

Genetic diversity of lowland sorghum landraces assessed by morphological and microsatellite markers

Beyene Amelework^{1*}, Hussien Shimelis¹, Pangirayi Tongoona², Mark Laing¹ and Fentahun Mengistu³

¹African Center for Crop Improvement, University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa

²West African Center for Crop Improvement, College of Basic and Applied Sciences, University of Ghana PMB LG 30, Legon, Ghana

³Ethiopian Institute of Agricultural Research, P.O. Box 2003, Addis Ababa, Ethiopia

*Corresponding author: Assefa@ukzn.ac.za (B. Amelework)

Abstract

Genetic diversity is one of the fundamental success factors in plant breeding programs. The objective of this study was to assess the genetic diversity in lowland sorghum landraces using 25 agro-morphological and 30 microsatellite markers. Phenotypic diversity of 267 genotypes originated from Ethiopia was determined using diversity indices, principal component, cluster, and discriminant analyses. High phenotypic diversity indices were recorded, ranging from 0.67 to 1.00, with a mean of 0.88. Principal component analysis and discriminant analyses identified four PCAs and five discriminant functions which contributed 82% and 92% of the total phenotypic variation among the landraces and their respective geographic origin, respectively. The PIC ranged from 0.26 to 0.88, with a mean of 0.61. The mean gene diversity was 0.69, which largely explained variation among genotypes within geographic origin. The SSR markers and phenotypic traits showed similar clustering patterns of landraces except some discrepancies. Information obtained in this study may be useful for future sorghum breeding improvement program.

Keywords: genetic diversity, microsatellite markers, phenotype, morphological traits, *Sorghum bicolor*.

Abbreviations: SSR_Simple Sequence Repeats, PCA_Principal Component Analysis, DA_Discriminant Analysis, PIC_Polymorphism Information Content, F_Fixation Index. UPGMA_Unweighted Pair Group Method with Arithmetic Mean.

Introduction

Sorghum (*Sorghum bicolor* (L) Moench) is the fifth most important food security crop serving as the staple dietary source for more than 500 million people worldwide. In 2012, the total annual worldwide sorghum production was 55 million tons from nearly 36 million ha; with a mean yield of 2.4 t ha⁻¹. Sixty percent of the production comes from Mexico, Nigeria, USA, India, Argentina and Ethiopia. In Ethiopia, sorghum is one of the staple food crops, ranking third after tef (*Eragrostis tef*) and maize (*Zea mays*), both in area of production and tonnage (FAO, 2012). In the country, sorghum is mainly used as a source of human food, animal feed, production of local beverages and the stalks for construction material at the household level (Teshome et al. 1999). Ethiopia is believed to be one of the centers of origin and diversity of sorghum (Doggett 1988). Four of the five sorghum races, wild, weedy and intermediate forms are all present in Ethiopia (Teshome et al., 1997). Stemler et al. (1977) described that sorghum as being grown over a wide range of agro-ecologies in Ethiopia. Detailed phenotypic and genetic diversity analyses of the crop need to be undertaken to understand the population dynamics of sorghum landraces across the major growing regions in Ethiopia. Ayana and Bekele (2000) sampled 415 sorghum accessions collected from different regions of Ethiopia and Eritrea and reported the existence of a high level of phenotypic variation in accessions from the regions of collections. Teshome et al. (2007) and Abdi et al. (2002) also found morphological

variation among sorghum landraces collected from North Shewa and South Welo regions. Geleta and Labuschagne (2005) evaluated the genetic diversity of some 64 sorghum accessions collected from the eastern highlands of Ethiopia.

A well-characterized germplasm is essential for crop improvement programs and strategic conservation of genetic resources. Germplasm characterization can be achieved using morphological traits and/or molecular markers. Phenotypic traits are agronomically important traits measured directly from the population. However, this method is time consuming, requires phenotyping skills, multi-locations and multi-years experimentation, to account for environmental and genotype by environment effects (Spooner et al., 2005). In sorghum, SSR markers have been widely applied for the assessment of genetic diversity and characterization of germplasm (Gamar et al., 2013) and estimation of genetic distances between and within populations (Kimani et al., 2014). SSR markers are currently the marker of choice for diversity analysis due to their ability to provide information on multi-allelic loci and greater genotypic differentiation (Anas and Yoshida, 2004). The landraces used in this study were mainly collected from the north eastern lowland region of the country where the largest genetic diversity of sorghum is present and known for being drought affected. In the region farmers developed and maintained sorghum for centuries owing to its adaptation to grow under erratic rainfall and recurrent drought stresses (Doggett, 1988).

Various authors indicated the presence of a wide morpho-agronomical variation in the Ethiopia sorghum landrace collections. However, most of the germplasm characterization in sorghum using phenotypic or limited molecular markers did not cover sorghum landraces from the lowland regions of Ethiopia where recurrent drought stress and high infestation of stalk borer are prevalent. The objective of this study was to assess genetic diversity present in lowland sorghum landraces using morphological traits and polymorphic SSR markers selected from 10 linkage groups.

