

Research Article

Validation of parents and estimation of molecular diversity through SSR markers in maize (*Zea mays* L.)

Dhairyashil M. Langade, J. P. Shahi, K. Srivastava and Amita Sharma

Department of Genetics and Plant Breeding, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi - 221005

E-mail: dhirajlangade@gmail.com

(Received: 15 Nov 2016; Revised: 10 Dec 2017; Accepted: 20 Dec 2017)

Abstract

India is the third largest consumer of edible oils; an increasing human population increases the demand of oil at an alarming rate. Therefore, there is a need to search new sources of oil in India to fulfill this goal. Molecular markers have facilitated the easy identification of diverse parents for development of superior high oil lines for commercial use. The parental material in the present experiment comprised of ten diverse inbred lines viz., high oil (HIGH OIL, DMHOC-09, DMHOC-15), QPM or intermediate oil (HUZQPM-01, HUZQPM-03, HUZQPM-05, HUZQPM-06) and low oil (HUZM-53, HUZM-265, HUZM-478) content containing lines were selected on the basis of yield, yield traits along with oil content at the Agricultural Research Farm, B.H.U., during *Kharif* 2011 and were subjected to validate the presence of QTL's linked to oil content and molecular diversity through SSR markers. A total number of 25 polymorphic SSR markers were used to evaluate molecular diversity among parental inbreds. The PIC of 25 SSR primers ranged from 0.42 to 0.50 with an average PIC of 0.46. The molecular validation studies revealed that SSR markers *bnlg2086*, *phi96100*, *umc1422*, *bnlg1452*, *bnlg2244*, *bnlg1711*, *umc2319*, *umc1360*, *phi119* and *phi050* were linked to QTLs for oil content. Therefore, these markers could be used in MAS for estimating high oil content in maize.

Key words

Maize, molecular diversity, SSR markers, validation

Introduction

Maize (*Zea mays* L.) is a versatile and most widely distributed crop of the world being grown in tropical, sub-tropical and temperate regions up to 50°C to more than 3000 m from sea level, under irrigated to semi-arid conditions. It is the highest productive crop among cereals; hence called as 'queen of cereals'. It is one of the most important food crops in the world with highest production and productivity.

In India, maize contributes nearly 8% in the national food basket and more than Rs.100 billion to the agricultural Gross Domestic Products (GDP). Maize provides food, feed, fodder, and serves as a source of basic raw material for a number of industrial products. Being a potential crop in India, maize occupies an important place as a source of human food (25%), animal (12%) and poultry feed (49%), industrial products mainly starch (12%) and 1% each in brewery and seed (QPM for Food and Nutritional Security in India, ICAR, 2015).

Development of superior high oil lines for commercial use in hybrid combination is one of the major goals of today's maize improvement programme(s). In India, very less emphasis is given for exploitation of heterosis for quality traits like high oil and starch content in maize. In the process of selection of parental lines, fixation of selection criteria is one of the crucial tasks that often a plant breeder encounters. Harvestable 'grain yield' is usually the primary trait for improvement. However, it is a complex quantitative character governed by polygenes, which are highly

influenced by environmental fluctuations. A number of characters are associated directly or indirectly with this and various approaches have been used to explain yield as a function of different traits (Johnson, 1973). Therefore, selection for yield attributing parameters would be more useful rather than yield *per se*. The study of genetic makeup of such quantitative traits is essential in deciding proper methodology for their improvement along with yield.

Molecular markers allow geneticists and plant breeders to locate and follow numerous interacting genes that determine a complex trait. Genetic linkage maps can provide a more direct method for selecting desirable genes *via* their linkage to easily detectable molecular markers (Tanksley *et al.*, 1989). Combining marker assisted selection methods with conventional breeding schemes can increase the overall selection and therefore, the efficacy of breeding programme(s).

The first step in the search for heterosis in crop improvement is a full characterization of available genetic diversity, which forms the basis for the analysis of combining ability of inbred lines (Verbitskaya *et al.*, 1999; Diniz *et al.*, 2005). Genetic diversity and levels of genetic variations in maize can be estimated by using both modern molecular markers together with early detailed studies on morphological markers (Goodman and Bird, 1977) especially for breeding programme(s) in crop plants.

