Australian Journal of Crop Science AJCS

AJCS 11(05):557-566 (2017) ISSN:1835-2707
doi: 10.21475/ajcs.17.11.05.p373

Screening, compiling and validation of informative microsatellite sets for marker-assisted breeding of key Ethiopian sorghum cultivars

Tesfaye Disasa[1,2*], Tileye Feyissa[2], Belayneh Admassu[1,3], Masresha Fetene[2], Santie M. Devillers[5], Venugopal Mendu[5]

[1]National Agricultural Biotechnology Research Center, P.O Box 2003, Addis Ababa, Ethiopia
[2]Faculty of Life Sciences, Addis Ababa University, P.O Box 1176, Addis Ababa, Ethiopia
[3]USDA-ARS, 1691 S2700W, ID 83210, USA
[4]Pwani University, PO Box 195-80108, Kilifi, Kenya
[5]Fiber and Biopolymer Research Institute (FBRI), Department of Plant & Soil Science, Texas Tech University, Lubbock, TX, USA

***Corresponding author: tesfayedisasa@yahoo.com**

Abstract

Sorghum is among the most important cereal crops produced globally due to its drought tolerance nature and multiple uses. Ethiopia is endowed with high genetic diversity for sorghum and the main sources of several genes that are responsible for biotic and abiotic stress tolerance. Despite this potential, sorghum productivity in the country remains very low. The study was designed to screen and compile the best simple sequence repeats (SSRs) that can be used for marker-assisted breeding of sorghum. Out of a total 304 SSRs markers screened used across eleven farmers preferred sorghum genotypes, nearly half of the markers 139 (45.7%) detected 543 alleles and a high degree of polymorphic information content (PIC) averaging 0.53. The overall observed heterozygosity (*Ho*) across all loci varied from 0.00 to 1.00 with an average of 0.16. Nearly 60 % (83 markers) showed no *Ho*. The gene diversity index (expected heterozygosity, *He*) ranged from 0.17 to 0.91 with a mean of 0.58. The weighted neighbor-joining cluster analysis grouped the genotypes into three distinct groups. All genotypes with stay-green features (B 35, Sorcoll 163/07, E 36-1 and Sorcoll 141/07) were clustered together. Genotypes such as Gambella, Macia, 76T#23 and Meko were clustered in the second group. The third group consists of Teshale and Sorcoll 146/07. Most of the used SSR markers were linked with stay-green traits and successfully discriminated genotypes with stay green (cluster I) from those genotypes with non-stay green features (cluster II and III). The result reveals the potential application of microsatellites in Ethiopian marker-assisted breeding program for further characterization and/or mapping of targeted traits in sorghum.

Keywords: Informative microsatellites; Marker-assisted breeding; Polymorphic information content; *Sorghum bicolor*.
Abbreviations: CTAB_Cetyltriethylammonium bromide; He_Expected heterozygosity; Ho_Observed heterozygosity, PIC_ Polymorphic information content; SBI_ *Sorghum bicolor* chromosome.

Introduction

The importance of sorghum (*Sorghum bicolor* (L.) Moench) (2n=20) across the world is well recognized, particularly in the lowland areas where rainfall is unreliable and crop failures due to recurrent drought occurrences are common. The crop plays a significant role for millions of food-insecure population and grows in a wide range of agro-ecological conditions. The depletion of water supplies, increased use of marginal farmlands, and global climatic trends suggest that the dry land crops such as sorghum will be of growing importance to feed the world's expanding population (Paterson, 2008). Since sorghum is a C4 grass, it has clear advantages over other grain crops because of its ability to return economic yields in hotter and drier environments (Bryden et al. 2009). Under favorable condition, sorghum has a high yield potential as compared to the other major cereals such as rice (*Oryza sativa*), wheat (*Triticum aestivum*) and maize (*Zea mays*) (Reddy et al., 2012). Besides, sorghum is a multipurpose crop with high biomass content with high scarification efficiency (Rai et al., 2016) which makes sorghum as a potential source to be used as animal feed as well as for the production of sustainable biofuel.

In Africa, sorghum is ranked the second most important cereal crop after maize (Borrell et al., 2001). In eastern African countries such as Sudan and Ethiopia, sorghum contributes about 40% of calories in the human diet (Kresovich et al., 2005). Sorghum is particularly a desirable crop for smallholder farmers in Ethiopia as it is relatively drought tolerant and will often yield a crop without irrigation. Besides, Ethiopia is the seventh producer of sorghum in the world and third in Africa contributing 12 % and 5 % of African and world annual production, respectively (Wani et al., 2011). The production and land coverage allocated for sorghum in Ethiopia accounted for about 18 % and 16 % of the total grain crop respectively (CSA, 2014).

Sorghum is believed to have been domesticated in Ethiopia and surrounding countries commencing around 4000-3000 BC (Dillon et al., 2007). There is evidence that the crop was first domesticated in a savanna between Chad and western Ethiopia (Doggett, 1988). This theory is further supported by the presence of a diverse number of wild sorghum relatives in the eastern African region, especially in Sudan and Ethiopia (Gebrekidan, 1982). Wild relatives are particularly good

sources of novel genes that can be used in crop improvement. Besides, the discovery of known stay-green genotypes (*E 36-1, B 35)* possessing novel genes for drought tolerance traits (Haussmann et al., 2002; Harris et al., 2007; Reddy et al., 2009) is a good demonstration of the importance of Ethiopian sorghum accessions in crop improvement. Besides, there is high genetic diversity among Ethiopian sweet sorghum accessions (Disasa et al., 2016a) which are characterized by high sugar content (Disasa et al., 2016b). This is a potential area to be used in bioethanol production as well as animal feed.

Despite the existence of a large genetic resource for both cultivated and wild sorghum in Ethiopia, there has been very little utilization of the resource in crop improvement. This is due to the reliance of the country on conventional based crop improvement program since the beginning of crop domestication. Hence, utilization of molecular approaches along with conventional breeding is an important step towards the improvement and full understanding of the contribution of Ethiopian sorghum germplasm to global breeding programs. A recent whole-genome sequencing study revealed untapped potential in Africa's indigenous sorghum (Mace et al., 2013), which global sorghum collections could greatly benefit from.