Results

Agro-morphological diversity

The compact-elliptical head types of sorghum were the most predominant at 51% followed by oval heads (24%). The proportion of semi-loose dropping, semi-compact elliptical and half broom corn phenotypic classes were comparatively low among the genotypes studied. Only two and four individuals were observed with loose dropping and semi-loose erect phenotypic head type classes from the entire sample, respectively. The majority (58%) of the accessions were found to be awned. Although awned types were predominant, awnless types were also observed in all the zones (Table 1). Of the five phenotypic classes observed for midrib colour, white phenotype was predominant followed by dull green and yellow. The proportion of landraces with white (38%), dull green (30%) and yellow (2%) midrib colours were higher than the proportions of landraces with brown and purple (2%). Out of the seven phenotypic classes of grain colour used in this study, white, red and yellow grain colours were predominant at 33%, 27% and 21%, respectively. Accessions with $\frac{3}{4}$ of the grain covered with a glume were predominant (32%), while the proportion of landraces with the grain fully covered glume longer than grain was low. Glume covered with $\frac{1}{4}$ and $\frac{1}{2}$ of the grain were observed in 29% and 28% of the entries, respectively. Panicle exertion was classified into four phenotypic classes; 48% of the accessions had panicles that were exerted (2 – 10 cm), and 35% had well-exserted (> 10 cm) panicles. Plants with a light green leaf colour were the most common across the three zones at 57%, with a slightly higher proportion being observed in the North Welo collection. Across genotypes, 69% had erect and 31% drooping leaf types. The Durra race was most predominant (51%) followed by caudatum race (27%). The Bicolor race and its hybrids were the least prevalent and were absent in the Wagemra collection (Table 1). All the phenotypic characters evaluated were highly polymorphic, with the maximum and minimum diversity index scores of 0.71 and 0.99 represented by ear head compactness, and grain and midrib colour, respectively. The mean phenotypic diversity index of genotypes was 0.87, showing high variability with respect to all phenotypic character classes. Very high polymorphism were recorded for the leaf rolling (0.90) and stay-green (0.82) traits. The average diversity index (H_c) pooled over the collection zones also showed a range of variation from 0.69 for ear head compactness to 0.98 for leaf colour. Although non-significant overall diversity was observed among the zones, landraces from Wagemra revealed the maximum diversity for seven qualitative characters, including the stay green and leaf rolling traits. South Welo accessions showed greater diversity for leaf orientation and leaf colour, whereas the North Welo collections displayed greater diversity for ear head shape and midrib colour (Table 2). The first four PCAs, each with eigenvalues greater than one explained 82% of the total variation among the studied landraces for all morphological

characters (Table 3). About 34% of the total variation accounted for by PCA1 alone was due to the mean effects of panicle exertion, stay-green, ear-head compactness and grain colour. In PCA2, leaf rolling and midrib colour made significant contributions. Likewise, PCA3 that accounted for about 16% of the total variance of the landraces, originated mainly from the mean effects of awn, glume cover and colour. Variation in head shape, leaf orientation and colour constituted a large part of the total variation explained by PCA4. Discriminant analysis identified 5 functions that explained 92% of the variation among the landraces and zones of collection (Table 3). DA1 that accounted for 39% of the total variation was found to be well-correlated with leaf rolling and ear head compactness. In DA2, glume colour and cover were the most important variables that explained 18% of the total variation. DA3 and DA4 explained about 26% of the total variation and well-correlated with leaf orientation and midrib colour; and ear-head shape and grain colour, respectively. The fifth discriminant function was highly correlated with the stay-green character.

SSR polymorphism

The 200 sorghum genotypes evaluated in this study were differentiated uniquely, using 30 SSR markers. The total number of polymorphic alleles amplified per locus over all genotypes ranged from 2 (Xcup61) to 24 (mSbCIR240), while the mean was 10.5. On the other hand, the effective numbers of alleles (N_e) for 17% of the loci were less than 2 and 27% of the loci had N_e values of more than 5, with a mean of 3.7 alleles per locus. The PIC values for SSR loci ranged from 0.26 for mSbCIR223 to 0.88 for Xgap206 and Xtxp145, with the mean value of 0.61. Sixty three percent of the markers used had a PIC value greater than 0.50. The results of χ^2 test showed significant differences in allele frequencies at all loci for all the genotype sets. The mean gene diversity (H_e) was observed to be 0.56, with maximum and minimum values recorded by microsatellite markers mSbCIR223 (0.21) and Xgap206 (0.81). The fixation indices (F) for all the markers except for Xtxp265 were more than 0.75 (Table 4). A total of 316 alleles were detected in the 200 sorghum genotypes, of which 143 (45%) were rare (allele with frequency < 5%) and 75 (24%) were private (alleles that occur in individuals from only one zone). Significant differences were detected in the mean number of observed and effective alleles per locus among zones. Fifty one percent of the private alleles were present in the South Welo genotypes and the least 13% in Wagemra genotypes. The mean number of observed alleles per zone was higher for South Welo (9.2) but lower for Wagemra (6.7). Similarly, the mean number of effective alleles per locus per zone was slightly higher in South Welo (3.7) followed by North Welo (3.6) (Table 5). Significant variation was observed in allelic richness at zone level, in which allelic richness in South Welo (7.5) and genotypes collected from Wagemra (6.7), revealed the lowest number, respectively. The mean number of alleles per locus per zone ranged from 6.7 (genotypes from Wagemra) to 9.2 in genotypes from South Welo. The possibility of two randomly sampled alleles in a given genotype to be different was higher than 63% (mean gene diversity = 0.63). The highest genetic diversity (0.646 \pm 0.033) was recorded from the South Welo genotypes and the least was from Wagemra (0.625 \pm 0.037). High genetic polymorphism was observed within genotypes, showing that on average, 98% of the loci were polymorphic.

