Vol. 19(2), pp.84-91, February, 2020
DOI: 10.5897/AJB2019.16981
Article Number: A510FEB62861
ISSN: 1684-5315

http://www.academicjournals.org/AJB

ACADEMIC JOURNALS
expand your knowledge

African Journal of Biotechnology

Full Length Research Paper

Genetic diversity among Ethiopian sorghum [*Sorghum bicolor* (L.) Moench] gene bank accessions as revealed by SSR markers

Alemu Tirfessa[1], Tesfaye Tesso[2], Asfaw Adugna[3*], Hussein Mohammed[4] and Dan Kiambi[5]

[1]Melkassa Agricultural Research Center, P. O. Box 436, Adama, Ethiopia.
[2]Department of Agronomy, Kansas State University, Manhattan, USA.
[3]Advanta Seeds Ltd., P. O. Box 10438, Eldoret, Kenya.
[4]Department of Plant Sciences, Hawassa University, P. O. Box 5, Hawassa, Ethiopia.
[5]Executive Director, African Biodiversity Conservation and Innovations Centre (ABCIC), Nairobi, Kenya.

Received 10 October 2019; Accepted 7 January 2020

The presence of genetic variation in plant populations is useful for conservation and use in breeding programs. This study was conducted to estimate the extent and patterns of genetic diversity among 200 sorghum accessions collected from different parts of Ethiopia and preserved in a gene bank. Using 39 polymorphic simple sequence repeat (SSR) markers, which were previously mapped, 261 alleles were produced with mean 6.7 alleles per SSR. Polymorphism Information Content (PIC) and Dice's similarity coefficient values ranged from 0.06 to 0.81 and from 0.062 to 0.96, respectively. Hierarchical clustering using UPGMA analysis revealed three major clusters with no clear distinction among geographical origins. Moreover, analysis of molecular variance (AMOVA) indicated that the majority of the variation (99.62%) observed was attributed to differences among accessions and only a small fraction of the total variation (0.38%) was related to regions of original collection, which may indicate that geographical origin is not a useful guide to follow for germplasm collection. Rather, agro-ecological classifications may be better for collection mission. Furthermore, crossing of accessions from the three distant clusters could result in promising genotypes for use as varieties or parents for the future breeding programs.

Key words: Allele, cluster, genetic variability, polymorphism information content, sorghum, SSR.

INTRODUCTION

Sorghum [*Sorghum bicolor* (L.) Moench] belongs to the Family Poaceae (the grass family) and the tribe

*Corresponding author email: asfaw123@rediffmail.com.

Andropogoneae (de Wet, 1978). The species is believed to have originated in northeastern quadrant of Africa, an area currently occupied by Ethiopia, Eritrea and Sudan (FAO, 1995). Evolutionary evidence exists to support this claim such as the wide range of sorghum types cultivated in the region and the diverse form of wild and weedy sorghum species still prevalent in the area (Ejeta and Grenier, 2005; Tesso et al., 2011). All the basic races of sorghum, except kaffir, are believed to be found in Ethiopia though the *durras* make the largest proportion of the race composition (Stemler et al., 1975).

Sorghum is the fifth most important cereal globally and is a staple for millions of people in Africa and South East Asia (Ejeta and Grenier, 2005). It is used for food, feed, construction and bioenergy. In Ethiopia, the grain is used for various purposes including the preparation of traditional foods in the form of different recipes, for the preparation of local beverages, and the straw is used as animal feed, for fuel and for construction. It is the second most important cereal in the country next to tef and maize (CSA, 2012).

Sorghum, in general, has tremendous genetic variability which is particularly evident in areas of early domestication or early introduction of the crop as are the situations of Ethiopia and Asia, respectively. This diversity in sorghum is readily apparent from the diverse morpho types grown in the country and the wide agro-ecological zones coverage of the crop (Doggett, 1988). Thousands of accessions of this species have been collected in the country and deposited in the international gene banks at ICRISAT in India (Reddy et al., 2006) and at USDA-ARS National Plant Germplasm System (NPGS) (Cuevas and Prom, 2013). Few of these have been sources of genes that contributed to some of the major breakthroughs in sorghum improvement programs, including high lysine (*hl*), high protein digestibility (*hpd*), and the stay green genes (Adugna, 2014).

