Genet Resour Crop Evol
https://doi.org/10.1007/s10722-024-02158-1

RESEARCH ARTICLE

Diversity among Bambara groundnut (*Vigna subterranea* L. Verdc) accessions using agro-morphological traits and diversity array technologies sequence low density markers in Malawi

Yalaukani Louis · Maliro Moses · Chipeta Michael · Gimode Davis · Sefasi Abel · Pungulani Lawrent · Morrone Vicki · Soko Peter · Changadeya Wisdom

Received: 28 June 2024 / Accepted: 31 August 2024

Abstract Bambara groundnut (*Vigna subterranea* (L.) Verdc) is a neglected and underutilized crop that plays a big role in improving livelihoods of smallholder farmers in Sub-Saharan Africa. Despite its importance, there is limited availability of commercially improved cultivars to smallholder farmers in Malawi. This study characterized selected Bambara groundnuts accessions for agro-morphological traits for germplasm discrimination. It also identified genetic variation using Single Nucleotide Polymorphism (SNP) markers through Diversity Array Technologies Sequence Low Density (DArTseqLD) that could be used to produce improved seed for crop improvement. Forty Bambara groundnuts accessions were evaluated at the Crops and Soil Sciences Department's farm of Lilongwe University of Agriculture and Natural Resources, Bunda College, Malawi. From the 40 accessions, 188 unique seed samples were selected for genotyping using DArTseqLD SNP markers. Data on agro-morphological traits were collected following the Bambara groundnut descriptor guidelines and multivariate analysis were performed. Principal Component Analysis revealed a total variation of 53%. The study generated 1048 DArTseqLD SNP markers. Analysis of molecular variance (AMOVA) identified 84% and 13% of genetic variation among and within the Bambara groundnut accessions respectively, whereas 3% genetic variation was observed among the total populations. Cluster analysis based on genotypic data grouped the 188 samples into 10 clusters. Based on phenotypic and genotypic data, it can be concluded that there is a significant degree of variation and genetic diversity in the accessions evaluated that can be used in crop improvement program as well as being directly used by farmers in seed production.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10722-024-02158-1.

Y. Louis · P. Lawrent
Department of Agricultural Research Services (DARS), Chitedze Research Station, P.O. Box 158, Lilongwe, Malawi

M. Moses (✉) · C. Michael · S. Abel
Lilongwe University of Agriculture and Natural Resources (LUANAR), P.O Box 219, Lilongwe, Malawi
e-mail: mmaliro@luanar.ac.mw

G. Davis
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), United Nations Avenue, Gigiri, Nairobi, Kenya

M. Vicki
Michigan State University (MSU), East Lansing, MI, USA

S. Peter
Self Help Africa, P/Bag B495, Lilongwe, Malawi

C. Wisdom
Chancellor College, Department of Biological Sciences, University of Malawi, P.O. Box 280, Zomba, Malawi

Published online: 09 September 2024

Keywords Bambara groundnut · Accessions · Smallholder farmers · Multivariate analysis · Phenotyping · Genotyping

Introduction

Bambara groundnut (*Vigna subterranea* (L.) Verdc.) is one of the underexploited crop with the opportunity to address some critical food requirements as it is a high-nutritional alternative food in Sub Saharan Africa (SSA) including Malawi (Khaliqi et al. 2021). It is a complete source of food as the seeds are high in protein (19–25%), carbohydrate (63%) and fiber (6.5%), as well as essential amino acids important for the human diet (Gbaguidi et al. 2018; Katungwe et al. 2015; Pungulani et al. 2012). Despite its importance in improving nutritional and economic status of smallholder farmers, the crop still remains neglected and underutilized. There is little research in Malawi on Bambara groundnut despite having a diverse genetic resources (Katungwe et al. 2015). For the crop to be fully used, there is need to conduct detailed research which will add value. Phenotyping and genotyping of Bambara groundnuts will help in coming up with good materials for crop improvement. Literature on neglected and underutilized crop species highlighted the importance of more molecular research to unlock the potentials (Mayes et al. 2013; Ntundu et al. 2006).

Molecular research is expected to shed new light on the population structure of native Bambara groundnut accessions which will aid the crop improvement through modern breeding techniques (Uba et al. 2021). Agromorphological characterization helps to reveal the genetic diversity that exists among Bambara groundnut useful by farmers and breeders to select the best performing germplasms for higher and more resilient yields (Umar & Kwon-ndung 2015). Quality Declared Seed Law is designed so that farmers can produce and market seed they produce, offering hospitable environment for farmer-led seed enterprise (Mastenbroek et al. 2021).To date this law is not applicable for use by farmers in Malawi, but this research that is part of an underutilized seed study with smallholder farmers, will contribute to operationalize this law through production of quality seed. Including farmers in the seed production process will not only assure user-preferred traits are included in selected cultivars but provide farmers with the means to practice appropriate seed production protocol, following the Quality Declared Seed Law regulations. Farmers assisting this process can improve availability of Bambara groundnut seed and help meet demand of the informal and formal seed systems (Almekinders & Louwaars, 2002, Mastenbroek et al. 2021; Pungulani et al. 2012; Owere et al. 2014). Farmer-led seed enterprises can lead to production and marketing of high-quality, fairly priced seed while promoting income security among smallholder farmers in Malawi. However, for the farmers wanting to produce and sell seed as a livelihood, a conducive policy and regulatory framework that is inclusive and less stringent is required as is the case with Malawi (Gregory et al. 2019).

Currently there is minimal production of Bambara groundnut in Malawi (Figure S3) due to lack of identification of high yielding germplasms in areas where this crop is commonly grown (MoAFS 2020). Farmers face difficulty to identify quality germplasms in part due to inadequate knowledge of plant taxonomy, the process of screening cultivars to selection for genetic improvement and ability to identify traits that increase adaptation to specific agro-ecological zones such as seed maturation time and day length sensitivity (Pungulani et al. 2012; Uba et al. 2021). Seeking genetic diversity includes seeking variation within individual gene loci/among alleles of a gene, or gene combinations, between individual plants or plant populations (Sood et al., 2016). Genetic variability is a critical component of any crop improvement program. When there are more diverse parents, more heterosis in the progeny is achieved and a higher chance of getting transgressive segregation. Breeders must identify diverse parents with high genetic variability in order to combine desirable characteristics for development of improved cultivars over those existing and grown by farmers (Patil & Modha 2022). Therefore, this research i) characterized Bambara groundnut's agro-morphological traits for the identification of traits of good performing germplasm for quality seed production in Malawi and ii) identified the genetic variations that exists within the Bambara groundnut germplasm genotyped using DArTseqLD makers. The information will be vital to plant breeders who seek to improve Bambara groundnut cultivars for crop improvement programs.

Materials and methods

Materials

A total of 40 accessions were collected in Eswazini, Chipala, Kandeu, Manjawira and Mombezi Extension Planning Area (EPA) in Malawi (Figure S1). Farmers selected their preferred Bambara groundnut seed types and then shared them at seed fairs with fellow farmers. They showcased their selections providing complete description of the plant's attributes. Three accessions were provided by the National Plant Genetic Resources Centre (NPGRC).

