

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# Application of Microsatellites in Genetic Diversity Analysis and Heterotic Grouping of Sorghum and Maize

---

Beyene Amelework, Demissew Abakemal,  
Hussein Shimelis and Mark Laing

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/65078>

---

## Abstract

Sorghum and maize are major cereal crops worldwide and key food security crops in Sub-Saharan Africa. The difference in the mating systems, maize as predominantly a cross-fertilizer and sorghum as a self-fertilizer is reflected in differences in visible phenotypic and genotypic variations. The reproductive differences dictate the level of genetic variation present in the two crops. Conventionally, a heterotic group assignment is made based on phenotypic values estimated through combining ability and heterosis analyses. However, phenotypic evaluation methods have their limitation due to the influence of the environment and may not reflect the heterotic pattern of the lines accurately. Therefore, more effective and complementary methods have been proposed for heterotic grouping of candidate lines. Estimation of molecular-based genetic distance has proven to be a useful tool to describe existing heterotic groups, to identify new heterotic groups, and to assign inbreds into heterotic groups. Among the molecular markers, microsatellites markers have proved to be a powerful tool for analyzing genetic diversity and for classifying inbred lines into heterotic groups. Therefore, the aim of this chapter was to elucidate the use of microsatellite markers in genetic diversity analysis and heterotic grouping of sorghum and maize.

**Keywords:** genetic diversity analysis, heterotic grouping, maize, microsatellites, simple sequence repeats, sorghum

## 1. Introduction

Maize and sorghum have been more widely evaluated in genetic and cytogenetic studies than other cereal crops. Maize is one of the domesticated crop species with the highest level of molecular polymorphism. Nucleotide diversity of more than 5% has been reported at some loci of the maize genome [1], and this has been confirmed by high genetic variability in maize. The molecular diversity of maize is approximately 3- to 10-fold higher than any other domesticated grass species [2]. Several factors have been suggested as reasons for the diversity in maize including: (1) differences in the growing environments, cultivation geared toward various production systems and varied consumption preferences [3] that influenced breeding of maize varieties to severe diverse human needs worldwide; (2) high level of cross-fertilization and independent assortment of genes that led to considerable gene transfer between populations, including wild relatives; (3) presence of duplications and recombination of genes leading to creation of mutations and ultimate phenotypic variability [4]; and (4) existence of transposons and retro-transposable genetic elements leading to marked genetic variation among maize populations [5].

Similarly, sorghum is one of the most genetically diverse self-fertilizing crops. Early domestication and selection of sorghum in response to environmental factors and human needs resulted in the wide variability. The environmental factors included day length, altitude, temperature, rainfall, and soil characteristics. Humans usually required a large panicle, a nonshattering habit, large grain, tall plant height, and early maturity. The greater genetic diversity is, therefore, partly due to the diverse physical environments and partly due to the interaction of man with the environment [6]. As a result, the new and stable sorghum biotypes that have emerged can be attributed to selection, adaptation, intercrossing, and the movement of plant material from place to place. Introduction of new genotypes that have evolved in other places may result in intercrossing with the native genetic resources leading to the development of new biotypes. This movement and evolution of germplasm gave rise to five major sorghum races: bicolor, caudatum, guinea, kafir, and durra [7].

Morpho-agronomic characters of crop plants have traditionally been used for assessment of genetic variability. These characters reflect genetic variations that are manifested as visible morphological traits [8]. However, assessments based on these characters are not efficient or reliable because they are strongly affected by environmental factors. Other genetic variations are compositional or chemicals that require various tests for evaluation [9]. Isozymes [10] and seed storage proteins [11] were the most widely used biochemical markers. Since the late 1980s, analyses using various electrophoretic [12] and reversed-phase high-performance liquid chromatography (RP-HPLC) [13] of seed storage proteins have been developed and are considered effective methods for cultivar identification. Often, the importance of these types of markers is inherently impeded by low polymorphism.

The application of DNA molecular markers as compared to morphological and biochemical markers overcomes the problem of low polymorphism. DNA markers are highly informative and have facilitated the identification of agronomic traits in wild, traditional, and improved germplasm through the dissection of quantitative traits [14]. DNA-based molecular markers

are independent to environmental factors. DNA markers are fast, efficient, and robust providing clear genetic differences than phenotypic markers [15]. Several DNA marker technologies are available for determining genetic variations. Nevertheless, selection of the best marker system depends on the target species, the aim of the marker analysis, and the resource capacity [14]. PCR-based markers are widely preferred for genotype characterization in diverse crop species, including sorghum and maize, as they are relatively simple to use, nondistractive, and require small quantity of DNA, thus permitting many reactions from a single sample [16]. In addition, genetic distance (GD) estimates using molecular markers are reportedly helpful to identify the best parent combinations for new pedigree starts and to assign lines into heterotic groups [17, 18].

Molecular markers, such as restricted fragment length polymorphisms (RFLPs), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLPs), and simple sequence repeats (SSRs) or microsatellites, have been proposed as tools not only to evaluate breeding lines and hybrids and cultivars [19] but also to facilitate the monitoring of introgression, mapping of quantitative trait loci (QTLs), and the assessment of genetic diversity [20, 21] in various crops, including sorghum and maize. SSR markers have been widely applied for the assessment of genetic diversity and characterization of germplasm [22–24], identification and fingerprinting of genotypes [14], and estimation of genetic distances between and within populations [22] and to assign inbred lines into heterotic groups [25, 26]. SSR data from a number of loci have the potential to provide unique allelic profiles or DNA fingerprints for precisely establishing genotypic identity. They also have greater discriminatory power than restricted fragment length polymorphisms markers and can exhibit genetic relations that are reflective of the pedigree of the inbred lines [27]. Genotyping of inbred lines using SSRs is a reliable way of germplasm characterization which, together with morphological descriptions, leads to unambiguous differentiation of genotypes that can be utilized for a hybrid breeding program [28]. Therefore, SSR markers are the efficient marker of choice due to their ability to provide informative multiallelic loci, highly reproducible test with great powers of genotypic differentiation, which are relatively simple to use [29].

Classification of the available complementary inbreds into distinct heterotic groups is crucial to the development of superior hybrids and in developing genetic pools and breeding populations for designed breeding and genetic analyses. Exploitation of heterosis, utilization of heterotic groups, and their patterns is well established and developed in maize [30]. However, efforts to determine heterotic groups in sorghum have not been successful in clearly delineating any patterns [31]. The phenomenon of heterosis between genetically distant or unrelated genotypes has been widely reported. A heterotic group is a group of related or unrelated genotypes displaying similar combining ability effects and providing a heterotic response when crossed to other genetically distinct and complementary group [32]. Classification of inbred lines into heterotic groups based on phenotypic values could be inaccurate due to the influence of the environment and may not truly reflect the heterotic pattern of the lines. Therefore, more effective methods have been proposed for genetic grouping of candidate lines, including the use of molecular markers, line by tester analysis, and diallel crosses, among others. Recently, the use of genetic distance as indices of genetic relatedness and as a tool for

defining potential heterotic groups has been used in numerous crop plants. Simple sequence repeats (SSRs) have proved to be a powerful tool for analyzing genetic diversity and for classifying inbred lines. Therefore, the aim of this chapter was to elucidate the use of SSR markers in the heterotic grouping of the two model crops using experimental data.

## 2. Use of SSR markers in assessing genetic diversity and heterotic grouping

### 2.1. Determination of genetic diversity using SSRs in sorghum

Assessment of genetic variability in crops has a strong impact on crop improvement programs and conservation of genetic resources [33]. SSR markers appear to be particularly useful for measuring diversity, for assigning genotypes to heterotic groups, and for genetic fingerprinting [34]. The study reported by our group [26], involving 36 sorghum lines, provided clear genetic differentiation using the 30 SSR markers (**Table 1**).

