Euphytica (2025) 221:34
https://doi.org/10.1007/s10681-025-03478-8

RESEARCH

Diversity analysis and genome-wide association studies of seed weight trait in Bambara groundnut (*Vigna subterranea* (L.) Verdc.) using diversity array technology sequence derived single nucleotide polymorphism markers

Phyllis Muturi · Martina Kyallo · Edmore Gasura · Nasser Yao

Received: 30 July 2024 / Accepted: 22 February 2025

Abstract Bambara groundnut is an indigenous drought tolerant legume cultivated in Africa. In Kenya, the crop is grown by women and the diversity of the accessions grown is unknown. Lack of information on the crop's diversity and population structure make genetic improvement of the crop difficult. The objectives of this study were to: (i) determine the genetic diversity and population structure of 86 Bambara groundnut accessions from Kenya using 4,399 SNP markers and, (ii) identify SNPs associated with the seed weight of the Bambara groundnut collection. The DArT complexity reduction approach in combination with Illumina short-read sequencing (HiSeq 2000) was applied. Population structure analysis suggested three genetic clusters. Accessions from the same county grouped into different clusters with exception of accessions from Kilifi that distinctly fell into the same cluster. Analysis of molecular variance indicated that 0.16% of the variance was due to genetic differentiation among the populations, 19% of the variance was as a result of differentiation among individuals within populations, and 81% variation was within individuals in a population. The study suggested significant DArTseq derived SNP markers ($p < 0.05$) distributed across Bambara groundnut contigs 1, 2, 3, and 5 associated with hundred seed weight, though the identified SNPs were not validated. This study provided an important foundation by offering valuable insights into the genetic diversity of Bambara groundnut in traditional cultivation areas in Kenya. These findings can serve as a starting point for further validation and application in Bambara groundnut breeding programs.

Keywords DArT markers · Genetic diversity · Population structure · Genome wide association · *Vigna subterranea* · SNP markers

Introduction

Bambara groundnut (*Vigna subterranea* (L.) Verdc.) is an indigenous legume well adapted in semi-arid lands of Sub-Saharan Africa (SSA) (Hamadou and Martin 2024). The crop is diploid, with 11 pairs of chromosomes ($2n = 2x = 22$) and is highly inbred as reported by Massawe et al. (2016). The crop is believed to originate from West Africa, with Burkina Faso being the highest producer, followed by Cameroon, Democratic Republic of Congo, Mali, Niger,

P. Muturi (✉)
Department of Water and Agricultural Resource Management, University of Embu, Embu, Kenya
e-mail: muturi.phyllis@embuni.ac.ke

M. Kyallo · N. Yao
Biosciences Eastern and Central Africa-International Livestock Research Institute (BecA-ILRI) Hub, Nairobi, Kenya

E. Gasura
Department of Plant Science, University of Zimbabwe, Harare, Zimbabwe

Published online: 10 March 2025

and Togo (FAOSTAT 2022). The crop is mainly grown by smallholder resource-constrained farmers, mainly women (Massawe et al. 2016). Bambara groundnut landraces have been found to have higher stability and adaptation to climate variability in low-input agriculture (Chimonyo et al. 2020). High seed weight is an important trait in Bambara groundnut that is considered by both consumers and farmers thus a major target trait in Bambara groundnut breeding. Seed weight positively correlates with high grain yield, improved seed germination, and enhanced seedling vigour (Somta et al. 2015). Seed weight is a quantitative trait with average heritability of over 90% in *Vigna* spp. (Thouseem et al. 2018).

Estimates of food insecurity indicate that Sub-Saharan Africa (SSA) has the highest proportion of in the world (Xie et al. 2021). There has been a growing awareness in recent years on the potential of Bambara groundnut to contribute to increased food, nutrition, and income securities in SSA (Hamadou and Martin 2024). Climate change variability and overdependence on a few exotic staples such as common beans that often fail due to inadequate rainfall may have contributed to food and nutritional insecurity in these areas (Hamadou and Martin 2024). Climate change brings in new challenges to semi-arid lands as it further deteriorates the prevailing dry and hot conditions translating to food and nutritional insecurities (Massawe et al. 2016). The types of crops cultivated in semi-arid lands have drastically reduced and few drought-tolerant indigenous crops such as Bambara groundnut are in the hands of farmers. Bambara groundnut is a nutritious climate-smart legume that is well adapted to arid and semi-arid lands (Chimonyo et al. 2020). The crop is rich in lysine (5.8–6.5 g/100 g) which is deficient in many cereals such as sorghum which are grown in semi-arid lands (Hamadou and Martin 2024).

Bambara groundnut is mainly grown in western Kenya highlands in Bungoma, Kakamega, Busia, Vihiga counties and at the coastal lowlands (Wakhungu et al. 2017). The seeds contain approximately 42.77–60.00% carbohydrate, 15.88–25.80% protein, and 1.4–12.0% fat thus referred to as complete food (Hamadou and Martin 2024). The crop holds considerable commercial potential since the crop fetches a premium in the market compared to other legumes such as common beans. Bambara groundnut has high levels of methionine with 1.0–1.3 g/100 g than most legumes thus can be a candidate to manage malnutrition (Hamadou and Martin 2024) and serves as an important source of protein among poor people who cannot afford animal protein. Bambara groundnut is a sustainable crop because; it requires no spraying to control pests and diseases, boosts soil fertility, conditions impoverished soils, tolerates acidic soils that are highly toxic for other types of legumes, and has antagonistic effects on *Striga* weed (Massawe et al. 2016). The crop is considered as one of the orphaned and underutilized species in Kenya (Wakhungu et al. 2017). Bambara groundnut has not received attention in research compared to other legumes such as common beans thus the reason why farmers grow unimproved genotypes and most communities do not know that this superior nut exists (Khan et al. 2021). The crop is a dual purpose whose grains are consumed either as roasted or boiled and the haulms are used as animal fodder (Massawe et al. 2016). The crop is important for the small-scale farmers in rural households due to its considerable commercial potential and can raise communities from poverty, food, and nutritional insecurities (Hamadou and Martin 2024).

Knowledge of genetic relationships between Bambara groundnut is important for improvement and may help to establish a core collection as part of the germplasm collection management (Pasipanodya et al. 2022). Moreover, the utilization of landraces is valued as they can contain favourable alleles for many agronomic traits. A core collection has the maximum genetic variation in a minimum number of accessions thus economical in conservation (Pasipanodya et al. 2022). Diversity in plant genetic resources provides breeders with opportunities to develop new cultivars with end-user desirable characteristics (Yu and Chung 2021). Scanty information on genetic diversity of Bambara groundnut has made the selection, genetic improvement, and conservation of the crop difficult. The development of improved varieties with desirable traits in crops depends on the available germplasm (Pasipanodya et al. 2022). Such collections are sources of novel genes desirable for enhancing productivity and resilience in crops (Yu and Chung 2021). Generally, there are often large numbers of uncharacterized accessions in gene banks in Sub-Saharan Africa. To overcome this bottleneck, it is advisable to work with core collections since they represent the crop's genetic diversity and is more cost-effective to work with (Boczkowska et al. 2016).

Traditional breeding of Bambara groundnut has been difficult since the crop is an extreme inbreeder and flowers are formed at the soil ground level. The use of molecular markers has revolutionized the precision of plant genetic analysis and molecular breeding of crops (Sun et al. 2024). Molecular markers such as DArT based markers have been used to reveal the genetic diversity in different populations in cowpea (Ketema et al. 2020). Single nucleotide polymorphisms are the best suited for diversity studies due to; their abundance in the genomes, amenability to high-throughput detection platforms, their stability, and are very near to or even within the gene of interest (Osundare et al. 2023).

