Received: 22 February 2024 | Accepted: 26 September 2024

DOI: 10.1002/csc2.21401

Crop Science

ORIGINAL ARTICLE

Crop Breeding & Genetics

Assessment of genetic diversity and heterotic alignment of CIMMYT and IITA maize inbred lines adapted to sub-Saharan Africa

Tinovonga Gonhi[1,2,3] | **Thomas Lapaka Odong**[1] | **Isaac Onziga Dramadri**[1,2] | **Mildred Ochwo-Ssemakula**[1] | **Zvenhamo Albert Chiteka**[2] | **Emmanuel Amponsah Adjei**[1,2] | **Dean Muungani**[4] | **Abebe Menkir**[4] | **Badu-Apraku Baffour**[4] | **Idris Adejumobi**[4] | **Brigitte Uwimana**[5] | **Thanda Dhliwayo**[6] | **Dagne Wegary**[7] | **John Derera**[4,8]

[1]Department of Agricultural Production, College of Agricultural and Environmental Sciences, Makerere University, Kampala, Uganda

[2]Makerere University Regional Centre for Crop Improvement (MaRCCI), Kampala, Uganda

[3]African Indigenous Vegetables Institute (AIVI), Marondera University of Agricultural Science and Technology (MUAST), Marondera, Zimbabwe

[4]Maize Improvement Program, International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria

[5]International Institute of Tropical Agriculture (IITA), Sendusu, Uganda

[6]International Maize and Wheat Improvement Center (CIMMYT), Mexico, Mexico

[7]International Maize and Wheat Improvement Center (CIMMYT), Harare, Zimbabwe

[8]One CGIAR, Ibadan, Nigeria

Correspondence
Tinovonga Gonhi, Department of Agricultural Production, College of Agricultural and Environmental Sciences, Makerere University, Kampala, Uganda.
Email: tinotendagonhi@gmail.com; tinovongagonhi@gmail.com

Assigned to Associate Editor Xue-Feng Ma.

Funding information
CGIAR'S initiative on Accelerated Breeding at IITA and CIMMYT

Abstract

Despite the breeding efforts by many institutions, maize (*Zea mays* L.) productivity in sub-Saharan Africa is still low. A limited number of productive maize hybrids have been developed partly due to a lack of knowledge on the diversity and heterotic relationship of the germplasm, especially in public breeding programs. Understanding the extent of diversity, structure, and heterotic grouping of available maize germplasm originating from different breeding programs is important to enhance long-term genetic gain in hybrid maize breeding programs by optimizing heterotic pools using modern breeding tools. Information about the genetic structure of the available germplasm could help breeders design effective breeding strategies to improve yield. This study was conducted to determine the genetic diversity, population structure, and heterotic alignment among 187 elite maize inbred lines from the IITA (International Institute of Tropical Agriculture) and CIMMYT

Abbreviations: AMOVA, analysis of molecular variance; CA, CIMMYT heterotic group A; CIMMYT, International Maize and Wheat Improvement Center; CMLs, CIMMYT maize lines; DArT-Seq, Diversity Array Technology sequencing; GD, genetic distance; IITA, International Institute of Tropical Agriculture; NGS, next-generation sequencing; OFP, off-patent; PCoA, principal coordinate analysis; POP, population; SNP, single nucleotide polymorphism; SSA, sub-Saharan Africa; TASSEL, Trait Analysis by aSSociation Evolution and Linkage.

(International Maize and Wheat Improvement Center) maize breeding programs. The inbred lines were genotyped with 9857 Diversity Array Technology sequencing based single nucleotide polymorphism markers. Hierarchical clustering revealed three major groups, with some subgroups consistent with the selection history, and pedigree of the inbred lines. Three broad groups were detected: two consisting of CIMMYT lines and a mixed group consisting of both CIMMYT and IITA inbred lines. The STRUCTURE analysis revealed six subpopulations (fixation index value = 0.58), which depicts a moderate genetic diversity among the materials. Population 2 comprises the highest number of genotypes (102) from both programs. More than 89% of the elite lines had homozygosity exceeding 95%, with the remaining lines requiring further inbreeding through repeated self-pollination. There was inconsistency in the predetermined heterotic groups' alignment between CIMMYT and IITA elite inbred lines. Analysis of molecular variance revealed that 96% of the total variation was accounted for by differences within groups, with the remaining 4% representing the variation between groups. This suggests that the two programs can benefit from germplasm exchange for the improvement of maize productivity.

Plain Language Summary

Due to climate change effects, a limited number of productive maize hybrids are able to withstand the prevailing conditions. This might be attributed to a lack of knowledge of the diversity and heterotic relationship of the available materials. Understanding the extent of diversity, structure, and heterotic grouping of available maize germplasm from different breeding programs will enable the development of climate-resilient crops and enhance long-term genetic gain in hybrid maize breeding programs through modern breeding tools. The genotyped lines revealed three major groups. Moderate genetic diversity among the materials revealed that it can be maximized. Note that 89% of the lines were pure, with the remaining requiring further repeated self-pollination. There is inconsistency in the predetermined heterotic groups, which breeders need to treat separately when sharing the materials. Overall, the two programs can benefit from sharing their material for the improvement of maize productivity.

1 | INTRODUCTION

In sub-Saharan Africa (SSA), "maize (*Zea mays* L.) is life" due to its role in the agro-food system (Erenstein et al., 2022). It is a primary food security crop in most farming systems, especially for smallholder farmers in mid-altitude and subhumid agroecologies of SSA (Mideksa et al., 2022). However, maize yields in SSA are still lower than 2 t/ha than global averages, despite significant achievements made by international agricultural institutes such as IITA (International Institute of Tropical Agriculture) and CIMMYT (International Maize and Wheat Improvement Center) through the Drought Tolerance Maize for Africa (DTMA) and Stress Tolerance Maize for Africa (STMA) projects from 2006 to 2019 (Abate et al., 2017; STMA, 2020). Maize production faces many challenges including water deficit for rain-fed crop production systems (Musimwa et al., 2022), soil degradation, pest and disease pressure, and stagnant yields (Pruitt, 2016). Genetic improvement provides a sustainable option to address some of these constraints, with its success driven by the availability and systematic use of diverse germplasm (Ertiro et al., 2017; Gedil & Menkir, 2019).

The development of high-yielding hybrids requires crossing diverse parents to maximize heterosis (Russell, 1909; Shull, 1908). Genetic diversity serves as a hub of many unique alleles for improving agronomic and quality traits, thereby promising parents can be selected in breeding programs (Singh et al., 2020). Information about the genetic

structure of the available germplasm could help breeders design effective breeding strategies to improve yield (Singh et al., 2020). For example, US breeders used source open-pollinated variety of a line as a basis for assigning inbreds to the now famous "Iowa Stiff Stalk Synthetic" and non-Stiff Stalk heterotic groups (HGs) that have been the basis of impressive genetic gain in the US Corn Belt (Duvick et al., 2004; Hallauer, 1999).

However, tropical maize germplasm has a wide genetic base and is genetically more complex than temperate material (de Faria et al., 2022; Lu et al., 2009; Xu & Crouch, 2008), and hence cannot be easily classified into two HGs (Akinwale, 2021). The complexity arises from germplasm management practices and breeding objectives that favor mixing of germplasm pools, such as development of open-pollinated varieties (OPVs) (Warburton et al., 2005) and use of commercial hybrids as source germplasm to develop inbred lines (Guimarães et al., 2018). For instance, CIMMYT maize populations and pools were formed with little regard to the racial origin of the landraces that were used to form them (CIMMYT, 1998). Still, an understanding of the genetic structure of germplasm pools, especially among breeding programs, is essential to facilitate germplasm exchange and use.

Studies to assess genetic diversity between CIMMYT and IITA germplasm have been conducted previously (Adebayo et al., 2015; Badu-Apraku et al., 2015, 2016; Dao et al., 2014; Dhliwayo et al., 2009; Ifie et al., 2015). However, to facilitate germplasm sharing, periodic assessment is necessary, especially because breeders occasionally introduce new sources of germplasm. For example, CIMMYT started introducing temperate germplasm around 2011 and released the first CIMMYT maize lines (CMLs) containing temperate germplasm in 2018. Additional CMLs containing temperate germplasm were released in 2023 (CIMMYT, 2017).

Adebayo et al. (2015) reported higher interpopulation genetic diversity than intrapopulation between CIMMYT and IITA inbred lines. They suggested that diversity might be due to the differences in genetic backgrounds and adaptive traits that breeders in each center selected during inbred line development.

It was also reported that CIMMYT material was more drought-tolerant than IITA material (Adebayo et al., 2015). More so, there was germplasm exchange that has been taking place between West Africa and Southern Africa maize breeding programs for their development of source population, OPVs, hybrids, as well as quantitative trait locus (QTL) mapping populations. Adebayo et al. (2015) reported that IITA maize materials contained more temperate germplasm compared to CIMMYT, which used mainly material from tropical germplasm. This made CIMMYT material more complex than IITA and then made materials from these two sister centers more complex when trying to combine them in direct single cross-hybrid formation as there will be no consistency in heterotic patterns (Badu-Apraku et al., 2016; Badu-Apraku & Fakorede, 2017).

Core Ideas

- There was moderate genetic diversity among the inbred lines evaluated from the two sister breeding programs.
- There was inconsistency in the predetermined heterotic alignment between International Maize and Wheat Improvement Center and International Institute of Tropical Agriculture elite inbred lines.
- The programs can benefit from germplasm exchange for the improvement of maize productivity.
- There is need to align heterotic groups to facilitate the ease utilization of available germplasm from two centers.
- Combining ability study recommended to determine potential productivity of crosses among lines from pops identified.

The combining ability (CA) theory developed by Sprague and Tatum (1942) allows the breeders to select and evaluate parental inbred lines, decreasing the blindness of parent selection, and increasing the efficiency of selecting excellent crosses. This process gave birth to the development of the maize HG theory (Lu & Xu, 2010). The existing maize HGs from both CIMMYT and IITA still need to be characterized and validated with the currently new inbred lines developed (Badu-Apraku et al., 2021). An HG is defined as the collection of related inbred lines, which tend to produce vigorous hybrids when crossed with lines from a different group, but not when crossed with lines of the same group (Suwarno et al., 2014). Dhliwayo et al. (2009) reported that the predefined heterotic grouping of the lines they used from CIMMYT and IITA did not consistently predict the performance of hybrids, which suggests that CIMMYT's and IITA's HGs do not correspond closely with each other. This was buttressed by their finding of the CA tests, which reviewed that CIMMYT A × IITA A hybrids yielded much better than the inter–HG CIMMYT B × IITA A. As highlighted before about the genetic complexity of tropical mid-altitude maize inbred lines due to mixed genetic composition and the broad genetic base of the source populations, Menkir et al. (2004) suggested that there might be challenges in classifying these inbred lines into distinct HGs based on the results of CA studies only. For that reason, merging CA (CIMMYT HG A) results and molecular markers that have the capacity to compare the similarity of inbred lines at the DNA level with testcross evaluation will

expedite the separation of these lines into well-defined HGs. Therefore, using high-resolution molecular markers becomes the best method to make inferences relating to genetic diversity among genotypes (da Silva et al., 2020) to maximize heterosis in hybridization. In addition, to refine the tropical maize HGs, it is important for the technology transfer from temperate maize breeding as well as capacity building in tropical countries (Guo et al., 2021).

Increased genetic gain can be achieved by either reducing the length of the breeding cycle or increasing the selection accuracy and efficiency of the breeding program. One of the ways that can be employed to increase the accuracy of selection in breeding is through the application of molecular markers, especially polymerase chain reaction (PCR)-based markers (Badu-Apraku et al., 2021), in conjunction with the next-generation sequencing (NGS) technology.

Among the PCR-based markers including amplified fragment length polymorphism and single sequence repeat (SSR), single nucleotide polymorphisms (SNPs) have been the markers of choice in analyzing genetic diversity of populations (Varshney et al., 2007). This is due to their abundance in the genome of plants and other organisms (Mammadov et al., 2012). More recently, the Diversity Array Technology sequencing (DArT-Seq) has been widely used by plant breeders to generate thousands of SNP markers and to assess the genetic diversity of panels and breeding populations in many crops. DArT-Seq is a hybridization-based high-throughput DNA sequencing modern technology, which is highly reproducible and of low-cost technology (Deschamps et al., 2012; Tomkowiak et al., 2021). Furthermore, this technology requires no prior sequence information for detecting SNPs associated with loci/alleles for traits of interest (Nadeem et al., 2018).

