International Journal of Agriculture and Biosciences 2025 14(3): 461-472.
https://doi.org/10.47278/journal.ijab/2025.025

RESEARCH ARTICLE eISSN: 2306-3599; pISSN: 2305-6622

Unveiling Cassava Diversity: Agronomic and SNP Marker Trait Profiling in the Democratic Republic of Congo

Mouritala Sikirou [1,2,*], Paterne A. Agre [3], Pierre Miafuntila [4], Najimu Adetoro [1], Eric Musungayi [4], Aboudou Hack Arouna [2], Chédrac Zokpon [2], Tony Bakelana [4], Abdul-Rasaq Adebowale [1] and Roland Bocco [5]

[1]International Institute of Tropical Agriculture (IITA), Democratic Republic of Congo
[2]National University of Agriculture (UNA), School of Horticulture and Green Landscaping (EHAEV), Ketou, P.O. Box 043, Benin
[3]IITA, Ibadan, Nigeria
[4]Institut National pour l'Étude et la Recherche Agronomiques (INERA), Mvuazi, Democratic Republic of Congo
[5]Texas AgriLife Research and Extension Center, Texas A&M University System, 1509 Aggie Dr., Beaumont, TX, 77713, USA
*Corresponding author: M.Sikirou@cgiar.org

ABSTRACT

Cassava is a perennial crop known for its adaptability to diverse agro-ecological conditions. In many tropical countries, including the Democratic Republic of Congo (DRC), it is a key contributor to food security, making it the most important staple food. Nearly 90% of cassava roots and leaves harvested in the DRC are consumed by humans, contributing significantly to the population's energy, protein, and micronutrient intake. This study evaluated 136 cassava varieties, including nine improved varieties developed by IITA and INERA-DRC along with 120 local varieties collected from various agro-ecological zones and maintained by INERA. Conducted at INERA Mvuazi in Kongo Central, DRC, the trial used an alpha lattice design with eight blocks and 17 plots replicated twice. Key agronomic traits such as CMD, plant height, biomass, harvest index, dry matter content, starch, and fresh root yield were profiled. Significant phenotypic variation was observed, particularly in plant height (2727.79cm), biomass (55.05kg), and yield (44.24t/ha), with high heritability estimates for traits like starch content (0.55) and biomass (0.98). Hierarchical clustering identified five distinct groups with varying traits and performance, highlighting the diversity within the germplasm. Genotypic analysis using SNP markers revealed substantial genetic diversity, with population structure analysis identifying four genetic clusters. The integration of phenotypic and genotypic data provided a comprehensive understanding of the relationships within the germplasm. The study identified 47 top-ranking cassava varieties with exceptional multi-trait performance, suitable for breeding programs aimed at improving yield, disease resistance, and dry matter content. This research offers valuable insights into the genetic and phenotypic diversity of cassava germplasm in the DRC, supporting future breeding efforts.

Keywords: Cassava germplasm, Agronomic Traits, SNP Markers, Hierarchical Clustering, Genetic Diversity.

Article History
Article # 25-009
Received: 10-Jan-25
Revised: 27-Jan-25
Accepted: 05-Feb-25
Online First: 11-Feb-25

INTRODUCTION

Cassava is a tuberous and starchy crop and the most grown in the world after cereals such as rice, wheat, and maize (Katz & Weaver, 2020; Adebayo, 2023). It plays an important role in the daily dishes of more than billions of people around the world and remains a crucial staple food of the Sub-Saharan African population, where more than 800 million rely on the cassava crop (Howeler et al., 2013; Munyahali et al., 2017). Generally, people process cassava to produce industrial products like ethanol, starchy, and high-quality cassava flour and starch, as well as several traditional products like lafun, fufu, gari, tapioca, chikwangue, and attieke (Adebayo & Silberberger, 2020; Adebayo, 2023). In addition, cassava offers the opportunity to use its leaves as stews, which leads to a high

Cite this Article as: Sikirou M, Agre PA, Miafuntila P, Adetoro N, Musungayi E, Arouna AH, Zokpon C, Bakelana T, Adebowale AR and Bocco R, 2025. Unveiling cassava diversity: Agronomic and SNP marker trait profiling in the Democratic Republic of Congo. International Journal of Agriculture and Biosciences 14(3): 461-472. https://doi.org/10.47278/journal.ijab/2025.025

A Publication of Unique Scientific Publishers

contribution to food security and potential income generation for farmers in Sub-Saharan Africa (Hauser et al., 2020; Malik et al., 2020).

In the Democratic Republic of Congo (DRC), the second-largest producer of cassava in the world after Nigeria, with an average production of 45 MT per year, cassava root and leaf are known as the primary food of more than 90 million Congolese. Cassava production significantly increased during the last decade with population growth, particularly with the COVID-19 crisis, when the wheat flour became rare in the African market. It is a famine reserve crop and cash crop that serves as a buffer against household food crises (Burns et al., 2010; FAOSTAT, 2022). Adebayo (2023) has demonstrated that the high cassava production is due to the significant increase in the harvested area. The yield in Africa remains relatively low, averaging 8.0 t and varying from 5.7 to 9.6 t, depending on the variety.

It is noted that cassava yield production is low in Africa, particularly in the DRC, and it is far below the world average of about 30t.ha^{-1} in Asia and Latin America (FAOSTAT, 2022; Malik et al., 2020). Even though cassava grows in rough terrain and there aren't many breeding programs for it, abiotic and biotic stresses make it hard to grow (Legg et al. 2011; Kintché et al. 2017; Bakelana et al. 2020).

Since 1933, the DRC has primarily driven its cassava breeding program by introducing advanced clones and implementing release-based projects in each province to address occasional issues (SENASEM, 2019). It has resulted in low adoption until recent years due to ecological diversity and the emergence of new diseases and pests across the country (IITA, 2024). To overcome these challenges, there is a need to establish a strong cassava program with a thorough knowledge of the diversity of the available germplasm and mostly local cultivars by characterizing and breeding in the local conditions for sustainability (Pierre et al., 2022; Amelework et al., 2021). We conducted this exercise across various countries and crops to bolster our breeding programs. The study of the yam cultivar collection in Benin (Agre et al., 2021) revealed two groups that could allow for their conservation before being replaced by improved and introduced varieties. In the meantime, yam landraces in Nigeria showed a lot of potential for crossing. They could be used as donors to improve many traits and target desirable genes of interest in Africa (Agre et al., 2023) using SNP markers along with phenotypic traits and socio-cultural factors (Loko et al., 2015). Researchers reported a similar trend when screening the entire germplasm of African rice (*Oryza glaberrima*) in various West African countries for resistance genes to abiotic stresses (Sikirou et al., 2018; Shaibu et al., 2018; Agbeleye et al., 2019), with genomic studies highlighting the underutilized potential in Africa and the domestication pathway (Wambugu et al., 2019). Additionally, researchers had made several attempts to uncover the genetic diversity of cassava, especially in Africa. In 2013, Kawuki et al., (2013) reported the genetic diversity among the cassava cultivars in central, eastern, and southern Africa after genotyping by collecting leaves for microsatellite analysis. During the last decade, single nucleotide polymorphisms (SNPs) remained the most used to approach the genetic variations in cassava germplasm (Nkhoma et al., 2020; Adu et al., 2021). Recently, SNP markers were used to find small genetic differences in Burundi's cassava germplasm. They described 118 cultivars based on their morphological traits and fresh yield roots (Pierre et al., 2022). During these years, Cassava Mosaic Disease (CMD) devastated the crop. The disease eroded the country's genetic diversity by wiping out the existing varieties on the ground, primarily because they were susceptible to the Ugandan variant's virus (Mahungu et al., 1994; Faostat, 2022).

Several attempts were made to restore the cassava germplasm in the DRC, with many clones introduced from West and East Africa and the local cultivar collected from farmers' fields during the socio-economic survey (Bisimwa et al., 2019). Therefore, this study aims to assess the genetic diversity of the current cassava germplasm in DRC based on phenotypic traits and fresh yield root, combining it with SNP markers through Diversity Arrays Technology and genotyping-by-sequencing (DArTseq). The goal is to identify the relationships within the germplasm to ensure its appropriate conservation and use in improving the cassava breeding pipelines, such as those producing white and yellow roots and leafy vegetables.

MATERIALS & METHODS

Genetics Materials

This study used a total of 136 cassava varieties, including nine improved varieties and 120 local varieties. The improved cassava ones were developed by the IITA and INERA-DRC, while the local cassava varieties were collected across different agro-ecological zones and are being maintained by the INERA.

Trials Design and Data Collection

A total of 136 genotypes were evaluated during the cropping season 2020-21 to assess the performance of DRC germplasm collection. The trial was established in INERA Mvuazi, Kongo Central, in the Democratic Republic of Congo, geographically located at Alt 990m, Lat: S 02°47.664', Long: E 028°58.137'. The trial was designed using an alpha lattice and laid out in 8 blocks with 17 plots (entries) replicated twice.