No.	Sorghum			Maize			
	Name	Origin	Type	Name	Origin	Heterotic group <sup>a</sup>	Type
1	72472	Ethiopia	Restorer	142-1-eQ	Ethiopia	Ecuador	QPM
2	72482	Ethiopia	Restorer	CML144	CIMMYT	Ecuador	QPM
3	72572	Ethiopia	Restorer	CML176	CIMMYT	Unknown	QPM
4	73059	Ethiopia	Restorer	CML491	CIMMYT	A	QPM
5	75454	Ethiopia	Restorer	F7215Q	Ethiopia	Kitale	QPM
6	200538	Ethiopia	Restorer	FS111	CIMMYT	Ecuador	QPM
7	200654	Ethiopia	Restorer	FS112	CIMMYT	Unknown	QPM
8	214855	Ethiopia	Restorer	FS151-3SR	CIMMYT	Pool 9A	QPM
9	237260	Ethiopia	Restorer	FS170N	CIMMYT	Unknown	Non-QPM
10	239156	Ethiopia	Restorer	FS170Q	CIMMYT	Unknown	QPM
11	239175	Ethiopia	Restorer	FS211-1SR	CIMMYT	Kitale	QPM
12	239208	Ethiopia	Restorer	FS232N	CIMMYT	Pool 9A	Non-QPM
13	242036	Ethiopia	Restorer	FS232Q	CIMMYT	Pool 9A	QPM
14	242047	Ethiopia	Restorer	FS2-3SR	CIMMYT	Unknown	QPM
15	244712	Ethiopia	Restorer	FS4-3SR	CIMMYT	Unknown	QPM
16	244715	Ethiopia	Restorer	FS45	CIMMYT	Ecuador	QPM
17	244727	Ethiopia	Restorer	FS48	CIMMYT	Kitale	QPM
18	244733	Ethiopia	Restorer	FS48-1SR	CIMMYT	Kitale	QPM
19	211239B	Ethiopia	Restorer	FS59-2	CIMMYT	Kitale	QPM
20	214838A	Ethiopia	Restorer	FS59-4N	CIMMYT	Ecuador	Non-QPM

No.	Sorghum			Maize			
	Name	Origin	Type	Name	Origin	Heterotic group <sup>a</sup>	Type
21	214838B	Ethiopia	Restorer	FS59-4Q	CIMMYT	Ecuador	QPM
22	239167A	Ethiopia	Restorer	FS60	CIMMYT	Pool 9A	QPM
23	242039B	Ethiopia	Restorer	FS67(BC1)	CIMMYT	Kitale	QPM
24	242049A	Ethiopia	Restorer	FS67(BC2)	CIMMYT	Kitale	QPM
25	242050B	Ethiopia	Restorer	FS67-N	CIMMYT	Kitale	Non-QPM
26	244725A	Ethiopia	Restorer	FS68(BC1)	CIMMYT	Kitale	QPM
27	244725B	Ethiopia	Restorer	FS68(BC2)	CIMMYT	Kitale	QPM
28	244735A	Ethiopia	Restorer	KIT12	CIMMYT	Ecuador	QPM
29	69286A	Ethiopia	Restorer	KIT29	CIMMYT	Unknown	QPM
30	71160A	Ethiopia	Restorer	KIT31	CIMMYT	Unknown	QPM
31	72578A	Ethiopia	Restorer	KIT32N	CIMMYT	Ecuador	Non-QPM
32	73056A	Ethiopia	Restorer	KIT32Q	CIMMYT	Ecuador	QPM
33	ICSA 101	ICRISAT	A1-CMS	KIT34	CIMMYT	Ecuador	QPM
34	ICSA 743	ICRISAT	A2-CMS	SRSYN20N	CIMMYT	Pool 9A	Non-QPM
35	ICSA 749	ICRISAT	A3-CMS	SRSYN20Q	CIMMYT	Pool 9A	QPM
36	ICSA 756	ICRISAT	A4-CMS	SRSYN48	CIMMYT	Ecuador	QPM

<sup>a</sup> Putative heterotic grouping based on phenotypic data of the non-QPM counterparts before conversion to QPM.

**Table 1.** Description of the 36 sorghum and maize genotypes.

The 32 lowland sorghum lines from Ethiopia were crossed with the four cytoplasmic male-sterile (CMS) lines using a line x tester mating design. The 128 single-cross hybrids, along with the parental genotypes plus four checks, were evaluated under rainfed and irrigated conditions. A 12 × 14 incomplete block design (alpha lattice), with three replications, was used for the evaluation of the hybrids and check varieties. To determine the magnitude of heterosis and combining ability effects, the maintainer lines were used in place of their male sterile counterparts. An interrow spacing was 0.75 m and intrarow spacing of 0.30 m. Each genotype was planted in three rows of 3 m long. A 1 m pathway was used to separate between plots. Two sorghum seeds were planted per hill, and two weeks after emergence the seedlings were thinned keeping one healthy and vigorous plant.

Performance data on 128 F1 hybrids generated from these parents were used for this study. Grain yield data were recorded on both rainfed and irrigated plots. Best linear unbiased estimates (BLUEs) were made from the grain yield performance of 128 hybrids. The BLUEs of hybrid performance were calculated using trial data from two environments (rainfed and irrigated) using Genstat for Windows 17th Edition [35]. The BLUEs were then used to calculate general combining ability (GCA), specific combining ability (SCA) effects, and the level of



heterosis. Heterotic group's specific and general combining ability (HSGCA) was computed as the sum of GCA and SCA. A phylogenetic tree was constructed from the genetic distance matrix and HSGCA value using the neighbor-joining method implemented in DARwin software ver 5.0 [36].

This study detected a total of 203 putative alleles and the number of alleles per locus detected was highly variable ranging from 2 (mSbCIR223, Xcup61, and Xtxp040) to 15 (Xtxp145), with a mean of 6.8 per locus (**Table 2**). Our results were slightly higher than Folkertsma et al. [37] and Ganapathy et al. [25] but lower than Wang et al. [38] and Mutegi et al. [34]. The higher level of allelic diversity of the SSR loci examined in this study was probably associated with the wide range of genetic diversity represented in sorghum R lines sampled. The results of a  $\chi^2$  test showed significant differences in major allele frequencies with a mean major allele frequency of 0.50. This result is in congruence with the results of Wang et al. [38]. A total of 60 rare alleles, those occurring at a frequency of  $\leq 5\%$ , were detected by the 30 SSR markers. The detection of a significant number of rare alleles could be attributed to the high genetic diversity within the sorghum lines. Polymorphism information content (PIC) values ranged from 0.15 (mSbCIR223) to 0.90 (Xtxp145) with a mean of 0.63 (**Table 2**). High PIC values have been reported by others [22, 39]. Among the tested SSRs, 26 markers (87%) revealed PIC values of greater than 0.5, indicating their usefulness in discriminating between the genotypes. Observed heterozygosity ( $H_o$ ) ranged from 0.0 to 0.03, with a mean of 0.01, indicating that the test lines used in the present study were genetically pure lines, which were maintained by continued self-fertilization. The mean expected heterozygosity ( $H_e$ ) was observed to be 0.64 with maximum and minimum  $H_e$  values recorded by SSR markers, Xtxp145 (0.91) and mSbCIR223 (0.15), respectively. Expected heterozygosity was higher for test materials, suggesting that 64% of individuals are expected to be heterozygous at a given locus under random mating conditions. This can be explained by the higher outcrossing rate (5%–50%) observed in sorghum [40]. The genetic distance between the lines ranged from 0.40 to 0.80, with overall mean of 0.63 [26].