Genome wide association study (GWAS) is a powerful tool for the identification of candidate regions associated with quantitative traits whereby the approach searches the whole genome for causal genetic variation and does not require previous information on candidate genes (Uba et al. 2023). GWAS is one of several approaches developed to identify marker-trait associations taking advantage of linkage disequilibrium and the approach has been adapted in plant research (Sahu et al. 2023). GWAS can reveal association between genomic loci and advantageous traits such as seed weight (Uba et al. 2023) that are targets for crop improvement. The main benefit of GWAS is the potential to capture greater diversity by including a large number of unrelated individuals with distinct genetic background (Uffelmann et al. 2021). GWAS usually narrows down the candidate regions using natural populations (Uba et al. 2023). Genotyping by sequencing (GBS) is one of the genotyping methods used for GWAS, and GBS-GWAS approaches have been successfully applied to the identification of candidate genes controlling quantitative traits in legumes (Susmitha et al. 2023).

There has been effort to study Bambara groundnut diversity using SNPs markers and conclusion that DArT SNP marker can be widely used for molecular analysis of Bambara groundnut (Uba et al. 2021). The next-generation sequencing technology has assisted the large-scale discovery of SNPs in various plant species (Khan et al. 2021). Genotyping-by-sequencing (GBS) is a next-generation sequencing (NGS) based method that employs reduced representation to enable high-throughput genotyping of a large number of SNP markers (Pasipanodya et al. 2022). Array-based assays such as Diversity Arrays Technology (DArT) have been utilized to provide the desired ultra-high throughput and cost-effectiveness (Adu et al. 2021). These microarray-based markers have been the markers of choice and have been used for the construction of high-density maps, quantitative trait loci mapping, and genetic diversity analysis with limited time and funds (Nazarul et al. 2021). DArTseq technology offers a gain in the reduction of genome complexity through intelligent selection of genome fraction corresponding predominantly to active genes in genome characterization (Salazar-Licea et al. 2022). The technology is among the most appropriate system to discover hundreds of polymorphic genomic loci, scoring thousands of unique genomic-wide DNA fragments in a single experiment, without requiring existing DNA sequence information (Bohra et al. 2021). The DArT complexity reduction approach in combination with Illumina short-read sequencing using Hiseq2000 was applied in this study. The objectives of this study were to: (i) determine the genetic diversity and population structure of 86 Bambara groundnut accessions collected from western and Coastal Kenya using 4,399 SNP markers, and (ii) identify potential SNPs associated with seed weight of the Bambara groundnut collection.

Materials and methods

Bambara groundnut germplasm

Eighty-six Bambara groundnut accessions were collected from small holder farmers from a total of five counties in Kenya where the crop is traditionally grown. It is worth noting that the materials were local popular landraces conserved by farmers in traditional cultivation areas and may not fully represent elite and improved lines used in Bambara groundnut commercial breeding programs in the world. Four counties were from western highlands namely Busia, Kakamega, Bungoma, Vihiga, and Kilifi in the coastal lowlands (Table 1). The seeds were sorted according to the seed coat colour/patterns in combination with hilum (eye) patterns as described in the Sect. 7.4.2. of the International Plant Genetic Resources Institute Bambara groundnut descriptor. The health of the seeds was determined visually, and healthy seeds were placed in separate dispensing envelopes. The accessions from each county were assigned accession

Table 1 Characteristics of Bambara groundnut accessions used in this study

IS no	Accession code	Seed coat colour	Seed source (county)	Mean of 100 seed weight (gm)
1	KAK101	Light brown red with black specks	Kakamega	60.33
2	KAK102	Cream whitish with black specks. Hilum surrounded a winged light grey brownish mark	Kakamega	105.33
3	KAK103	Brown hilum surrounded by an all-round black layer	Kakamega	54.93
4	KAK104	Light brown with reddish vein like structures. Hilum surrounded by a dark region	Kakamega	69.93
5	KAK105	Light brown	Kakamega	58.35
6	KAK107	Dark brown beige seed with hilum surrounded by a dark region	Kakamega	72.79
7	KAK108	Brown seed with lots of black specks	Kakamega	84.73
8	KAK109	Cream white with light brown specks. Hilum surrounded by a brownish grey winged pattern	Kakamega	76.17
9	KAK110	Light brown red	Kakamega	53.87
10	KAK111	Light brown	Kakamega	42.93
11	KAK112	Dark brown red	Kakamega	74.05
12	KAK113	Dark brown with deep red vein like structures. Hilum surrounded by a very dark region	Kakamega	69.91
13	KAK114	Beige colour with hilum surrounded by a dark region	Kakamega	67.11
14	KAK115	Cream with few black specks longitudinal on the hilum. Hilum surrounded by a light greyish brown wing	Kakamega	83.83
15	KAK116	Light brown. Hilum surrounded by a thin black layer	Kakamega	51.99
16	KAK117	Dark purple	Kakamega	60.05
17	KAK118	Black	Kakamega	55.23
18	KAK119	Cream with a few black specks longitudinal to the hilum	Kakamega	79.07
19	KAK120	Black brown	Kakamega	54.93
20	KAK121	Cream. Hilum surrounded by a brown winged pattern	Kakamega	84.85
21	KAK122	Dark red	Kakamega	46.57
22	KAK124	Cream with long black specks on one side. Hilum surrounded by a black thin layer with light brown winged region	Kakamega	84.74
23	KAK125	Black	Kakamega	59.29
24	KAK126	Dark brown with hilum surrounded by a dark black region that supports the veins light markings	Kakamega	71.01
25	KAK127	Light brown with some veins like structures	Kakamega	49.59
26	BGM201	Light brown	Bungoma	74.01
27	BGM202	Dark brown	Bungoma	59.53
28	BGM203	Cream with hilum surrounded by a thin dark layer and a light brown winged region	Bungoma	91.51
29	BGM204	Cream with light brown specks. Hilum surrounded by a dark brown winged pattern	Bungoma	89.5
30	BGM205	Light brown	Bungoma	56.82
31	BGM206	Greyish with many black small specks	Bungoma	99.63
32	BGM207	Dark brown	Bungoma	133.17
33	BGM208	Dark brown	Bungoma	85.88
34	BGM209	Black	Bungoma	48.18
35	BGM210	Cream seed with brown small specks. Hilum surrounded by a dark brown winged pattern	Bungoma	92.56

Table 1 (continued)

