Genetic diversity assessment of hydrogen cyanide, total carotenoid content, and dry matter content in biofortified cassava using trait-linked SNP markers

Bismark Anokye[1,4]*, *Peter Amoah*[2], *Bardee Wrojay Potter Jr.*[3], *Abdoul-Razak Oumarou Mahamane*[5], *Theophilus Adu-Gyamfi*[6], *Levitikos Dembure*[2], *Nezif Abajebal Abadura*[2], *Bunmi Olasanmi*[7], *Elizabeth Parkes*[1]*

[1]*International Institute of Tropical Agriculture (IITA), Ibadan, Oyo State, Nigeria*
[2]*Departement of Crop Production and Landscape Management, University of Ebonyi State, Abakaliki, Nigeria*
[3]*Pan African University Life and Earth Sciences Institute (including Health and Agriculture), Ibadan, Nigeria*
[4]*Department of Soil and Crop Sciences, Texas A&M University, College Station, TX, USA*
[5]*Department of Crop Sciences, North-West University (Mahikeng Campus), Mmabatho, South Africa*
[6]*Department of Plant Sciences, University of Cambridge, Cambridge, United Kingdom*
[7]*Department of Agronomy, University of Ibadan, Ibadan, Nigeria*

**Corresponding authors: E.parkes@cgiar.org; Bismark.anokye@ag.tamu.edu*

Citation: Anokye B., Amoah P., Potter Jr. B.W., Oumarou Mahamane A.R., Adu-Gyamfi T., Dembure L., Abajebal Abadura N., Olasanmi B., Parkes E. (2026): Genetic diversity assessment of hydrogen cyanide, total carotenoid content, and dry matter content in biofortified cassava using trait-linked SNP markers. Czech J. Genet. Plant Breed., 62: 102–114.

Abstract: Assessing of genetic diversity is essential for identifying useful alleles for crop improvement. This study evaluated genetic diversity among two cassava breeding populations for total carotenoid content (TCC), dry matter content (DMC), and hydrogen cyanide (HCN) concentration using trait-linked single nucleotide polymorphism (SNP) markers. A total of 360 genotypes were analysed, including 261 from the IITA breeding programme (Population 1), 23 progenitor lines, and 76 from the University of Ibadan Cassava (UIC) breeding programme (Population 2). Minor allele frequency (MAF), gene diversity (GD), observed heterozygosity (H_e), and polymorphic information content (PIC) were computed. Principal component analysis (PCA) and hierarchical clustering were performed to examine genetic variation and population structure. Call rates were high (96–100%). MAF ranged from 0.00 to 0.50, with mean values of 0.28, 0.28, and 0.29 for Population 1, Population 2, and progenitors, respectively. GD averaged 0.36, 0.36, and 0.35 across these groups. Observed heterozygosity was 0.42, 0.41, and 0.43, while PIC values averaged 0.29, 0.27, and 0.27 for Population 1, Population 2, and progenitors, respectively. PCA and clustering analyses grouped the genotypes into three clusters containing 257, 88, and 15 genotypes. The first two principal components explained 39.1% of the total genetic variation. The results indicate substantial genetic diversity among the studied genotypes, suggesting strong potential for allele pyramiding and highlighting the informativeness of the SNP markers used.

Keywords: biofortification; cassava; dry matter content; genetic diversity; hydrogen cyanide; SNP markers; total carotenoid

---

Supported by African Union [Grant No. PAU020110MB].

© The authors. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

https://doi.org/10.17221/121/2025-CJGPB

Cassava (*Manihot esculenta* Crantz; Euphorbiaceae) is a perennial, dicotyledonous tuber crop of global socio-economic and nutritional significance. Globally, it ranks fourth as an essential food crop only to rice, wheat and maize, playing a central role in food security and income generation for smallholder farmers (Mtunguja et al. 2019; Mohidin et al. 2023). Over 1 billion people in Sub-Saharan Africa (SSA), Asia and Latin America subsist on cassava as a staple food and a source of dietary energy (Adebayo 2023; Nyaika 2024). Additionally, cassava functions as an important cash crop, provides raw materials for agro-industries, serves as a famine reserve and a source of animal feed, and contributes to national foreign exchange earnings through export (Tumuhimbise et al. 2014). Thus, cassava plays an indispensable role in sustaining food systems, supporting rural livelihoods, and strengthening the agricultural resilience of millions worldwide.

Over the past decade, concerted breeding efforts have led to the development and release of cassava varieties combining traits such as provitamin A carotenoids (pVAC), through biofortification, high dry matter content (DMC), and low hydrogen cyanide (HCN) (Ilona et al. 2017; Eyinla et al. 2019; Machado et al. 2023; Kongsil et al. 2024). However, there remains a dearth of information on the genetic diversity of these improved varieties (Kamara et al. 2018). Consequently, this lack of information is hindering breeding efforts aimed at exploring useful alleles for population improvement. Comprehensive knowledge of the extent of genetic diversity within a source population is essential for broadening the genetic base to develop new varieties and for identifying genetic differences among closely related individuals (Tiago et al. 2016). Therefore, genetic analysis of cassava breeding lines is essential to enhance yield and nutritional potential while maximising the utilisation of desirable traits for the development of improved varieties (Andrade et al. 2017).

The assessment of genetic diversity in cassava populations can be approached through a variety of methods, including biochemical, morphological, and molecular characterisation (Okogbenin et al. 2012). Each approach has its advantages and limitations. Morphological markers, for instance, are strongly influenced by environmental conditions, which can reduce their reliability in estimating genetic diversity (Xiao et al. 2025). In contrast, DNA-based molecular markers offer a more precise alternative, unaffected by environmental factors, and enable plant breeders to correlate genotypic variation with phenotypic traits, thereby providing robust insights into genetic diversity (Amesh et al. 2020).

The use of modern breeding tools such as molecular markers has provided useful insight for genetic characterization (Ren et al. 2013; Xiao et al. 2025). This permits the detection of genetic polymorphisms among cassava genotypes by revealing useful alleles that can be exploited in cassava biofortification breeding pipeline (Rabbi et al. 2014; Udoh et al. 2017). Such information guides the breeder to select unique clones to broaden the genetic base and develop superior varieties since the information revealed is based on the plant genotypes and independent of environmental variation.

There are several DNA-based molecular markers available that plant breeders can leverage for genetic diversity studies. Some of these DNA-based molecular markers encompass restriction fragment length polymorphism (RFLPs), random amplified polymorphic DNA (RAPD) markers, amplified fragment length polymorphism (AFLP), inter simple sequence repeats (ISSRs), simple sequence repeats (SSRs), and single-nucleotide polymorphisms (SNPs), among others (Amesh et al. 2020). Previous studies have leveraged some of these markers over the years to characterise cassava germplasm for genetic variation (Ceballos et al. 2015; Karim et al. 2020). Nonetheless, SNP markers are increasingly becoming markers of choice for genetic diversity studies and plant breeding applications (Rabbi et al. 2020; Amoah et al. 2025). SNPs are useful in detecting and identifying specific genetic differences even in a low-diverse population (Ferri et al. 2010). A marker system based on SNPs provides extra advantages due to the low genotyping cost, abundance in the genomes, biallelic nature, and their amenability for high-throughput detection formats and platforms (De Oliveira et al. 2012; Ousmeal et al. 2023). These advantages make SNP markers particularly suitable for genetic diversity research (Varshney 2010; Welsch et al. 2010; Karim et al. 2020). Thus, the objective of the present study was to assess the genetic diversity and relationships between two cassava populations using SNP markers to explore useful alleles to broaden the population base.