Attempts have been made to estimate genetic diversity among Ethiopian sorghum gene bank accessions (Ayana et al., 2000), *in situ* landrace populations (Adugna, 2014) as well as sorghum germplasm pools from the African region (Agrama and Tuinstra, 2003). Although these studies have generated useful information and contributed to knowledge of sorghum germplasm in the region, further investigation to unravel the extent of genetic variability and establish the pattern of genetic relationship among the accessions is essential. The power of molecular marker technology and the steady improvement of its application will enhance further understanding of the pattern of genetic relationship in sorghum and facilitate its utilization in breeding programs. Therefore, this study was conducted to determine the extent and pattern of genetic variability among the Ethiopian sorghum germplasm accessions using simple sequence repeat markers (SSRs).

MATERIALS AND METHODS

Description of the study areas and plant materials

The sorghum growing environments in Ethiopia are traditionally classified into four broad categories representing the major agro-ecologies. These are the dry and wet lowlands with altitudes of less than 1600 m, the mid-altitude areas with elevation of between 1601 m and 1900 m and the highlands with altitudes of over 1900 m (Gebrekidan, 1981). For the sake of simplicity, the dry and the wet lowland agro-ecologies in this study were merged and represented as “lowland”. Figure 1 represents the map of Ethiopia showing the seven sites where the germplasm set was originally collected. The geographical characteristics including altitude, geographical coordinates (latitude and longitude) and agro-ecological classifications are presented in Table 1. A total of 200 sorghum accessions representing seven geographical regions within the above agro-ecological categories were received from the Ethiopian Institute of Biodiversity (EIB) for this study (Table 1, and Supplementary Table S1). The accessions were randomly drawn from over 9000 accessions maintained at EIB during 2008, but the sampling was set to represent the three major agro-ecologies.

DNA extraction, PCR amplification and genotyping

Fifty seeds from each of the 200 sorghum accessions were planted in a pot in the green house at Melkassa Agricultural Research Center, Ethiopia in May 2008. Leaf tissue samples of 15 individual four-week-old seedlings from every accession were collected, placed in medicinal bags containing silica gel, sealed and shipped to the Biosciences for East and Central Africa (BECA) laboratory in Nairobi, Kenya for DNA analysis. Total DNA was extracted from bulk sample of 15 individuals per accession using a modified CTAB protocol.

A total of 39 polymorphic sorghum SSR primer pairs, which were previously mapped, were chosen for genotyping the accessions (Table 2) based on their level of polymorphism and consistent amplification across accessions and genome coverage observed previously. PCR was run either in 96-well or 384-well plates with a total reaction volume of 5 µL that consisted of 5 ng genomic DNA, 1X magnesium free PCR buffer, and 2.0 mM $MgCl_2$, 2 pmoles of fluorescent dye-labelled forward primer labelled with FAM, VIC, NED and PET fluorescent dyes and 2 pmoles of un-labelled reverse primer, 2 mM of dNTPs, 5 Unit AmpliTaq Gold DNA polymerase (Applied Biosystems).

Touch-down PCR amplifications were performed for each primer pair separately using Gene-Amp PCR System 9600 (Applied Biosystems). This consisted of 15 min at 94°C for the initial denaturation followed by ten cycles of 94°C for 10 s, 61°C for 20 s (ramp of 1°C per cycle) and 72°C for 30 s, then by 31 cycles of 94°C for 10 s, 54°C for 20 sand 72°C for 30 s, and a final extension of 20 min at 72°C. After the PCR amplification few samples from each primer pair were randomly selected and checked for proper amplification and product intensity in 2% agarose gels after electrophoresis. PCR products were separated by co-loading post-PCR products based on fluorescent dye and/or fragment size. Depending on band intensity on agarose gel, 0.5 to 1.0 µL of PCR products from each of the 6-FAM, VIC, NED and PET-labelled PCR products were pooled together and the final volume adjusted to 10 µL by adding the required volume from a mix of an injection solution (HiDi) and size standard (GS500LIZ) (1 mL HiDi and 12 µL GS500 LIZ for 384-well plates). DNA fragments were denatured and size-fractioned using ABI 3730 Capillary DNA sequencer (PE-Applied