IS no	Accession code	Seed coat colour	Seed source (county)	Mean of 100 seed weight (gm)
36	BGM211	Cream with dark specks. Hilum surrounded by dark brown winged pattern	Bungoma	75.93
37	BGM212	Dark brown	Bungoma	87.43
38	BGM213	Cream. Hilum surrounded by a greyish brown winged pattern	Bungoma	77.15
39	BGM214	Black	Bungoma	47.67
40	BGM215	Light red	Bungoma	47.67
41	BGM216	Cream with long black specks running longitudinal on the hilum	Bungoma	78.73
42	VHG301	Cream with hilum surrounded by a black brown winged pattern	Vihiga	64.25
43	VHG302	Light brown red with black specks	Vihiga	81.31
44	VHG303	Cream	Vihiga	79.99
45	VHG304	Dark red brown	Vihiga	75.69
46	VHG305	Dark red	Vihiga	59.69
47	VHG306	Dark red	Vihiga	58.45
48	VHG307	Cream with hilum surrounded by a brown winged pattern	Vihiga	58.75
49	VHG308	Light brown	Vihiga	48.73
50	VHG309	Light brown	Vihiga	63.81
51	VHG310	Light brown red with many small black specks	Vihiga	68.39
52	VHG311	Light brown	Vihiga	53.88
53	VHG312	Red	Vihiga	55.13
54	VHG313	Brown with hilum surrounded by a black winged pattern	Vihiga	52.67
55	VHG314	Light red	Vihiga	49.01
56	KLF401	Red brown	Kilifi	49.01
57	KLF402	Light brown	Kilifi	93.91
58	KLF403	Light brown	Kilifi	62.96
59	KLF404	Dark brown	Kilifi	66.68
60	KLF405	Light brown red greyish with many small black specks	Kilifi	68.74
61	KLF406	Cream with hilum surrounded by maroon winged pattern	Kilifi	96.11
62	KLF407	Cream with hilum surrounded by a black winged pattern	Kilifi	77.29
63	KLF408	Cream with long black specks. Hilum surrounded by a greyish brown winged region	Kilifi	62.39
64	KLF409	Black	Kilifi	65.37
65	KLF411	Cream	Kilifi	89.43
66	KLF412	Cream with hilum surrounded by a brown winged region	Kilifi	115.07
67	KLF413	Red	Kilifi	83.83
68	KLF414	light brown	Kilifi	90.93
69	KLF415	cream with hilum surrounded by greyish winged region	Kilifi	97.31
70	BSA501	Dark purple with vein like structures and hilum surrounded by a black region	Busia	62.84
71	BSA502	Black	Busia	53.55
72	BSA504	dark brown	Busia	54.16
73	BSA505	Light brown	Busia	60.92
74	BSA506	Light brown with many black specks	Busia	96.65
75	BSA508	Light brown	Busia	54.25
76	BSA509	Light brown	Busia	58.00
77	BSA510	Light brown with hilum surrounded by a black region	Busia	78.29

Table 1 (continued)

IS no	Accession code	Seed coat colour	Seed source (county)	Mean of 100 seed weight (gm)
78	BSA511	Red	Busia	60.63
79	BSA513	Red	Busia	70.9
80	BSA514	Black	Busia	69.00
81	BSA515	Light brown with hilum surrounded by a black region	Busia	56.33
82	BSA517	Light brown	Busia	55.92
83	BSA518	Dark brown with dark red veins. Hilum surrounded by a black region supporting the dark veins	Busia	73.27
84	BSA519	Light brown with dark red veins with hilum surrounded by a black region	Busia	66.33
85	BSA521	Light brown with dark red veins with hilum surrounded by a dark region	Busia	61.53
86	BSA522	Black	Busia	57.91

codes for identification purposes. The geographical variation of the five counties is shown in Table 2.

DNA extraction

Three seed samples of Bambara groundnut from each of the 86 accessions were grown in pots in the screen house at the BecA-ILRI Hub in Nairobi, Kenya for DNA extraction. Total DNA was extracted from approximately 30 mg of leaves from three weeks old Bambara groundnut seedlings using the Quick-DNA Plant/Seed kit (Zymo Research, USA) following the manufacturer's protocol. The integrity of DNA was checked using agarose gel electrophoresis while the concentration of DNA was estimated using the NanoDrop 1000 spectrophotometer (NanoDrop Technologies, Thermo Scientific, Wilmington, DE, USA). The extracted DNA was normalized to 50 ng/µl with triple distilled water and stored at −20 °C before use. DArTseq was done by the Integrated Genotyping Support and Service (IGSS) unit hosted at BecA-ILRI Hub.

DArTseq SNP data scoring and filtering

Since Bambara groundnut and cowpea (*Vigna unguiculata*) are close relatives, the genome of the latter was used as the reference genome for SNP mining. A total of 8,798 SNPs was generated before filtering. DArTseq detected silico DArT markers are dominant and were scored in a binary fashion where 1 represented presence and 0 represented absence of a restriction fragment in the genomic representation of each Bambara groundnut sample (Kilian et al. 2012). The markers were screened according to call rate range from 0.50 to 1.00, polymorphism information content (PIC) range from 0.23 to 0.5, marker reducibility, missing SNP data, and minor allele frequency according to Valdisser et al., (2017). A total of 4399

Table 2 Geographical information of counties sampled in this study

County	Elevation (meters above sea level (MASL)	Ecology	Latitude	Longitude	No. samples collected
Kilifi	30	Coastal lowland	3.5107°S	39.9093°E	14
Vihiga	1200	Western highland	0.0816°N	34.7229°E	14
Busia	1227	Western highland	0.4608°N	34.1115°E	17
Bungoma	1385	Western highland	0.5695°N	34.5584°E	16
Kakamega	1535	Western highland	0.2842°N	34.7523°E	25

Source: Jaetzold and Schmidt (1982)

DArTseq informative SNPs were considered for the current study after quality control.

Population structure analysis

Bayesian population structure analysis was done using Structure software version 2.3.3 (Falush et al. 2003). A Bayesian clustering approach was used to estimate the number of subpopulations (K) using 100,000 Markov Chain Monte Carlo (MCMC) simulations with ten replicates for each K-value from 1 to 10 with a burn-in length of 10,000. The admixture model and correlated allele frequencies were applied (Pritchard et al. 2003). The online STRUCTURE HARVESTER application (http://taylor0.biology.ucla.edu/structureHarvester/) was used to analyse the structure results. The two model choice criterions used to detect the most probable value of K were the LnP(D) value for each given K and ΔK, an ad hoc quantity related to the second-order change of the log probability of data with respect to the number of clusters inferred by Structure (Evanno et al. 2005). Population differentiation due to genetic structure was assessed using a Neighbour Joining (NJ) tree method and Principal Component Analysis (PCoA) generated by R statistical software (R Core Team 2021).

Genetic diversity and relationships analysis

Genetic distances among the evaluated Bambara groundnut genotypes were calculated from the proportion of shared alleles obtained from DArTseq SNPs by using Euclidean genetic distance coefficients according to Nadeem et al. (2018). The genotypic data was subjected to various within and among groups genetic diversity measures using GenAlEx version 6.5 software (Peakall and Smouse 2012). Genetic diversity parameters namely, number of different alleles (Na), number of effective alleles (Ne), Shannon's information index (I), observed heterozygosity (Ho), expected heterozygosity (He), and unbiased expected heterozygosity (uHe) were calculated using the same software (Fatokun et al. 2018). Total genetic variation was partitioned into within and among populations through Analysis of Molecular Variance (AMOVA) using GenAlEx genetic software (Fatokun et al. 2018). Genetic distances between each pair of accessions and between pre-grouped populations were measured based on Nei's genetic distance using the same software (Nei et al. 1983). Genetic distance matrices for each locus were summed across loci assuming statistical independence (Nei et al. 1983). Pair-wise genetic frequency-based dissimilarity or distance matrix between individuals was calculated according to Nei et al. (1983). The resulting dissimilarity matrix was subjected to tree construction using the unweighted pair group method analysis (UPGMA) using R software (R Core Team 2021).