Two-row plots with five plants spaced one meter apart and within rows were used for each genotype in the trial. No fertilizer, irrigation, or pesticide was applied and weeds were managed throughout the growing period. The genotypes were profiled using key agronomic traits such as cassava mosaic disease (CMD), plant height (Phe), biomass, harvest index (HI), dry matter content (DMC), starch, and fresh yield root (Yield). Data were collected at different plant growth stages. The agronomic data were recorded following the protocol described by IITA (1990), while the CMD severity was assessed according to Legg et al. (2015), as detailed below in Table 1.

Table 1: Phenotypic traits used for cassava profiling

Traits	Code	Sampling period	Trait Nature	methods
Cassava Mosaic Disease	CMD	06 MAP	Qualitative	Measure using the scale 1 to 5, 1 means no infection, 2 is mild infection, 3 is moderate infection, 4 is severe infection, and 5 is very severe infection.
Harvest index	HI	12 MAP	Quantitative	Ratio of the weight of the fresh roots (marketable and non-marketable) to the total quantity of biomass.
Biomass	biomass	12 MAP	Quantitative	It is a measurement of the aboveground of the shoot portion including leaves, stems and branches in each plot.
Plant height	Phe	12 MAP	Quantitative	The measurement was taken using a graduated slat placed along the stem from the ground of the soil to the top of the plant.
Dry matter content	DMC	12 MAP	Quantitative	Dry 100g of cassava fresh root at 70 °C for 72 hours. The dried sample was weighed, and the root dry matter content (percentage) was calculated as the ratio between fresh weight and dry weight (Ekanayake, 1996).
Starch	Starch	12 MAP	Quantitative	It is a linear regression between the specific gravity and the dry matter content as described by Safo-Kantanka & Owusu-Nipah (1992)
Yield	Yield	12 MAP	Quantitative	It implies the weighing in tons of all recovered storage roots in the net plot at the harvesting.

Genotypic Study

In this study, we genotyped a total of 136 cassava using the quality control SNP markers developed by IITA-Ibadan. Cassava leaves were collected from young plants as discs and then shipped to Intertek Swaziland for DNA extraction and genotyping.

Statistical Analysis

Analysis of variance was performed to determine differences among the accessions for the various traits using the variability R package (Popat & Porma, 2020).

The generated LSmeans for the 136 varieties were then used for principal component analysis in the FactorMiner and MissMDA R packages (Josse & Husson, 2016). The optimal number of factors to be retained was determined using the dimdesc function in R, and the contribution of each trait was determined using the principle of Peres-Neto et al. (2003). Each trait's contribution was determined using Peres-Neto et al. (2003) principle. A dissimilarity matrix was generated using the Gower method implemented in cluster R packages (Maechler et al., 2023). The final hierarchical cluster was then performed using the ward.D2 method in the cluster R package 42. Correlation among the different phenotypic variables was performed using the R software and the result was displayed as a heatmap.

Genotypic Data Assessment

The entire population was genotyped using 40 sets of Kompetitive alleles-specific markers developed by the cassava breeding unit at the International Institute of Tropical Agriculture, IITA-Ibadan. Summary statistics, including allele frequencies and heterozygosity for both SNP markers and genotypes, were calculated using PLINK 1.9 version. Population structure was inferred using Admixture, with the optimal number of clusters determined by k-means clustering. Discriminant analysis of principal components (DAPC) was applied to identify distinct groups, and membership probabilities were assigned based on a 70% cutoff. Genetic differentiation between groups was assessed using Fst and pairwise genetic distances calculated with VCFtools and Phylentropy, respectively. Finally, a hierarchical clustering dendrogram was constructed using Ward's method to visualize population relationships.

Joint Analysis of Phenotypic and Molecular Data

A joint dissimilarity matrix was constructed by combining phenotypic and genotypic distance matrices to integrate phenotypic and genotypic information. Hierarchical clustering was then performed using the combined matrix of the phenotypic and genotypic data. The resulting dendrograms were compared using tanglegram analysis to assess their similarity.

Identification of Top-ranking Cassava Varieties

The total genetic values (BLUP) were used to calculate a multi-trait index based on the factor analysis and ideotype-design (FAI-BLUP) index using the Metan package (Olivoto & Nardino, 2020) implemented in R. The FAI-BLUP index was used to identify the best cassava varieties (ideotypes) based on multi-traits and free from multicollinearity. A radar chart was generated to visualize the selected cassava varieties using the radar chart function implemented in the Metan R package. The radar chart function also visualized the strengths and weaknesses of the selected cassava ideotypes (Olivoto & Nardino, 2020).

RESULTS

Phenotypic Traits Profiling

The assessment of phenotypic variation within a common population is an important step for breeders in developing a new variety that combines genetic and environmental factors to fulfill the target market. Table 2 displays the traits measured in this study. It also shows illustrates the performance of the genotypes performed and the variations in the genetic parameters for CMD, plant height, biomass, dry matter content, starch, and yield in the cassava germplasm in the. Significant variation was observed in the trait within the germplasm. High phenotypic variation was observed in the traits such as plant height (2727.79), biomass (55.05), and yield (44.24) compared to dry matter content (18.62) and starch content (11.29). The same trend was reported for the genotypic variation. However, there was almost no variation between the harvest index (0.01). High-to-very-high heritability ranged from 0.55 to 0.98 for starch content and biomass, respectively. It is worth noting that the average mean for the CMD is 1.61 and that the germplasm included genotypes with high average yield (37.47), dry matter (56.57%), and starch (41.44%).

The correlation showed that there was a strong positive link between starch and dry matter content. There was also a highly positive link between fresh yield root,

Table 2: Summary statistic and phenotypic traits variation

Source of variation	CMD	Phe (cm)	HI	Biomass	DMC (%)	Starch (%)	Yield (t/ha)
Min	1	41.0	0.07	0.43	18.80	7.47	2.6
Max	5	200.50	0.94	43.82	56.57	41.44	77.73
Mean	1.61	115.97	0.47	12.05	35.29	19.25	37.46
Genotypic Variance	0.75	2456.24	0.01	54.05	11.06	6.28	43.54
Phenotypic Variance	1.02	2727.79	0.01	55.05	18.62	11.29	44.24
* Heritability	0.76	0.90	0.96	0.98	0.59	0.55	0.96

Min: Minimum, Max: Maximum, CMD: Cassava mosaic disease; Phe: Plant height; HI; Harvest index; DMC: Dry matter content, *: Broad sense

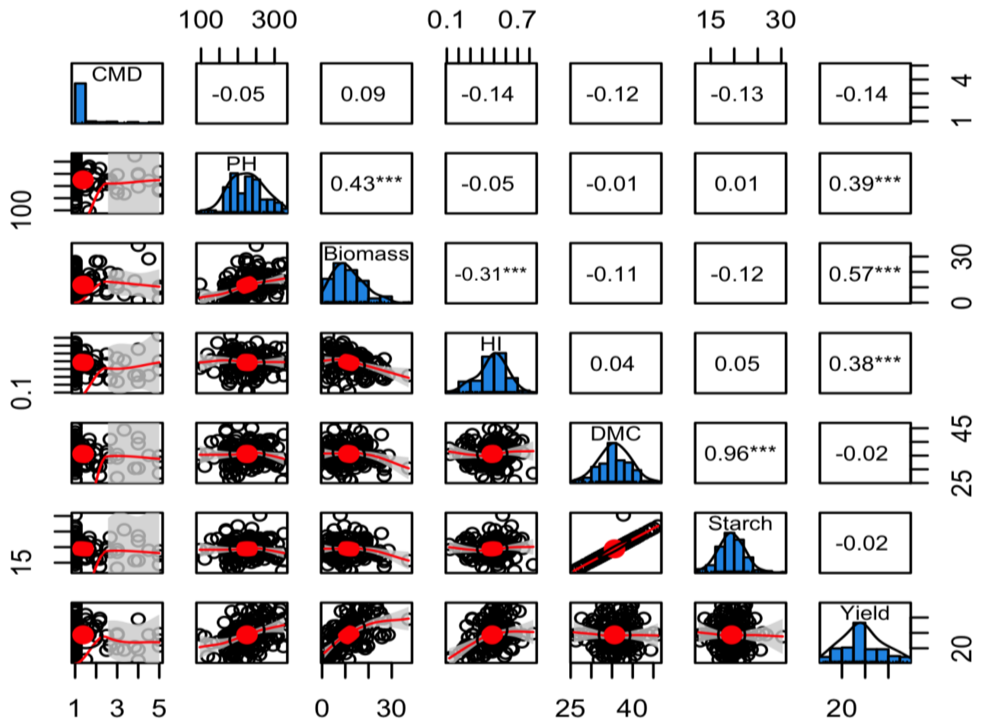
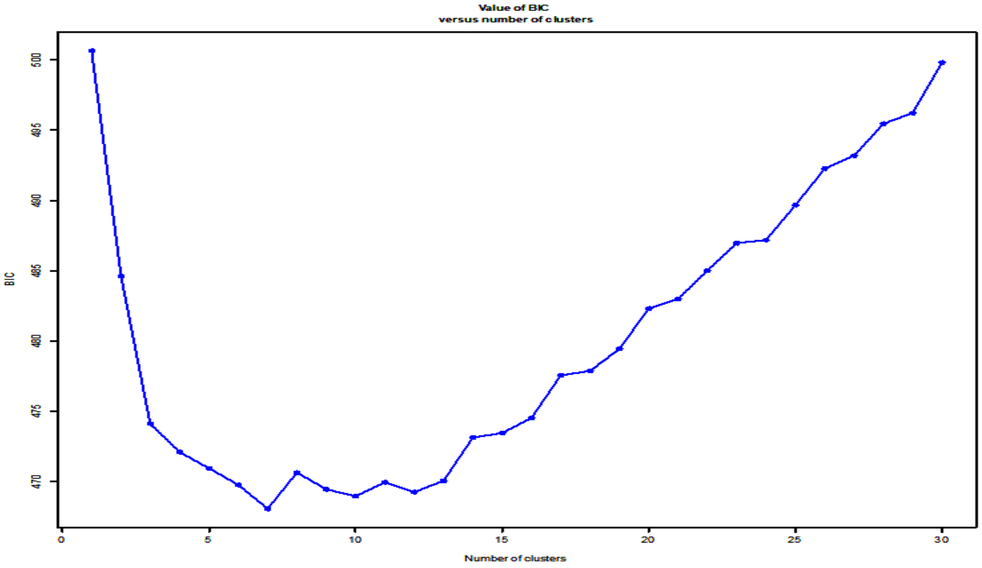


Fig. 1: Traits correlation analysis.