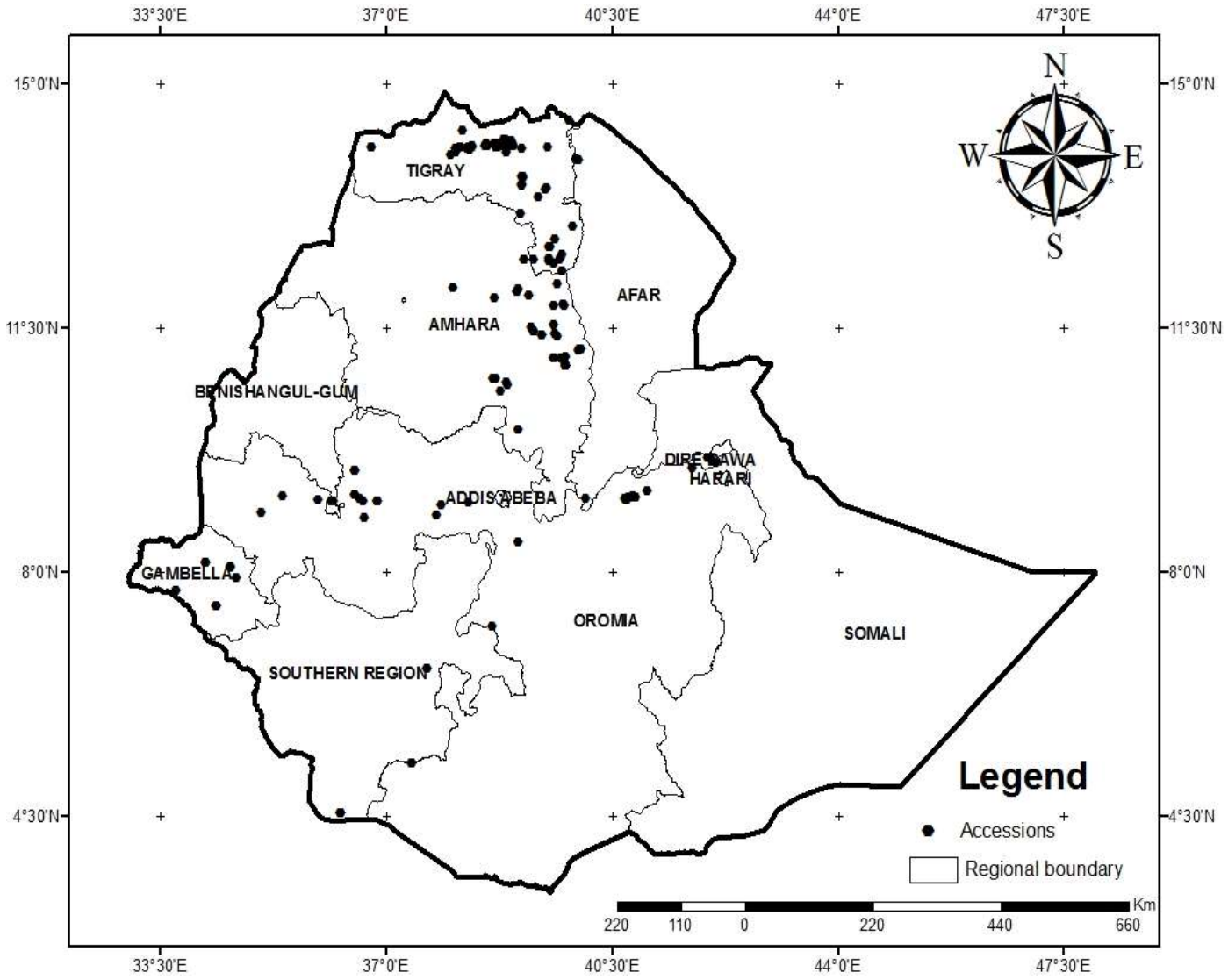

Figure 1. Map of Ethiopia showing the origin of accessions

Biosystems) as described in the user's manual. The peaks were sized and the alleles called using GeneMapper software version 3.7 and the internal size standard GS500LIZ-3730. A positive control sample (genotype BTx623) was included in all PCR for verifying the repeatability of each PCR and genotyping data (allele calls).

Data scoring and statistical analysis

Quality index, total number of alleles, maximum, minimum and abundant alleles (s) for 39 SSR markers were calculated from AlleloBin program. Those markers which showed unacceptable quality index were excluded from final analysis. Since the maximum number of alleles in a given bulked sample was higher than the two alleles expected for a diploid individual, the allele sizes data were converted into "1" and "0" where a "1" indicates the presence of a specific allele and "0" indicates its absence. Thus, all statistical analyses were performed after converting the adjusted allele calls into binary format.