Table 1. Proportion of lowland sorghum genotypes between and within collection zones for 13 morphological traits.

Traits	Variation	Zones and percentage				Traits	Variation	Zones and percentage			
		South Welo	North Welo	Wag Hemra	Overall			South Welo	North Welo	Wag Hemra	Overall
Ear head compactness	Compact	75.8	73.5	75.7	74.9	Leaf orientation	Erect	33.1	28.6	29.7	31.0
	Loose	16.7	8.2	13.5	13.1		Dropping	66.9	71.4	70.3	69.0
	Semi compact	3.8	9.2	0.0	5.2	Leaf colour	Dark green	47.4	36.7	43.2	42.9
	Semi loose	3.8	9.2	10.8	6.7		Light green	52.6	63.3	56.8	57.1
Ear head orientation	Erect	74.2	71.4	67.6	72.3	Glume colour	White	12.0	13.3	16.2	13.1
	Re-curved	25.8	28.6	32.4	27.7		Red	0.0	4.1	0.0	1.5
	Elliptical	34.6	29.6	48.6	34.7		Purple	26.3	29.6	16.2	26.1
Ear head shape	Oblong	18.0	27.6	13.5	20.9		Black	9.0	14.3	16.2	11.9
	Round	25.6	18.4	13.5	21.3		Gray	39.1	22.4	32.4	32.1
	Semi-loose	17.3	18.4	24.3	18.7		Brown	9.8	6.1	13.5	9.0
	Loose	4.5	6.1	0.0	4.5		Dark brown	3.8	10.2	5.4	6.3
Grain colour	Red	21.1	33.7	32.4	27.2	Glume cover	25%	32.3	26.5	27.0	29.5
	Yellow	25.6	16.3	21.6	21.6		50%	30.8	28.6	16.2	28.0
	Brown	11.3	5.1	10.8	9.0		75%	27.1	33.7	43.2	31.7
	White	33.1	30.6	29.7	31.7		100%	6.8	9.2	13.5	8.6
	Light orange	5.3	7.1	0.0	5.2		> 100%	3.0	2.0	0.0	2.2
	White with orange	1.5	7.1	0.0	3.4	Panicle exsertion	< 2 cm	3.8	4.1	0.0	3.4
	White with red	2.3	0.0	5.4	1.9		2 - 10 cm	51.9	44.9	51.4	49.3
Awns	Absent	37.6	44.9	48.6	41.8	> 10 cm	33.1	38.8	37.8	35.8	
	Present	62.4	55.1	51.4	58.2	re-curved	11.3	12.2	10.8	11.6	
Midrib colour	White	40.6	31.6	48.6	38.4	Stay-green	No senescent	6.8	6.1	2.7	6.0
	Dull green	22.6	38.8	21.6	28.4		25% senescent	27.1	35.7	16.2	28.7
	Yellow	33.1	27.6	27	30.2		50% senescent	52.6	46.9	59.5	51.5
	Brown	1.5	1.0	2.7	1.5		75% senescent	13.5	11.2	21.6	13.8
	Purple	2.3	1.0	0.0	1.5		100% senescent	0.0	0.0	0.0	0.0
Race	Bicolor	5.3	8.2	0.0	5.6	Leaf rolling	Not rolled	36.1	38.8	0.0	32.1
	Caudatum	28.0	27.6	21.6	27.0		25% rolled	33.1	43.9	73.0	42.5
	Durra	54.5	44.9	56.8	51.3		50% rolled	18.8	8.2	16.2	14.6
	Durra-bicolor	3.8	5.1	0.0	3.7		75% rolled	12.0	9.2	10.8	10.8
	Durra-caudatum	8.3	14.3	21.6	12.4		100% rolled	0.0	0.0	0.0	0.0