MATERIAL AND METHODS

Description of plant material. A total of 360 cassava genotypes from two populations were used in this study. Population 1 comprised 261 genotypes derived from open-pollinated crosses among 23 advanced International Institute of Tropical Agriculture (IITA) breeding lines combining high carotenoid content, high dry matter, high starch, and Cassava mosaic virus (CMD) resistance. Population 2 consisted of 76 genotypes from the University of Ibadan Cassava (UIC) breeding program with storage root colour variability ranging from cream to deep yellow. The parental materials for population 2 included yellow-root cassava varieties (IITA-TMS 1070593, IITA-TMS 1011371, IITA-TMS 1070539, IITA-TMS 1011368, and IITA-TMS 1011412) selected for their high β-carotene content, CMD resistance, and storage root yield. Botanical seeds were deliberately used to generate the IITA-derived clones (Population 1), as sexual reproduction in cassava introduces recombination and allelic variation that are essential for assessing genetic diversity and identifying useful alleles for population improvement, rather than for varietal maintenance or clonal fidelity.

Experimental layout and management. Botanical seeds for Population 1 were sown in a greenhouse, and seedlings (20–25 cm tall) were transplanted to the field. Plants were harvested eight months after planting (MAP). Population 1 trial was established using an augmented design with a spacing of 1 m between rows and 0.25 m within rows, with clone TMEB117 as a common check. Seedlings were transplanted in the first week of November 2024 at IITA, Ibadan, Nigeria (7°24'N, 3°54' E). Population 2 was established in May 2024 at the Teaching and Research Farm of the Department of Crop and Horticultural Sciences, University of Ibadan (7°27'N, 3°45'E), using a randomised complete block design (RCBD). Twenty stem cuttings were planted per 20 m^2 plot at a spacing of 1 × 1 m, with two replications. Field management at both experimental stations followed standard cassava cultivation practices and technical recommendations.

Genotyping. The study evaluated 261 IITA-derived clones (Population 1), 76 University of Ibadan UIC clones (Population 2), and 23 parental lines used in generating Population 1, all of which were genotyped using nine trait-linked kompetitive allele-specific PCR (KASP) markers previously developed for cassava improvement (Udoh et al. 2017; Ogbonna et al. 2020; Rabbi et al. 2020) (Table 1) and flanking sequences of the SNPs presented in Table 2. Leaf samples were collected by punching four 6-mm disks from each of the 360 samples into 96-well plates using single-hole punchers and forceps. The plates were labeled, covered with paraffin oil, and stored at –50 °C in a LABCONCO FreeZone 18 Liter freeze-dryer at the IITA Bioscience Center. Samples were freeze-dried at –51 °C and 5.0 Pa for 72 h until completely dry, and then shipped to Intertek Laboratory, Australia, for genotyping using the KASP method.

Two non-template controls (NTCs) were included per plate. Genotyping was conducted using a high-throughput PCR platform in 384-well plates. Each KASP reaction contained genomic DNA, a marker-specific assay mix with two allele-specific forward primers and a common reverse primer, and the KASP-TF™ Master Mix. The master mix includes two FRET-based fluorescent cassettes (FAM and

Table 1. Description of the nine trait-linked KASP markers used for genotyping in the two populations

Traits	SNP name	Favourable allele	Unfavourable allele	Reference
Dry matter content	S1_24197219	C	T	Rabbi et al. (2020)
	S6_20589894	G	A	Rabbi et al. (2020)
	S12_5524524	C	T	Rabbi et al. (2020)
Carotenoid content	S1_24155522	A	C	Udoh et al. (2017)
	S13_0543962	G	A	Rabbi et al. (2020)
	S5_3387558	T	C	Rabbi et al. (2020)
	S8_25598183	T	G	Rabbi et al. (2020)
Hydrogen cyanide	S16_773999	A	G	Ogbonna et al. (2020)
	S14_6050078	G	A	Ogbonna et al. (2020)

The table lists each marker, the trait it targets, and the favourable and unfavourable alleles associated with improved or reduced phenotypic performance; these single nucleotide polymorphism (SNP) markers were previously validated in cassava breeding programs and are widely used for marker-assisted selection; the favourable alleles represent variants linked to desirable trait expression, while unfavourable alleles correspond to less preferred phenotypic outcomes; KASP – kompetitive allele-specific PCR

Table 2. Markers and flanking sequences of the SNPs used for genotyping cassava genotypes in two populations

Marker_allele_trait	Chr.	SNPs and sequence
S1_24197219 (C/T)	1	GATGTAGGCATGTTACATATAAGGGCTACATACACATTAGCAGCTAAAATGAGACCCGGATACCGAGCAATGCCATCAATTGAGAGATGAACTCAGGGTG[C/T]CCTGGCCATGCAGCTCCAGTAACCAAATTTTCATGAGTGTAGCAACGATGTATTGGATCAGGTTCTAGCCATGTTGCCCCAGCCAAGACCACGTTAATCT
S6_20589894 (A/T)	6	ATTGATGATTTTTTATTCATGATATGTAGCTATCAAAGTTACTCAGCAATGTCCTTGTTTTAGCCATGCTAGCAGCATGTTTTGTTGCGACAACAGTTGG[A/G]AGTTGTATGAATATTGTTTTATCTTGTATGCAGAATATCATTGGGCAGGAAGCAGGGAAAAGCGTGATTGAGGAATATTTACGTCGTAGGGGTCACTCAG
S13_0543962 (A/G)	1	GGAGGTTTTTTTATGTGGCATTCTCAGCAGCTGCAGGAATCTCATTGTTCTTTACAATTCCAAGGCTCTTTCTTGCAATTAAAGGTGGGGAAGGTGCCCC[A/G]GACCTCTGGGGAACTGCTGGAAATGCTGCCATTAATATTGGTGGTAAATGCTTTAACCTTTCTCTGTCATATGAAGAAAATGAGTTAATTGATGTATAAT
S5_3387558 (C/T)	5	GTTACACTTAGACCCTTGTCATTAAACATTACTGAGGCTGCAGTTGAAGTGTAAACAACTCTTTTCACTGTCTTTGATTCCAAGCATGTCCTTAATATCC[C/T]TAGCAATCCATCCACGGCTATTTTGGTCACACTTTCTTCAGGTTCTTTTCCATAATGATCCATTGGGTGAGCCACATGGAAGACTCCAATACAACCTTCA
S8_25598183 (G/T)	8	TAAATTCTGACTGTCTTGGCATGACTGTCCAGGTAGTCCCCGAAAATGAGAATGCTGCTCTCTACTCCACTCATTCATTCAAGATTTTGTTCAAGGAAGG[G/T]GGTTGTGGAACCTTCATTCCGCTCTTTTTCAACTTGCTCTCTTCAGTAAGGCAATACAATCAGCAACAAACCTCTGGAATGGGGCCCCAGATGAACCCTT
S1_24155522 (A/C)	1	GACAGATGAGCTTGTTGATGGACCTAATGCTTCACACATAACGCCAACAG[A/C]TTTAGATAGGTGGGA AGCAAGGTTGGAAGATATGTTTCGAGGTCGTCCCT
S12_5524524 (T/C)	12	TGAATTATTTTAACTCTTTGATTGCTTCGCCAGTGCCTGGTCTCCAGAATGTGTGTGTTGCTTTGGTTTGTAGTTCCAAAGGTGAGCTGTGGCAATTTTA[T/C]TGCAGCCCCACTGGCATTAGACGCAGTAAATTATATCAGGACGAAGTAAGTTCATCCTTCAAAGGAAATGATAATGGTCAATTTGTGGGGAGCAAAGGTT
S16_773999 (A/G)	16	GTGAGCAACGAGCTAGGAGCAGGACATCCAAAATCAGCAGCATTTTCTGTGATAATTGTGAATCTATG CTCTTTCATTATCGCTGTWRTCGCAGCCATCRTCGTGATGATWTTGCGAGACTATCTCAGCTACGCTTT YACTGATGGTGAA[A/G]CTGTTTCCAAAGCRGTCTCCGATCTYACCCCCTWCTTGGCTGYCACCCTCATTCTAAATGGSGTGCAGCCTGTTTTGTCCGGTRAGATAATYYCCCCAACCTCCCCCCTGTTCCCTGTTATTA TTAYGATTCCTTCATAGCAGTACTTTT
S14_6050078 (G/A)	14	CTCKGCAGGTATCTTACCCATAATTTTTTTTAWAAGGAAATACRGAGACTAAGATTTTCAAATATTTTATTCATATATGYCTAAYGTMTTAAAAATTGGCAAATTTATAGGAGCATTTGCAGCAATGGCAATATTCGGAAAGATGGACCA[A/G]TTACTTGCGCCTAAAGGAATTTCTATGACAATTGCGCCCTTAGGAGCTGTTTGTGCCGTCCTCTTTGCCACTCCTTCCTCYCCTGCTGCTCGGGTACCCTCCWCTGCTCYGTTTTATGGCAG TAGATAAAAATGAGAACTTTTTAACTTT

