Genetic Variability and Classification of Highland Adapted Quality Protein Maize Inbred Lines Using SSR Markers

Demissew Abakemal[1,*] and Shimelis Hussein[2]

[1]*Ethiopian Institute of Agricultural Research, Ambo-PPRC, Ethiopia, P.O.Box 37, Ambo, Ethiopia*

[2]*African Center for Crop Improvement, University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa.*

[*]*Corresponding author, Email: d.abakemal@gmail.com*

Abstract

Development of improved quality protein maize (QPM) varieties/hybrids would complement strategies for reducing problems of malnutrition in developing countries such as Ethiopia.The highland maize breeding program in Ethiopia, in collaboration with CIMMYT, has developed QPM inbred lines adapted to highland sub-humid maize agro-ecology. However, there is limited information on the genetic variability and interrelationship among the QPM inbred lines. The present study was, therefore, conducted to assess the genetic variability and thereby classify elite QPM inbred lines developed for tropical-highlands and highland transition maize agro-ecologies using microsatellite (SSR) markers. A total of 36 white-grained maize inbred lines, including 30 QPM and six non-QPM were genotyped using 26 simple sequence repeat (SSR) markers. Estimates of the average number of alleles per locus, gene diversity, and polymorphism information content (PIC) were 3.8, 0.53, and 0.49, respectively. Pair-wise Euclidean genetic distances ranged from 0.11 to 1.10 with mean of 0.74.Three major genetic groups were also identified, which are generally consistent with available pedigree information except a few discrepancies. Therefore, the genetic classification using the SSR markers could assist in strategic QPM breeding for tropical-highland and highland transition agro-ecologies. The outputs also form the basis for future studies aimed at confirming heterotic groups and identifying any new heterotic patterns that can emerge in the highland QPM germplasm.

Key words: Genetic diversity, Heterotic group, Highland maize, microsatellite markers, QPM

Introduction

Maize (*Zea mays* L., 2n=2x=20) is the principal grain crop in tropical environments, and globally ranks third after wheat and rice. It forms the basis for food security in many tropical countries including Ethiopia and other eastern and central African (ECA) countries (www.asareca.org; Atlin *et al*., 2011). However, it still needs improvement for adaptation to major agro-ecologies in order to meet the increasing gap between demand and supply of maize grain. Maize production is also increasingly important in the Tropical-highland regions of the world. The highland maize mega-environments include tropical highlands (2000-3600 m.a.s.l.), tropical highland transition zones (1500-2000 m.a.s.l), and temperate highlands (1000-2500 m.a.s.l.) (Bjarnason, 1994). In Ethiopia, the highland sub-humid agro-ecology is estimated to cover 20% of the land devoted to annual maize

cultivation (Twumasi-Afriyie *et al*., 2002).

In spite of its high potential for maize production, the highland zone in Ethiopia is generally characterized by high human population density and consequently high levels of malnutrition and poverty. Hence, alleviation of the nutritional problem of millions of people whose staple food is maize may depend upon the improvement of the inherent nutritional quality of the maize (Prasanna *et al*., 2001; Sofi *et al*., 2009). A nutritionally-enhanced quality protein maize germplasm doubles the two limiting amino acids (lysine and tryptophan) in maize. QPM contains 90% of the protein quality of case in milk compared with 40% for conventional maize (NRC, 1988).

Farmers in the highland zone generally grow unimproved maize varieties characterized by long maturity, vulnerability to frost attack, tall plant height and susceptible to lodging, which together contribute to low yield potential (Twumasi-Afriyie et al., 2002).This calls for the development of suitable maize cultivars to boost productivity in this agro-ecology. Most importantly, previous breeding efforts did not focus on QPM characterization and development because of challenges associated with QPM breeding despite its nutritional advantages (Atlin et al., 2011).Therefore, there is a need to efficiently characterize and develop highland QPM germplasm for breeding and genetic conservation. To this end, highland-adapted QPM inbred lines were developed in collaboration with CIMMYT-Ethiopia by converting the non-QPM counterparts following a backcross breeding procedure described by Vivek *et al*. (2008) using QPM donor lines sourced from CIMMYT. Results of grain sample analysis at CIMMYT-Mexico showed that the converted QPM inbred lines were phenotypically stable with protein levels of 8–14 g 100 g^{-1} grain and tryptophan levels of 0.65–0.90 g 100 g^{-1} protein (Twumasi-Afriyie *et al*., 2012).

Although better understanding on the genetic diversity of maize germplasm ensures the breeder in planning crosses for hybrid and line development, assigning lines to heterotic groups, and in plant variety protection, this aspect has not been sufficiently addressed specifically for the highland maize germplasm except a few reports (including- Liu *et al*., 2003; Beyene *et al*., 2005, 2006; Legesse *et al*., 2007, 2008). Understanding the genetic diversity and heterotic pattern among the existing genetic pool is important to exploit heterosis through proper choice of parents and well-designed crosses (Kalia *et al*., 2011).

Several researchers applied molecular markers such as microsatellites in their studies of QPM and non-QPM germplasm to determine genetic variation and relationships and for heterotic grouping (Reif *et al*., 2003; Bantte and Prasanna, 2003, Liu *et al*., 2003; Beyene *et al*., 2006; Legesse *et al*., 2007; Wu *et al*., 2010; Makumbi *et al*., 2011; Kalia *et al*., 2011; Wende *et al*., 2012; Kalyana-Babu *et al*., 2012). For germplasm characterization, simple sequence repeat (SSR) markers have been reported to provide much better information than single nucleotide polymorphisms (SNPs) (Hamblin *et al*., 2007). The present study was, therefore, conducted to assess the genetic variability and classes of elite QPM inbred lines developed for tropical-highlands and highland transition

maize agro-ecologies using SSR markers.