This allows rapid identification of SNP (Cruz et al., 2013; Raman et al., 2014). The DArT-Seq was successfully used in different genetic diversity studies such as *Solanum lycopersicum* (Pailles et al., 2017), *Solanum tuberosum* (Berdugo-Cely et al., 2017), *Allium sativum* (Egea et al., 2017), coffee (Spinoso-Castillo et al., 2020), cowpea (Edema et al., 2023), yams (Amponsah Adjei et al., 2023; Bhattacharjee et al., 2020), and cassava (Adu et al., 2021; Xia et al., 2005). In maize, this technology has gained the momentum in genetic diversity analysis (Ayesiga et al., 2023; Badu-Apraku et al., 2021; Obeng-Bio et al., 2020; Tomkowiak et al., 2021). However, few studies have been done to dissect on genetic variation, population structure, heterotic alignment, and purity status among maize germplasm adapted to tropical lowland and tropical mid-altitude from West and Southern Africa, respectively, using this highly informative and performance genome marker technology of DArT-Seq on maize breeding in Zimbabwe.

The objective of this study was to assess the genetic diversity, population structure, and heterotic alignment among inbred lines originating from the IITA-Nigeria and CIMMYT-Zimbabwe maize breeding programs using the DArT-Seq-based SNP markers.

2 | MATERIALS AND METHODS

2.1 | Genetic materials and leaf sampling

A total of 187 elite lines (Supporting Information) developed for tolerance to abiotic and biotic stress factors at CIMMYT and IITA were grown in Mzarabani in Zimbabwe during the 2022 winter season (May–October). The maize leaf sampling was carried out in accordance with the recommended LGC plant sample collection kit (KBS-9370-001) protocol. Four leaf discs from each plot were collected into 96-well tube plates 4 weeks after planting using a leaf puncher. Next, the leaves were freeze-dried in a vertical freeze dryer with a vacuum pump (Freeze Dryer BK-FD10P, Biobase LLC) for 18 h at −58.7°C.

2.2 | DNA extraction and SNP discovery by DArT-Seq technology

Maize leaf samples were sent to DArT in Australia for genotyping. DNA extraction was done using Nucleomag Plant DNA extraction kit (Mag-Bind Plant DNA DS 96 Kit). The genomic DNA extracted was in the range of 50–100 ng/uL. DNA quality and quantity were checked on 0.8% agarose gel. Libraries were constructed following the DArT-Seq complexity reduction method (Carling et al., 2015; Kilian et al., 2014) through the digestion of genomic DNA using a combination of *Pst*I and *Mse*I enzymes. This was followed by the ligation of barcoded adapters and common adapters followed by PCR amplification of adapter-ligated fragments. Libraries were then sequenced using single-read sequencing runs for 77 bases.

NGS was carried out using Hiseq2500 (Kilian et al., 2016). DArT-Seq marker scoring was achieved using DArTsoft14, which is an in-house marker scoring pipeline (Kilian et al., 2016). Two types of DArT-Seq markers were scored, SilicoDArT markers scored as binary for presence/absence (1 or 0, respectively), whereas the SNP markers were scored as co-dominantly, that is, 0–1–2, where 2 represents the heterozygous genotype. Both SilicoDArT markers and SNP markers were aligned to the *Z. mays* references genomes ("Maize_NAM_v5" and "Maize_v329") to identify chromosome positions.

2.3 | Filtering, analysis of genetic diversity, and population structure

The data quality control and filtering were performed using TASSEL (Trait Analysis by aSSociation Evolution and

Linkage) (v5.2.90) (Bradbury et al., 2007). A total of 9857 SNP markers were filtered where SNPs with greater than 10% of missing data, less than 0.05 minor allele frequency, and unknown positions on the genome were removed. The SNP data were further imputed using the *k*-nearest neighbor genotype imputation method (Bradbury et al., 2007), resulting in a total of 2025 SNPs (24.2%) selected for further analysis.

SNP marker and diversity statistics including average number of alleles per locus, *F* statistics (Wright) among others, were calculated using TASSEL (v5.2.90) (Bradbury et al., 2007). The polymorphism information content was assessed using PowerMarker (v3.25) (Liu & Muse, 2005). In addition, observed and expected heterozygosity was determined using the "adegenet" package in R (R Core Team, 2020). The Bayesian model-based clustering that is implemented in the STRUCTURE software (v2.3.4) (Evanno et al., 2005; Pritchard et al., 2000) was used to assess the genetic structure of the germplasm. The STRUCTURE analysis was run considering a burn-in period of 10,000 Markov-chain Monte Carlo iterations and a 100,000-run length with an admixture model following the Hardy–Weinberg equilibrium and its correlated allele frequencies. Furthermore, for each number of clusters, 10 independent runs were performed, which ranged from 1 to 10. The hypothetical number of subpopulations (*K*) was estimated through the application of a Bayesian clustering approach.

To identify the optimum *K* value, which is the distinct peak in the change of likelihood (ΔK), the output from STRUCTURE was uploaded to STRUCTURE HARVESTER (Earl & vonHoldt, 2012). Inbred lines with a membership probability of $\geq$0.80 were assigned into the same group, while those with <0.80 were designated as a mixed group (Kumar et al., 2022).

Population differentiation assessment among the genotypes from the two programs was done using GenAlEx 6.503 (Peakall & Smouse, 2012) so as to perform the analysis of molecular variance (AMOVA) (Reyes-Valdés, 2013). The marker datasets were coded numerically prior to AMOVA, as A = 1, C = 2, G = 3, and T = 4 (Blyton & Flanagan, 2006) and later subjected to AMOVA with 999 permutations. This leads to the genetic variations being partitioned into two groups: variation among the population (PhiPR) and variation within the population (PhiPT) (Amponsah Adjei et al., 2023; Frank et al., 2018). Also, phylogenetic relationships among genotypes based on program source were built using a Euclidean distance matrix in PowerMarker (v3.25) (Liu & Muse, 2005). The resulting trees were visualized using DARwin 6.0.21 (Perrier & Jacquemoud-Collet, 2006) (https://darwin.cirad.fr/). To assess the heterotic alignment between material from CIMMYT and IITA, the filtered and imputed SNP data were analyzed with DARwin (v6.0.21) software. The cluster analysis was performed with unweighted pair group method using arithmetic averages, using a distance matrix of maximum rank distance in Paleontological statistics (PAST) (Hammer et al., 2001) to visualize how the materials clustered in a graphical presentation.

3 | RESULTS

3.1 | Quality, diversity, and functional characterization of DArT-Seq-SNPs

A total of 9857 unfiltered SNPs (and 19,714 SNP and SilicoDart markers) were generated from the DArT-Seq genotyping of 187 maize inbred lines (Table S1). The filtering and imputation process was done, and a large set of SNPs was removed, and 2025 good-quality SNPs (24.2%) distributed across the 10 chromosomes of maize were retained (Figure 1) and used for further analyses. The highest number of SNPs for marker density across chromosomes was on chromosome 1 (276 SNPs; 13.6%) followed by chromosome 5 (262 SNPs; 12.9%) and chromosome 2 (254 SNPs; 12.5%), while the lowest number of SNPs were mapped on chromosome 10 (138 SNPs; 6.8%) of the physical map ("Maize_NAM_v5" and "Maize_v329") (Figure 1). Quality parameters such as average PIC value across all the markers was 0.2852 and ranged from 0.0782 to 0.5825. The observed heterozygosity ranged from 0 to 0.36 with a mean of 0.04, while the expected heterozygosity ranged from 0.08 to 0.50 and the mean was 0.31 (Table 1).

3.2 | Population differentiation, genetic relation, and genetic purity of the maize inbred lines

The genotypic data of the 187 inbred lines were subjected to cluster analysis using TASSEL (v5.2.90). The genetic diversity analysis using the distance matrix in TASSEL (v5.2.90) revealed an average distance between the materials evaluated to be 0.31 (Table 2). The hierarchical and radical clustering revealed three major groups (Figures S1 and S2), with some subgroups consistent with the selection history, and pedigree of the inbred lines. Three broad groups were detected where two groups consist of CIMMYT lines and a mixed group consisting of both CIMMYT and IITA inbred lines (Figure 2). Clusters 1 and 3 were unique to the CIMMYT-Southern Africa breeding program (about 95% of the CIMMYT materials) and cluster 2 was a mixture of CIMMYT and IITA-originated inbred lines (Figure 2). Even though there was a mixture of materials clustering from the two programs, the materials from each center tended to group closer together forming two subclusters (Figure 2, subclusters 2a and 2b). Surprisingly, in cluster 1, there was an IITA entry (TZMI2085) that reflected to be tightly related to CIMMYT materials although it was revealed to be a pro-vitamin A inbred line.

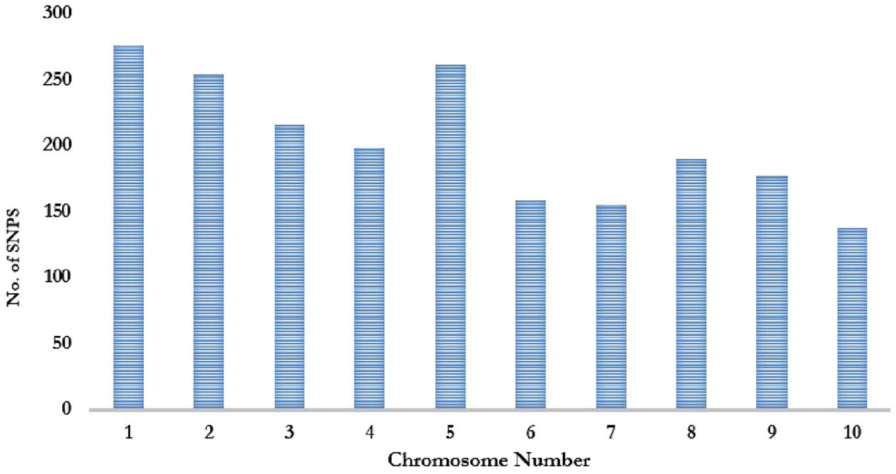

FIGURE 1 Distribution of 2025 single nucleotide polymorphisms (SNPs) across 10 *Zea mays* chromosomes based on physical map.

TABLE 1 Quality and summary statistics of 2025 DArT-Seq-SNPs (Diversity Array Technology sequencing-single nucleotide polymorphisms) markers on maize chromosomes.

	PIC value	Call rate	Observed heterozygosity	Expected heterozygosity
Maximum	0.58	1.00	0.36	0.50
Average	0.29	0.65	0.04	0.31
Minimum	0.08	0.20	0.00	0.08

Abbreviation: PIC, polymorphism information content.

TABLE 2 Summary statistics of the genetic diversity between the 187 inbred lines evaluated using genetic distance.

Description	Genetic distance
Minimum	0.01
Maximum	0.38
Average	0.31

The AMOVA among the two populations based on geographic origin (IITA-Nigeria and CIMMYT-Zimbabwe) was performed and revealed 96% of the variability within each population group, while only 4% was between them and this was statistically significant ($p < 0.001$) (Table 3).

The genetic purity test was performed in TASSEL (v5.2.9) to determine the level of heterozygosity and homozygosity to know the lines to use in future line breeding or hybrid development as well as action to be taken to purify them if needed. The results revealed that more than 89% of the elite lines had homozygosity exceeding 95%, with the remaining lines being moderate to high heterozygosity, which requires further inbreeding through repeated self-pollination (Figure 3).

The highly heterozygous lines (maximum of 22% heterogeneity) include the CIMMYT popular public lines such as CML445 (G148) and CML590 (G009) (17% and 22% heterogeneity, respectively) and IITA public lines: TZMI1268 (G005) and TZEI2725 (G037) with 17% and 22% heterogeneity, respectively (Table 4).