This table presents the nine trait-linked single nucleotide polymorphism (SNP) markers used for genotyping cassava populations, showing their chromosome positions, allele variants, and full flanking sequences; these markers target genomic regions associated with dry matter content (DMC), total carotenoid content (TCC), and hydrogen cyanide (HCN) concentration; bracketed nucleotides indicate the biallelic SNP site, where the first allele represents the favourable allele and the second denotes the unfavourable allele for the respective trait; chromosomal positions are provided for reference to the cassava genome assembly, and the flanking sequences represent the genomic regions targeted for kompetitive allele-specific PCR (KASP) genotyping assays; Chr. – chromosome position

HEX), a reference dye (ROX™), Taq DNA polymerase, dNTPs, and $MgCl_2$ in a proprietary buffer. After thermal cycling, fluorescence signals were read, and genotype calls were made using KRAKEN™ software. Detailed protocols for KASP setup and execution are provided in the KASP user guide (Genomics LGC 2013).

Genotypic data and marker-trait association analysis. SNP calls were generated using SNP Viewer, and genotypic data were exported as a marker-by-genotype matrix for downstream analyses. A raw HapMap genotype file was produced in R Statistical Software (R Core Team 2020). This file was imported into TASSEL 5.0 (Bradbury et al. 2007) to calculate key genetic diversity indices, including call rate, minor allele frequency (MAF), polymorphic information content (PIC), gene diversity (GD), observed heterozygosity (H_e), which were used to assess the extent of genetic differences between the two populations and parental lines. Genotypes were further coded according to allele calls at nine SNP loci associated with key agronomic traits in cassava.

Genotypic data were curated by removing ambiguous calls (e.g., 'NN' or missing values), while retaining all SNP loci due to the limited number of informative markers. SNPs were treated as categorical variables rather than being forced into numeric encodings to preserve allelic information. Genetic dissimilarity among genotypes was quantified using an allele mismatch distance, calculated as the proportion of non-matching alleles across loci, excluding missing comparisons.

Genetic relationships among the cassava genotypes were examined using hierarchical clustering based on an allele-mismatch distance matrix calculated from the nine trait-linked SNP markers. Clustering was performed using the unweighted pair group method with arithmetic mean (UPGMA). To evaluate the statistical robustness of the inferred clusters, multiscale bootstrap resampling was conducted using the pvclust package in R (Suzuki & Shimodaira 2006). A total of 1 000 bootstrap iterations were performed, and approximately unbiased (AU) *P*-values were computed to assess the support of internal branches in the dendrogram. Branches with high AU support values were interpreted as relatively stable, providing a conservative evaluation of cluster robustness and genetic relationship among the cassava genotypes (Daemo et al. 2023).

Phylogenetic relationships among the cassava genotypes were visualised as a circular (fan-type) dendrogram, generated using the ape package in R (Paradis et al. 2004). Genotypes were labelled according to their assigned identifiers (G1–G360) and colour-coded to indicate major genetic clusters identified from the hierarchical analysis. This visualisation enabled a clear interpretation of genetic structure across genotypes drawn from the different source populations.

Principal component analysis (PCA) was additionally performed using the prcomp function in R to summarise overall patterns of genetic variation across the nine SNP markers and to complement the clustering results (R Core Team 2020). Scripts used for PCA visualisation and hierarchical clustering, including bootstrap resampling and dendrogram construction, are provided as Figure S1 in Electronic Supplementary Material (ESM).

RESULTS

Diversity assessment among genotypes in two breeding populations; Population 1 and 2. The summary statistics of the nine SNP markers used to estimate the MAF, GD, level of H_e, and the PIC values are presented in Table 3. The MAF, PIC, H_e, and GD are key population genetic indices that measure the informativeness of a marker and genetic differences among any studied populations. Genetic variation was evident among the 360 cassava genotypes**,** which include both breeding populations and their 23 progenitor lines**,** highlighting the diversity present within the germplasm set.

The MAF ranged from 0.00 (S1_24197219, S1_24155522) to 0.43 (S8_25598183), with an average value of 0.28 in Population 2, while in Population 1 the MAF ranged from 0.11 (S1_24197219) to 0.45 (S5_3387558), with an average of 0.28 across the SNP markers. The progenitors MAF ranged from 0.00 to 0.5 with a mean of 0.29.

The mean H_e across all loci was 0.42 and 0.41 in Population 2 and Population 1, respectively. In Population 1, H_e values ranged from 0.15 (S1_24197219) to 0.88 (S5_3387558), whereas in Population 2 they ranged from 0.00 (S1_24197219, S1_24155522) to 0.73 (S5_3387558). GD was highest (0.50) for SNP S5_3387558 and S16_773999 and lowest (0.00) for S1_24197219 and S1_24155522, with the mean GD of 0.35 across the parents of Population 1. The average gene diversities obtained for the two populations, including the progenitors, were somewhat lower than the mean heterozygosity observed.

