Original Paper Open Access

Genetic diversity analysis of elite maize inbred lines of diverse sources using SSR markers

Lameck Nyaligwa[1,2], Shimelis Hussein[1], Beyene Amelework[1]*, Habteab Ghebrehiwot[1]

[1]African Center for Crop Improvement, University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa
[2]Selian Agricultural Research Institute (SARI), P.O. BOX 6024 Arusha, Tanzania
*Corresponding author: E-mail: assefa@ukzn.ac.za

Abstract

Genetically diverse and complementary elite inbred lines are fundamental for hybrid maize breeding or for strategic conservation. The objectives of this study were to determine the genetic diversity present among 79 elite maize inbred lines collected from five different sources using 30 SSR markers and to identify unique genotypes for breeding. The mean numbers of observed and effective alleles were 4.7 and 2.4 per marker, respectively. The polymorphic information content (PIC) of the 30 SSR markers ranged from 0.01 to 0.92, with a mean of 0.51. Six markers: phi 031, phi 308707, phi 96100, phi 114, phi 299852 and phi 063 revealed the highest PIC values greater than 0.70 suggesting the potential of these loci for genetic diversity analysis. The mean value of total gene diversity was 0.51 which partitioned 72% among inbred lines, 26% within inbred lines and 7% was attributed to variation between the five sources. Cluster analysis allocated the inbred lines into three main genetic groups. The principal coordinate analysis explained 67% of the total genetic variation detected among inbred lines and separated them into two main clusters. The study identified elite maize inbred lines such as TL2012-2, TL2012-14, TL2012-20, TL2012-31, TL2012-54 and TL2012-55 from Cluster I; TL2012-6, TL2012-35, TL2012-53 and TL2012-75 from cluster II; TL2012-12, TL2012-27, TL2012-29 and TL2012-58 from cluster III with marked genetic differences for hybrid breeding to exploit heterosis.

Keywords: cluster analysis, genetic diversity, inbred line, principal coordinate analysis, Zea mays, SSR markers

Introduction

Maize (Zea mays L; 2n = 2x = 20) is one of the most important cereal crops in the world providing food, feed and bioenergy (Linehan et al, 2013; Ranum et al, 2014). It is predominantly a cross pollinated crop showing the highest phenotypic and genotypic variability (Molin et al, 2013; Li et al, 2014). Development of maize hybrids to exploit heterosis relies on genetically diverse and complementary elite inbred lines (Singh et al, 2013). Subsequently, inbred lines should be accurately characterized and selected using molecular and phenotypic markers for effective breeding (Kage et al, 2012; Xu et al, 2013). Molecular markers are increasingly important and complementary to phenotypic traits and pedigree records to assessing the genetic diversity and relationship present among various maize germplasm for efficient selection of parents for breeding. Furthermore, molecular markers facilitate classification of breeding lines into well-defined heterotic groups aided by combining ability tests (Bidhendi et al, 2012).

The most common molecular markers used to assess genetic diversity in maize include restriction fragment length polymorphism (RFLP), random amplified polymorphic (RAPD), microsatellite or simple sequence repeats (SSRs), amplified fragment length polymorphism (AFLP) and single nucleotide polymorphism (SNP) (Molin et al, 2013). The SSR markers are known for their dominant inheritance, locus specificity, extensive genome coverage and simple detection of locus using labeled primers (Wu et al, 2010; Xu et al, 2013). SSR markers have been successfully used to genotype diverse maize germplasm collections (Semagn et al, 2012).

Productivity of maize is considerably low in Tanzania and sub-Saharan Africa due to a number of stress factors, predominantly by diseases and insect pests and unavailability of improved cultivars. Foliar diseases of maize including maize streak virus (MSV) caused by Geminivirus, leaf blight caused by *Exserohilum turcicum* Pass Leonard and Suggs, grey leaf spot (*Cercospora zeae-maydis* Tehon & Daniels) or common leaf rust (*Puccinia sorghi* Schr) inflict significant yield losses in the humid tropical regions (Lamia et al, 2010). Targeted hybrid breeding using elite, locally adapted and genetically diverse inbred lines may boost productivity of maize. The objectives of this study were to determine the genetic diversity present among 79 elite maize inbred lines collected from five different sources using 30 SSR markers and to identify unique genotypes for breeding.

 RECEIVED 06/23/2015

Materials and Methods

Maize inbred lines

The study used 79 maize inbred lines of different sources of geographic origins. The description of the inbred lines is presented in Supplementary Table 1. Thirty six inbred lines were obtained from the International Maize and Wheat Improvement Center (CIMMYT), of which 21 were from CIMMYT/Kenya and 15 from CIMMYT/Zimbabwe. Eighteen inbreds were acquired from Selian Agricultural Research Institute (SARI) of Tanzania, 21 from the International Institute of Tropical Agriculture (IITA) of Nigeria and four from the University of KwaZulu-Natal (UKZN), South Africa. The inbred lines were selected for their reported tolerance to maize steak virus (MSV). The lines are stable and kept homozygous through continuous controlled selfing.