Genome-wide association for Bambara groundnut seed weight

Seed weight was considered for genome-wide association studies (GWAS) using 4,399 DArTseq SNP-derived markers. Association analysis was performed with R software using Mixed Linear Model (MLM) method (R Core Team 2021). Manhattan plots for seed weight were generated in GWAS and showed the most significant associations ($-\log$ (p-value) > 2). Marker alleles with p-values ≤ 0.001 in MLM model were declared significantly associated with 100 seed weight in Bambara groundnut.

Results

Means of 100 Bambara groundnut seed weight

The mean weight of 100 Bambara groundnut seeds (in grams) tested in this study is presented in Table 1. The grand 100 seed weight mean was 69.73 gm. Accession BGM207 from Bungoma scored the highest 100 seed weight of 133.17 gm while KAK111 accession from Kakamega scored the least seed weight of 42.93 gm. As mentioned earlier, the seeds were popular local landraces conserved and utilised by smallholder farmers in traditional Bambara groundnut cultivation areas in Kenya and may not fully represent elite or improved lines used in commercial breeding programs.

Population structure

The results of population structure of the 86 Bambara groundnut genotypes from the five counties sampled was inferred using STRUCTURE 2.3.4. The most suitable K value for determining the genetic cluster was found to be K = 3 suggesting the presence of

three main populations. The three genetic clusters were symbolised by the three colours namely; blue, green and red as shown in the Bambara groundnut panel (Fig. 1). Each colour represents a different gene pool. The mixing of colours in the STRUCTURE plot may indicate weak genetic differentiation among populations. The bar plot represents each accession as a single vertical bar broken into K colour segments, with lengths proportional to the estimated probability of membership in each inferred cluster. The population structure diagram indicated varying levels of genetic admixture across the five sampled counties.

Phylogenetic analysis

A cluster dendrogram was constructed to visualize the genetic relationships among populations using 4,399 DArTseq markers and results are shown in Fig. 2. The resulting dendrogram illustrated the genetic similarity and divergence among individuals/

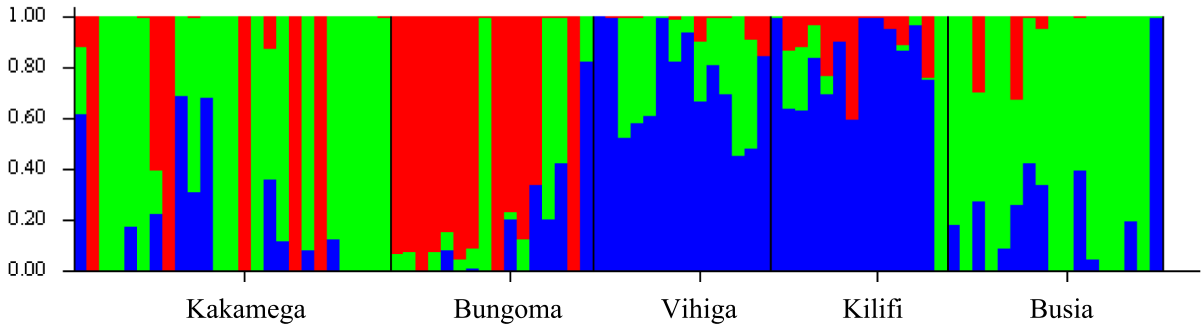

Fig. 1 Distribution pattern of 86 Bambara groundnut accessions based on 4,399 DArTseq-derived SNP marker

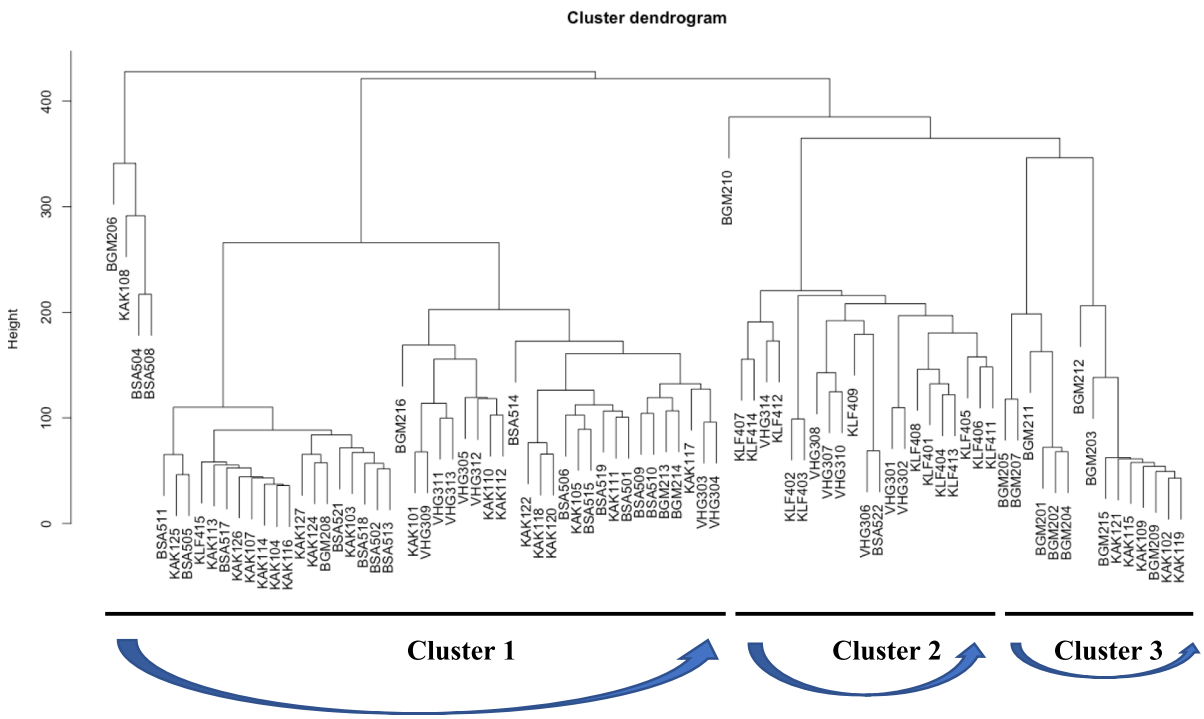

Fig. 2 Hierarchical clustering of 86 Bambara groundnut accessions based on 4,399 DArTseq-derived SNP marker

populations. The dendrogram suggested three major clusters indicating distinct genetic groupings among the samples. Bambara groundnut populations in clusters 2 and 3 grouped closely, suggesting high genetic similarity while population in cluster 1 formed a distinct cluster indicating genetic divergence. Cluster 1 included populations from all the counties sampled except Kilifi suggesting genetic admixture and possible shared ancestry. Interestingly, cluster 2 comprised populations from Kilifi (coastal lowland) and Vihiga (western highland), indicating possibility of adaptation to different environments, historical gene flow or shared ancestry. Within cluster 1 and 2, there were subgroups distinct from others, potentially suggesting recent isolation. Cluster 3 had populations exclusively from Bungoma and Kakamega (western highlands). The grouping of the three clusters in the dendrogram supported the findings from the STRUCTURE analysis, which identified shared ancestry between these populations.

Principal component analysis

The results of principal component analysis (PCoA) is presented in Fig. 3. The first principal coordinate (PCA1) accounted for 25.27% of the variation observed while the second principal coordinate (PCA2) accounted for 16.41% accounting for a cumulative 41.68 of the variance. The primary source of this variation could be attributed to geographic and environmental differences between western highlands and coastal lowland. Bambara groundnut populations grouped into three clusters in the PCoA plot. Cluster 1 and cluster 3 consisted of populations from western highlands, while cluster 2 included populations from both the coastal lowland and western highlands. The clustering pattern observed in the PCoA was consistent with results from the STRUCTURE analysis, which identified three genetic clusters.