The PIC values ranged from 0.17 (S1_24197219) to 0.37 (S5_3387558, S13_0543962) with a mean of 0.29 in Population 1, and from 0.00 (S1_24155522, S1_24197219) to 0.37 (S8_25598183, S16_773999), with a mean of 0.27 in Population 2. The PIC values in the parents of Population 1 ranged from 0.00 to 0.38, with an average of 0.27. About 78% of the SNP markers in this study had PIC values greater

https://doi.org/10.17221/121/2025-CJGPB

Table 3. Summary statistics of the nine SNP markers for genetic diversity studies of the three studied populations

SNP marker	Population 1				Population 2				Parents of Population 1			
	MAF	GD	H_e	PIC	MAF	GD	H_e	PIC	MAF	GD	H_e	PIC
S1_24155522	0.14	0.24	0.22	0.21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S1_24197219	0.11	0.19	0.15	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S1_30543962	0.43	0.49	0.63	0.37	0.37	0.47	0.52	0.36	0.37	0.47	0.57	0.36
S5_3387558	0.45	0.50	0.88	0.37	0.36	0.46	0.73	0.36	0.46	0.49	0.91	0.37
S6_20589894	0.40	0.48	0.41	0.36	0.34	0.45	0.53	0.35	0.32	0.43	0.36	0.34
S8_25598183	0.24	0.36	0.39	0.30	0.43	0.49	0.57	0.37	0.37	0.47	0.48	0.36
S12_5524524	0.15	0.26	0.30	0.22	0.37	0.46	0.58	0.35	0.37	0.46	0.64	0.36
S14_6050078	0.21	0.33	0.36	0.28	0.28	0.38	0.40	0.31	0.20	0.31	0.39	0.27
S16_773999	0.29	0.41	0.37	0.32	0.41	0.49	0.42	0.37	0.50	0.50	0.48	0.38
Mean	0.28	0.36	0.41	0.29	0.28	0.36	0.42	0.27	0.29	0.35	0.43	0.27

This table summarises key genetic diversity statistics (MAF, GD, H_e, and PIC) for the nine single nucleotide polymorphism (SNP) markers across two cassava breeding populations and their progenitors; the values highlight differences in allele frequencies and marker informativeness among populations; mean values provide an overall assessment of diversity within each group; these statistics correspond to the same markers whose sequences and allelic variants are presented in Table 2; MAF – minor allele frequency; GD – gene diversity; H_e – heterozygosity; PIC – polymorphic information content

than 0.28, indicating that the majority of the markers were moderately informative compared to those with lower PIC values. GD, H_e, and PIC values exhibited a consistent pattern across the SNP markers, indicating that the same markers tended to show higher or lower levels of genetic informativeness. Different PIC values were observed across the SNP loci for the entire populations.

The PIC, MAF, H_e, and GD values varied substantially across the nine SNP markers, reflecting clear genetic differentiation among the cassava clones. With the exception of markers S1_24197219 and S1_24155522, which recorded PIC, GD, and H_e values of 0.00 in Population 2 and the progenitors, all other markers showed moderate to high diversity across the populations.

Notably, several markers, including S5_3387558, S1_30543962, S6_20589894, and S16_773999, exhibited comparatively higher PIC (0.31–0.37), GD (0.45–0.50), and H_e (0.52–0.91) values, indicating strong discriminatory power. When averaged across SNP loci, Population 1 showed the highest mean genetic diversity (MAF = 0.28; GD = 0.36; H_e = 0.41; PIC = 0.29), followed closely by Population 2 (MAF = 0.28; GD = 0.36; H_e = 0.42; PIC = 0.27). The progenitor set displayed the lowest diversity (MAF = 0.29; GD = 0.35; H_e = 0.43; PIC = 0.27), consistent with their narrower genetic base relative to the segregating breeding populations.

Population stratification. Figure 1 represents the population structure of cassava genotypes characterised by PCA and diversity among 360 cassava genotypes based on the nine markers across the two populations and the progenitors. The total amount of variation explained by the first two PCs was 39.1%. The first and second PCs (PC1 and PC2) explained respectively 22.3% and 16.8% of the total variation observed across the populations. There are no distinct cluster patterns along the two PCs based on the genotype background or source.

Cluster membership and genetic differentiation. Cluster analysis grouped the genotypes into three distinct clusters (Figure 2). Cluster 1 contained the highest number of genotypes (340), followed by cluster 2 (12), while cluster 3 had the fewest (8). Members of cluster 1 comprised clones from populations 1 and 2, and progenitors, whereas clusters 2 and 3 included a mixture of genotypes from both Populations 1 and 2 only. The greatest pairwise genetic distance was observed between clones belonging to different major UPGMA clusters, indicating substantial divergence among the cassava germplasm, whereas the smallest genetic distances were detected among closely related clones that clustered together at low dissimilarity levels, reflecting high genetic similarity within these groups.

Genotypes in cluster 3 were characterised by comparatively low cyanide potential (range = 9.31 to 49.00 mg/100 g; mean = 27.18 mg/100 g) and low total carotenoid content (TCC) (range = 0.59–7.15 µg/g; mean = 4.40 µg/g), along with low to moderate DMC relative in clusters 1 and 3. DMC among cluster 1 genotypes ranged from 16% (TMEB3) to 30.9% (UIC-17-646), with an average of 23.94%.

Based on marker properties, cluster 3 genotypes were either heterozygous (carrying both favourable and unfavourable alleles) or homozygous (two unfavourable alleles) for the four markers associated with pVAC accumulation, as well as for locus S14_6050078 (genotypes AA and AG), which is linked to hydrogen cyanide (HCN) content (Table 1). Cluster 2 comprised population 1 and the remaining genotypes from population 2. Notably, IITA-TMS-IBA141092 was the female parent used to generate IBASN201149 (Population 1), and both the parent and offspring were consistently grouped within clusters 2 and 1, indicating shared genetic similarity.

A similar clustering pattern was observed for variety IITA-TMS-IBA180158 when used as female to generate IBASN201160 (Population 1). However, when variety IITA-TMS-IBA180158 was used as male (IITA-TMS-IBA154810/IITA-TMS-IBA180158), all the progenies (family IBASN201098) were grouped in cluster 1. The majority of the parents used as both males and females were found in cluster 3, including their progenies. Cluster 2 members are generally high cyanogenic potential genotypes with high carotenoids and dry matter content.

Cluster 1 is the largest consisting of clones from both populations and the progenitors of Population 1. The genotypes consist of low, moderate, and high TCC, DMC, and HCN content. All three genotype classes for 7 SNP markers were observed within cluster 3, except markers S1_24155522 (AA) and S1_24197219 (TT), which had only one type of genotype class.