The application of informative microsatellite markers in sorghum improvement program plays an important role in facilitating the breeding program. Large numbers of SSR primers have been assembled in the public domain that can be used for various purposes such as diversity analysis, construction of linkage and QTL mapping, marker assisted selection and other related activities (Brown et al., 1996; Bhattramakki et al., 2000; Kong et al., 2000; Schloss et al., 2002; Menz et al., 2002; Wu et al., 2006; Li et al., 2009; Ramu et al., 2009). However, most of them developed using exotic germplasm with very few number of sorghum genotypes of Ethiopian origin (Beyene et al., 2014; Disasa et al., 2016a). Therefore, there is a need to screen and compile the most polymorphic markers that can be used for marker assisted selection in the future using selected Ethiopian sorghum genotypes.

Results

Optimization of DNA isolation protocol

Isolation of good quality and high concentration of DNA is one important aspect in the study of molecular biology and related disciplines. The effectiveness of each protocol varies from plant species to species. Some crop plants will yield good DNA with a commonly used protocol whereas others require a sophisticated protocol to isolate genomic DNA. In this experiment, genomic DNA was isolated using CTAB and Kit methods and the result was compared for quality as well as quantity. Both methods have yielded high concentration of DNA (Supplementary Table 1). The concentration of genomic DNA obtained using CTAB method is reasonably high ranging from 123.1 to 433.3 ng/µl. The concentration obtained using Kit method is nearly double as compared to CTAB method of DNA isolation with the lowest and highest score of 434.7 ng/µl for *76T#23* and 758.6 ng/µl for *Sorcoll 163/07* genotypes (Supplementary Table 1). However, the concentration of isolated genomic DNA through CTAB method is still very sufficient and economical to be used for genotyping using large number of markers. Similarly, the quality of the PCR products obtained using both methods was higher and remains very similar with no significant difference (Supplementary Fig 2).

SSR polymorphism

In the current study, attempt has been made to screen and compile the most polymorphic markers using selected Ethiopian sorghum genotypes to be used in marker assisted breeding in the future. Fragment analysis result showed that among 304 SSR primer pairs tested, a total of 142 were found to be polymorphic markers. These markers are evenly distributed across all linkage groups (Table 1). However, three markers were excluded from the final analysis due to their poor PCR product quality. A total of 139 (45.7%) showed clear band and found to be polymorphic across the selected eleven sorghum genotypes. On the contrary, 97 (31.9%) were unable to discriminate the genotypes and considered as monomorphic markers for the selected genotypes and the remaining 68 (22.4%) didn't work at all. Most of the polymorphic and monomorphic markers were clearly separated during fragment scoring (Fig 1) and those monomorphic markers were also excluded from the final analysis.

One hundred thirty-nine informative SSR markers produced a total of 543 alleles across the eleven key Ethiopian genotypes with an average of 3.91 alleles per marker. The observed number of alleles ranged from two to ten per primer pair. *Xtxp008* produced the highest number of alleles (10) followed by *Xtxp032* (9) and *Xtxp033* (8). These markers also presented higher values of gene diversity (*He* = 0.9).

Majority of the markers (56%) have a polymorphism information content (PIC) value of greater than 0.5 and considered as very informative. Similarly, 58 out of 139 markers (42%) showed PIC value between 0.25 and 0.5 which is categorized as informative markers. Only three markers *Xisep0805*, *Xisep0444* and *Xisep0815* grouped as less informative with PIC value below 0.25. The highest PIC value of the scored SSR loci was obtained from *Xtxp343* which was 0.91 whereas the lowest PIC value (0.15) was obtained from *Xisep0815* with an average value of 0.53 (Table 2).

Nearly 60 % (83 markers) have showed zero value of the observed heterozygosity (*Ho)*. Only four markers *Xisep0938, Xisep0704, Xtxp159* and *Xtxp283* had *Ho* value of 1.0. In contrast, the gene diversity index (expected heterozygosity, *He*) was ranging from 0.17 for *Xisep0805* to 0.91 for *Xtxp343* with a mean value of 0.58. The lowest gene diversity index and PIC value was scored from *Xisep0805*. Among the tested SSR markers, primers named as *Xtxp* showed higher PIC and hence it is considered as highly informative markers (Table 2).

Genetic relationship among the genotypes

Molecular based knowledge of the genetic relationships among these selected sorghum genotypes is very essential to intensively utilize further in breeding program. The 139 polymorphic markers discriminated the selected Ethiopian sorghum genotypes with stay-green features in one group including the worldwide known stay-green sources of sorghum, *B 35* and *E 36-1* (Figs 2 and 3). Neighbor-joining cluster analysis grouped the genotypes into three distinct groups. The first group constituted genotypes with stay-green feature, *B 35, Sorcoll 163/07, E 36-1and Sorcoll 141/07*. The second group composed of farmers preferred Ethiopian genotypes *Gambella, Macia, 76T#23, Meko* and *Melkam*. The third group includes *Teshale* and *Sorcoll 146/07*. Both the second and third groups are characterized by non-stay greenness.

Table 1. Distribution of the polymorphic markers across the ten linkage groups.