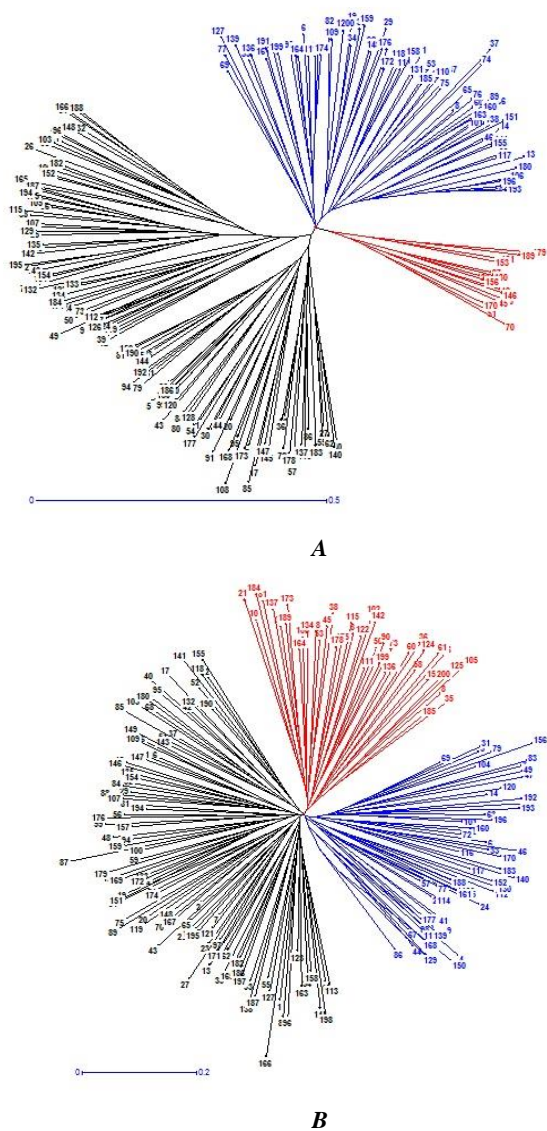


Fig 1. Neighbor-joining dendrograms based on UPGMA genetic dissimilarity depicting genetic relationship between 200 lowland sorghum genotypes: A. Classification based on phenotypic qualitative traits, B. Classification based on SSR markers. Blue line denote South Welo, Red – North Welo, Green – Waghemra.

Genetic and morphological relationships

Cluster analysis based on genetic dissimilarity using the neighbor-joining method in DARwin 5.0, based on SSR and morphological markers classified the 200 sorghum genotypes into three distinct groups (Fig 1). However, the cluster analysis using SSR markers further clustered the genotypes into clearer and more defined subgroups than morphological traits. The clustering patterns in both cases did not correspond with the putative population structure based on the zones of collection. This could be expected because most of the genotypes were collected from lowland agro-ecologies of the different zones in the North East Ethiopia. Group I consisted of 49% (97) of the total genotypes and composed of three sub-clusters (IA, IIA and IIIA) (Fig 1B). Group II contained 38% (75) of the total genotypes and it also comprised of two sub-clusters (IIA and IIB). In Group I, South Welo constituted 53%, North Welo (41%) and

Waghemra (50%) of the genotypes tested. On the other hand, the cluster analysis based on morphological traits grouped 40, 32 and 29% of the genotypes into three separate groups (Fig 1A). Group I was dominated by sorghum genotypes with compact elliptical ear-head types at 89%. Semi-loose sorghum genotypes (51%) were common in Group II, whereas in Group III genotypes which were compact round and with re-curved panicles (81%) were the predominate types.

Discussion

A wide range of variability was observed in qualitative traits among lowland sorghum landraces and zones of collection. Panicle and grain related traits are frequently used by farmers in varietal selection and in the naming of sorghum landraces in Ethiopia (Teshome et al., 1997). The predominance of compact head types was related to the prevalence of the durra race which is characterized by large seed size, high panicle weight and grain yield. The present results are similar to those of Geleta and Labuschagne (2005). It was also indicated that very compact panicles were a distinctive characteristic of extremely dry conditions where compact durra sorghum races are frequent (Grenier et al., 2004). White, red and yellow grain colours were predominant in the north eastern parts of Ethiopia. In the present study, awned sorghum types were common in sorghum accessions of lowland origins. This suggested that awns may have a positive link with drought tolerance. In small cereals, ear awns, had a positive and significant correlation with grain yield and is a useful morphological marker while selecting lines for high yield under moisture stress condition (Taketa et al., 2011). Stay-green was scored as described by Xu et al. (2000), the majority of the landraces showed slightly to ½ of the leaves senescent. No landrace was observed where complete foliar death. This suggests that, in these landrace samples, there was a strong presence of stay-green genes, which could be used as phenotypic markers for future breeding for drought tolerance. Midrib colour is one of the most important traits used by the farmers to differentiate between sweet stem and grain sorghum varieties. In Ethiopia some varieties of sorghum were grown for their food and brewing quality. Other varieties are grown for their sweet stalk and high fresh and dry biomass. In recent years, brown midrib sorghum types have been reported to have reduced lignin content and increase forage digestibility for animals (Rao et al., 2012). The present study identified four genotypes (72498-A, 239146-C, 244740-A and 202508) with a brown midrib, and these could be used to develop high biomass cultivars with good fodder quality. The diversity indices differed among the zones, and all zones showed from high to average scores for all the phenotypic characters. The South Welo Zone is a relatively lowland wet area with a mean annual rainfall of 600-1000 mm, and a growing duration of 110-150 days, whereas North Welo and Waghemra Zones are lowland dry areas with a mean annual rainfall of less than 600 mm, and a growing duration of 90-130 days (Ayana and Bekele, 2000). The maximum variation in stay-green, awns and leaf rolling traits observed in North Welo and Waghemra could be attributed to frequent moisture stresses events at these environments. These genotypes also showed slightly lower mean numbers of observed and effective alleles. This may have resulted from contentious selection for genotypes with their superior drought tolerance. The continued selection of these genotypes by farmers would result in a narrowing of genetic diversity, and a reduced

Table 2. Diversity indices estimated using 13 morphological traits of lowland sorghum landraces collected from three zones.

