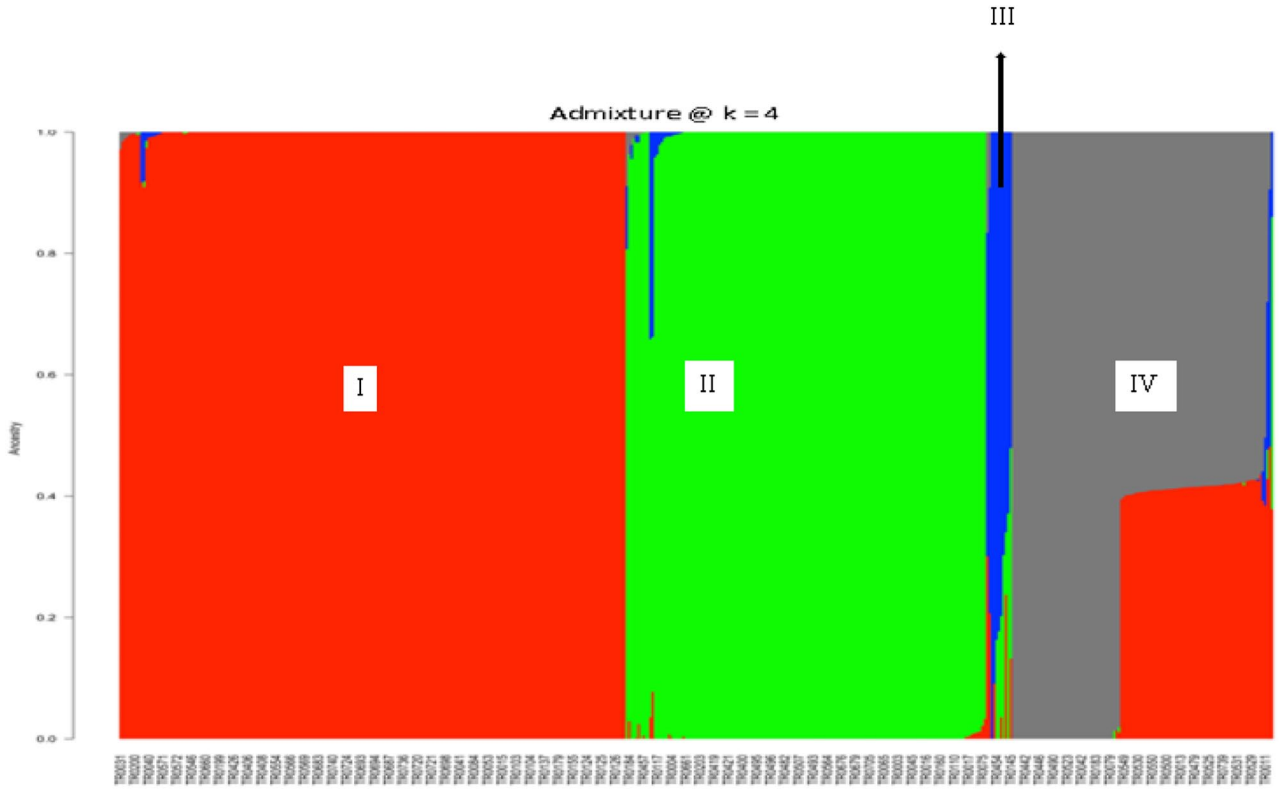

Fig. 3 Graphical representation of population structure of 490 taro cultivars' at K = 4, based on the membership coefficient of ≥ 50%

Eddoes), Ebonyi (8 Eddoes), Ekiti (16 Eddoes), Kwara (5 Dasheens), Ondo (6 Eddoes), Oyo [17 (11 Dasheens and 6 Eddoes)]. The cultivars of this cluster showed moderate agronomic performance with average TY (0.92 Kg/plant), NCMPH (22 cormels), LA (616.56 m^2), and EM was 32 days.

Genetic diversity indices and population structure

For filtered 3047 DArT-SNPs, the minor allele frequency (MAF) ranged from 0.050 to 0.490 with an average of 0.249. The expected heterozygosity (H_e) ranged from 0.002 to 0.533 with an average of 0.347, whereas the observed heterozygosity (H_o) varied from 0.160 to 0.500 with an average of 0.150. The polymorphic information content (PIC) ranged from 0.140 to 0.380 with an average of 0.281. Based on gene pools, the Eddoes recorded higher diversity metrics than Dasheens (Table 4). The diversity indices of the Eddoes exceeded the average of the entire collection, indicating that the Eddoes were the main contributors to the diversity observed among 490 taro cultivars (Table 4).

Inference on the population structure of 490 cultivars was made at K = 4 (Fig. 3). All 215 Dasheen cultivars grouped in Cluster I, while 272 Eddoes were distributed among Cluster II (153), III (11) and IV (108). Three Eddoe cultivars, TR0455, TR0537 and TR0113 from Ondo, Akwa Ibom and Ekiti, were not classified into any of the four clusters because their ancestry percentage was less than 50%, and were considered as admixed.

Similarly, based on BIC and using DAPC, the 490 taro cultivars were grouped into 4 clusters (Fig. 4). Cluster 1 had 46 Eddoe cultivars [40 from Oyo and 6 from Akwa Ibom]. Cluster 2 had the highest number of cultivars (221), with 110 Dasheens from Kwara, 105 Dasheens from Oyo, two Eddoes from Ekiti, one Eddoe from Ondo and three Eddoes from Akwa Ibom. Clusters 3 and 4 comprised only Eddoes. Cluster 3 had cultivars from Oyo (26), Ekiti (31), Ondo (45), Akwa Ibom (7), Ebonyi (36), and Anambra (15), while Cluster 4 represented cultivars