DNA extraction

Inbred lines were planted in the northern Tanzania at Ngaramtoni research site of SARI in 2012. Leaf samples of 5-6 cm long were sampled from 10 young maize plants of about 3 to 4 weeks old. DNA extraction was done using the modified CTAB procedure (Irfan et al, 2013). The leaf samples were stored overnight at -80˚C for DNA extraction. About 100 mg freeze-dried leaf samples were crushed into fine powder using GenoGrinder-2000 (SPEX Sample Prep, LLC, NJ,USA) at a speed of 500 strokes per minute by shaking for 4 minutes. The samples were grinded for an additional 2 minutes after the addition of 600 µl of freshly prepared modified CTAB DNA extraction buffer. The samples were incubated at 65˚C water bath for 30 minutes with continuous gentle shaking. The tubes were then removed and cooled for 5 - 10 minutes in a fume hood and subjected to centrifugation at 3,500 rpm for 10 minutes at 15˚C. The supernatant was transferred into new micro-tubes and 400 µl chloroform: isoamylalcohol (24:1) was added into the side of the tubes and mixed gently. Samples were shaken for up to 30 minutes at room temperature. The corrosive chloroform was removed carefully using pipette. The aqueous layer was transferred to fresh strip tubes and the chloroform: isoamylalcohol was washed repeatedly to produce a clean DNA solution.

To each sample 300 µl of isopropanol was added and mixed very gently for DNA precipitation while keeping the tubes in the -20˚C freezer overnight. The samples were then centrifuged at 3500 rpm for 30 minutes and the supernatant was discarded to obtain the DNA pellet. Each DNA pellet was further washed with 70% ethanol and centrifuged for 15 minutes. The supernatant was discarded and traces of ethanol were removed by air dry the DNA pellet for about 15 - 20 minutes. The DNA was suspended in 150 µl of 10 mM Tris-HCl ph 8.3 and the samples were incubated for about 45 minutes at 45˚C water bath with gentle tapping every 10 minutes. Each sample was treated with 3µl RNAse and the RNAse was spanned down with centrifuge (3,500 rpm for 1 - 2 minutes). The DNA was incubated at 37˚C water bath for 3 hours and finally the extracted DNA was stored at 4˚C for further use.

SSR analysis

Thirty SSR markers were used in this study. The markers were selected from the maize genome database (http://www.agron.missouri.edu) based on their degree of polymorphisms and distribution among the maize genome. Genotyping was done using a standard PCR protocol for maize SSR markers (CIMMYT, 2005). The SSR analysis involved preparation of the cocktail mix composed of PCR products, highly deionized (Hi-Di) formamide and GENESCAN 500 internal lane size standard (LIZ-500) labeled with N, N, N', N'-tetramethyl-6-carboxyrhodamine (TAMARA) (Perkin Elmer - Applied Biosystems). The PCR mix was prepared in a total reaction volume of 15 µl containing 2 µl (50 ng) genomic DNA, 2.5 mM magnesium chloride ($MgCl_2$), 0.4 mM of dNTPs, 50 ng of each forward and revers primers, 2 µl of 1x reaction buffer, 0.1 ml Taq DNA polymerase and distilled water to bring volume to 15 µl. Samples containing 1.2 µl of the PCR products, 1.0 ml (Hi-Di) formamide and 12 µl of LIZ-500 internal lane size standard was denatured at 95°C for 3 minutes and placed on ice for 5 minutes. DNA samples were electrophoresed on an ABI-3730 automatic DNA sequencer (Applied Biosystems, USA) equipped with GENESCAN 672 software v. 1.2 (PE-Applied Biosystems). The 2 µl of each DNA sample was loaded in the polymerase chain reaction (PCR) and the resultant PCR fragments were resolved on the genetic analyzer, the ABI 3730. A total of 2,362 data points were captured out of the expected 2,370 data points using the Genscan®software giving an overall success rate of 99.7%.

Data analysis

Genetic diversity analysis

Genetic diversity of 79 maize inbred lines was analyzed using GenAlex version 6.5 (Peakall and Smouse, 2007) software program. The χ^2 test was also performed to determine if the allelic frequencies among the 30 SSR markers used were significant. The genetic diversity parameters considered in this study were: the total number of alleles per locus (N_a), the number of effective alleles per locus (N_e), observed and expected heterozygosity denoted by (H_o) and (H_e), respectively. Other genetic parameters estimated were: polymorphic information content (PIC), and fixation index (F_{IS}) (Nei's, 1978). The polymorphic information content was estimated according to Smith et al (1997):

$$PIC = \sum_{i=1}^{N} p_i^2$$

where, p_i is the frequency of the i^{th} allele.

Genetic distance and cluster analysis

To examine the degree of population differentia-

tion among the study material, gene flow, the Nei's unbiased genetic distance and identity were estimated according to Nei (1978) using GenAlex. The genetic relationships or relatedness of 79 sampled inbred lines were estimated using neighbour-joining algorithm using the unweighted pair group method (UPGMA) in DARwin 5.0 software (Perrier and Jacquemoud-Collet, 2006). A dendrogram for 79 inbred lines was then generated based on the dissimilarity matrix to visualize pattern of clusters within and among inbred lines. Further, a principal coordinate analysis was performed to complement clustering or grouping patterns revealed by the dendrogram. The genetic structure was investigated as described by Nei's (1978) analysis.