Materials and Methods

Plant materials

Thirty-six white-grained maize inbred lines comprised of 31highland adapted lines from the highland Maize Breeding, two mid-altitude adapted lines from Mid-Altitude Breeding Programs in Ethiopia, and three CIMMYT maize lines (CMLs) were used for this study (Table 1). All the 30 inbred lines are converted elite QPMs while six are non-QPM. The initial conversion and development of the highland QPM inbred lines was done according to the procedure described by Vivek *et al*. (2008). Briefly, a total of 51 highland-adapted non-QPM inbred lines from three heterotic groups (21 Ecuador, 10 Pool 9A, and 20 Kitale) were backcrossed with two CIMMYT derived QPM lines (CML144 and CML176). The non-QPM maize lines served as recurrent parents while the CIMMYT QPM lines were donors. CML144 (a tropically-adapted inbred line) was used as QPM donor for most lines in the Kitale Group, whereas CML176 (a sub-tropical adapted inbred line) was used for majority of the lines in Ecuador and Pool 9A heterotic groups. The BC_1F_1 and BC_2F_1 generations were selfed for five generations to develop the QPM inbred lines following rigorous field and light box selection in the laboratory. This facilitated selection for grain modification, good agronomic characters and resistance to the common highland maize diseases, such as common rust caused by *Puccinia sorghi* and northern corn leaf blight (*Exserohilum turcicum*). The non-QPM versions of the two mid-altitude QPM inbred lines (F7215Q and 142-1-eQ) are well-adapted tester lines used for the Mid-altitude and highland transition materials in Ethiopia.

DNA extraction

Seedlings of the 36 genotypes were raised in plastic seed trays for 3 weeks in a screen house at the Biosciences for east and central Africa (BecA) hub in Nairobi, Kenya. About equal leaf tissue from 10 plants per genotype was bulked, cut into pieces with scissors, and transferred into 1.2 ml strip tubes that contained one 4 mm stainless steel grinding balls. The tissue was freeze-dried (lyophilized) for three days using a Labconco freeze dryer (http://www.labconco.com). The lyophilized leaf samples were ground into fine powder using GenoGrinder-2000 at 500 strokes per minute for six minutes. Genomic DNA was extracted using a modified version of the mini-prep Cetyl-Trimethyl Ammonium Bromide (CTAB) method of CIMMYT protocol (http://www.generationcp.org/capcorner/chile_wksp_2005/manuals/manual_01.pdf). The quality of the isolated DNA was checked after running aliquots of DNA samples on a 0.8% agarose gel that contained 0.3 µg/ml GelRed (Biotium). DNA concentration was measured using Nano Drop ND-800 Spectrophotometer.

Table 1. Summary of the 36 QPM and non-QPM inbred lines used in the study.

S/N	Name	Pedigree	Source	QPM	Heterotic group*
1	142-1-eQ	Unknown (derived from Ecuador-573)	ETHIOPIA	QPM	Ecuador
2	CML144	Pob62c5HC182-2-1-2-B-B-3-1-#-#	CIMMYT	QPM	B
3	CML176	(P63-12-2-1/P67-5-1-1)-1-2-B-B	CIMMYT	QPM	Unknown
4	CML491	(6207QB/6207QA)-1-4-#-2-2-B-B	CIMMYT	QPM	A
5	F7215Q	Unknown (derived from Kitale Syn. II)	ETHIOPIA	QPM	Kitale
6	FS111	[POOL9Ac7-SR(BC2)]FS111-6-1-1-2-1-#/CML176BC1F1-8-1-2-1-1-#-#-#	CIMMYT	QPM	Ecuador
7	FS112	[POOL9Ac7-SR(BC2)]FS112-4-2-1-1-2-#/CML144(BC2)-25-8-2-1-3-1-#-#	CIMMYT	QPM	Unknown
8	FS151-3SR	[POOL9Ac7-SR(BC2)]FS151-3SR-1-2-1-1-#/CML176BC1F1-2-3-1-#-#	CIMMYT	QPM	Pool 9A
9	FS170N	[POOL9Ac7-SR(BC2)]FS170-2-1-3-2-2-1-#-#-#	CIMMYT	Non QPM	Unknown
10	FS170Q	[POOL9Ac7-SR(BC2)]FS170-2-1-3-1-#/CML176(BC2)-5-2-1-3-1-#-#	CIMMYT	QPM	Unknown
11	FS211-1SR	[POOL9Ac7-SR(BC2)]FS211-1SR-1-1-1-#/CML144(BC2)-14-21-1-3-2-1-#-#-#	CIMMYT	QPM	Kitale
12	FS232N	[POOL9Ac7-SR(BC2)]FS232-4-1-3-1-2-1-3-##-#-#-#-#	CIMMYT	Non QPM	Pool 9A
13	FS232Q	[POOL9Ac7-SR(BC2)]FS232-4-1-3-1-#/CML176(BC2)-17-1-1-1-#-#	CIMMYT	QPM	Pool 9A
14	FS2-3SR	[POOL9Ac7-SR(BC2)]FS2-3SR-2-1-2-#/CML176BC1F1-18-2-2-1-1-#-#-#	CIMMYT	QPM	Unknown
15	FS4-3SR	[POOL9Ac7-SR(BC2)]FS4-3SR-1-1-1-#/CML176(BC2)-8-2-1-1-1-#-#-#	CIMMYT	QPM	Unknown
16	FS45	[POOL9Ac7-SR(BC2)]FS45-3-2-2-1-#/CML144(BC2)-8-14-2-1-4-1-#-#-#	CIMMYT	QPM	Ecuador
17	FS48	[POOL9Ac7-SR(BC2)]FS48-1-1-1-1-1-#/CML144(BC2)-6-25-5-2-1-4-#	CIMMYT	QPM	Kitale
18	FS48-1SR	[POOL9Ac7-SR(BC2)]FS48-1SR-2-1-2-1-#/CML144(BC2)-7-4-1-3-2-1-#	CIMMYT	QPM	Kitale
19	FS59-2	[POOL9Ac7-SR(BC2)]FS59-2-2-1-1-#/CML144(BC2)-9-9-3-2-2-1-#	CIMMYT	QPM	Kitale
20	FS59-4N	[POOL9Ac7-SR(BC2)]FS59-4-1-2-1-1-#-B-B-B-B-B	CIMMYT	Non QPM	Ecuador
21	FS59-4Q	[POOL9Ac7-SR(BC2)]FS59-4-1-2-1-1-#/CML176BC1F1-3-2-3-#-#	CIMMYT	QPM	Ecuador
22	FS60	[POOL9Ac7-SR(BC2)]FS60-2-1-1-1-#/CML176BC1F1-5-3-1-2-1-#	CIMMYT	QPM	Pool 9A
23	FS67(BC1)	[POOL9Ac7-SR(BC2)]FS67-1-2-1-1-1-#/CML144(BC1)F1-11-1-2-2-2-#	CIMMYT	QPM	Kitale
24	FS67(BC2)	[POOL9Ac7-SR(BC2)]FS67-1-2-3-1-#/CML144(BC2)-10-11-2-4-1-2-#	CIMMYT	QPM	Kitale
25	FS67-N	[POOL9Ac7-SR(BC2)]FS67-1-2-3-1-#-B-B-B-B-B	CIMMYT	Non QPM	Kitale
26	FS68(BC1)	[POOL9Ac7-SR(BC2)]FS68-1-1-2-1-1/CML144(BC1)F1-1-1-2-1-1-#-#	CIMMYT	QPM	Kitale
27	FS68(BC2)	[POOL9Ac7-SR(BC2)]FS68-1-1-2-1-1/CML144(BC2)-33-1-1-1-#-#	CIMMYT	QPM	Kitale
28	KIT12	[KIT/SNSYN[N3/TUX]]c1F1-##(GLS=1)-12-2-1-#/CML176(BC2)-6-2-3-3-1-#-#-#	CIMMYT	QPM	Ecuador
29	KIT29	[KIT/SNSYN[N3/TUX]]c1F1-##(GLS=2)-29-35-2-3/CML144(BC2)-29-24-1-1-2-1-#-#-#	CIMMYT	QPM	Unknown
30	KIT31	[KIT/SNSYN[N3/TUX]]c1F1-##(GLS=1.5)-31-17-1-1/CML144(BC2)-31-14-1-1-1-2-#-#-#	CIMMYT	QPM	Unknown
31	KIT32N	[KIT/SNSYN[N3/TUX]]c1F1-##(GLS=2.5)-32-1-1-1-#-#-#-#-#-#-#-#	CIMMYT	Non QPM	Ecuador
32	KIT32Q	[KIT/SNSYN[N3/TUX]]c1F1-##(GLS=2.5)-32-1-1-#/CML176BC1F1-12-1-3-1-1-#-#-#	CIMMYT	QPM	Ecuador
33	KIT34	[KIT/SNSYN[N3/TUX]]c1F1-##(GLS=2.5)-34-2-1-#/CML176BC1F1-6-1-1-1-1-#	CIMMYT	QPM	Ecuador
34	SRSYN20N	SRSYN95[ECU//SC/ETO]F1-##(GLS=3.5)-20-2-1-1-#-#-#-#-#-#	CIMMYT	Non QPM	Pool 9A
35	SRSYN20Q	SRSYN95[ECU//SC/ETO]F1-##(GLS=3.5)-20-2-1-#/CML176(BC2)-4-2-2-3-2-#-#-#	CIMMYT	QPM	Pool 9A
36	SRSYN48	SRSYN95[ECU//SC/ETO]F1-##(GLS=3.5)-48-1-1-#/CML176(BC2)-11-2-1-1-1-#-#	CIMMYT	QPM	Ecuador