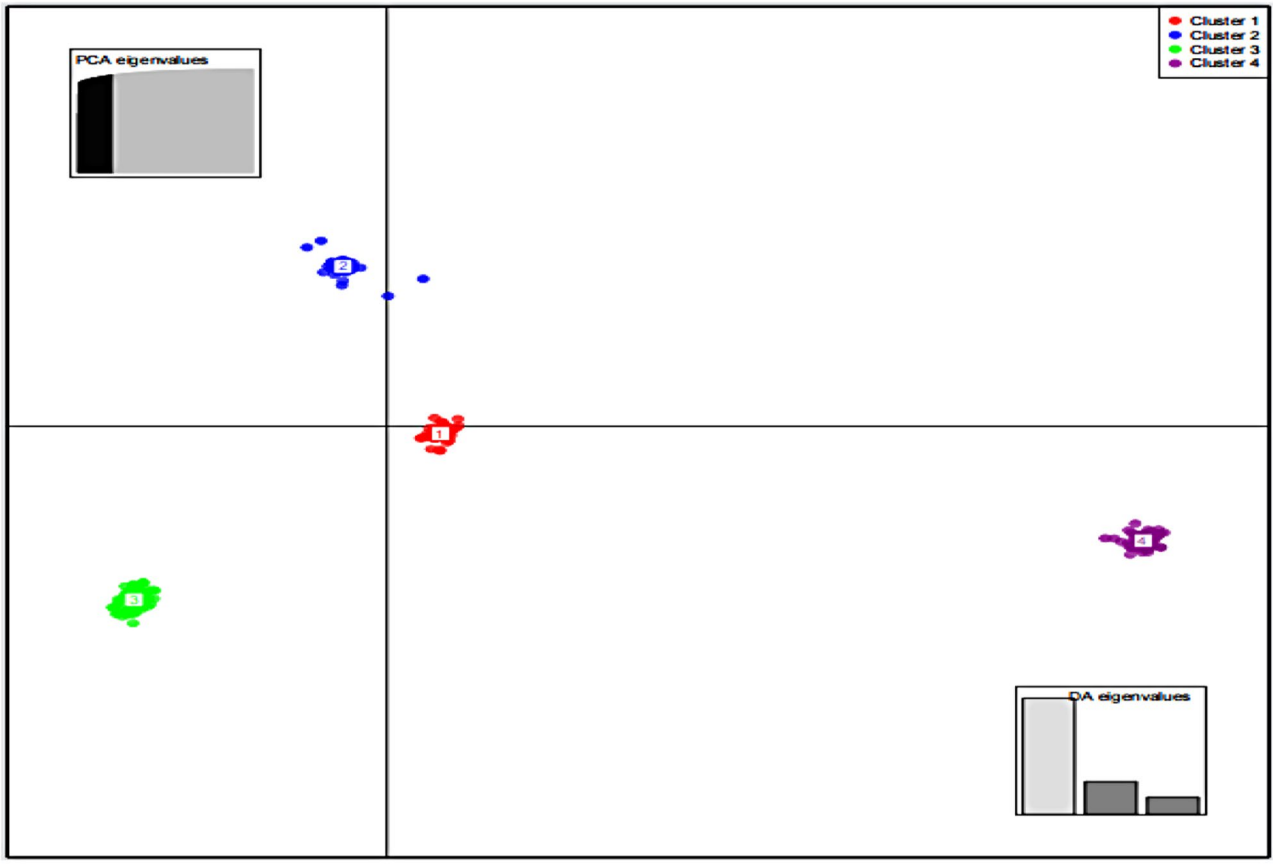

Fig. 4 Discriminant analysis of principal components (DAPC) using 3047 DArT-SNP markers. The axes represent the first two linear discriminants (LD). Each color depicts a cluster, and each dot represents a cultivar. The numbers represent the different subpopulations identified by DAPC analysis

from Oyo (10), Ondo (1), Akwa Ibom (47) and Ebonyi (5). All Dasheens (215 cultivars) grouped in Cluster 2, while the Eddoes (275 cultivars) were distributed across all four clusters.

The hierarchical clustering based on 3047 SNP markers on 490 cultivars also revealed four major sub-groups: I, II, III, and IV (Fig. 5). Similar to admixture and DAPC, all 215 Dasheen cultivars grouped together in sub-group IV, while the Eddoes were distributed in the other three sub-groups. Among Eddoes, 162 cultivars grouped together in sub-group I, 64 grouped in sub-group III, and 46 grouped in sub-group II. The three admixed cultivars in population structure, were found in Cluster 2 (1), Cluster 3 (1), and Cluster 4 (1) of DAPC, and sub-group I (1) and III (2) of HC.

The genetic distances (GD) among the cultivars (Supplementary Table S4) were lowest (0.0087) among Dasheens, between TR0724 and TR0744, and TR0724 and TR0664. TR0724 and TR0744 were collected from the same farmer's field in Kwara State, while TR0664 was collected from a farmer's field in Oyo State. The highest GD (0.57631) was observed among Eddoes, between TR0035 and TR0449, and TR0035 and TR0203. The cultivars were collected from Ebonyi (TR0035), Oyo (TR0449) and Ekiti (TR0203) States.

The AMOVA analysis revealed significant genetic variation between the two gene pools (49%), within gene pools (32%), and among cultivars within gene pools (18%). The F_{ST} was 0.492, and gene flow (Nm) was 0.2403 (Table 5). The AMOVA among States of collection revealed that 41% of the variation was among the States, 39% within the States, and 19% among the cultivars within the States. The F_{ST} and Nm was 0.415 and 0.353, respectively (Table 6). The pairwise F_{ST} ranged from 0 (Ondo and Anambra cultivars) to 0.9 (Kwara and Anambra cultivars), while the gene flow was highest between Ondo and Anambra cultivars (1458.98) and lowest between Kwara and Anambra cultivars (0.028) (Table 7). The pairwise genetic distance was lowest between cultivars from Ondo and Anambra (0.01), and peaked between Anambra and Kwara cultivars (0.413) (Table 8).

Combined analysis of phenotypic and molecular data

The hierarchical clustering based on composite matrix [combining both morphological (10 quantitative traits) and molecular data (3047 DArT-SNPs)] delineated 114