Locus	LG	Genetic parameter				
		N	A	$H_o$	$H_e$	PIC
gpsb067	H (8)	6	0.49	0.03	0.68	0.67
gpsb123	H (8)	3	0.47	0.00	0.58	0.57
mSbCIR223	B (2)	2	0.92	0.00	0.15	0.15
mSbCIR240	H(8)	10	0.46	0.03	0.76	0.75
mSbCIR276	C (3)	3	0.64	0.00	0.53	0.52
mSbCIR283	G (10)	10	0.42	0.00	0.79	0.78
mSbCIR286	A (1)	7	0.67	0.00	0.54	0.53
mSbCIR306	A (1)	3	0.50	0.00	0.56	0.55
SbAGB02	E (7)	7	0.36	0.00	0.79	0.78
Xcup02	F (9)	3	0.56	0.00	0.54	0.54

Locus	LG	Genetic parameter				
		<i>N</i>	<i>A</i>	<i>Ho</i>	<i>He</i>	PIC
Xcup14	C (3)	3	0.81	0.00	0.33	0.33
Xcup53	A (1)	3	0.67	0.00	0.51	0.50
Xcup61	C (3)	2	0.51	0.03	0.51	0.50
Xgap206	F (9)	10	0.17	0.03	0.89	0.88
Xgap72	I (6)	7	0.44	0.00	0.69	0.68
Xgap84	B (2)	10	0.33	0.00	0.81	0.80
Xtxp010	F (9)	5	0.67	0.00	0.52	0.51
Xtxp012	D (4)	9	0.24	0.03	0.84	0.83
Xtxp015	J (5)	10	0.49	0.00	0.73	0.72
Xtxp021	D (4)	6	0.60	0.03	0.61	0.60
Xtxp040	E (7)	2	0.81	0.00	0.32	0.31
Xtxp057	I (6)	6	0.44	0.00	0.73	0.72
Xtxp114	C (3)	3	0.81	0.00	0.33	0.33
Xtxp141	G (10)	11	0.29	0.03	0.84	0.83
Xtxp145	I (6)	15	0.19	0.00	0.91	0.90
Xtxp265	I (6)	8	0.28	0.00	0.81	0.80
Xtxp273	H (8)	12	0.22	0.00	0.87	0.86
Xtxp278	E (7)	4	0.68	0.03	0.51	0.50
Xtxp320	A (1)	11	0.42	0.00	0.79	0.78
Xtxp321	H (8)	12	0.42	0.00	0.79	0.78
Mean		6.77	0.50	0.01	0.64	0.63
SE		0.68	0.04	0.00	0.04	0.03

*N* = number of alleles, *A* = major allele frequency, *Ho* = observed heterozygosity, *He* = expected heterozygosity, PIC = polymorphism information content.

**Table 2.** Summary statistics for the 30 SSR loci screened across 36 sorghum genotypes.

## 2.2. Determination of genetic diversity using SSRs in maize

Comparing two marker systems (SSRs and RAPDs), researchers [23] reported that the RAPDs produced several polymorphic bands, although the resolution power of the agarose gel electrophoresis was not good enough to allow the bands of both marker systems to be seen clearly. In the study by Demissew et al. [23], the 25 RAPD markers yielded a total of 31 alleles, with an average of 1.24 alleles per locus. Only 7.5% of the RAPD primers exhibited polymorphic bands, while the majority of the markers were monomorphic. The results were consistent with the findings of Asif et al. [41]. The application of a given marker in characterizing



genotypes can be determined by the level of polymorphism it can detect and its discriminatory potential to distinguish individuals. Higher PIC value was observed for the SSR markers as compared to RAPD, reflecting the better discriminating power of SSR markers over RAPDs that makes them ideal for use in fingerprinting of maize lines as was reported by Liu et al. [42]. Garcia et al. [43] also found that the RFLP and SSR polymorphism information content means were higher than the RAPD and AFLP means.

Marker	Chrom.	RPL	N	Ho	A	PIC
nc130	5	3	3	0.000	0.333	0.404
nc133	2	5	3	0.000	0.343	0.454
phi029	3	4	3	0.029	0.443	0.410
phi046	3	4	3	0.000	0.472	0.412
phi056	1	3	4	0.030	0.561	0.633
phi065	9	5	4	0.056	0.611	0.604
phi072	4	4	4	0.056	0.306	0.401
phi075	6	2	3	0.028	0.236	0.354
phi076	4	6	6	0.143	0.600	0.663
phi079	4	5	5	0.028	0.625	0.690
phi084	10	3	2	0.056	0.333	0.346
phi102228	3	4	3	0.000	0.222	0.337
phi114	7	4	4	0.000	0.515	0.524
phi123	6	4	3	0.000	0.417	0.505
phi299852	6	3	7	0.028	0.681	0.735
phi308707	1	3	3	0.000	0.528	0.541
phi331888	5	3	4	0.028	0.458	0.512
phi374118	3	3	4	0.000	0.417	0.542
phi96100	2	4	4	0.083	0.597	0.659
umc1161	8	6	8	0.091	0.409	0.577
umc1304	8	4	3	0.143	0.386	0.380
umc1367	10	3	4	0.000	0.194	0.303
umc1545	7	4	5	0.000	0.314	0.423
umc1917	1	3	4	0.029	0.357	0.497
umc2250	2	3	2	1.000	0.500	0.375
Mean			3.9	0.073	0.434	0.491

N = number of alleles, A = minor allele frequency, Ho = observed heterozygosity, PIC = polymorphism information content.

**Table 3.** Summary statistics for the 25 SSR loci screened across 36 maize genotypes.

In another study, a total of 98 alleles, with a mean of 3.9 alleles per marker, were detected across 30 quality protein maize (QPM) and 6 non-QPM maize inbred lines using 25 SSR markers [24] (**Table 3**). The number of alleles detected in this study was in agreement with other studies [44]. Beyene et al. [45] genotyped 62 traditional Ethiopian highland maize accessions with 20 SSRs and reported a total of 98 alleles and a mean of 4.9 alleles per marker. Legesse et al. [20] reported an average of 3.9 alleles per marker by genotyping 56 highland and mid-altitude non-QPM inbred lines using 27 SSRs. Krishna et al. [15] reported a mean of 4.1 alleles using 48 SSR loci and 63 QPM inbred lines. The mean number of alleles in these studies were, however, lower than the 5.4 and 6.4 alleles previously reported by Wu et al. [46] and Yao et al. [47], respectively, but higher than the 3.3 alleles reported by Kassahun and Prasanna [48] and the 2.4–3.4 alleles reported by Babu et al. [49, 50]. The differences in mean numbers of alleles among different studies could be attributed to the type of germplasm, sample size, and repeat length of the SSRs used [24].

Demissew et al. [24] reported PIC values ranging from 0.30 (less discriminative marker, umc1367) to 0.735 (highly discriminative marker, phi299852) with a mean of 0.491 (**Table 3**). According to Botstein et al. [51] PIC guideline, 14 markers from Demissew et al. [24] were reasonably informative ( $0.30 < \text{PIC} < 0.50$ ) and the remaining 11 markers were highly informative ( $\text{PIC} > 0.50$ ). The values were comparable with previous reports by Dhliwayo et al. [52] and Mahar et al. [53] but lower than those of reported by Krishna et al. [15]. Smaller PIC values may have been due to the presence of relatively few dinucleotide repeat SSR markers [24] as opposed to a greater number of dinucleotides used in other studies [49, 50] or the presence of little genetic variability among the genotypes used in that particular study [52].