Experimental site and field layout

During the 2020/2021 growing season, the experiment was carried out at the Lilongwe University of Agriculture and Natural Resources (LUANAR)—Bunda College, Department of Crops and Soil Sciences Research Farm. The farm is located at an elevation of 1158 m above sea level with a latitude of 14º35'S and a longitude of 33° 50'E. The average annual temperature ranges from 18 °C to 27 °C and the area receives between 800 and 1031 mm of rain per year. The experiment was set up in an Alpha lattice design with three replications planted at a ridge length of 3 m spaced at 0.75 m between rows and 15 cm between planting stations with one seedling per station. Weeding, banking, pest and disease management was all carried out as provided in the Guide to Agricultural Production (MoAFS 2020).

Phenotypic data collection and analysis

Phenotypic data collection

Qualitative and quantitative data on all agro-morphological traits of Bambara groundnut (Table S1) were collected from vegetative to the processing stage following the Bambara nut descriptor by International Plant Genetic Resources Institute (IPGRI, 2000).

Phenotypic data analysis

Quantitative phenotypic data were subjected to Analysis of Variance (ANOVA) and multivariate analysis in the R software environment (R Core Team 2021). Mean separation were done at 5% level of probability. Normality test was done using quantile—quantile (Q-Q) plot using qqman package (Castillo-Gutiérrez et al. 2021). Principal component analysis was done using the factoextra package (Kassambara and Mundt, 2016) and biplot of principal components were generated using ggbiplot package in R software environment (R Core Team 2021. Pictures taken in the field were used and percentages were calculated for qualitative data.

Genotypic data collection and analysis

Leaf sampling for DNA extraction A total of 40 Bambara groundnut germplasm from farmers and NPGRC were used as the source of samples for DNA extraction. Based on the combination of four qualitative traits including seed shape, seed colour, eye pattern and testa pattern 188 unique seed samples were selected from the 40 Bambara groundnut accessions. These were planted on 23rd December, 2021 in the screen house of the Lilongwe University of Agriculture and Natural Resources (LUANAR). The plants were allowed to grow for three weeks to allow full establishment. Tender leaves from 188 plants were collected for DNA extraction. Sampling of leaves followed a procedure described by Intertek (Parmar et al. 2021). Two leaf discs were collected using a leaf puncher from a fresh clean leaf and placed in each of the 96 well Abgene plates while two wells were left blank as control. Then the samples were oven dried at 48 °C for 24 h. After 24 h the plates were sealed with silicone mats and sent to the Intertek laboratory in Sweden for DNA extraction process.

DNA extraction and genotyping High Molecular Weight (HMW) DNA was extracted from the tissue samples and quality checked in Intertek Sweden and then sent for the high-throughput DArTseqLD Single Nucleotide Polymorphism (SNP) genotyping at the Diversity Arrays Technology Pty Ltd. Canberra, Australia. This entailed subjecting the DNA samples to complexity reduction by digestion/ligation reactions as described by Kilian et al. (2012) but replacing the single PstI-compatible adaptor with two adaptors corresponding to PstI and MseI restriction enzymes' overhangs. The PstI-compatible adapter were designed to include Illumina flowcell attachment sequence, barcode region, while the reverse adapter contained flowcell attachment region and MseI-com-

patible overhang sequence. Only mixed fragments (PstI-MseI) were effectively amplified in PCR. Subsequently, equimolar amounts of PCR product from each sample were bulked and sequenced by the Hiseq2500/Novaseq6000 (Illumina® Inc., San Diego, CA, USA). Thereafter, the generated sequences were processed using proprietary DArT analytical pipelines. Approximately 250,000 sequences per barcode/sample were used in the marker call. Identical sequences were collapsed into FASTQCOL files followed by SNPs calling using the software package DArTsoft14.

Genotypic data analysis Germplasm sample data were filtered for quality using a minor allele frequency (MAF) threshold of 0.05 and call rate cut-off of 0.7 with the rest of the missing calls imputed by mean. Basic diversity metrics for the accessions were inferred using the snpReady package (Granato & Fritsche-Neto, 2018). These included the polymorphism information content (PIC), MAF, Nei's genetic diversity (GD), and observed heterozygosity (Ho). Grouping of the genotypes was evaluated using principal component analysis (PCA) as computed using the factoextra package (Kassambara & Mundt, 2016). An Euclidian distance matrix was used for dendrogram plotting and further clustering was done by discriminant analysis of principle components (DAPC) using the Adegenet package (Jombart et al., 2010). The analyses were done in the R software environment (R Core Team 2021). Analysis of Molecular Variance (AMOVA) was done using GeneAlEx v 6.5 (Peakall and Smouse 2012) with population groupings deduced from dendrogram and DAPC clustering.

Results

Terminal leaflet shape and growth habit of Bambara groundnut.

The study revealed that 22.5% of the 40 accessions had elliptic terminal leaflet type (Fig. 1a) and 37.5% had oval terminal leaflet type (Fig. 1b) while lanceolate terminal leaflet was exhibited by 40% of the accessions (Fig. 1c). Semi-bunch growth habit exhibited 52.5% of the total Bambara groundnut accessions (Fig. 1d) while 5% had a bunch type growth habit (Fig. 1e) and 42.5% had a spreading growth habit (Fig. 1f).

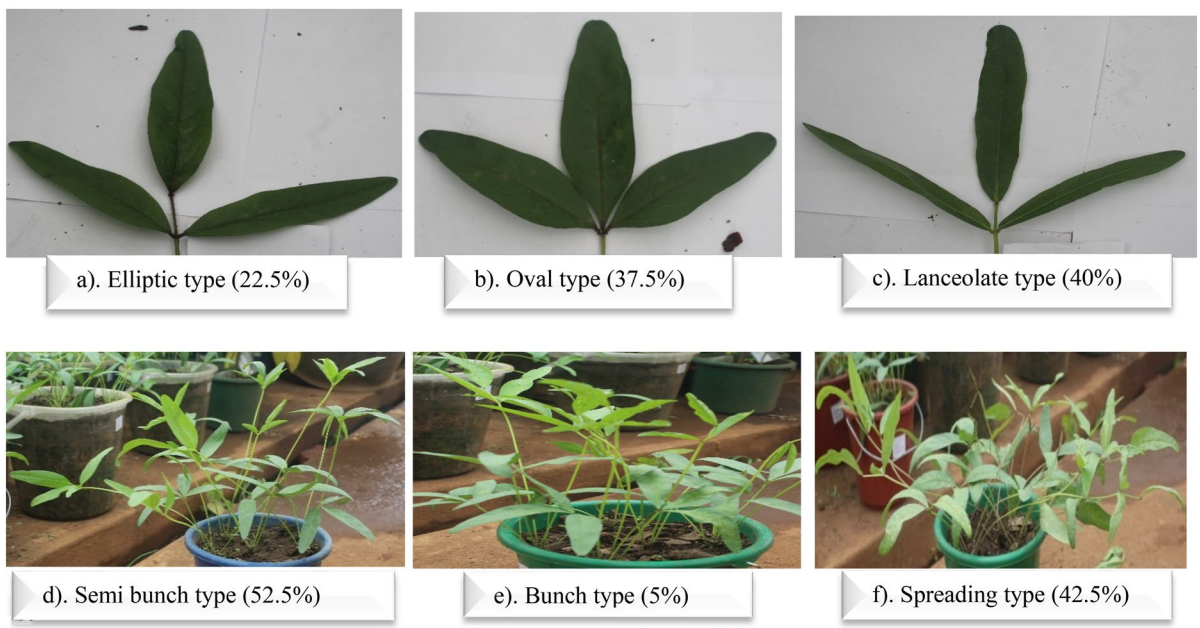

Fig. 1 Percent variation terminal leaflet shape (**1a**) Elliptic terminal leaflet type, **1b** Oval terminal leaflet type and **1c** Lanceolate terminal leaflet type and growth habit **1d** Semi bunch type, **1e** Bunch and **1f** Spreading type