Similarity coefficient was calculated as a measure of genetic similarity among accessions and used to generate a dendrogram using the hierarchical clustering and Unweighted Pair Group Method using Arithmetic Averages (UPGMA) algorithm of DARwin version 5.0 (Perrier and Jacquemoud-Collet, 2006). Principal coordinate analysis (PCoA) was used to investigate the overall variation and patterns of relationship among the accessions based on the broad agro-ecologies using DARwin. Analysis of molecular variance (AMOVA) was used to partition SSR variation among-groups and among accessions within-groups components. Significance levels for variance component estimates were computed by a non-parametric permutation procedure, using 1000 permutations. AMOVA and F_{ST} indices were calculated using the

Table 1. Geographical characteristics of the regions of collection of the sorghum germplasm accessions used in the study

Region of collection	Altitudinal class	Number of accessions	Altitude range m	Latitude	Longitude
Gamo Gofa	Mid-altitude	3	1600	05°16'	37°23'
	Lowland	7	500-1270	04°11'-06°38'	36°16'-37°38'
Hararghe	Highland	42	2090	09°05'	40°48'
	Mid-altitude	11	1640-1850	09°03'-09°30'	40°42'-41°44'
	Lowland	63	1250-1430	09°35'-09°39'	41°59'-42°06'
Illubabor	Lowland	12	400-600	07°31'-08°09'	33°45'-34°41'
Shewa	Highland	4	1980-2370	08°49'-10°03'	37°47'-39°02'
	Mid-altitude	2	1600-1900	07°14'-08°26'	38°38'-39°02'
Tigray	Highland	37	1920-2485	12°31'-14°50'	38°16'-39°60'
	Mid-altitude	2322	1600-1900	12°20'-14°21'	38°00'-39°59'
	Lowland	16	700-1570	11°57'-14°60'	36°05'-39°50'
Wollega	Highland	1	2260	08°47'	36°39'
	Mid-altitude	78	1600-1880	09°02'-09°40'	35°24'-36°51'
	Lowland	7	1340-1480	08°47'-09°40'	35°04'-36°51'
Wollo	Highland	2421	1930-2540	10°36'-12°30'	34°01'-39°45'
	Mid-altitude	14 17	1600-1880	11°02'-12°09'	39°15'-40°01'
	Lowland	23	1470-1580	10°58'-12°26'	39°25'-39°46'
Total		200			

Table 2. Summary of the 39 SSR markers used to characterize the 200 sorghum accessions in 2008.

Locus	Repeat motif	SBI chromosome	Allele size range bp		PIC	No. of alleles in this study
			Previous studies	This study		
gpsb067	GT10	SBI-08	160-190	168 - 184	0.57	6
gpsb123	CA7+GA5	SBI-08	284-304	282 -294	0.40	5
mSbCIR223	AC6	SBI-02	101-124	103 -115	0.44	4
mSbCIR238	AC26	SBI-02	69-129	73 -111	0.50	10
mSbCIR240	TG9	SBI-08	102-180	105 -201	0.29	4
mSbCIR246	CA7.5	SBI-07	86-114	93 -101	0.52	2
mSbCIR248	GT7.5	SBI-05	79-111	89 -101	0.28	6
mSbCIR262	CATG3.25	SBI-10	208-446	212 -221	0.42	3
mSbCIR276	AC9	SBI-03	221-252	227 -231	0.19	3
mSbCIR283	CT8 GT8.5	SBI-10	111-157	113 -155	0.75	12
mSbCIR286	AC9	SBI-01	104-150	104 -128	0.28	7
mSbCIR300	GT9	SBI-07	74-118	100 -112	0.18	4
mSbCIR306	GT7	SBI-01	118-126	119 -123	0.37	3
mSbCIR329	AC8.5	SBI-05	73-121	105 -115	0.35	5
SbAGB02	AG_{35}	SBI-07	92-176	92 -126	0.65	9
Xcup02	GCA6	SBI-09	186-216	194 -203	0.16	3
Xcup14	AG_{10}	SBI-03	209-251	204 - 246	0.34	3
Xcup53	$TTTA_5$	SBI-01	182-202	178 -194	0.51	3

Table 2. Contd.