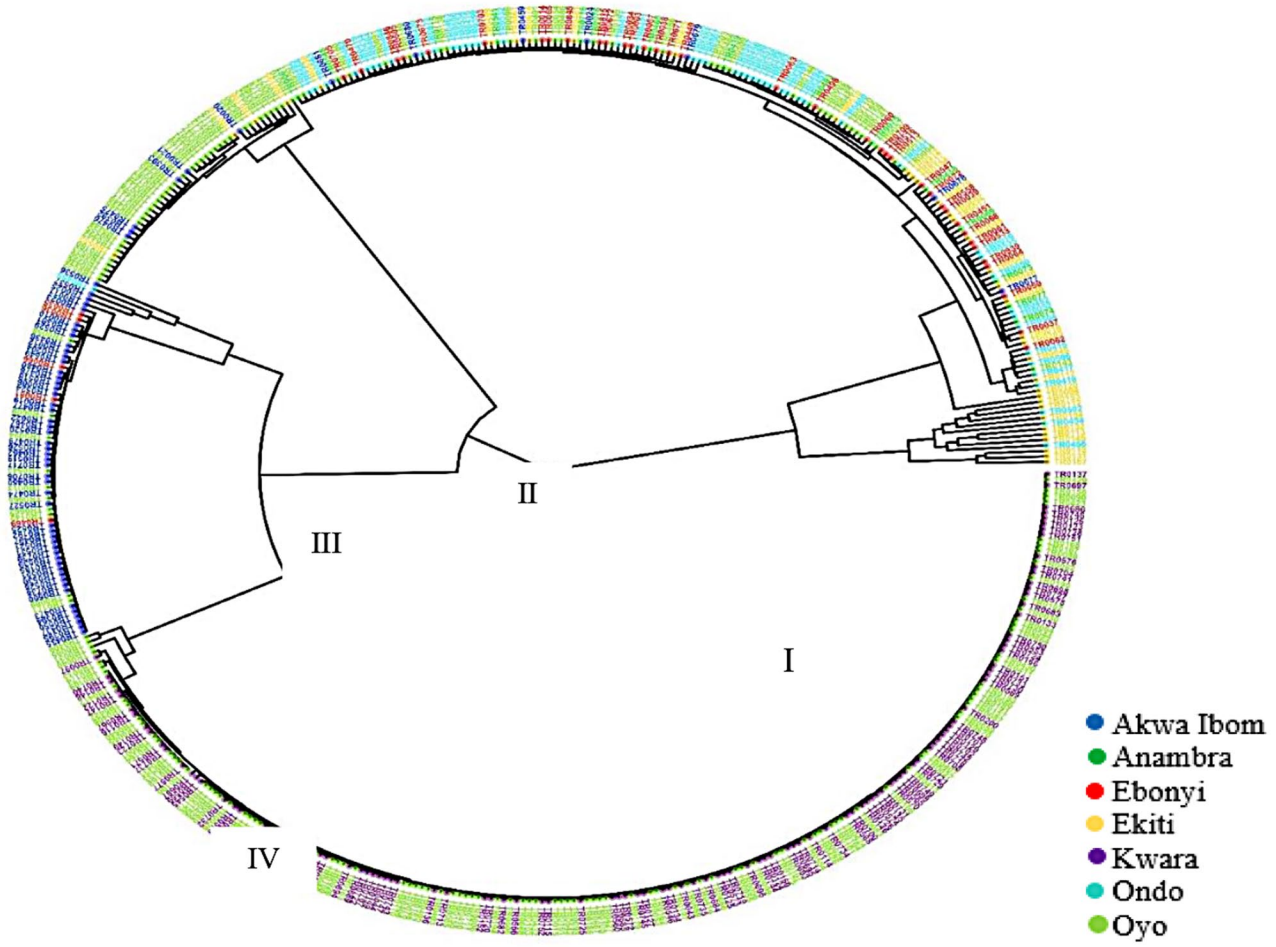

Fig. 5 Ward's minimum variance hierarchical cluster of 490 taro cultivars using 3047 DArT-SNP markers

Table 5 Analysis of molecular variance for 490 cultivars across two gene pools using 3047 DArT-SNP markers

Source	Degree of freedom	Sum of squares	Mean squares	Estimated variance	Percentage variance	*p* value
Among Gene Pools	1	168252.10	168252.10	347.60	49%	0.001
Among Cultivars	488	237892.60	487.50	129.30	18%	0.001
Within Gene Pools	490	112154.50	228.90	228.90	32%	0.001
Total	979	518299.20		705.80	100%	0.001
F_{ST}	0.492					0.001
Nm	0.258					0.001
GD among pops	0.240					0.001

Table 6 Analysis of molecular variance among the States of collection of 490 taro cultivars using 3047 DArT-SNP markers

Source	Degree of freedom	Sum of squares	Mean squares	Estimated variance	Percentage variance	*p* value
Among States	6	187318.16	31219.70	241.60	41%	0.001
Among Cultivars	483	218811.19	453.00	112.07	19%	0.001
Within States	490	112154.50	228.9	228.89	39%	0.001
Total	979	518283.85		582.56	100%	0.001
F_{ST}	0.415					0.001
Nm	0.353					0.001

Table 7 Pairwise F_{ST} (lower diagonal) and estimates of gene flow (upper diagonal) among states of collection of 490 taro cultivars using 3047 DArT-SNP markers

	Oyo	Ekiti	Ondo	Kwara	Akwa Ibom	Ebonyi	Anambra
Oyo	-	0.549	0.33	1.195	0.944	0.398	0.301
Ekiti	0.313	-	7.445	0.088	0.372	16.331	5.824
Ondo	0.431	0.032	-	0.06	0.234	39.087	1458.980
Kwara	0.173	0.74	0.806	-	0.167	0.07	0.028
Akwa Ibom	0.209	0.402	0.516	0.599	-	0.311	0.21
Ebonyi	0.386	0.015	0.006	0.781	0.446	-	20.705
Anambra	0.453	0.041	0.000	0.900	0.544	0.012	-

Table 8 Pairwise genetic distance among States of collection

	Oyo	Ekiti	Ondo	Kwara	Akwa Ibom	Ebonyi	Anambra
Oyo	-						
Ekiti	0.088	-					
Ondo	0.157	0.019	-				
Kwara	0.085	0.279	0.349	-			
Akwa Ibom	0.05	0.127	0.192	0.167	-		
Ebonyi	0.132	0.016	0.008	0.323	0.143	-	
Anambra	0.215	0.043	0.01	0.413	0.25	0.023	-

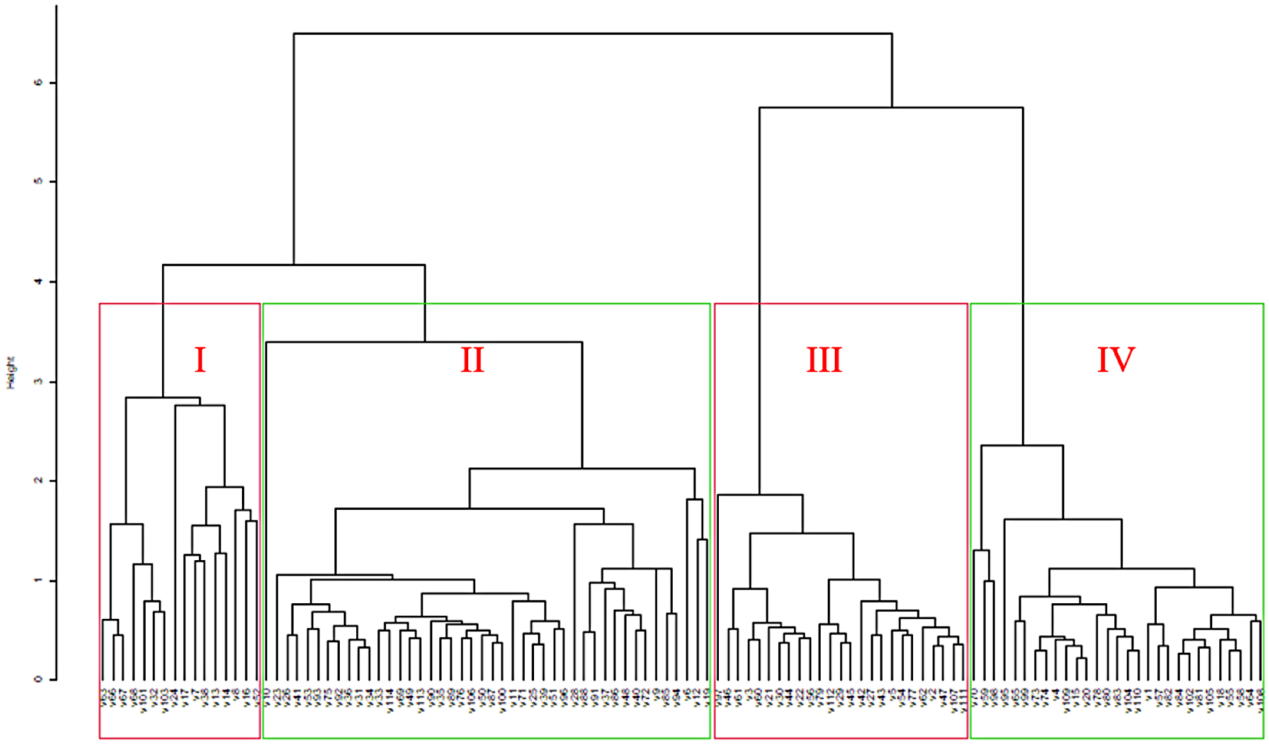

Fig. 6 Ward's minimum variance hierarchical clustering based on composite distance matrix of 114 taro cultivars