No	Chromosomes (Linkage groups)										
	A (SBI-01)	B (SBI-02)	C (SBI-03)	D (SBI-04)	E (SBI-07)	F (SBI-09)	G (SBI-10)	H (SBI-08)	I (SBI-06)	J (SBI-05)	Not well known
1	*Xcup33*	*Xtxp008*	*Xtxp033*	*Xtxp012*	*Xtxp159*	*Xtxp0289*	*Xtxp0141*	*Xisep108*	*Xisep0429*	*Xisep1029*	*Xisep0643*
2	*xgap342*	*Xgap084*	*Xtxp285*	*Xtxp041*	*Xisep704*	*Xtxp0296*	*Xgap001*	*Xisep809*	*Xisep0444*	*Xisep1127*	*Xisep0648*
3	*Xtxp032*	*Xisep310*	*Xisep0114*	*Xtxp021*	*Xisep0716*	*Xgap032*	*Xisep314*	*Xisep0815*	*Xisep0502*	*Xtxp014*	*Xisep0830*
4	*Xtxp357*	*Xisep0612*	*Xisep0117*	*Xtxp024*	*Xisep0829*	*Xisep110*	*Xisep0621*	*Xisep1150*	*Xisep346*	*sbKAKG1*	*Xisep0901*
5	*Xisep327*	*Xisep0938*	*Xisep0132*	*Xisep202*	*Xisep831*	*Xisep0517*	*Xisep0622*	*Xisep1225*	*Xisep0422*	*Xisep1208*	*Xisep0905*
6	*Xisep0839*	*Xisep1013*	*Xisep138*	*Xisep224*	*Xisep0328*	*Xisep537*	*Xisep0624*	*Xtxp205*	*Xisep423*	*Xtxp303*	*Xisep1717*
7	*Xisep1032*	*Xcup26*	*Xcup24*	*Xisep0234*	*Xisep0805*	*Xisep0539*	*Xisep0639*	*Xisep1231*	*Xisep0443*	*Xisep1107*	*Xtxp113*
8	*Xisep1039*	*Xcup40*	*Xcup61*	*Xisep0242*	*Xtxp312*	*Xisep0543*	*Xcup50*	*Xtxp0210*	*Xisep0449*	*Xisep1140*	*Xtxp160*
9	*Xisep1046*	*Xtxp211*	*Xisep0101*	*Xcup048*	*SBGEF06*	*Xcup002*	*Xisep0630*		*Xisep0617*	*Xtxp091*	*Xtxt267*
10	*Xtxp080*	*Xtxp315*	*Xtxp205*	*Xtxp328*		*Xisep511*	*Xtxp270*		*Xcup12*	*Xtxp023*	*Xtxp355*
11	*Xtxp329*	*Xtxp348*	*SABGEF08*	*Xisep1103*		*Xisep0506*	*SBGE01*		*Xisep0435*	*Xtxp283*	*Xtxp307*
12	*Xcup53*	*SABAGA04*	*Xisep1042*	*Xtxp343*		*Xisep0550*	*Xcup49*		*Xtxp274*		*Xtxp361*
13	*Xcup22*	*SBAGAB03*	*Xtxp031*	*Xtxp026*		*Xtxp324*					
14	*Xtxp279*	*Xcup36*	*Xtxp034*			*Xtxp067*					
15	*Xtxp284*	*Xisep1145*				*Xtxp258*					
16	*Xtxp335*	*Xtxp072*				*Xtxp324*					
17		*Xtxp286*				*Xtxp358*					
18		*Xtxp298*									
Total	**16**	**18**	**14**	**13**	**9**	**17**	**12**	**8**	**12**	**11**	**12**

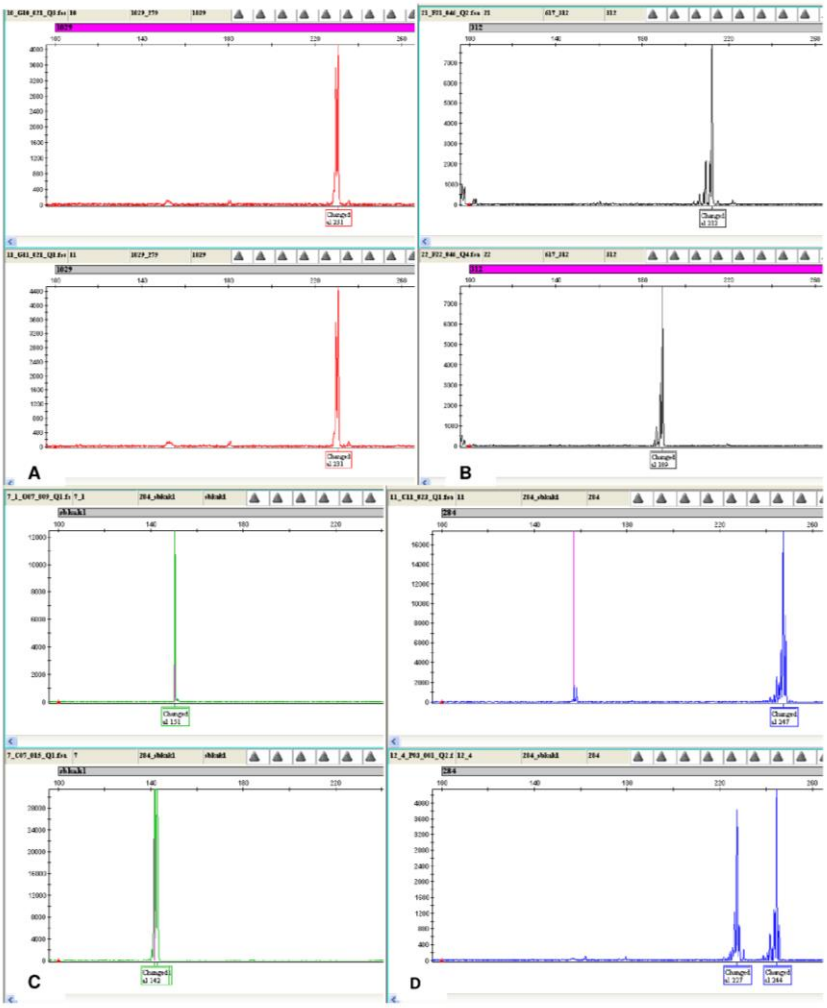

Fig 1. The fragment sizes were manually scored using GeneMapper 4.0 software (Applied Biosystems). Scored bins showing (A) a monomorphic marker, uniform across the genotypes. (B, C and D) polymorphic markers. The two bins on the same locus (D) indicates heterozygosity while the remaining loci (A, B and C) are considered as homozygous type. Each marker was labeled M-13 tagged forward primer labels (A = PET, B = NED, C = VIC and D = FAM).

The highest genetic distance (0.86) was observed between genotypes *Sorcoll 141/07* and *76 T1#23* followed by *Sorcoll 141/07* and *Meko* (0.79). On the contrary, the lowest genetic distance was found between *Gambella* and *Melkam* (0.20). Most of them appeared to be distantly related and they have genetic distance of greater than 0.50 (Table 3).

Discussion

The ratio of 260 to 280 using the CTAB method (Mace et al., 2003) was between 1.83 and 2.01 which is within the expected range of good quality DNA. Similarly, high quality genomic DNA with a 260 to 280 ratio ranging from 1.78 to 1.89 was also isolated using Kit method (Supplementary Table 1). This suggested that the isolated genomic DNA was free from any impurities. Sorghum is among those crops that yielded a good DNA with commonly used protocols like CTAB. Our result showed that there was no significant difference between CTAB and KIT based methods for the quality of both genomic DNA and PCR product suggesting that CTAB method would be an ideal method to isolate sorghum genomic DNA for SSR analysis. Therefore, the current optimized CTAB protocol can also be used for genotyping by sequencing (GBS) purposes which requires relatively good quality DNA with fair concentration.