Analysis of molecular variance (AMOVA)

Analysis of Molecular Variance was performed to estimate population genetic structure and differentiation among and within the sets of inbred lines based on their region of origin. AMOVA uses the estimated F statistics such as F_{ST}, F_{IS}, and F_{IT} to compare the genetic structure among and within populations. For easy management and utilization, the total molecular variance was dissected into within and among population variations. The AMOVA procedures were done using GenALex based on (Nei, 1978).

Results

Polymorphism and allelic diversity of SSR markers

Polymorphism among the 79 maize inbred lines was investigated using 30 SSR markers. These markers, revealed a total of 140 putative alleles. The observed number of alleles (N_a) varied from 2 (umc 2250, umc 1266, and phi 062) to 11 (phi 96100), with a mean of 4.7. The number effective alleles (N_e) detected varied from 1.0 (umc 1266) to 4.7 (phi 031), with a mean of 2.4 per locus (Table 1). More than half of the total numbers of observed alleles in this study was detected by 13 (43%) of the markers used suggesting the existence of significant polymorphism among the markers. The results further showed that the observed heterozygosity (H_o) varied from zero (umc 2250) to 0.86 (phi 031), with a mean of 0.14, signifying that 90% of the loci were fixed and reached in relatively homozygous state. The expected heterozygosity also varied slightly similar to the observed heterozygosity with values ranging between 0.013 (umc 1266) and 0.793 (phi 031).

The polymorphic information content (PIC) values for all markers ranged from 0.01 (umc 1266) to 0.92 (umc 96100), with a mean value of 0.51. Similarly, half of the loci used revealed a PIC value of greater than the overall mean of 0.51 indicating that most of the markers used had high discriminatory power and are useful for genetic diversity studies. Marker phi 96100 was the most polymorphic locus with a PIC value of 0.92 followed by phi 031 (PIC = 0.79) and phi 063 (PIC

Table 1 - Genetic parameters of the 30 SSR markers used in the study of 79 maize inbred lines collected from five sources.

Locus	Repeats	Bin#	Genetic parameters					
			N_a $^{\$}$	N_e	H_o	H_e	F_{IS}	PIC
phi 056	CCG	1.00	8	2.52	0.40	0.61	0.33	0.61
umc 1917	(CTG)6	1.04	5	1.48	0.10	0.33	0.69	0.32
umc 2047	(GACT)4	1.09	3	1.65	0.05	0.40	0.87	0.39
phi 308707	AGC	1.10	5	3.62	0.11	0.73	0.84	0.72
phi 227562	ACC	1.11	6	2.74	0.25	0.64	0.61	0.64
phi 96100	ACCT	2.01	11	4.37	0.10	0.78	0.87	0.92
umc 2250	(ACG)4	2.04	2	1.16	0.00	0.14	1.00	0.14
phi 374118	ACC	3.02	5	2.33	0.08	0.58	0.87	0.57
phi 029	AG/AGCG	3.04	4	2.07	0.08	0.52	0.85	0.52
phi 102228	AAGC	3.06	3	1.34	0.09	0.26	0.65	0.26
umc 1266	(CAG)4	3.06	2	1.01	0.01	0.01	0.01	0.01
phi 046	ACGC	3.08	4	1.46	0.09	0.31	0.72	0.31
phi 072	AAAC	4.00	5	3.21	0.14	0.69	0.80	0.69
phi 093	AGCT	4.08	4	3.03	0.08	0.67	0.89	0.67
nc 130	AGC	5.00	4	1.71	0.33	0.42	0.21	0.41
phi 331888	AAG	5.04	4	2.41	0.05	0.59	0.91	0.59
phi 075	CT	6.00	4	1.99	0.08	0.5	0.85	0.50
phi 031	GTAC	6.04	7	4.71	0.86	0.79	0.09	0.79
phi 299852	AGC	6.07	9	3.67	0.14	0.73	0.81	0.73
phi 112	AG	7.01	4	1.98	0.03	0.5	0.95	0.50
phi 114	GCCT	7.03	5	3.52	0.13	0.72	0.82	0.72
phi 069	GAC	7.05	3	2.71	0.10	0.63	0.84	0.63
umc 1304	(TCGA)4	8.02	4	1.58	0.05	0.37	0.86	0.37
phi 014	GGC	8.04	4	2.27	0.19	0.56	0.66	0.56
phi 108411	AGCT	9.05	3	1.92	0.09	0.48	0.81	0.48
phi 041	AGCC	10.00	5	3.61	0.17	0.73	0.77	0.72
phi 063	TATC	10.02	6	4.05	0.16	0.76	0.78	0.75
umc 1367	(CGA)6	10.03	5	1.20	0.05	0.17	0.70	0.17
phi 062	ACG	10.04	2	1.24	0.04	0.19	0.80	0.19
phi 084	GAA	10.05	4	1.77	0.04	0.44	0.91	0.43
Overall mean			4.67	2.41	0.14	0.51	0.72	0.51
SE			0.37	0.19	0.03	0.04	0.05	0.04

$^{\$}N_a$ - number of observed alleles; N_e - Number of effective alleles; H_o - Observed heterozygosity; H_e - Average gene diversity within genotypes; F_{IS} - Fixation index; PIC - Polymorphic information content.