cultivars into four clusters (Fig. 6). Cluster I comprised 16 cultivars including three Eddoes and two Dasheens from Oyo, four Eddoes from Ekiti, three Eddoes from Ebonyi, two Eddoes from Ondo, one Eddoe from Anambra, and one Dasheen from Kwara. Cluster II had 44 cultivars consisting of six Dasheens and five Eddoes from Oyo, eight Eddoes from Akwa Ibom, eight Eddoes from Ekiti, eight Dasheens from Kwara, four Eddoes from Ebonyi, three Eddoes from Ondo, and two Eddoes from Anambra. Cluster III had 25 cultivars consisting of seven Eddoes from Akwa Ibom, six Dasheens and one Eddoe from Oyo, three Eddoes from Ebonyi, three Eddoes from Ekiti, three Dasheens from Kwara, one Eddoe from Anambra, and one Eddoe from Ondo. Twenty-nine cultivars which

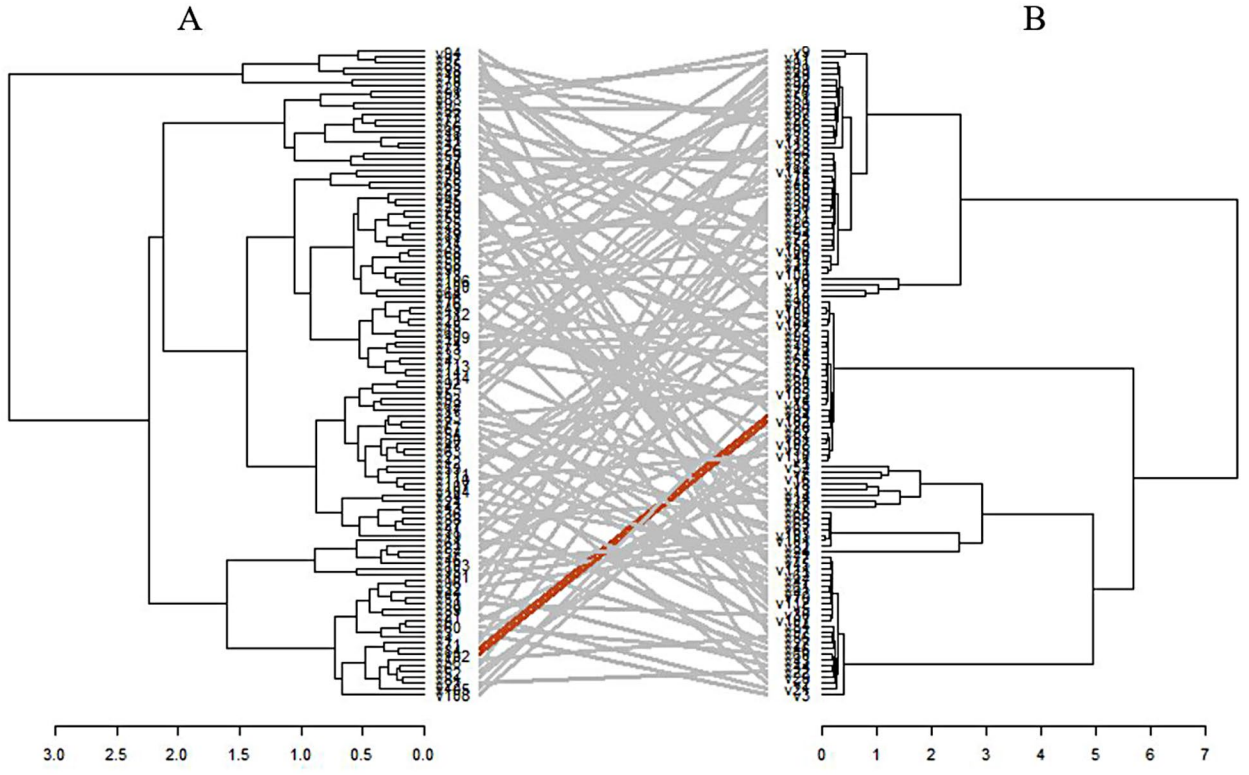

Fig. 7 Comparison of phenotypic (**A**) and genotypic (**B**) hierarchical cluster dendrograms of 114 taro cultivars. The grey lines between the dendrograms are the mismatched cultivars, while the orange lines are the cultivars in the same position in the phenotypic and genotypic hierarchical clusters

comprised 10 Eddoes from Akwa Ibom, six Eddoes from Ekiti, five Eddoes from Ondo, four Eddoes from Ebonyi, two Dasheens from Kwara, and one Eddoe and one Dasheen from Oyo grouped in Cluster IV.

The Mantel test between phenotypic and genotypic genetic distances showed a correlation (r) value of 0.01, while the correlation between the phenotypic and composite distance matrices was 0.00. However, the r-value between the molecular and composite genetic distances was 0.70. There was little (1.7%) similarity between phenotypic and molecular hierarchical cluster dendrograms, as shown in the tanglegram (Fig. 7).

Discussion

Knowledge of genetic diversity for the available germplasm is a pre-requisite for genetic improvement of taro in Nigeria. In this study, we collected taro cultivars across major taro growing States representing most important agro-ecological regions in Nigeria. A total of 490 cultivars were collected from seven States, and were characterized with both agro-morphological and molecular markers to assess the extent of genetic diversity and population structure.

The high level of variability was majorly observed for qualitative traits such as Lamina orientation (LO), Corm bud color (CBC), Corm shape (CS) and Corm skin surface (CSS). These four traits are important for cultivation practices and crop utilization; plant health, growth stage and visual appeal; yield, storage and culinary use; and disease resistance [22, 59, 60]. Amadi et al. [21] reported similar variability for LO, CBC and CS on Nigerian taro cultivars.

Although Eddoes showed more morphotypes than Dasheens, Corm fiber color (CFC) was the main distinguishing morphological trait between these two gene pools, the Eddoes had light yellow fibers while the Dasheens had purple fibers. Similar variability in CFC among the two taro gene pools was reported from the Republic of Benin by Quenum et al. [61]. Furthermore, the Dasheen cultivars in this study were characteristically similar to Dasheen cultivars from China [24], South-East Asia and Oceania [62] and Republic of Benin [61]. Additionally, farmers and consumers prefer specific traits such as Petiole color, Corm shape and Corm bud color in a cultivar since these traits contribute towards plant health, vigor and culinary attributes [22]. Therefore, these traits should be considered as important selection criteria in any breeding program. For example, in this study, the "drooping" and "erect-apex down" type of Lamina Orientation (LO) was most prominent among the taro cultivars. These LOs are desirable morphological traits as "drooping" LO allows for better light penetration to lower

leaves increasing photosynthesis and water also falls easily from the leaves reducing the risk of fungal infection, while the "erect apex down" LO reduces weediness due to quick canopy formation [59, 63]. This is in agreement with Oladimeji et al. [64] who in their review reported the prominence of these two lamina orientations among Nigerian taro. Similarly, Lebot et al. [62] also reported the prominence of both LOs in taro cultivars of Asian origin.