3.3 | Population structure and principal coordinate analysis of the germplasm from the two breeding institutions

The highest expected heterozygosity was observed on POP4 (where POP is population) followed by POP2 with 0.46 and 0.30, respectively, with the least fixation index (Fst) of 0.11 and 0.31, respectively (Table 5), whereas the least expected heterozygosity was on POP1 and POP5 with 0.07 and 0.08, respectively, with high Fst values of 0.85 and 0.83, respectively (Table 5). To understand the pattern of population structure of the germplasm from the two breeding programs, a Bayesian information criterion was used. The determination of the populations at each K-value and membership coefficients (qi) in STRUCTURE analysis was very informative in coming up with the number of populations of the genotyped materials. Simulations (logarithm probability relative to standard deviation, ΔK) estimated from the SNP markers showed a sharp peak at $\Delta K = 6$ (Figure 4), which explained the optimum number of subpopulations ($K = 6$). At this $\Delta K = 6$, the distribution of genotypes in each assumed subpopulations were as follows: POP1 = 13 genotypes (6.9%) POP2 = 101 genotypes (53.7%), POP3 = 18 genotypes (9.6%), POP4 = 11 genotypes (5.9%), POP5 = 13 genotypes (6.9%), and POP6 = 10 genotypes (5.3%) (Table 6). In POP2 with the highest number of genotypes, there was almost equal number of materials from both centers (IITA = 44 genotypes and CIMMYT = 55 genotypes). CIMMYT materials were found in all the subpopulations assumed, whereas IITA materials were only found in POP2 and POP4 (Table 6). The majority of IITA materials were found in POP2 with 44 genotypes out of 49 IITA materials genotyped. These results from

TABLE 3 Analysis of molecular variance (AMOVA) of 187 genotypes based on distance method using 2025 single nucleotide polymorphism (SNP) markers.

Source	*df*	SS	MS	Estimated variance	Percentage[a]	Fst value
Among pops	1	1298.9	1298.9	13.3***	4	0.58
Within pops	186	62,367.3	335.3	335.3***	96	
Total	187	63,666.2		348.6		

Abbreviations: df, degree of freedom; Fst, fixation index; MS, means squares; SS, sum of squares.

[a]Percentage of the total variation.

***$p < 0.001$.

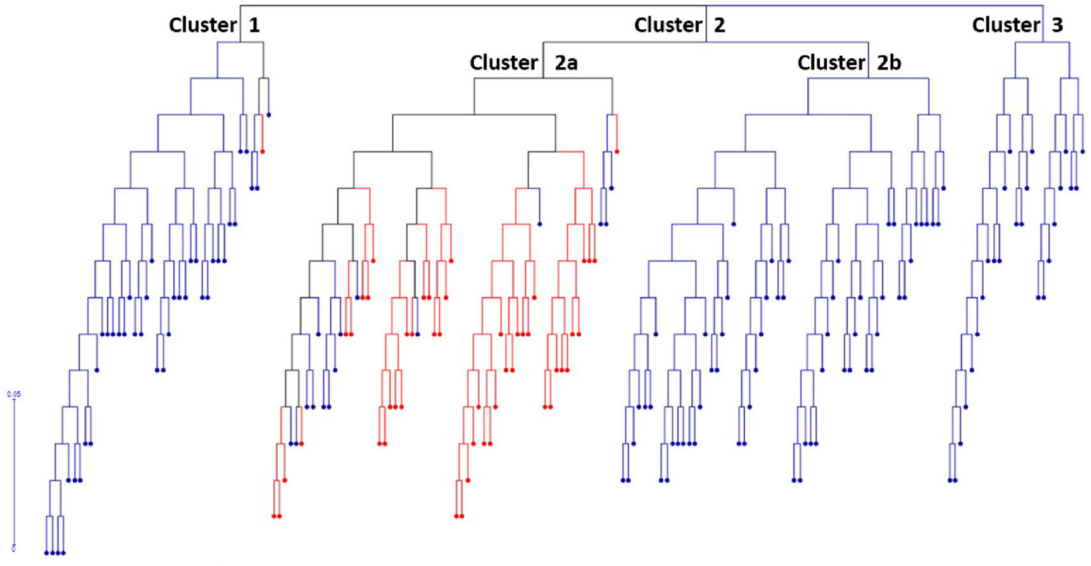

FIGURE 2 Hierarchical clustering of genetic similarity of the 187 inbred lines of maize (*Zea mays* L.) on the basis of single nucleotide polymorphism (SNP) markers observations into three main clusters (blue = CIMMYT [International Maize and Wheat Improvement Center], red = IITA [International Institute of Tropical Agriculture] materials).

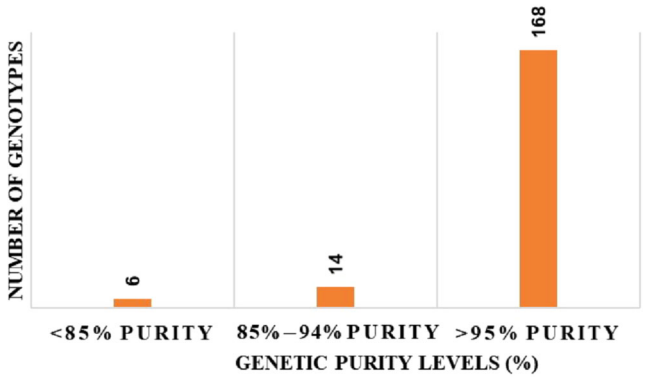

FIGURE 3 Bar graph showing the genetic purity level (proportion of homozygous) of the 187 inbred lines genotyped.

STRUCTURE concur with the hierarchical clustering analysis (Figure 2) where IITA materials were only found in two clusters and also the majority of its materials were also found in one cluster (cluster 2a). It was observed that popular public CMLs that are known to be sister lines (CML312 and CML311; CML444 and CML566) were clustered in the same subpopulation, hence confirmed the accuracy of the grouping (Table 6). There was an inverse correlation between the expected heterozygosity and the fixation index.

More so, 6% of the total sampled materials were grouped with the popular CIMMYT HG B tester CML444 in POP1, 9% grouped with CML312 (HG A) in POP3, and 5.3% grouped with CML395 in POP5 (Table 6). The rest percentage were not aligned to the common A and B testers, hence attest to the possibility of the AB HG to be present. This will be validated in a CA study. This corroborates with previous findings in other studies involving CIMMYT and IITA lines. Lemi and Diro (2022) reported that two quality protein maize inbred lines from IITA-Nigeria and CIMMYT- Zimbabwe showed positive s10 Marchpecific combining ability (SCA) with both testers from the different HGs and proposed that they belong to HG AB.

Furthermore, the principal coordinate analysis (PCoA) based on the pairwise Euclidean genetic distance (GD) matrix among the materials was done to untangle the genetic

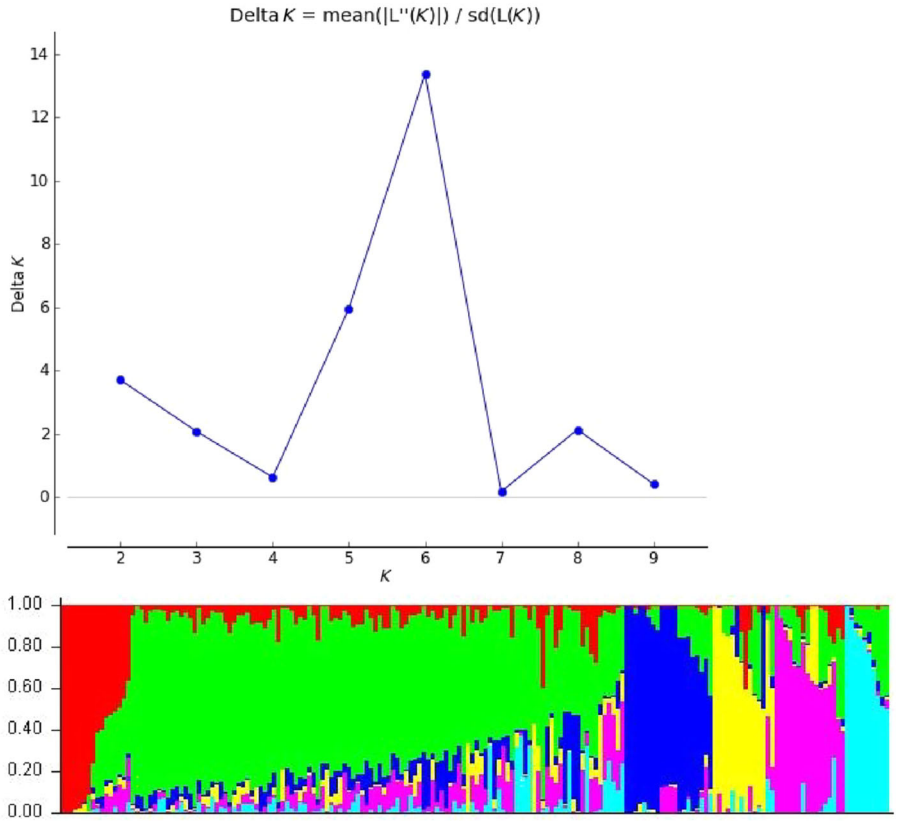

FIGURE 4 Population structure in the maize germplasm: Top = likelihood of ΔK showing the optimum K value ($K = 6$); bottom = population structure obtained for 187 maize genotypes based on 2025 single nucleotide polymorphisms (SNPs) at different $\Delta K = 6$.

TABLE 4 The list of the genotypes with high heterozygosity levels that needs breeders' attention.

Taxa code	Proportion heterozygous	Proportion homozygous	Proportion homozygous	Comments on purity and actions to be done
G173	0.22	0.78	78	High heterozygosity, ear-to-row selection
G009	0.22	0.78	78	High heterozygosity, ear-to-row selection
G132	0.18	0.82	82	High heterozygosity, ear-to-row selection
G037	0.17	0.83	83	High heterozygosity, ear-to-row selection
G148	0.17	0.83	83	High heterozygosity, ear-to-row selection
G005	0.16	0.84	84	High heterozygosity, ear-to-row selection

TABLE 5 Pairwise comparison of allele-frequency divergence among pops (net nucleotide distance), fixation index (Fst), and expected heterozygosity among the six subpopulations using computed point estimates of P.

	POP1	POP2	POP3	POP4	POP5	POP6	Expected He	Fst	Percentage membership
POP1	–	0.16	0.23	0.17	0.28	0.28	0.07	0.85	10.8
POP2	0.16	–	0.10	0.05	0.15	0.15	0.30	0.31	47.4
POP3	0.23	0.10	–	0.10	0.22	0.18	0.19	0.59	13.9
POP4	0.17	0.05	0.10	–	0.15	0.16	0.46	0.11	8.6
POP5	0.28	0.15	0.22	0.15	–	0.27	0.08	0.83	11.5
POP6	0.28	0.15	0.18	0.16	0.27	–	0.12	0.77	7.8

Abbreviations: He, heterozygosity; POP, population.

TABLE 6 Inferred ancestry of individuals in each assumed population (Red = IITA [International Institute of Tropical Agriculture], black = CIMMYT [International Maize and Wheat Improvement Center]).