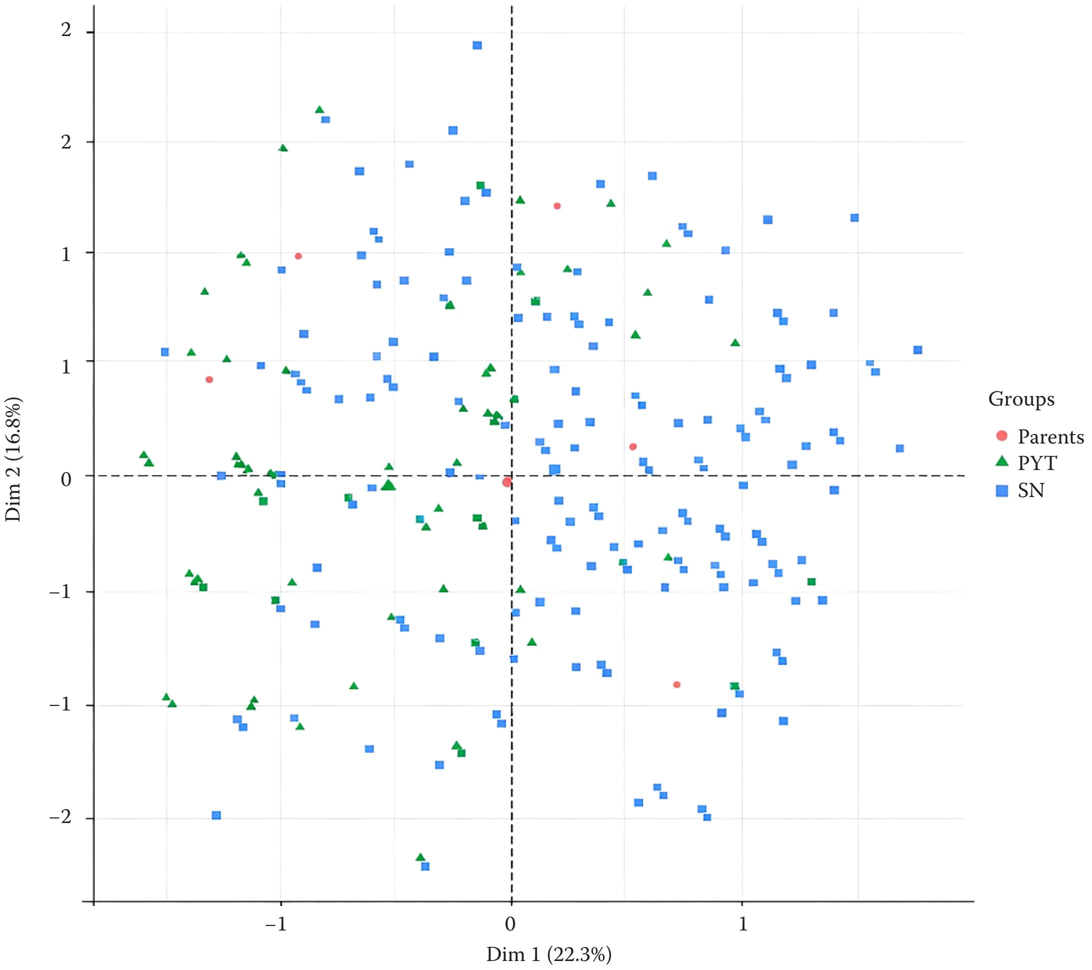

Figure 1. Principal component analysis of 360 cassava genotypes derived from nine SNP markers
PYT – Population 1; SN – Population 2

https://doi.org/10.17221/121/2025-CJGPB

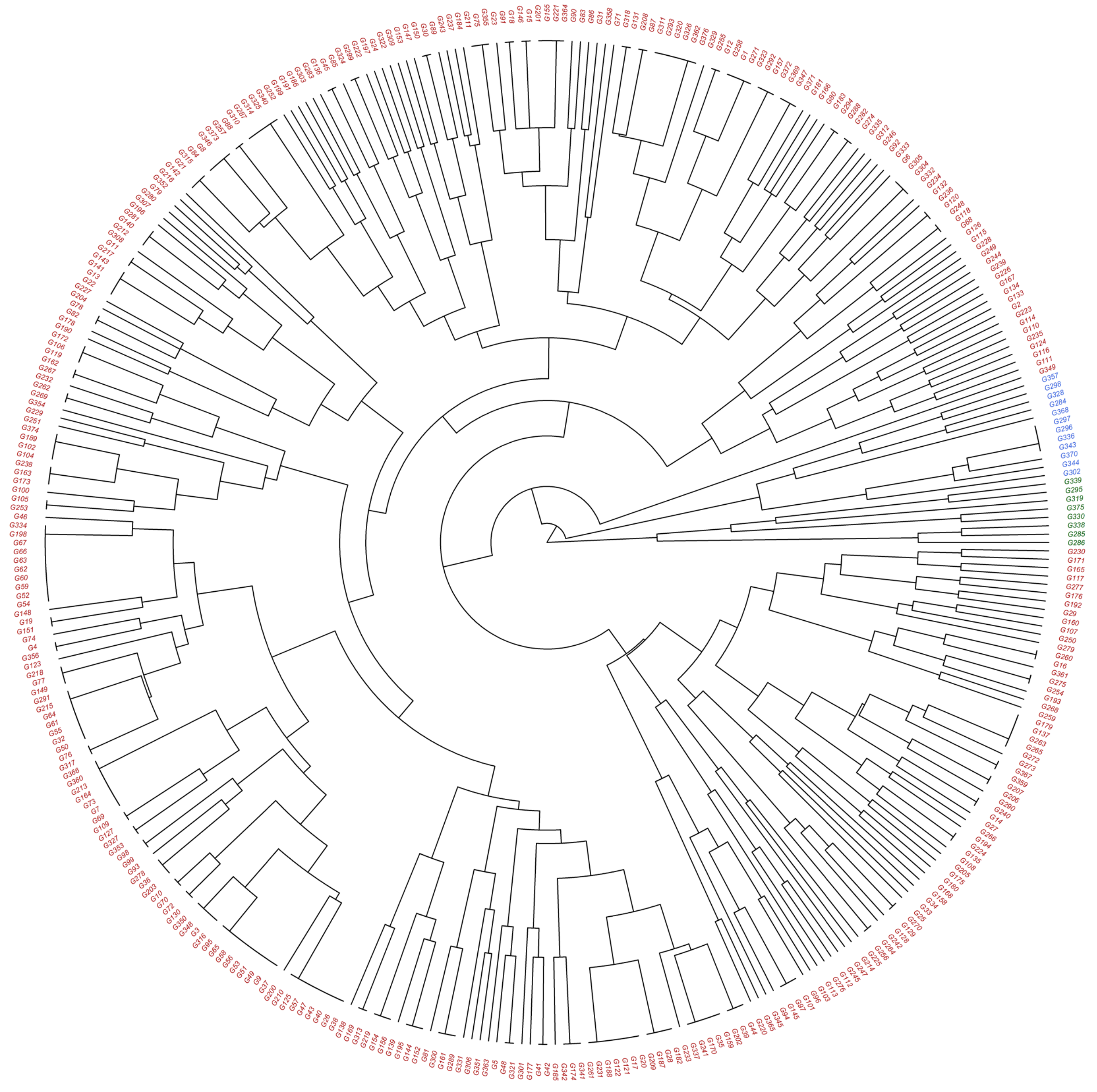

Figure 2. Circular (fan-type) hierarchical clustering dendrogram of genotypes (G1–G360) based on allele mismatch distance calculated from nine single nucleotide polymorphism (SNP) markers; branch lengths represent genetic dissimilarity; genotype labels correspond to sample identifiers listed in Table 1; the fan layout provides a comprehensive visual overview of genetic relationships and cluster memberships among genotypes, with branch lengths representing genetic dissimilarity; statistical robustness of the hierarchical cluster was evaluated using multiscale bootstrap resampling (1 000 iterations) implemented in pvclust package in R; full approximately unbiased (AU) and bootstrap probability (BP) values are reported in Table S1 and Figure S1 in ESM

Samples were grouped into three clusters (clusters 1–3), indicated by colour coding; cluster 1 (red) contained the largest number of genotypes, followed by cluster 2 (blue), while cluster 3 (green) comprised the fewest genotypes

DISCUSSION

SNP markers were employed to estimate genetic diversity and assess the extent of genetic variation for TCC, DMC, and HCN concentration among two cassava populations: Population 1 (IITA-derived) and Population 2 (UIC-derived), along with the progenitors used in developing Population 1. The efficient utilization of source populations in crop improvement programs largely depends on a clear understanding of the genetic differences and the magnitude of variation existing within and between them. The variation in MAF observed across the SNP markers reflects the genetic heterogeneity between the two populations, consistent with the findings of Ioannidis (2009). The range of MAF values obtained confirms that the SNP markers used in this study were highly polymorphic and, therefore, suitable for assessing genetic diversity among the cassava genotypes. MAF represent a key indicator of allelic richness and balance within a population that plays a critical role in determining the informativeness of SNP markers for genetic diversity assessment (Wang et al. 2009a). SNP loci with moderate to high MAF values contribute more effectively to estimates of GD, H_e, and PIC, whereas loci low MAF values provide limited discriminatory power due to allele fixation (Prakash et al. 2019). In SNP-based diversity studies, MAF thresholds of ≥0.50 are generally considered acceptable while values between 0.20 and 0.50 are regarded as highly informative for population genetic inference (Wang et al. 2009b; Udoh et al. 2017; Zhang et al. 2018; Rabbi et al. 2020). In the present study, MAF values ranged from 0.00 to 0.50, with mean values of approximately 0.28 across populations, signifying a predominance of moderately frequent alleles and substantial genetic variation. The predominance of SNP markers within this informative range in the current study supports their effectiveness for assessing genetic diversity and genetic relationship among cassava breeding populations (Kanaka et al. 2023).