Seed colour and eye pattern of Bambara groundnut

The study revealed that 67.5% had seed with cream testa colour (Fig. 2a) while 17.5% had light red (Fig. 2b) and 15% had dark red seed colour (Fig. 2c). The study revealed that 15% of the total Bambara groundnut accessions had an eye like butterfly (Figs. 2d) while 75% had an eye with 2 thin lines on both sides of the hilum (Fig. 2e) and 10% had no eye (Fig. 2f).

Testa pattern and seed shape of Bambara groundnut

The study revealed that 52.5% of the total Bambara groundnut accessions had a dotted testa pattern (Fig. 3a) while 20% had little rhomboid spotting on both sides of the hilum (Figs. 3b) and 5% had no testa pattern (Fig. 3c). The study revealed that 62.5% had an oval seed shape (Fig. 3d), 30% had a round seed shape (Fig. 3e) and 7.5% had a seed with a flat base (Fig. 3f).

Analysis of variance showing effects of accessions on qualitative traits of Bambara groundnut

Analysis of variance showed that significant differences were recorded among the Bambara groundnut accession in qualitative traits. The differences were observed in pod colour ($P<0.001$), pod texture ($P<0.001$), seed colour ($P<0.001$), seed shape ($P<0.001$), terminal leaflet length ($P<0.001$), and testa pattern ($P<0.001$) (Table 1).

Analysis of variance showing effects of accessions on quantitative traits of Bambara groundnut

Analysis of variance showed that significant differences were recorded among Bambara groundnut accessions on quantitative traits (Table 2). These differences were observed in yield ($Kgha^{-1}$), ($P<0.001$), harvest index ($P<0.05$), 100 seed weight in grams ($P<0.001$), number of leaves per plant ($P<0.001$), peduncle length ($P<0.05$), plant hieght ($P<0.05$),

Fig. 2 Percent variation in seed colour **2a** Cream, **2b** Dark red and **2c** Light red seed colour and eye pattern **2d** Eye like a butterfly, **2e** Eye as 2 thin lines on both slides of hilum and **2f** No eye

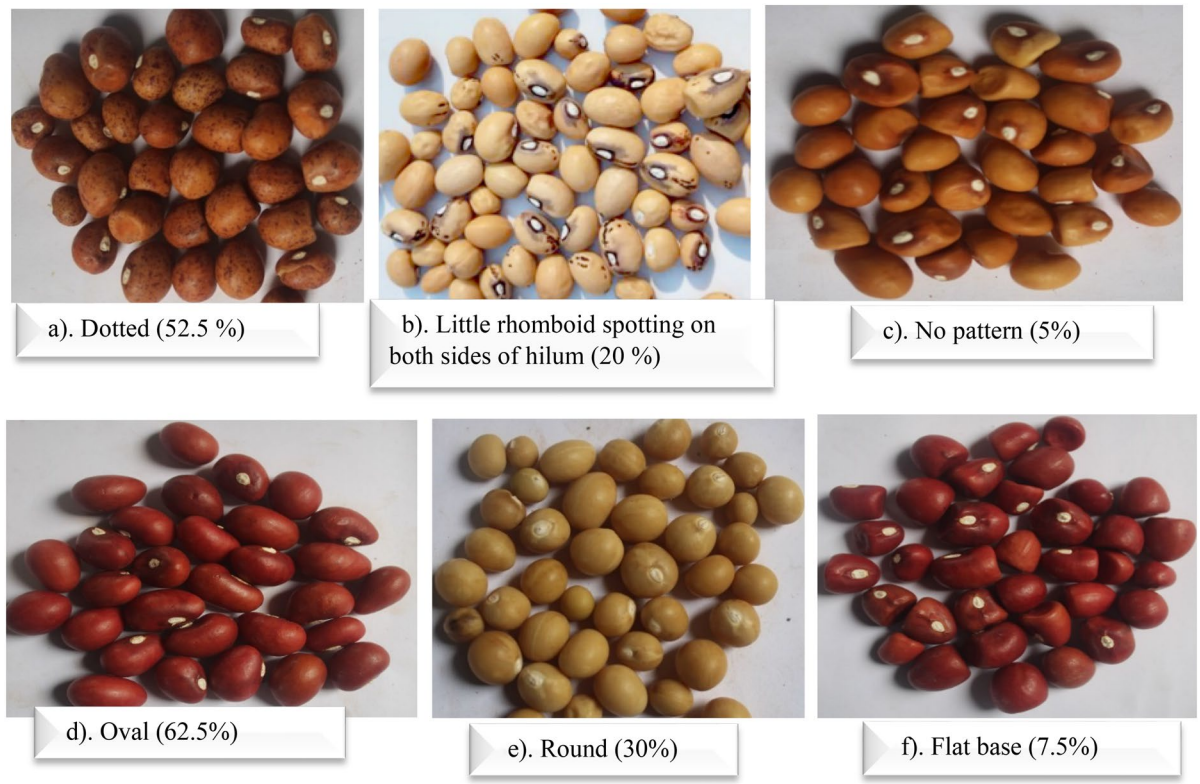

Fig. 3 Percent variation in testa pattern **3a** Dotted type, **3b** Little rhomboid spotting on both sides of hilum, **3c** No pattern, **3d** Oval seed shape, **3e** Round seed shape and **3f** Seed shape with flat base

Table 1 Analysis of variance showing effects of accessions on qualitative traits of Bambara groundnut

Source of variation	Degrees of freedom	Eye pattern	Growth habit	Pod colour	Pod texture	Seed colour	Seed shape	Terminal leaflet shape	Testa p attern
Replicates	2	0.0333	0.5333	0.325	0.3083	4.658	1.083	0.3583	1.5083
Accessions	39	0.2476 ^{ns}	0.3229 ^{ns}	3.5479*	1.4453*	5.742*	2.02*	0.9212*	6.8699*
Residual	78	0.1701	0.2684	0.5985	0.2229	1.428	8.269	0.3071	0.816

*$P < 0.001$ Significant, **$P < 0.05$ Highly Significant, ns, not significant

number of pods per plant ($P < 0.05$), petiole length ($P < 0.001$), and terminal leaflet length ($P < 0.05$).