The majority of the markers (40%) had fragment size between 201 to 250 bp. Similarly, second highest frequency (29%) was observed for allele size ranged from 151 to 200 bps followed by allele sizes between 100 to 150 bps between ranging (Table 1). In general, the majority of the screened polymorphic markers (94%) allele sizes ranged from 100 to 300 bp which is in agreement with many studies undertaken in the past (Bhattramakki et al., 2000; Kong et al., 2000; Menz et al., 2002; Scholes et al., 2002; Wu et al., 2006; Li et al., 2009; Ramu et al., 2009 Disasa et al., 2016a). Besides, most of the markers that their fragment sizes lie in this range showed clear and longer beaks while scoring using GenMapper software (Fig 1) indicated that they should be the choice of marker for the application in sorghum improvement programs. Previous studies showed that SSR markers have been extensively used to detect the variability in grain sorghum and to evaluate their genetic diversity (Ali et al., 2008; Muraya et al., 2011; Adugna et al., 2012; Billot et al., 2013). The mean number of alleles per locus (3.91) detected in this study was lower than the average number of

Table 2. Diversity statistics of key Ethiopian sorghum genotypes computed with 139 polymorphic SSR loci.

Marker	MAF	NA	Ho	He	PIC
Xcup033	0.55	4.00	0.59	0.09	0.52
Xgap342	0.36	5.00	0.74	0.00	0.70
Xtxp008	0.20	10.00	0.88	0.20	0.87
Xtxp012	0.45	4.00	0.64	0.73	0.58
Xtxp041	0.50	4.00	0.65	0.36	0.59
Xtxp021	0.45	3.00	0.63	0.00	0.55
Xtxp024	0.41	7.00	0.77	0.18	0.74
Xtxp032	0.20	9.00	0.87	0.80	0.85
Xtxp033	0.14	8.00	0.87	0.43	0.85
Xtxp0141	0.45	4.00	0.64	0.00	0.58
Xtxp159	0.50	4.00	0.61	1.00	0.53
Xtxp285	0.73	4.00	0.45	0.00	0.42
Xtxp0289	0.50	4.00	0.66	0.00	0.60
Xtxp0296	0.36	4.00	0.69	0.00	0.64
Xtxp357	0.45	3.00	0.63	0.00	0.55
Xgap001	0.30	6.00	0.79	0.10	0.75
Xgap032	0.25	7.00	0.84	0.00	0.82
Xgap084	0.71	3.00	0.45	0.00	0.41
Xisep0110	0.73	2.00	0.40	0.00	0.32
Xisep0114	0.82	3.00	0.31	0.00	0.29
Xisep0117	0.82	3.00	0.31	0.00	0.29
Xisep0132	0.38	4.00	0.71	0.13	0.66
Xisep0138	0.40	4.00	0.66	0.20	0.59
Xisep0202	0.40	5.00	0.67	0.30	0.60
Xisep0224	0.60	3.00	0.54	0.00	0.47
Xisep0234	0.50	4.00	0.64	0.00	0.58
Xisep0242	0.38	4.00	0.69	0.00	0.63
Xisep0310	0.39	5.00	0.75	0.33	0.71
Xisep0314	0.25	6.00	0.81	0.00	0.79
Xisep0327	0.65	3.00	0.49	0.10	0.41
Xisep0429	0.50	3.00	0.55	0.11	0.45
Xisep0444	0.89	2.00	0.20	0.00	0.18
Xisep0502	0.50	3.00	0.61	0.11	0.54
Xisep0517	0.38	3.00	0.66	0.00	0.58
Xisep0537	0.80	3.00	0.34	0.00	0.31
Xisep0539	0.43	3.00	0.61	0.00	0.53
Xisep0543	0.60	3.00	0.56	0.00	0.50
Xisep0621	0.40	4.00	0.72	0.20	0.67
Xisep0101	0.30	5.00	0.78	0.00	0.74
Xisep0622	0.50	4.00	0.66	0.00	0.60
Xisep0624	0.56	3.00	0.54	0.78	0.44
Xisep0643	0.50	5.00	0.67	0.75	0.63
Xisep0648	0.75	3.00	0.40	0.33	0.36
Xisep0704	0.50	4.00	0.60	1.00	0.53
Xisep0716	0.50	3.00	0.60	0.91	0.52
Xisep0809	0.63	2.00	0.47	0.00	0.36
Xisep0815	0.88	2.00	0.22	0.00	0.19
Xisep0829	0.50	3.00	0.58	0.00	0.49
Xisep0830	0.60	2.00	0.48	0.00	0.36
Xisep0831	0.60	3.00	0.54	0.00	0.47
Xisep0839	0.50	2.00	0.50	0.00	0.38
Xisep0901	0.63	2.00	0.47	0.00	0.36
Xisep0905	0.64	3.00	0.52	0.14	0.46
Xisep0938	0.40	4.00	0.69	1.00	0.63
Xisep1013	0.50	2.00	0.50	0.00	0.38
Xisep1029	0.56	4.00	0.62	0.00	0.57
Xisep1032	0.73	4.00	0.44	0.09	0.41
Xisep1039	0.60	2.00	0.48	0.00	0.36
Xisep1046	0.80	3.00	0.34	0.10	0.30
Xisep1127	0.82	2.00	0.30	0.00	0.25
Xisep1150	0.33	4.00	0.72	0.00	0.67
Xisep1717	0.50	3.00	0.62	0.78	0.55
Xtxp080	0.43	5.00	0.73	0.00	0.70
Xtxp329	0.44	4.00	0.69	0.00	0.64
Xcup002	0.64	3.00	0.51	0.00	0.44
Xcup024	0.50	3.00	0.62	0.00	0.55
Xcup048	0.57	4.00	0.61	0.00	0.57
Xcup053	0.60	5.00	0.60	0.00	0.57
Xcup061	0.59	2.00	0.48	0.09	0.37
Xisep0108	0.60	4.00	0.58	0.00	0.54
Xisep0328	0.70	3.00	0.46	0.00	0.41
Xisep0346	0.73	3.00	0.43	0.00	0.39
Xisep0422	0.45	3.00	0.63	0.00	0.55
Xisep0423	0.45	4.00	0.63	0.30	0.55
Xisep0443	0.45	5.00	0.71	0.00	0.67
Xisep0612	0.38	4.00	0.72	0.00	0.67

Table 2. continued.