Maize breeding depends on the level of genetic diversity available which aids in the estimation of

the degree of expected heterosis from inbred lines together with heritability and variation during breeding (William and Michael, 2002; Duan *et al.*, 2006). In order to broaden the genetic base for commercial hybrids, there is a need to have more diversity among inbreds (Darrah and Zuber, 1986; James *et al.*, 2002). Genetic diversity is the most important factor limiting the average number of alleles identified per Simple Sequence Repeats (SSR) locus during screening programmes (Legesse, 2007). These markers can be effectively used to validate various traits such as oil content using genetic markers (James *et al.*, 2002; Sherry *et al.*, 2002).

Materials and methods

A total of 25 SSR markers, selected from maize GDB (Maize Genetics and Genomics Database, 2010) were used for genetic diversity analysis among ten maize inbreds. The details of SSR primers used are listed in Table 1. SSR markers allow detecting polymorphisms at the DNA level, which facilitates separating genotypes into well-defined groups based on genetic distance estimates (Enoki *et al.*, 2002).

DNA from young leaves were collected from 20 - 25 days old maize seedlings and immediately stored at -20°C till further processing. The DNA was extracted following CTAB extraction method according to Doyle and Doyle, 1987. The DNA quality estimation was done using Biophotometer plus. Polymerase Chain Reaction was performed to selectively amplify *in-vitro* a specific segment of the total genomic DNA to billion folds (Mullis *et al.*, 1986).

The amplified DNA fragments generated through SSR primers were resolved through electrophoresis in 2.5% agarose gel. Ethidium bromide solution at a final concentration of 0.03 ng/μl was added to the agarose solution. For electrophoresis, 15μl of the PCR product was mixed with 2μl of 6X loading dye (0.25% bromophenol blue in 30% glycerol) and loaded in slot of agarose gel. In order to determine the molecular size of the amplified products, each gel was also loaded with 1μg DNA of a 100 bp DNA size marker (Fermentas, USA). Gel electrophoresis was performed at a constant voltage of 65 V for about 3.5 hours. Finally, the gels were visualized under a UV light source in a gel documentation system (Gel Doc™ XR+, BIO-RAD, USA) and the images of amplification products were captured and stored in a computer for further analysis and future use.

SSR markers generated clear and unambiguous bands of various molecular weight sizes, which were scored for the presence (1) and absence (0) of the corresponding band among the 10 genotypes.

The marker data was used to generate a data matrix in Microsoft Excel 2007. This data matrix was subjected to further analysis using NTSYS-pc version 2.11 (Rohlf, 1997). The SIMQUAL program was used to calculate the Jaccard's similarity coefficients. The resulting similarity matrix was used to construct UPGMA (Unweighted Pair Group Method with Arithmetic Mean) based dendrogram. Polymorphic information content (PIC) for each SSR marker was calculated as per the formula:

$$PIC = 1 - \sum_{i=1}^k P_i^2$$

Where, P_i is the frequency of the i^{th} allele and k is the total number of different alleles at the specific locus. The binary data was subjected to principal component analysis (PCA) using the EIGEN and PROJ modules of NTSYSpc.

Results and discussion

Genetic diversity is the most important factor which contributes towards the realization of heterosis. SSR's are more potent than any other marker system and are robust tools to detect polymorphism at the molecular level. Initially, 58 SSR markers were selected randomly. Out of 58 SSR's, 25 SSR's produced a reproducible and polymorphic banding pattern which is exhibited in Table 1.

These 25 SSR's were further used to study molecular diversity among 10 inbreds which yielded a total of 25 polymorphic bands. The number of polymorphic bands per primer was with an average of two (Table 1). The PIC of 25 SSR primers ranged from 0.42 to 0.50 with an average PIC of 0.46 which was found greater than those in the previous studies by Shah *et al.* (2009), Babu *et al.* (2012), Kumar *et al.* (2012), Mishra and Singh (2012), Wasala and Prasanna (2013) which means that the selected primers were highly polymorphic and the degree of diversity among genotypes is very high.