*= Putative heterotic grouping based on phenotypic data of the non-QPM counterparts before conversion to QPM

Genotyping

Thirty SSR markers were used initially for this study. The microsatellites were chosen based on prior information, including chromosomal distribution, minor allele frequency (MAF), polymorphic information content (PIC), and repeat length. Polymerase chain reaction (PCR) was performed in 96-wells plates in a total reaction volume of 10 µl that consisted of 50 ng template DNA, 1× magnesium-free PCR buffer, 2.3 mM $MgCl_2$, 0.20 µM of the forward primer labeled either with 6-FAM, PET, VIC or NED fluorescent dyes, 0.20 µM of the reverse primers, 0.20 mM each dNTP, and 0.375 unit Taq DNA polymerase. PCR amplifications were performed for each primer-pair separately using Gene-Amp PCR System 9600 (PE-Applied Biosystems) according to the following protocol: 3 min initial denaturation at 94 ^{o}C, followed by 35 cycles of 94 ^{o}C for 30 sec, one min annealing at 52 to 60 ^{o}C (depending on the recommended temperature for the primer) and 72 ^{o}C for two min, and a final extension of 10 min at 72 ^{o}C. For high throughput and low cost genotyping, PCR products were separated by pooling 1.2 µl of the PCR products from each of the 6-FAM, VIC, PET and NED-labeled markers and 9 µL of a mix of an injection solution (HiDi) and GS-500LIZ size standard (1 mL HiDi and 12 µl GS500 LIZ for a 96-well plate). DNA fragments were denatured for three minutes at 94 ^{o}C and size-fractioned using ABI 3730 Capillary DNA Sequencer with GeneScan software. Allele sizes were called using GeneMapper version 4 software (©2005 Applied Biosystems). Both DNA extraction and SSR genotyping were done at the BecA hub.

Data Analysis

Multivariate analyses were performed on the SSR data of the 26 markers (Table 2), since the data of the remaining four markers (phi031, phi041, phi112 and phi227562) were excluded because they showed ambiguous allele calls with high proportion of heterozygosity (markers that were not homozygous), high missing data points or they were monomorphic. For each SSR marker, number of alleles per marker, allele frequency, observed heterozygosity, and polymorphic information content (PIC) were computed using Power-Marker version 3.25 (Liu and Muse, 2005). Polymorphism information content (PIC), which indicates how informative each SSR marker was estimated following the expression applied by Smith et al. (1997): PIC= $1-\sum f_i^2$, where f_i is the frequency of the i^{th} allele. Estimates of genetic similarity among all pairs of inbred lines were calculated from the data matrices in the form of dissimilarity units and expressed as Euclidean genetic distance (Liu and Muse, 2005) using Power Marker software. SAHN - Sequential agglomerative hierarchical nested cluster analysis method (Sneath and Sokal, 1973) of the NTSYS-pc 2.10 software (Rohlf, 2000) was used to compute Euclidean distance matrix, and cluster the distance matrix in the form of dendrogram by employing UPGMA (Unweighted Pair Group Method based on Arithmetic Averages) with average linkage (Sneath and Sokal, 1973). Finally, 'cophenetic correlation coefficient' was estimated using COPH and MXCOMP options of the NTSYS-pc 2.10 to test the goodness of fit of the cluster analysis to the dissimilarity matrix on which it was based. The Mantle's (1967) two-way test was used for this purpose.

Results

Analysis of SSR Polymorphism

As suggested by several authors, in choosing marker loci for diversity study, it is important to consider representativeness of a particular locus in the entire genome to avoid biases due to sampling. Accordingly, there were two to four pairs of primers corresponding to each chromosome except chromosome nine that had only single primer pair per locus. Among the 26 pairs of SSR primers evenly distributed on the 10 chromosomes, 25 pairs produced polymorphic products while only one (phi062) was monomorphic. This is equivalent to 96% polymorphic loci. The number of alleles scored for each marker varied from 2 in phi084 and umc2250 to 8 in umc1161 excluding the monomorphic marker. The 26 markers amplified a total of 99 alleles, with an average of 3.80 alleles per marker. Major allele frequency was the lowest (0.32) in phi299852 and highest (0.88) in umc1367 (excluding the monomorphic marker), and the overall average was 0.59. The polymorphism information content ranged from 0.22 (umc1367) to 0.74 (phi299852) and the overall average was 0.49 excluding phi062. Major allele frequency showed strong negative correlation ($r = -0.91$; $p < 0.001$) with polymorphic information content (Table 2).