The delineation of 114 cultivars into four groups based on ten quantitative traits is important for initiating taro breeding program in the country as parents can be selected from different clusters for targeted traits such as Cluster I cultivars for high yield and Cluster II for weed control due to early emergence. Parental lines can be selected from Clusters I, II and III to generate bi-parental mapping populations for targeted traits such as yield and days to emergence. Cluster IV had the most diverse group of cultivars, therefore parental selection from this cluster will meet the aim of adaptation to varying climatic conditions and consumer preferences. There was a high positive and significant correlation between Total yield and the Number of cormels per hill in this study. This implies that there is enhanced yield potential and farmers can manipulate planting density and nutrient application to increase the number of cormels for yield increase. Similar high correlation between Yield and Number of cormels was reported by Amadi et al. [21] (0.54) and by Cheema et al. [65] (0.83). The majority (90%) of the quantitative traits measured in this study showed high heritability, consistent with the report of Mulualem et al. [66] on Ethiopian taro.

This study used 3047 DArT-SNP markers to dissect the genetic diversity among 490 taro cultivars. Three complementary approaches were combined here: admixture ancestry (population structure), the DAPC, and hierarchical cluster dendrogram to assess the genetic diversity and population structure among 490 cultivars. All approaches grouped cultivars into four distinct clusters with Dasheen cultivars grouped together in the same cluster across all three methods indicating low variability among the cultivars. The DAPC and hierarchical clustering showed high consonance, validating the presence of genetically related individuals which can help guide the selection of parents in breeding programs. Similar consistency between hierarchical clustering and DAPC in genetic diversity studies has been reported on yam [67, 68] and maize [31]. The level of admixture among the taro cultivars in this study was very low, possibly because taro is clonally propagated, resulting in high genetic uniformity [69]. It has been well documented that clonality reduces genetic variation as a result of the absence of segregation and genetic recombination [70, 71]. This is coupled with the lack of active breeding program in Nigeria that could introduce genetic variation through gene flow and/or genetic recombinations.

The 3047 DArT-SNPs used in the study was able to distinguish among the cultivars. Low genetic diversity observed among Dasheens is in agreement with the study from Thailand (0.007) wherein AFLP markers were used [2]. On the other hand, the Eddoes were highly diverse. Miyasaka et al. [4] reported that triploid taro (Eddoes) is common in Africa while diploids (Dasheen) are not, this could be the reason for the higher diversity among Eddoes when compared to Dasheens in Nigeria and the latter could be a recent introduction. This is also supported by AMOVA analysis, which revealed that about half of the variation was between the gene pools which may be mainly contributed due to the genetic diversity observed among the Eddoes compared to Dasheens. High molecular differences between the Dasheens and Eddoes is well established and reported in several studies [2, 4, 64]. Furthermore, high genetic divergence has also been observed between the Eddoes and Dasheens among the States of collection. The genetic distances (GD) among Kwara cultivars (all Dasheen) and other states with only Eddoe cultivars were high. However, low GD between cultivars from Kwara and Akwa Ibom, despite belonging to different gene pools, indicated some degree of similarity (for most qualitative traits except Corm shape and Corm fiber color) and lack of movement of cultivars. Lebot et al. [62] classified the two gene pools (Dasheen and Eddoes) as different ideotypes, and suggested that breeders should make improvement independently, and it is needless to interbreed the two owing to their morphological differences and consumer preferences. Additionally, Dasheens are diploids and Eddoes are triploids [4], making intermating difficult among them.

The analysis of molecular variance among the cultivars from different States also revealed substantial variation indicating that geographical isolation may be the reason for genetic diversity. It also indicated that differences among the cultivars from the States was the greatest cause of genetic variation, which are unique to their geographical environment meeting required consumers' preferences. Fufa et al. [23] reported 1% among-population molecular variance in their study while in this study, the among-population variance was to the extent of 41% and variation among individuals was 19% indicating that collection of taro cultivars from diverse agro-ecological regions covering larger number of States contributed towards this variation.

A low correlation was found between the genotypic and phenotypic-trait-based distance matrices which is in agreement with other studies in maize (0.05) [31] and sweetpotato (0.13) [72], while it was moderate in winged-yam (0.4) [62]. This suggests that it is always better to assess genetic diversity based on both phenotypic traits

and molecular markers [31]. The low correlation can be further explained by the phenotypic plasticity exhibited in phenotypic traits wherein a genotype can produce several phenotypes due to environmental effects [73, 74].

Conclusion

In this study, taro cultivars representing major growing regions and States of Nigeria were collected to assess genetic diversity and population structure. This is the first study that included taro cultivars from all the agro-ecological regions of the country where taro is currently grown and marketed. The characterization of taro cultivars using both morphological and molecular markers provided a broader perspective on genetic diversity within and between the two gene pools presently available in the country. Generally, higher genetic diversity was observed among Eddoes than Dasheens based on both morphological as well as molecular characterization. This could be attributed to the recent introductions of Dasheens into Nigeria (personal communications with farmers) compared to Eddoes. It is believed that Eddoes were the first introductions to Nigeria 2000 years ago. Furthermore, additional data based on multi-location evaluation of the germplasm collected in this study is necessary to validate and understand the factors responsible for low genetic correlations between phenotypic and genotypic diversity observed in this study. Nevertheless, the findings are useful to facilitate conservation efforts and initiate a strong breeding program for taro improvement in Nigeria and beyond. The corms and cormels of 490 taro cultivars are currently conserved at the Genetic Resource Center (GRC) of IITA, Ibadan, Nigeria.

Abbreviations

Nm	Gene flow
DArT-SNP	Diversity array technology single nucleotide polymorphism
DArTseq	Diversity array technology sequencing
NRCRI	National root crops research institute
H_e	Expected heterozygosity
H_o	Observed heterozygosity
MAF	Minor allele frequency
PIC	Polymorphic information content
VCF	Variant call format
BIC	Bayesian information criterion
DAPC	Discriminant analysis of principal component
AMOVA	Analysis of molecular variance
F_{ST}	Fixation index
HC	Hierarchical cluster dendrogram
LD	Linear discriminants
GD	Genetic distances
r	Correlation

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-024-05791-1.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

The authors would like to thank the staff of the Germplasm Health Unit and Virology and Molecular Diagnostics Unit, the Yam Breeding Unit, Mr. Peter Illuebbey, and the Cassava Unit at IITA for their immense support during the collection expedition and agro-morphological characterization. Special thanks also go to the Bioscience Center at IITA for their support during the sample preparation for genotyping, Dr. Stanley Adekemi for her support during the analysis and Mr. Oyedele Oluwafemi for the generation of the collection sites' map.