Traits	Zone			Diversity index ^a			
	South Welo	North Welo	Wag Hemra	H'	H _{cl}	H _{cl} /H'	(H'-H _{cl})/H'
Head compactness	0.67	0.70	0.76	0.71	0.69	0.97	0.03
Head orientation	0.82	0.86	0.91	0.85	0.87	1.02	-0.02
Head shape	0.91	0.94	0.89	0.92	0.91	1.00	0.00
Grain colour	0.82	0.87	0.90	0.83	0.86	1.04	-0.04
Awns	0.96	0.99	1.00	0.98	0.98	1.00	0.00
Glume colour	0.86	0.91	0.94	0.87	0.90	1.04	-0.04
Glume cover	0.85	0.85	0.92	0.85	0.88	1.03	-0.03
Leaf orientation	0.92	0.86	0.88	0.89	0.89	0.99	0.01
Leaf colour	1.00	0.95	0.99	0.99	0.98	0.99	0.01
Midrib colour	0.98	0.99	0.95	0.99	0.97	0.98	0.02
Panicle exertion	0.78	0.80	0.87	0.78	0.82	1.05	-0.05
Stay-green	0.75	0.82	0.83	0.82	0.80	0.97	0.03
Leaf rolling	0.70	0.83	0.94	0.90	0.82	0.91	0.09
Overall diversity index (H')	0.87	0.86	0.88	0.90	0.87	1.00	0.00

^a H' = Diversity index, H_{cl} = Average diversity index of each character for the collection zones, H_{cl}/H' = Proportion of diversity within zone, (H'-H_{cl})/H' = Proportion of diversity between zones

Table 3. Principal component (PC) and discriminant function (DA) scores for 13 qualitative characters in sorghum landraces

Characters	Principal component				Discriminant function				
	PC1	PC2	PC3	PC4	DA1	DA2	DA3	DA4	DA5
Leaf rolling	0.08	0.87	0.20	-0.05	0.57	0.19	0.15	0.24	-0.39
Head compactness	0.89	-0.10	0.22	-0.12	0.55	-0.34	0.17	-0.31	0.31
Glume cover	-0.05	0.12	-0.90	-0.39	-0.24	0.45	0.31	0.23	-0.32
Glume colour	-0.06	-0.72	0.55	-0.07	0.16	0.41	-0.01	-0.25	-0.21
Leaf orientation	0.10	0.49	0.03	0.63	0.06	0.04	0.55	0.11	0.11
Midrib colour	-0.25	0.86	-0.11	0.01	-0.30	0.02	0.47	-0.02	0.11
Panicle exertion	-0.72	-0.24	-0.19	-0.02	-0.12	0.22	-0.46	0.11	0.43
Head shape	-0.16	-0.03	0.03	0.94	-0.10	0.27	-0.19	0.62	0.06
Grain colour	0.55	-0.47	0.40	0.03	-0.12	-0.31	-0.14	0.50	0.32
Stay-green	0.95	-0.17	-0.12	0.18	0.42	0.32	-0.11	-0.28	0.54
Awns	0.31	0.11	0.92	-0.07	0.15	-0.27	0.12	-0.13	0.32
Leaf colour	-0.44	0.29	-0.29	-0.69	0.00	0.37	0.21	0.08	0.25
Head orientation	0.02	-0.07	-0.10	0.09	0.05	0.01	0.18	0.06	0.12
Eigenvalues	4.06	2.40	1.92	1.50	0.21	0.09	0.08	0.06	0.05
% total variance	33.82	20.01	15.98	12.50	38.70	18.00	14.90	10.70	9.70
% cumulative variance	33.82	53.83	69.81	82.31	38.70	56.70	71.60	82.30	92.00

allelic richness when compared to the other landraces from other areas. The PIC values of loci provide an estimate of the discriminatory power of loci, considering the number of alleles and their relative frequencies. Forty percent of the loci used in this study had PIC values of more than 0.75, indicating that these loci have high discriminatory powers that distinctively classified most of the genotypes. The polymorphism values reported in this study were higher than those reported by Ganapathy et al. (2012). In general, the current results concur with the results of Smith et al. (2000) and Geleta et al. (2006), who also showed that markers with moderate PIC values can classify most of the inbred lines and detect the polymorphism rate at a specific locus (Smith et al., 2000). The morphological traits of ear-head shape, compactness and orientation played a vital role in the delineation of the diversity into different groups. Based on the above mentioned traits, two groups of genotypes were distinguished. The first group was dominated by elliptical and/or round compact heads. The second group was mainly composed of semi-loose heads types. Genetic analysis employed in this study also showed good correspondence in dividing the genotypes into two groups. The first group was dominated by long cycles (6 months to mature), late maturing and relatively high yielding landraces. The second group was mainly composed of short cycle (4 months to mature), early maturing and low yielding landraces. The two distinct groups may have resulted from farmers' selection for adaptation for

specific production niches, cooking values and market orientation. The clustering of the genotypes based on phenotypic traits and SSR markers showed some discrepancies in terms of type and number of genotypes. SSR markers were better in clustering the genotypes according to races, whereas morphological markers were better in clustering genotypes according to their panicle related traits. The observed differences in the clustering patterns might be due to the fact that a combination of alleles can produce similar phenotypes. Additionally, none of the SSR markers used in this study are known to be linked to any of the morphological traits. A positive correlation ($r = 0.122$) was observed between genetic similarities and morphological similarities. The weak correlation may be due to the fact that morphological markers based on widely used descriptors provide different information than molecular markers. However, the positive correlation implies that morphological markers can be used as a useful tool in characterizing the genetic diversity of sorghum genotypes in areas where sophisticated genotyping laboratories are not available. Generally, the use of two or more different methods helps to better understand the genetic diversity, and the relationships within and among genotypes of crop species. In this study, the use of both SSR and morphological markers revealed the existence of a substantial genetic diversity in the lowland sorghum germplasm of Ethiopia.