Marker	MAF	NA	Ho	He	PIC
Xisep0449	0.55	2.00	0.50	0.00	0.37
Xisep0511	0.41	7.00	0.76	0.36	0.73
Xisep0639	0.27	7.00	0.83	0.00	0.80
Xtxp014	0.30	6.00	0.80	0.00	0.77
sbKAKG01	0.55	4.00	0.63	0.00	0.58
Xcup22	0.73	2.00	0.40	0.00	0.32
Xcup26	0.82	3.00	0.31	0.00	0.29
Xcup40	0.55	3.00	0.56	0.00	0.48
Xcup50	0.45	3.00	0.64	0.09	0.57
Xisep0506	0.64	3.00	0.51	0.00	0.44
Xisep0550	0.73	4.00	0.45	0.00	0.42
Xisep0617	0.70	2.00	0.42	0.00	0.33
Xisep0630	0.56	3.00	0.59	0.00	0.53
Xisep0805	0.91	2.00	0.17	0.00	0.15
Xisep1208	0.50	3.00	0.59	0.00	0.51
Xtxp211	0.27	7.00	0.80	0.55	0.77
Xtxp270	0.55	5.00	0.64	0.00	0.61
Xtxp279	0.27	6.00	0.81	0.00	0.78
Xtxp284	0.36	6.00	0.78	0.00	0.75
Xtxp303	0.68	3.00	0.48	0.27	0.42
Xtxp312	0.27	6.00	0.81	0.00	0.78
Xtxp315	0.45	3.00	0.58	0.00	0.49
Xtxp328	0.75	3.00	0.41	0.00	0.37
Xtxp348	0.70	2.00	0.42	0.00	0.33
SABAGAH04	0.27	7.00	0.79	0.55	0.76
SABGEF08	0.64	2.00	0.46	0.18	0.36
SBAGAB03	0.60	5.00	0.57	0.40	0.52
SBGE01	0.72	5.00	0.46	0.22	0.43
SBGEF06	0.61	2.00	0.48	0.56	0.36
Xcup12	0.64	4.00	0.55	0.00	0.50
Xcup36	0.32	5.00	0.78	0.36	0.74
Xcup49	0.77	3.00	0.38	0.09	0.34
Xisep1042	0.68	3.00	0.48	0.09	0.43
Xisep1103	0.82	2.00	0.30	0.00	0.25
Xisep1107	0.73	3.00	0.43	0.00	0.39
Xisep1140	0.45	4.00	0.69	0.70	0.63
Xisep1145	0.85	3.00	0.27	0.10	0.25
Xisep1225	0.68	5.00	0.50	0.18	0.46
Xtxp343	0.18	5.00	0.91	0.73	0.91
Xtxp026	0.27	7.00	0.83	0.00	0.80
Xtxp067	0.64	2.00	0.46	0.00	0.36
Xtxp072	0.33	7.00	0.80	0.78	0.77
Xtxp091	0.45	5.00	0.71	0.60	0.67
Xtxp113	0.45	5.00	0.71	0.82	0.67
Xtxp160	0.68	5.00	0.51	0.36	0.48
Xtxp205	0.70	2.00	0.42	0.20	0.33
XtXt267	0.56	5.00	0.64	0.00	0.61
Xtxp335	0.56	4.00	0.57	0.11	0.50
Xisep1231	0.60	2.00	0.48	0.00	0.36
Xisep0435	0.77	2.00	0.35	0.45	0.29
Xtxp021	0.55	3.00	0.60	0.00	0.53
Xtxp023	0.56	3.00	0.59	0.00	0.53
Xtxp031	0.67	3.00	0.50	0.00	0.45
Xtxp034	0.64	5.00	0.56	0.00	0.53
Xtxp258	0.27	5.00	0.76	0.00	0.72
Xtxp274	0.45	6.00	0.73	0.00	0.70
Xtxp283	0.50	3.00	0.60	1.00	0.52
Xtxp286	0.73	3.00	0.43	0.00	0.39
Xtxp298	0.33	5.00	0.77	0.00	0.73
Xtxp307	0.65	5.00	0.55	0.10	0.52
Xtxp324	0.59	3.00	0.57	0.45	0.50
Xtxp358	0.56	5.00	0.63	0.13	0.60
Xtxp361	0.36	6.00	0.76	0.00	0.73
Total	**74.49**	**543**	**81.24**	**22.07**	**73.49**
Mean	**0.54**	**3.91**	**0.58**	**0.16**	**0.53**

MAF –Major allele frequency; NA –Allele number; Ho – observed heterozygosity; He –expected heterozygosity; PIC –Polymorphic information content

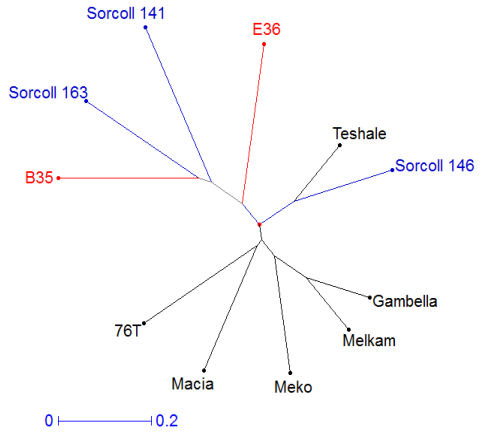

Fig 2. Neighbor-joining cluster analysis dendrogram showing the genetic relationship among the farmers preferred sorghum genotypes based on 139 SSR markers. Red colors indicate globally known genotypes for their stay-green features while blue colors stands for Ethiopian genotypes with potential stay green characteristics.