Population	POP1			POP2			POP3	POP4	POP5	POP6
	CML593	TZEI 2778	CZL16031	CZL1368	CML592	CZL16027	CML537	CML590	**CML395**	CML545
Genotypes	CZL15076	TZEI 2758	TZEEI 214	BSL210859	TZEEI 231	TZEI 2775	CML539	DJ737-505	CZL16066	BSL210837
	CZL15131	TZEI 2756	TZEEI 357	TZMI1267	G176	CZL16174	CZL1359	TZEI 2725	CZL16124	BSL210836
	CZL15081	TZEI 2779	CZL20004	BSL210865	DJ853-5	TZEEI 247	CZL1360	TZMI1268	CZL16125	BSL210835
	CML444	TZEI 2792	TZEI 2722	TZMI1167	CZL15013	TZMI1252	**CML312**	CZL15038	CML543	BSL210834
	CZL16051	TZEEI 319	CML571	DJL182071	TZMI1270	CZL15084	CZL1424	CZL1358	CML607	CZL16145
	CML566	TZMI1152	DJ853-16	CZL15178	CZL17021	TZEI 2726	CZL15167	DJL184822	DJ737-166	CZL17015
	CZL15130	TZMI1151	TZMI1294	CZL15123	DJ737-513	CML538	CZL15019	CML197	CZL16014	DJ854-28
	CZL15191	TZMI1240	TZEI 2766	DJ737-682	CZL20003	CML572	CZL1112	CZL20013	CZL16077	CML536
	CZL16057	TZEI 2790	BSL210854	CML548	DJ854-7	CML547	CML589	TZMI1273	CZL15085	CZL15110
	DJ737-348	TZEEI 320	TZMI1278	BSL210867	DJ737-62	CZL16115	CZL1349	TZMI863	CZL15168	
	DJ737-349	CZL15224	CZL16144	CZL1461	CML576	CML489	CZL20002	CZL15007		
	CZL16040	TZEI 2768	CZL20004	G178	CZL15181	CML445	CZL15003	DJ737-4		
		TZEI 2760	TZMI1184	CZL1310	DJL191378	CML550	CZL15225			
		CZL15128	TZEEI 269	CZL16132	CZL20001	DJ853-15	**CML311**			
		TZMI2030	TZMI2040	CZL1254	DJ854-21	CML591	CZL15206			
		TZMI1301	TZEI 2753	TZEEI 273	TZMI1298	CZL15034	CZL15208			
		CZL15205	TZMI2037	TZMI878	DJ854-13	DJ737-7	BSL210843			
		TZMI2012	TZMI1271	TZMI2085	DJ737-153	DJ737-469				
		TZEI 2733	TZEEI 310	DJ737-648	CML549	TZMI1307				
		TZEEI 386								
Proportion of the total sample (%)	6.9	53.7					9.6	5.9	6.9	5.3
No. of IITA lines	0	44					0	4	0	0
No. of CIMMYT lines	13	55					18	7	13	10

Abbreviation; POP, population; SCA, specific combining ability.

divergence in the maize inbred lines using the DArT-Seq-SNP markers. The PCoA results revealed the total genetic variation, which was explained by the first two axes of the PCoA (Figure 5). Moreover, the analysis revealed that IITA materials (red) are closely related compared to CIMMYT materials (blue) that were scattered in all the quadrants, showing high genetic diversity among the materials (Figure 5). The wide divergence of CIMMYT genotypes was retained with some clear clusters being shown that might add up to six subpopulations and this concurs with the output from the STRUCTURE analysis (Figure 2).

3.4 | Heterotic grouping alignment

The grouping alignment of the materials was done using DARwin 6.0.21 (Perrier & Jacquemoud, 2006). The visualization revealed that there is no consistent alignment between materials from West Africa (IITA) and those from Southern Africa (CIMMYT) (Figure 6). Surprisingly, none of the genotypes from HG A of IITA were grouped with the materials from HG A of CIMMYT (Figure 6 CA). CIMMYT materials from HG B were clustered into two groups (CB1 and CB2). Although it was observed that there was no consistency in aligning materials from the two centers into their existing groups (A and B), in CB2 cluster, there were two lines from IITA HG B (IB) that well aligned with CIMMYT HG B (CB) (Figure 6, CB2). More so, there was a group of early maturity material from IITA that were not known (In/a) about the HG they belong. These lines have been developed recently and they were at S6, therefore we did not have much information known about them including genetic diversity, population structure, and heterotic grouping, and they were to be advanced to S7 stage after which they would be

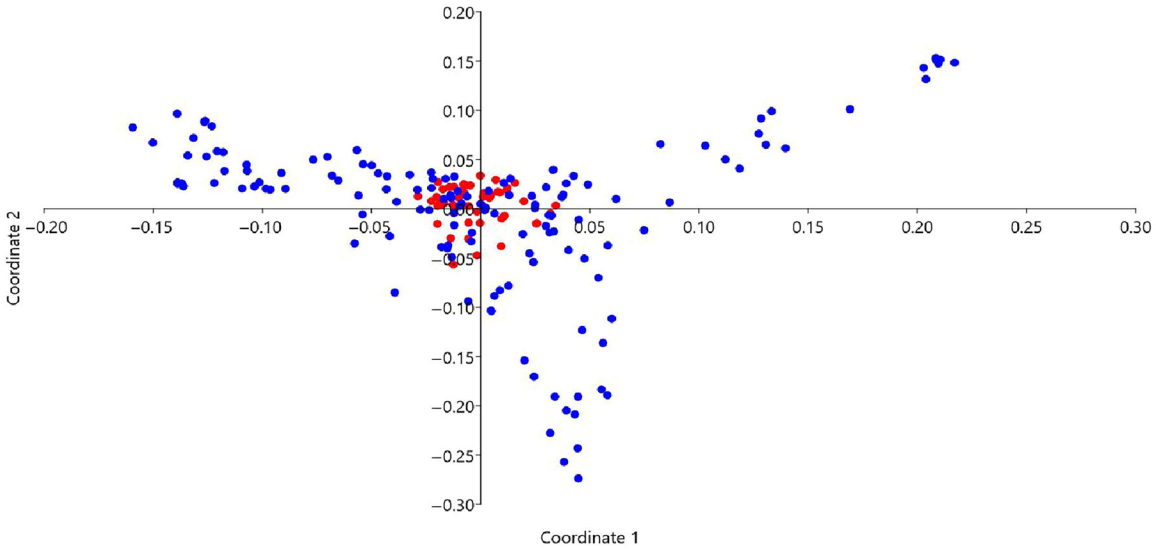

FIGURE 5 Principal coordinate analysis (PCoA) of 187 maize inbred lines genotypes based on 2025 DArT single nucleotide polymorphism (SNP) markers. International Maize and Wheat Improvement Center (CIMMYT)-Zimbabwe = blue; International Institute of Tropical Agriculture (IITA)-Nigeria = red.

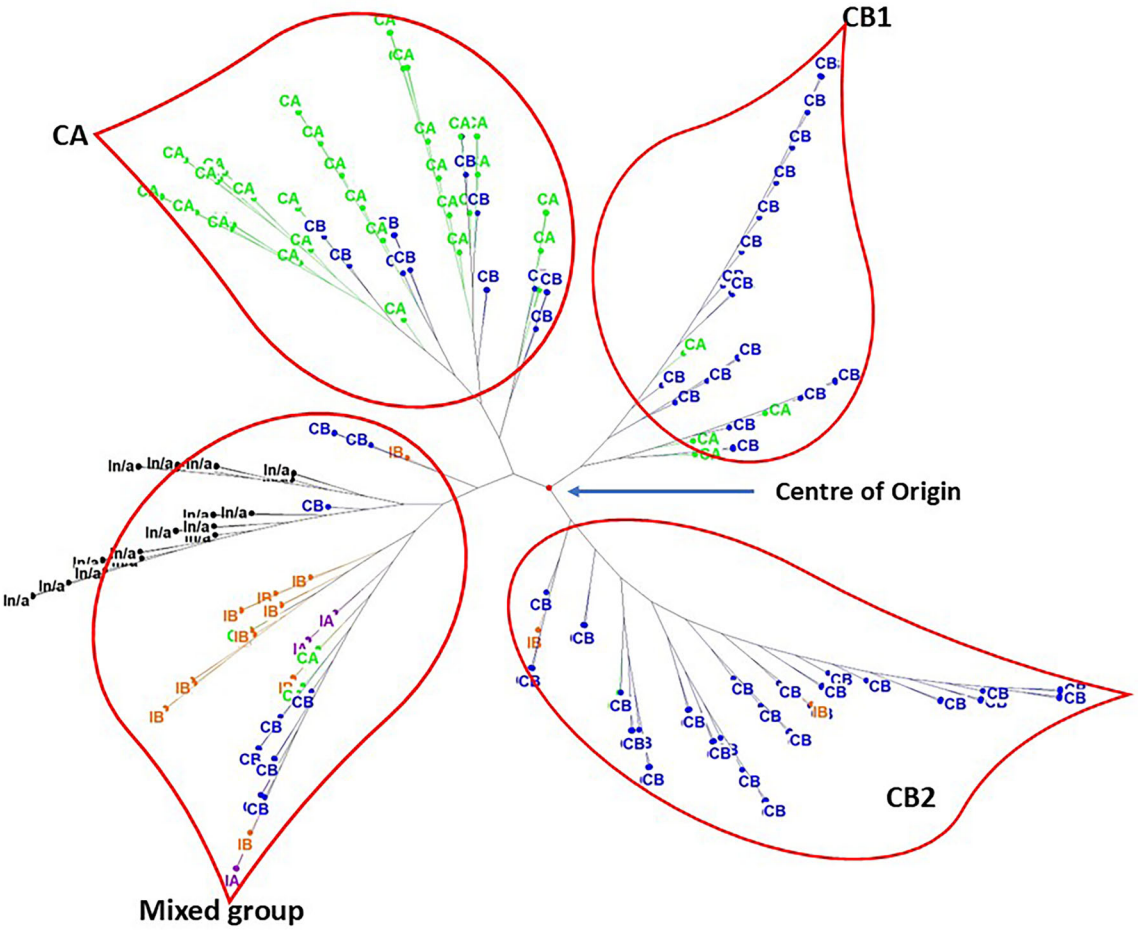

FIGURE 6 Heterotic grouping alignment of materials from CIMMYT (International Maize and Wheat Improvement Center) and IITA (International Institute of Tropical Agriculture) using DArT-Seq-SNPs (Diversity Array Technology sequencing single nucleotide polymorphisms) markers. CA = CIMMYT heterotic group A, CB1 = CIMMYT heterotic B group 1; CB2 = CIMMYT heterotic group B 2; IA = IITA heterotic group A; IB = IITA heterotic group B; In/a = IITA not available heterotic group information.

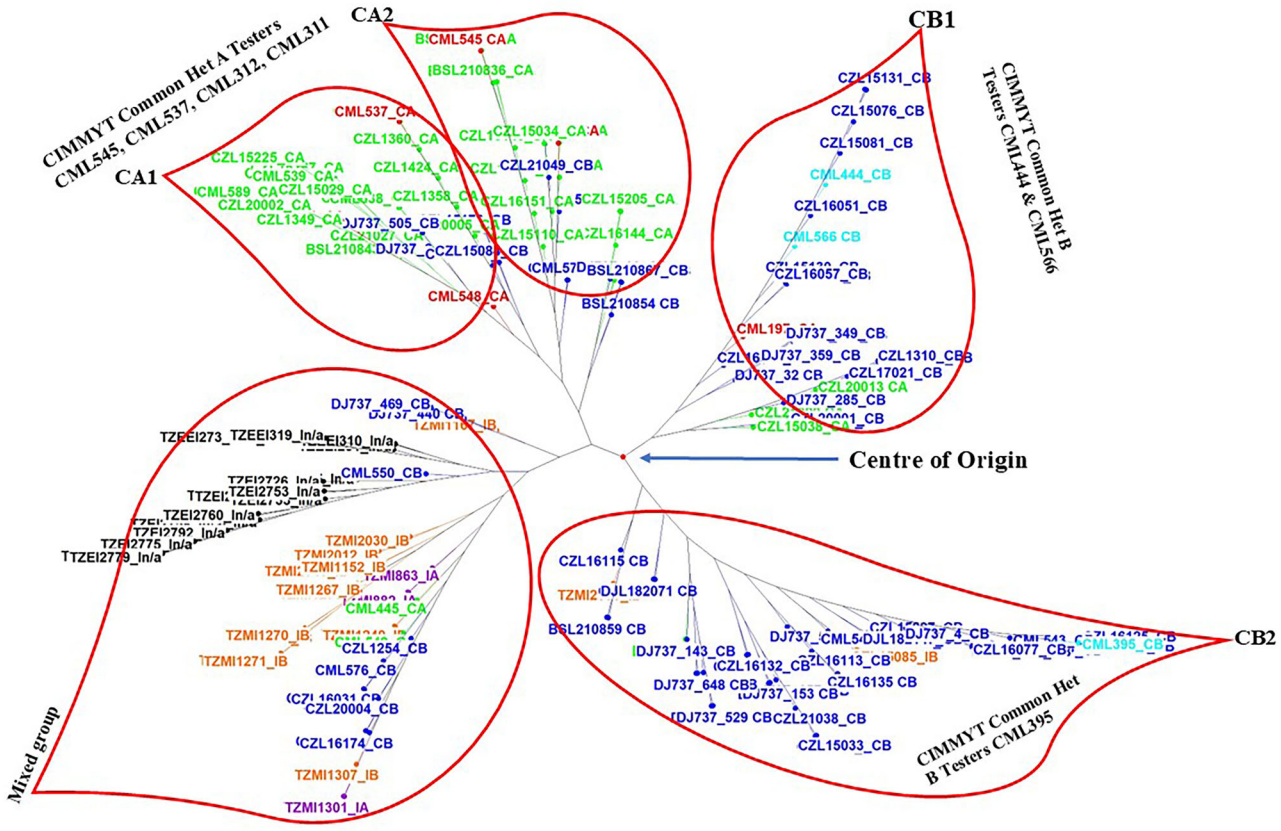

FIGURE 7 Heterotic grouping alignment of materials from CIMMYT (International Maize and Wheat Improvement Center) and IITA (International Institute of Tropical Agriculture) using DArT-Seq-SNPs (Diversity Array Technology sequencing single nucleotide polymorphisms) markers with their actual names showing common inbred lines from CIMMYT used in defining CIMMYT heterotic groups, for example, CML444 = HetB, CML312, CML545 = Het A, CA1 = CIMMYT heterotic group A cluster 1, CA2 = CIMMYT heterotic group A cluster 2, CB1 = CIMMYT heterotic group B cluster 1; CB2 = CIMMYT heterotic group B cluster 2; IA = IITA heterotic group A; IB = IITA heterotic group B, In/a = IITA not available heterotic group information. Blue color = materials from CIMMYT in hetB, green color = materials from CIMMYT in het A, **black color** = early materials from IITA no available heterotic group information (In/a), brown = intermediate to later materials from IITA het A, purple = intermediate to later materials from IITA het B.