The observed MAF distribution across populations reflects several contributing factors, including the use of sexually derived segregating breeding populations, selection pressure for quality-related traits (TCC, DMC, and HCN concentration), and the trait-linked nature of the SNP markers employed. SNP markers with MAF values approaching 0.50 suggest balanced allele frequencies and high informativeness, while loci with near-zero MAF values reflect fixation resulting from directional selection or physical linkage to major-effect loci (Hernandez et al. 2019).

According to Wang et al. (2009a), the quality of a SNP marker is primarily determined by its MAF and H_e proportion. Although cassava is predominantly an outcrossing species and the use of botanical seeds is expected to generate high levels of heterozygosity, the breeding populations analysed in this study do not only represent randomly mating natural populations in Hardy-Weinberg equilibrium. Rather, they consist of segregating progenies derived from selected parental lines and were genotyped using trait-linked SNP markers associated with quality traits. Under such conditions, directional selection, family structure, and linkage to major-effect loci can lead to deviations from Hardy–Weinberg expectations, including heterozygote excess and reduced average gene diversity. Similar patterns have been reported in cassava and other outcrossing, clonally propagated crops when functional or trait-linked markers were used for genetic diversity assessment (Alves 2002; Kawuki et al. 2009; Rabbi et al. 2020; Kanaka et al. 2023). The observed pattern is primarily explained by the outcrossing nature of cassava and the breeding structure of the populations analysed, although the historically diversity genetic background of cassava, shaped in part by introgression from wild *Manihot* relatives, may also contribute to the overall high heterozygosity observed (Byrne 1984; Alves 2002; Alves-Pereira et al. 2022).

More than 78% of the SNP markers analysed exhibited a PIC value greater than 0.28, indicating that these markers were informative. The PIC value reflects the discriminatory power and informativeness of a marker, thereby quantifying the degree of polymorphism among genotypes at a given SNP locus (Ganapathy et al. 2012; Makueti et al. 2015; Prempeh et al. 2020; Abadura et al. 2025). Thus, the SNP markers employed in this study were of sufficiently high quality to detect genetic variation among the cassava genotypes (Oliveira et al. 2014).

The variations in PIC values across loci may be attributed to several factors, including the breeding behaviour of the species, evolutionary pressures, mutation rates within the population, the level of genetic diversity and sample size, the sensitivity of the genotyping method, and the genomic location of the primers used in the study (Choudhury et al. 2013; Ogbonna et al. 2021).

The bi-allelic nature of SNPs leads to low PIC values, reported to range from 0 to 0.5, as opposed to SSRs, which are multi-allelic and can have a PIC value that goes above 0.50 to 1.0 (Choudhury et al. 2013; Prempeh et al. 2020). Thus, the two populations studied showed moderate to high genetic diversity for TCC, DMC, and HCN content for high PIC, GD, and H_e values recorded (Prempeh et al. 2020; Olasanmi et al. 2021). This suggests the availability of useful alleles that could be exploited in cassava biofortification programs. The determination of the extent of genetic diversity for traits among breeding populations using SNPs is hence, highly recommended (Wolfe et al. 2016; Andrade et al. 2017). Markers with relatively higher PIC revealed high GD and H_e, as also reported by Olasanmi et al. (2021). The polymorphism and the distribution of DNA segments across the genome differ among different marker types for the genotypic background.

The values observed for markers S1_24197219 and S1_24155522 were expected since the two markers co-segregated in the studied population due to physical linkage on chromosome 1, which was also observed by Zhang et al. (2015). A genome-wide association study (GWAS) report by Rabbi et al. (2020) found the two SNP loci in the same position on chromosome 1. Although the markers used in this study were mainly designed for marker-assisted selection (MAS), however, they can be used for genetic diversity studies.

The clustering of genotypes along the principal component (PC) axes indicates extensive sharing of genetic backgrounds among the studied populations. The PCA revealed the key traits contributing to variation across 360 cassava genotypes. The close relatedness between UIC-derived clones (Population 2) and IITA HarvestPlus clones (Population 1), as well as historical bottlenecks in the breeding history of IITA progenitors, likely shaped the observed clustering.

Cluster 1 contained parents used as both male and female lines and their progenies, suggesting a shared pedigree structure consistent with parental contribution observed in reciprocal crosses (Akinwale et al. 2010). The diversity of genotype classes corresponds with the wide phenotypic range observed for assessed traits. The co-occurrence of both populations and progenitors in cluster 1 further indicates extensive genomic exchange between IITA- and UIC-derived lines. The predominance of progenitors and Population 1 genotypes in cluster 1 confirms their role as the primary pedigrees of Population 1. Moreover, genotypes sharing the same parental clones tended to cluster together (clusters 1 and 2), highlighting the influence of pedigree on genetic structure (Kawuki et al. 2009). The predominance of high approximately unbiased (AU) *P*-values coupled with low bootstrap probability (BP) values (Table S1 in ESM) reflects the advantage of multiscale bootstrap resampling in hierarchical clustering, even when a limited number of markers is used. Previous studies have shown that BP values tend to underestimate branch support due to selection bias inherent in standard bootstrap procedures, whereas AU values provide a less biased and more reliable assessment of cluster stability, even when a limited number of markers is used (Wang et al. 2016; Flores et al. 2021). Accordingly, the AU-supported clusters observed in this study likely represent robust genetic groupings within the cassava germplasm.

CONCLUSION

This study identified substantial genetic diversity among the 360 cassava genotypes drawn from two breeding populations and their progenitors, which can be exploited in future breeding programs to facilitate the development of low-HCN clones with high TCC and DMC. Most SNP markers (78%) were moderately to highly informative, as reflected by their MAF, GD, H_e, and PIC values. PCA and hierarchical clustering revealed three major genetic groups reflecting shared ancestry and breeding history. The observed genetic differentiation and allele distribution patterns indicate the presence of useful variation for improving TCC, DMC, and HCN concentration. These findings highlight valuable alleles for cassava biofortification and demonstrate the effectiveness of trait-linked SNP markers for accelerating genetic improvement through marker-assisted and genomic selection.

Acknowledgements. The authors extend their colossal gratitude to the Pan African University of Earth and Life Science (including Health and Agriculture), University of Ibadan, and the International Institute of Tropical Agriculture.

REFERENCES

Abadura N.S., Abebe A.T., Rabbi I.Y., Beyene T.M., Abtew W.G. (2025): DArTseq SNP-based genetic diversity analyses in cassava (*Manihot esculenta* Crantz) genotypes sourced from different regions revealed high levels of within-population diversity. PLoS ONE, 20: e0308027.

Adebayo W.G. (2023): Cassava production in africa: A panel analysis of the drivers and trends. Heliyon, 9: e19939.