alleles reported by Wang et al. (2009) but slightly higher than the result documented by Ali et al. (2008) who documented 3.2 on an average. The result was also lower than mean value obtained from genotyping of large number of sorghum accessions collected around the world (Billot et al., 2013). This variation is most probably due to the diverse and massive accession used for genotyping and the polymorphic nature of the selected SSR markers. Most of the markers used in the present study were highly informative as well as highly polymorphic. The polymorphic information content values of markers play an important role in estimating the discrimination power in a set of accession based on the number of alleles as well as the frequencies of each allele (Smith et al., 2000). The computed average PIC value (0.53) in this experiment was very close to most of the previously reported values using both grain and sweet sorghum (Agrama and Tuinstra,2003; Caniato et al., 2007; Ali et al., 2008; Wang et al., 2009; Ramu et al., 2013). However, the result was somewhat lower than the average PIC value reported in wild sorghum (Muraya et al., 2011; Adugna et al., 2012) and sweet sorghum (Disasa et al., 2016a) populations. This is possibly due the polymorphic nature of the selected SSR markers used for genotyping as well as type of germplasm studied. Since the selected markers are highly polymorphic, they will serve as background markers for further molecular characterization and mapping of Ethiopian sorghum germplasm lines in the future. The overall mean gene diversity (0.58) in this study is slightly lower than previously reported studies (Deu et al., 2008; Adugna et al., 2012; Billot et al., 2013). However, this shouldn't be used for comparison purposes because estimates of such type of genetic parameters depends on various factors such as the type of marker used (Barakat et al., 2011), the size of the SSR repeats and the location of the SSR on the genome (between coding or non-coding DNA regions), the sampling schemes (single plant or DNA bulk) and the number of surveyed SSR (Deu et al., 2008). The principle coordinate analysis based on the dissimilarity of 139 SSR markers and cluster analysis clearly showed that *Sorcoll 163/07* clustered tightly with *B 35*. Similarly, Neighbor-joining cluster analysis grouped the accessions (*Sorcoll 163/07* and *Sorcoll 141/07)* in between the two stay-green genotypes (*E 36-1* and *B 35*). This suggests that the accessions have similar genome composition to that of the stay-green materials and may be originated from the same environment. This information helps to validate previous reports which are based on morphological screening of genotypes for the trait of interest like drought tolerance in sorghum. Mengistie (2009) also reported that three accessions (*Sorcoll 163/07, Sorcoll 141/07* and *Sorcoll 146/07*) collected from different parts of Ethiopia showed stay-green features after evaluating morpho-physiologically in different water deficit environment.

Materials and Methods

Sorghum germplasm selection

Eleven Ethiopian farmers preferred sorghum genotypes were used to screen and compile informative microsatellite sets that can be used for marker-assisted breeding of Ethiopian sorghum cultivars. These genotypes were *Teshale, Gambella, Meko, Melkam, Macia, 76T#1, B 35, E 36-1, Sorcoll 141/07, Sorcoll 146/07* and *Sorcoll 163/07*. They were selected for their essential traits such as: early maturity, seed quality, high yield, high sugar content and drought tolerance.

Preparation of plant samples and DNA extraction

Seeds of eleven selected sorghum genotypes were sown in a greenhouse at the National Agricultural Biotechnology Research Center (NABRC) of the Ethiopian Institute of Agricultural Research, Ethiopia. Leaf tissues were collected from two to three-week-old seedlings followed by genomic DNA extraction using Promega Kit (Madison, USA) and CTAB method (Mace et al., 2003) in order to compare the two protocols for further sorghum genotyping. Quantity and quality of the DNA was checked using *Qubit®2.0*(*Life Technologies***,** Grand Island, NY) and by running on 0.8% agarose gel stained with GelRed® (Biotium, USA), respectively.

Polymerase chain reaction and fragment analysis

SSR genotyping and data analysis was done at ICRISAT-Nairobi in Kenya. A total of 304 polymorphic SSR primers that were evenly distributed across the whole sorghum nuclear genome were selected for use in genotyping the eleven genotypes. The markers were selected from previous reports (Brown et al., 1996; Bhattramakki et al., 2000; Kong et al., 2000; Menz et al., 2002; Wu et al., 2006; Li et al., 2009; Ramu et al., 2009) and were obtained from ICRISAT-India. All forward primers contained an M13-tag (5'-CACGACGTTGTAAAACGAC - 3') on the 5' end that was

Table 3. Pair wise population Nei's genetic distance showing the magnitude of genetic differentiation among key Ethiopian sorghum genotypes.

Genotypes	*Sorcoll 141/07*	*Sorcoll 146/07*	*Sorcoll 163/07*	*76T1#23*	*B-35*	*E 36*	*Gambella*	*Macia*	*Meko*	*Melkam*	*Teshale*
Sorcoll 141/07	0.000										
Sorcoll 146/07	0.667	0.000									
Sorcoll 163/07	0.619	0.676	0.000								
76T1#23	0.855	0.565	0.731	0.000							
B-35	0.594	0.730	0.495	0.743	0.000						
E 36-1	0.690	0.620	0.757	0.648	0.716	0.000					
Gambella	0.779	0.505	0.728	0.509	0.743	0.643	0.000				
Macia	0.754	0.599	0.750	0.487	0.746	0.648	0.492	0.000			
Meko	0.786	0.514	0.654	0.464	0.728	0.662	0.380	0.578	0.000		
Melkam	0.763	0.518	0.756	0.517	0.725	0.623	0.204	0.512	0.395	0.000	
Teshale	0.681	0.227	0.638	0.565	0.667	0.581	0.462	0.496	0.496	0.452	0.000

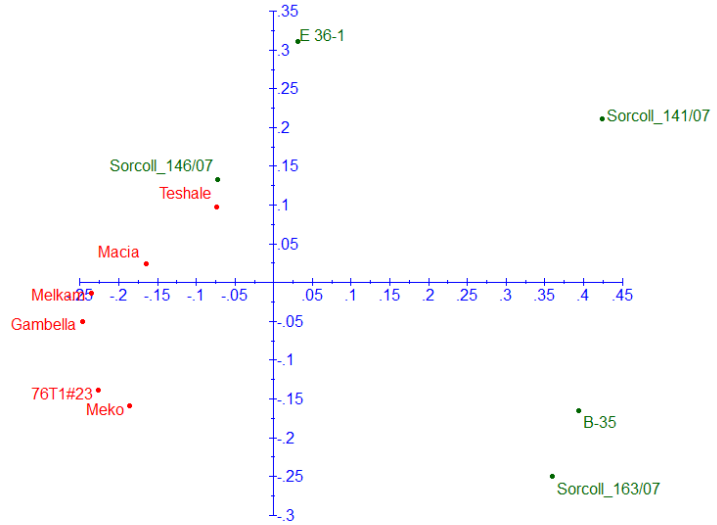

Fig 3. Biplot of the axis 1 and 2 of the principle coordinate analysis based on the dissimilarity of 139 SSR markers for key Ethiopian sorghum genotypes. Green colors indicate genotypes with potential stay green gene source where as red colors stand for non-stay green gene source.