Principal Components Analysis (PCA) of Bambara groundnut quantitative traits

The quantitative traits for Bambara groundnut were further examined to determine which traits account for the majority of the variation between accessions using the principal component analysis (PCA). The PCA revealed a total variation of 72.8% (Table 3). The first principal component (PC1) with an Eigen value of 6.791, represents a proportion of 27.2% equivalent to four individual variables (Yield ($Kgha^{-1}$), number of pods per plant, number of nodes per stem and number of stems per plant). The second principal component (PC2) with an Eigen value of 3.756, represents a cumulative proportion of 15% of total variation, been contributed only by number of days to flowering. The third principal component (PC3) had an Eigen value of 2.654 and reflects 10.6% variations exhibited significant positive effects for

Table 2 Analysis of variance showing effects of accessions on quantitative traits of Bambara groundnut

Source of variation	Degrees of freedom	Yield ($Kgha^{-1}$)	Harvest Index	Branches per stem	100 seed weight (grams)	Leaves per plant	Peduncle length (mm)	Plant height (cm)	Pod per plant	Petiole length (mm)	Terminal leaflet length (cm)
Replicates	2	590,842	0.17135	6.758	972.1	28,876	0.718	4.72	785.1	20.036	1.511
Accessions	39	691,792*	0.03074**	1.588 ns	529.3*	23,674*	3.551**	31.92**	768.1**	20.088*	1.813**
Residual	78	116,088	0.0164	1.425	122.2	9064	1.707	15.2	386	7.928	1.069

*$P < 0.001$ Significant, **$P < 0.05$ Highly Significant, ns, not significant

terminal leaflet length and internode length identifying it as the major contributor to the significant germplasm differences of agronomic traits among the Bambara groundnut accessions. Analysis of Variance revealed that there was a significant difference among the accessions in traits like petiole length (cm), number of leaves per plant, plant height (cm), peduncle length (cm), number of pods per plants, 100 seed weight and yield in kilograms per hectare.

Biplot of principal component analysis for Bambara groundnut quantitative traits.

The biplot depicts the distribution of accessions based on different Principal Component (PC) scores (PC1, PC2, and PC3), as well as the relationship between different traits and PC scores. Biplot of the 40 Bambara groundnut germplasms evaluated at Bunda during the 2020/2021 growing season is presented in Figure 4. Based on PC 1 and 2, yield ($Kgha^{-1}$), was positively correlated with 100 seed weight in grams, number of leaves per plant, and petiole length. Therefore, these traits are the standard features that can be used to determine similarities and diversity of species phenotype. Terminal leaflet width, terminal leaflet length, pods per plant and flowers per peduncle were negatively correlated with yield therefore, these traits cannot be used to identify similarities and diversity of species phenotype.

Analysis of molecular variance (AMOVA) for the 188 Bambara groundnut accessions samples genotyped

Percent variation exists within and between the Bambara groundnut accessions (Table 4). Genetic analysis revealed that there is a 13% genetic variation within the genotyped Bambara groundnut accessions and 84% genetic variation among the genotyped Bambara groundnut accessions while 3% genetic variations was observed among the ten populations.

Pairwise Nei's standard genetic distance between the ten populations of Bambara groundnut evaluated using 1048 SNP markers

Nei genetic distance reveals the closest relationship among the accessions. The smaller the distance, the closer the relationship while the larger the distance

Table 3 Principal Components Analysis (PCA) of Bambara groundnut quantitative traits

Principal components (PC)	PC1	PC2	PC3	PC4	PC5	PC6
Eigen values	6.791	3.756	2.654	1.981	1.795	1.232
Proportion of Variance	27.2	15	10.6	8	7	5
Cumulative Proportion %	27.2	42.1	52.8	60.8	68	73
Days to first flowering	−0.183	0.318	0.199	0.228	−0.176	0.138
Days to 50 flowering	−0.158	0.287	−0.143	0.247	0.133	−0.247
Flowers per peduncle	0.129	−0.358	0.195	0.2	−0.109	0.42
Petiole length (cm)	0.224	−0.242	0.174	−0.186	−0.169	−0.197
Leaves per plant	0.183	0.237	0.216	−0.105	−0.249	0.137
Terminal leaflet length(cm)	−0.219	−0.217	0.374	0.293	0.284	−0.512
Plant height (cm)	0.175	−0.264	0.149	0.208	0.116	−0.213
Peduncle length (cm)	0.196	−0.192	0.123	0.149	−0.189	−0.203
Internode length (cm)	0.12	0.127	0.652	1.148	0.109	0.205
Nodes per stem	0.424	−0.234	0.152	−0.219	0.266	−0.369
Branches per stem	0.37	−0.261	−0.413	−0.165	−0.366	0.142
Stems per plant	0.268	0.2	−0.128	0.491	0.199	0.176
Pods per plant	0.38	−0.178	−0.163	0.323	0.158	−0.147
Seeds per pods	−0.179	−0.146	−0.307	−0.432	0.171	−0.212
100Seed weight (g)	0.229	−0.133	−0.229	0.224	0.124	−0.222
Harvest index	0.224	−0.365	−0.183	−0.214	−0.179	0.151
Yield (Kgha^{-1})	0.351	0.118	−0.175	0.207	0.187	0.139

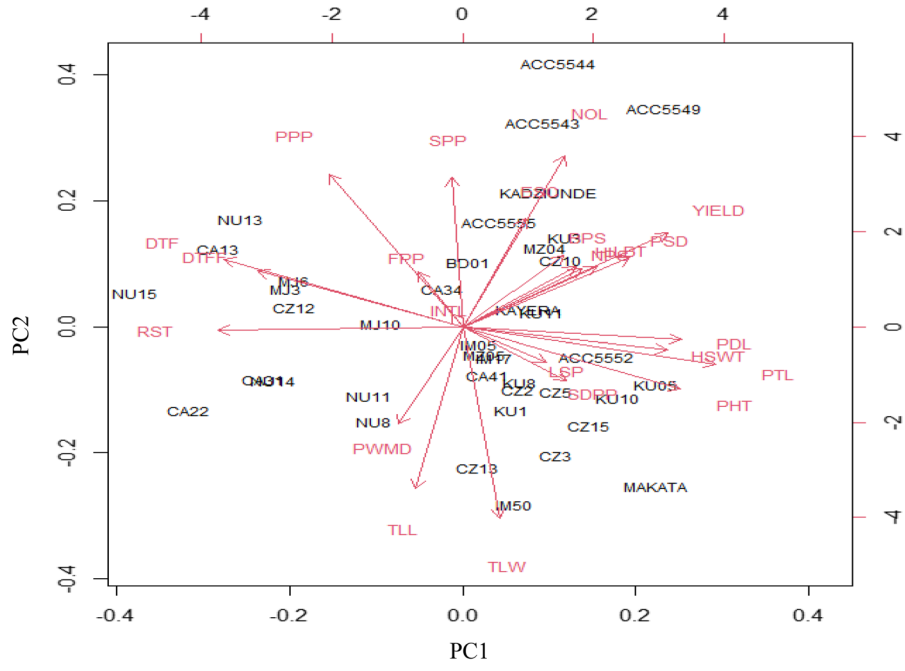

Fig. 4 Biplot of principal component for Bambara groundnut accessions quantitative traits. PPP (Pods per plant), SPP (Seeds per pod), NOL (Number of leaves), FPP (Flowers per peduncle), INTL (Internode length), DTF (Days to 50% flowering), DTFF (Days to first flower), RST (Rust), PWMD (Powderly mildew), TLL (Terminal leaflet length), TLW (Terminal leaflet width), PHT (Plant height), PLT (Petiole length), PDL (Peduncle length), HSWT (Hundred seed weight), PSD (Plant spread), LSP (Leaf shape)

the wider the relationship (Kumari & Pande 2010). The average Nei genetic distances between accessions within each population revealed Pop 5 vs Pop3 and Pop 6 vs Pop 3 had the smallest genetic distance of 0.001 (Table 5). The study also revealed that the closest distance was observed in Pop 7 vs Pop 6, Pop 8

Table 4 Analysis of Molecular Variance for the 188 Bambara groundnut accessions samples genotyped

Source	df	SS	MS	Est. Var	%
Among population	9	2751.094	305.677	3.323	3
Among accessions	178	32,851.707	184.210	84.217	84
Within accessions	188	2377.500	12.646	12.646	13
Total	375	37,980.301		101.926	100

vs Pop 3 and Pop 8 Vs Pop 5 with a genetic distance of 0.001. The largest distance was observed in Pop 7 vs Pop 2 with a genetic distance of 0.111 indicating a wider relationship among the two populations.