Xcup61	CAG_7	SBI-03	189-204	195 -198	0.20	2
Xcup63	$GGATGC_4$	SBI-02	127-163	141 -147	0.14	2
Xgap072	AG16	SBI-06	168-229	172 -208	0.55	8
Xgap084	AG14	SBI-02	171-235	175 -211	0.73	13
Xgap206	AC13/AG20	SBI-09	86-164	98 -146	0.79	12
XISEP0310	CCAAT4	SBI-02	158-219	179 -204	0.36	2
Xtxp010	CT_{14}	SBI-09	119-155	132 -150	0.48	7
Xtxp012	CT_{22}	SBI-04	143-215	169 -203	0.64	15
Xtxp015	TC_{16}	SBI-05	197-273	199 -225	0.52	8
Xtxp021	AG_{18}	SBI-04	145-227	167 -193	0.60	11
Xtxp040	GGA_7	SBI-07	108-144	128 -137	0.32	3
Xtxp057	GT_{21}	SBI-06	213-285	220 -254	0.56	8
Xtxp114	AGG_8	SBI-03	196-245	212 -236	0.18	3
Xtxp136	GCA_5	SBI-05	238-246	237 -243	0.06	3
Xtxp141	GA_{23}	SBI-10	126-175	129 -185	0.72	12
Xtxp145	AG_{22}	SBI-06	204-278	178 -248	0.81	16
Xtxp265	GAA_{19}	SBI-06	163-246	178 -226	0.72	12
Xtxp273	TTG_{20}	SBI-08	148-243	195 -237	0.53	8
Xtxp278	TTG_{12}	SBI-07	225-319	242 -254	0.17	4
Xtxp320	AAG_{20}	SBI-01	250-329	253 -289	0.48	8
Xtxp321	$GT_4 + AT_6 + CT_{21}$	SBI-08	180-252	188 -218	0.60	12

ARLEQUIN program, version 3.11(Excoffier et al., 2005).

RESULTS

SSR polymorphism

The analysis of SSR products in this study showed fragment lengths similar to that reported in the earlier studies (Table 2). The 39 SSR markers produced a total of 261 alleles among the 200-sorghum accessions. The number of alleles per polymorphic locus ranged from two (mSbCIR246, Xcup061 and Xcup63) to 16 (Xtxp145) with mean 6.7. The Polymorphic Information Content (PIC) ranged from 0.06 (Xtxp136) to 0.81 (Xtxp145) with mean 0.45 (Table 2).

Extent and pattern of genetic diversity among accessions

Dice's similarity coefficient calculated based on 39 SSR markers ranged from 0.062 between Acc#225836 from Gamo Gofa and Acc # 226058 from Wello to 0.96 between Acc#242049 from Wello and Acc#69228 from Hararghe. These indicated that there was a fair amount of variation among the accessions studied. In addition to this, there were about 20 pairs of accessions, which showed dissimilarity value greater than 0.6.

Hierarchical clustering using UPGMA classified the accessions into three major groups (Figure 2). The first cluster (blue) consisted of 130 accessions, the second (green) 37 accessions, and the third (red) 30 accessions all mixed from different regions of collection. Moreover, the PCoA (Figure 3) revealed that there was no clear grouping of accessions according to the old agro-ecological classification, which bases altitude. Analysis of molecular variance (AMOVA) in this study, using collection region as a grouping criterion, 99.62% of the variation was accounted to among accessions while variation among regions of collection contributed only 0.38 % to the total variability observed (Table 3).

DISCUSSION

Ethiopia is endowed with rich natural resources including immense genetic diversity in cultivated plant species. *S. bicolor* is among the most diverse indigenous crop species in Ethiopia. The present study tried to uncover the extent and patterns of genetic diversity in sorghum gene bank accessions collected country wide.