fluorescently labeled to allow detection of amplification products (Schuelke, 2000). PCR amplification was performed in 10 µl reaction volume comprising of 1 x PCR buffer (20 mM Tris-HCl, pH 7.6; 100 mMKCl; 0.1 mMEDTA; 1 mMDTT; 0.5% (w/v) Triton X-100; 50% (v/v) glycerol), 2 mM$MgCl_2$, 0.16 mM dNTPs, 0.16 µM fluorescent labeled M13-forward primer, 0.04 µM forward primer, 0.2 µM reverse primer, 0.2 units of Taq DNA polymerase (SibEnzyme Ltd, Russia) and 30 ng of template DNA. Forward primers were labeled with FAM, PET, NED or VIC (Applied Biosystems, USA). PCR was carried out in a GeneAmp® PCR System 9700 thermal cycler (Applied Biosystems) programed for initial denaturation at 94°C for 15 min, followed by second denaturation at 94°C for 30 sec, annealing at 50°C for 1 min, extension at 72°C for 2 min and final elongation at 72°C for 20 min.

Successful amplification was confirmed by running 2.0 µl of the PCR products on a 2% (w/v) agarose gel stained with GelRed® (Biotium) and visualized under UV. Depending on the nature of fluorescent label and strength of the amplification bands used, a volume ranging from 2.5 µl to 3.5 µl of four different amplification products were co-loaded along with the internal size standard, GeneScan™ –500 LIZ® (Applied Biosystems) and Hi-Di™ Formamide (Applied Biosystems, USA). The fragments were separated by capillary electrophoresis using an ABI Prism® 3730 Genetic analyzer (Applied Biosystems). PCR fragment sizes were manually scored using GeneMapper 4.0 software (Applied Biosystems). Power Marker v.3.25 (Liu and Muse 2005) was used to compute PIC, heterozygosity (*Ho*) and gene diversity (expected heterozygosity, *He*) for each marker as well as the average across markers. Polymorphism information content (PIC) was calculated using the method of Botstein et al. (1980).

$$PIC = 1 - \sum_{i=0}^{k} p_i^2 - \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} 2 p_i^2 p_j^2$$

Where, p_i and p_j are the frequencies of alleles *i* and *j*, respectively

Phylogenetic analysis

To identify the pair-wise genetic relationships among accessions, a genetic dissimilarity matrix was analyzed using Neighbor Joining (NJ) method, as implemented in DARwinv5 (Perrier and Jacquemoud-Collet, 2006). The dendrogram was also constructed using the same software.

Conclusion

Detection of a large number of SSR markers that are polymorphic among Ethiopian key sorghum genotypes plays a significant role in a marker-assisted breeding program of the crop. It helps to discern contrasting parents to be used as mapping population on the DNA level and hence improves the efficiency of sorghum improvement program. Validation of morph-physiological characterized stay-green genotypes collected from different regions has been successfully achieved in this experiment. It was very important to compare these genotypes with globally known stay-green sorghum genotypes using large numbers of molecular markers. The selected markers are capable of categorizing Ethiopian key sorghum genotypes according to their adaptation to various biotic and abiotic stresses. The selected microsatellites sets would be useful resources for marker-assisted breeding program of the country as well as the region.

Acknowledgements

The study was conducted by financial support provided by the Swedish International Development Agency (SIDA) through Bio-Innovate project "Delivering new sorghum and millets innovations for food security and improving livelihoods in Eastern Africa"- project No. 01/2010. The authors acknowledge the Lab facilities support provided by the International Crop Research for Semi-Arid Tropics (ICRISAT-Nairobi).

References

Adugna A, Snow AA, Sweeney PM, Bekele E, Mutegi E (2012) Population genetic structure of *in situ* wild *Sorghum bicolor* in its Ethiopian center of origin based on SSR markers. Genet Resour Crop Ev. 60:1313-1328.

Agrama HA, Tuinstra MR (2003) Phylogenetic diversity and relationships among sorghum accessions using SSRS and RAPDs. Afr J of Biotechnol. 2:334-340.

Ali ML, RajewskiJF, Baenziger PS, Gill KS, Eskridge KM, Dweikat I (2008) Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. Mol Breeding. 21:497-509.

Baraket G, Chatti K, Saddoud O, Abdelkarim AB, Mars M, Trifi M, Hannachi AS (2011) Comparative assessment of SSR and AFLP markers for evaluation of genetic diversity and conservation of fig, Ficuscarica L., genetic resources in Tunisia. Plant Mol Biol Rep. 29:171-184.

Beyene A, Hussien S, Pangirayi T, Mark L, Fentahun M (2014) Genetic variation in lowland sorghum (*Sorghum bicolor* (L.) Moench) landraces assessed by simple sequence repeats. Plant Genet Resour. 14:1-11.

Bhattramakki D, Dong J, Chhabra AK, Hart GE (2000) An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. Genome. 43:988-1002.

Billot C, Ramu P, Bouchet S, Chantereau J, Deu M, Gardes L, NoyerJL, Rami JF, Rivallan R, Li Y, Lu P, Wang T, Folkertsma RT, Arnaud E, Upadhyaya HD, Glaszmann JC, Hash CT (2013) Massive sorghum collection genotyped with SSR markers to enhance use of global genetic resources. PLoS One. 8:e59714.

Borrell AK, Hammer GL, Douglas ACL (2000) Does maintaining green leaf area in sorghum improve yield under drought? I. Leaf growth and senescence. Crop Sci. 40: 11026-11030.

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J of Hum Genet. 32:314-331.

Brown SM, Hopkins MS, Mitchell SE, Senior ML, Wang TY, Duncan RR, Gonzalez-Candelas F, Kresovich S (1996) Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. Theor Appl Genet. 93:190-198.

Bryden WL, Selle PH, Cadogan DJ, Li D, Muller ND, Jordan DR, Gidley MJ, Hamilton WD (2009) A review of the nutritive value of sorghum for broilers. RIDRC Publication No 09/077. pp: 7.

Caniato F, Guimara˜es C, Schaffert R, Alves V, Kochian L, Borém A, Klein PE, Magalhaes JV (2007) Genetic diversity for aluminium tolerance in sorghum. Theor Appl Genet. 114:863-876.

CSA (2014) Centeral Statistical Agency of Ethiopia Agricultural sample survey report onarea and production of major crops, Addis Ababa, Ethiopia. pp 121.