Supplementary Fig. 1: Bayesian Information Criteria displaying the maximum clustering.

plant height, harvest index, and biomass. However, a negative relationship was observed between the yield, CMD, dry matter content, and starch (Fig. 1). Overview: the CMD was negatively correlated with all the traits measured in this study except with biomass.

The principal component analysis showed that the first three components varied by 75.69, with 29.80 and 56.57 contributing to the variation from the first to the second component (Table 3). Traits like dry matter and starch content were identified to be highly positive (0.83) and associated with the first component. In contrast, yield, biomass, plant height, starch, and dry matter content were associated with the second component. The third component was characterized negatively by the harvest index but positively associated with CMD.

In addition to the principal component analysis with the traits, hierarchical clustering was applied to group individuals based on their similarities in this study (Fig. 2). We identified five distinct groups with different traits and performance. Group 1 comprised the smallest subpopulation, and it was constituted by the introduction of cassava clones, including DSC260, which was recently introduced by DSMZ for the resistance to CBSD. Group 2 is mostly made up of biofortified and released white cassava varieties, such as Pululu, which was chosen for its resistance to CAD. It made up 23.52% of the total population. Group 3 (27.94%) had the most cultivated varieties (OBAMA and OBAMA1) in DRC, as well as the wild cassava *M. glaziovii* (GRAZ20210005). Lastly, we have Group 5, with varieties like Zizila, known for its long storage capacity, and Litoy, with its high yield and high dry matter.

Table 3: Principal component analysis with traits contribution on the first six components

Traits	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Dim.6
CMD	-0.19	-0.26	0.42	0.83	0.10	-0.03
PH	-0.37	0.62	0.16	-0.11	0.65	0.01
Biomass	-0.58	0.58	0.38	0.01	-0.33	0.22
HI	0.13	0.185	-0.87	0.36	0.10	0.18
DMC	0.83	0.47	0.22	0.08	-0.03	-0.01
Starch	0.83	0.48	0.20	0.085	-0.01	0.06
Yield	-0.39	0.76	-0.33	0.17	-0.24	-0.23
Eig	2.08	1.87	1.33	0.89	0.62	0.14
%variance	29.80	26.77	19.12	12.73	8.95	2.03
Cumulative (%)	29.80	56.57	75.69	88.43	97.38	100

Genetic Diversity Assessment

A total of 36 SNP markers were used to access the genetic summary statistics, such as the polymorphism information content (PIC), the observed and expected heterozygosity, and the minor allele frequency (MAF). The expected heterozygosity varies from 0.04 to 0.5, while the observed heterozygosity varies from 0.3 to 0.5 (Fig. 3). For the polymorphism information content and the allele frequency, it was reported to vary from 0.05 to 0.5 (Fig. 3), respectively.

Population Structure and Diversity Assessment using the QC Markers

Using the Bayesian information criteria, k=4 was identified as the optimal cluster (Supplementary Fig. 1). The discriminant analysis showed that the 136 varieties of cassava were not evenly spread across the four genetic groups (Fig. 4). Cluster 1 had 37 cassava varieties of all local landraces collected from the Kongo central regions. In group 2, we identified 46 cassava varieties collected from the different agroecological. A total of 18 cassava landraces were grouped in cluster 3. Cassava varieties of these clusters were all collected from the Kongo Central region. Cluster 4 had a total of 35 cassava varieties. In this cluster, we reported the presence of both landraces and breeding lines. Using the IBS matrix, the genetic distance varies from 0 to 0.42, with the lowest genetic distance reported among several landraces collected from different agroecological zones. In contrast, the highest genetic distance was reported between landraces and breeding lines. The IBS genetic distance revealed several cassava varieties with low genetic distance, potentially leading to their classification as duplicates. Clustering analysis-based IBS classified the 136 cassava varieties into four clusters (Fig. 5). Using the posterior probability and through the admixture analysis, 38 cassava varieties were identified as admixed ones (Fig. 6). We identified the admixt varieties with a low posterior probability. Most of the cassava were identified as homogenic with an ancestry probability greater than 70%. It's worth noting that the admixed cassava varieties were the improved cassava varieties.

Using the Gower dissimilarity matrix generated from the phenotypic data and the Jaccard dissimilarity matrix originated from the SNP markers, few genotypes were properly aligned across the two clusters (Fig. 7).

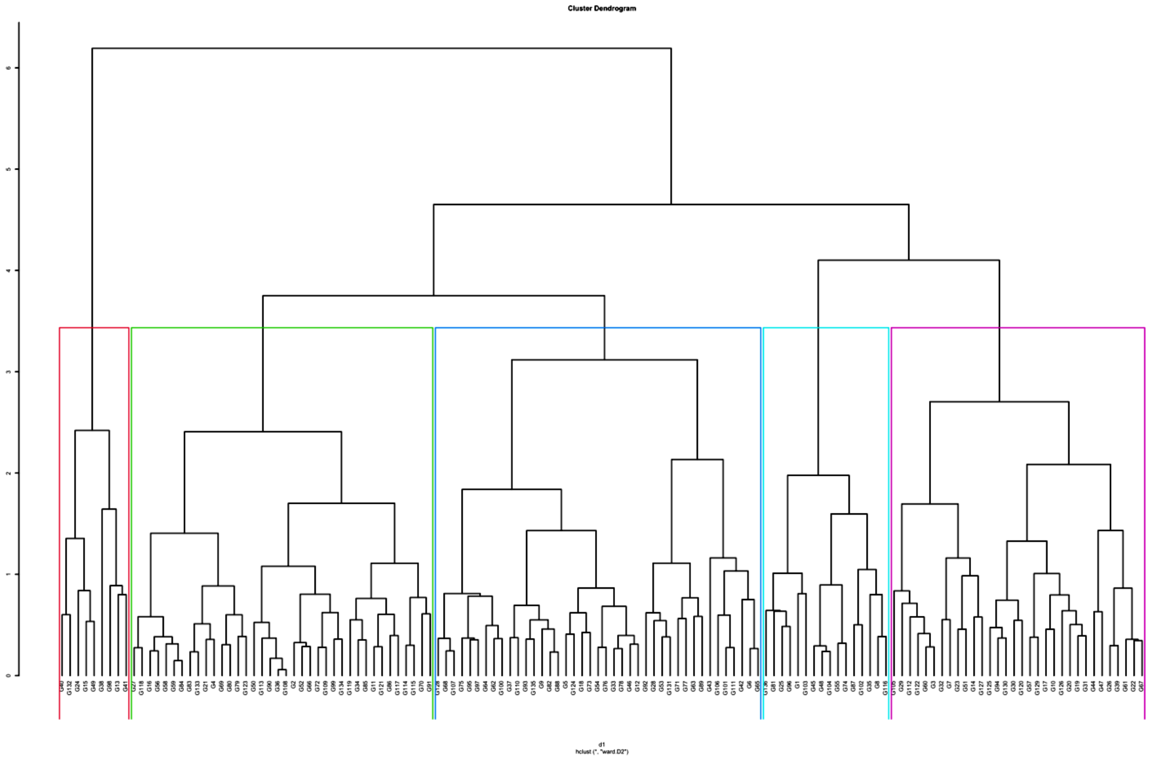


Fig. 2: Hierarchical clustering based on seven phenotypic traits.