characterized (B. Badu-Apraku, IITA Breeder, personal communication, March 10, 2022). The alignment showed that all these materials are 50% closely related and aligned closer to their IITA HGs A and B (hetB) than their A group (Figure 6, mixture). In addition, all the materials from IITA are generally closer to CIMMYT hetB than A, as in the mixed cluster there were more materials from CIMMYT hetB and also in one of the CIMMYT hetB cluster 2 (CB2), there were two inbred lines from IITA hetB. The most popular CIMMYT inbred line that shows highly related to IITA materials was CML550, which is in CIMMYT hetB (Figure 7, cluster mixture). However, although those lines from IITA clustered together with CB, because of the inconsistency in heterotic alignment, it is difficult to conclude that heterotic Bs from both programs are aligned as in the mixed cluster there were also some materials from CIMMYT heterotic A (CA) and no IITA heterotic A (IA) materials were found in the CA (Figure 7 cluster CA).

Although the materials within the CIMMYT clusters seemed to group according to their heterotic grouping (e.g., CML545, CML537, CML312, and CML311) (Figure 6, cluster CA1; hetA testers), there were some discrepancies; for example, CML197 (type A) and CML566 (type B) (Figure 6, cluster B1) grouped together regardless of them being unrelated by pedigree, hence continuous refining of the HGs is still needed. The clustering of CML444 with CML566 was consistent with pedigree information (Figure 6, cluster B1). There was a cluster of some inbred lines (e.g., BSL210854, CML571, and DJL184822) from CIMMYT hetB, but they were found to be closely related to CA.

4 | DISCUSSION

The characterization for genetic diversity among the genotypes available for breeding purposes is of uttermost importance toward efficient germplasm conservation and maximizing heterosis in future hybridization breeding programs (Amponsah Adjei et al., 2023; Melchinger, 2015). DArT-

Seq technology is among the increasingly used technologies in plant breeding that enable faster genotyping using NGS (Deschamps et al., 2012; Tomkowiak et al., 2021). In this study, to understand the genetic diversity among the materials from the Western Africa tropical lowland breeding program and the Southern Africa medium attitude maize breeding program, 2025 informative DArT-Seq SNP markers well distributed across all 10 chromosomes of maize were identified, as this provides a better estimate of diversity in the panel (Kumar et al., 2022).

The purity test done to the genotyped materials revealed a greater percentage (89%) of the materials being highly homozygous and the remaining being highly heterozygous (maximum of 22% heterozygous) including popular public lines such as CML445 [G148] and CML590 [G009], which recorded 17% and 22% heterozygous, respectively). These results are similar to the preceding study by Kumar et al. (2022) where they reported public lines like CML220-1 showing high heterozygosity (up to 43.8%) and the authors suspected either pollen contamination or seed mixture during seed production and maintenance. However, on the overall heterozygosity, our study revealed a very low heterozygosity (3%) compared to the previous studies (Ertiro et al., 2017; Kumar et al., 2022). Kumar et al. (2022) reported an overall average of 17% heterozygosity for their 384 inbred lines, while Ertiro et al. (2017) reported a range of 12.5%–31.5% heterozygosity in 265 maize inbred lines from CIMMYT, the Ethiopian Institute of Agricultural Research, and the IITA using 220,878 polymorphic SNPs. This difference might be because in our study we just used materials from two breeding programs compared to germplasm from 18 breeding centers used by Kumar et al. (2022) and three programs used by Ertiro et al. (2017). Regardless of the low overall heterozygosity reported in this study, the purification of those heterozygous inbred lines by ear-to-row selection from the original seeds source of the lines is still required. Therefore, homozygosity analysis for breeding lines is very important as it directly determine the line to be used as a parent in new line development and/or hybrid formation.

The genetic variability analysis for inter-institution revealed moderate genetic diversity with mean of 0.31 between their materials. This finding concurs with the previous studies done on tropical germplasm between IITA and CIMMYT that reported average diversity of 0.34 (Adebayo et al., 2015; Dhliwayo et al., 2009). Also, similar GDs (0.20–0.34) among 384 tropical maize inbred lines were reported (Kumar et al., 2022). The low to moderate genetic diversity reported in this study also concurs with the phenotypic observations witnessed between the materials in this study (data not shown). Most of the materials have erectile, broad shiny leaves and deep green color. These traits contribute to abiotic stress tolerance, especially drought stress such as the stay-green trait that reduces the rate of leaf senescence. This is not surprising given that these materials were all sampled from the DTMA and STMA projects. As a rule of thumb, CIIMYT and IITA elite materials are expected to undergo stress tolerance evaluation (Dagne Wegary, CIMMYT Breeder, personal communication, October 2, 2023). However, most of the IITA materials were susceptible to northern corn leaf blight or turcicum leaf blight (TLB), a foliar disease caused by *Exserohilum turcicum*. This means, IITA breeding program can maximize on CIMMYT materials in TLB tolerance improvement. However, the reported GD in the current study was slightly smaller than those reported in several previous studies (Adetimirin et al., 2008; Laakili et al., 2016; Liu et al., 2003; Matsuoka et al., 2002). The observed lower values in our study could arise from the small number of lines used for genotyping compared to 439 lines used by Badu-Apraku et al. (2021). The pedigree information of lines that were sampled for the current study shows that most IITA lines have a common parentage. Furthermore, the proportion in terms of representation from the two institutions was not balanced (138 CIMMYT and 49 IITA). Therefore, the small number of lines sampled from IITA could not capture the whole diversity within the breeding program, hence recommending increasing the number of lines that are drawn from different genetic backgrounds in future studies. However, some previous studies reported a relatively high diversity between CIMMYT and IITA materials regardless of the number and proportion of materials used. For instance, Ifie et al. (2015) assessed the genetic diversity among nine CIMMYT and 87 IITA early maturing inbred lines using 31 polymorphic SSR and 261 SNP markers and reported high genetic diversity between the evaluated materials. In contrast, Dhliwayo et al. (2009) evaluated a total of 20 lines from IITA and CIMMYT (10 from each institution) using SSR markers and reported low to moderate genetic diversity.

Furthermore, the moderate genetic variation between the genotypes from the two sister institutions might be attributed to the same genetic pool used, especially when breeding for abiotic stress tolerance such as drought. An analysis of the pedigree information of lines that were sampled for the current study shows common parentage for most IITA lines, and that both IITA and CIMMYT programs used common temperate off-patent (OFP) inbred lines (ex-plant variety protection) in crosses with tropical lines to develop breeding populations (new starts). Consistent with previous studies, the abiotic stress–tolerant parents, such as CML444, CML312, and LaPosta Sequia population, were commonly used as parents of the inbred lines characterized in the current study. Previous studies reported that the La Posta Sequia population was widely used in developing inbred lines and maize varieties tolerant to drought in SSA; hence, it was a common parent for most drought-tolerant materials (Adebayo et al., 2015). This lead to the clustering of germplasm into the same subpopulation with this common parentage of LaPosta Sequia

(Adebayo et al., 2015). Guei and Wassom (1996) described La Posta Sequia germplasm as a late maturing, lowland, tropical, and white dent maize that was improved at CIMMYT for drought tolerance through eight cycles of recurrent selection.

The results of this study revealed that, although between the programs there was moderate genetic diversity, AMOVA revealed that there was higher genetic diversity within each population (96% within population). These findings corroborate with findings from previous studies (Abu et al., 2021; Adebayo et al., 2015; Badu-Apraku et al., 2021; Dhliwayo et al., 2009). Warburton et al. (1998) reported that the high genetic diversity within the population than interpopulation indicates that the populations are heterogeneous at the molecular level. More so, genetic diversity studies on other crops besides maize reviewed higher genetic diversity within each institutions or population than between them, for example, in sugar cane *Saccharum* spp (Singh et al., 2020) and yams *Dioscorea* spp (Amponsah Adjei et al., 2023). This means that breeders from these centers can benefit from sharing germplasm for the improvement of specific traits and designing new hybrids (e.g., utilization of IITA X CIMMYT patterns). For instance, the West and Central Africa breeding programs emphasized more on biotic tolerant screening like *Striga* and maize streak virus, which are the most problematic stresses (Adebayo et al., 2015), while the breeding programs in Southern Africa invested more in breeding for tolerance to abiotic stresses like drought and low soil fertility. Also, the existence of high diversity within the programs reflects the possibility of maximizing heterosis within the respective programs (Adebayo et al., 2015). Adebayo et al. (2015) further reiterated that the diversity could result from the differences in genetic backgrounds and adaptive traits, which is given more priority by breeders during inbred line development. Kumar et al. (2022) also suggested that greater genetic diversity can still be available among sister lines that shared common parentage.

The study showed six subpopulations being assumed to be present among the genotyped materials. This depicts the existence of genetic diversity that still exists within the breeding programs (Guo et al., 2021) that breeders can maximize in developing productive inbred lines as well as making hybrids among lines from the different clusters. These results corroborate with the previous reports (Kumar et al., 2022) where $K = 6$ was reported out of 384 inbred lines genotypes using over 60,000 SNPs using the genotyping by sequencing platform.

In the present study, the PCoA revealed that materials originating from IITA, regardless of their maturity categories, are closely related compared to the diversity within CIMMYT materials. This agrees with the pedigree analysis done (not shown but available on special request for confidential purposes) where it was revealed that most of the IITA lines used in this study were closely related or sister lines that are derived from the same population or populations with one common parent. Most of the IITA lines used in the current study are related since they were derived from the same parents or populations with a common parent. This explains the lack of diversity among them as many of them have a common parentage in the form of drought-tolerant parent LaPosta, and temperate OFP lines from the United States. The major populations noted were TZE-W Pop DT C5 STR C5 and TZEE-W POP HDT C1 STR C5 S6 for the early maturing group, which were extracted from drought-tolerant and striga-resistant early maturing populations and combined heat- and drought-tolerant and striga-resistant extra-early maturing populations, respectively, whereas for the intermediate to late maturing inbred lines, it was noted that most lines were derived from crosses of TZMI407 × M017 (which is tropical × temperate crosses) and La Posta Sequia population, which is a common parent widely adopted in SSA in breeding for drought tolerance in maize; hence, majority of the materials were related or sister lines. These findings agree with the report by Badu-Apraku et al. (2021) where they found low to moderate genetic diversity being observed in the 439 inbred lines of IITA maize germplasm. Akinwale et al. (2014) suggested that this might be attributed due to the breeding strategies adopted at IITA, which cut across the extra-early, early, intermediate, and late maturing groups. The close clustering of materials originating from West Africa might have negative implications in the event that there is an outbreak of a stress (biotic or abiotic) because the whole or a bigger portion of the materials will be flagellated due to lack of diversity that is the source of noble alleles for crop stress tolerance. More so, the highly dispersed and closely distributed CIMMYT and IITA materials, respectively, reflected by the PCoA corroborate with what was reported by Adebayo et al. (2015) that IITA maize materials contained more temperate germplasm compared to CIMMYT, which uses mainly material from tropical germplasm. However, after doing the pedigree analysis of the CIMMYT materials used in this current study, it was observed that some of the CIMMYT inbred lines also contain both tropical and temperate materials. This might be the reason why the overall diversity between the materials from the two programs revealed less diversity. There are many temperate lines such as LH132, PHG39, and some coded "OFP" that were used in line breeding at CIMMYT. These populations are the major sources for stress tolerance, especially drought.