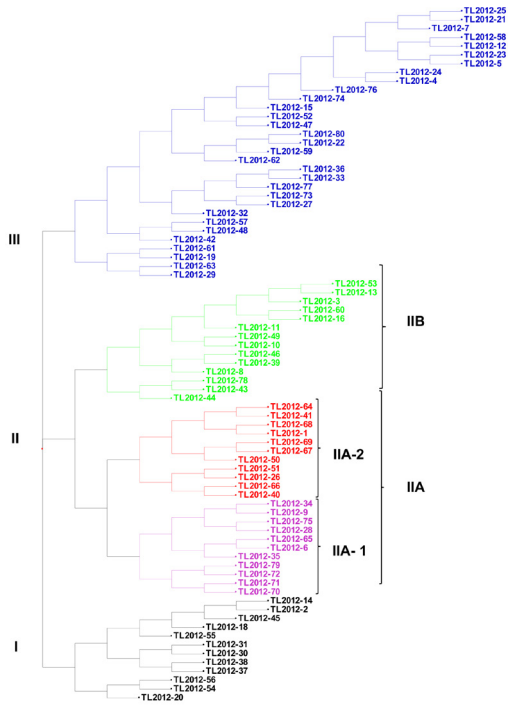

Figure 1 - Dendrogram showing the genetic relationship of 79 maize inbred lines using 30 SSR markers.

= 0.75). The results of the χ^2 test showed significant differences in major allele frequencies at all loci for all sets of inbred lines. The fixation index level (F_{IS}) that measures the level of inbreeding among and within inbred lines varied significantly from 0.01 (umc 1266) to 1.0 (umc 2250) with a mean of 0.72, indicating the presence of appreciable levels of homozygosity among the study materials (Table 1).

Cluster and principal component analyses

The dendrogram discriminated and clustered the inbred lines into three major clusters (clusters I, II and III) with few sub-clusters found in cluster II (Figure 1). The distribution of the inbred lines into these three main clusters was not homogeneous. Cluster I consisted of 12 inbred lines collected from SARI/ Tanzania, CIMMYT/Kenya or Zimbabwe and UKZN/ South Africa. Cluster II consisted of the largest number of inbred lines (36) and it was further sub divided into two sub-clusters (IIA and IIB) (Figure 1). The sub-cluster IIA contained 29% of the genotypes studied and it further sectioned into two sub-sub clusters (IIA-1 and IIA-2). Sub-clusters IIA-1 and IIA-2 contained 11 genotypes each, which were originated from CIMMYT/Kenya and IITA/Nigeria, respectively. Sub-cluster IIB consisted of 14 inbred lines mainly from CIMMYT/Kenya and SARI/Tanzania. Cluster III comprised of 31 geographically diverse inbred lines, of which 36% were from CIMMYT/Zimbabwe, 26% from SARI/Tanzania, 23% from IITA/Nigeria and 15% from CIMMYT/Kenya. Further, principal coordinate analysis was constructed to examine the genetic clustering of the 79 inbred lines using genetic distances (Figure 2). The percentages of variance for the first 2 principal coordinates were 52.7% and 14.3% with a total variance of 67.0%. Based on the principal coordinate analysis, the 79 inbred lines were classified into two major groups (Group A and B) (Figure 2). Group B consisted of 86% of the inbred lines and Group B had 11 inbred lines. Of the 11 inbred lines allocated in Group B, 8 were sourced from CIMMYT/ Zimbabwe signifying that these inbred lines are genetically unique.

Genetic variability within and among populations

The mean values of observed (N_a) and effective (N_e) number of detected alleles were 3.8 and 2.4, respectively. The lowest values of 2.6 and 1.8 for the same parameters, respectively, were recorded by inbred lines sourced from Zimbabwe. Similarly, the highest N_a and N_e values of 3.8 and 2.4, respectively, were recorded from IIITA/ Nigeria inbred line collections. The mean observed heterozygosity (H_o) and expected heterozygosity (H_e) across inbred lines were 13 and 47%, respectively. Their respective lowest values of 0.10 and 0.38 were recorded from inbred lines collected from CIMMYT/Zimbabwe. The highest value of H_o = 0.15 and H_e = 0.51 were recorded from IITA/Nigeria inbred lines (Table 2). In general, the probability of two randomly sampled alleles in a given genotype to be different was 50%. Highest polymorphism among the inbred lines within the geographic origins was observed and on average, 93% of the loci were polymorphic (Table 2).

Analysis of molecular variances

AMOVA partitioned the total molecular variances into within and among the sets of inbred lines evaluated based on their source of collection or geographic origins. There was a highly significant difference ($P \leq 0.001$) of molecular variance within and among inbred lines and within geographic origins. The largest genetic variability (67%) was attributed to variation among inbred lines within sources of collections,

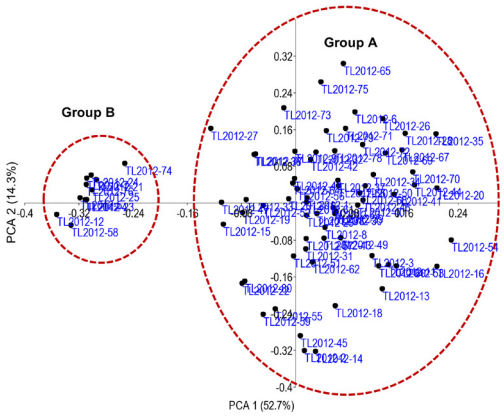

Figure 2 - Principal coordinate analysis of genetic grouping of 79 maize inbred lines assessed by 30 SSR markers

Table 2 - Genetic diversity of 79 inbred lines among five regions of origins.