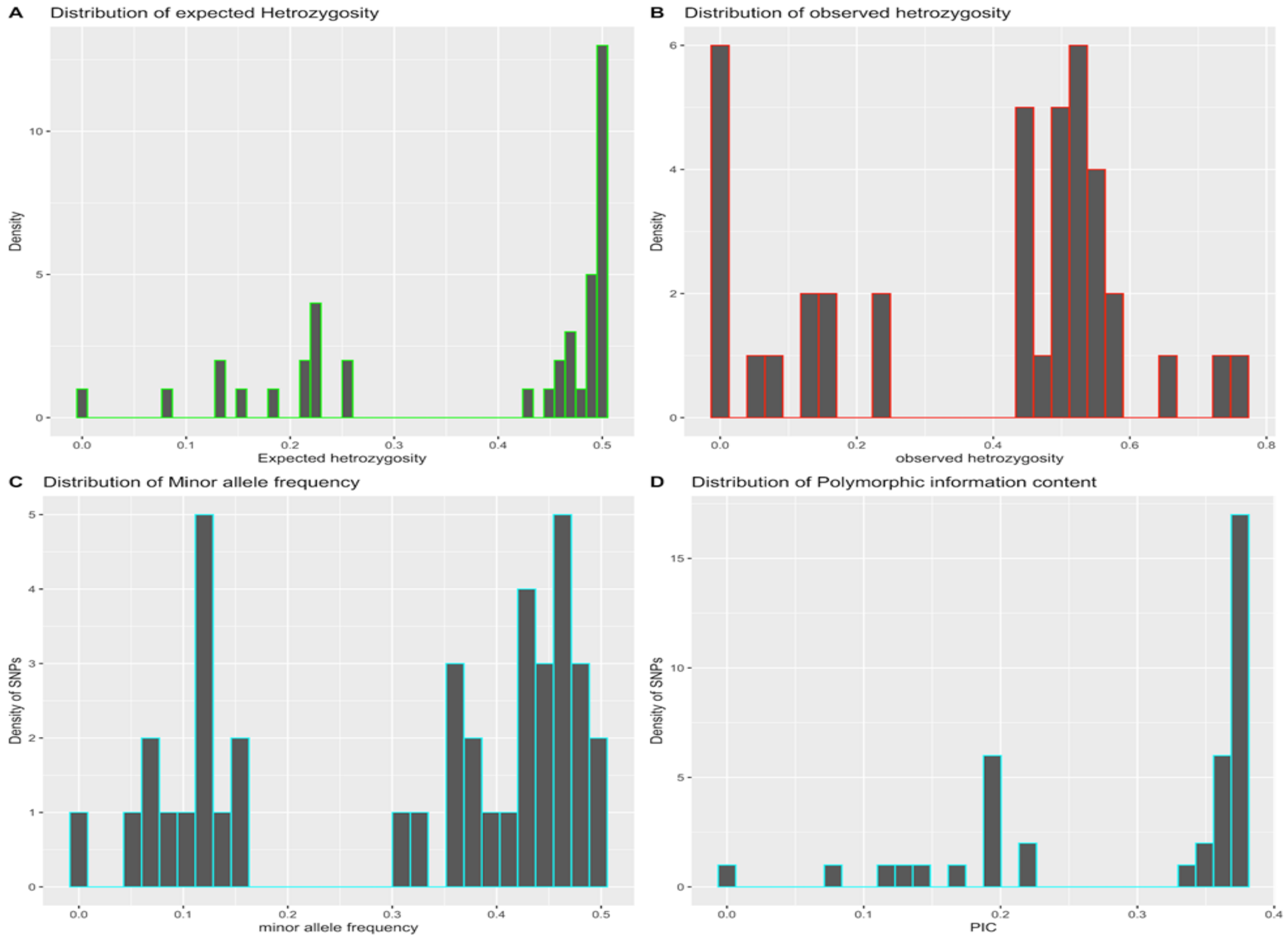


Fig. 3: Genetic statistic summary of the QC markers.

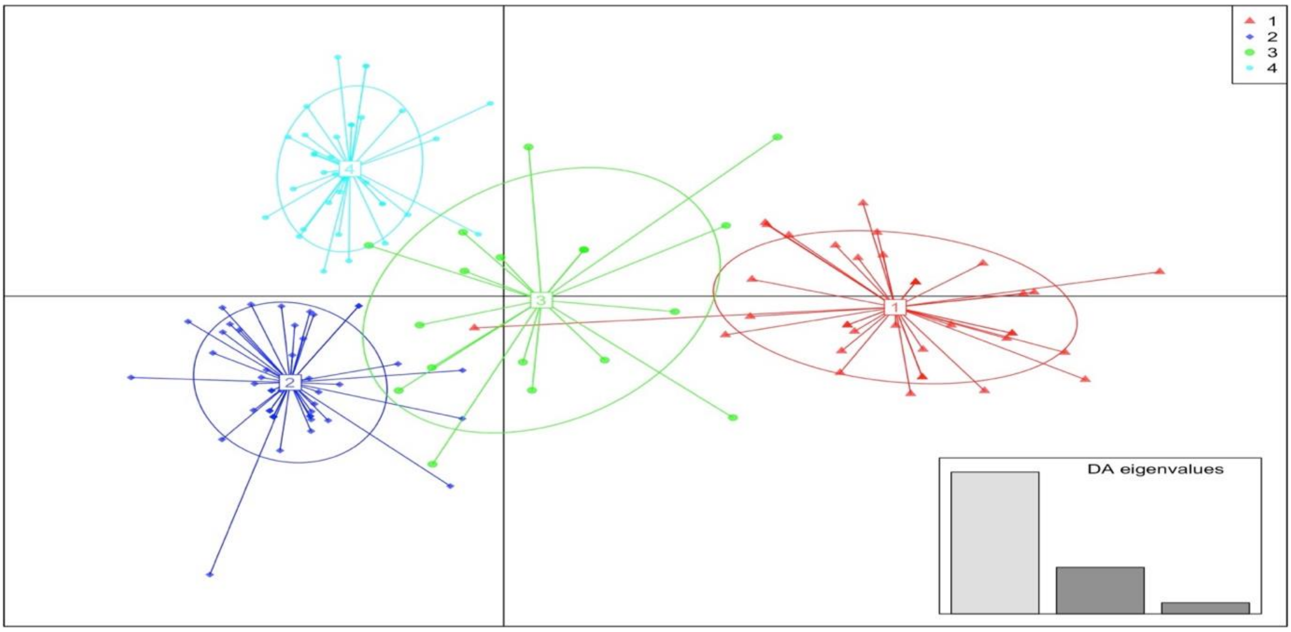


Fig. 4: Discriminant analysis displaying the 4 genetic groups using the QC markers.

Ideotype Cassava Varieties as Progenitors

Using the MGDI technique, we identified 47 cassava varieties as top-ranking out of 136 (Supplementary Fig. 8). The MGDI index ranged from 9 to 14, and of the 136 studied, 47 had low MGDI index values. The 47 top-ranking were chosen for their exceptional multi-trait performance, including high yield, low disease, and high dry matter content with high starch. The selected 47's strengths and weaknesses were presented in a radar plot that accounted for each factor's proportion (Fig. 8).

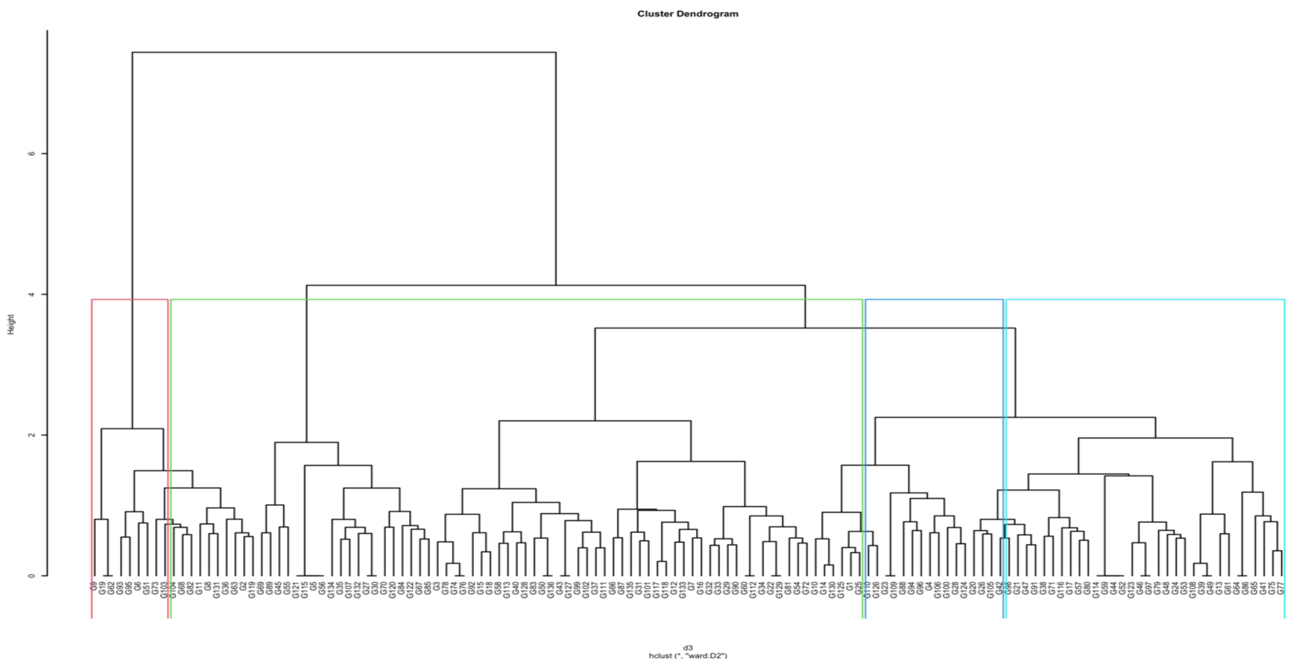


Fig. 5: Hierarchical clustering-based QC SNP markers.