Table 2. Summary of characterization for the 26 SSR markers used in the present study.

SSR locus	Bin no.	Repeat motif	Allele Frequency	No. of Alleles	Gene Diversity	*Hetero-zygosity	PIC	$Ann. Temp
nc130	5.00	AGC	0.67	3	0.48	0.00	0.40	54
nc133	2.05	GTGTC	0.66	3	0.51	0.00	0.45	54
phi029	3.04	AGCG	0.56	3	0.52	0.03	0.41	56
phi046	3.08	ACGC	0.53	3	0.52	0.00	0.41	60
phi056	1.01	CCG	0.44	4	0.69	0.03	0.63	56
phi062	10.04	ACG	1.00	1	0.00	0.00	0.00	56
phi065	9.03	CACTT	0.39	4	0.67	0.05	0.60	54
phi072	4.01	AAAC	0.69	4	0.46	0.05	0.40	56
phi075	6.00	CT	0.76	3	0.39	0.02	0.35	54
phi076	4.11	GAGCGG	0.40	6	0.71	0.14	0.66	60
phi079	4.05	AGATG	0.38	5	0.73	0.02	0.69	60
phi084	10.04	GAA	0.67	2	0.44	0.05	0.35	54
phi102228	3.06	AAGC	0.78	3	0.37	0.00	0.34	54
phi114	7.02	GCCT	0.48	4	0.60	0.00	0.52	60
phi123	6.07	AAAG	0.58	3	0.57	0.00	0.50	54
phi299852	6.07	AGC	0.32	7	0.77	0.02	0.74	58
phi308707	1.01	AGC	0.47	3	0.62	0.00	0.54	56
phi331888	5.04	AAG	0.54	4	0.59	0.02	0.51	58
phi374118	3.02	ACC	0.58	4	0.59	0.00	0.54	54
phi96100	2.10	ACCT	0.40	4	0.71	0.08	0.66	56
umc1161	8.06	GCTGGG	0.59	8	0.61	0.09	0.58	56
umc1304	8.02	TCGA	0.61	3	0.48	0.14	0.38	54
umc1367	10.03	CGA	0.88	4	0.23	0.13	0.22	62
umc1545	7.00	AAGA	0.69	5	0.48	0.00	0.42	54
umc1917	1.03	CTG	0.64	4	0.54	0.03	0.50	52
umc2250	2.00	ACG	0.50	2	0.50	1.00	0.38	58
Mean			0.59	3.8	0.53	0.07	0.49	

* = Observed heterozygosity of alleles at a specific locus; $ = Annealing temperature in ^{o}C.

Genetic distance and cluster analysis

The pairwise genetic dissimilarity values based on Euclidean distance measures among all the 36 inbred lines ranged from 1.10 between the inbred pairs FS211-1SR and FS111 to 0.11 between inbred lines CML144 and CML491with overall mean of 0.74 (Table 3). Out of the 41 values recorded as long genetic distances in the pairwise comparisons, 18 (44%) were Euclidian genetic distance measures ≥ 1.00. It was also interesting to note that the second minimum genetic distance (0.25) was found between FS67-N and FS67(BC1) when relatively the maximum distance was recorded between the other version [FS67(BC2)] and FS67-N. High genetic distances were also recorded between inbred lines which were same prior to the conversion into QPM, *viz*., KIT32N and KIT32Q as well as FS170N and FS170Q. In contrast, FS59-4N and FS59-4Q had relatively short distances.

The results of relationships based on the genetic distances among inbreds were consistent with the dendrogram (Figure 1) constructed by UPGMA cluster. The result of the UPGMA clustering algorithm was able to broadly group the 36 inbred lines into three major clusters designated as Group 1, 2 and 3. The cophonetic correlation coefficient computed as a measure of goodness of fit was 0.78. Concurring with the cluster analysis, the genetic distances among genotypes from different groups were higher than that within same group. For e.g., inbred pairs CML144 and CML491 as well as FS67-N and FS67(BC1) were observed among the most closely related genotypes as they had relatively small genetic distance values and thus classified under the same cluster Group 1. On the contrary, inbred pairs FS211-1SR and FS111 as well as KIT12 and FS170Q were the most distantly related genotypes among others. This was substantiated both by maximum genetic distances and differences in their categories such that FS211-1SR and KIT12 were from cluster Group 2 while FS111 and FS170Q were found in the same cluster Group 3. In general, considering individual genetic distance values among pairs of inbred lines in the different cluster groups, it was found out that maximum distances were identified more often between cluster Groups 1 and 2 than others.

Discussion

Liu et al. (2003) reported that the diversity in tropical-highland maize was poorly represented among available inbred lines, suggesting that these germplasm could be tapped to identify new alleles of agronomic importance. In the present study, a total of 99 alleles with an average of 3.8 per locus were detected in 30 QPM and 6 non-QPM inbred lines using 26 SSRs. The average number of alleles found in this study was on par with the reports of Beyene *et al*. (2006), Legesse *et al*. (2007), Dhliwayo *et al*. (2009), and Makumbi *et al*. (2011). From the work of Beyene *et al*. (2006), a total of 98 alleles and an average of 4.9 alleles per locus were detected among 62 traditional Ethiopian highland maize accessions using 20 SSR markers. Similarly, Legesse *et al*. (2007) reported that 27 SSR primer sets identified 104 alleles with mean of 3.85 alleles among 56 highland and mid-altitude maize lines of CIMMYT origin. The mean number of alleles per locus found in this study was also within the range of other studies in QPM germplasm. For example, Bantte and Prasanna (2003) found mean of 3.25 alleles using 36 SSR

Table 3. Euclidian genetic distance matrix among selected 30 QPM and six non-QPM inbred lines analyzed using 26 SSR markers.