Deu M, Sagnard F, Chantereau J, Calatayud C, He´rault D, Mariac C, Pham JL, Vigouroux Y, Kapran I, Traore PS, Mamadou A, Gerard B, Ndjeunga J, Bezançon G (2008) Niger-wide assessment of in situ sorghum genetic diversity with microsatellite markers. Theor Appl Genet. 116:903-913.

Dillon SL, Shapter FM, Henry RJ, Cordeiro G, Izquierdo L, Lee S (2007) Domestication to crop improvement: genetic resources for sorghum and saccharum (Andropogoneae). Ann Bot. 100:975-989.

Disasa T, Feyissa T, Admassu A, Paliwal R, De Villers SM, Odeny DA (2016a) Molecular evaluation of Ethiopian sweet sorghum germplasm and their contribution to regional breeding programs. Aust J Crop Sci. 10:520-527.

Disasa T, Feyissa T, Admassu A (2016b) Characterization of Ethiopian sweet sorghum accessions for °Brix, morphological and grain yield traits. Sugar Tech. DOI 10.1007/s12355-016-0440-3

Doggett H (1988) Sorghum (2nd edn). London: Longman.

Gebrekidan B (1982) Utilization of accession in sorghum improvement. In: House LR, Mughogho LK and Peacock JM (eds) Sorghum in the eighties. Proceedings of the international symposium on sorghum, ICRISAT, Patancheru, 1982.

Harris K, SubudhiPK, Borrell A, Jordan D, Rosenow D, Nguyen H, Klein P, Klein R, Mullet J (2007) Sorghum stay-green QTL individually reduce post-flowering drought induced leaf senescence. J Exp Bot. 58:327-338.

Haussmann BI, Mahalakshmi V, Reddy BV, Seetharama N, Hash CT, Geiger HH (2002) QTL mapping of stay-green in two sorghum recombinant inbred populations. Theor Appl Genet. 106:133-142.

Kong L, Dong J, Hart GE (2000) Characteristics, linkage-map positions and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple sequence repeats (SSRs). Theor Appl Genet. 101:438-448.

Kresovich S, Barbazuk B, Bedell JA (2005) Toward sequencing the sorghum genome. A US national science foundation-sponsored workshop report. Plant Physiol. 138:1898-1902.

Li M, Yuyama N, Luo L, Hirata M, Cai H (2009) In silico mapping of 1758 new SSR markers developed from public genomic sequences for sorghum. Mol Breeding. 24:41-47.

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics. 21:2128-2129.

Mace ES, Buhariwalla HK, Crouch JH (2003) A high-throughput DNA extraction protocol for tropical molecular breeding programs. Plant Mol Biol Rep. 21:459-460.

Mace ES, Tai S, Gilding EK, Li Y, Prentis PJ, Bian L, Campbell BC, Hu W, Innes DJ, Han X, Cruickshank A, Dai C, Fre`re C, Zhang H, Hunt CH, Wang X, Shatte T, Wang M, Su Z, Li J, Lin X, Godwin ID, Jordan DR,Wang J (2013) Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. Nat Commun. 4: 2320.

Mengistie Z (2009) Evaluation of post-flowering drought resistance property of Ethiopian sorghum accessions collected from the drought prone northern part. M. Sc. Thesis, Addis Ababa University, Addis Ababa.

Menz MA, Klein RR, Mullet JE, Obert JA, Unruh NC, Klein PE (2002) A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP, RFLP and SSR markers. Plant Mol Biol. 48:483-499.

Muraya MM, de Villiers S, Parzies HK, Mutegi E, Sagnard F, Kanyenji BM, Kiambi D, Geiger HH (2011) Genetic structure and diversity of wild sorghum populations (*Sorghum spp*.) from different eco-geographical regions of Kenya. Theor Appl Genet. 123:571-583.

Paterson, AH (2008) Genomics of sorghum. Int J Plant Genomics. 2008:1-6.

Perrier X, Jacquemoud-Collet JP (2006) DARwin software, http://darwin.cirad.fr/darwin

Rai KM, Thu SW, BalasubramanianVK, Cobos CJ, Disasa T, Mendu V (2016)

Identification, characterization, and expression analysis of cell wall related genes in *Sorghum bicolor* (L.) Moench, a food, fodder, and biofuel Crop. Front Plant Sci. 7:1-19.

Ramu P, Kassahun B, Senthilvel S, Kumar CA, Jayashree B, Folkertsma RT, Reddy LA, Kuruvinashetti MS, Haussmann BIG, Hash CT (2009) Exploiting rice–sorghum synteny for targeted development of EST-SSRs to enrich the sorghum genetic linkage map. Theor Appl Genet. 119:1193-1204.

Ramu P, Billot C, Rami JF, Senthilvel S, Upadhyaya HD, Reddy L A, Hash CT (2013)

Assessment of genetic diversity in the sorghum reference set using EST-SSR markers. Theor Appl Genet. 126:2051-2064.

Reddy BVS, Ramesh S, Reddy PS, Kumar AA (2009) Genetic enhancement for drought tolerance in sorghum. Plant Breed Rev. 31:189-222.

Reddy BVS, Kumar AA, Sharma HC, Srinivasa RP, Blümmel M, Ravinder, Reddy Ch, Sharma R, Deshpande S, DattaMazumdar S, Dinakaran E (2012) Sorghum improvement (1980–2010): Status and way forward. J SAT Agricultural Research.10:1-14.

Schloss SJ, Mitchell SE, White GM, Kukatla R, Bowers JE, Paterson AH, Kresovich S (2002) Characterization of RFLP clone sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. Theor Appl Genet. 105:912-920.

Schuelke M (2000) An economic method for the fluorescent labelling of PCR fragments. Nature Biotechnol. 18:233-234.

Smith JSC, Kresovich S, Hopkins MS, Mitchell SE, Dean RE, Woodman WL, Lee M, Porter K (2000) Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. Crop Sci. 40:226-232.

Wang ML, Zhu C, Barkley NA, Chen Z, Erpelding JE, Murray SC, Tuinstra MR, Tesso T, Pederson GA, Yu J (2009) Genetic diversity and population structure analysis of accessions in the US historic sweet sorghum collection. Theor Appl Genet. 120:13-23.

Wani SPR, Albrizio NR, Vaija (2011) Sorghum crop yield response to water. FAO, Rome, Italy, 2011.

Wu YQ, Huang Y (2006) An SSR genetic map of *Sorghum bicolor* (L.) Moench and its comparison to published genetic map. Genome.50:84-89.