Table 4. Summery statistics for the 30 polymorphic SSR loci screened across 200 sorghum genotypes.

Marker	Genetic parameter					
	N_a	N_e	A_r	H_e	F	PIC
gpsb067	8	2.74	4.62	0.54	0.86	0.60
gpsb123	6	2.69	3.53	0.58	0.92	0.56
Xcup02	4	2.20	2.69	0.52	0.90	0.44
Xcup14	3	1.57	2.78	0.35	0.95	0.33
Xcup53	4	2.01	2.96	0.43	0.84	0.44
Xcup61	2	1.97	2.00	0.45	0.90	0.37
Xgap84	16	6.71	5.13	0.80	0.77	0.83
Xgap72	11	3.41	7.43	0.67	0.85	0.66
Xgap206	19	8.87	7.72	0.83	0.88	0.88
Xtxp010	9	1.74	3.73	0.39	0.85	0.41
Xtxp012	14	5.18	6.59	0.76	0.82	0.78
Xtxp015	10	2.77	5.53	0.63	0.70	0.61
Xtxp021	12	2.26	5.20	0.49	0.85	0.54
Xtxp040	4	1.53	2.27	0.32	0.95	0.30
Xtxp057	13	4.76	5.84	0.74	0.88	0.76
Xtxp114	4	1.51	2.20	0.30	0.94	0.29
Xtxp141	15	5.76	6.94	0.77	0.86	0.81
Xtxp145	17	9.23	8.83	0.84	0.90	0.88
Xtxp265	21	5.77	8.61	0.78	0.11	0.81
Xtxp273	16	6.91	7.72	0.79	0.85	0.84
Xtxp278	9	1.91	3.26	0.43	0.81	0.44
Xtxp320	14	3.56	6.60	0.68	0.86	0.70
Xtxp321	15	3.31	6.09	0.63	0.84	0.68
SbAGB02	12	5.19	6.11	0.75	0.89	0.78
mSbCIR223	3	1.40	2.09	0.21	0.92	0.26
mSbCIR240	24	4.95	9.07	0.78	0.82	0.79
mSbCIR276	3	2.22	2.98	0.53	0.82	0.49
mSbCIR283	14	5.08	7.51	0.67	0.83	0.79
mSbCIR286	11	2.15	4.39	0.47	0.95	0.52
mSbCIR306	3	2.35	2.80	0.52	0.88	0.48
Overall mean	10.5	3.72	5.11	0.59	0.83	0.60
SD	6.03	2.20	2.25	0.18	0.04	0.20

N_a Total number of alleles per locus, N_e Numbers of effective alleles per locus, A_r Allelic richness, H_e Gene diversity within genotypes, F Fixation index, PIC Polymorphic information content, SD Standard deviation

Table 5. Genetic diversity estimates of lowland sorghum genotypes collected from three zones assessed by 30 SSR markers.

Genetic parameters	South Welo	North Welo	Wag Hemra	Overall
Total alleles (At)	275	254	202	316
Private alleles (Ap)	38	27	10	75
Rare alleles (Ar)	142	129	73	143
No. of alleles per locus (N_a)	9.17	8.47	6.73	10.53
No. effective alleles per locus (N_e)	3.72	3.61	3.53	3.72
Allelic richness (AR)	7.49	7.39	6.71	10.52
Gene flow (Nm)	14.18	11.58	6.17	10.64
Shannon information index (I)	1.450	1.402	1.329	1.461
Observed heterozygosity (H_o)	0.108	0.102	0.091	0.100
Gene diversity (H_e)	0.646	0.632	0.625	0.640
Fixation index (F)	0.843	0.845	0.865	0.849
Population differentiation (F_{ST})	0.017	0.021	0.039	0.030

Materials and Methods

Plant materials, study sites and experimental design

A total of 267 lowland sorghum landraces collected across a wide altitudinal range (900-1540 meter above sea level [masl]) from the northeastern parts of Ethiopia were used for this study (Supplementary table 1). The landraces were obtained mainly from the Ethiopian Institute of Biodiversity Conservation (IBC) along with their passport data. The landraces were subjected to two cycles of selection to maintain their genetic purity. The landraces were originally collected from three zones: Waghembra (37 entries), North

Welo (98) and South Welo (132). Phenotyping was conducted at two locations, Sirinka and Kobo under rainfed conditions. The locations represent the lowland agro-ecologies in Ethiopia. The landraces were planted in an alpha lattice design with two replications with inter- and intra- row spacing of 0.75m and 0.3m, respectively. Five entries were included as comparative controls viz., Girana-1, Misikir, and Raya, ICSV 111 and Jigury. All standard agronomic practices were followed as required. Observations were made on five plants of each genotype. Then 200 landraces were selected for genotyping on the basis of their morphological distinctiveness and drought tolerance. From each genotype, a bulk sample of leaf tips from five seedlings was harvested 4

weeks after germination. The leaf samples were dried using a silica gel protocol (Rogstad, 2003). The samples were arranged in 3 x 96-well plates and sent to DNA LandMarks Inc., Canada for genotyping. Genomic DNA was extracted from the dried leaf samples and the quality of the extracted DNA was evaluated on 1% agarose gel.