Populations	Genetic parameters						
	N[$]	N_a	N_e	H_o	H_e	F_{IS}	P_p
CIMMYT/Kenya	21	3.53	2.34	0.12	0.50	0.75	0.93
CIMMYT/Zimbabwe	15	2.60	1.78	0.10	0.38	0.76	0.93
IITA/Nigeria	21	3.77	2.37	0.15	0.51	0.70	0.97
SARI/Tanzania	18	3.37	2.17	0.18	0.47	0.57	0.97
UKZN/South Africa	4	2.37	2.12	0.11	0.50	0.72	0.83
Overall mean	15.75	3.13	2.15	0.13	0.47	0.70	0.93
SE	0.51	0.11	0.08	0.01	0.02	0.03	0.02

[$]N = Population size; N_a - number of observed alleles; N_e - Number of effective alleles; H_o - Observed heterozygosity; H_e - Average gene diversity within genotypes; F_{IS} - Fixation index; Pp - Percent polymorphism.

while 26% of the total variation was explained by variation within inbred lines and 7% of variation was explained by variation between sources of collection (Table 3). This signifies that the sources of inbred lines had small but significant contribution to the total molecular variances detected and that the maximum variation was among the inbred lines within sources of collection.

Population differentiation based on source of collection or geographic origin

The genetic differentiation (F_{ST}) among inbred lines within source of collection or region of origin ranged from 0.03 between SARI/Tanzania and CIMMYT/Kenya inbred lines to 0.15 between UKZN/South Africa and CIMMYT/Zimbabwe lines (Table 4). The low values of F_{ST} imply that there is high frequency of identical alleles among inbred lines between regions of origins, hence, low genetic differentiation of inbred lines among sources or regions. Gene flow (N_m) or gene migration varied considerably among the sources of collection with consistently low values between UKZN/South Africa and CIMMYT/Zimbabwe lines with the rest of the sources. A considerable amount of gene flow existed among IITA/Nigeria, SARI/Tanzania and CIMMYT/Kenya collections probably due to germplasm exchange among the CGIAR centers. The genetic distances (GD) of inbred lines among sources of collection were small that ranged from 0.04 between SARI/Tanzania and CIMMYT/Kenya to 0.25 between CIMMYT/Zimbabwe and UKZN/South Africa. This suggests that there were substantial genetic relationships of inbred lines despite the varied source of collections. Similarly, the genetic identity (GI) varied from 0.83 to 0.96 signifying the close relationship among inbred lines of diverse sources (Table 4).

Discussion

In this work, the genetic diversity of maize inbred lines from five diverse sources or geographic origins were assessed based on 30 SSR markers. The present study found a total number of alleles of 140 with mean of 4.7 alleles. Gichuru (2013) assessed the genetic diversity of 40 MSV resistant maize inbred lines using 28 SSR markers and reported a total number of observed alleles of 135 with a mean of 4.8 per SSR locus. Mora et al (2013) found effective number of alleles of 4.8 when assessing 28 Brazilian popcorn populations with 11 SSR markers. A mean of 4.9 alleles per locus were reported by Warburton et al (2002) after genotyping 57 inbred lines derived from CIMMYT/Mexico with 85 SSRs and by Lu and Bernardo (2001) when determining the genetic diversity of 40 inbred lines sourced from the USA with 83 SSR markers.

The mean H_e value in the present study (0.51) and effective number of alleles (2.4) per locus were similar to Rupp et al (2009) in their studies of genetic diversity between inbred lines of popcorn progenies. However, the mean H_e in this study was lower than reported previously. Patto et al (2004), Enoki et al (2002), and Pejic et al (1998) reported higher H_e values of 0.53, 0.69, and 0.72, respectively. This may probably be due to a higher average number of detected alleles per locus. The low H_o and high F_{IS} values observed in this study indicated the occurrence of maximum level of inbreeding (Labate et al, 2003) and the tested inbred lines showed considerably high level of homozygosity in 90% of the loci useful for hybrid breeding.

Xu et al (2013) reported that genetic diversity if well managed can be used to enhance biotic and abiotic stress tolerance in any crop species. Kanagarasu et al (2013) also indicated that successful adaptation

Table 3 - Analysis of molecular variance (AMOVA) among 79 maize inbred lines assembled from five geographic origins using 30 SSR markers.

Source of variation	df[$]	SS	MS	Est var	Per var	F-Statistics
Among sources or geographic origins/regions	4	114.7	28.7	0.5	7	0.001
Among lines within region	74	929.3	12.6	5.3	67	0.001
Within line	79	160.0	2.0	2.0	26	0.001
Total	157	1204.1	43.3	7.8	100	

[$]DF - Degree of freedom; SS - sum of squares; MS - mean sum of squares; Est var - estimated variance; Per Var - Percentage variation.

Table 4 - Pair-wise estimates of gene flow (N_m; above diagonal off brackets), genetic differentiation (F_{ST}; above diagonal within brackets); genetic distance GD (lower diagonal off brackets) and genetic identity (GI; lower diagonal within brackets).