Discriminant analysis of principle components (DACP) of the 188 Bambara groundnut accessions samples genotyped

DAPC determined the optimal grouping of the 188 accessions at 10 clusters (Fig. 5). There was generally good correspondence between the DAPC deduced clusters and the clusters obtained by the dendrogram at K = 10. All but one DAPC cluster (cluster 3) had > 60% correspondence with the dendrogram clusters (Figures S2). This DAPC cluster was predominated by members of dendrogram cluster 5, which also predominated DAPC cluster 4, suggesting high similarity between DAPC cluster 3 and 4.

Geographical distribution of the 188 Bambara groundnut accessions based on DAPC analysis

Germplasm collected in each geographical area, revealed by genetic analysis reported that out of the 188 accession samples genotyped, NPGRC had a total of 62 accession samples followed by Chipala EPA with 32 accession samples and the least was from Kasungu with 7 accessions samples (Table 6). Cluster 1 has a total of 7 accession samples while cluster 10 had 44 accessions which is the highest number of accessions samples.

Discussion

Bambara nut qualitative trait variation

The findings revealed that there is significant morphological variability in qualitative characteristics such as terminal leaflet shape, growth habit, seed colour, eye pattern, testa pattern and seed shape. Variations on leaflet shape obtained in this study are in agreement with results reported by other researchers. For example, Beket et al. 2019 reported that the majority of the accessions had 74.26% leaflets with elliptic shape, followed by accessions with 17.82% of the leaflets had an oval shape. In a similar study, the terminal leaflet shape were reported to be 64% oval, 24% elliptic and 12% lanceolate (Valombola et.al., 2022). The study also reported that 52.5% of the Bambara groundnut germplasm had a semi bunch growth habit while 42.5% had a growth habit which is spreading and 5% had a growth habit which are bunch type. This is in agreement with the findings by

Table 5 Pairwise Nei's standard genetic distance between the nine populations of Bambara groundnut evaluated using 1048 single nucleotide polymorphism (SNP) markers

NB: Pop means Population which are the clusters based on geographical distribution

	Pop1	Pop2	Pop3	Pop4	Pop5	Pop6	Pop7	Pop8	Pop9
Pop1	–	–	–	–	–	–	–	–	–
Pop2	0.026	–	–	–	–	–	–	–	–
Pop3	0.046	0.047	–	–	–	–	–	–	–
Pop4	0.008	0.003	0.039	–	–	–	–	–	–
Pop5	0.022	0.035	0.000	0.019	–	–	–	–	–
Pop6	0.050	0.035	0.000	0.051	0.006	–	–	–	–
Pop7	0.078	0.113	0.019	0.095	0.022	0.001	–	–	–
Pop8	0.029	0.057	0.001	0.034	0.001	0.009	0.025	–	–
Pop9	0.093	0.009	0.078	0.097	0.058	0.073	0.142	0.080	–

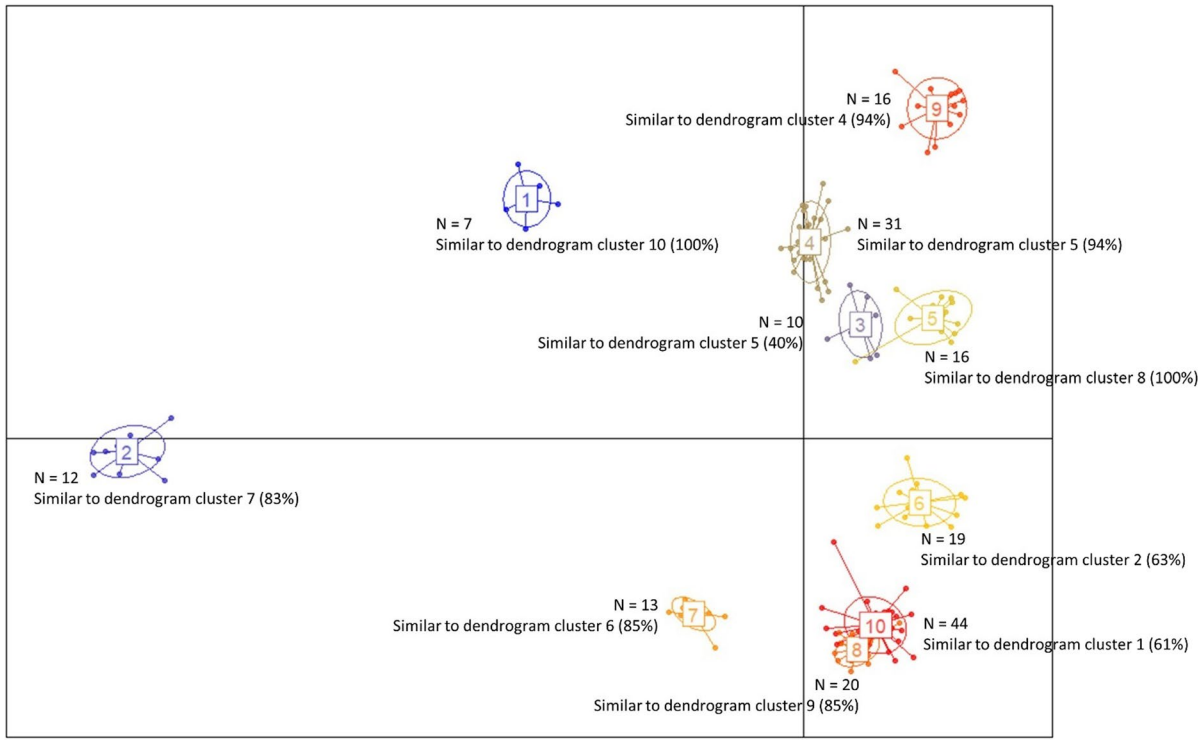

Fig. 5 Discriminant analysis of principle components (DAPC) scatter plot of the 188 Bambara groundnut accessions genotyped. The 'N indicates the number of accessions in the DAPC cluster. Comparisons were made with the dendrogram clusters and the percentages indicate the proportion of accessions in dendrogram clusters that were also in the DAPC clusters

Table 6 Geographical distribution of the 188 Bambara groundnut accessions based on discriminant analysis of principal components (DAPC)

DAPC Cluster	Total	Kandeu EPA	Manjawira EPA	Mombezi EPA	Chipala EPA	Eswazini EPA	Genebank	Mitundu	Kasungu
1	7	0	0	2	2	0	2	1	0
2	12	2	9	0	0	0	0	0	1
3	10	0	0	0	0	1	7	0	2
4	31	12	6	6	0	0	7	0	0
5	16	0	0	2	3	3	7	1	0
6	19	0	1	0	3	5	9	1	0
7	13	0	0	2	3	0	5	3	0
8	20	0	2	2	6	0	8	2	0
9	16	0	0	6	2	0	6	0	2
10	44	3	0	5	13	4	11	6	2
Total	188	17	18	25	32	13	62	14	**7**

Gbaguidi et al. 2018, in Benin who reported that there were more bunch landraces among the three types of growth habits. In a similar research, Beket et al. 2019 reported that the spreading type was the most common with 59.41% among the accessions studied, followed by the bunch type with 31.68%.