The 25 maize SSR markers were capable of detecting 50 alleles with an average of 2.0 alleles per locus. The amplification of lower number of alleles per locus may be due to the poor resolution of agarose gel as compared to polyacrylamide gel. Within 25 SSR markers, 10 were used to validate oil content QTL's. Gel images showing typical SSR banding patterns generated by primers *bnlg* 2286 (Plate 1), *bnlg* 2244 (Plate 2) and *umc* 1360 (Plate 3) have been presented. Therefore, most of the primers were selected on the basis of positions of QTL's for oil content on corresponding

chromosomes. The SSR marker *bnlg2286*, *phi96100*, *umc1422*, *bnlg1452*, *bnlg2244*, *bnlg1711*, *umc2319*, *umc1360*, *phi119* and *phi050* linked with QTL's for oil content were amplified corresponding high and low oil content bands in most of high oil inbreds and low oil inbreds. Therefore, these markers could be used in marker aided selection for oil content in maize. Similar reports were confirmed by Zhang *et al.* (2008) and Yang *et al.* (2010).

The relationship between ten maize inbreds was studied by UPGMA based cluster analysis of data developed by 25 SSR markers. The Jaccard's similarity (Table 2) was used to produce a dendrogram to obtain clustering of maize inbreds (Fig. 1). The Jaccard's similarity coefficients ranged from 0.00 to 0.55. The highest similarity coefficient (1.00) was observed between genotypes HIGH OIL and DMHOC-09, DMHOC-09 and DMHOC-15 and HUZQPM-03 and HUZQPM-06, whereas, lowest similarity coefficient (0.00) was observed between genotypes HUZQPM-05 and HUZM-478, HUZQPM-05 and HUZQPM-01. In cluster analysis, all the inbreds were grouped broadly into five clusters, viz., Cluster I, II, III, IV and V.

The major Cluster I consisted of total of 4 inbreds, which could be further divided into 2 sub clusters, viz., IA (3 inbreds) and IB (1 inbred). Cluster II consisted of 2 inbreds HUZQPM-03 and HUZQPM-06. Cluster III consisted of two inbreds HUZM-478 and HUZQPM-01. Cluster IV and cluster V consisted of a single inbred each viz., HUZM-265 and HUZQPM-05 respectively. The results revealed that a rich diversity existed between the germplasm collections from different geographical regions of the world. Earlier reports predominantly support association of molecular data with clustering and pedigree records (Senior *et al.*, 1996; Liu *et al.*, 2003; Legesse *et al.*, 2007).

PCA analysis is used to estimate distance among the selected parents. Greater is the distance between two parents, greater is the heterosis. Results of Principal Component Analyses (PCA) using SSR markers are depicted in Fig. 2. The EIGEN values obtained from SSR markers revealed that the first three principal components cumulatively accounted for 61.17% of total variation, in which 47.31% was accounted by component 1 and 8.24% by component 2. In PCA analysis, broadly four clusters were observed. Clear grouping was observed among the inbreds on the basis of pedigree and other traits.

Acknowledgment

The corresponding author is very much grateful to the University Grants Commission and Banaras Hindu University for providing fellowship and resources for carrying out this research work.