Author contributions

R.B., A.A., P.L.K., R.R.V. and R.B. contributed to the conception; J.J.O., A.A., P.L.K., R.R.V. and R.B. contributed to the methods; J.J.O., P.A. and O.J.I. analyzed the data; J.J.O. prepared the original draft; J.J.O., A.A., P.L.K., R.R.V., P.A., O.J.I. and R.B. reviewed and edited the manuscript; A.A., P.L.K., R.R.V. and R.B. were involved in supervision. All authors read and approved the final manuscript.

Funding

This work has been funded by The Swedish Research Council (Grant No. 2019–04270); the Pan African University Life and Earth Sciences Institute (PAULESI), Ibadan, Nigeria; the International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria; and the Swedish University of Agricultural Sciences (SLU), Sweden.

Data availability

All data supporting the findings of this study are included in this article and its supplementary files. The metadata used for the analysis are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

The plant materials used in the study were purchased from the farmers and marketers. Sampling was done on private land with due permission from the owners.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Plant Breeding Program, Pan African University Life and Earth Sciences Institute (Including Health and Agriculture), Ibadan, Oyo State 200284, Nigeria
[2]International Institute of Tropical Agriculture (IITA), PMB 5320, Oyo Road, Ibadan, Oyo State 200001, Nigeria
[3]Department of Crop and Horticultural Sciences, University of Ibadan, Ibadan, Oyo State 200284, Nigeria
[4]Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Lomma SE-23422, Sweden

Received: 30 May 2024 / Accepted: 6 November 2024
Published online: 14 November 2024

References

1. Kuruvilla K, Singh A. Karyotypic and electrophoretic studies on taro and its origin. Euphytica. 1981;30(2):405–13. https://doi.org/10.1007/BF00034004.
2. Kreike CM, Van Eck HJ, Lebot V. Genetic diversity of taro, *Colocasia esculenta* (L.) Schott, in Southeast Asia and the Pacific. Theor Appl Genet. 2004;109:761–8. https://doi.org/10.1007/s00122-004-1691-z.
3. Chaïr H, Traore RE, Duval MF, Rivallan R, Mukherjee A, Aboagye LM, Van Rensburg WJ, Andrianavalona V, Pinheiro de Carvalho MA, Saborio F, Sri Prana M. Genetic diversification and dispersal of taro (*Colocasia esculenta* (L.) Schott). PLoS ONE. 2016;11(6):e0157712. https://doi.org/10.1371/journal.pone.0157712.
4. Miyasaka SC, Bellinger MR, Kantar MB, Helmkampf M, Wolfgruber T, Paudel R, Shintaku M. Genetic diversity in taro (*Colocasia esculenta*). Genet Divers Hortic Plants. 2019;191–215. https://doi.org/10.1007/978-3-319-96454-6_7.