Another study reported that 44% were bunch type, 24% were semi-bunch, and 32 percent were slightly spread (Valombola et.al., 2022). That could be explained by the fact that the bunch landraces were more advantageous to the farmer particularly during harvest when the roots and stems were unearthed easily and with less pods left in the soil (Gbaguidi et al. 2018).

This study also revealed variations in seed colour among the Bambara groundnut accessions and this is in agreement with a study conducted by Judicaëlle et al. 2020 who reported that seed colour had a high degree of variability with 31.11% of the seed colour to be cream. The lowest proportions included dark brown seeds, black, dotted dark brown on a cream background, light brown marbled spots on a cream background and black rhombic spots on a cream background on the micro hair but not on the hairy ends with 1.11%. Similarly, this study revealed variations in eye patterns among Bambara groundnut and this was in tandem with the research of Valombola et al., (2022), who also reported that 64% of the accessions had no eye color, 16% had amber color, and 4% had brownish, brown, grey, black, or purple color. The study also revealed variation in testa pattern among Bambara groundnuts similar to the report of Valombola et.al., (2022). He identified 20% for the tan color, 12% for the red and cream, 8% for the dark tan and purple, and a minimum of 4% for speckled brown, cream, brown, and black for seed testa pattern. The results for testa pattern and seed colour obtained in this study are important to plant breeder if they want to develop Bambara groundnut cultivars that are resistant to diseases and storage pests. Seed colour and pattern are important traits since it has been reported that dark-colored and red coloured seeds attract fewer beetles for oviposition, indicating a higher resistance to beetle infestation. Additionally, have some amount of phenolic derivatives which have been reported to have antimicrobial activity and disease resistance (Baidoo et al. 2015). By selecting for these traits, breeders can improve Bambara groundnut cultivars, contributing to better crop quality and yield. Testa pattern and seed shape will also help in discriminating Bambara groundnut germplasm for seed production since this will help to differentiate off-types with pure seed (Khan et al. 2020).

Analysis of variance showing effects of accessions on qualitative and quantitative traits of Bambara groundnut

Analysis of variance showed that significance difference were observed in pod colour ($P<0.001$), pod texture ($P<0.001$), seed colour ($P<0.001$), seed shape ($P<0.001$), terminal leaflet length ($P<0.001$), and testa pattern ($P<0.001$). In a study conducted by Gbaguidi et al., (2018), significance differences were also observed in agro-mophological traits inclusive of yield ($Kgha^{-1}$), ($P<0.001$), harvest index ($P<0.05$), 100 seed weight in grams ($P<0.001$), number of leaves per plant ($P<0.001$), peduncle length ($P<0.05$), plant hieght ($P<0.05$), number of pods per plant ($P<0.05$), petiole length ($P<0.001$), and terminal leaflet length ($P<0.05$). traits among the studied Bambara ground nut accessions. Mohammed et al., (2020) also reported a significance difference in most of the quantitative traits of Bambara groundnuts like petiole length, seed weight and plant height. Agro-morphological traits are important since it will help to characterize Bambara groundnuts and identify desirable genotypes for seed production, breeding and conservation.

Principal components analysis for Bambara groundnut quantitative traits

Principal Component Analysis (PCA) reported a 53% total phenotypic variation for the qualitative traits among the Bambara groundnuts germplasm. The results are in agreement with the findings by Olanrewaju et al. 2021, who reported that PC1 and PC2 accounted for 42.3% of all observed variances. Khan et al. 2020 reported 45.8% variations in Bambara groundnut and some of the traits contributing to the variation were yield ($Kgha^{-1}$). In another study by Jonah et al., (2014), it was reported that a 71% of the observed variation in the first three principal components with 100 seed weight as one of the traits contributing positively to the variations while a research by Mohammed et al., (2020) reported a 79% of the total variations with seed weight and biomass contributing to the total variations among the Bambara groundnut accessions. A similar study reported that the first three axes expressed 84.01% of the total variability with yield as one of the traits that contributed positively to the total variation (Touré et al. 2012).

The current findings of PCA are important because they suggest that, in order to make the selection more effective and balanced, indices based on principal component analysis (PCAs) should be generated. This would allow for the simultaneous genetic selection of the desired characteristics for germplasm improvement in breeding programs (Vargas et al. 2018; Viana et al. 2020).

Biplot of principal component analysis for Bambara groundnut quantitative traits

PCA biplot loading for both variables and germplasm revealed how strongly each trait influences a PC and how they are related. Based on Principal Component (PC) 1 and 2, yield ($Kgha^{-1}$), was positively correlated with 100 seed weight in grams, number of leaves per plant, and petiole length. Therefore, these traits are the standard features that can be used to determine similarities and diversity of species phenotype. In other words, they show large variations across the germplasms studied, this implies that these traits are the best discriminators of the morphological traits under consideration. In a similar study, it was reported that yield and 100 seed weight were the major contributing traits and these can be used to discriminate among Bambara groundnut germplasm. Internode length and flowers per peduncle on the other hand, had the shortest vectors, indicating that they are the least discriminatory of the morphological traits as noted under our study and by Olanrewaju et al., (2021) and Valombola et al., (2019). A study by Mohammed et al., (2020) also reported that the PCA biplot revealed a significant variation among Bambara groundnut genotypes.

Analysis of molecular variance for the 188 Bambara groundnut germplasm samples

Analysis of molecular variance revealed 13% within and 84% among the Bambara groundnut germplasm genotyped while 3% were observed among the populations. This is in agreement with the phenotypic data, which revealed a total genetic variation of 72.8% among the Bambara groundnut germplasm using Principal Component Analysis. The results are also in agreement with the study by Uba et al. 2021 where the AMOVA reported that 89% of genetic variation occurred among the populations, 8% between regions, and 3% between populations. The greater the genetic variations of the germplasm, the greater the likelihood of success in breeding desirable traits. The high percentage value of genetic variation within population obtained from AMOVA could be due to natural adaptation or extensive seed exchange among farmers between collection sites, or it could be due to the population's common origin, which could have resulted in Bambara groundnut growers using the same seed continuously, without new introductions. This is also possible because in Malawi, seed sources for planting Bambara groundnut by farmers includes; farmer-saved seeds, exchange and market purchase (Pungulani et al. 2012). Therefore, due to the crop's autogamous breeding system, this is likely to result in a heterogeneous population of landraces and thus higher intra-landrace diversity, as opposed to the homogeneous population that would be expected.