Genetic diversity

Gene diversity, polymorphism information content, and heterozygosity were determined and shown in Table 3. The mean polymorphism information content (PIC) values for each SNP locus in Bambara groundnut accessions from Bungoma, Vihiga, Kilifi, Busia, and Kakamega were 77.13%, 57.58%, 59.90%, 69.99%, and 38.92%, respectively. The mean number of different alleles for each population was 1.77, 1.57, 1.60, 1.70, and 1.38, respectively. Bambara groundnut population from Busia had the highest number

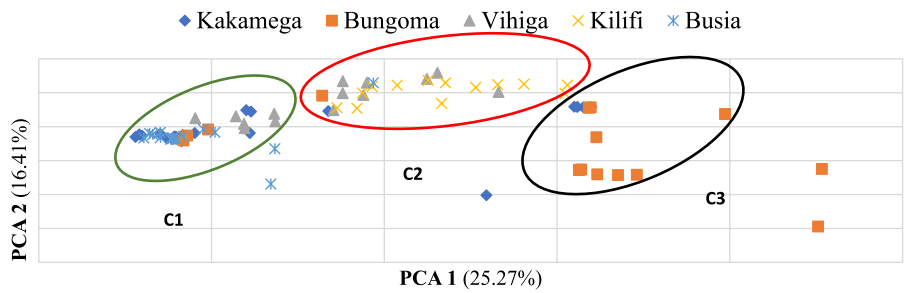

Fig. 3 Principal component analysis (PCoA) based on 4,399 DArTseq-derived SNP markers

Table 3 Estimation of population diversity indices in the 86 Bambara groundnut accessions studied

Population	(N)	(Na)	(Ne)	(P %)	(Ho)	(He)	(uHe)	(I)
Bungoma	16	1.77	1.27	77.13%	0.19	0.17	0.17	0.28
Vihiga	14	1.57	1.3	57.58%	0.21	0.18	0.19	0.28
Kilifi	14	1.6	1.2	59.90%	0.15	0.14	0.14	0.23
Busia	17	1.7	1.33	69.99%	0.24	0.2	0.21	0.32
Kakamega	25	1.38	1.18	38.92%	0.15	0.12	0.13	0.18

N, No. of samples analysed; Na, Number of different alleles; Ne, number of effective alleles; P%, Percentage of polymorphic loci; Ho, Observed Heterozygosity; He, expected heterozygosity; uHe, Unbiased expected heterozygosity; I, Shannon's Information index

of effective alleles (1.33), heterozygosity (0.21), and Shannon Index (0.32). Nevertheless, populations from Bungoma showed the highest level of different alleles (1.77) and the highest percentage of polymorphic loci (77.13%). Populations from Kakamega showed the lowest level of all the measures of genetic diversity measured in this study. The genetic diversity measure of populations from Bungoma and Vihiga were similar in this study (0.28).

Genetic distance among populations

The genetic distance among the different Bambara groundnut populations was estimated with 4,399 DArTseq derived SNP markers and the results are shown in Table 4. The greatest genetic distance was observed between populations from Vihiga and Kilifi (0.045), Vihiga and Kakamega (0.045), and Vihiga and Busia (0.042) in that order. The least genetic distance was observed between populations from Bungoma and Vihiga (0.021), Kilifi and Busia (0.021), Kilifi and Kakamega (0.021 and Bungoma and Kakamega (0.023) in that order.

Analysis of molecular variance

The Analysis of Molecular Variance (AMOVA) was conducted to evaluate the distribution of genetic variation among and within the 86 Bambara groundnut accessions. The results of the analysis is shown in Table 5. The AMOVA was calculated using R software. It was observed that 0.16% of the variance was due to genetic differentiation among the populations and 19% of the variance was as a result of genetic differentiation among individuals within populations, and 81% variation was within individuals in a population. These findings suggest the predominance of genetic diversity within accessions compared to divergence among them. The mean fixation index (FST) was estimated at 0.063, suggesting low genetic differentiation among the accessions.

Table 4 Nei genetic distance between populations

Populations	Bungoma	Vihiga	Kilifi	Busia	Kakamega
Bungoma	0				
Vihiga	0.021	0			
Kilifi	0.027	0.045	0		
Busia	0.035	0.042	0.021	0	
Kakamega	0.023	0.045	0.021	0.036	0

Genome-wide association study on seed weight in Bambara groundnut

A genome-wide association study (GWAS) was performed to identify genomic regions associated with 100 seed weight in Bambara groundnut, utilizing 4,399 DArTseq SNP-derived markers. Manhattan plots for seed weight was generated in GWAS indicating the most significant associations ($-\log(p\text{-value}) > 2$) (Fig. 4). A quantile–quantile (Q-Q) plot confirmed a normal distribution of the trait and the model effectively controlled false positives (Fig. 4).

Association mapping was performed using R statistical software using Genome Association and Prediction Integrated Tool (GAPIT) as described by Lipka et al., (2012) (Table 6). The significant SNPs, their allelic effects, and their respective minor allele frequencies (MAF) in the population are shown in Table 6. *p* values were used to determine the significant association of the trait with markers (Table 6). The results suggested nine significant ($p < 0.001$) associations were detected for grain seed weight on contigs 1, 2, 3, and 5 as shown in Table 6 contributing to 10–15% of phenotypic variance. These loci

Table 5 Analysis of molecular variance based on 4399 DArTseq SNP derived markers

Source of variation	Degrees of freedom	Sum of squares	Mean squares	Estimated variance	Variation (%)
Among populations	4	8976.097	2244.024	100.370	0.16
Among individuals	4	16,759.03	4189.7579	201.8491	19
Within individuals	79	67,216.13	850.8371	850.8371	81
Total	87	92,951.257	7284.619	1153.056	100

Mean fixation Index (Fst): 0.063

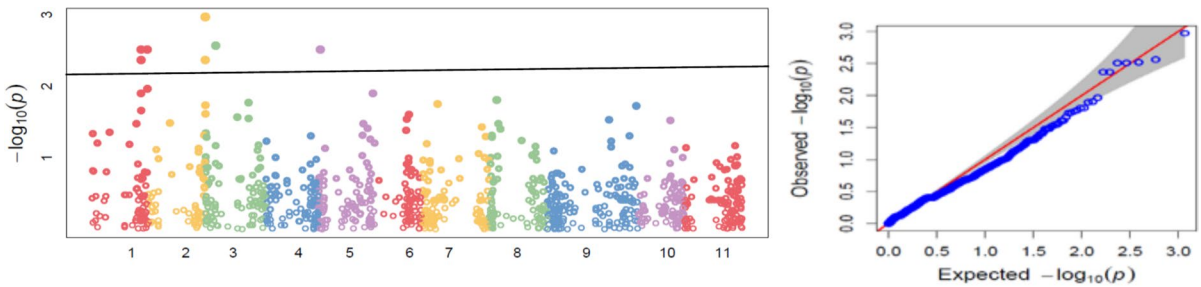

Fig. 4 Manhattan plot and respective quantile–quantile (Q-Q) plot of hundred seed weight

Table 6 Genome wide associations of single nucleotide polymorphisms (SNPs) with seed weight