5. Singh S, Singh DR, Faseela F, Kumar N, Damodaran V, Srivastava RC. Diversity of 21 taro (*Colocasia esculenta* (L.) Schott) accessions of Andaman Islands. Genet Resour Crop Evol. 2012;59:821–9. https://doi.org/10.1007/s10722-011-9721-6.
6. Fujimoto T. Taro (*Colocasia esculenta* [L.] Schott) cultivation in vertical wet-dry environments: farmers' techniques and cultivar diversity in Southwestern. Ethiopia Econ Bot. 2009;63:152–66. https://doi.org/10.1007/s12231-009-9074-7.
7. Ermayanti TM, Rantau DE, Wulansari A, Martin AF, Al Hafiizh E. Variasi Jumlah Kromosom Talas Bentul (*Colocasia esculenta* (L.) Schott) in Vitro Hasil Perlakuan Orizalin. JBI. 2019;15(1).
8. Nguyen Xuan V, Pham Thi Viet A, Nguyen Thi Quynh H, Le Thi Tuyet M, Vu Thi Bich H, Le Thi T. Study on chromosome number and karyotype in the north taro germplasm preserved at the Plant Resources Center - Vietnamese Academy of Agricultural Sciences. J Nat Sci. 2021.
9. Wang JK, Steinke WE, Carpenter JR. Food, feed and fuel from taro. In Proceedings of the International Conference on Agricultural Engineering and Agro-Industries in Asia, Asian Institute of Technology, Bangkok, Thailand, 10–13 November 1981. Asian Institute of Technology. 1983:454–470.
10. Rashmi DR, Raghu N, Gopenath TS, Palanisamy P, Bakthavatchalam P, Karthikeyan M, Gnanasekaran A, Ranjith MS, Chandrashekrappa GK, Basalingappa KM. Taro (*Colocasia esculenta*): an overview. J Med Plants Stud. 2018;6(4):156–61.
11. Ramanatha RV, Matthews PJ, Eyzaguirre PB, Hunter D. The Global Diversity of Taro: ethnobotany and conservation. 2010.
12. Okon B, Ibon LA, Nsa EE, Ubua JA. Reproductive and growth traits of parents and F1 hatchings of *Achatina achatina* (L.) snails under mixed feeding regime with graded levels of swamp taro cocoyam (*Cyrtosperma chamissonis*) and pawpaw leaves (*Carica papaya*). J Agric Sci. 2012;4(11):289–98.
13. Bammite D, Matthews PJ, Dagnon DY, Agbogan A, Agre P, Akintayo TO, Odah K, Dansi A, Abberton M, Tozo K. Genetic and morphological diversity in *Colocasia esculenta* and *Xanthosoma maffafa* in Togo, West Africa. Adv Hort Sci. 2021;35(3):255–67. https://doi.org/10.36253/ahsc-9689.
14. Lee W. Taro (*Colocasia esculenta*). Ethnobotanical Leaflets. 1999;1999(1):4.
15. Okonkwo CA. Polymorphism in taro (*Colocasia esculenta* (L.) Schott) and its implication in taro breeding and production in Nigeria. J Hortic Sci. 1987;62(3):395–400. https://doi.org/10.1080/14620316.1987.11515797.
16. Ubalua AO, Ewa F, Okeagu OD. Potentials and challenges of sustainable taro (*Colocasia esculenta*) production in Nigeria. J appl biol. 2016;4(1):053–9. https://doi.org/10.7324/JABB.2016.40110.
17. Grimaldi IM, van Andel TR. Food and medicine by what name? Ethnobotanical and linguistic diversity of taro in Africa. Econ Bot. 2018;72(2):217–28. https://doi.org/10.1007/s12231-018-9413-7.
18. Matthews PJ. An introduction to the history of taro as a food: Global Diversity of Taro. 2010. minpaku.repo.nii.ac.jp
19. Statistical Database FAOSTATFAO. 2023. https://www.fao.org/faostat/en/#data/QCL. Accessed 8 March 2023.
20. Onyeka J. Status of cocoyam (*Colocasia esculenta* and *Xanthosoma spp*) in West and Central Africa: production, household importance and the threat from leaf Blight. CGIAR Research Program on Roots, Tubers and Bananas (RTB). 2014:1–39.
21. Amadi CO, Onyeka J, Chukwu GO, Okoye BC, Ezeji L, Ezigbo EC. Evaluation of exotic genotypes of taro (*Colocasia Esculenta*) in Nigeria. Niger Agric J. 2015;46(1):36–42.
22. Fufa TW, Oselebe HO, Nnamani CV, Afiukwa CA, Uyoh EA. Systematic review on farmers' perceptions, preferences and utilization patterns of Taro [*Colocasia esculenta* (L.) Scott] for food and nutrition security in Nigeria. J Plant Sci. 2021;9(4):224–33.
23. Fufa TW, Abtew WG, Amadi CO, Oselebe HO. DArTSeq SNP-based genetic diversity and population structure studies among taro [*Colocasia esculenta* (L.) Schott] accessions sourced from Nigeria and Vanuatu. PLoS ONE. 2022;17(11):e0269302. https://doi.org/10.1371/journal.pone.0269302.
24. Jianchu X, Yongping Y, Yingdong P, Ayad WG, Eyzaguirre PB. Genetic diversity in taro (*Colocasia esculenta* Schott, Araceae) in China: an ethnobotanical and genetic approach. Econ Bot. 2001;55:14–31.
25. Mandal R, Mukherjee A, Mandal N, Tarafdar J, Mukharjee A. Assessment of genetic diversity in Taro using morphometrics. Curr Agric Res J. 2013;1(2):79. https://doi.org/10.12944/CARJ.1.2.02.
26. Vinutha KB, Devi AA, Sreekumar J. Morphological characterization of above ground characters of Taro (*Colocasia esculenta* (L.) Schott.) Accessions from North East India. J Root Crops. 2015;41(1):3–11.
27. Bammite D, Matthews PJ, Dagnon DY, Agbogan A, Komi OD, Dansi A, Koffi TO. Agro morphological characterization of taro (*Colocasia esculenta*) and yautia (*Xanthosoma mafaffa*) in Togo, West Africa. Afr J Agric Res. 2018;13(18):934–45. https://doi.org/10.5897/AJAR2018.13043.
28. Boampong R, Aboagye LM, Nyadanu D, Esilfie M. Agro-morphological characterization of some taro (*Colocasia esculenta* (L.) Schott.) Germplasms in Ghana. J. Plant Breed. Crop Sci. 2018;10(8):191–202. https://doi.org/10.5897/JPBCS2018.0734.
29. Ouédraogo N, Sombié D, Eric PA, Traoré RE, Sama H, Bationo Kando P, Sawadogo M, Lebot V. Nutritional and phytochemical characterization of taro (*Colocasia esculenta*) germplasm from Burkina Faso. J. Plant Breed. Crop Sci. 2023;15(1):32–41. https://doi.org/10.5897/JPBCS2022.0999.
30. Vogt L, Bartolomaeus T, Giribet G. The linguistic problem of morphology: structure versus homology and the standardization of morphological data. Cladistics. 2010;26(3):301–25. https://doi.org/10.1111/j.1096-0031.2009.00286.x.
31. Stanley A, Menkir A, Paterne A, Ifie B, Tongoona P, Unachukwu N, Meseka S, Mengesha W, Gedil M. Genetic diversity and population structure of maize inbred lines with varying levels of resistance to *Striga hermonthica* using agronomic trait-based and SNP markers. Plants. 2020;9(9):1223. https://doi.org/10.3390/plants9091223.
32. Sansaloni C, Franco J, Santos B, Percival-Alwyn L, Singh S, Petroli C, Campos J, Dreher K, Payne T, Marshall D, Kilian B. Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. Nat Commun. 2020;11(1):4572. https://doi.org/10.1038/s41467-020-18404-w.
33. von Cruz M, Kilian A, Dierig DA. Development of DArT marker platforms and genetic diversity assessment of the U.S. Collection of the new oilseed crop lesquerella and related species. PLoS ONE. 2013;8:e64062. https://doi.org/10.1371/journal.pone.0064062.
34. Kilian A, Graner A. NGS technologies for analyzing germplasm diversity in genebanks. Brief Funct Genom. 2012;11:38–50. https://doi.org/10.1093/bfgp/elr046.
35. Jighly A, Alagu M, Makdis F, Singh M, Singh S, Emebiri LC, Ogbonnaya FC. Genomic regions conferring resistance to multiple fungal pathogens in synthetic hexaploid wheat. MOLB. 2016;36:1–9. https://doi.org/10.1007/s11032-016-0541-4.
36. Mahboubi M, Mehrabi R, Naji AM, Talebi R. Whole-genome diversity, population structure and linkage disequilibrium analysis of globally diverse wheat genotypes using genotyping-by-sequencing DArTseq platform. 3 Biotech. 2020;10(2):48. https://doi.org/10.1007/s13205-019-2014-z.
37. Nelimor C, Badu-Apraku B, Garcia-Oliveira AL, Tetteh A, Paterne A, N'guetta AS, Gedil M. Genomic analysis of selected maize landraces from Sahel and Coastal West Africa reveals their variability and potential for genetic enhancement. Genes. 2020;11(9):1054. https://doi.org/10.3390/genes11091054.
38. Bocianowski J, Tomkowiak A, Bocianowska M, Sobiech A. The use of DArTseq technology to identify markers related to the heterosis effects in selected traits in maize. CIMB. 2023;45(4):2644–60. https://doi.org/10.3390/cimb45040173.
39. Gbedevi KM, Boukar O, Ishikawa H, Abe A, Ongom PO, Unachukwu N, Rabbi I, Fatokun C. Genetic diversity and population structure of cowpea [*Vigna unguiculata* (L.) Walp.] Germplasm collected from Togo based on DArT markers. Genes. 2021;12(9):1451. https://doi.org/10.3390/genes12091451.
40. Adu BG, Akromah R, Amoah S, Nyadanu D, Yeboah A, Aboagye LM, Amoah RA, Owusu EG. High-density DArT-based SilicoDArT and SNP markers for genetic diversity and population structure studies in cassava (*Manihot esculenta* Crantz). PLoS ONE. 2021;16(7):e0255290. https://doi.org/10.1371/journal.pone.0255290.
41. Bredeson JV, Lyons JB, Oniyinde IO, Okereke NR, Kolade O, Nnabue I, Nwadili CO, Hřibová E, Parker M, Nwogha J, Shu S. Chromosome evolution and the genetic basis of agronomically important traits in greater yam. Nat Commun. 2022;13(1):2001. https://doi.org/10.1038/s41467-022-29114-w.
42. Amponsah Adjei E, Esuma W, Alicai T, Bhattacharjee R, Dramadri IO, Edema R, Chamba EB, Odong TL. Genetic diversity and population structure of Uganda's yam (*Dioscorea spp.*) genetic resource based on DArTseq. PLoS ONE. 2023;18(2):e0277537. https://doi.org/10.1371/journal.pone.0277537.
43. Grimaldi IM, Leke WN, Borokini I, Wanjama D, Van Andel T. From landraces to modern cultivars: field observations on taro *Colocasia esculenta* (L.) Schott in sub-saharan Africa. Genet Resour Crop Evol. 2018;65(7):1809–28. https://doi.org/10.1007/s10722-018-0651-4.
44. IPGRI. Descriptors for Taro (*Colocasia esculenta*). Rome, Italy: International Plant Genetic Resources Institute; 1999.