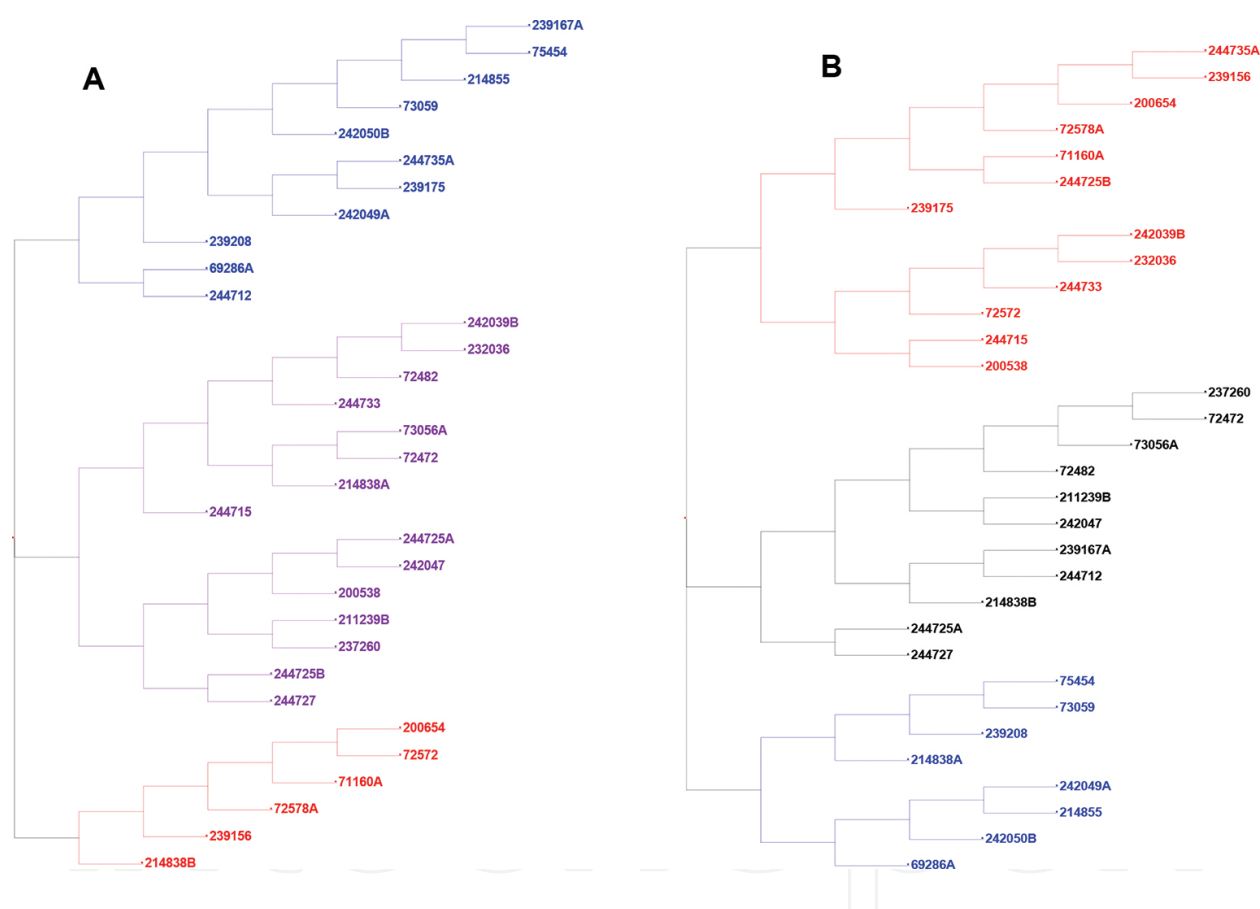
### **3. Use of SSR markers in population structure analysis and heterotic grouping**

#### **3.1. Population structure and heterotic grouping in sorghum**

In sorghum, a predominantly self-pollinated crop, the exploitation of heterosis began in the USA in the 1950s. There have been few studies on the mechanism of heterosis, heterotic grouping, and the use of molecular markers as selection criteria for parents in sorghum when compared to other crops such as maize [54]. Heterosis in sorghum has been reported in the form of increased grain, hastened flowering and maturity, increased height, and larger stems and panicles [54]. Enhanced grain yield was reported by Kambal and Webster [55] to be a product of an increased number of seeds per panicle and increased seed weight. Hybrid sorghum cultivars have been demonstrated to be more productive than pure line varieties [56]. Significant heterosis for grain yield and other agronomic traits has been reported in sorghum [57]. It has also been reported that  $F_1$  hybrids have superior buffering capacity across variable environments than pure lines in sorghum [58]. Consequently, breeding for hybrid cultivars is a better option than pure line varieties while improving sorghum grain yield.

SSR marker data have frequently been used as a tool to examine the dynamics of differentiation and population structures within germplasm collections [34, 38]. Cluster analysis using

neighbor-joining tree analysis and structure analysis can estimate the number of subpopulations and the genetic relatedness among assessed genotypes. The study by Amelework et al. [22] investigated the extent of genetic differentiation, population structure, and patterns of relationship among 200 sorghum landraces collected from lowland agro-ecology. The results obtained from both model-based population structure analysis and neighbor-joining tree analysis revealed that two group patterns existed. The two distinct subgroups resulted from farmers' selection for adaptation for the two main seasons. The results obtained from these two separate analyses support each other, with small discrepancy between groupings. Out of the 200 landraces, 32 genotypes were selected based on prior study on the basis of their relatively better yield performance and better adaptability in a moisture stress environment. They were kept homogenous through continued selfing and selection.



**Figure 1.** Dendrograms using neighbor-joining based on UPGMA depicting genetic relationship between 32 sorghum lines: (A) genetic relationship based on HSGCA value under irrigated conditions. (B) Genetic relationship based on HSGCA value under rainfed conditions. The different groups identified by specific colors (blue for lines that revealed high and positive HSGCA with ICSA 743, red for ICSA 756, purple for ICSA 749, and black ICSA 101).

Estimation of molecular-based genetic distance have been proven to be a useful way to describe existing heterotic groups, to identify new heterotic groups, and to assign inbreds of unknown genetic origin to established heterotic groups [25]. The cluster analysis carried out on the 32 lines and 4 A/B female lines, based on SSR markers, revealed three distinct groups among the 36 parental genotypes [26]. Cluster I consisted of a large number of landraces (15 genotypes).

This cluster consisted of R lines such as 242039B, 244733, 242036, 244735A, 73059, and 214855 that showed highest significant HSGCA in cross-combination with ICSA 749 and ICSA 756. This group was dominated by late flowering and high biomass lines. The high biomass was, in turn, expressed as large numbers of leaf and larger leaf width per plant. The second cluster was composed of 11 landraces and 1 CMS line. All the R lines clustered in this group except 244727, revealing positive HSGCA in cross-combination with ICSA 756. Cluster II was dominated by early flowering with small panicle and high 100 seed weight genotypes. It was reported that heterosis in sorghum is expressed as a high plant or crop growth rate as compared with the parents [59]. The third cluster (III) composed of six landraces and three CMS. This cluster consisted of three R lines (75454, 239208, and 242049A) with high and positive HSGCA in cross-combination with ICSA 743 and ICSA 756.

Heterotic groups comprise sets of genotypes that perform well when crossed with genotypes from a different heterotic group [30]. Heterotic groups in sorghum have been defined by the milo-kafir cytoplasmic genetic male-sterility system where lines are grouped either as A/B-lines or R-lines [25]. The independent cluster analysis carried out based on HSGCA value for grain yield under irrigation and rainfed condition revealed three heterotic patterns based on the distribution of the 32 lines across environments (**Figure 1A** and **B**). In this study, R and B lines did not show distinct heterotic grouping. The groupings that appeared were mainly based on the female parents. For example, in **Figure 1A**, the genotypes assigned to the first cluster (blue) had high and positive HSGCA value in a cross-combination with ICSA 743. The second group (purple) composed of 15 genotypes that showed positive HSGCA values in a cross-combination with ICSA 749. The third group (blue) was mainly represented by genotypes that revealed positive HSGCA values in a cross-combination with ICSA 756.

The extent of genetic diversity between the two parents has been proposed as a possible measure of the prediction of heterosis [60]. Although it has been suggested that the genetic distance between parents is positively correlated with heterosis of  $F_1$  hybrids, strong association has rarely been observed between heterosis and genetic distance between parents [61]. However, studies in different crops have shown moderate to strong correlation between combining ability and *per se* performance [62]. Even though this method is extensively used for prediction of heterosis, it is hypothetical and relies heavily on field evaluation. In the study of Amelework et al. [26], it was found that there were significant variations for grain yield, SCA, HSGCA, mid-, and better-parent heterosis among the 128  $F_1$  hybrids and 36 parental lines for grain yield. However, the results of the correlation analysis revealed that SSR-based genetic distance had no significant association with any of the grouping methods across environments (**Table 4**). Better-parent heterosis (BPH) under irrigated conditions had no significant correlation with SCA and HSGCA under both irrigated and rainfed conditions. On the contrary, both mid- and better-heterosis under rainfed conditions showed significant association with SCA, HSGCA, mid-parent heterosis (MPH) under irrigation, and across the two environments. The lack of significant association between genetic distance and other hybrid performance indicator in this study is also supported by other studies. In studies on rice [63], wheat [64], and grain sorghum [65], there were also nonsignificant relationships between whole genome-based genetic distance and hybrid performance. However, Boppenmaier et al.