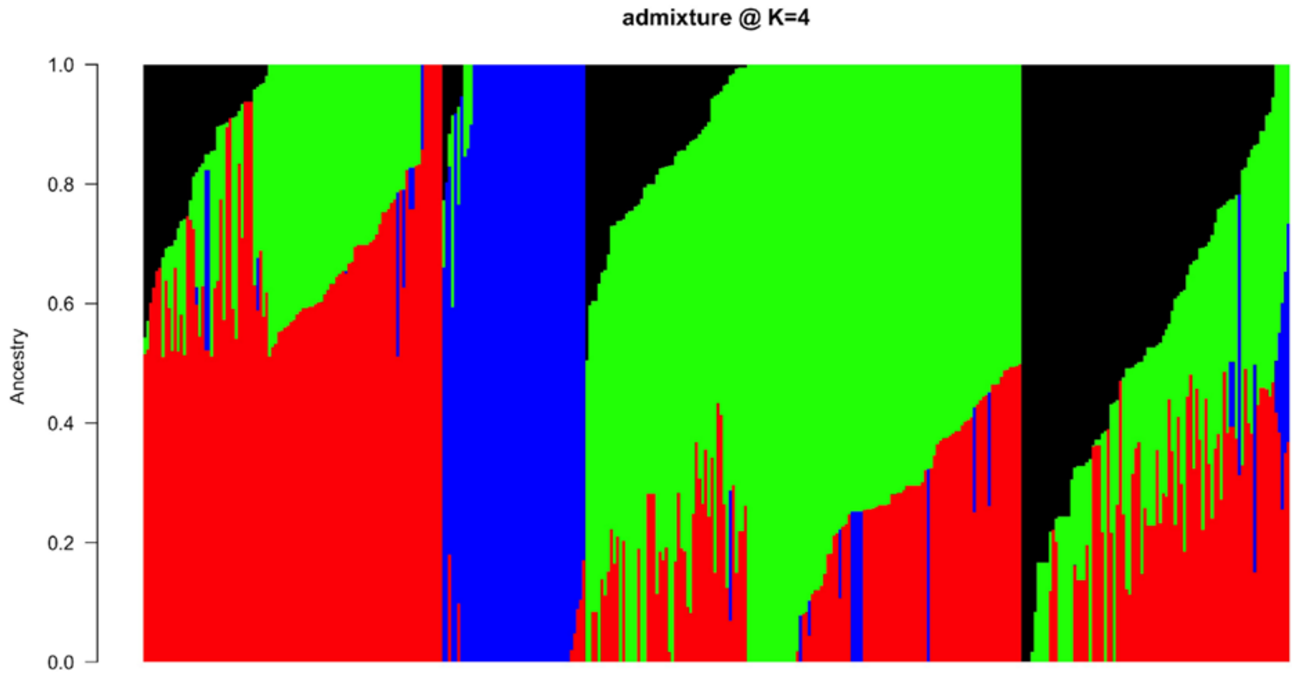


Fig. 6: Population structure using QC SNP markers: Each color stands for different group, while each bar represents a single cassava variety.

DISCUSSION

Understanding the existing genetic resources is fundamental for the long-term sustainability of cassava (Manihot esculenta) breeding programs. Germplasm collections and evaluations play an essential role in conserving genetic resources and improving the development of new varieties that are adapted to target population environments. Genetic diversity provides the raw material for breeders to withstand various stresses by identifying desirable traits such as high yield, disease resistance (to biotic stresses like cassava mosaic disease [CMD], cassava brown streak disease [CBSD] and cassava anthracnose disease [CAD]), abiotic stress resilience (e.g., drought, submergence, cold and soil fertility) and improved root quality (Ceballos et al., 2020; Liang et al., 2024).

Since the 2000s, cassava production in the Democratic Republic of Congo (DRC) has faced significant challenges due to CMD outbreaks, leading to a substantial loss of germplasm. As noted by Mahungu et al. (1994), efforts to restore cassava production primarily relied on germplasm introductions from IITA’s West and East African breeding programs. However, the lack of regular germplasm updates has limited the genetic diversity available to DRC breeding initiatives. This study aims to evaluate the genetic diversity of cassava germplasm in the DRC using phenotypic and molecular data, providing insights for sustainable breeding strategies.

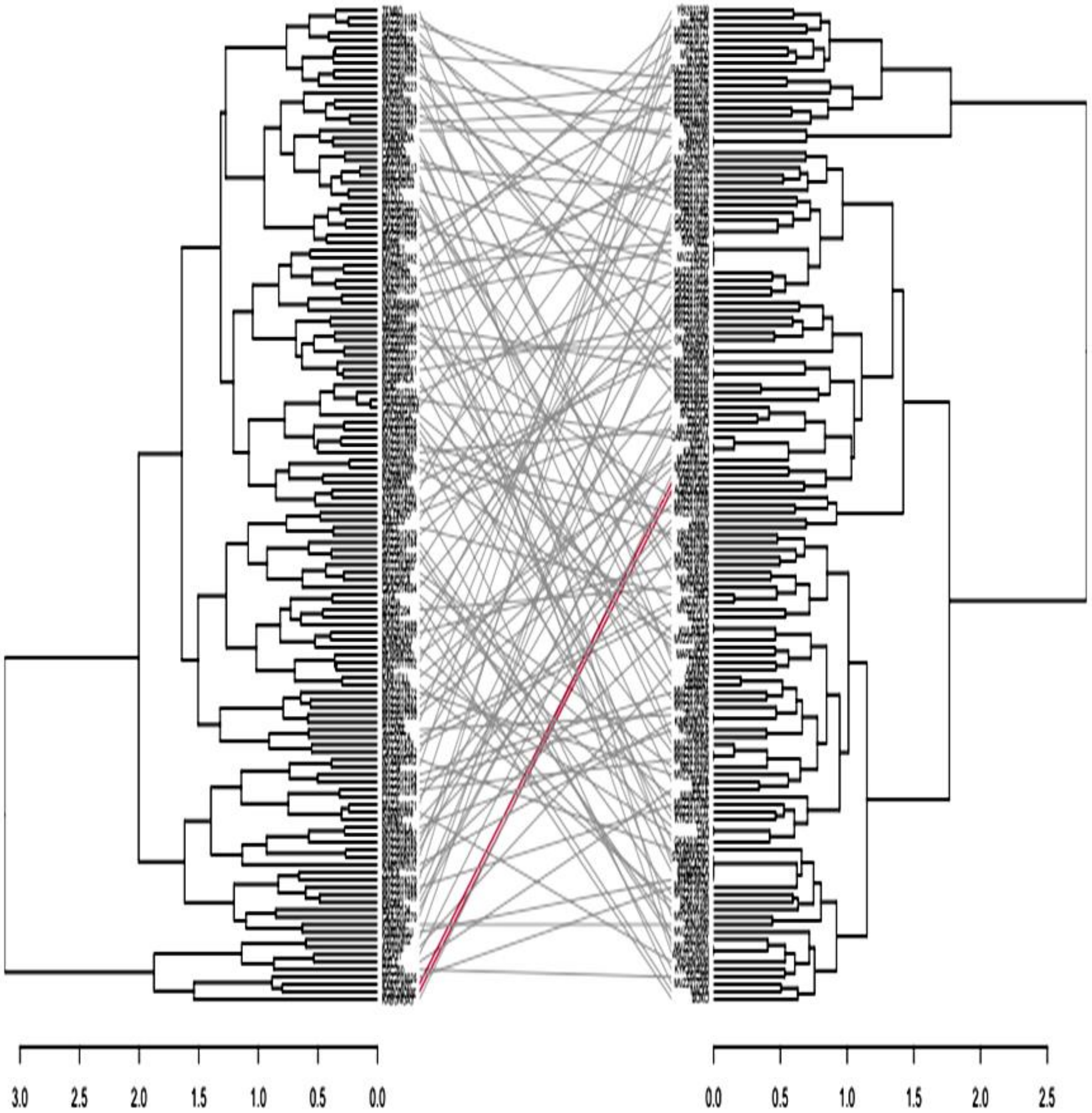


Fig. 7: Phenotypic and genomic comparative cluster showing few genotypes (highlighted red) identified in the same group.

Phenotypic variation is an essential aspect of breeding programs, as it reflects observable traits that influence crop productivity and farmer preferences. In this study, significant variation was observed in key agronomic traits such as fresh root yield, CMD severity, dry matter content (DMC) and starch content. Phenotypic diversity was highlighted by clustering analysis, which identified five distinct groups of cassava varieties. Such diversity underscores the importance of exploring and utilizing locally adapted and farmer-preferred varieties in breeding programs (Simiyu, 2023).