In addition, the higher diversity in CIMMYT materials compared to the IITA set in this current study might be due to other factors such as (1) sample size, which was more for CIMMYT than IITA, (2) diverse sources of germplasm, (3) intentional breeding activities to create genetic diversity, and (4) complexity associated with tropical germplasm as mentioned before. However, the larger sample size might be the major factor that has made the CIMMYT materials to reveal

a larger genetic diversity that IITA. This was noted after doing the pedigree analysis of the CIMMYT materials used in this study. It was observed that most of the lines shared a common parentage such as the popular old public CMLs such as CML444, CML312, CML395, and other temperate lines such as LH132 and the OFP lines like OFP105, OFP106, OFP1, and OFP14. This reflects that there is a lot of material recycling in line development as the breeders were developing heat-tolerant and drought stress tolerant products.

The larger genetic divergence in CIMMYT materials revealed by PCoA in PAST software was in accordance with the hierarchical clustering of the materials that was revealed by DARwin software for heterotic grouping. CIMMYT genotypes from heterotic A and B were split into two making A1 and A2, and B1 and B2, respectively. Guo et al. (2021) also reported higher genetic diversity among CIMMYT inbred lines than among expired US Plant Variety Protection. This diversity can be very important to both centers and private partners in mining beneficial alleles leading to the development of vigorous inbred lines within the HG, and when crossed to the opposite HG line, high heterosis can be realized in hybridization (Makumbi et al., 2011; Shull, 1952). IITA breeding program can benefit from these materials from Southern Africa, especially in breeding for abiotic stress tolerance (Adebayo et al., 2015).

IITA materials were observed to be closely related to the CIMMYT HG B as compared to CIMMYT type A. This might not be surprising as these materials were extracted from their preceding projects of STMA projects and DTMA projects where the programs collaborated. As highlighted before, La Posta Sequia is a common parent widely adopted in SSA in breeding for drought tolerance in maize and being categorized in the HG B of the CIMMYT program, hence clusters materials closer to each other. The close relationship of IITA materials to CIMMYT HetB might be caused by the presence of a popular inbred line CML395, that is, a type B in CIMMYT grouping, which is the common parentage to the majority of materials from both programs. Dhliwayo et al. (2009) highlighted that CML395 was developed from one of several hundred S2 lines received by CIMMYT breeders from IITA's breeding program in Cameroon in the mid-1990s. It is known to originate from a cross involving 50% IITA Midaltitude Streak Resistant (TZMSR; Dhliwayo et al., 2009); hence, the genetic similarity between IITA materials with CIMMYT type B is understandable.

The current study revealed that the materials from West Africa maize breeding program whose HGs were not known were highly closely related to each other. This concurred with their pedigree information where all of them are shown to be derived from the common population

"TZEE-W POP HDT" to IITA HG B than A: This can be accepted to a greater extent, as the B type contains more of the LaPosta Sequia, which is the major drought-tolerant drought donor population used in developing adaptable inbred lines in SSA (Adebayo et al., 2015). Genetic variation between germplasm within the programs can be maximized by breeders to produce productive and vigorous inbred lines (Makumbi et al., 2018). This will be of great importance in making single crosses using two vigorous parents from opposite or heterotic populations (Shull, 1908). This will be even much better in the formation of three-way or other types of hybrids. However, these findings contradict with the previous study done by Adebayo et al. (2015) where they reported high genetic variation between CIMMYT by IITA than among the materials within the program. This might be due to the fact that the studied inbred lines were being developed toward abiotic stress tolerance, that is, drought; hence, breeders were taking germplasm from the same pool (LaPosta Sequia) during the Drought Tolerance Maize for Africa and Stress Tolerance Maize for Africa projects (Abate et al., 2017; STMA, 2020). Hence breeders were mandated to develop materials with drought tolerance in their respective regions.

The STRUCTURE analysis of the materials from the two breeding centers revealed the presence of six subpopulations. This depicts the presence of genetic diversity existing among and between the breeding materials. These findings agree with the recently reported study of tropical maize in India where 384 inbred lines genotyped with 60,227 SNPs were grouped into three to six groups with no clear patterns of clustering by centers-wise breeding lines and any other agronomic traits (Kumar et al., 2022). However, genotypes were grouped partially based on their source germplasm from where they derived. All these results bounce back to the reflection of the complexity and diversity in tropical maize germplasm, which breeders can still untangle the hidden treasures of it.

Furthermore, the materials from the West African maize breeding program were not able to clearly differentiate themselves into their own unique HGs. This might be caused by their breeding background as stated by Adebayo et al. (2015) that IITA maize improvement program used to be focusing on the development of OPVs rather than hybrids, so shifting from OPV-oriented to hybrid-oriented programs will always face these challenges of mixing up and inconsistences, especially when applying the molecular classification method. Badu-Apraku et al. (2021) reported many heterotic clusters that were produced after evaluating early to extra-early maize germplasm with SNP markers. This reiterated the complexity that can be faced when coming from OPV-oriented program. Therefore, continuous refining of the new and already developed materials is still important so as to develop a cost-effective and manageable number of HGs where the lesser the number, the better (Badu-Apraku et al., 2021). However, for the heterotic grouping of materials from Southern Africa, the majority were able to group together according to their respective pedigree and CA background except for a few discrepancies. This might be due to the fact that CIMMYT has

invested a lot in hybrid programs long back than IITA; hence, more work has been done in refining their HG. These results are in agreement with the recent study by Tomkowiak et al. (2021), where the 94 maize inbred lines were grouped according to their pedigree using the DArT-Seq technology. The lack of enough distinction in the heterotic grouping of inbred lines reported in this study also concurs with their pedigree information where one inbred line contains parents from different HGs. For instance, genotype G174 contains CML395, CML444 are B type, and CML442 and CML197 that are A type. This might be caused by the selfing of commercial hybrids, which is a common practice and allowed in most countries to develop maize inbred lines (Guimarães et al., 2018).

The inconsistency of CIMMYT and IITA predetermined HGs in their alignment depicts that breeders from these two breeding programs and other partners when exchanging germplasm from these two centers have to treat them as differently. This inconsistency corroborates with previous studies where the performances of the developed hybrids from the predefined heterotic grouping of CIMMYT's and IITA's HGs were not able to be predictable. This suggests that their HG do not correspond closely with each other (Dhliwayo et al., 2009). For this reason, this heterotic orientation will be validated using the line × tester (L×T) evaluation study, which is a follow-up study to this research. Although Dhliwayo et al. (2009) recommended to increase the number of inbred materials to be evaluated to buttress their finding, this study still concurs with their results regardless of the increment in germplasm that was evaluated (a total of 187 inbred lines comprising 138 from CIMMYT and 49 from IITA) versus 20 lines (10 from each center) they evaluated. This might lead to the crossing of closely related parents in developing crosses leading to less heterosis being realized (Oyetunde et al., 2020). This might be the reason why yield in SSA seems to be stagnant, regardless of germplasm shared between breeding programs. Therefore, it is suggested that breeders must treat these HGs as different and study them before making final crosses. All germplasm lines should be genotyped and allocated to common HGs to facilitate sharing and their utilization in both new hybrid designing and new starts population that will be used for new inbred extraction.

5 | CONCLUSION AND RECOMMENTATIONS

Genetic diversity analysis is an important step toward maximizing the use of the available IITA and CIMMYT germplasm to realize maximum heterosis with the potential to enhance genetic gain. This study revealed that there was moderate genetic diversity among the inbred lines evaluated from the two sister programs at IITA and CIMMYT in SSA. However, CIMMYT germplasm reflects a higher genetic divergence within its program, which it can still utilize together with its interested partners. In addition, this study also found that the population structure of the assessed materials has three main clusters with six subpopulations, which reflects the greater genetic variation among the materials. The 2025 DArT-Seq SNP markers were capable to group the materials in agreement with their pedigree background qualifying as high-quality markers that can be used in the preliminary classification of new unknown materials or poorly characterized materials in the maize program. Our study also provided more evidence that the predetermined HGs used at CIMMYT and IITA do not align well with each other regardless of increasing the number of entries as suggested by previous researchers. There is need to align the HGs to facilitate the ease utilization of germplasm that is obtained from both programs by the partners enable them to design IITA × CIMMYT hybrid patterns. The information generated from this study is very useful for different purposes, including parental selection for new breeding programs, inferences about heterotic grouping, and genome-wide association studies. This study would assist the breeders from the two programs in identifying inbred lines that are relatively highly heterogeneous and require purification depending on the purpose of the breeding program and the intended use of the inbred lines toward genetic gain realization. Further studies are recommended to use high-density markers to untangle all the differences between the materials at higher resolution. Also, due to the few molecular markers used, it is recommended that the results from this study be validated with the follow-up L×T CA study to justify the level of heterosis being observed from the crosses. So, it is also recommended to sample an equal number of materials from different genetic backgrounds from both centers for use in the future diversity and heterotic alignment study. Also, the CA study is recommended to determine the potential productivity of crosses among the inbred lines that will be drawn from the six subpopulations that were identified in the current study.

AUTHOR CONTRIBUTIONS

Tinovonga Gonhi: Conceptualization; data curation; formal analysis; methodology; visualization; writing—original draft; writing—review and editing. **Thomas Lapaka Odong**: Conceptualization; supervision; writing—review and editing. **Isaac Onziga Dramadri**: Conceptualization; data curation; methodology; supervision; writing—review and editing. **Mildred Ochwo-Ssemakula**: Methodology; supervision; writing—review and editing. **Zvenhamo Albert Chiteka**: Conceptualization; methodology; supervision; writing—review and editing. **Emmanuel Amponsah Adjei**: Conceptualization; data curation; formal analysis; methodology; software; supervision; visualization; writing—review and editing. **Dean Muungani**: Resources;

writing—review and editing. **Abebe Menkir**: Resources; validation; writing—review and editing. **Badu-Apraku Baffour**: Resources; validation; writing—review and editing. **Idris Adejumobi**: Writing—review and editing. **Brigitte Uwimana**: Conceptualization; data curation; investigation; methodology; project administration; resources; supervision; validation; writing—review and editing. **Thanda Dhliwayo**: Conceptualization; methodology; validation; writing—review and editing. **Dagne Wegary**: Conceptualization; formal analysis; funding acquisition; investigation; methodology; project administration; resources; supervision; validation; writing—review and editing. **John Derera**: Conceptualization; funding acquisition; investigation; methodology; project administration; resources; supervision; writing—review and editing.

ACKNOWLEDGMENTS

The study was funded by CGIAR's initiative on Accelerated Breeding at IITA and CIMMYT. Our sincere appreciation for the technical staff of Makerere University, CIMMYT-ZimbabweIITA-Uganda, and IITA-Nigeria. The first author also expresses his gratitude to the German Academic Exchange Service (DAAD) In-Region scholarship for funding his PhD tuition and stipend at Makerere University during this project. We also thank Seed Co Ltd for hosting some of the field trials in Zimbabwe.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The datasets analyzed during the current study are available at https://drive.google.com/drive/folders/1v-I3bFy5C7VleTjWFuVnhi7eSZeVt-Ee together with the Supporting Information.