*Line ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1	0.00																																		
2	0.67	0.00																																	
3	0.62	0.86	0.00																																
4	0.44	0.68	0.61	0.00																															
5	0.68	0.81	0.11	0.68	0.00																														
6	0.88	0.74	0.76	0.61	0.73	0.00																													
7	0.61	0.76	0.51	0.92	0.57	0.87	0.00																												
8	0.62	0.46	0.82	0.68	0.76	0.96	0.96	0.00																											
9	0.43	0.70	0.54	0.40	0.59	0.81	0.80	0.59	0.00																										
10	0.56	0.54	0.79	0.77	0.73	0.77	0.93	0.63	0.68	0.00																									
11	0.80	0.75	0.71	0.86	0.76	0.75	0.74	0.82	0.67	0.63	0.00																								
12	0.85	0.84	1.00	0.70	1.05	0.89	1.10	0.76	0.94	0.95	0.89	0.00																							
13	0.74	0.84	0.85	0.77	0.91	0.72	0.94	0.74	0.74	0.71	0.57	0.76	0.00																						
14	0.52	0.51	0.69	0.45	0.81	0.82	0.76	0.59	0.48	0.73	0.86	0.67	0.82	0.00																					
15	0.61	0.86	0.73	0.83	0.81	0.89	0.65	0.75	0.57	0.87	0.73	0.96	0.71	0.81	0.00																				
16	0.93	0.84	0.71	0.80	0.65	0.78	0.68	0.92	0.73	0.84	0.60	0.94	0.79	0.86	0.86	0.00																			
17	0.61	0.78	0.73	0.40	0.84	0.68	0.82	0.84	0.59	0.87	0.79	0.72	0.76	0.47	0.87	0.57	0.00																		
18	0.64	0.65	0.76	0.49	0.82	0.76	0.68	0.82	0.56	0.95	0.87	0.78	0.85	0.54	0.57	0.82	0.57	0.00																	
19	0.56	0.41	0.71	0.61	0.82	0.82	0.62	0.60	0.50	0.68	0.82	0.71	0.79	0.28	0.68	0.82	0.63	0.44	0.00																
20	0.50	0.59	0.89	0.60	0.87	0.75	0.79	0.67	0.60	0.68	0.82	0.62	0.57	0.44	0.82	0.78	0.70	0.53	0.40	0.00															
21	0.69	0.54	0.99	0.80	0.94	0.69	0.78	0.70	0.76	0.71	0.83	0.94	0.75	0.85	0.57	0.70	0.78	0.67	0.73	0.70	0.00														
22	0.56	0.54	0.65	0.61	0.60	0.67	0.74	0.65	0.59	0.79	0.87	1.00	0.91	0.73	0.54	0.82	0.79	0.60	0.71	0.84	0.48	0.00													
23	0.61	0.45	0.92	0.77	0.82	0.86	0.59	0.73	0.76	0.70	0.87	1.04	0.87	0.76	0.81	0.89	0.86	0.71	0.56	0.56	0.72	0.82	0.00												
24	0.65	0.71	0.79	0.51	0.68	0.47	0.77	0.79	0.62	0.82	0.91	0.96	0.65	0.68	0.73	0.74	0.65	0.51	0.57	0.40	0.54	0.62	0.64	0.00											
25	0.61	0.49	0.97	0.77	0.97	0.90	0.67	0.68	0.65	0.82	0.75	0.81	0.82	0.64	0.80	0.81	0.65	0.63	0.52	0.52	0.65	0.75	0.54	0.65	0.00										
26	0.71	0.54	0.79	0.71	0.68	0.70	0.65	0.73	0.74	0.71	0.84	1.01	0.71	0.68	0.84	0.79	0.87	0.63	0.52	0.41	0.71	0.73	0.48	0.25	0.65	0.00									
27	0.76	0.70	0.87	0.37	0.92	0.73	1.02	0.60	0.76	0.95	0.98	0.75	0.91	0.76	0.84	0.92	0.57	0.60	0.71	0.80	0.75	0.71	0.78	0.68	0.68	0.90	0.00								
28	0.58	0.60	0.92	0.46	0.87	0.88	0.87	0.65	0.54	0.79	0.89	0.77	0.81	0.58	0.81	0.78	0.54	0.54	0.56	0.53	0.71	0.82	0.57	0.59	0.54	0.75	0.51	0.00							
29	0.61	0.90	0.62	0.90	0.68	0.96	0.77	0.91	0.85	0.88	1.07	0.89	0.82	0.99	0.90	1.02	0.88	1.02	0.85	0.98	0.90	0.79	0.79	0.88	0.82	0.93	0.91	1.02	0.00						
30	0.71	0.49	0.73	0.65	0.68	0.99	0.76	0.41	0.57	0.65	0.73	0.92	0.71	0.59	0.67	0.73	0.92	0.68	0.57	0.65	0.64	0.73	0.70	0.71	0.76	0.54	0.84	0.65	1.05	0.00					
31	0.58	0.67	0.68	0.37	0.68	0.91	0.85	0.62	0.62	0.71	0.68	0.75	0.88	0.67	0.90	0.79	0.71	0.62	0.74	0.83	0.97	0.74	0.79	0.82	0.88	0.76	0.62	0.62	0.82	0.54	0.00				
32	0.76	0.83	0.74	0.64	0.80	0.61	0.71	0.86	0.76	0.95	0.80	0.70	0.86	0.61	0.61	0.68	0.52	0.55	0.68	0.70	0.73	0.61	1.07	0.58	0.79	0.83	0.74	0.76	1.03	0.83	0.71	0.00			
33	0.32	0.78	0.52	0.55	0.52	0.75	0.57	0.84	0.57	0.76	0.95	1.05	0.88	0.76	0.67	0.97	0.73	0.70	0.73	0.75	0.64	0.57	0.70	0.59	0.78	0.71	0.75	0.75	0.54	0.87	0.85	0.85	0.00		
34	0.47	0.70	0.67	0.37	0.67	0.64	0.72	0.76	0.65	0.84	0.92	0.84	0.88	0.71	0.63	0.98	0.73	0.51	0.71	0.75	0.67	0.44	0.78	0.53	0.92	0.75	0.51	0.73	0.81	0.73	0.57	0.68	0.43	0.00	
35	0.62	0.70	0.65	0.74	0.60	0.89	0.57	0.71	0.59	0.84	0.87	1.05	0.91	0.69	0.75	0.76	0.84	0.87	0.76	0.84	0.99	0.65	0.65	0.91	0.86	0.79	1.09	0.76	0.85	0.68	0.74	0.98	0.73	0.78	0.00
36	0.70	0.92	0.76	0.61	0.71	1.02	0.96	0.76	0.73	0.90	0.92	0.80	0.91	0.86	0.75	0.92	0.90	0.71	0.82	0.78	0.77	0.76	0.82	0.79	0.95	0.84	0.82	0.65	0.91	0.68	0.74	0.92	0.75	0.60	0.82