[66] and Mosar and Lee [67] reported significant genetic relationships between genetic distance and hybrid performance of maize and oats, respectively. The prediction power of genetic distance has been inconsistent in many studies using different species and different germ-plasm [68]. This may be because of the peculiarities of many agronomic traits and lack of common phenotypic assaying methods across environments.

	SCA_C	SCA_I	SCA_R	HSGC A_C	HSGC A_I	HSGC A_R	MPH_C	MPH_I	MPH_R	BPH_C	BPH _I	BPH _R
GD	-0.07ns	-0.04ns	-0.13ns	-0.07ns	-0.05ns	-0.11ns	-0.07ns	-0.00ns	-0.09ns	-0.09ns	0.10ns	-0.14ns
SCA_C		0.89**	0.78**	0.93**	0.84**	0.74**	0.67**	0.66**	0.51**	0.61**	0.03ns	0.36**
SCA_I			0.49**	0.82**	0.93**	0.48**	0.62**	0.70**	0.36**	0.58**	0.10ns	0.23*
SCA_R				0.74**	0.51**	0.91**	0.55**	0.45**	0.53**	0.46**	0.12ns	0.35**
HSGCA_C					0.90**	0.79**	0.80**	0.78**	0.59**	0.71**	0.11ns	0.43**
HSGCA_I						0.54**	0.72**	0.83**	0.41**	0.65**	0.21*	0.29*
HSGCA_R							0.68**	0.53**	0.70**	0.58**	0.11ns	0.50**
MPH_C								0.88**	0.78**	0.95**	0.38**	0.70**
MPH_I									0.48**	0.83**	0.56**	0.41**
MPH_R										0.75**	0.11ns	0.92**
BPH_C											0.40**	0.73**
BPH_I												0.16ns
BPH_R												

GD = SSR-based genetic distance; SCA\_C = specific combining ability effects across irrigation and rainfed conditions; SCA\_I = specific combining ability effects under irrigated condition; SCA\_R = specific combining ability effects under rainfed conditions; HSGCA\_C = general plus specific combining ability effects across irrigated and rainfed conditions; HSGCA\_I = general plus specific combining ability effects under irrigated conditions; HSGCA\_R = general plus specific combining ability effects under rainfed conditions; MPH\_C = mid-parent heterosis across irrigated and rainfed conditions; MPH\_I = mid-parent heterosis under irrigated conditions; MPH\_R = mid-parent heterosis under rainfed conditions; BPH\_C = better-parent heterosis across irrigated and rainfed conditions; BPH\_I = better-parent heterosis under irrigated conditions; BPH\_R = better-parent heterosis under rainfed conditions; ns, nonsignificant.

\* Significant at 5% level of probability.

\*\* Significant at 1% level of probability.

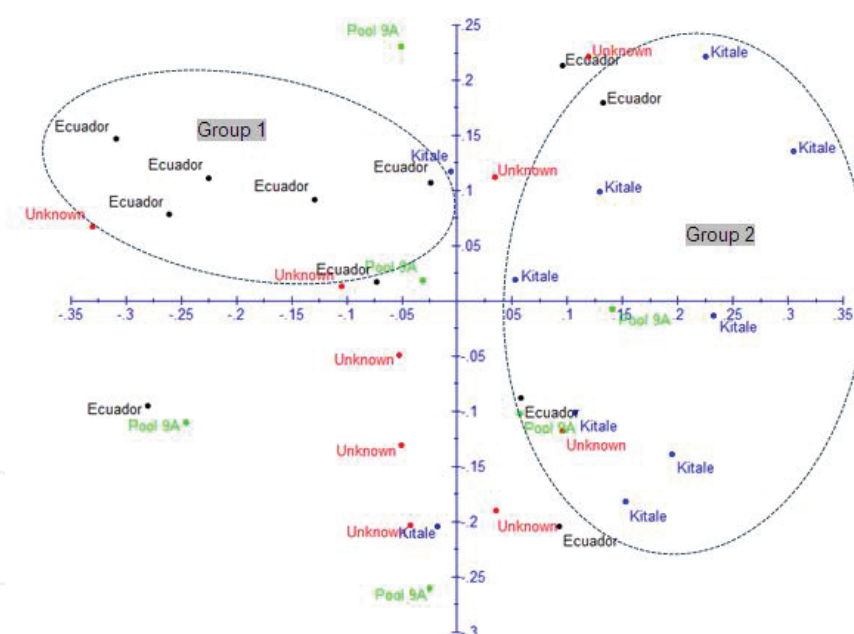
**Table 4.** Correlation matrix of the methods of heterotic grouping of the 128 hybrids yield performance under irrigated and rainfed conditions.

### 3.2. Use of SSR markers in delineation of maize population structures and heterotic groups

Genetic distance estimates are indicators of the presence or absence of relationships among genotypes. The estimates can be made using various types of molecular markers. Heterotic group assignment is often made through combining ability experiments. Also, several authors suggested the use of molecular markers in heterotic grouping [17, 18]. A comparison of SSRs and SNPs markers were carried out by Hamblin et al. [69] to characterize maize inbred lines,

to elucidate the population structure, and the genetic relationships among individuals. The authors reported that the SSRs were markers of choice than SNPs by clustering the test germplasm into populations and providing more resolution in measuring genetic distance.

A study by Demissew et al. [24] indicated the extent of genetic differentiation, population structure, and patterns of relationship among 36 maize inbred lines developed from CIMMYT source germplasm (**Table 1**). This study used 25 SSRs and applied a model-based population structure, neighbor-joining cluster, and principal coordinate analyses. All these different multivariate methods revealed the presence of two to three primary cluster groups, which was in general agreement with prior pedigree information and partly with the putative heterotic groups. The model-based population structure analysis in the same study assigned about half of the inbred lines into their putative heterotic group previously defined by breeders. There were 17, 14, and 5 inbred lines in cluster groups I, II, and III, respectively (**Figure 2**). Cluster Group I was dominated by six lines from the Ecuador heterotic group, four from the Kitale group, two from the Pool 9A group, three from previously uncategorized lines, and two CMLs (CML144 and CML491). Out of the 17 lines in Group 1, 8 of them were converted to QPM using CML176 as donor, whereas only 3 lines out of 17 were converted to QPM using CML144 as donor.



**Figure 2.** Plot of PC1 (13.0%) and PC2 (10.5%) from principal coordinate analysis of 36 inbred lines genotyped using 25 SSR markers. Lines that belong to the same heterotic group are indicated with the same color (Ecuador = black; Pool9A = green; Kitale = blue and unknown = red). *Source:* Demissew et al. [24].

Three lines in Group I were non-QPM counterparts. A mid-altitude line (F7215Q), which was converted into QPM using CML159 as donor parent, was also found in this group. Similarly, the cluster in Group II was dominated by five lines extracted from the Kitale heterotic group, four from Ecuador, four Pool9A, and one previously uncategorized line. Six lines in Group II



were converted to QPM using CML144. Five lines were converted using CML176 and the remaining three lines were again non-QPM counterparts. The other mid-altitude line (142–1eQ) which was converted into QPM using CML176 as donor parent was also found in this group. As regards cluster Group III, it included two previously uncategorized lines with CML144 being used for their conversion to QPM, one from Kitale with CML144 again used as the QPM donor, one from Pool9A where CML176 was the QPM donor, and CML176 itself. However, in the report of Bantte and Prasanna [70], it was noted that CML176 and CML144 were categorized together into one cluster group. Such incongruities with the results of other investigators in assigning inbred lines into heterotic groups may occur due to error in seed handling or pollination [71]. It may also be caused by differential selection of the different lines in different environments or genetic drift and mutation [27].