ORCID

Tinovonga Gonhi https://orcid.org/0009-0000-5155-2081
Thomas Lapaka Odong https://orcid.org/0000-0002-2801-4549
Isaac Onziga Dramadri https://orcid.org/0000-0002-9423-7126
Emmanuel Amponsah Adjei https://orcid.org/0000-0002-3138-6060
Abebe Menkir https://orcid.org/0000-0002-5907-9177
Badu-Apraku Baffour https://orcid.org/0000-0003-0113-5487
Idris Adejumobi https://orcid.org/0000-0002-6533-0612
Brigitte Uwimana https://orcid.org/0000-0001-7460-9001
Thanda Dhliwayo https://orcid.org/0000-0001-8583-129X
John Derera https://orcid.org/0000-0003-3715-0689

REFERENCES

Abate, T., Fisher, M., Abdoulaye, T., Kassie, G. T., Lunduka, R., Marenya, P., & Asnake, W. (2017). Characteristics of maize cultivars in Africa: How modern are they and how many do smallholder farmers grow?. *Agriculture & Food Security*, *6*, 1–17. https://doi.org/10.1186/s40066-017-0108-6

Abu, P., Badu-Apraku, B., Ifie, B. E., Tongoona, P., Ribeiro, P. F., Obeng-Bio, E., & Offei, S. K. (2021). Genetics of extra-early-maturing yellow and orange quality protein maize inbreds and derived hybrids under low soil nitrogen and Striga infestation. *Crop Science*, *61*(2), 1052–1072. https://doi.org/10.1002/csc221401.20384

Adebayo, M. A., Menkir, A., Gedil, M., Blay, E., Gracen, V., Danquah, E., & Funmilayo, L. (2015). Diversity assessment of drought tolerant exotic and adapted maize (*Zea mays* L.) inbred lines with microsatellite markers. *Journal of Crop Science and Biotechnology*, *18*(3), 147–154. https://doi.org/10.1007/s12892-014-0076-3

Adetimirin, V. O., Vroh-Bi, I., The, C., Menkir, A., Mitchell, S. E., & Kresovich, S. (2008). Diversity analysis of elite maize inbred lines adapted to West and Central Africa using SSR markers. *Maydica*, *53*(2), 143–149.

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

Akinwale, R. O. (2021). Heterosis and heterotic grouping among tropical maize germplasm. *Cereal Grains*, *2*, 59.

Akinwale, R. O., Badu-Apraku, B., Fakorede, M. A. B., & Vroh-Bi, I. (2014). Heterotic grouping of tropical early-maturing maize inbred lines based on combining ability in Striga-infested and Striga-free environments and the use of SSR markers for genotyping. *Field Crops Research*, *156*, 48–62. https://doi.org/10.1016/j.fcr.2013.10.015

Amponsah Adjei, E., Esuma, W., Alicai, T., Bhattacharjee, R., Dramadri, I. O., Edema, R., Chamba, E. B., & Odong, T. L. (2023). Genetic diversity and population structure of Uganda's yam (*Dioscorea* spp.) genetic resource based on DArTseq. *PLoS One*, *18*(2), e0277537. https://doi.org/10.1371/journal.pone.0277537

Ayesiga, S. B., Rubaihayo, P., Oloka, B. M., Dramadri, I. O., Edema, R., & Sserumaga, J. P. (2023). Genetic variation among tropical maize inbred lines from NARS and CGIAR breeding programs. *Plant Molecular Biology Reporter*, *41*(2), 209–217. https://doi.org/10.1007/s11105-022-01358-2

Badu-Apraku, B., & Fakorede, M. A. B. (2017). Genetic diversity, heterotic grouping, and testers in hybrid maize production. *Advances in genetic enhancement of early and extra-early maize for sub-Saharan Africa* (pp. 139–184). Springer. https://doi.org/10.1007/978-3-319-64852-1_7

Badu-Apraku, B., Fakorede, M. A. B., Gedil, M., Annor, B., Talabi, A. O., Akaogu, I. C., Oyekunle, M., Akinwale, R. O., & Fasanmade, T. Y. (2016). Heterotic patterns of IITA and CIMMYT early-maturing yellow maize inbreds under contrasting environments. *Agronomy Journal*, *108*(4), 1321–1336. https://doi.org/10.2134/agronj2015.0425

Badu-Apraku, B., Fakorede, M. A. B., Gedil, M., Talabi, A. O., Annor, B., Oyekunle, M., Akinwale, R. O., Fasanmade, T. Y., Akaogu, I. C., & Aderounmu, M. (2015). Heterotic responses among crosses of IITA and CIMMYT early white maize inbred lines under multiple stress environments. *Euphytica*, *206*, 245–262. https://doi.org/10.1007/s10681-015-1506-0

Badu-Apraku, B., Garcia-Oliveira, A. L., Petroli, C. D., Hearne, S., Adewale, S. A., & Gedil, M. (2021). Genetic diversity and population structure of early and extra-early maturing maize germplasm adapted to sub-Saharan Africa. *BMC Plant Biology*, *21*(1), Article 96. https://doi.org/10.1186/s12870-021-02829-6

Berdugo-Cely, J., Valbuena, R. I., Sánchez-Betancourt, E., Barrero, L. S., & Yockteng, R. (2017). Genetic diversity and association mapping in the Colombian central collection of *Solanum tuberosum* L. Andigenum group using SNPs markers. *PLoS One*, *12*(3), e0173039. https://doi.org/10.1371/journal.pone.0173039

Bhattacharjee, R., Agre, P., Bauchet, G., De Koeyer, D., Lopez-Montes, A., Kumar, P. L., & Asiedu, R. (2020). Genotyping-by-sequencing to unlock genetic diversity and population structure in white yam (*Dioscorea rotundata* Poir.). *Agronomy*, *10*(9), 1437. https://doi.org/10.3390/agronomy10091437

Blyton, M. D. J., & Flanagan, N. S. (2006). A comprehensive guide to GenAlEx—Google Scholar. (n.d.). https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Blyton%2C+M.+D.+J.%2C+%26+Flanagan%2C+N.+S.+%282006%29.+A+comprehensive+guide+to%3A+GenAlEx+6.5.+Australian+National+University%2C+131.&btnG=

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*(19), 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Carling, J., Heller-Uszyńska, K., Jaccoud, D., Machado, A., Hopper, C., Xia, L., Vippin, C., Caig, V., Uszyński, G., & Kilian, A. (2015, January 10-14). DArTTM and DArTseqTM genome profiling for breeding, prebreeding and population genetics applications [Workshop presentation]. IXXIII *Plant and Animal Genome*, San Diego, CA, USA.

CIMMYT. (1998). *A complete listing of improved maize germplasm from CIMMYT*. CIMMYT. https://repository.cimmyt.org/server/api/core/bitstreams/22f5898c-e8c9-4f6b-9253-e23492f40ef5/content

CIMMYT. (2017). *New publications: Maize variety replacement lags in sub-Saharan Africa*. CIMMYT. https://www.cimmyt.org/news/new-publications-maize-variety-replacement-lags-in-sub-saharan-africa/

Cruz, V. M. V., Kilian, A., & Dierig, D. A. (2013). Development of DArT marker platforms and genetic diversity assessment of the U.S. collection of the new oilseed crop Lesquerella and related species. *PLoS One*, *8*(5), 1–13. https://doi.org/10.1371/journal.pone.0064062

Dao, A., Sanou, J., Gracen, V., & Danquah, E. Y. (2014). Heterotic relationship between INERA, CIMMYT and IITA maize inbred lines under drought and well-watered conditions. *Maydica*, *59*(3), 201–210.

da Silva, L. P., Mata, V. A., Lopes, P. B., Lopes, R. J., & Beja, P. (2020). High-resolution multi-marker DNA metabarcoding reveals sexual dietary differentiation in a bird with minor dimorphism. *Ecology and Evolution*, *10*(19), 10364–10373. https://doi.org/10.1002/ece3.6687

de Faria, S. V., Zuffo, L. T., Rezende, W. M., Caixeta, D. G., Pereira, H. D., Azevedo, C. F., & DeLima, R. O. (2022). Phenotypic and molecular characterization of a set of tropical maize inbred lines from a public breeding program in Brazil. *BMC Genomics*, *23*(1), 54. https://doi.org/10.1186/s12864-021-08127-7

Deschamps, S., Llaca, V., & May, G. D. (2012). Genotyping-by-sequencing in plants. *Biology*, *1*(3), 460–483. https://doi.org/10.3390/biology1030460

Dhliwayo, T., Pixley, K., Menkir, A., & Warburton, M. (2009). Combining ability, genetic distances, and heterosis among elite CIMMYT and IITA tropical maize inbred lines. *Crop Science*, *49*(4), 1201–1210. https://doi.org/10.2135/cropsci2008.06.0354

Duvick, D. N., Smith, J. S. C., & Cooper, M. (2004). Long-term selection in a commercial hybrid maize breeding program. In J. Janick (Ed.), *Plant breeding reviews* (Vol. 24, pp. 109–151). John Wiley & Sons.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Edema, R., Adjei, E. A., Ozimati, A. A., Tusiime, S. M., Badji, A., Ibanda, A., & Dramadri, I. O. (2023). Genetic diversity of cowpea parental lines assembled for breeding in uganda. *Plant Molecular Biology Reporter*, *41*(4), 713–725. https://doi.org/10.1007/s11105-023-01394-6

Egea, L. A., Mérida-García, R., Kilian, A., Hernandez, P., & Dorado, G. (2017). Assessment of genetic diversity and structure of large garlic (*Allium sativum*) germplasm bank, by diversity arrays technology "genotyping-by-sequencing" platform (DArTseq). *Frontiers in Genetics*, *8*(JUL), Article 98. https://doi.org/10.3389/fgene.2017.00098

Erenstein, O., Jaleta, M., Sonder, K., & Mottaleb, K. (2022). Global maize production, consumption and trade: Trends and R & D implications. *Food Security*, *14*, 1295–1319. https://doi.org/10.1007/s12571-022-01288-7

Ertiro, B. T., Semagn, K., Das, B., Olsen, M., Labuschagne, M., Worku, M., Wegary, D., Azmach, G., Ogugo, V., Keno, T., Abebe, B., Chibsa, T., & Menkir, A. (2017). Genetic variation and population structure of maize inbred lines adapted to the mid-altitude sub-humid maize agro-ecology of Ethiopia using single nucleotide polymorphic (SNP) markers. *BMC Genomics*, *18*(1), Article 777. https://doi.org/10.1186/s12864-017-4173-9

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Frank, K., Symphorien, A., Arfang, B., Natasha, M., Angele, I., Michael, A., Thomas, O., Peter, W., Mildred, O.-S., Geoffrey, T., Moses, B., Sadik, K., & Patrick, R. (2018). Genetic diversity and population structure of *Peronosclerospora sorghi* isolates of sorghum in Uganda. *International Journal of Environment, Agriculture and Biotechnology*, *3*(5), 1656–1667. https://doi.org/10.22161/ijeab/3.5.11

Gedil, M., & Menkir, A. (2019). An integrated molecular and conventional breeding scheme for enhancing genetic gain in maize in Africa. *Frontiers in Plant Science*, *10*(November), Article 1430. https://doi.org/10.3389/fpls.2019.01430

Guei, R. G., & Wassom, C. E. (1996). Genetic analysis of tassel size and leaf senescence and their relationships with yield in two tropical lowland maize populations. *African Crop Science Journal*, *4*(3), 275–281.

Guimarães, L. J. M., Trindade, R. S., Parentoni, S. N., & Guimarães, P. E. O. (2018). Development of maize inbred lines. *Maize Breeding*, 102–129.

Guo, R., Chen, J., Petroli, C. D., Pacheco, A., Zhang, X., San Vicente, F., Hearne, S. J., & Dhliwayo, T. (2021). The genetic structure of CIMMYT and US inbreds and its implications for tropical maize breeding.

Crop Science, *61*(3), 1666–1681. https://doi.org/10.1002/csc221401.20394

Hallauer, A. R. (1999). Temperate maize and heterosis. In J. G. Coors & S. Pandey (Eds.), *The genetics and exploitation of heterosis in crops* (pp. 353–361). ASA and CSSA.