Pairwise Nei's standard genetic distance between the ten populations of Bambara groundnut evaluated using 1048 SNP markers

The average Nei genetic distance between germplasm within each population revealed Pop 5 vs Pop3, Pop 6 vs Pop 3 and Pop 10 vs Pop 2 had the smallest genetic distance of 0.001. The study also revealed that the closest distance was observed in Pop 7 vs Pop 6, Pop 8 vs Pop 3 and Pop 8 Vs Pop 5 with a genetic distance of 0.001. The largest distance were observed in Pop 7 vs Pop 2 with a genetic distance of 0.111. Uba et al. 2021 reported similar findings where the closest distance based on geographical regions were reported to be 0.002.

Discriminant analysis of principle components of the 188 Bambara groundnut germplasm samples genotyped

DAPC grouped the genotypic data for 188 germplasm samples into 10 clusters. The fairly high degree of correspondence between DAPC and dendrogram analysis suggests that 9–10 clusters are reasonable for optimally grouping the accessions. The largest cluster had 44 germplasm while the smallest cluster had 7 germplasm. DAPC clustering based on geographical distribution of the germplasm revealed that most of the Bambara groundnut germplasm from the same region did not cluster accurately based on their

collection site. This could be due to low genetic differentiation between populations, implying that the genetic background of Bambara groundnut population does not always correlate with their geographical origin or region. Uba et al. 2021 reported that the grouping of some accessions from different regions into the same cluster may indicate the degree of relatedness between accessions from different regions, which can be attributed in part to the transfer and exchange of seeds between regions via gene banks and human migration.

Geographical distribution of the 188 Bambara groundnut germplasm based on DAPC analysis

This study identified the number of germplasms per geographical area as revealed by genetic analysis and it has been discovered that out of the 188 germplasm samples genotyped, NPGRC had a total of 62 germplasm samples followed by Chipala EPA with 32 germplasm samples and the least was Kasungu with 4 germplasm samples. Similar study by Olukolu et al. 2012 reported that Bambara groundnut germplasm were grouped into several clusters based on geographical region. Ontong et al., (2021) also reported similar observation on geographical distribution of Bambara groundnut accessions using simple sequence repeats (SSR) markers.

Conclusion

The study has revealed that there is a significant degree of phenotypic and genotypic variation in the Bambara groundnut germplasm in Malawi as evidenced from 78.2% phenotypic variation by PCA. And 84% genetic variation revealed by (AMOVA). The study findings indicated that the DArT SNP marker is informative and selective in identifying genetic variations therefore, it could be widely used for molecular analysis of Bambara groundnut. Based on molecular characterization of germplasm from various collection sites, the DArT SNP marker revealed that variation exists among the germplasm and that the pattern of genetic variation varied across the collection sites. There are ten major subpopulations of Bambara groundnut germplasm based on molecular characterization. Both phenotypic and genotypic characterization is important because it can accelerate variety development and deployment, particularly for neglected and underutilized crops species. This could significantly shorten breeding cycles and allow cultivars to be released more quickly. As a result, these findings will provide a significant contribution to the future variety improvement. Information on the phenotypic and genotypic variation of various Bambara groundnut germplasm will be critical for strategic production and recommendations to farmers and breeders who embark on breeding programs and seed production respectively.

Acknowledgements The authors are thankful to the McKnight foundation through its Collaborative Crop Research Program (CCRP) that funded the Seed Systems' research project (Grant no. 19-265). Thankful to all collaborators and supervisors for technical support at the time of project implementation. I would like to acknowledge Frank Tchuwa and Daimon Kambewa for their technical support during the project implementation. We would also like to acknowledge Ric Coe and the entire CCRP team for their technical support during data processing and analysis and not forgetting farmers who provided the germplasms used in this study.

Author contributions Data collection and cleaning: Louis Yalaukani. Phenotypic data analysis : Louis Yalaukani. Genotypic sample collection and processing: Louis Yalaukani and Michael Chipeta. Genotypic data analysis : Davis Gimode. Supervision : Michael Chipeta, Abel Sefasi, Wisdom Changadeya and Moses Maliro. Writing—original draft : Louis Yalaukani. Writing – Review and editing: Louis Yalaukani, Maliro Moses, Chipeta Michael, Gimode Davis, Morrone Vicki, Sefasi Abel, Pungulani Lawrent, Soko Peter.

Funding This research was supported by The McKnight foundation Collaborative Crop Research Program (CCRP) Funded Research Project on Seed Systems Grant # 19–265 (Strengthening farmer-managed seed systems for improved seed quality and access to preferred varieties in Malawi). The funders had the role in research design, research implementation, data collection and analysis, payment of genotyping services, preparation of the manuscript and decision to publish.

Data availability All relevant data are available within the manuscript and it's supporting information files in the supplementary section.

Declarations

Conflict of interest The author declares that there is no conflict of interest exist.

References

Baidoo PK, Kwansa NA, Annin CP (2015) The role of seed coat and its pigmentation on the acceptance of Bambara

groundnut (*Vigna subterranea* L. Verdc.) cultivars by the cowpea beetle, *Callosobruchus maculatus* (F.). Adv Entomology 03(04):125–131. https://doi.org/10.4236/ae.2015.34015

Beket SB, Dagou S, Koffi A, Kouamé GK, Léonie CK, Raoul SS (2019) Evaluation of the diversity in qualitative traits of Bambara groundnut germplasm (*Vigna subterranea* (L.) Verdc.) of Cte dIvoire. Afr J Biotechnol 18(1):23–36. https://doi.org/10.5897/AJB2018.16691

Castillo-Gutiérrez S, Estudillo-Martínez MD, Lozano-Aguilera ED (2021) Influence of the fitted straight line for confidence bands algorithm in Q-Q plots. Open J Statis 11(06):925–930. https://doi.org/10.4236/ojs.2021.116054

Gbaguidi AA, Dansi A, Dossou-Aminon I, Gbemavo DSJC, Orobiyi A, Sanoussi F, Yedomonhan H (2018) Agro-morphological diversity of local Bambara groundnut (*Vigna subterranea* (L.) Verdc.) collected in Benin. Genet Resour Crop Evol 65(4):1159–1171. https://doi.org/10.1007/s10722-017-0603-4

Gregory PJ, Mayes S, Hui C, Ebrahim H, Advina J, Giva J (2019) Crops For the Future (CFF): An overview of research efforts in the adoption of underutilised species. Planta 250(3):979–988. https://doi.org/10.1007/s00425-019-03179-2

Jonah PM, Abimiku OE, Adeniji OT (2014) Multivariate analysis and character association on the growth and yield of Bambara groundnut in Mubi, Adamawa State, Nigeria. Int J Manag Soc Sci Res 3(2):58–65

Judicaëlle, K. D., Hervé, N., Adjima, O., Golo, K. M. N., Nerbéwendé, S., Mahama, O., & Mahamadou, S. (2020). Agro-morphological characterization of Bambara nut accession*s [Vigna subterranea (L)* Verdcourt] from Burkina Faso. L. 15727–15744