The inconsistent results in identifying heterotic pools following phenotypic evaluations during the initial phase of development of the inbred lines might have contributed to the failure of the SSR markers to categorize the remaining 50% of the inbred lines into the known heterotic groups [24]. Partial or unclear heterotic patterns were previously reported by Semagn et al. [72] in tropical and subtropical CIMMYT maize inbred lines. It was also noted from the present study that prior conversions of conventional maize inbred lines into QPM counterparts were not done systematically leading to disruption of the original heterotic system. The inbred lines from the three known heterotic groups (Kitale, Ecuador, and Pool 9A) were spread throughout the three genetic clusters (**Figure 2**).

The conversions had been done using phenotypic selections without monitoring the genetic backgrounds using molecular markers. Consequently, recombinants were selected and only a small portion of the genome of the recurrent parent was recovered. This suggested the need to use marker-assisted backcrossing (MAB) or marker-assisted selection (MAS) in the development of QPM lines through backcross procedures. Marker-assisted breeding and/or MAS can be used to facilitate background selection and to avoid disruption of newly established heterotic groups. Furthermore, earlier phenotypic selection methods used by CIMMYT could have contributed for the lack of genetic information and the partial success of the SSR markers to recognize all the available heterotic groups. In the early 1990s, broad-based genetic pools and populations were utilized by CIMMYT breeders to develop inbred lines and open pollinated varieties (OPV). Consequently, the classification of CIMMYT populations and inbred lines into heterotic groups through various mating designs has been intensified to exploit hybrid technologies using different representative testers. However, it is not easy to cluster inbred lines into their respective heterotic groups if they are extracted from similar genetic pool or source population without considering origin or heterotic pattern of inbred lines [73]. Therefore, many generations of reciprocal recurrent selection may be necessary before the lines from each heterotic group begin to significantly diverge [74].

### **3.3. Genetic purity analysis of maize lines using SSRs and implications for heterotic grouping**

In a previous study Demissew et al. [23], the genetic variability of quality protein maize (QPM) inbred lines were investigated using SSR and RAPD markers. A single SSR amplification

product (allele) per locus was expected from all the inbred lines given the high level of expected homozygosity. However, “double bands” were detected using SSR markers, which could have been masked should RAPD markers were only used in that study. The “double bands” or SSR heterozygosity indicated that some of the QPM inbred lines were not homozygous at the specific locus. This genetic background is not expected for inbred lines given that these individuals are a product of continuous and controlled selfing yielding high levels of homozygosity. The SSR markers used in the present study facilitated differentiation of homozygotic and heterozygotic alleles in the tested inbred lines sourced from the same genetic pool. The SSR profile observed in this study concurs with the reports of Bantte and Prasanna [70]. A study by Shehata et al. [75] used SSRs and analyzed the molecular diversity and heterozygosity. The authors reported that different seed sources of the same inbreds were important source of genetic variations. Also, there is a limited genetic variability that can be expected within inbred lines sourced from the same genetic sources suggesting the danger of ignoring this during sampling of inbred lines yet evolved through continued selfing. This is not uncommon in cross-fertilizing crops such as maize where a wide range of genetic variability is expected due to random crosses or mutational events over time [76].

In a related study conducted by Demissew et al. [24], the genetic purity and classification of maize inbred lines were tested using SSR markers. The authors reported 4.0–16.7% heterozygosity present among the tested inbred lines showing higher than the expected value after four generations of continuous selfing. In another study (B. Tadesse, unpublished), a total of 88 maize inbred lines were genotyped using a subset of 191 SNPs, identified for a routine quality control analysis [77]. This result showed that nearly 78% of the inbred lines showed high levels of heterozygosity. Factors such as seed admixture, pollen contamination, mislabeling of seed sources, and mixing of different seed stocks for planting are reported to be the source of heterozygous-inbred lines (K. Semagn, unpublished). The study by Warburton et al. [73] reported that bulking during maintenance breeding, seed regeneration, and contamination with seeds or pollen of other samples could possibly cause small changes in allelic frequencies. However, high levels of heterozygosity can significantly change phenotypic uniformity, heterotic patterns, and hence performance of hybrids. These may result in the distribution of mixed hybrids lacking proper genetic identity. Consequently, additional generations of selfing for all lines with high levels of heterozygosity are essential. The levels of homozygosity should be monitored frequently, especially in QPM materials, because opaque2 is a recessive gene that is liable to contamination. For new pedigree starts, such problems could be minimized by implementing a routine quality control genotyping using a subset of informative markers at different stages in a breeding program [77].

## 4. Conclusions

SSRs have been proved to be a valuable tool for diversity analysis and to assign inbred lines into heterotic groups in both sorghum and maize [22–24, 26]. SSR have greater discriminatory power than RAPDs markers, and can identify genetic relations that are reflective of the pedigree of the inbred lines. SSR markers were also found to be useful in studying the genetic

purity and the level of heterozygosity in inbred lines. Genotyping of inbred lines using SSRs is a reliable way of germplasm characterization which, together with morphological descriptions, leads to unambiguous differentiation of genotypes that can be utilized for hybrid breeding programs.

Heterotic groups in sorghum have been defined either as A/B-lines or R-lines. However, recent molecular marker-based diversity studies that utilize more detailed analyses have indicated the existence of a more complex system of genetic relationships among elite parental lines. In this study, although nonsignificant association between genetic distance and hybrid performance was observed, some patterns were detected in the distribution of sorghum genotypes. The challenges of using SSR markers as a tool for heterotic grouping in sorghum is that the genetic distance estimates can be affected by several factors such as the distribution of markers in the genome, the number of markers used, and the nature of the evolutionary mechanism underlying the variation measured. Additionally, the basic assumption for molecular diversity to predict hybrid performance is the existence of high levels of gametic phase linkage disequilibrium between yield quantitative trait loci and marker alleles. QTLs influencing heterosis in grain yield are located in certain chromosomal regions, which are unevenly distributed over the genome. Therefore, future research should focus on combined use of field-based progeny tests for yield and yield components, and molecular-based distance measurements to improve breeding efficiency. To improve prediction efficiency of molecular markers, dissecting the diversity of individual linkage groups will be exploited.

## Author details

Beyene Amelework<sup>1\*</sup>, Demissew Abakemal<sup>1,2</sup>, Hussein Shimelis<sup>1</sup> and Mark Laing<sup>1</sup>

\*Address all correspondence to: amele\_g@yahoo.com

1 African Center for Crop Improvement, University of KwaZulu-Natal, Pietermaritzburg, South Africa

2 Ethiopian Institute of Agricultural Research, Ambo-PPRC, Ethiopia, Ambo, Ethiopia

## References

- [1] Henry AM, Damerval C. High rates of polymorphism and recombination at the opaque-2 locus in cultivated maize. *Molecular and General Genetics*. 1997;256:147–157.
- [2] Buckler ES, Thornsberry JM, Kresovich S. Molecular diversity, structure and domestication of grasses. *Genetic Resources*. 2001;77:213–218.