Akinwale M., Aladesanwa R., Akinyele B., Dixon A., Odiyi A. (2010): Inheritance of ß carotene in cassava (*Manihot esculenta* Crantz). International Journal of Genetics and Molecular Biology, 2: 198–201.

Alves A.A.C. (2002): Cassava botany and physiology. Chapter 5. In: Hillocks R.J., Thresh J.M., Bellotti A.C. (eds): Cassava: Biology, Production, and Utilization. Wallingford, CABI: 67–89.

Alves-Pereira A., Zucchi M.I., Clement C.R., Viana J.P.G., Pinheiro J.B., Veasey E.A., de Souza A.P. (2022): Selective signatures and high genome-wide diversity in traditional Brazilian manioc (*Manihot esculenta* Crantz) varieties. Scientific Reports, 12: 1268.

Amoah P., Adetimirin V.O., Anokye B., llesanmi O.J., Unachukwu N., Aprova E.I. (2025): Characterization of carotenoid profiles and presence of functional markers in sub-tropical maize (*Zea mays* L.) inbred lines. Czech Journal of Genetics and Plant Breeding, 61: 86–99.

Amesh P.A.R., Allikarjuna G.U.M., Ameena S.H.S., Umar A.N.K., Urulakshmi K.O.L.A.G., Eddy B.V.I.R., Eddy P.C.H.O.B.U.L.R., Ekhar A.K.C.H.S. (2020): Advancements in molecular marker technologies and their applications in diversity studies. Journal of Biosciences, 45: 123.

Andrade E.K.V., Andrade J.V.C., Laia M.L., Fernandes J.S.C., Oliveira A.J.M., Azevedo A.M. (2017): Genetic dissimilarity among sweet potato genotypes using morphological and molecular descriptors. Acta Scientiarum. Agronomy, 39: 447–455.

Bradbury P.J., Zhang Z., Kroon D.E., Casstevens T.M., Ramdoss Y., Buckler E.S. (2007): TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics, 23: 2633–2635.

Byrne D. (1984): Breeding cassava. Chapter 3. In: Janick J. (ed.): Plant Breeding Reviews, Vol. 2, Westport, Avi Publishing Company: 73–133.

Ceballos H., Kawuki R.S., Gracen V.E., Yencho G.C., Hershey C.H. (2015): Conventional breeding, marker-assisted selection, genomic selection and inbreeding in clonally propagated crops: A case study for cassava. Theoretical and Applied Genetics, 128: 1647–1667.

Choudhury B., Khan M.L., Dayanandan S. (2013): Genetic structure and diversity of indigenous rice varieties (*Oryza sativa*) in the Eastern Himalayan region of Northeast India. Springer Plus, 2: 228–237.

Daemo B.B., Yohannes D.B., Beyene T.M., Abtew W.G. (2023): Genetic variability and divergence analysis in cassava (*Manihot esculenta* Crantz) accessions for storage root yield and other agronomic characters. Journal of Agriculture and Food Research, 14: 100874.

De Oliveira E.J., Vilela de Resende M.D., Da Silva Santos V., Fortes Ferreira C., Alvarenga Fachardo Oliveira G., Suzarte Da Silva M., Aguilar-Vildoso C.I. (2012): Genome-wide selection in cassava. Euphytica, 187: 263–276.

Eyinla T.E., Maziya-Dixon B., Alamu O.E., Sanusi R.A. (2019): Retention of pro-vitamin A content in products from new biofortified cassava varieties. Foods, 8: 177.

Ferri L., Perrin E., Campana S., Tabacchioni S., Taccetti G., Cocchi P., Ravenni N., Dalmastri C., Chiarini L., Bevivino A., Manno G., Mentasti M., Fani R. (2010): Application of multiplex single nucleotide primer extension (mSNuPE) to the identification of bacteria: the *Burkholderia cepacia* complex case. Journal of Microbiological Methods, 80: 251–256.

Flores J.R., Lehtonen S., Hyvönen J. (2021): Exploring the impact of unstable terminals on branch support values in paleontological data. Paleobiology, 47: 432–445.

Ganapathy K.N., Gomashe S.S., Rakshit S., Prabhakar B. (2012): Genetic diversity revealed the utility of SSR markers in classifying parental lines and elite genotypes of sorghum (*Sorghum bicolour* (L.) Moench). Australian Journal of Crop Sciences, 6: 1486–1493.

Hernandez R.D., Uricchio L.H., Hartman K., Ye C., Dahl A., Zaitlen N. (2019): Ultrarare variants drive substantial cis heritability of human gene expression. Nature Genetics, 51: 1349–1355.

Ioannidis J.P.A. (2009): Population-wide generalizability of genome-wide discovered associations. Journal of the National Cancer Institute, 101: 1297–1299.

Ilona P., Bouis H.E., Palenberg M., Mourisi M., Oparinde A. (2017): Vitamin A cassava in Nigeria: Crop development and delivery. Chapter 9. African Journal of Food, Agriculture, Nutrition and Development, 17: 12000–12025.

Kamara L., Kassoh F.A., Kamara E.G., Masssaquoi F.B., Yila K.M., Contech A.R. (2018): Evaluation of improved cassava genotypes for storage root yield and post traits in the Njala soil series. Journal of Interdisciplinary Biosciences, 2: 12–20.

Kanaka K.K., Sukhija N., Goli R.C., Singh S., Ganguly I., Dixit S.P., Dash A., Malik A.A. (2023): On the concepts and measures of diversity in the genomics era. Current Plant Biology, 33: 100278.

Karim K.Y., Ifie B., Dzidzienyo D., Danquah E.Y., Blay E.T., Whyte J.B.A., Kulakow P., Rabbi I., Parkes E.Y., Omoigui L., Norman E.P., Iluebbey P. (2020): Genetic characterization of cassava (*Manihot esculenta* Crantz) genotypes using agro-morphological and single nucleotide polymorphism markers. Physiology and Molecular Biology of Plants, 26: 8.

Kawuki R., Ferguson M., Labuschagne M., Herselman L., Kim D.J. (2009): Identification, characterization, and application of single nucleotide polymorphisms for diversity assessment in cassava (*Manihot esculenta* Crantz). Molecular Breeding, 23: 669–684.

Kongsil P., Ceballos H., Siriwan W., Vuttipongchaikij S., Kittipadakul P., Phumichai C., Wannarat W., Kositratana W., Vichukit V., Sarobol E., Rojanaridpiched C. (2024): Cassava breeding and cultivation challenges in Thailand: Past, present, and future perspectives. Plants, 13: 1–29.

Machado R., Fernandes A.M., Leonel M., Figueiredo R.T.De., Antonio M., Rangel S., Ringenberg R., Oliveira

L.A.De, Santos S., Vieira E.A. (2023): Measurement of dry matter and starch in modern cassava genotypes during long harvest cycles. Horticulturae, 9: 733.

Makueti T.J., Gordon O., Zac T., Alice M., Alain T., Ebenezer A., Robert K. (2015): Genetic diversity of *Dacryodes edulis* provenances used in controlled breeding trials. Journal of Plant Breeding and Crop Sciences, 7: 327–339.

Mohidin S.R.N.S.P., Moshawih S., Hermansyah A., Asmuni M.I., Shafqat N., Ming L.C. (2023): Cassava (*Manihot esculenta* Crantz): A systematic review for the pharmacological activities, traditional uses, nutritional values, and phytochemistry. Journal of Evidence-Based Integrative Medicine, 28: 1.

Mtunguja M.K., Beckles D.M., Laswai H.S., Ndunguru J.C., Sinha N.J. (2019): Opportunities to commercialize cassava production for poverty alleviation and improved food security in Tanzania. African Journal of Food Agriculture, Nutrition and Development, 19: 13928–13946.