Extent and patterns of genetic diversity

Wide range of variation was observed among accessions

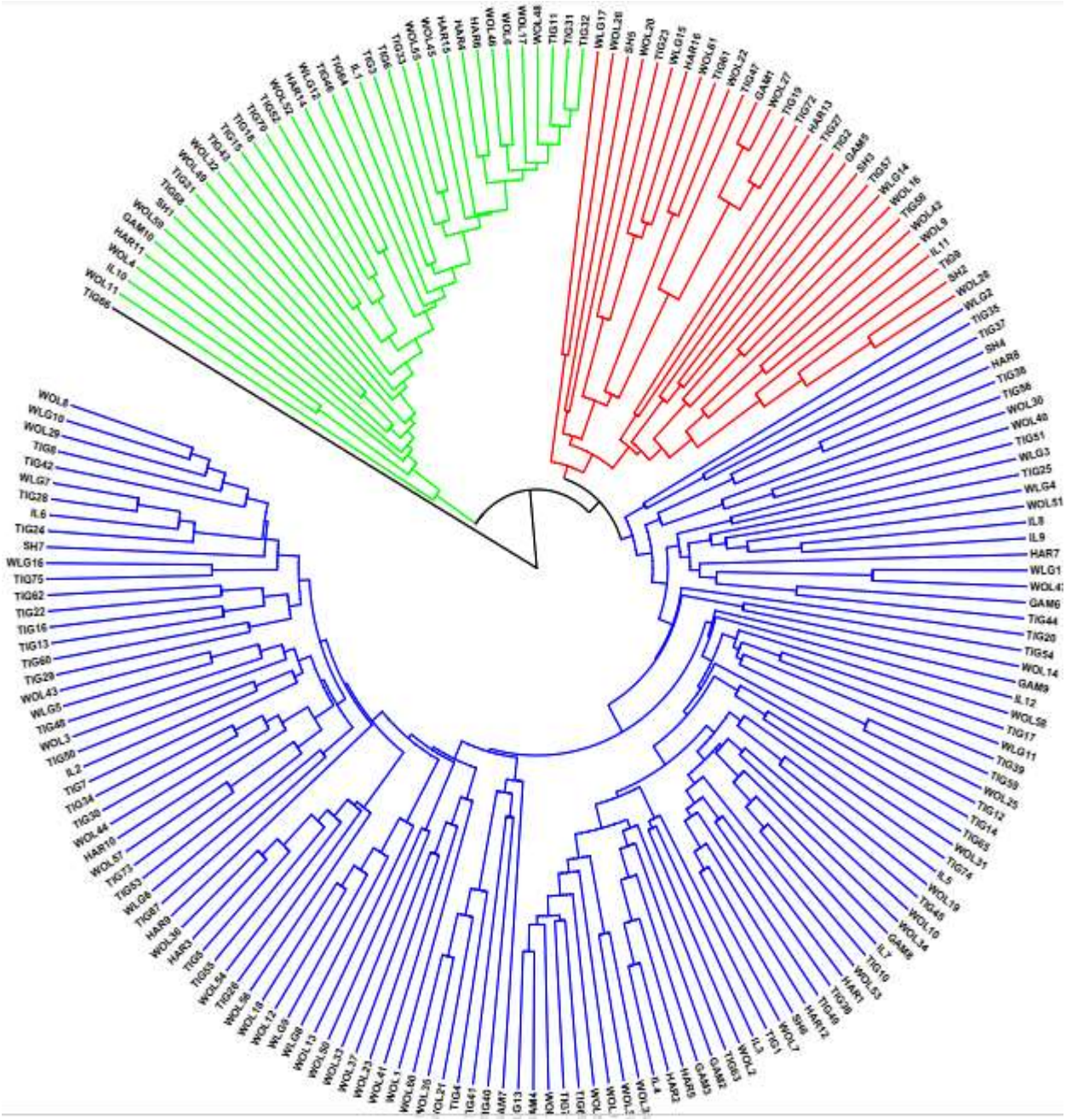

Figure 2. UPGMA analysis clustered the 200 sorghum accessions in to three broad groups, which are not related to geographical regions of origin.

for Dice's similarity coefficient calculated from 39 SSR markers. The fact that both hierarchical clustering and PCoA failed to revealed clear grouping of accessions based on either altitude or region of origin indicates these parameters are not good guides to follow for sorghum germplasm collection (Figures 2 and 3). Similar to this study, lack of complete differentiation in Ethiopian and Eritrean sorghum accessions based on geographic origin was previously reported. Although variation in the growing environment among the different regions was substantial, genetic differences attributed to region of collection was not significant. This is because