*1= KIT32N, 2= 142-1-eQ, 3= CML144, 4= CML176, 5= CML491, 6= F7215Q, 7= FS111, 8= FS112, 9= FS151-3SR, 10= FS170N, 11= FS170Q, 12= FS211-1SR, 13= FS232N, 14= FS232Q, 15= FS2-3SR, 16= FS4-3SR, 17= FS45, 18= FS48, 19= FS48-1SR, 20= FS59-2, 21= FS59-4N, 22= FS59-4Q, 23= FS60, 24= FS67(BC1), 25= FS67(BC2), 26= FS67-N, 27= FS68(BC1), 28= FS68(BC2), 29= KIT12, 30= KIT29, 31= KIT31, 32= KIT32Q, 33= KIT34, 34= SRSYN20N, 35= SRSYN20Q, and 36= SRSYN48

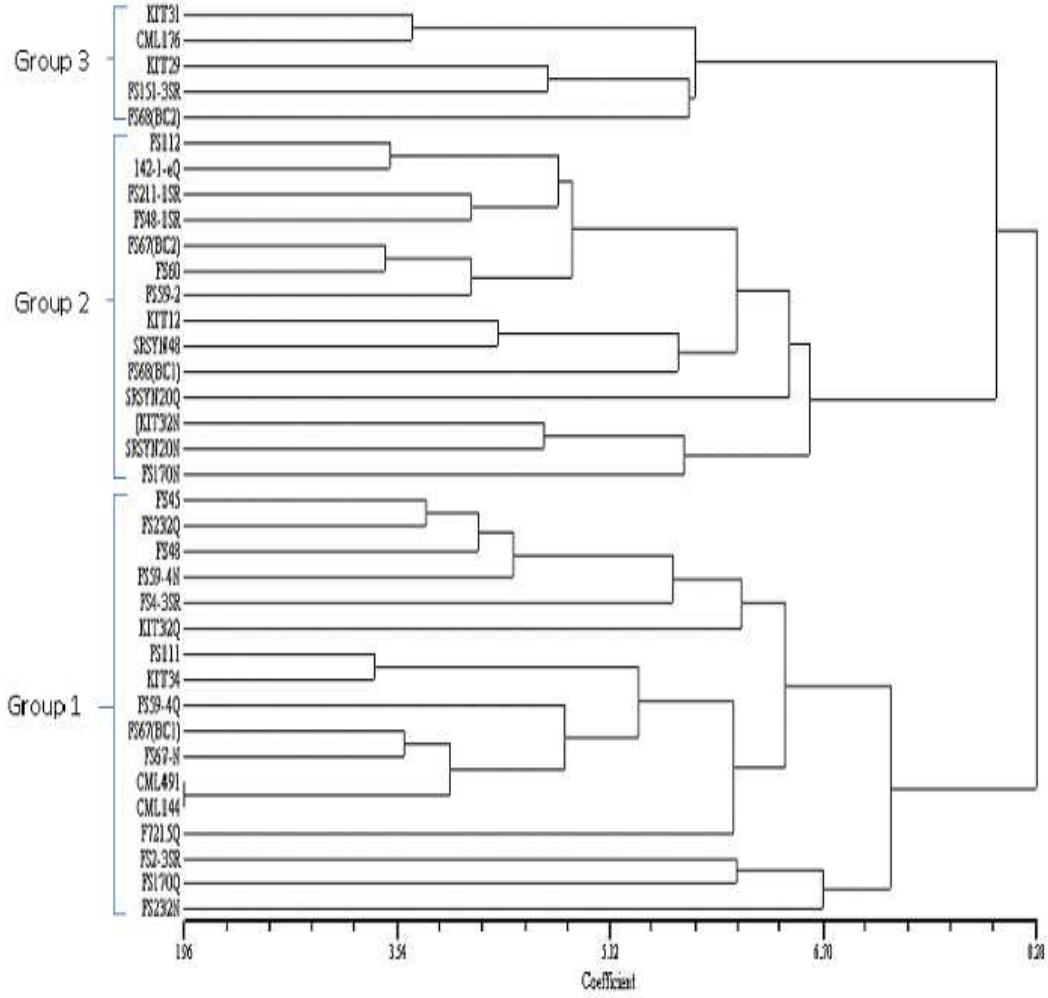

Figure 1. Dendrogram showing three cluster groups using UPGMA cluster analysis based on 26 SSR genetic distance estimates among 30 QPM and 6 non-QPM inbred lines.

loci and 23 QPM inbred lines; Babu *et al*. (2009) reported mean of 2.42 alleles using 47 SSR loci and 19 QPM inbred lines; Kalyana-Babu *et al*. (2012) obtained mean of 3.44 alleles using 75 SSR loci and 48 QPM and normal inbreds; and Krishna *et al*. (2012) had mean of 4.08 alleles using 48 SSR loci and 63 QPM inbred lines. The slight variations in mean numbers of alleles among germplasm could be attributed to differences in sample sizes, type of germplasm, and repeat types of SSRs used (Rajab *et al*., 2006). Contrary to the above and present studies, high average numbers of alleles per locus were previously reported by Yao *et al*. (2008), Liu *et al*. (2005), and Wu *et al*. (2004) for 124 maize landraces (6.4), 38 waxy-maize landraces (4.1), and popcorn landraces (5.4), respectively. As expected, such disparity suggested a wider genetic diversity of maize landraces than inbred lines.

PIC provides an estimate of how informative is a particular marker by considering both the number of alleles that are expressed and the relative frequencies of those alleles (Smith *et al*., 1997). In the present study, PIC values ranged from 0.22 (less discriminative primer-*umc1367*) to 0.74 (highly discriminative primer- *phi299852*) with a mean of 0.49. These values agreed with the report of Dhliwayo *et al*. (2009) who estimated PIC values ranging from 0.00 to 0.77 and mean of 0.43. Other findings reported different mean PIC values such as 0.60 by Kamalesh *et al*. (2009) and 0.91 by Krishna *et al*. (2012). The relatively small PIC value in the present study could primarily be explained by either the presence of few di-nucleotide repeats as opposed to more di-nucleotides (Smith *et al*., 1997; Bantte and Prasanna, 2003; Kalyana-Babu *et al*., 2012) or lower genetic variability among the germplasm used for the study (Dhliwayo *et al*., 2009). A significant positive correlation (r= 0.69) was also observed between PIC values and number of alleles amplified per primer.