References

- Babu, B.K., Pooja, P., Bhatt, J.C. and Agrawal, P.K. 2012. Characterization of Indian and exotic quality protein maize (QPM) and normal maize (*Zea mays* L.) inbreds using simple sequence repeat (SSR) markers. *Afri. J. Biotechnol.*, **11**(41): 9691 – 9700.
- Darrah, L.L. and Zuber, M.S. 1986. The United States farm corn germplasm base and commercial breeding strategies, *Crop Sci.*, **26**: 1109 – 1113.
- Diniz, L.E.C., Ruas, C.F., Carvalho, V.P., Torres, F.M., Ruas, E.A., Santos, M.O., Sera, T. and Ruas, P.M. 2005. Genetic diversity among forty coffee varieties assessed by RAPD markers associated with restriction digestion, *Brazilian Arch. Bio. Tech.*, **48**(4): 511 – 521.
- Doyle, J.J. and Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue, *Phytochemical Bulletin*, **19**: 11 – 15.
- Duan, Y.P., Chen, W.G., Li, M.S., Li, X.H., Liu, X., Tian, Q.Z., Bail and Zhang, S.H. 2006. The genetic diversity among 27 maize populations based on SSR data, *Sci. Agric. Sinica*, **39**: 1102 – 1113.
- Enoki, H., H. Sato, and K. Koinuma. 2002. SSR analysis of genetic diversity among maize inbred lines adapted to cold regions of Japan. *Theo. Appl. Gen.*, **104**: 1270 – 1278.
- Goodman M.M. and Bird, R. Mc. K. 1977. The races of maize IV: tentative grouping of 219 Latin American races, *Eco. Bot.*, **31**: 204 – 221.
- James, G.G., Joanne, A.L., Kendall, R.L., Margaret, E.S. and Stephen, K. 2002. SSR variation in important US maize inbred lines, *Crop Sci.*, **42**: 951 – 957.
- Johnson, G.R. 1973. Relationship between yield and several yield components in a set of maize hybrids, *Crop Sci.*, **13**: 649 – 651.
- Kumar, V., Singh, P.K. and Gupta, A. 2012. Molecular diversity in inbreds of quality protein maize (*Zea mays* L.), *Current Advances in Agric. Sci.*, **4**(1): 57 – 59.
- Legesse, B.W., Myburg, A.A., Pixley, K.V. and Botha, A.M. 2007. Genetic diversity of African maize inbred lines revealed by SSR markers, *Hereditas*, **144**: 10 – 7.
- Liu, J., Liu, G., She, Z., Zhi, Q. and Chen, G. 2003. Assessment of genetics relationship among 15



- elite maize inbred lines of China using SSR markers: comparison with data from heterosis and pedigree. *Xibie Zhiwu Xubeao*, **23**(4): 741 – 750.
- Mishra, P. and Singh, N.K. 2012. Allelic diversity among short duration maize (*Zea mays* L.) genotypes using SSR markers, *Madras Agric. J.*, **99**(4/6): 232 – 236.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. 1986. Specific enzymatic amplification of DNA *in vitro*: The polymerase chain reaction, *Cold Spring Harbor Symposium on Quantitative Biology*, **51**: 263 – 273.
- Quality Protein Maize for Food and Nutritional Security in India. 2009. (eds. M.L. Jat, Sain Dass, V.K. Yadav, J.C. Sekhar and D.K. Singh), Directorate of Maize Research, ICAR
- Rohlf, F.J. 1997. NTSYS-pc 2.1, Numerical Taxonomy and Multivariate Analysis System. Setauket, NY: Exeter Software.
- Senior, M.L., Chin, E.C.L. and Lee, M. 1996. Simple sequence repeat markers developed from maize sequences found in the GENE BANK database: Map construction, *Crop Sci.*, **36**: 1676 – 1683.
- Shah, Z.I., Munir, S.A., Arshad, I.S., Mumtaz, R.N. and Swati, Z.A. 2009. Genetic diversity of Pakistani maize genotypes using chromosome specific simple sequence repeats (SSR) primer sets, *Afri. J. Biotech.*, **8**(3): 375 – 379.
- Sherry, R.W., Larissa, M.W., Maud, I.T., Brandon, S.G. and Edward, S.B. 2002. Genetic diversity and selection in the maize starch pathway, *Proc. Natl. Acad. Sci.*, **99**: 12959 – 12962.
- Tanksley, S.D., Young, N.D., Paterson, A.H. and Bonierable, M.W. 1989. RFLP mapping in plant breeding: new tools for an old science, *Bio Techniques*, **7**: 257-264.
- Verbitskaya, T.G., Sivolap, Y.M., Guzhva, D.V. and Sokolov, V.M. 1999. Analysis of the relation between the genetic divergence of the DNA of corn lines and hybrid productivity, *Russ. Agric. Sci.*, **5**: 28 – 33.
- Wasala, S.K. and Prasanna, B.M. 2013. Microsatellite marker-based diversity and population genetic analysis of selected lowland and mid-altitude maize landrace accessions of India, *J. Plant Biochem. Biot.*, **22**(4): 392 – 400.
- William, S.K. and Michael, R.C. 2002. Essentials of genetics, Higher Education Press, Beijing.
- acid composition and oil content in high-oil maize, *Theo. Appl. Genet.*, **120**: 665 – 678.
- Zhang, J., Lu, X.Q., Song, X.F., Yan, J.B., Song, T.M., Dai, J.R., Rocheford, T. and Li, J.S. 2008. Mapping quantitative trait loci for oil, starch and protein concentrations in grain with high-oil maize by SSR markers, *Euphytica*, **162**: 335 – 344.
- Yang, X.H., Guo, Y.Q., Yan, J.B., Zhang, J., Song, T.M., Rocheford, T. and Li, J.S. 2010. Major and minor QTL and epistasis contribute to fatty