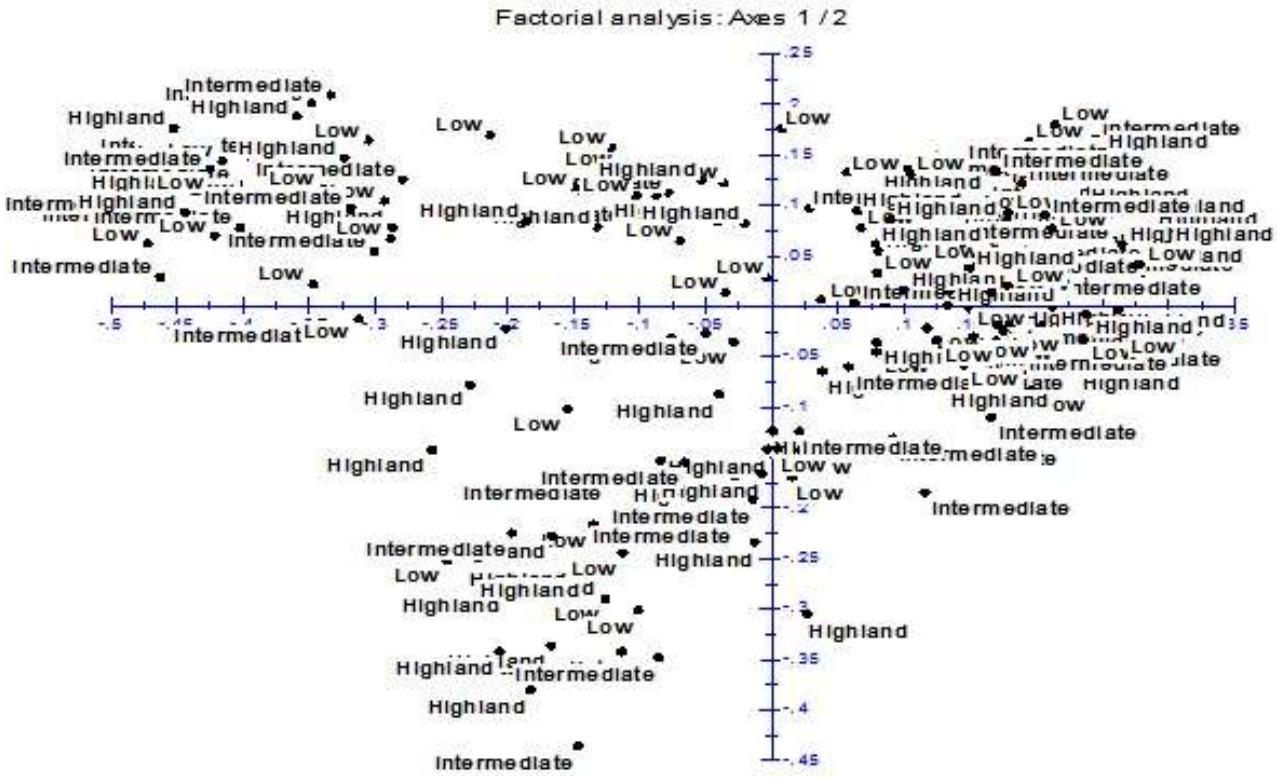

Figure 3. Principal coordinates analysis PCoA showing the clustering pattern of the 200 sorghum accessions based on altitudinal classes

Table 3. Analysis of molecular variance AMOVA describing the between and within region variability for SSR marker alleles

Source of variation	DF	Sum of squares	Variance components	Variation %	F_{ST}
Among regions	6	144.027	0.08431	0.38	
Within regions	189	4132.183	21.97970	99.62	0.00382
Total	195	4276.210	22.06400		

geographical regions in Ethiopia are not completely different in terms of climate and other soil variables, rather they share similar agro-ecologies. Adugna (2012) suggested that the latest agro-ecological zonation, which classified Ethiopia into 32 agro-ecologies out of which sorghum grows in 12 of them, were more informative to follow for sorghum germplasm collection than geographical regions of origin and the traditional classification using the four agro-ecologies (dry lowland, wet lowland, intermediate, and highland). Another reason could be because traits for adaptation to different conditions might have been the result of point mutations that are difficult to detect with the regular SSR marker assays.

The high variation among accessions observed in this study was in agreement with previous studies and confirms that sorghum is predominately inbreeder (Doggett, 1988). Ejeta and Grenier (2005) reported that genetic variation among sorghum accessions accounted for 86% of the total variation. Similarly, diversity study among 'Orange' accessions in the U.S. national collection using SSR markers showed that 90% of the total genetic variation was partitioned among accessions, while 10% of the variation was found within accessions also reported that a large proportion of genetic variation was observed within regions (88%) rather than between regions (12%) in wild sorghum from Ethiopia and Eritrea. Ng'uni et al. (2011)

also found that 82% of the total genetic variation was attributable to among accessions whereas 18% was within accessions. Adugna et al., (2013) studied *in situ* genetic diversity analysis in wild sorghum and reported that 58.8% of the variation was within populations. Similarly, Adugna (2014) reported 54.4% of the variation in cultivated sorghum landraces to be within populations.

Implications for future breeding and conservation efforts

Understanding of the extent and pattern of genetic variation can be useful for several reasons. Such information can be used to design effective *in situ* germplasm conservation and for setting germplasm collection mission as well as to estimate or predict the risk of genetic erosion in certain areas. From breeding standpoint, knowledge of the pattern of genetic variability is useful for defining heterotic groups in hybrid breeding and for relating the observed pattern with presence of certain economically important traits. Previously, the extent and pattern of genetic variability among the 200 Ethiopian sorghum [*Sorghum bicolor* (L.) Moench] germplasm accessions used in the present study were estimated using 26 morpho-agronomic traits at two locations and the result revealed the presence of wide range of genetic variability for all the 14 quantitative characters measured (Tesso et al., 2011).