Sn	SNP	Contig	Position	*Allele	*p* value	MAF	R^2 with SNP	Allelic effect estimate
1	100,003,989	1	36,158,329	G/C	0.004	0.212	0.12	7.05
2	100,006,122	1	36,148,794	A/T	0.003	0.136	0.128	-3.63
3	**100,014,020**	**1**	**4.1E + 07**	G/T	0.003	0.011	0.127	3.79
4	**100,014,021**	**1**	**4.1E + 07**	G/A	0.011	0.071	0.1	-4.64
5	4,182,577	2	40,978,951	T/C	0.001	0.293	0.152	2.16
6	4,184,000	2	40,967,085	T/C	0.004	0.31	0.12	-0.36
7	4,184,163	3	7,342,052	A/G	0.003	0.223	0.13	-6.85
8	4,182,351	5	39,550,886	C/G	0.013	0.299	0.097	6.85
9	1E + 08	5	1,657,936	C/G	0.003	0.011	0.127	-8.72

SNP Single nucleotide polymorphism, *Allele corresponding to seed weight based on 86 Bambara groundnut genotypes studied in this study, *MAF* minor allele frequency

collectively could contribute to the genetic control of hundred seed weight, highlighting the polygenic nature of this trait. These findings suggest that hundred seed weight in Bambara groundnut is influenced by loci distributed across multiple genomic regions. These markers ranged in minor allele frequency from 0.01 to 0.299. Two SNP markers bolded in Table 6 overlapped on the same position on contig 1. While functional annotations were not directly applied to the significant SNPs in this study, the regions surrounding the significant loci are likely to harbour genes involved in seed development. Further validation is necessary to identify specific genes linked to seed weight.

Discussion

Genetic diversity and relationship

Results of the diversity analysis indicated that Bambara groundnut accessions from Bungoma had the highest polymorphic information content (77%) (Table 3). The values were high in Busia and intermediary for Vihiga, Kilifi, and Kakamega. These results suggested that most of the Bambara groundnut accessions cultivated by farmers in western highlands and Coastal lowlands of Kenya possibly originated from one source (Bungoma). The mixture of accessions in the three clusters implied that the accessions formed a heterogenic, uneven group with variable genetic backgrounds. This suggested that farmers might have exchanged germplasm across the sampled regions and interbreeding of the various accessions might have occurred over time. This observation is supported by the fact that Bambara groundnut originates from West Africa and was brought to East Africa by Bantus as they were moving south. The highest genetic distance was observed to be similar between populations from Vihiga and Kilifi and Vihiga and Kakamega (0.045) respectively (Table 4). This implies that though Vihiga is geographically distant from Kilifi and Kakamega, Bambara groundnut germplasm sampled from these areas are genetically similar. This

could be explained by the fact that Bambara groundnut is grown by farmers as a minor crop and since it is indigenous, the same germplasm must have been exchanged by farmers in the early years and even today during barter trade, cultural functions such as dowry payment, trade among other events. The least genetic distance of 0.021 was observed between populations from Bungoma and Vihiga; Kilifi and Busia; and Kilifi and Kakamega, respectively, implying a close genetic relatedness between the populations. It is worth noting that the Bambara ground nut accessions s used in this study were collected from farmers who use their farm generated and saved seeds using informal seed system. Bambara groundnut breeding is young in Kenya and there is no improved variety developed by plant breeders thus limited genetic diversity of the crop.

Population structure and analysis of molecular variance

Information about the structure of accessions collections informs conservation and utilization of genetic resources. STRUCTURE analysis, hierarchical clustering and principal component analysis (PCoA) were used to infer the population structure of the collected Bambara groundnut accessions. These analyses methods consistently gave similar results of three clusters. It was observed that accessions sampled from Kilifi distinctly clustered together. This implied that the germplasm sampled from Kenya western highlands and Coastal lowlands showed a high level of homogeneity and evenness. A low mean fixation index (Fst) estimate value of 0.063 and a small percentage variation (6%) among populations as revealed by analysis of molecular variance were recorded (Table 5). The low mean fixation index suggested a low degree of differentiation among populations and increased levels of admixtures. This could be attributed to the likelihood that alleles in the studied Bambara groundnut population were related owing to the fact that the seeds evaluated in this study were unimproved accessions conserved by farmers in traditional cultivation areas. The low degree of differentiation among sampled populations suggested the need for widening the genetic base. As mentioned earlier, this lower level of variation among accessions could be attributed to germplasm exchange among farmers in the sampled counties coupled by limited introduction of new accessions since the crop is traditionally cultivated in the sampled counties. According to Wright (1977), FST of 0.00–0.05 indicates low differentiation, 0.05–0.15 indicates moderate differentiation and 0.15–0.25 high levels of differentiation, while an FST > 0.25 indicates a very high level of differentiation. Broadening of the genetic base could be achieved by the introduction of landraces from other countries such as West Africa which is the primary centre of origin of the crop. The low mean fixation index could be attributed to the fact that Bambara groundnut is cleistogamous and highly inbreeding resulting in high percentages of selfing which tend to maintain genetic diversity within individual populations as reported by several authors (Kuan et al. 2017; Mayes et al. 2019; Onwubiko et al. 2011). The current findings suggest that the accessions retained a rich genetic base within populations, making them valuable for breeding and conservation efforts, although the low differentiation among accessions may limit opportunities for creating highly divergent Bambara groundnut hybrids.

The low level of variation observed in this study is supported by other studies. For example, Odongo et al. (2015) who studied genetic diversity in Bambara groundnut using SSR markers reported that the highest proportion of the total variation (98%) was among individuals within accessions, and variation among accessions accounted for less than 2%. A study conducted by Ntundu et al. (2004) on analysis of genetic diversity in Bambara groundnut landraces in Tanzania using amplified fragment length polymorphism (AFLP) markers reported that genetic distances between all accessions based on Jaccard's variability index ranged from 0.1 to 0.68, with a total average of 0.3. Pasquet et al. (1999) used isozyme markers and reported that both wild and domesticated Bambara groundnut accessions were characterized by low genetic diversity, indicating that wild Bambara groundnut is the progenitor of the domesticated type. However, it is worth noting that isozymes are generally limited by the low levels of polymorphism detectable and may fail to discriminate cultivars differing only slightly in the genetic make-up (Pasquet et al. 1999; Amadou et al. (2001); ; ; also reported considerable genetic diversity among 25 African Bambara groundnut accessions from International Institute for Tropical Agriculture (IITA), Ibadan, Nigeria, using Random Amplified Polymorphic DNA

(RAPD) markers, and demonstrated two main groups of accessions mainly along the lines of their geographic origin.

Genome wide association studies

Genome wide association studies (GWAS) was conducted to detect regions of the scaffolds associated with hundred seed weight trait in Bambara groundnut. The MLM model in GAPIT was selected for the analysis because the model takes into account both population structure and relative kinship—there is a significant reduction of false positives compared with the general linear model (GLM) that considers population structure only and doesn't take into account population structure or relative kinship (Kamfwa et al. 2015). The Manhattan plots for hundred seed weight were generated in genome wide association study (GWAS) indicating the most significant associations (Table 6 and Fig. 4). The identification of nine significant loci distributed across contigs 1, 2, 3, and 5 underscores the polygenic nature of seed weight trait in Bambara groundnut. This observation is consistent with previous findings in legumes, where seed weight is controlled by multiple loci of small to moderate effect. The significant loci identified in this study are likely linked to key biological pathways regulating seed development and growth.