Table 1. List of polymorphic SSR markers

SI. No.	Primer	Forward sequence (5-'3')	Reverse sequence (5-'3')	For. Tm ⁰	Rev. Tm ⁰	PIC	Bin	Trait
1	bnlg 2277	TTACGGTACCAATTCGCTCC	GACGACGCCATTTTCTGATT	57.3	55.3	0.42	2.02	Random
2	bnlg 1523	GAGCACAGCTAGGCAAAAGG	CTCGCACGCTCTCTCTTCTT	59.4	59.4	0.48	3.03	Random
3	bnlg 1043	TTTGCTCTAAGGTCCCCATG	CATACCCACATCCCGGATAA	57.3	57.3	0.42	6.01	Random
4	nc 009	CGAAAGTCGATCGAGAGACC	CCTCTCTTCACCCCTTCCTT	59.4	59.4	0.48	6.04	Random
5	bnlg1070	TTCCAGTAAGGGAGGTGCTG	TAAGCAACATATAGCCGGGC	59.4	57.3	0.50	7.03	Random
6	umc1380	CTGCTGATGTCTGGAAGAACCCT	AGCATCATGCCAGCAGGTTTT	62.4	57.9	0.42	10.01	Random
7	phi084	AGAAGGAATCCGATCCATCCAAGC	CACCCGTACTTGAGGAAAACCC	62.7	62.1	0.48	10.04	Random
8	bnlg 1754	CCATCGCTGTACACATGAGG	TACCCGAAGGATCTGTTTGC	59.4	57.3	0.50	3.09	Protein Content
9	bnlg 2082	GACGGAAGGTGGAGCATAGA	ACGAACGTGATACGGGTCTC	59.4	59.4	0.50	8.03	Protein Content
10	umc 1652	GAGAGCAGTAGCACTGACCCTTTC	CACTCGACCTCGATCGGAAC	64.4	61.4	0.32	4.04	Plant Height
11	bnlg 2286	CGGAACCTGCTGCAGTTAAT	GAGATGCAGGAATGGGAAAA	57.3	55.3	0.50	1.04	Oil Content
12	phi 96100	AGGAGGACCCCAACTCCTG	TTGCACGAGCCATCGTAT	61.0	53.7	0.50	2.01	Oil Content
13	umc 1422	GAGATAAGCTTCGCCCTGTACCTC	CTCATCGCGATCTCCCAGTC	64.4	61.4	0.42	2.02	Oil Content
14	bnlg1452	CTCCTCTCCTCCACGATCAC	CGCAAACGATCTCTGACCTT	61.4	57.3	0.48	3.04	Oil Content
15	bnlg 2244	CAGGAAAACGAAAACCCAGA	CTACGCGGGTCTCATCTCAT	55.3	59.4	0.50	4.08	Oil Content
16	bnlg 1711	TAATCTTGGGGGGTTTAGGG	GACATGTCCCATTCCCATT	57.3	57.3	0.48	5.07	Oil Content
17	umc2319	GATCCACGCGAGGTTCAGT	GCTCTCACTAGCCTCGCATTC	61.4	64.0	0.50	6.05	Oil Content
18	umc1360	GCTAGTTGAGTTCGACACCAGGT	TGACTGTGACTGTGACTATGACCG	62.7	62.7	0.42	8.02	Oil Content
19	phi119	GGGCTCCAGTTTTAGTCATTGG	ATCTTTCGTGCGGAGGAATGGTCA	62.4	62.7	0.50	8.02	Oil Content
20	phi050	TAACATGCCAGACACATACGGACAG	ATGGCTCTAGCGAAGCGTAGAG	63.0	62.1	0.42	10.03	Oil Content
21	phi112	TGCCCTGCAGGTTACATTGAGT	AGGAGTACGCTTGGATGCTCTTC	62.4	62.4	0.42	7.01	Opaque 2
22	umc1035	CTGGCATGATCACGCTATGTATG	TAACATCAGCAGGTTTGCTCATTC	60.6	59.3	0.42	1.06	Leaf Length
23	umc1593a	CATGTTGATCATATGCACGAGAGA	CAGCCTGGTACTCATGGTTAAT	59.3	60.6	0.48	7.03	Leaf Length
24	umc1395	TGAATGAGTGGCATTCAAAATCTG	CAGATTGCATGTGTGAGTGTGTGT	57.6	61.0	0.48	1.05	Ear Height
25	bnlg 1014	CACGCTGTTTCAGACAGGAA	CGCCTGTGATTGCACTACAC	57.3	59.4	0.50	1.01	Ear Diameter