- [3] Gomez JAA, Bellon MR, Smale M. A regional analysis of maize biological diversity in South-Eastern Guanajuato, Mexico. *Economic Botany*. 2000;54:60–72.
- [4] Helentjaris T, Weber D, Wright S. Identification of genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. *Genetics*. 1998;118:353–363.
- [5] Bennetzen JL, Jianxin MA, Devos KM. Mechanisms of recent genome size variation in flowering plants. *Annals of Botany*. 2005;95:127–132.
- [6] Rao NGP, Murty UR, Rana BS. Sorghum. In: Chapra VL, Prakash S, editors. *Evolution and Adaptation of Cereal Crops*. Science Publishers Inc., Enfield, USA; 2002. pp. 213–238. DOI: 10.1093/AOB/MCG049
- [7] Rooney WL, Smith CW. Techniques for developing new cultivars. In: Smith CY, Fredericksen RA, editors. *Sorghum, History, Technology, and Production*. Wiley, New York; 2000. pp. 329–347.
- [8] Ayana A, Bekele E. Multivariate analysis of morphological variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. *Genetic Resource and Crop Evolution*. 1999;46:378–384.
- [9] Shechter Y. Biochemical systematic study in *Sorghum bicolor*. *Bulletin of the Torrey Botanical Club*. 1975;102:334–339.
- [10] Zong JD, Gouyon PH, Sarr A, Sandmeier M. Genetic diversity and phylogenetic relations among Sahelian sorghum accessions. *Genetic Resources and Crop Evolution*. 2005;52:869–878.
- [11] Gepts P. Genetic diversity of seed storage proteins in plants. In: Brown AHD, Clegg MT, Kahler AL, Weir BS, editors. *Plant Population Genetic, Breeding and Genetics Resources*. Sinauer Associates, Inc., USA; 1990. pp. 98–115.
- [12] Tkachuk R, Mellish VJ. Wheat cultivar identification by high voltage electrophoresis. *Annals of Technology in Agriculture*. 1980;29:207–212.
- [13] Scanlon MG, Sapirstein HD, Bushuk W. Computerized wheat varietal identification by high-performance liquid chromatography. *Cereal Chemistry*. 1989;66:439–443.
- [14] Lorz H, Wenzel G. *Biotechnology in Agriculture and Forestry: Molecular Marker Systems in Plant Breeding and Crop Improvement*. Vol. 55. Springer-Verlag, Berlin; 2008. DOI: 10.1007/b137756.
- [15] Krishna MSR, Reddy SS, Naik VCB. Assessment of genetic diversity in quality protein maize (QPM) lines using simple sequence repeat (SSR) markers. *African Journal of Biotechnology*. 2012;11:16427–16433.
- [16] Soleimani VD, Baum BR, Johnson DA. Identification of Canadian durum wheat [*Triticum turgidum* L. subsp. durum (Desf.) Husn.] cultivars using AFLP and their STS markers. *Canadian Journal Plant Sciences* 2002;82:35–41.

- [17] Flint-Garcia SA, Buckler ES, Tiffin P, Ersoz E, Springer NM. Heterosis is prevalent for multiple traits in diverse maize germplasm. *PLoS One* 2009;4:e7433.
- [18] Lu Y, Yan J, Guimaraes C, Taba S, Hao Z, Gao S, Chen S, Li J, Zhang S, Vivek B, Magorokosho C, Mugo S, Makumbi D, Parentoni S, Shah T, Rong T, Crouch J, Xu Y. Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. *Theoretical and Applied Genetics*. 2009;120:93–115.
- [19] Mohanty A, Martin JP, Aguinalalde I. Chloroplast DNA study in wild populations and some cultivars of *Prunus avium* L. *Theoretical and Applied Genetic*. 2001;103:112–117.
- [20] Legesse BW, Myburg AA, Pixley KV, Botha AM. Genetic diversity of African maize inbred lines revealed by SSR markers. *Hereditas*. 2007;144:10–17.
- [21] Pooja D, Singh NK. Heterosis, molecular diversity, combining ability and their inter-relationships in short duration maize (*Zea mays* L.) across the environments. *Euphytica*. 2011;178:71–81.
- [22] Amelework B, Shimelis H, Tongoona P, Laing M, Mengistu F. Genetic variation in lowland sorghum (*Sorghum bicolor* (L.) Moench) landraces assessed by simple sequence repeats. *Plant Genetic Resources: Characterization and Utilization*. 2015;13:131–141.
- [23] Demissew A, Watson G, Shimelis H, Derera J, Twumasi-Afryie S. Comparison of two PCR-based DNA markers with high resolution melt analysis for the detection of genetic variability in selected quality protein maize inbred lines. *African Journal of Agricultural Research*. 2012;742:5692–5700.
- [24] Demissew A, Shimelis H, Derera J, Kassa S. Genetic purity and patterns of relationships among tropical highland adapted quality protein and normal maize inbred lines using microsatellite markers. *Euphytica*. 2015;204:49–61.
- [25] Ganapathy KN, Gomashe SS, Rakshit S, Prabhakar B, Ambekar SS, Ghorade RB, Biradar BD, Saxena U, Patil JV. Genetic diversity revealed utility of SSR markers in classifying parental lines and elite genotypes of sorghum (*Sorghum bicolor* L. Moench). *Australian Journal of Crop science*. 2012;6:1486–1493.
- [26] Amelework B, Shimelis H, Laing M. Genetic variation in sorghum as revealed by phenotypic and SSR markers: implications for combining ability and heterosis for grain yield. *Plant Genetic Resources* 2016:1–13. DOI:10.1017/S1479262115000696
- [27] Senior ML, Murphy JP, Goodman MM, Stuber CW. Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. *Crop Science*. 1998;38:1088–1098.
- [28] Dagne W, Vivek B, Abuschagne ML. Association of parental genetic distance with heterosis and specific combining ability in quality protein maize. *Euphytica*. 2013;191:205–216.



- [29] Van-Inghelandt D, Melchinger AE, Lebreton C, Stich B. Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. *Theoretical and Applied Genetics*. 2010;120:1289–1299.
- [30] Hallauer AR, Russell WA, Lamkey KR. Corn breeding. In: Sprague GF and Dudley JW, editors. *Corn and Corn Improvement*. ASA, Madison, WI, USA; 1988. pp. 469–564.
- [31] Gilbert ML. Identification and search for heterotic patterns in sorghum. In: Wilkinson D, editor. *Proceedings of 49th Annual Corn and Sorghum Industrial Research Conference*, 9–10 December 1994, Chicago, IL. American Seed Trade Associations, Washington, DC; 1994. pp. 117–126.
- [32] Melchinger AE. Genetic diversity and heterosis. In: Coors JG, Pandey S, editors. *The Genetics and Exploitation of Heterosis in Crops*. ASA, CSSA, and SSSA, Madison, WI; 1999. pp. 9–118.
- [33] Assar AHA, Uptmoor R, Abdelmula AA, Salih M, Ordon F, Friedt W. Genetic variation in sorghum germplasm from Sudan, ICRISAT, and USA assessed by simple sequence repeats (SSRs). *Crop Science*. 2005;45:1636–1644.
- [34] Mutegi E, Sagnard F, Semagn K, Deu M, Muraya M, Kanyenji B, deVilliers S, Kiambi D, Herselman L, Labuschagne M. Genetic structure and relationships within and between cultivated and wild sorghum (*Sorghum bicolor* (L.) Moench) in Kenya as revealed by microsatellite markers. *Theoretical Applied Genetics*. 2011;122:989–1004.
- [35] Payne RW, Murray DA, Harding SA, Baird DB, Soutar DM. *GenStat for Windows (14th Edition) Introduction*. VSN International, Hemel Hempstead, UK; 2011.
- [36] Perrier X, Jacquemoud-Collet JP. DARwin software. Dissimilarity analysis and representation for windows [Internet]. 2006. Available from: <http://www.darwin.cirad.fr/darwin.html> [Accessed: 2016-02-15].
- [37] Folkertsma RT, Frederick H, Rattunde W, Chandra S, Raju GS, Hash CT. The pattern of genetic diversity of Guinea-race *Sorghum bicolor* (L.) Moench landraces as revealed with SSR markers. *Theoretical Applied Genetic*. 2005;111:399–409.
- [38] Wang ML, Zhu CS, Barkley NA, Chen ZB, Erpelding JE, Murray SC, Tuinstra MR, Tesso T, Pederson GA, Yu JM. Genetic diversity and population structure analysis of accessions in the US historic sweet sorghum collection. *Theoretical and Applied Genetics*. 2009;120:13–23.
- [39] Muraya MM, Mutegi E, Geiger HH, de Villiers SM, Sagnard F, Kanyenji BM, Kiambi D, Parzies HK. Wild sorghum from different eco-geographic regions of Kenya display a mixed mating system. *Theoretical and Applied Genetics*. 2011, 122:1631–1639.
- [40] Barnaud A, Trigueros G, MvKey D, Joly HI. High outcrossing rates in fields with mixed sorghum landraces: how are landraces maintained? *Heredity*. 2008;101:445–452.