The amount of genetic variance explained by these loci ranged from 10 to 15%. This indicated that these associations had significant effects and the genomic regions associated with seed weight in Bambara groundnut need to be investigated further. The R^2 values reported in this study that ranged from 10 to 15% are consistent with the genetic complexity of traits such as legume seed yield traits controlled by several genes with small but cumulative effect as described by Kamfwa et al., (2015). Nine significant ($p < 0.001$) associations were detected for grain seed weight on contigs 1, 2, 3, and 5. This observation is supported by Somta et al. (2015) who reported that QTLs on contigs 2 and 5 were associated with seed weight in mung bean *Vigna radiate* (L.) Wilczek, Azuki bean (*V. angularis* (Ohwi) Ohwi and Ohashi) and rice bean (*V. umbellata* (Thunb.) Ohwi and Ohashi) suggesting that the region of the chromosome conditioning seed weight in *Vigna* species is conserved. The significant SNPs identified in this study provide a valuable resource for Bambara groundnut breeding programs. These markers can be used to develop marker-assisted selection (MAS) strategies to improve seed weight. However, the practical application of these SNPs requires further validation to confirm their robustness and relevance across diverse Bambara groundnut populations and environmental conditions. Previous GWAS reports on seed weight in common bean had reported that seed weight is controlled by regions in contigs 2, 3, 6, 7, 8, and 11 (Wen et al. 2019).

Limitations of the current study

A key limitation of this study is the lack of experimental or bioinformatics validation for the identified candidate SNPs and contigs, which was not conducted due to resource constraints. The lack of validation may introduce some uncertainty regarding the detected variations, potentially limiting their immediate application in downstream applications such as marker-assisted selection or functional genomics. Furthermore, the seeds evaluated in this study were unimproved accessions conserved by farmers in traditional cultivation areas in Kenya, which, while rich in within population genetic diversity, may not fully represent elite or improved germplasm used in commercial breeding programs in the world. Despite these limitations, the study makes a significant contribution by exploring the genetic diversity present in farmer-conserved accessions, offering a foundational dataset that can guide future research and validation efforts. This work provides a stepping stone for understanding and utilizing the untapped genetic potential of local Bambara groundnut germplasm for crop improvement.

Conclusion and recommendation

Results of the present study suggested that DArTseq derived SNP markers can be used for genotyping of indigenous species such as Bambara groundnut whose sequence information is not yet available. The study suggested that Bambara groundnut accessions used constituted of three gene pools. The study indicated low genetic variation among the populations but a high genetic variability among individuals and within population that can be exploited by breeders to improve Bambara groundnut. The study suggested candidate SNPs associated with seed weight on contigs 1, 2, 3, and 5 providing a foundation for

future research and breeding efforts. While the findings highlight the genetic basis of seed weight, further experimental and bioinformatics validation is necessary to translate these associations into practical applications.

Future research

To build upon the findings of this study, future work could focus on the following: conducting experimental validation of significant SNPs identified in this study using techniques such as CRISPR-based functional studies; using bioinformatics tools to annotate significant loci and identify candidate genes with putative roles in hundred seed weight; and incorporating multi-environment phenotypic experiments to identify stable associations and account for environmental effects on hundred seed weight in Bambara groundnut.

Authors' contributions PM conceived, conceptualized the idea of the research, and wrote the manuscript, MK and EG did data analyses, MK assisted in running the laboratory experiments and NY supervised the work. All authors reviewed the manuscript.

Funding The laboratory work was supported by the BecA-ILRI Hub through the Africa Biosciences Challenge Fund (ABCF) program. Farmers from the sampled counties where the accessions were collected are appreciated.

Data availability Most of the data used to support the findings of this study are included in the article. Additional data are available from the corresponding author upon request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Adu BG, Akromah R, Amoah S, Nyadanu D, Yeboah A, Aboagye LM et al (2021) High-density DArT-based SilicoDArT and SNP markers for genetic diversity and population structure studies in cassava (*Manihot esculenta* Crantz). PLoS ONE 16(7):e0255290. https://doi.org/10.1371/journal.pone.0255290

Amadou HI, Bebeli PJ, Kaltsikes PJ (2001) Genetic diversity in Bambara groundnut (*Vigna subterranea* L.) germplasm revealed by RAPD markers. Genome 44(6):995–999

Boczkowska M, Łapiński B, Kordulasińska I, Dostatny DF, Czembor JH (2016) Promoting the use of common oat genetic resources through diversity analysis and core collection construction. PLoS ONE 11(12):e0167855. https://doi.org/10.1371/journal.pone.0167855

Bohra A et al (2021) Genomics: shaping legume improvement. In: Saxena KB, Saxena RK, Varshney RK (eds) Genetic enhancement in major food legumes. Springer, Cham

Chimonyo VGP, Wimalasiri EM, Kunz R, Modi AT, Mabhaudhi T (2020) Optimizing traditional cropping systems under climate change: a case of maize landraces and Bambara groundnut. Front Sustain Food Syst. https://doi.org/10.3389/fsufs.2020.562568

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure: Extensions to linked loci and correlated allele frequencies. Genetics 164:1567

FAOSTAT (2022) Rome: wold agricultural. Information Center. Available online: https://www.fao.org/faostat/en/#data/QCL. Accessed on 26 March 2024

Fatokun C, Gezahegn G, Michael A, Melaku G, Nnanna U, Olaniyi O, Muyideen Y, Ismail R, Ousmane B (2018) Genetic diversity and population structure of a mini-core subset from the world cowpea (*Vigna unguiculata* (L.) Walp.) germplasm collection. Sci Rep 8:16035. https://doi.org/10.1038/s41598-018-34555-9

Hamadou M, Martin AM (2024) Investigating Bambara bean (*Vigna subterranea* (Verdc.) L.) protein and hydrolysates: a comprehensive analysis of biological and biochemical properties. Appl Food Res 4(2): 100489, ISSN 2772-5022, https://doi.org/10.1016/j.afres.2024.100489

Jaetzold R, Schmidt H (1982) Farm management handbook of Kenya. Natural conditions and farm management information. Part A, B and C. Ministry of Agriculture, Kenya

Kamfwa K, Cichy KA, Kelly JD (2015) Genome-wide association study of agronomic traits in common bean. Plant Genome. https://doi.org/10.3835/plantgenome2014.09.0059

Ketema S, Tesfaye B, Keneni G, Amsalu FB, Assefa E, Greliche N et al (2020) DArTSeq SNP-based markers revealed high genetic diversity and structured population in Ethiopian cowpea [*Vigna unguiculata* (L.) Walp] germplasms. PLoS ONE 15(10):e0239122. https://doi.org/10.1371/journal.pone.0239122

Khan MMH, Rafii MY, Ramlee SI, Jusoh M, Al-Mamun M (2021) Bambara groundnut (*Vigna subterranea* L. Verdc): a crop for the new millennium, its genetic diversity, and improvements to mitigate future food and nutritional challenges. Sustainability 13(10):5530. https://doi.org/10.3390/su13105530

Kilian A et al (2012) Diversity Arrays Technology: a generic genome profiling technology on open platforms. In: Pompanon F, Bonin A (eds) Data Production and Analysis in Population Genomics. Methods in Molecular Biology (Methods and Protocols), vol 888. Humana Press, Totowa NJ

Kuan HW, Chai HH, Kendabie P, Ahmad NS, Jani J, Massawe F, Kilian A, Mayes S (2017) Integrating genetic maps in Bambara groundnut [*Vigna subterranea* (L.) Verdc.] and their syntenic relationships among closely related legumes. BMC Genet 18:192. https://doi.org/10.1186/s12864-016-3393-8