Hammer, Ø., Harper, D. A. T., & Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, *4*(1), 4–9. http://palaeo-electronica.org/2001_1/past/issue1_01.htm

Ifie, B. E., Badu-Apraku, B., Gracen, V., & Danquah, E. Y. (2015). Genetic analysis of grain yield of IITA and CIMMYT early-maturing maize inbreds under Striga-infested and low–soil-nitrogen environments. *Crop Science*, *55*(2), 610–623. https://doi.org/10.2135/cropsci2014.07.0470

Kilian, A., Sanewski, G., & Ko, L. (2014). The application of DArT-seq technology to pineapple. In *XXIX International Horticultural Congress on Horticulture: Sustaining Lives, Livelihoods and Landscapes* (IHC2014): IV 1111 (pp. 181–188). https://doi.org/10.17660/ActaHortic.2016.1111.27

Kumar, B., Rakshit, S., Kumar, S., Singh, B. K., Lahkar, C., Jha, A. K., Kumar, K., Kumar, P., Choudhary, M., Bir Singh, S., Amalraj, J. J., Prakash, B., Khulbe, R., Kamboj, M. C., Chirravuri, N. N., & Hossain, F. (2022). Genetic diversity, population structure and linkage disequilibrium analyses in tropical maize using genotyping by sequencing. *Plants*, *11*(6), 1–14. https://doi.org/10.3390/plants11060799

Laakili, A., Belkadi, B., Gaboun, F., Yatrib, C., Makhloufi, M., El Antry, S., Medraoui, L., Laamarti, A., & Filali-Maltouf, A. (2016). Analysis of dendrometric diversity among natural populations of cork oak (*Quercus suber* l.) from Morocco. *Turkish Journal of Agriculture and Forestry*, *40*(2), 127–135. https://doi.org/10.3906/tar-1407-147

Lemi, Y., & Diro, D. (2022). Heterotic groupings, perse performance and standard heterosis of quality protein maize (*Zea mays* L.) for yield and yield contributor traits adapted at mid altitude of Ethiopia. *International Journal of Research in Agronomy*, *5*(2), 42–52. https://doi.org/10.33545/2618060X.2022.v5.i2a.112

Liu, K., Goodman, M., Muse, S., Smith, J. S., Buckler, E., & Doebley, J. (2003). Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics*, *165*(4), 2117–2128. https://doi.org/10.1093/genetics/165.4.2117

Liu, K., & Muse, S. V. (2005). PowerMaker: An integrated analysis environment for genetic maker analysis. *Bioinformatics*, *21*(9), 2128–2129. https://doi.org/10.1093/bioinformatics/bti282

Lu, Z. M., & Xu, B. Q (2010). On significance of heterotic group theory in hybrid rice breeding. *Rice Science*, *17*(2), 94–98. https://doi.org/10.1016/S1672-6308(08)60110-9

Lu, Y., Yan, J., Guimaraes, C. T., Taba, S., Hao, Z., Gao, S., & Xu, Y. (2009). Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. *Theoretical and Aapplied Genetics*, *120*, 93–115. https://doi.org/10.1007/s00122-009-1162-7

Makumbi, D., Assanga, S., Diallo, A., Magorokosho, C., Asea, G., Worku, M., & Bänziger, M. (2018). Genetic analysis of tropical midaltitude-adapted maize populations under stress and nonstress conditions. *Crop Science*, *58*(4), 1492–1507. https://doi.org/10.2135/cropsci2017.09.0531

Makumbi, D., Betrán, J. F., Bänziger, M., Ribaut, J., & Ba, M. (2011). Combining ability, heterosis and genetic diversity in tropical maize (*Zea mays* L.) under stress and non-stress conditions. *Euphytica*, *180*(2), 143–162. https://doi.org/10.1007/s10681-010-0334-5

Mammadov, J., Aggarwal, R., Buyyarapu, R., & Kumpatla, S. (2012). SNP markers and their impact on plant breeding. *International Journal of Plant Genomics*, *2012*, 728398. https://doi.org/10.1155/2012/728398

Matsuoka, Y., Mitchell, S. E., Kresovich, S., Goodman, M., & Doebley, J. (2002). Microsatellites in Zea—Variability, patterns of mutations, and use for evolutionary studies. *Theoretical and Applied Genetics*, *104*(2–3), 436–450. https://doi.org/10.1007/s001220100694

Melchinger, A. E. (2015). Genetic diversity and heterosis discussion session. In J. G. Coors & S. Pandey (Eds.), *The genetics and exploitation of heterosis in crops* (pp. 163–171). Crop Science Society of America. https://doi.org/10.2134/1999.GeneticsAndExploitation.C15

Menkir, A., Melake-Berhan, A., The, C., Ingelbrecht, I., & Adepoju, A. (2004). Grouping of tropical mid-altitude maize inbred lines on the basis of yield data and molecular markers. *Theoretical and Applied Genetics*, *108*(8), 1582–1590. https://doi.org/10.1007/s00122-004-1585-0

Mideksa, L. Y., Alamerew, S., & Tadesse, B. (2022). Hybrid performance and heterosis for yield and agronomic traits of quality protein maize (*Zea mays* L.) inbred lines adapted to mid-altitude agroecology of Ethiopia. *Agro Bali: Agricultural Journal*, *5*(2), 219–239. https://doi.org/10.37637/AB.V5I2.791

Musimwa, T. R., Molnar, T. L., Dutta, S., Dhliwayo, T., Trachsel, S., & Lee, M. (2022). Phenotypic assessment of genetic gain from selection for improved drought tolerance in semi-tropical maize populations. *Journal of Agronomy and Crop Science*, *209*(1), 71–82. https://doi.org/10.1111/jac.12592

Nadeem, M. A., Nawaz, M. A., Shahid, M. Q., Doğan, Y., Comertpay, G., & Yıldız, M. (2018). DNA molecular markers in plant breeding: Current status and recent advancements in genomic selection and genome editing. *Biotechnology & Biotechnological Equipment*, *32*(2), 261–285. https://doi.org/10.1080/13102818.2017.1400401

Obeng-Bio, E., Badu-Apraku, B., Ifie, B. E., Danquah, A., Blay, E. T., Dadzie, M. A., Noudifoulè, G. T., & Talabi, A. O. (2020). Genetic diversity among early provitamin A quality protein maize inbred lines and the performance of derived hybrids under contrasting nitrogen environments. *BMC Genetics*, *21*(1), Article 78. https://doi.org/10.1186/s12863-020-00887-7

Oyetunde, O. A., Badu-Apraku, B., Ariyo, O. J., & Alake, C. O. (2020). Efficiencies of heterotic grouping methods for classifying early maturing maize inbred lines. *Agronomy Journal*, *10*(8), 1–27. https://doi.org/10.3390/agronomy10081198

Pailles, Y., Ho, S., Pires, I. S., Tester, M., Negrão, S., & Schmöcke, S. M. (2017). Genetic diversity and population structure of two tomato species from the Galapagos islands. *Frontiers in Plant Science*, *8*(February), Article 138. https://doi.org/10.3389/fpls.2017.00138

Peakall, R., & Smouse, P. E. (2012). GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, *28*(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Perrier, X., & Jacquemoud-Collet, J. P. (2006). *DARwin: Dissimilarity analysis and representation for Windows* [Computer software]. DARwin. http://darwin.cirad.fr/darwin

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Pruitt, J. D. (2016). *A brief history of corn: Looking back to move forward* [Doctoral dissertation, University of Nebraska–Lincoln]. Doctoral

Documents from Doctor of Plant Health Program. https://www.proquest.com/openview/501a70e86170f87c27c17a175fa3302d/1?pq-origsite=gscholar&cbl=18750

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

Raman, R., Cowley, R. B., Raman, H., & Luckett, D. J. (2014). Analyses using SSR and DArT molecular markers reveal that Ethiopian accessions of white lupin (*Lupinus albus* L.) represent a unique genepool. *Open Journal of Genetics*, *4*(2), 87–98. https://doi.org/10.4236/ojgen.2014.42012

Reyes-Valdés, M. H. (2013). Informativeness of microsatellite markers. *Methods in Molecular Biology*, *1006*, 259–270. https://doi.org/10.1007/978-1-62703-389-3_18

Russell, W. A. (1909). Corn breeding. *Botanical Gazette*, *48*(5), 396–398. https://doi.org/10.1086/330056

Shull, G. H. (1908). The composition of a field of maize. *Journal of Heredity*, *os-4*(1), 296–301. https://doi.org/10.1093/jhered/os-4.1.296

Shull, G. H. (1952). Beginnings of the heterosis concept. *Heterosis*, *23*, 31–33.

Singh, R. B., Mahenderakar, M. D., Jugran, A. K., Singh, R. K., & Srivastava, R. K. (2020). Assessing genetic diversity and population structure of sugarcane cultivars, progenitor species and genera using microsatellite (SSR) markers. *Gene*, *753*(December), 144800. https://doi.org/10.1016/j.gene.2020.144800

Spinoso-Castillo, J. L., Escamilla-Prado, E., Aguilar-Rincón, V. H., Ramos-Morales, V., García de los Santos, G., Pérez-Rodríguez, P., & Corona-Torres, T. (2020). Genetic diversity of coffee (*Coffea* spp.) in Mexico evaluated by using DArTseq and SNP markers. *Genetics Resources and Crop Evolution*, *5*(67), 1795–1806. https://doi.org/10.1007/s10722-020-00940-5

Sprague, G. F., & Tatum, L. A. (1942). General vs. specific combining ability in single crosses of corn. *Agronomy Journal*, *34*(10), 923–932. https://doi.org/10.2134/AGRONJ1942.00021962003400100008X

Stress Tolerance Maize for Africa (STMA). (2020). Stress tolerant maize builds the resilience of Africa's smallholder farmers. *STMA Bulletin*. https://stma.cimmyt.org/wp-content/uploads/sites/46/2020/05/STMA-I-FINAL-Bulletin-May-14-2020.pdf

Suwarno, W. B., Pixley, K. V., Palacios-Rojas, N., Kaeppler, S. M., & Babu, R. (2014). Formation of heterotic groups and understanding genetic effects in a provitamin a biofortified maize breeding program. *Crop Science*, *54*(1), 14–24. https://doi.org/10.2135/cropsci2013.02.0096

Tomkowiak, A., Bocianowski, J., Spychała, J., Grynia, J., Sobiech, A., & Kowalczewski, P. Ł. (2021). DArTseq-based high-throughput SilicoDArT and SNP markers applied for association mapping of genes related to maize morphology. *International Journal of Molecular Sciences*, *22*(11), 5840. https://doi.org/10.3390/ijms22115840

Varshney, R. K., Chabane, K., & Hendre, P. S. (2007). Comparative assessment of EST-SSR, EST-SNP and AFLP markers for evaluation of genetic diversity and conservation of genetic resources using wild, cultivated and elite barleys. *Plant Science*, *173*(6), 638–649. https://doi.org/10.1016/j.plantsci.2007.08.010

Xia, L., Peng, K., Yang, S., Wenzl, P., Carmen de Vicente, M., Fregene, M., & Kilian, A. (2005). DArT for high-throughput genotyping of cassava (Manihot esculenta) and its wild relatives. *Theoretical and Applied Genetics*, *110*, 1092–1098. https://doi.org/10.1007/s00122-005-1937-4

Xu, Y., & Crouch, J. H. (2008). Genomics of tropical maize, a staple food and feed across the world. *Genomics of Tropical Crop Plants*, 333–370. https://doi.org/10.1007/978-0-387-71219-2_14

Warburton, M. L., Ribaut, J. M., Franco, J., Crossa, J., Dubreuil, P., & Betrán, F. J. (2005). Genetic characterization of 218 elite CIMMYT maize inbred lines using RFLP markers. *Euphytica*, *142*, 97–106. https://doi.org/10.1007/s10681-005-0817-y

Warburton, M. L., Xianchun, X., Crossa, J., Franco, J., Melchinger, A. E., Frisch, M., Bohn, M., & Hoisington, D. (1998). Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. *Crop Science*, *42*(6), 1832–1840. https://doi.org/10.2135/cropsci2002.1832

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Gonhi, T., Odong, T. L., Dramadri, I. O., Ochwo-Ssemakula, M., Chiteka, Z. A., Adjei, E. A., Muungani, D., Menkir, A., Baffour, B.-A., Adejumobi, I., Uwimana, B., Dhliwayo, T., Wegary, D., & Derera, J. (2024). Assessment of genetic diversity and heterotic alignment of CIMMYT and IITA maize inbred lines adapted to Sub-Saharan Africa. *Crop Science*, 1–19. https://doi.org/10.1002/csc2.21401