- [41] Asif M, Rahman M, Zafar Y. Genotyping analysis of six maize (*Zea mays* L.) hybrids using DNA fingerprinting technology. *Pakistan Journal of Botany*. 2006;38:1425–1430.
- [42] Liu K, Goodman M, Muse S, Smith JS, Buckler E, Doebley J. Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics*. 2003;65:2117–2128.
- [43] Garcia AFA, Benchimol LL, Barbosa MMA, Geraldi OI, Souza LC, Souza PA. Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. *Genetic and Molecular Biology*. 2004;27:579–588.
- [44] Makumbi D, Betran JF, Banziger M, Ribaut JM. Combining ability, heterosis and genetic diversity in tropical maize (*Zea mays* L.) under stress and non-stress conditions. *Euphytica*. 2011;180:143–162.
- [45] Beyene Y, Botha AM, Myburg AA. Genetic diversity among traditional Ethiopian highland maize accessions assessed by simple sequence repeat (SSR) markers. *Genetic Resource and Crop Evolution*. 2006;53:1579–1588.
- [46] Wu Y, Zheng Y, Sun R, Wu S, Gu H, Bi Y. Genetic diversity of waxy corn and popcorn landraces in Yunnan by SSR markers. *Acta Agronomica Sinica*. 2004;30:36–42.
- [47] Yao Q, Fang P, Kang K, Pan G. Genetic diversity based on SSR markers in maize (*Zea mays* L.) landraces from WuLing mountain region in China. *Journal of Genetics*. 2008;87:287–291.
- [48] Kassahun B, Prasanna BM. Simple sequence repeat polymorphism in Quality Protein Maize (QPM) lines. *Euphytica*. 2003;129:337–344.
- [49] Babu BK, Agrawal PK, Mahajan V, Gupta HS. Molecular and biochemical characterization of short duration quality protein maize. *Journal of Plant Biochemistry and Biotechnology*. 2009;18:93–96.
- [50] Babu BK, Pooja P, Bhatt JC, Agrawal PK. Characterization of Indian and exotic quality protein maize (QPM) and normal maize (*Zea mays* L.) inbreds using simple sequence repeat (SSR) markers. *African Journal of Biotechnology*. 2012;11:9691–9700.
- [51] Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*. 1980;32:314–331.
- [52] Dhliwayo T, Pixley K, Menkir A, Warburton M. Combining ability, genetic distances, and heterosis among elite CIMMYT and IITA tropical maize inbred lines. *Crop Science*. 2009;49:1201–1210.
- [53] Mahar K, Agrawal PK, Babu BK, Gupta HS. Assessment of genetic diversity among the elite maize (*Zea mays* L.) genotypes adapted to North-Western

Himalayan region of India using microsatellite markers. *Journal of Plant Biochemistry and Biotechnology*. 2009;18:217–220.

- [54] Quinby JR. Manifestation of hybrid vigor in sorghum. *Crop Science*. 1963;3:288–291.
- [55] Kambal AE, Webster OJ. Manifestation of hybrid vigor in grain sorghum and relations among the components of yield, weight per bushel and height. *Crop Science*. 1966;6:513–516.
- [56] Kenga R, Alabi SO, Gupta SC. Combining ability studies in tropical sorghum [*Sorghum bicolor* (L.) Moench]. *Field Crops Research*. 2004;88:251–260.
- [57] Haussmann BIG, Obilana AB, Ayiecho PO, Blum A, Schipprack W, Geiger HH. Quantitative-genetic parameters of sorghum (*Sorghum bicolor* (L.) Moench) grown in semi-arid areas of Kenya. *Euphytica*. 1999;105:109–118.
- [58] Reddy BVS, Sharma HC, Thakur RP, Ramesh S, Kumar AA. Characterization of ICRISAT-bred sorghum hybrid parents. *International Sorghum and Millets Newsletter*. 2007;48:1–123.
- [59] Blum A. Heterosis, stress, and the environment: a possible road map towards the general improvement of crop yield. *Journal of Experimental Botany*. 2013;64:4829–4837.
- [60] Zhang Q, Gao YJ, Yang S, Ragab R, Saghai Maroof MA, Li ZB. A diallele analysis of heterosis in elite hybrid rice based on RFLPs and microsatellites. *Theoretical and Applied Genetics*. 1994;89:185–192.
- [61] Rao M, Reddy GL, Kulkarni RS, Ramesh S, Lalitha RSS. Prediction of heterosis based on genetic divergence of parents through regression analysis in sunflower (*Helianthus annuus* L.). *Helia*. 2004;27:51–58.
- [62] Bertan I, Carvalho FIF, Oliveira AC. Parental selection strategies in plant breeding programs. *Journal of Crop Science and Biotechnology*. 2007;10:211–222.
- [63] Hua JP, Xing YZ, Xu CG, Sun XL, Yu SB, Zhang QF. Genetic dissection of an elite rice hybrid revealed that heterozygotes are not always advantageous for performance. *Genetics*. 2002;162:1885–1895.
- [64] Corbellini M, Perenzin M, Accerbi M, Vaccino P, Borghi B. Genetic diversity in breed wheat as revealed by coefficient of parentage and molecular markers, and its relationship to hybrid performance. *Euphytica*. 2002;123:273–285.
- [65] Jordan D, Tao Y, Godwin I, Henzell R, Cooper M, McIntyre QF. Prediction of hybrid performance in grain sorghum using RFLP markers. *Theoretical and Applied Genetics*. 2003;106:559–567.
- [66] Boppenmaier J, Melchinger AE, Brunklaus-Jung E, Geiger HH, Herrmann RG. Genetic diversity for RLFP in European maize inbreds. III. Performance of crosses with versus between heterotic groups for grain traits. *Plant Breeding*. 1992;111:217–226.

- [67] Mosar H, Lee M. RFLP variation of genealogical distance, multivariate distance, heterosis and genetic variation in oats. *Theoretical and Applied Genetics*. 1994;87:947–956.
- [68] Yu CY, Hu SW, Zhao HX, Guo AG, Sun GL. Genetic distance revealed by morphological characters, isozymes, proteins and RAPD markers and their relationships with hybrid performance in oilseed rape (*Brassica napus* L.). *Theoretical and Applied Genetics*. 2005;110:511–518.
- [69] Hamblin MT, Warburton ML, Buckler ES. Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. *PLoS One*. 2007;2:e1367.
- [70] Bantte K, Prasanna BM. Simple sequence repeat polymorphism in Quality Protein Maize (QPM) lines. *Euphytica*. 2003;129:337–344.
- [71] Rajab C, Abdolahadi H, Ghannadha MR, Warburton ML, Talei AR, Mohammadi SA. Use of SSR data to determine relationships and potential heterotic groupings within medium to late maturing Iranian maize inbred lines. *Field Crops Research*. 2006;95:212–222.
- [72] Semagn K, Magorokosho C, Vivek BS, Makumbi D, Beyene Y, Mugo S, Prasanna BM, Warburton ML. Molecular characterization of diverse CIMMYT maize inbred lines from eastern and southern Africa using single nucleotide polymorphic markers. *BMC Genomics*. 2012;13:113. DOI:10.1186/1471-2164-13-113.
- [73] Warburton ML, Ribaut JM, Franco J, Crossa J, Dubreuil P, Betran FJ. Genetic characterization of 218 elite CIMMYT maize inbred lines using RFLP markers. *Euphytica*. 2005;142:97–106.
- [74] Xia XC, Reif JC, Melchinger AE, Frisch M, Hoisington DA, Beck D, Pixley K, Warburton ML. Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers: II. Subtropical, tropical mid-altitude, and highland maize inbred lines and their relationships with elite U.S. and European maize. *Crop Science*. 2005;45:2573–2582.
- [75] Shehata AI, Al-Ghethar HA, Al-Homaidan AA. Application of simple sequence repeat (SSR) markers for molecular diversity and heterozygosity analysis in maize inbred lines. *Sudan's Journal of Biological Sciences*. 2009;16:57–62.
- [76] Hallauer AR, Carena MJ, Filho JBM. *Quantitative genetics in maize breeding*. Springer, New York; 2010.
- [77] Semagn K, Beyene Y, Makumbi D, Mugo S, Prasanna BM, Magorokosho C, Atlin G. Quality control genotyping for assessment of genetic identity and purity in diverse tropical maize inbred lines. *Theoretical and Applied Genetics*. 2012;125:1487–1501.