Phenotyping and data analysis

Thirteen qualitative characters, the standard sorghum descriptors for characterization, were used for phenotyping (IBPGR/ICRISAT, 1993) (Table 1). The phenotypic data were subjected to statistical analysis using GenStat (Payne et al. 2008) and SAS statistical software (SAS Institution, 2004). Shannon-Weaver index of diversity (H') was computed for each character from frequency distributions observed in the different classes (Perry and McIntosh, 1991) as follows: $H' = 1 - \sum_{i=1}^n p_i \log_e p_i$; where H' = Shannon

diversity Index; p_i = the proportion of accessions in the i^{th} class of an n -class character; n = the number of phenotypic classes of traits. Each diversity index value was divided by its maximum value ($\log_e n$) and normalized to keep the values between 0 and 1. The diversity index for each character was computed from the complete data set while the average diversity index was computed for each character for the collection zones. In addition, levels of diversity within localities and between localities were estimated in relation to the total variation. The data were also subjected to the Principal Component Analysis (PCA) and Discriminant Analysis (DA) procedures using SAS statistical software.

Genotyping and data analysis

For molecular diversity assessment, a set of 30 microsatellite markers were used. The markers used in this study were selected from SSR diversity kit (Billot et al., 2012) from all the linkage groups of sorghum. Genotyping was conducted at the genotyping service laboratory of DNA LandMarks Inc., Canada. For electrophoresis, an M13 forward primer method (5' CACGACGTTGTAAAA-CGAC3') was used at the 5' end of each primer. An M13-forward primer was labelled with one of four fluorescence dyes (6FAM, PET, NED or VIC) (Applied Biosystems, USA) for polymerase chain reaction (PCR) products detection using the ABI 3730xl Genetic Analyzer. Multiplex PCR reactions were performed in a total of 20 μ l reaction volume using a GeneAmp PCR system 9700 thermal cycler. PCR products were pooled together along with internal size standards (GeneScanTM, 500 LIZ[®] from Applied Biosystems). The data was generated through capillary electrophoresis and then analyzed using GeneMapper software. Product size was scored in base pairs based on the relative migration of the internal size standard. Information generated from the GeneMapper software was then used to determine the diversity parameters. Genotypic data were subjected to various measures of the genetic diversity of within and among genotypes using GenAlex version 6.5 (Peakall and Smouse, 2007). Genetic diversity parameters, such as the number of alleles per locus (N_a), the number of effective alleles per locus (N_e), allele richness (A_R) and average gene diversity (H_g) were determined. Furthermore, the fixation index (F) and polymorphic information content (PIC) were estimated for each locus and pre-determined group, based on the zone of collection. Allelic richness was corrected for sample size differences. Genotypic and phenotypic relationships within and among genotypes from

the three zones were assayed with a neighbor-joining algorithm using the unweighted pair group method (UWPGM) in DARwin 5.0 software (Perrier & Jacquemoud-Collet, 2006). A dendrogram was then generated on the dissimilarity matrix. Bootstrap analysis was performed for node construction using 10000 bootstrap values.

Conclusion

The present study showed that there is considerable genetic variability among sorghum landraces found in the drought-prone lowlands of Ethiopia. There was very high diversity for all the morphological traits studied. The proportion of total diversity obtained among the collection zones was less than that of within each zone. It would be possible to make selections for any of the traits within a particular region, assuming that a significant portion of the phenotypic variation is genetic. It was also found that the landraces showed a wide range of variation for stay-green and leaf rolling characters that would be useful in sorghum breeding for drought tolerance. From a conservation point of view, examining the way in which genetic diversity within a species is distributed among its populations and regions of origin always need to be considered. The detection of a significant number of private and rare alleles could be attributed to the high genetic diversity in the sorghum germplasm. This conforms that Ethiopia was one of the primary centers for the domestication of sorghum (Doggett, 1988). The diversity available in the studied germplasm would allow for future exploitation by sorghum improvement programs.

Acknowledgments

The Alliance for a Green Revolution in Africa, the Institute of Biodiversity Conservation of Ethiopia, Generation Challenge Program, DNA Landmark Laboratory and the University of KwaZulu-Natal are sincerely thanked for technical and financial support for the study.

References

- Abdi A, Bekele E, Asfaw Z, Teshome A (2002) Patterns of morphological variation of sorghum [*Sorghum bicolor* (L.) Moench] landraces in quantitative characters in North Shewa and South Welo, Ethiopia. *Hereditas*. 137: 161-172.
- Anas T, Yoshida T (2004) Genetic diversity among Japanese cultivated sorghum assessed with simple sequence repeat markers. *Plant Prod Sci*. 7: 217-223.
- Ayana A, Bekele E (2000) Geographic patterns of morphological variation in sorghum [*Sorghum bicolor* (L.) Moench] germplasm from Ethiopia and Eritrea: quantitative characters. *Euphytica*. 115: 91-104.
- Billot C, Rivallan R, Sall MN, Fonckea D, Deu M, Glaszmann JC, Noyer JL, Rami JF, Risterucci AM, Wincker P, Ramu P, Hash CT (2012) A reference microsatellite kit to assess for genetic diversity of *Sorghum bicolor* (Poaceae). *Amer J Bot*. 99: e245-250.
- Doggett H (1988) *Sorghum*. 2nd ed. Longman Scientific and Technical, London
- FAO (2012) Database of agricultural production Food and Agriculture Organization of the United Nation, FAOSTAT. <http://faostat.fao.org/default.aspx>. Accessed July 2014.
- Gamar YA, Kiambi D, Kairichi M, Kyallo M, Elgada MH (2013) Assessment of genetic diversity and structure of Sudanese sorghum accessions using simple sequence repeat (SSRs) markers. *Greener J Plant Breed Crop Sci*. 1: 16-24.