Origins	CIMMYT/KEN[$]	CIMMYT/ZIM	IITA/NIG	SARI/TAZ	UKZN/SA
CIMMYT/KEN		3.85 (0.06)	6.10 (0.04)	7.39 (0.03)	2.25 (0.10)
CIMMYT/ZIM	0.09 (0.92)		2.54 (0.09)	3.50 (0.07)	1.39 (0.15)
IITA/NIG	0.06 (0.94)	0.16 (0.85)		6.14 (0.04)	1.84 (0.12)
SARI/TAZ	0.04 (0.96)	0.10 (0.91)	0.06 (0.95)		2.11 (0.11)
UKZN/SA	0.12 (0.89)	0.25 (0.78)	0.19 (0.83)	0.15 (0.86)	

[$]CIMMYT - International Maize and Wheat Improvement Centre; IITA - International Institute of Tropical Agriculture; UKZN - University of KwaZulu-Natal; SARI - Selian Agricultural Research Institute; KEN - Kenya; NIG - Nigeria; TAZ - Tanzania; SA - South Africa; ZIM - Zimbabwe.

to certain agro-climatic conditions and improvement of any crop species depends on the availability of genetic diversity within the available breeding material. Information about the genetic diversity and relationships among diverse genetic resources is very valuable in crop improvement programmes and for strategic conservation of genetic resources (Abera et al, 2012; Kage et al, 2013; Wu et al, 2014). Therefore, the high level of genetic diversity identified in the current work will assist maize breeders to set out their breeding objectives and to select potential parents to be used in their breeding programmes.

PIC demonstrates the ability of the SSR loci to detect differences among the inbred lines based on the number of allele per locus and evenness of allele frequencies. The mean PIC value determined using the 30 SSR loci in this work (PIC = 0.51) agrees with the earlier reports of Oppong et al (2014) who estimated a mean PIC value of 0.50 in bulk genetic characterization of Ghanaian maize landraces using microsatellite markers. Some markers showed significant discrimination power than others, in this study six SSR loci (phi 031, phi 308707, phi 96100, phi 114, phi 299852, and phi 063) identified the largest mean number of alleles (6.9) and mean PIC (0.76) reflecting their potential to detect differences between the inbred lines.

Based on the information available in MaizeGDB (http://www.maizegdb.org) phi 96100 tags the gene *os1* (*opaque-endosperm 1*) and *rgh2* (*rough-endosperm 2*) located on chromosome 2 (bin 2.01), while phi 308707 tags the gene *sen3* (*soft-endosperm 3*) on chromosome 1 (bin 1.10). SSR markers phi 299852 and phi 114 located in the chromosomal regions with genes *tsh1* (*tassel sheath1*) and *tsh4* (*tassel sheath 4*), respectively control male and female inflorescent development. Similarly, phi 063 located on the vicinity of gene *aasr1* (*abscisic acid stress ripening 1*) increases kernel yield under water deficit. A major QTL for anthesis-silking interval was reported in the vicinity of phi014 and phi031 (Ribaut et al 1996). Thus, the highly polymorphic SSR alleles identified in this study appear to be associated with some important characters. Consequently, these may be useful for further association mapping of the underlying genes.

The SSR loci umc1367 and phi 041 both detected the same number of alleles (N_a = 5), but they had different PIC values of 0.17 and 0.72, respectively. The tri and tetra nucleotide SSR loci (Phi 041, phi 96100, phi 031, phi 308707, phi 299852, phi 114) identified the highest number of alleles (> 5) and PIC (> 0.70), as compared to higher repeat motifs. This implies that there is significant correlation between the repeat size and number of allele (Legesse et al, 2007; Abera et al, 2012). Discriminatory power of a given locus depends on many factors such as rate of amplification, length of repeat and detection ability of the loci. Missing values in the dataset could also affect differently the polymorphism and hence discrimination ability of markers (Inghelandt et al, 2010).

Cluster analysis grouped the 79 inbreds into three main clusters. However, classifications of inbred lines were independent to source of collection. The principal coordinate analysis revealed that maize lines from CIMMYT/Zimbabwe displayed unique clustering pattern than the rest of the collections, although they could possibly share common genetic background being sourced from CIMMYT. The relatively high differentiation of CIMMYT/Zimbabwe lines against other collections could be attributed to high inbreeding (F_{IS} = 0.38) and low number of effective alleles (2.6) (Saavedra et al, 2013).

In the present study AMOVA revealed 67% of the molecular variation in the maize inbred lines accounted by among inbred lines variations while 26% and 7% of the total variation was partitioned within inbred lines and among source of collection, respectively. The magnitudes of between and within population differentiation were quantified using the F-statistics of Wright (1951). Genetic differentiation among the regions of origin was low (F_{ST} = 0.07) but significant with low effects of non-random mating within the populations (F_{IS} = 0.72). The low genetic differentiation among sources of collection may be due to the high degree of inbreeding and high gene flow. Van Heerwaarden et al (2010) reported relatively lower value of F_{ST} = 0.02. On the other hand, Kashiani et al (2012) reported strong genetic isolation with a mean values F_{ST} = 0.96 and N_m = 0.01 among the different populations of tropical sweet corn inbred lines using microsatellite markers.