The study has shown the presence of high genetic distance among some of the accessions. Moreover, UPGMA analysis clustered the 200 accessions into three large groups irrespective of the geographical regions of collection. Therefore, future crossing program can make use of the distant accessions identified in this analysis, and/or among accessions belonging to the three broad clusters. Based on the result of the current genetic diversity revealed by molecular genetic markers among the accessions, the following conclusions can be drawn: (1) the wide range of variability observed can be exploited in plant breeding to develop varieties with improved yield potential; (2) Although variation in the growing environment between the different regions was substantial, genetic differences for sorghum attributed to region of collection was not significant. Therefore, the old system of agroecological zonation may not be helpful in collection of germplasm. Rather, the latest agro-ecological zonation can serve better for this purpose.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

Adugna A (2012). Population genetics and ecological studies in wild sorghum in Ethiopia: implications for germplasm conservation. PhD Dissertation, Addis Ababa University.

Adugna A (2014). Analysis of in situ diversity and population structure in Ethiopian cultivated *Sorghum bicolor* L. landraces using phenotypic traits and SSR markers. SpringerPlus 3:212.

Adugna A, Snow AA, Sweeney PM, Bekele E, Mutegi E (2013). Population genetic structure of *in situ* wild Sorghum bicolor in its Ethiopian center of origin based on SSR markers. Genetic Resources and Crop Evolution 60:1313-1328.

Ayana A, Bryngelsson T, Bekele E (2000). Genetic variation of Ethiopian and Eritrean sorghum *Sorghum bicolor* L. Moench germplasm assessed by random amplified polymorphic DNA RAPD. Genetic Resources and Crop Evolution 47:471-482.

Agrama HA, Tuinstra MR (2003). Phylogenetic diversity and relationships among sorghum accessions using SSRs and RAPDs. African Journal of Biotechnology 2:334-340.

Central Statistical Agency (CSA) (2012). Agricultural sample survey, report on area and production of crops (private peasant holdings, Meher season), vol VI. Addis Ababa, Ethiopia.

Cuevas HE, Prom LK (2013). Assessment of molecular diversity and population structure of the Ethiopian sorghum [*Sorghum bicolor* L. Moench] germplasm collection maintained by the USDA–ARS National Plant Germplasm System using SSR markers. Genetic Resources and Crop Evolution 60:1817-1830.

de Wet JMJ (1978). Systematics and evolution of sorghum sect. Sorghum (Gramineae). American Journal of Botany 65:477-484.

Doggett H (1988). Sorghum Long Mans. 2nd ed. Green and Co ltd. London pp. 40-41.

Ejeta G, Grenier C (2005). Sorghum and its weedy hybrids. In: Gressel J (ed) Crop ferality and volunteerism. Taylor and Francis, Florida pp. 123-135.

Excoffier L, Laval G, Schneider S (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics 1:47-50.

Food and Agricultural Organization (FAO) (1995). Sorghum and millets in human nutrition. Rome: Food and Agricultural Organization, Rome, Italy. FAO Food and Nutrition Series No.27. www.**fao**.org/docrep/t0818e/t0818e00.htm Accessed 2 February 1968

Gebrekidan B (1981). Salient features of the sorghum breeding strategies used in Ethiopia. Ethiopian Journal of Agricultural Science 3:97-104.

Perrier X, Jacquemoud-Collet JP (2006). DARwin software. http://darwin.cirad.fr/darwin. Accessed on 10 August 2008.

Ng'uni D, Geleta M, Bryngelsson T (2011). Genetic diversity in sorghum *Sorghum bicolor* L. Moench accessions of Zambia as revealed by simple sequence repeats SSR. Hereditas 148:52-62.

Reddy GV, Upadhyaya HD, Gowda CLL (2006). Current status of sorghum genetic resources at ICRISAT: Their sharing and impacts. International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India.

Stemler Ann BL, Harlan JR, de Wet JMJ (1975). Evolutionary History of Cultivated Sorghums *Sorghum bicolor* [Linn.] Moench of Ethiopia. Bulletin of Torrey Botanical Club 102:325-333.

Tesso T, Tirfessa A, Grenier C, Mohammed H (2011). Association between morphological and yield components in the durra sorghums of Ethiopia. Hereditas 148:98-109.