Previous studies have consistently highlighted the critical role of phenotypic diversity in cassava breeding. For example, Elias et al. (2000) demonstrated significant genetic variability in traditional cassava farming systems, while Kamanda et al. (2020) reported similar findings in biofortified cassava cultivars. In this study, cassava varieties exhibited high yields (maximum average yield of 77.73%), high dry matter content (56.57%) and high starch content (41.44%), aligning with the findings of Ceballos et al. (2015), who emphasized the integration of genetic diversity with phenotypic evaluation in cassava improvement.

Importantly, the CMD severity score in this study ranged from 1 to 5, with a median severity of 1.61, indicating that most germplasm exhibited resistance or tolerance to CMD. This finding aligns with similar studies on disease resistance in cassava germplasm. For instance, Liang et al. (2024) reported genetic improvements in CMD resistance in cassava breeding populations, emphasizing the need for regular germplasm evaluation to address emerging disease pressures.

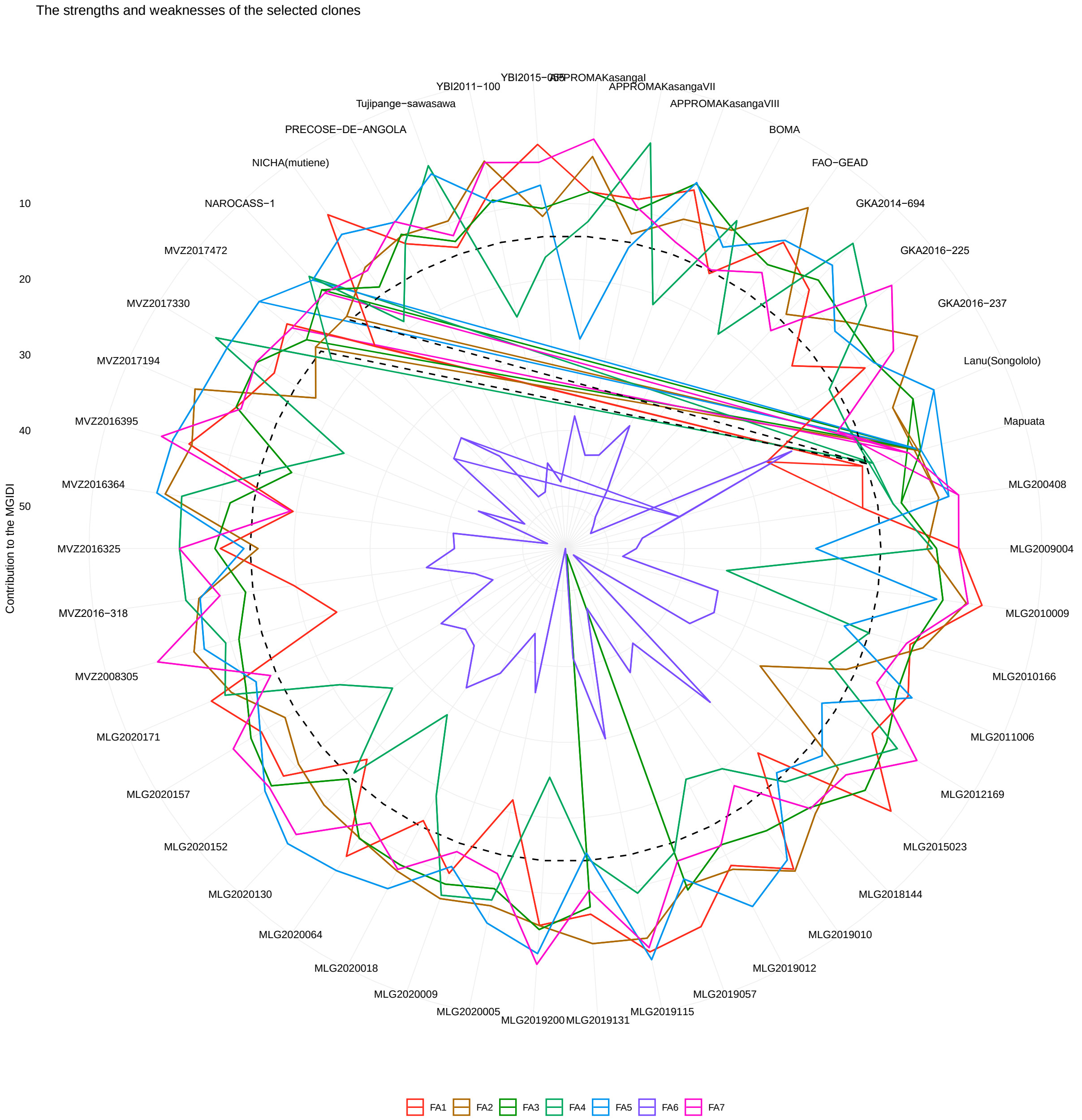
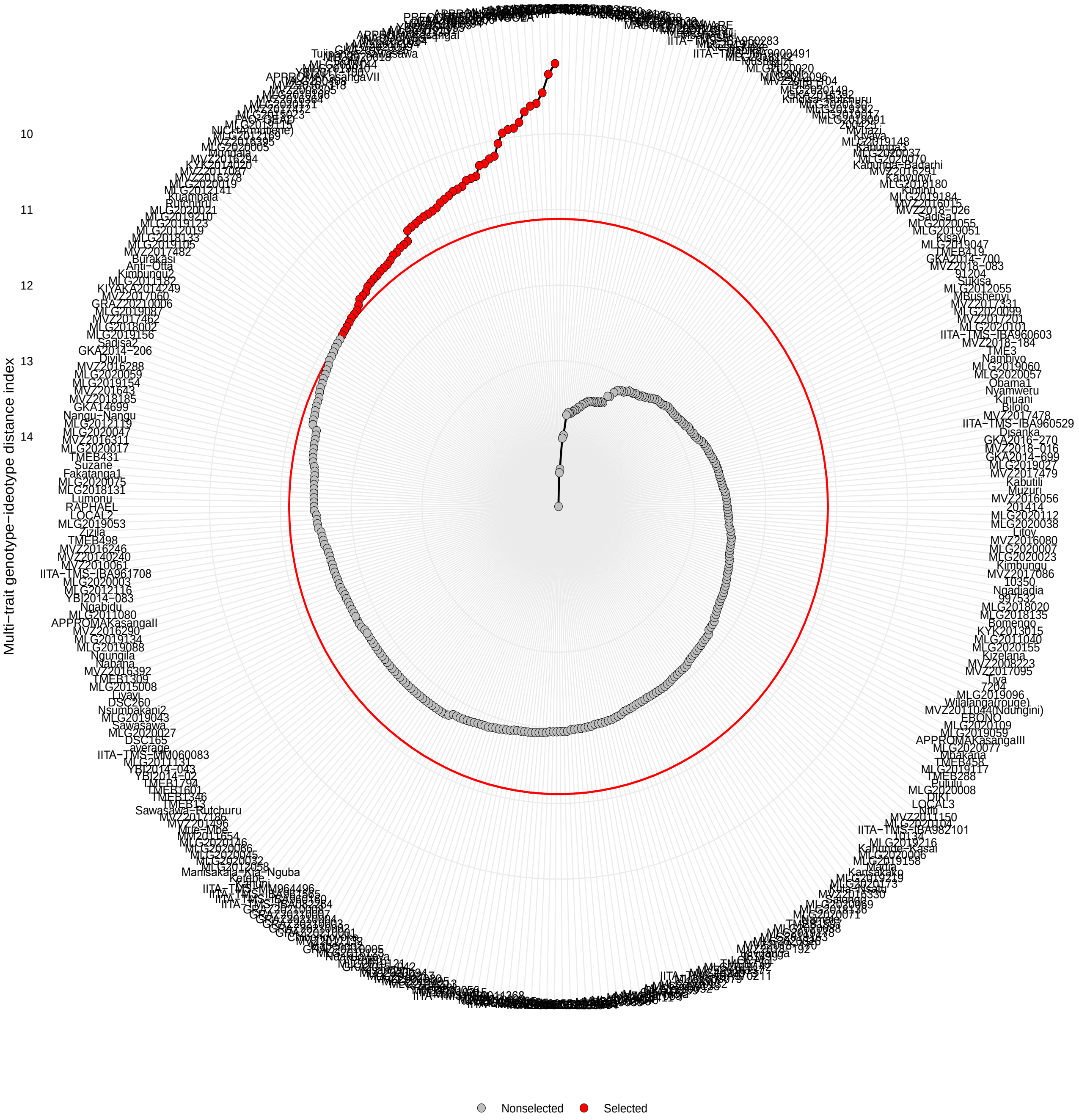


Fig. 8: Radar chart displaying weakness and strength of the selected cassava varieties.

Heritability estimates are crucial for determining the potential for genetic improvement in breeding programs. The high heritability observed in traits such as yield, DMC, and starch content (ranging from 0.55 to 0.98) reflects the robustness of the trials and the reliability of the results. High heritability simplifies the selection process and accelerates genetic gains, as documented by Ceballos et al. (2020). Furthermore, the observed negative correlation between fresh root yield and CMD severity score suggests that CMD severity can serve as a secondary trait for selecting high-yielding and disease-resistant varieties. Similar relationships have been documented in other studies, such as Darkwa et al. (2020), who emphasized the importance of indirect selection strategies in cassava breeding.