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhiwu Z (2012) GAPIT: genome association and prediction integrated tool. J Bioinform 28:2397–2399. https://doi.org/10.1093/bioinformatics/bts444

Massawe F, Mayes S, Cheng A (2016) A Crop diversity: an unexploited treasure for food security. Trends Plant Sci 21(5):365–368. https://doi.org/10.1016/j.tplants

Mayes S, Kuan WH, Hui H, Gao X, Kundy AC, Mateva KI, Zahrulakmal M, Hahiree MKIM, Kendabie P, Licea LCS, Massawe F, Mabhaudh T, Modi AT, Berchie N, Amoah S, Faloye B, Abberton M, Olaniyi O, Azam-Ali SN (2019) Bambara groundnut: an exemplar underutilised legume for resilience under climate change. Planta Org. https://doi.org/10.1007/s00425-019-03191-6

Nadeem MA, Habyarimana EC, Çiftçi V, Nawaz MA, Karakoy T, Comertpay G et al (2018) Characterization of genetic diversity in Turkish common bean gene pool using phenotypic and whole-genome DArTseq-generated silicoDArT marker information. PLoS ONE 13(10):e0205363. https://doi.org/10.1371/journal.pone.0205363

Nazarul H, Sana C, Neha N, Nidhi S, Rafiul AL (2021) Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. J Genet Eng Biotechnol 19(1): 128, ISSN 1687-157X. https://doi.org/10.1186/s43141-021-00231-1

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data II. Gene Frequency Data J Mol Evol 19:153–170. https://doi.org/10.1007/bf02300753

Ntundu WH, Inga CB, Jørgen L, Christiansen Sven B, Andersen (2004) Analysis of genetic diversity in Bambara groundnut [*Vigna subterranea* (L.) Verdc] landraces using amplified fragment length polymorphism (AFLP) markers. Afr J Biotechnol 3(4):220–225

Odongo FO, Oyoo ME, Wasike V, Owuoche JO, Karanja L, Korir P (2015) Genetic diversity of Bambara groundnut (*Vigna subterranean* (L.) verdc.) landraces in Kenya using microsatellite markers. Afr J Biotechnol 14(4):283–291

Onwubiko NC, Uguru MI, Ngwuta AA, Inyang ET, Nnajiemere OJ (2011) Floral biology of Bambara groundnut [*Vigna subterranea* (L.) Verdc.]. J Crop Sci 3(11):293–295

Osundare OT, Akinyele BO, Odiyi AC et al (2023) (2023) Genetic diversity and population structure of some Nigerian accessions of Bambara groundnut (*Vigna subterranea* (L.) Verdc.) using DArT SNP markers. Genet Resour Crop Evol 70:887–901. https://doi.org/10.1007/s10722-022-01472-w

Pasipanodya JT, Horn LN, Achigan-Dako EG, Musango R, Sibiya J (2022) Utilization of plant genetic resources of Bambara groundnut conserved ex situ and genetic diversification of its primary genepool for semi-arid production. Agriculture 12(4):492. https://doi.org/10.3390/agriculture12040492

Pasquet RS, Schwedes S, Gepts P (1999) Isozyme diversity in Bambara groundnut. Crop Sci 39:1228–1236

Peakall R, Smouse PE (2012) GenAlEx 6.5. Genetic analysis in Excel. Population genetic software for teaching and research—an update. J Bioinform 28:2537–2539

Pritchard JK, Wen W, Falush D (2003) Documentation for structure software: 2003. Version 2

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

Sahu TK, Singh M, Kalia S, Singh AK (2023) Genome-wide association study (GWAS): concept and methodology for gene mapping in plants. In: Raina A, Wani MR, Laskar RA, Tomlekova N, Khan S (eds) Advanced crop improvement, vol 2. Springer, Cham. https://doi.org/10.1007/978-3-031-26669-0_17

Salazar-Licea L et al (2022) The Bambara groundnut genome. In: Chapman MA (ed) Underutilised crop genomes. Compendium of plant genomes. Springer, Cham. https://doi.org/10.1007/978-3-031-00848-1_11

Somta P, Sompong C, Alisa K, Peerasak S (2015) QTLS controlling seed weight and days to flowering in mungbean [*Vigna radiata* (L.) Wilczek], their conservation in Azuki bean [*V. angularis* (Ohwi) Ohwi & Ohashi] and rice bean [*V. umbellata* (Thunb) Ohwi & Ohashi]. Agrivita. https://doi.org/10.17503/Agrivita-2015-37-2-p159-168

Sun L, Lai M, Ghouri F, Nawaz MA, Ali F, Baloch FS, Nadeem MA, Aasim M, Shahid MQ (2024) Modern plant breeding techniques in crop improvement and genetic diversity: from molecular markers and gene editing to artificial intelligence—a critical review. Plants 13(19):2676. https://doi.org/10.3390/plants13192676

Susmitha P, Kumar P, Yadav P, Sahoo S, Kaur G, Pandey MK, Singh V, Tseng TM, Gangurde SS (2023) Genome-wide association study as a powerful tool for dissecting competitive traits in legumes. Front Plant Sci 14:1123631. https://doi.org/10.3389/fpls.2023.1123631

Thouseem N, Beena T, Namitha E (2018) Genetic parameters in seed yield components of Cowpea (*Vigna unguiculata* (L.) Walp). N Int J Curr Microbiol Appl Sci 7(2):2268–2274

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

Uba CU, Oselebe HO, Tesfaye AA et al (2023) Association mapping in bambara groundnut [*Vigna subterranea* (L.) Verdc.] reveals loci associated with agro-morphological traits. BMC Genom 24:593. https://doi.org/10.1186/s12864-023-09684-9

Uffelmann E, Huang QQ, Munung NS et al (2021) Genome-wide association studies. Nat Rev Methods Primers. https://doi.org/10.1038/s43586-021-00056-9

Valdisser PAMR, Wendell PJ, Almeida JE, Filho MBSF, Coelho GRC, Menezes IPP, de Vianna JPG, Zucchi MI, Lanna AC, Coelho ASG, Oliveir JP, Moraes AC, Brondani C, Vianello RP (2017) In-depth genome characterization of a Brazilian common bean core collection using DArT-seq high-density SNP genotyping. BMC Genom 18:423. https://doi.org/10.1186/s12864-017-3805-4

Wakhungu CN, Tabu IM, Otaye OD (2017) Wasike WV (2017) distribution of fusarium wilt of Bambara nut (*Vigna subterranea* (L.) verdc.) in farmers' fields' of Busia county in western Kenya and its management using farmyard manure. Arch Phytopathol Plant Protect 50(7–8):398–414. https://doi.org/10.1080/03235408.2017.1319645

Wen L, Chang H-X, Brown PJ, Domier LL, Hartman GL (2019) Genome-wide association and genomic prediction identifies soybean cyst nematode resistance in common bean including a syntenic region to soybean Rhg1 locus. Hortic Res 6:9. https://doi.org/10.1038/s41438-018-0085-3

Wright S (1977) Evolution and the genetics of populations. experimental results and evolutionary deductions, vol 3. The University of Chicago Press, London, p 622

Xie H, Wen Y, Choi Y, Zhang X (2021) Global trends on food security research: a bibliometric analysis. Land 10(2):119. https://doi.org/10.3390/land10020119

Yu JK, Chung YS (2021) Plant variety protection: current practices and insights. Genes 12(8):1127. https://doi.org/10.3390/genes12081127