A low F_{ST} value of 0.03 was observed among IITA/Nigeria, CIMMYT/Kenya and SARI/Tanzania inbred lines of in the present study. This can be explained by the high level of gene flow, which leads to genetic homogeneity (Amelework et al, 2015). The high level of gene flow observed among the inbreds may be at-

tributed to an exchange of genetic materials between CIMMYT and collaborating countries for germplasm evaluation, breeding and release. In addition, the germplasm used in this study were selected based on their reported MSV resistance which may enhance accumulation of other desirable alleles of phenotypic traits (Carvalho et al, 2012).

In conclusion, the present study demonstrated the existence of considerable degree of genetic diversity among maize inbred lines adapted to the mid altitude maize growing conditions of northern Tanzania and exotic lines using 30 SSR markers. Using a dendrogram based on genetic distance estimates, the following inbred lines were selected: TL2012-2, TL2012-14, TL2012-20, TL2012-31, TL2012-54 and TL2012-55 from Cluster I; TL2012-6, TL2012-35, TL2012-53 and TL2012-75 from cluster II; TL2012-12, TL2012-27, TL2012-29 and TL2012-58 from cluster III useful for strategic maize breeding or conservation. Crosses involving these inbred lines may result in significant and measurable yield gains due to hybrid vigour or heterosis. Further, these inbred lines had considerable phenotypic diversity with regards to plant height, ear height, number of tassel branches, grain yield, days to 50% tasseling and silking and resistance to MSV (data not presented). The variability of the inbred lines for these agronomic traits could be attributed to differences in genetic constitution, and source of collection or origins suggesting the value of classification of the lines for effective breeding and efficient management of the collections. The microsatellites loci phi 031, phi 308707, phi 96100, phi 114, phi 299852 and phi 063 revealed the highest PIC values suggesting the effectiveness and potential of these markers for genetic analysis.

Acknowledgements

The Alliance for a Green Revolution in Africa (AGRA) is sincerely acknowledged for financial support for the study. The Ministry of Agriculture, Food Security and Cooperatives of Tanzania through the Department of Research and Development of Tanzania is also thanked for providing leave of absence and research facilities to the first author. The International Maize and Wheat Improvement Center (CIMMYT) in Kenya and Zimbabwe and the International Institute of Tropical Agriculture (IITA)/Nigeria are sincerely thanked for providing some of the inbred lines.

References

Abera W, Shimelis H, Derera J, Mosisa W, Danson J, Laing MD, 2012. Genetic interrelationships among medium to late maturing tropical maize inbred lines using selected SSR markers. Euphytica 10: 1-9

Amelework B, Shimelis H, Tongoona P, Laing M, Mengistu F, 2015. Genetic variation in lowland sorghum (*Sorghum bicolor* (L.) Moench) landraces assessed by simple sequence repeats. Plant Genet Resour 13:131-141

Bidhendi MZ, Choukan R, Darvish F, Mostafavi K, Majidi E, 2012. Classifying of maize inbred iines into heterotic groups using diallel analysis. World Acad Sci Eng Technol 6: 1161-1164

CIMMYT, 2005. Laboratory protocols: CIMMYT applied molecular genetics laboratory. Third edition. Mexico DF

Enoki H, Sato H, Koinuma K, 2002. SSR analysis of genetic diversity among maize inbred lines adapted to cold regions of Japan. Theor Appl Genet 104: 1270-1277

Gichuru LM, 2013. Breeding investigations on utility of maize streak virus resistance germplasm for hybrid development in the tropics. PhD thesis, University of KwaZulu-Natal, Pietermaritzburg, South Africa

Inghelandt DV, Melichinger AE, Lebreton C, Stich B, 2010. Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. Theor Appl Genet 7: 1289-1299

Irfan M, Ting ZT, Yang W, Chunyu Z, Qing M, Lijun Z, Feng L, 2013. Modification of CTAB protocol for maize genomic DNA extraction. Res J Biotech 81: 41-45

Kage U, Madalageri D, Malakannavar L, Ganagashetty P, 2013. Genetic diversity studies in newly derived inbred lines of maize (*Zea mays* L.). Mol Plant Breed 4: 77-83

Kanagarasu S, Nallathambi G, Ganesan KN, Kannan S, Shobhana VG, Senthil N, 2013. Determination of genetic polymorphism among indigenous and exotic maize inbreds using microsatellite markers. Afr J Biotech 12: 5723-5728

Kanyamasoro MG, Karungi J, Asea G, Gibson P, 2012. Determination of the heterotic groups of maize inbred lines and the inheritance of their resistance to the maize weevil. Afr Crop Sci J 20: 99-104

Kashiani P, Saleh G, Panandam JM, Abdullah NAP, Selamat A, 2012. Molecular characterization of tropical sweet corn inbred lines using microsatellite markers. Maydica 57: 154-163

Labate JA, Lamkey KR, Mitchell SE, Kresovich S, Sullivan H, Smith JSC, 2003. Molecular and historical aspects of corn belt dent diversity. Crop Sci 43: 80-91

Lamia K, Hedia B, Jean-Marc A, Neila TF, 2010. Comparative analysis of genetic diversity in Tunisian apricot germplasm using AFLP and SSR markers. Scie Horticult 127: 54-63

Legesse WB, Myburg AA, Pixley KV, Botha AM, 2007. Genetic diversity of maize inbred lines revealed by SSR markers. Hereditas 144: 10-17

Li L, Eichten SR, Shimizu R, Petsch K, Cheng-Ting Y, Wu W, Chettoor AM, Givan SA, Cole RA, Fowler JE, Evans MMS, Scanlon MJ, Yu J, Schnable PS,