On the other hand, the mean observed heterozygosity of the inbred lines (7%) was also found in an acceptable range for running the marker data analysis. The inbreds used in the current study are advanced generations and should have high level of homozygosity and as such the residual heterozygosity (Bantte and Prasanna, 2003; Demissew *et al*., 2012) could most likely be the cause for the observed heterozygosity per locus. The other reason could also be attributed to human errors during maintenance breeding (e.g. seed admixture, pollen contamination, mislabeling of seed sources and mixing of different seed stocks for planting).

The genetic distance values with mean of 0.74 determined in this study is closer to the mean Euclidean distance (0.59) among Highland and mid-altitude non-QPM inbreds reported by Legesse *et al*. (2007). The dendrogram generated from the Euclidean genetic distance matrix categorized the 36 lines into three groups. The clustering revealed an evidence of associations for part of the inbred lines to their putative heterotic groups previously suggested using combining ability analyses of phenotypic data. The failure of association of some of the inbreds to their putative heterotic groups was that when the backcross procedure of incorporation of *opaque-2* gene was carried out initially, there was no stringent selection to recover characters of the recurrent parents. The process of conversion had been carried out using pedigree selection mainly focusing on the development of new and adapted highland QPM lines from two different backcross generations (Twumasi-Afriyie *et al*., 2002).

It was also observed that there were 17, 14 and 5 inbred lines in cluster groups one, two and three, respectively. Cluster Group I was dominated by six lines from Ecuador heterotic group, four from Kitale group, two from Pool 9A group, three from previously uncategorized lines, and two CMLs (CML144 and CML491). Out of 17 lines in Group 1, eight of them were converted to QPM using CML176 as donor while only three lines out of 17 were converted to QPM using CML144 as donor. Besides, three lines in Group I were non-QPM counterparts. A mid-altitude line (F7215Q) which was converted into QPM using CML159 as donor parent was also found in this group. As opposed to this grouping, Dagne (2008) assigned CML144 and

CML491 in two different groups. Similarly, cluster Group II was dominated by five lines extracted from Kitale heterotic group, four from Ecuador, four Pool9A, and one previously uncategorized line. Six lines in Group II were converted to QPM using CML144. Whereas five lines were converted using CML176 and the remaining three lines were again non-QPM counterparts. The other mid-altitude line (142-1eQ) which was converted into QPM using CML176 as donor parent was also found in this group. As regards cluster Group III, two previously uncategorized lines with CML144 being used for their conversion to QPM, one from Kitale with CML144 again used as QPM donor, one from Pool9A where CML176 was a QPM donor, and CML176 itself were all included in this group. However, in the report of Bantte and Prasanna (2003), it was noted that CML176 and CML144 were categorized together into one cluster group. Such incongruities with the results of other investigators of assigning inbred lines into heterotic groups may occur due to seed handling or pollination errors (Rajab *et al*., 2006). This may also be caused by differential selection of the different lines in different environments, genetic drift and mutation (Senior *et al*., 1998). The outputs of the present study, therefore, could form the basis for future studies aimed at confirming heterotic groups and identifying any new heterotic groups that can emerge in the highland QPM germplasm.

Conclusion

The 30 QPM and six non-QPM advanced inbred lines adapted to the highland environments were successfully genotyped using the 26 SSR markers. The markers were also effective in classifying the diverse and closely related inbred lines into genetic groups. However, some lines from different genetic backgrounds were fitted in the same cluster. Two reasons could be suggested for this discrepancy. Firstly, the population testers initially used to classify the non-QPM lines based on phenotypic data had not been good testers in discriminating the lines into their respective heterotic groups. Secondly, due attention was not paid to the heterotic system during selection of lines for use as donors to convert the non-QPM lines into QPM. Therefore, it was concluded that the SSR classification information generated in this study can serve as a platform for detailed characterization and heterotic grouping with large number of QPM genotypes followed by making crosses and field evaluations for confirmation. Besides, increased allelic diversity among the lines may provide high degree of hetroszygosity in the hybrid combination which may translate to increased heterosis.

Acknowledgments

The authors thank the Generation Challenge Program (GCP) for funding this work. The protein quality analysis of the inbred lines was financed by CIMMYT-Ethiopia.

References

Atlin, G.N., N. Palacios, R. Babu, B. Das, S. Twumasi-Afriye, D.K. Friesen, H. De Groote, B. Vivek and K.V. Pixley. 2011. Quality protein maize: Progress and prospects. *Plant Breed. Rev.* 34:83–129.

Bantte, K. and B.M. Prasanna. 2003. Simple sequence repeat polymorphism in Quality Protein

Maize (QPM) lines. *Euphytica* 129:337-344.

Beyene, Y., A.M. Botha and A.A. Myburg. 2005. Phenotypic diversity for morphological and agronomic traits in traditional Ethiopian highland maize accessions. *S. Afr. J. Plant Soil* 22:100–105.

Beyene, Y., A.M. Botha and A.A. Myburg. 2006. Genetic diversity among traditional Ethiopian highland maize accessions assessed by simple sequence repeat (SSR) markers. *Genet. Resour. Crop Evol.* 53:1579–1588.

Bjarnason, M. (ed.). 1994. *The Subtropical, Midaltitude, and Highland Maize Subprogram.* Maize Program Special Report. Mexico, D.F.: CIMMYT.

Dagne, W.G. 2008.Genotypic variability and combining ability of quality protein maize inbred lines under stress and optimal conditions. PhD Thesis. University of the Free State, South Africa.

Demissew, A., G. Watson, H. Shimelis, J. Derera and S. Twumasi-Afriyie. 2012. Comparison of two PCR-based DNA markers with High Resolution Melt Analysis for the detection of genetic variability in selected quality protein maize inbred lines. *Afr. J. Agric. Res.* 742: 5692-5700.

Dhliwayo, T., K. Pixley, A. Menkir and M. Warburton. 2009. Combining ability, genetic distances and heterosis among elite CIMMYT and IITA tropical maize inbred lines. *Crop Sci*. 49:1201-1210.

Hamblin, M.T., M.L. Warburton and E.S. Buckler. 2007. Empirical comparison of Simple Sequence Repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. *PLoS One* 2:e1367.

Kalia, R.K., M.K. Rai, S. Kalia, R. Singh and A.K. Dhawan. 2011. Microsatellite markers: an overview of the recent progress in plants. *Euphytica* 177:309–334.