The molecular analysis using SNP markers provided complementary insights into the genetic diversity of cassava germplasm. By integrating clustering techniques such as Discriminant Analysis of Principal Components (DAPC), hierarchical clustering, and structure analysis, four well-defined genetic groups were identified. These findings highlight the effectiveness of molecular markers in uncovering genetic relationships that may not be evident through phenotypic observations alone (Agre et al., 2021). Molecular markers, particularly SNPs, have revolutionized cassava breeding by enabling high-resolution genetic analyses. In this study, the integration of phenotypic and molecular data revealed inconsistencies between phenotypic and molecular clustering, emphasizing the need for combined approaches. Previous studies have reported similar discrepancies in clonal crops such as provitamin-A cassava (Kamanda et al., 2020) and yam (Agre et al., 2023) and sugarcane (Alves et al., 2017). The importance of integrated approaches in breeding programs cannot be overstated, as they provide a comprehensive understanding of genetic resources and their potential applications.



Supplementary Fig. 8: MGDI displaying the 47 selected cassava varieties.

Using the MGDI (Multi-Trait Genotypic Diversity Index) method, 47 superior cassava varieties were identified, combining phenotypic and molecular diversity. This approach offers a robust framework for selecting progenitors in targeted breeding programs. As suggested by Ouattara et al. (2024), identifying superior progenitors is critical for enhancing productivity, resilience, and the overall utility of cassava as a staple crop, particularly in regions like the DRC where cassava serves as a primary source of food and income.

Germplasm exploration in diverse environments, including wild relatives, local landraces, and farmer-selected varieties, is pivotal for restoring genetic diversity and addressing challenges such as CMD and climate change. This study's findings highlight the effectiveness of germplasm restoration efforts in the DRC, as evidenced by the collection of disease-resistant and high-yielding varieties. Recent studies, such as Liang et al. (2024) and Simiyu (2023), have emphasized the role of diverse germplasm in developing climate-resilient and high-yielding cassava varieties.

The findings of this study underscore the need for continuous germplasm evaluation and integration of modern breeding tools in cassava improvement programs. Enhanced germplasm collection and conservation should involve expanding exploration efforts to include underrepresented regions and wild relatives while establishing long-term conservation strategies to preserve genetic diversity. Integration of genomic tools is essential, leveraging genomic selection, GWAS, and CRISPR-based tools for targeted trait improvement (Liang et al., 2024) and enhancing the use of molecular markers to accelerate breeding cycles. Developing climate-resilient varieties with enhanced tolerance to abiotic stresses such as drought and salinity, as highlighted by Ceballos et al. (2020), remains a key priority. Incorporating farmer preferences into breeding objectives ensures the adoption of new varieties, while collaborative research efforts should strengthen partnerships between research institutions, breeders, and farmers to facilitate knowledge exchange and resource sharing.

Conclusion

This study evaluated a total of 136 cassava varieties, including nine improved varieties developed by IITA and INERA-DRC and 120 local varieties collected from various agro-ecological zones and maintained by INERA.

Significant phenotypic variation was observed among the genotypes, particularly in traits like plant height, biomass, and yield. High heritability estimates for traits such as starch content and biomass underscore the potential for genetic improvement. Hierarchical clustering identified five distinct groups with varying traits and performance, highlighting the diversity within the germplasm.

Genotypic analysis using SNP markers revealed substantial genetic diversity, with population structure analysis identifying four genetic clusters. The integration of phenotypic and genotypic data through hierarchical clustering provided a comprehensive understanding of the relationships within the germplasm. The study identified 47 top-ranking cassava varieties with exceptional multi-trait performance, suitable for breeding programs aimed at improving yield, disease resistance and dry matter content.

Overall, this study provides valuable insights into the genetic and phenotypic diversity of cassava germplasm in the DRC, offering a foundation for future breeding efforts to develop improved cassava varieties.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Acknowledgment: The authors are grateful for the financial support of the donor United States Agency for International Development (USAID) through the funded project named ''Action to Control Cassava Brown Streak Disease in the Democratic Republic of Congo (CBSD-RDC)''. We also thank Kayondo Ismail for his technical support in-leaf samples for genotyping through the support of the OneCGIAR Initiative Research Program on Roots, Tubers, and Bananas (CRP-RTB).

Author's Contribution: Conceptualization, MS; methodology, MS, NA, PM and EM; software, PA and ID; validation, MS and PA; formal analysis, PA; investigation, MS, NA, MP and EM; resources, MS, NA, EM and PM; data curation, MS, PA, AHA, CZ, TB, ARA and RB; writing-original draft preparation, MS and PA; writing-review and editing, ARA, RB; visualization, MS, PA, AHA, CZ, RB; supervision, MS; project administration, MS; funding acquisition, MS; All authors have read and agreed to the published version of the manuscript.

REFERENCES

Adebayo, W.G. (2023). Cassava production in Africa: A panel analysis of the drivers and trends. *Heliyon*, 9 (9), 2405-8440. https://www.cell.com/heliyon/fulltext/S2405-8440

Adebayo, W.G., & Silberberger, M. (2020). Poverty reduction, sustainable agricultural development, and the cassava value chain in Nigeria. *The Palgrave Handbook of Agricultural and Rural Development in Africa*, 525-551. https://doi.org/10.1007/978-3-030-41513-6_24

Adu, B.G., Akromah, R., Amoah, S., Nyadanu, D., Yeboah, A., Aboagye, L.M. & Owusu, E.G. (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

Agbeleye, O., Olubiyi, O.A., Ehirim, B.O., Shittu, A.O., Jolayemi, O.L., Adetimirin, V.O., Ariyo, O.J., Sanni, K.A., & Venuprasad, R. (2019). Screening African rice (*O. glaberrima* Steud.) for tolerance to abiotic stress. III. Flooding. *Journal of Breeding and Genetics*, 51(2), 128-150.

Agre, P.A., Dassou, A.G., Loko, L.E., Idossou, R., Dadonougbo, E., Gbaguidi, A., & Asfaw, A. (2021). Diversity of white Guinea yam (*Dioscorea rotundata* Poir.) cultivars from Benin as revealed by agro-morphological traits and SNP markers. *Plant Genetic Resources*, *19*(5), 437-446. https://doi.org/10.1017/S1479262121000526

Agre, P.A., Edemodu, A., Obidiegwu, J.E., Adebola, P., Asiedu, R., & Asfaw, A. (2023). Variability and genetic merits of white Guinea yam landraces in Nigeria. *Frontiers in Plant Science*, *14*, 1051840. https://doi.org/10.3389/fpls.2023.1051840.

Alves, R. M., de Sousa, S. C. R., de Albuquerque, P. S. B. & dos Santos, V. S. (2017). Phenotypic and genotypic characterization and compatibility among genotypes to select elite clones of *cupuassu*. *Acta Amazonica* 47, 175–184. https://doi.org/10.1590/1809-4392201602104

Amelework, A.B., Bairu, M.W., Maema, O., Venter, S.L. & Laing, M. (2021). Adoption and Promotion of Resilient Crops for Climate Risk Mitigation and Import Substitution: A Case Analysis of Cassava for South African Agriculture. *Frontier Sustainable Food System*, 5, 617783. https://doi.org/10.3389/fsufs.2021.617783

Bakelana, Z., Laura, M.B., Kehoe, M., Pita, J., Monde, G., Mahungu, N., & Kalonji, M. (2020). Cassava Root Necrosis Disease (CRND): A new crop disease spreading in western Democratic Republic of Congo and in some central African countries. *Journal of Agricultural Science*; Vol. 12, No. 3; ISSN 1916-9752. https://doi.org/10.5539/jas.v12n3p105

Bisimwa, E.B., Birindwa, D.R., Yomeni, M.O., Rudahaba, N., Byamungu, K., & Bragard, C. (2019). Multiple cassava viruses' co-infections and resurgence of pests are leading to severe symptoms and yield losses on Cassava in the South-Kivu Region, Democratic Republic of Congo. *American Journal of Plant Sciences*, *10*(11), 1969. 10.4236/ajps.2019.1011138

Burns, A., Gleadow, R., Cliff, J., Zacarias, A., & Cavagnaro, T. (2010). Cassava: the drought, war and famine crop in a changing world. *Sustainability*, *2*(11), 3572-3607. https://doi.org/10.3390/su2113572.

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. https://doi.org/10.1007/s00122-015-2555-4.

Darkwa, K., Olasanmi, B., Asiedu, R., & Asfaw, A. (2020). Review of empirical and emerging breeding methods and tools for yam (Dioscorea spp.) improvement: Status and prospects. *Plant Breeding*, *139*(3), 474-497. https://doi.org/10.1111/pbr.12783

Ekanayake, I. (1996). Cassava: procedures for growth analysis: procedures manual. https://cgspace.cgiar.org/server/api/core/bitstreams/3e63fec4-5805-46a4-976c-5e7c888a748d/content

Elias, M., Panaud, O., & Robert, T. (2000) Assessment of genetic variability in a traditional cassava (Manihot esculenta Crantz) farming system, using AFLP markers. *Heredity* 85, 219–230. https://doi.org/10.1046/j.1365-2540.2000.00749.x

Faostat (2022). FAOSTAT database. Available at: http://bit.ly/NmQzZf.