- Ganapathy KN, Gomashe SS, Rakshit S, Prabhakar B, Ambekar SS, Ghorade RB, Biradar BD, Saxena U, Patil JV (2012) Genetic diversity revealed utility of SSR markers in classifying parental lines and elite genotypes of sorghum (*Sorghum bicolor* L. Moench). *Aust J Crop Sci.* 11: 1486-1493.
- Geleta N, Labuschagne MT (2005) Quantitative trait variation in sorghum [*Sorghum bicolor* (L.) Moench] germplasm from eastern highlands of Ethiopia. *Biodiver Conserv.* 14: 3055-3064.
- Geleta N, Labuschagne MT, Viljoen CD (2006) Genetic diversity analysis in sorghum germplasm as estimated by AFLP, SSR and morpho-agronomical markers. *Biodiver Conserv.* 15: 3251-3265.
- Grenier C, Bramel PJ, Dahlberg JA, El-Ahmadi A, Mahmoud M, Peterson GC, Rosenow DT, Ejeta G (2004) Sorghum of the Sudan: analysis of regional diversity and distribution. *Genet Resour Crop Evol.* 51: 489-500.
- IBPGR/ICRISAT (1993) Descriptors for Sorghum [*Sorghum bicolor* (L.) Moench]. International Board of Plant Genetic Resources. Rome, Italy/International Crop Research Institute for Semi- Arid Tropics, Patancheru, India.
- Kimani PM, Wachira F, Cheruiyot EK, Owuochi J, Kimani E (2014) Genetic diversity of African sorghum (*Sorghum bicolor* L. Moench) accessions based on microsatellite markers. *Aust J Crop Sci.* 8: 171-177.
- Payne RW, Murray DA, Harding SA, Baird DB, Soutar DM (2008) *GenStat for Windows*, 11th Edition, VSN International, Hemel Hempstead.
- Peakall R, Smouse PE (2007) GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes.* 6: 288-295.
- Perrier X, Jacquemoud-Collet JP (2006) DARwin software: Dissimilarity analysis and representation for windows. Available from <http://www.darwin.cirad.fr/darwin.html>
- Perry MC, McIntosh MS (1991) Geographic patterns of variation in the USDA soybean germplasm collection: I. Morphological traits. *Crop Sci.* 31: 1350-1355.
- Rao PS, Deshpande S, Blummel M, Reddy BVS, Hash T (2012) Characterization of brown midrib mutants of sorghum (*Sorghum bicolor* (L.) Moench). *Euro J Plant Sci Biotech.* 6: 71-75.
- Rogstad SH (2003) Plant DNA extraction using silica. *Plant Mol Biol Rep.* 21: 463.
- SAS Institute Inc. (2004) *SAS/STAT Users Guide 9.1.3*. SAS Institute, Cary, NC.
- Smith JSC, Kresovich S, Hopkins MS, Mitchell SE, Dean RE, Woodman WL, Lee M, Poster K (2000) Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. *Crop Sci.* 40: 226-232.
- Spoooner D, van Treuren R, de Vicente MC (2005) Molecular markers for gene bank management. IPGRI Technical Bulletin No. 10, International Plant Genetic Resources Institute, Rome, Italy.
- Stemler ABL, Harlan JR, de Wet JMJ (1977) The sorghum of Ethiopia. *Econ Bot.* 31: 446-460
- Taketa S, Yuo T, Sakurai Y, Miyake S, Ichii M (2011) Molecular mapping of the short awn 2 (*lks2*) and dense spike 1 (*dsp1*) genes on barley chromosome 7H. *Breed Sci.* 61: 80-85.
- Teshome A, Patterson D, Asfaw Z, Torrance JK, Arnsan JT (2007) Changes of *Sorghum bicolor* landraces diversity and farmers' selection criteria over space and time, Ethiopia. *Genet Resour Crop Evol.* 54: 1219-1233.
- Teshome A, Fahrig L, Torrance JK, Lambert JD, Arnason JT, Baum BR (1999) Maintenance of sorghum (*Sorghum bicolor* Poaceae) landrace diversity by farmers' selection in Ethiopia. *Econ Bot.* 53: 79-88.
- Teshome A, Baum BR, Fahrig L, Torrance JK, Arnason TJ, Lambert JD (1997) Sorghum [*Sorghum bicolor* (L.) Moench] landrace variation and classification in North Shewa and South Welo, Ethiopia. *Euphytica.* 97: 255-263.
- Xu W, Subudhi PK, Crasta OR, Rosenow DT, Mullet JE, Nguyen HT (2000) Molecular mapping of QTLs conferring stay-green in sorghum (*Sorghum bicolor* L. Moench). *Genome.* 43: 461-469.