Kalyana-Babu, B., P. Pooja, J.C. Bhatt and P.K. Agrawal. 2012. Characterization of Indian and exotic quality protein maize (QPM) and normal maize (*Zea mays* L.) inbreds using simple sequence repeat (SSR) markers. *Afr. J. Biotechnol.* 11:9691-9700.

Kalyana-Babu, B., P.K. Agrawal, Vinay Mahajan and H.S. Gupta. 2009.Molecular and biochemical characterization of short duration quality protein maize (QPM). *J. Plant Biochem. Biot.* 18: 93-96.

Kamalesh, S.M., P.K. Agrawal, B. Kalyana-Babu and H.S. Gupta. 2009. Assessment of Genetic Diversity among the Elite Maize (*Zea mays* L.) Genotypes from North-Western Himalayan Region of India using Microsatellite Markers. *J. Plant Biochem. Biot.* 18: 217-220.

Krishna, M.S.R., R.S. Sokka and V.C.B. Naik. 2012. Assessment of genetic diversity in quality protein maize (QPM) lines using simple sequence repeat (SSR) markers. *Afr. J. Biotechnol.* 11:16427-16433.

Legesse, B.W., A.A. Myburg, K.V. Pixley and A.M. Botha. 2007. Genetic diversity of African maize inbred lines revealed by SSR markers. *Hereditas* 144:10-17.

Legesse, B.W., A.A. Myburg, K.V. Pixley, S. Twumasi-Afriyie and A.M. Botha. 2008. Relationship between hybrid performance and AFLP based genetic distance in highland maize inbred lines. *Euphytica* 162:313–323.

Liu, K. and S.V. Muse. 2005.PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21:2128-2129.

Liu, K., M. Goodman, S. Muse, J.S. Smith, Ed. Buckler and J. Doebley. 2003. Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics* 165: 2117–2128.

Liu, Y.J., Y.B. Huang, T.Z. Rong, M.L. Tian and J.P. Yang. 2005. Comparative analysis of genetic diversity in land-races of waxy maize from Yunnan and Guizhou using SSR markers. *Sci Agric Sinica* 4:648–653.

Makumbi, D., J.F. Betran, M. Banziger and J.M. Ribaut. 2011.Combining ability, heterosis and genetic diversity in tropical maize (Zea mays L.) under stress and non-stress conditions. *Euphytica* 180:143–162.

Mantel, N.A. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27:209-220.

National Research Council (NRC). 1988. Quality protein maize. National Academic Press, Washington, DC.

Prasanna, B.M., S.K. Vasal, B. Kassahun and N.N. Singh. 2001. Quality Protein Maize. *Current Sci.* 81:1308-1319.

Rajab, C.R., A. Hossainzadeh, R.M. Ghannadha, M.L. Warburton, A.R. Talei and A.S. Mohammadi. 2006. Use of SSR data to determine relationships and potential heterotic groupings within medium to late maturing Iranian maize inbred lines. *Field Crop Res*. 95:212–222.

Reif, J.C., A.E. Melchinger, X.C. Xia, M.L. Warburton, D.A. Hoisington, S.K. Vasal, *et al*. 2003. Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. *Crop Sci.* 43:1275-1282.

Rohlf, F.J. 2000. NTSYS-pc: *Numerical Taxonomy and Multivariate Analysis System.* ver. 2.1, Exeter Software. Satauket, New York, USA.

Senior, M.L., J.P. Murphy, M.M. Goodman and C.W. Stuber. 1998. Utility SSRs for determining genetic similarities and relationships in maize using an agarose gel system. *Crop Sci.* 38:1088-1098.

Smith, J.S.C. and E.C. Chin. 1997. An evaluation of the 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.

Sneath, P.H.A. and R.R. Sokal. 1973. *Numerical taxonomy*. Freeman, San Francisco.

Sofi, P.A., S.A. Wani, A.G. Rather and S.H. Wani. 2009. Quality protein maize (QPM): Genetic manipulation for the nutritional fortification of maize. *J. Plant Breed. Crop Sci.* 1:244-253.

Twumasi-Afriyie, S., H. Zelleke, K. Yihun, B. Assefa and S. Tariku. 2002. Development and Improvement of Highland Maize in Ethiopia. pp. 31-38. *In:* M. Nigusse and D. Tanner (eds.). Proceedings of the Second National Maize Workshop of Ethiopia, Addis Ababa, Ethiopia. 12-16 November 2001. EARO (Ethiopian Agricultural Research Organization) and CIMMYT, Addis Ababa, Ethiopia.

Twumasi-Afriyie, S., A.K. Demissew, B. Gezahegn, A. Wende, N. Gudeta, N. Demoz, D. Friesen, Y. Kassa, A. Bayisa, A. Girum and F. Wondimu.

2012. A Decade of Quality Protein Maize Research Progress in Ethiopia (2001–2011). pp. 47-57. *In*: M. Worku *et al*. (Eds.). *Meeting the Challenges of Global Climate Change and Food Security through Innovative Maize Research*. *Proceedings of the Third National Maize Workshop of Ethiopia.*Addis Ababa, Ethiopia. 18-20 April 2011. Ethiopian Institute of Agricultural Research (EIAR) and CIMMYT, Addis Ababa, Ethiopia.

Vivek, B.S., A.F. Krivanek, N. Palacios-Rojas, S. Twumasi-Afriyie and A.O. Diallo. 2008. *Breeding quality protein maize (QPM): Protocols for developing QPM cultivars.* CIMMYT, Mexico D.F.

Wende, A., H. Shimelis, J. Derera, J. Danson and M.D. Laing. 2012. Genetic interrelationships among medium to late maturing tropical maize inbred lines using selected SSR markers. *Euphytica*: DOI 10.1007/s10681-012-0826-6.

Wu, C.L., S.F. Li, B.X. Dong, Q.Q. Zhang and C.Q. Zhang. 2010. Determination of the number of SSR alleles necessary for the analysis of genetic relationships between maize inbred lines. *Agr. Sci. China* 9:1713-1725.

Wu, Y.S., Y.L. Zheng, R. Sun, S.Y. Wu, H.B. Gu and Y.H. Bi. 2004. Genetic diversity of waxy corn and popcorn land-races in Yunnan by SSR markers. *Acta Agronomica Sinica* 30:36–42.

Yao, Q.L., P. Fang, K.C. Kang and G.T. Pan. 2008. Genetic diversity based on SSR markers in maize (*Zea mays* L.) landraces from Wuling mountain region in China. *J. Genet.* 87:287-291.