Timmermans MCP, Springer NM, Muehlbauer GJ, 2014. Genome-wide discovery and characterization of maize long non-coding RNAs. Genome Biol 15: 1-15

Linehan V, Thorpe S, Gunning-Trant C, Heyhoe E, Harle K, Hormis M, Harris-Adams K, 2013. Global food production and prices to 2050: Scenario analysis under policy assumptions, ABARES conference paper 13.6, Canberra

Lu H, Bernardo R, 2001. Molecular marker diversity among current and historical maize inbreds. Theoret Appl Genet 103: 613-617

Molin D, Coelho CJ, Máximo DS, Ferreira FS, Gardingo JR, Matiello RR, 2013. Genetic diversity in the germplasm of tropical maize landraces determined using molecular markers. Genet Mol Res 12: 99-114

Mora F, Saavedra J, Silva TA, Scapim CA, 2013. Bayesian analysis of the genetic structure of a Brazilian popcorn germplasm using data from simple sequence repeats (SSR). Chil J Agric Res 73: 99-107

Nei M, 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genet Mol Res 89: 583-590

Oppong A, Bedoya CA, Ewool MB, Asante MD, Thompson RN, Adu-Dapaah H, Lamptey JNL, Ofori K, Offei SK, Warburton ML, 2014. Bulk genetic characterization of Ghanaian maize landraces using microsatellite markers. Maydica 59: 1-8

Patto MCV, Satovic Z, Pego S, Fevereiro P, 2004. Assessing the genetic diversity of Portuguese maize germplasm using microsatellite markers. Euphytica 137: 63-72

Peakall R, Smouse PE, 2007. GENALEX6: genetic analysis in excel. Population genetic software for teaching and research. Mol Ecol 6: 288-295

Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Castiglioni P, Taramino G, Motto M, 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs and AFLPs. Theor Appl Genet 97: 1248-1255

Perrier X, Jacquemoud-Collet JP, 2006. DARwin software. http://www.darwin.cirad.fr/darwin.html. Accessed 13 July 2014

Ranum P, Pena-Rosas JP, Garcia-Casal MN, 2014. Global maize production, utilization, and consumption. Ann NY Acad Sci 1312: 105-112

Ribaut JM, Hoisington DA, Deutsch JA, Jiang C , Gonzalez-de-Leon D, 1996. Identification of quantitative trait loci under drought conditions in tropical maize: 1. Flowering parameters and the anthesis-silking interval. Theor Appl Genet 92: 905–914

Rupp JP, Mangolin CA, SCapim MF, Machado PDS, 2009. Genetic structure and diversity among sweet corn (Su1-germplasm) progenies using SSR markers. Maydica 54: 125-132

Saavedra J, Silva TA, Mora F, Scapim CA, 2013. Bayesian analysis of the genetic structure of a Brazilian popcorn germplasm using data from simple sequenec repeates (SSR). Chil J Agric Res 73: 99-107

Semagn K, Magorokosho C, Vivek BS, Makumbi D, Beyene Y, Mugo S, Prasanna BM, Warbuon ML, 2012. Molecular charcterization of diverse CIMMYT maize inbred lines from eastern and souther Africa using single nucleotide polymorphic markers. BMC Genomics 13: 113-121

Singh RK, Jena SN, Khan S, Yadav S, Banarjee N, Raghuvanshi S, Bhardwaj V, Dattamajumder SK, Kapur R, Solomon S, Swapna M, Srivastava S, Tyagi AK, 2013. Development, cross-species/genera transferability of novel EST-SSR markers and their utility in revealing population structure and genetic diversity in sugarcane. Gene 524: 309-329

Smith JSC, Chin ECL, Shu H, Smith OS, Wall SJ, Senior ML, Mitchell SE, Kresovich S, Ziegle J, 1997. An evaluation of utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. Theor Appl Genet 95: 163-173

van Heerwaarden J, Ross-Ibarra J, Doebley J, Glaubitz JC, Gaut BS, Eguiarte LE, 2010. Fine scale genetic structure in the wild ancestor of maize (*Zea mays ssp. parviglumis*). Mol Ecol 19: 1162-1173

Warburton ML, Xianchun X, Crossa J, Franco J, Melchinger AE, Frisch M, Bohn M, Hoisington D, 2002. Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. Crop Sci 42: 1832-1840

Wright S, 1951. The genetical structure of populations. Ann Eugenics 15: 323-354

Wu K, Yang M, Liu H, Tao Y, Mei J, and Zhao Y, 2014. Genetic analysis and molecular characterization of Chinese sesame (*Sesamum indicum* L.) cultivars using insertion-deletion (indel) and simple sequence repeat (SSR) markers. BMC Genomics 35: 1-15

Wu CL, Li SF, Dong BX, Zhang QQ, Zhang CQ, 2010. Determination of the number of SSR alleles necessary for the analysis of genetic relationships between maize inbred lines. Agric Sci China 9: 1713-1725

Xu J, Liu L, Xu Y, Chen C, Rong T, Alli F, Zhou S, Wu F, wang LYJ, Cao M, 2013. Development and characterization of simple sequence repeat markers providing genome- wide coverage and high resolution in maize. DNA Res 20: 497-509