Kassambara A, Mundt F (2018) Factoextra: extract and visualize the results of multivariate data analyses. R Package version 1.0.7. https:// cran. rproj ect. org/ web/ packages/ factoextra/ index. html

Katungwe P, Mwangwela A, Geresomo N (2015) Dietary adequacy of rural school children among bambara groundnut growing farmers in Ntchisi district of Malawi. Afr J Food Agric Nutr Dev 15(1):9620–9634

Khaliqi A, Rafii MY, Mazlan N, Jusoh M, Oladosu Y (2021) Genetic analysis and selection criteria in bambara groundnut accessions based yield performance. Agronomy 11(8):1–13. https://doi.org/10.3390/agronomy11081634

Khan MH, Rafii MY, Ramlee SI, Jusoh M, Mamun A (2020) Genetic variability, heritability, and clustering pattern exploration of Bambara groundnut (*Vigna subterranea* L. Verdc) accessions for the perfection of yield and yield-related traits. BioMed Res Int 2020(1):2195797

Kumari K, Pande A (2010) Study of genetic diversity in finger millet (*Eleusine coracana* L. Gaertn) using RAPD markers. Afr J Biotechnol 9(29):4542–4549

Lule D, Tesfaye K, Fetene M, de Villiers S (2012) Multivariate analysis for quantitative traits in finger millet (Eleusine coracana subsp coracana) population collected from eastern and Southeastern Africa: Detection for patterns of genetic diversity. Int J Agri Res 7(6):303–314. https://doi.org/10.3923/ijar.2012.303.314

Mastenbroek A, Otim G, Ntare BR (2021) Institutionalizing quality declared seed in Uganda. Agronomy. https://doi.org/10.3390/agronomy11081475

Mayes S, Basu SM, Molosiwa O, Redjeki ES, Ahmad NS, Khan F, Zehra S, Noah S, Mayes K, Roberts J, Stadler F, Massawe F, Kilian A, Azam-Ali S (2013) Molecular analysis of Bambara groundnut, an underutilised African legume crop as part of the BAMLINK project - What lessons can we learn? Acta Hort 979:451–458. https://doi.org/10.17660/ActaHortic.2013.979.48

Minnaar-Ontong A, Gerrano AS, Labuschagne MT (2021) Assessment of genetic diversity and structure of Bambara groundnut [*Vigna subterranea* (L.) verdc.] landraces in South Africa. Sci Rep 11(1):1–9. https://doi.org/10.1038/s41598-021-86977-7

MoAFS (2020). Guide to Agricultural Production and Natural Resources Management in Malawi. Agricultural Communication Branch; Department of Agriculture Extension Services, Ministry of Agriculture, Lilongwe; Malawi.,424. ISBN999–34–00–9

Mohammed MS, Shimelis HA, Laing MDD (2019) Preliminary morphological characterization and evaluation of selected bambara groundnut [*Vigna subterranea* (L.) verdc.] genotypes for yield and yield related traits. Legume Res 43(2):157–164. https://doi.org/10.18805/LR-475

Ntundu WH, Shillah SA, Marandu WYF, Christiansen JL (2006) Morphological diversity of bambara groundnut [*Vigna subterranea* (L.) Verdc.] landraces in Tanzania. Genet Resour Crop Evol 53(2):367–378. https://doi.org/10.1007/s10722-004-0580-2

Olanrewaju OS, Oyatomi O, Babalola OO, Abberton M (2021) Genetic diversity and environmental influence on growth and yield parameters of Bambara groundnut. Front Plant Sci 12:1–15. https://doi.org/10.3389/fpls.2021.796352

Olukolu BA, Mayes S, Florian Stadler N, Ng Q, Fawole I, Dominique D, Azam-Ali SN, Abbott AG, Kole C (2012) Genetic diversity in Bambara groundnut (*Vigna subterranea* (L.) Verdc.) as revealed by phenotypic descriptors and DArT marker analysis. Genet Resour Crop Evol 59(3):347–358. https://doi.org/10.1007/s10722-011-9686-5

Owere L, Tongoona P, Derera J, Wanyera N (2014) Farmers ' perceptions of finger millet production constraints. Var Prefer and Their Implic to Finger Mill Breed Uganda 6(12):126–138. https://doi.org/10.5539/jas.v6n12p126

Parmar S, Deshmukh DB, Kumar R, Manohar SS, Joshi P, Sharma V, Chaudhari S, Variath MT, Gangurde SS, Bohar R, Singam P, Varshney RK, Janila P, Pandey MK (2021) Single Seed-based high-throughput genotyping and rapid generation advancement for accelerated groundnut genetics and breeding research. Agronomy 11(6):1226. https://doi.org/10.3390/agronomy11061226

Patil HE, Modha KG (2022) Genetic diversity in finger millet [*Eleusine coracana* (L.) Gaertn] using Genetic diversity in finger millet [*Eleusine coracana* (L.) Gaertn] using principal component analysis. The Pharma Innov J 10(9):6

Pungulani L, Kadyampakeni D, Nsapato L, Kachapila M (2012) Selection of High Yielding and Farmers' Preferred Genotypes of Bambara Nut (t;Vigna subterranea (L.) Verdc) in Malawi. Am J Plant Sci 03(12):1802–1808. https://doi.org/10.4236/ajps.2012.312A221

R Core Team (2021) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https:// www.R- proje ct. org

Touré Y, Koné M, Tanoh HK, Koné D (2012) Materials and methods areas of origin and main characteristics of ten native Bambara groundnut landraces from the ivory coast. Tropicultura 2:216–221

Uba CU, Oselebe HO, Tesfaye AA, Abtew WG (2021) Genetic diversity and population structure analysis of bambara groundnut (Vigna subterrenea L) landraces using DArT SNP markers. PLoS ONE 16(7):e0253600. https://doi.org/10.1371/journal.pone.0253600

Umar ID, Kwon-ndung EH (2015) ScienceDirect Assessment of variability of finger millet (Eleusine coracana (L) Gaertn) landraces germplasm in Northern Nigeria. Niger J Genet 28(2):48–51. https://doi.org/10.1016/j.nigjg.2015.09.005

Valombola JS, Akundabweni LM, Awala SK, Hove K (2019) Agronomic and morphological diversity of Bambara groundnut (*Vigna subterranea* (L.) Verdc) accessions in North-Central Namibia. Welwitschia Int J Agri Sci 99:88–99

Valombola JS, Awala SK, Hove K (2022) Morphological characterisation of Bambara groundnut (Vigna subterranea (L.) Verdc.) germplasm collections: A basis for crop improvement. J Agri Appl Biol 1(2):100–117

G Vargas FS Schenkel LF Brito HHR Neves de DP Munari AA Boligon.R Carvalheiro 2018 Unravelling biological biotypes for growth, visual score and reproductive traits in Nellore cattle via principal component analysis Livest Sci 217 37 43 https://doi.org/10.1016/j.livsci.2018.09.010

Viana AFP, Rorato PRN, Mello FCB, Machado DS, Figueiredo AM, Bravo AP, Feltes GL (2020) Principal component analysis of breeding values for growth, reproductive and visual score traits of Nellore cattle. Livest Sci 241:104262. https://doi.org/10.1016/j.livsci.2020.104262