45. Andarini YN, Afza H. Sutoro. Estimation of Leaf Area of Taro (*Colocasia esculenta*). JIPI. 2020;24(4):610–7. https://doi.org/10.18343/jipi.25.4.610.
46. Diversity Array Technology Pty Ltd. https://ordering.diversityarrays.com/files/DArT_DNA_isolation.pdf/. Accessed 25 May 2023.
47. DArT – DArTseq services. https://www.diversityarrays.com/services/dartseq/. Accessed 25 May 2023.
48. Yin J, Jiang L, Wang L, Han X, Guo W, Li C, Zhou Y, Denton M, Zhang P. A high-quality genome of taro (Colocasia esculenta (L.) Schott), one of the world's oldest crops. Mol Ecol Res. 2020;21(1). https://doi.org/10.1111/1755-0998.13239.
49. Aravind J, Wankhede DP, Kaur V, Aravind MJ, Package. 'augmentedRCBD'. 2021. https://aravind-j.github.io/augmentedRCBD/https://cran.r-project.org/package=augmentedRCBD
50. Drost HG. Philentropy: information theory and distance quantification with R. J Open Source Softw. 2018;3(26):765. https://doi.org/10.21105/joss.00765.
51. Danecek P. 1000 genomes project analysis group. The variant call format and VCFtools. Bioinf (Oxford England). 2011;27:2156–8. https://doi.org/10.1093/bioinformatics/btr330.
52. Liu K. PowerMarker: new genetic data analysis software. Version 3.0. Free program distributed by the author over the Internet at https://www.powermarker.net.2003
53. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009;19(9):1655–64. https://doi.org/10.1101/gr.094052.109.
54. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 2010;11:1–5. https://doi.org/10.1186/1471-2156-11-94.
55. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20(2):289–90. https://doi.org/10.1093/bioinformatics/btg412.
56. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2023. https://www.R-project.org/
57. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics. 2012;28(19):2537–9. https://doi.org/10.1093/bioinformatics/bts460.
58. Galili T. Dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. Bioinformatics. 2015;31(22):3718–20. https://doi.org/10.1093/bioinformatics/btv428.
59. Thakur P, Ram D, Naik U. Morphological characterization of Taro [*Colocasia esculenta* (L.) Schott] germplasm. JPP. 2021;10(2):1264–8.
60. Pasiona LC. Morphological characterization of taro (*Colocasia esculenta*) cultivars in the Bicol region, Philippines. Agrik CRI J. 2021;1:1–3.
61. Quenum ZN, Kumar PL, Akoroda MO, Dansi A, Vetukuri RR, Bhattacharjee R. Ethnobotany and perceptions on the value of Taro (*Colocasia esculenta*) among farmers in the Benin Republic. Afr J Agric Res. 2023;19(5). https://doi.org/10.5897/AJAR2022.16240.
62. Lebot V, Hartati S, Hue NT, Viet NV, Nghia NH, Okpul T, Pardales J, Prana MS, Prana TK, Thongjiem M, Krieke CM. Characterizing taro using isozymes and morpho-agronomic descriptors. In: Ramanatha RV, Matthews PJ, Eyzaguirre PB, Hunter D, editors. The global diversity of Taro: Ethnobotany and Conservation. Rome, Italy: Bioversity International; 2010. pp. 39–55.
63. Vinutha YN, Lohakare AS, Wankhede SR. Performance of taro (*Colocasia esculenta* L.) genotypes for qualitative traits under Marathwada conditions. Int J Adv Bioc Res. 2024;8(2):383–5. https://doi.org/10.33545/26174693.2024.v8.i2e.571.
64. Oladimeji JJ, Kumar PL, Abe A, Vetukuri RR, Bhattacharjee R. Taro in West Africa: Status, challenges, and opportunities. Agronomy. 2022;12(9):2094. https://doi.org/10.3390/agronomy12092094.
65. Cheema DS, Singh H, Dhatt AS, Sidhu AS, Garg N. Studies on genetic variability and correlation for yield and quality traits in Arvi [*Colocasia esculenta* (L.) Schott]. Inl International Conference on Indigenous Vegetables and Legumes. Prospectus for Fighting Poverty, Hunger and Malnutrition. 2007;752. pp. 255–260. https://doi.org/10.17660/ActaHortic.2007.752.41
66. Mulualem T, WeldeMichael G, Belachew K. Genetic diversity of Taro (*Colocasia esculenta* (L.) Schott) genotypes in Ethiopia based on agronomic traits. Time J Agric Vet Sci. 2013;1(2):23–30.
67. Agre P, Asibe F, Darkwa K, Edemodu A, Bauchet G, Asiedu R, Adebola P, Asfaw A. Phenotypic and molecular assessment of genetic structure and diversity in a panel of winged yam (*Dioscorea alata*) clones and cultivars. Sci Rep. 2019;9(1):18221. https://doi.org/10.1038/s41598-019-54761-3.
68. Bakayoko L, Pokou DN, Kouassi AB, Agre PA, Kouakou AM, Dibi KE, Nzue B, Mondo JM, Adebola P, Akintayo OT, Asfaw A. Diversity of Water Yam (*Dioscorea alata* L.) accessions from Côte d'Ivoire based on SNP markers and agronomic traits. Plants. 2021;10(12):2562. https://doi.org/10.3390/plants10122562.
69. Nybom H, Lācis G. Recent large-scale genotyping and phenotyping of plant genetic resources of vegetatively propagated crops. Plants. 2021;10(2):415. https://doi.org/10.3390/plants10020415.
70. O'Brien EK, Denham AJ, Ayre DJ. Patterns of genotypic diversity suggesta long history of clonality and population isolation in the Australian aridzone shrub Acacia Carneorum. Plant Ecol. 2014;215:55–71.
71. Huan R, Wang Y, Li K, Ying-Qiang W. Genetic variation and population structure of clonal Zingiber zerumbet at a fine geographic scale: a comparison with two closely related selfing and outcrossing Zingiber species. BMC Ecol Evol. 2021;21(1):116. https://doi.org/10.1186/s12862-021-01853-2.
72. Andrade EK, Andrade VC, Laia ML, Fernandes JS, Oliveira AJ, Azevedo AM. Genetic dissimilarity among sweet potato genotypes using morphological and molecular descriptors. Acta Sci Agron. 2017;39:447–55.
73. Zhao F, Xu S. Genotype by environment interaction of quantitative traits: a case study in barley. G3: Genes|. Genomes| Genet. 2012;2(7):779–88. https://doi.org/10.1534/g3.112.002980.
74. Uhlarik A, Ceran M, Živanov D, Grumeza R, Skøt L, Sizer-Coverdale E, Lloyd D. Phenotypic and genotypic characterization and correlation analysis of pea (*Pisum sativum* L.) diversity panel. Plants. 2022;111321. https://doi.org/10.3390/plants11101321.

Publisher's note