Nyaika J. (2024): Cyanide in cassava : Understanding the drivers, impacts of climate variability, and strategies for food security. Food and Energy Security, 13: e573.

Ogbonna A.C., De Andrade L.R.B., Rabbi I.Y., Mueller L.A, De Oliveira E.J., Bauchet G. (2020): A large-scale genome-wide association study, using historical data, identifies the conserved genetic architecture of cyanogenic glucoside content in cassava (*Manihot esculenta* Crantz) root. Plant Journal, 105: 754–770.

Ogbonna A.C., De Andrade L.R.B., Mueller L.A, De Oliveira E.J., Bauchet J.G. (2021): Comprehensive genotyping of Brazilian cassava (*Manihot esculenta* Crantz) germplasm bank: Insights into diversification and domestication. Theoretical and Applied Genetics, 134: 1343–1362.

Okogbenin E., Egesi C.N., Olasanmi B., Ogundapo O., Kahya S., Hurtado P., Marin J., Akinbo O., Mba C., Gomez H., De Vicente C., Baiyeri S., Uguru M., Ewa F., Fregene M. (2012): Molecular marker analysis and validation of resistance to cassava mosaic disease in elite cassava genotypes in Nigeria. Crop Science, 52: 2576–2586.

Olasanmi B., Kyallo M., Yao N. (2021): Marker-assisted selection complements phenotypic screening at the seedling stage to identify cassava mosaic disease-resistant genotypes in African cassava populations. Scientific Reports, 11: 2850.

Oliveira E.J., De Ferreira C.F., Santos V.S., Jesus O.N., Oliveira G.A.F., Silva M.S. (2014): Potential of SNP markers for the characterization of Brazilian cassava germplasm. Theoretical and Applied Genetics, 127: 1423–1440.

Ousmael K., Whetten R.W., Xu J. (2023)*:* Identification and high-throughput genotyping of single nucleotide polymorphism markers in a non-model conifer (*Abies nordmanniana* (Steven) Spach). Scientific Reports, 13: 22488.

Paradis E., Claude J., Strimmer K. (2004): APE: Analyses of phylogenetics and evolution in R language. Bioinformatics, 20: 289–290.

Prakash J., Herlin M., Kumar J., Garg G., Åkesson K.E., Grabowski P.S., McGuigan F.E. (2019): Analysis of *RAMP3* gene polymorphism with body composition and bone density in young and elderly women. Gene, 721: 100009.

Prempeh R.N., Manu-Aduening J., Quain M.D., Asante I.K., Samuel Kwasi Offei S.K., Danquah E.Y. (2020): Assessment of genetic diversity among cassava landraces using single nucleotide polymorphic markers. African Journal of Biotechnology, 19: 383–391.

R Core Team (2020): R: A language and environment for statistical computing. Vienna, R Foundation for Statistical Computing. Available at https://www.R-project.org/

Rabbi I., Hamblin M., Gedil M., Kulakow P., Ferguson M., Kpan I., Delphine A.S., Jean-Luc J. (2014): Genetic mapping using genotyping-by-sequencing in the clonally propagated cassava. Crop Science, 54: 1384–1396.

Rabbi I.Y., Kayondo S.I., Bauchet G., Yusuf M., Aghogho C.I., Ogunpaimo K., Uwugiaren R., Smith I.A., Peteti P., Agbona A., Parkes E., Ezenwaka L., Wolfe M., Jannink J-L., Egesi C., Kulakow P. (2020): Genome-wide association analysis reveals new insights into the genetic architecture of defensive, agro-morphological, and quality-related traits in cassava. Plant Molecular Biology, 109: 195–213.

Ren J., Sun D., Chen L., You F.M., Wang J., Peng Y., Nevo E., Sun D., Luo M.C., Peng J. (2013): Genetic diversity was revealed by single nucleotide polymorphism markers in a worldwide germplasm collection of durum wheat. International Journal of Molecular Sciences, 14: 7061–7088.

Suzuki R., Shimodaira H. (2006): Pvclust : An R package for assessing the uncertainty in hierarchical clustering. Bioinformatics, 12: 1540–1542.

Tiago A.V., Rossi A.A.B., Tiago P.V., Carpejani A.A., Silva B.M., Hoogerheide E.S.S., Yamashita O.M. (2016): Genetic diversity in cassava landraces grown on farms in Alta Floresta-MT, Brazil A.V. Genetics and Molecular Research, 15: 1–10.

Tumuhimbise R., Shanahan P., Melis R., Kawuki R. (2014): Combining ability analysis of storage root yield and related traits in cassava at the seedling evaluation stage of breeding. Journal of Crop Improvement, 28: 530–546.

Udoh L.I., Maleku G., Parkes Y. E., Kulakow P., Adesoye A., Nwuba C., Rabbi Y. I. (2017): Candidate gene sequencing and validation of SNP markers linked to carotenoid content in cassava (*Manihot esculenta* Crantz). Molecular Breeding, 37: 123.

Varshney R.K. (2010): Gene-based marker systems in plants: High throughput approaches for discovery and genotyping. Molecular Techniques in Crop Improvement, 2: 119–142.

Wang H.Z., Wu Z.X., Lu J.J., Shi N.N., Zhao Y., Zhang Z.T., Liu J.J. (2009a): Molecular diversity and relationships among *Cymbidium goeringii* cultivars based on inter-simple sequence repeat (ISSR) markers. Genetica, 136: 391–399.

Wang H.-C., Susko E., Roger A.J. (2016): Split-specific bootstrap measures for quantifying phylogenetic stability and the influence of taxon selection. Molecular Phylogenetics and Evolution, 105: 114–125.

Wang S., Meyer E., McKay J.K., Matz M.V. (2009b): 2b-RAD: A simple and flexible method for genome-wide genotyping. PLoS ONE, 7: e37137.

Welsch R., Arango J., Bär C., Salazar B., Al-Babili S., Beltrán J., Chavarriaga P., Ceballos H., Tohme J., Beyer P. (2010): Provitamin A accumulation in cassava (*Manihot esculenta*) roots is driven by a single nucleotide polymorphism in a phytoene synthase gene. Plant Cell, 22: 3348–3356.

Wolfe M.D., Rabbi I.Y., Egesi C., Hamblin M., Kawuki R., Kulakow P., Lozano R., Del Carpio D.P., Ramu P., Jannink J. (2016): Genome-wide association and prediction reveal genetic architecture of cassava mosaic disease resistance and prospects for rapid genetic improvement. Plant Genome, 9: 1–13.

Xiao L., Cheng D., Ou W., Chen X., Rabbi I.Y., Wang W., Li K., Yan H. (2025): Advancements and strategies of genetic improvement in cassava (*Manihot esculenta* Crantz): From conventional to genomic approaches. Horticulture Research, 12: uhae341.

Zhang S., Chen X., Lu C., Ye J., Zou M., Lu K. (2018): Genome-wide association studies of 11 agronomic traits in cassava (*Manihot esculenta* Crantz). Frontiers in Plant Science, 9: 503.

Zhang X., Perez-Rodriguez P., Semagn K., Beyene Y., Babu R., Lopez-Cruz M.A.F., San Vicente O.M., Buckler E., Jannink J-L., Prasanna B.M., Crossa J. (2015): Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. Heredity, 114: 291–299.

Received: December 3, 2025
Accepted: March 5, 2026
Published online: March 23, 2026