Hauser, S., Bakelana, Z., Bungu, D.M., Mwangu, M.K., Ndonda, A. (2020). Storage root yield response to leaf harvest of improved and local cassava varieties in DR Congo. *Archives of Agronomy and Soil Science*, 1-15. https://doi.org/10.1080/03650340.2020.1800647

Howeler, R.H., Lutaladio, N., & Thomas, G. (2013). *Save and grow: cassava: a guide to sustainable production intensification*. Rome: Food and Agriculture Organization of the United Nations. Italie, 1-129. ISBN 9251076413.

IITA (1990). Cassava in Tropical Africa. A reference manual. IITA, Ibadan, Nigeria. 108.

IITA (2024) Farmer-preferred, CBSD-resistant cassava varieties released by Central Africa breeding program. Accessed February 2024. Available from https://www.iita.org/news-item/farmer-preferred-cbsd-resistant-cassava-varieties-released-by-central-africa-breeding-program/.

Josse, J., & Husson, F. (2016). missMDA: a package for handling missing values in multivariate data analysis. *Journal of Statistical Software*, *70*, 1-31. https://doi.org/10.18637/jss.v070.i01.

Kamanda, I., Blay, E. T., Asante, I. K., Danquah, A., Ifie, B. E., Parkes, E., & Sesay, S. (2020). Genetic diversity of provitamin-A cassava (Manihot esculenta Crantz) in Sierra Leone. *Genetic resources and crop evolution*, *67*, 1193-1208. https://doi.org/10.1007/s10722-020-00905-8

Katz, S.H., & Weaver, W.W. (2003). *Encyclopedia of food and culture*. Charles Scribner's Sons. ISBN 0684805685. https://thuvienso.hoasen.edu.vn/handle/123456789/8262.

Kawuki, R.S., Herselman, L., Labuschagne, M.T., Nzuki, I., Ralimanana, I., Bidiaka, M., & Ferguson, M. E. (2013). Genetic diversity of cassava (Manihot esculenta Crantz) landraces and cultivars from southern, eastern and central Africa. *Plant Genetic Resources*, *11*(2), 170-181. https://doi.org/10.1017/S1479262113000014.

Kintché, K., Hauser, S., Mahungu, N.M., Ndonda, A., Lukombo, S., Nhamo, N., Uzokwe, V.N.E., Yomeni, M., Ngamitshara, J. & Vanlauwe, B. (2017). Cassava yield loss in farmer fields was mainly caused by low soil fertility and suboptimal management practices in two provinces of the Democratic Republic of Congo. *European Journal of Agronomy*, *89*, 107-123. https://doi.org/10.1016/j.eja.2017.06.011.

Legg, J. P., Kumar, P. L., Makeshkumar, T., Tripathi, L., Ferguson, M., Kanju, E., ... & Cuellar, W. (2015). Cassava virus diseases: biology, epidemiology, and management. In *Advances in virus research*, (91), 85-142. https://doi.org/10.1016/bs.aivir.2014.10.001

Legg, J.P., Jeremiah, S.C., Obiero, H.M., Maruthi, M.N., Ndyetabula, I., Okao-Okuja, G., & Kumar, P.L. (2011). Comparing the regional epidemiology of the cassava mosaic and cassava brown streak virus pandemics in Africa. *Virus Research*, *159*(2), 161-170. https://doi.org/10.1016/j.virusres.2011.04.018.

Liang, X., Dong, C., Wenjun, O., Xin, C., Ismail, Y.R., Wenquan, W., Kaimian, L. & Huabing, Y. (2024), Advancements and strategies of genetic improvement in cassava (*Manihot esculenta* Crantz): from conventional to genomic approaches, *Horticulture Research*, uhae 341, https://doi.org/10.1093/hr/uhae34.

Loko, Y.L., Adjatin, A., Dansi, A., Vodouhè, R., & Sanni, A. (2015). Participatory evaluation of Guinea yam (Dioscorea cayenensis Lam.–D. rotundata Poir. complex) landraces from Benin and agro-morphological characterization of cultivars tolerant to drought, high soil moisture and chips storage insects. *Genetic Resources and Crop Evolution*, *62*, 1181-1192. https://doi.org/10.1007/s10722-015-0221-y

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K. (2023). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.8. https://CRAN.R-project.org/package=cluster

Mahungu, N.M., Dixon, A.G.O., & Kumbira, J.M. (1994). Breeding cassava for multiple pest resistance in Africa. *African Crop Science Journal*, *2*(4), 539-552.

Malik, A.M., Kongsil, P., Nguyễn, V.A., Ou Sholihin, W., Srean, P., Sheela, M.N., Augusto, L., & Ishitani, M. (2020). Cassava breeding and agronomy in Asia: 50 years of history and future directions. *Breeding Science*, 2020(70), 145-166. https://doi.org/10.1270/jsbbs.18180

Munyahali, W., Pypers, P., Swennen, R., Walangululu, J., Vanlauwe, B., & Merckx, R. (2017). Responses of cassava growth and yield to leaf harvesting frequency and NPK fertilizer in South Kivu, Democratic Republic of Congo. *Field Crops Research*, *214*, 194-201. https://doi.org/10.1016/j.fcr.2017.09.018.

Nkhoma, N., Shimelis, H., Laing, M.D., Shayanowako, A., & Mathew, I. (2020). Assessing the genetic diversity of cowpea [Vigna unguiculata (L.) Walp.] germplasm collections using phenotypic traits and SNP markers. *BMC Genetics*, *21*, 1-16. https://doi.org/10.1186/s12863-020-00914-7.

Olivoto, T., & Nardino, M. (2020). MGIDI: A novel multi-trait index for genotype selection in plant breeding. *bioRxiv*, 2020-07. https://doi.org/10.1101/2020.07.23.217778.

Ouattara, F., Agre, P. A., Adejumobi, I. I., Akoroda, M. O., Sorho, F., Ayolié, K., & Bhattacharjee, R. (2024). Multi-trait selection index for simultaneous selection of water yam (Dioscorea alata L.) genotypes. *Agronomy*, *14*(1), 128. https://doi.org/10.3390/agronomy14010128

Peres-Neto, P. R., Jackson, D. A. & Somers, K. M. (2003). Giving meaningful interpretation to ordination axes: assessing loading significance in principal component analysis. *Ecology* 84, 2347–2363.

Pierre, N., Wamalwa, L.N., Muiru, W.M., Simon, B., Kanju, E., Ferguson, M.E., Ndavi, M.M., & Tumwegamire, S. (2022). Genetic diversity of local and introduced cassava germplasm in Burundi using DArTseq molecular analyses. *PLos ONE* 17(1), e0256002. https://doi.org/10.1371/journal.pone.0256002.

Popat R, Patel R, Parmar D (2020). _variability: Genetic Variability Analysis for Plant Breeding Research_. R package version 0.1.0, https://CRAN.R-project.org/package=variability.

Safo-Kantanka, O., & Owusu-Nipah, J. (1992). Cassava varietal screening for cooking quality: relationship between dry matter, starch content, mealiness and certain microscopic observations of the raw and cooked tuber. *Journal of the Science of Food and Agriculture*, *60*(1), 99-104. https://doi.org/10.1002/jsfa.2740600116

SENASEM (2019). *Catalogue national varietal des cultures vivrières. Répertoire des variétés homologuées de plantes à racines, tubercules et du bananier*, 124, 93-117.

Shaibu, A.A., Uguru, M.I., Sow, M., Maji, A.T., Ndjiondjop, M.N., & Venuprasad, R. (2018). Screening African rice (Oryza glaberrima) for tolerance to abiotic stresses: II. Lowland drought. *Crop Science*, *58*(1), 133-142. https://doi.org/10.2135/cropsci2017.04.0255.

Sikirou, H., Rojanaridpiched, C., Phumichai, C., Becerra, L. A., Kittipadakul, P., Iglesias, C., & Gracen, V. E. (2020). Excellence in cassava breeding: perspectives for the future. *Crop Breeding, Genetics and Genomics*, *2*(2). https://doi.org/10.20900/cbgg20200008

Sikirou, M., Shittu, A., Konaté, K.A., Maji, A.T., Ngaujah, A.S., Sanni, K.A., Ogunbayo, A.S., Akintayo, I., Saito, K., Ahanchédé, A., & Venuprasad, R. (2018). Screening African rice (*Oryza glaberrima*) for tolerance to abiotic stresses: I. Fe toxicity. *Field Crops Research*, *220*, 3-9. https://doi.org/10.1016/j.fcr.2016.04.016.

Simiyu (2023). Genetic diversity and resistance to bacterial blight among Kenyan cassava (Manihot Esculent Crantz) germplasm. *Faculty of Agriculture, University of Nairobi.* 105. http://erepository.uonbi.ac.ke/handle/11295/164589.

Wambugu, P.W., Ndjiondjop, M.N. & Henry, R. (2019). Advances in molecular genetics and genomics of African Rice (*Oryza glaberrima* Steud). *Plants*, *8*(10), 376. https://doi.org/10.3390/plants8100376.