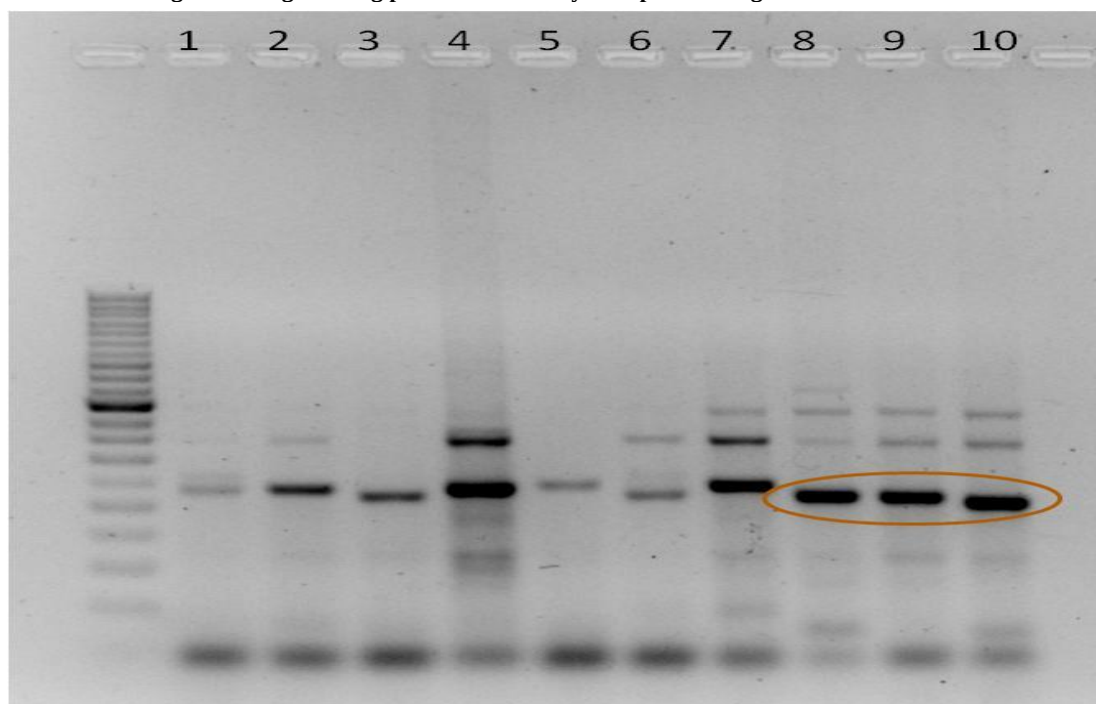
Note: For. Tm⁰ = Forward primer temperature, Rev. Tm⁰ = Reverse primer temperature, PIC= Polymorphic information content.



Table 2. Jaccard's similarity coefficient of the 10 maize inbreds based on 25 polymorphic SSR markers

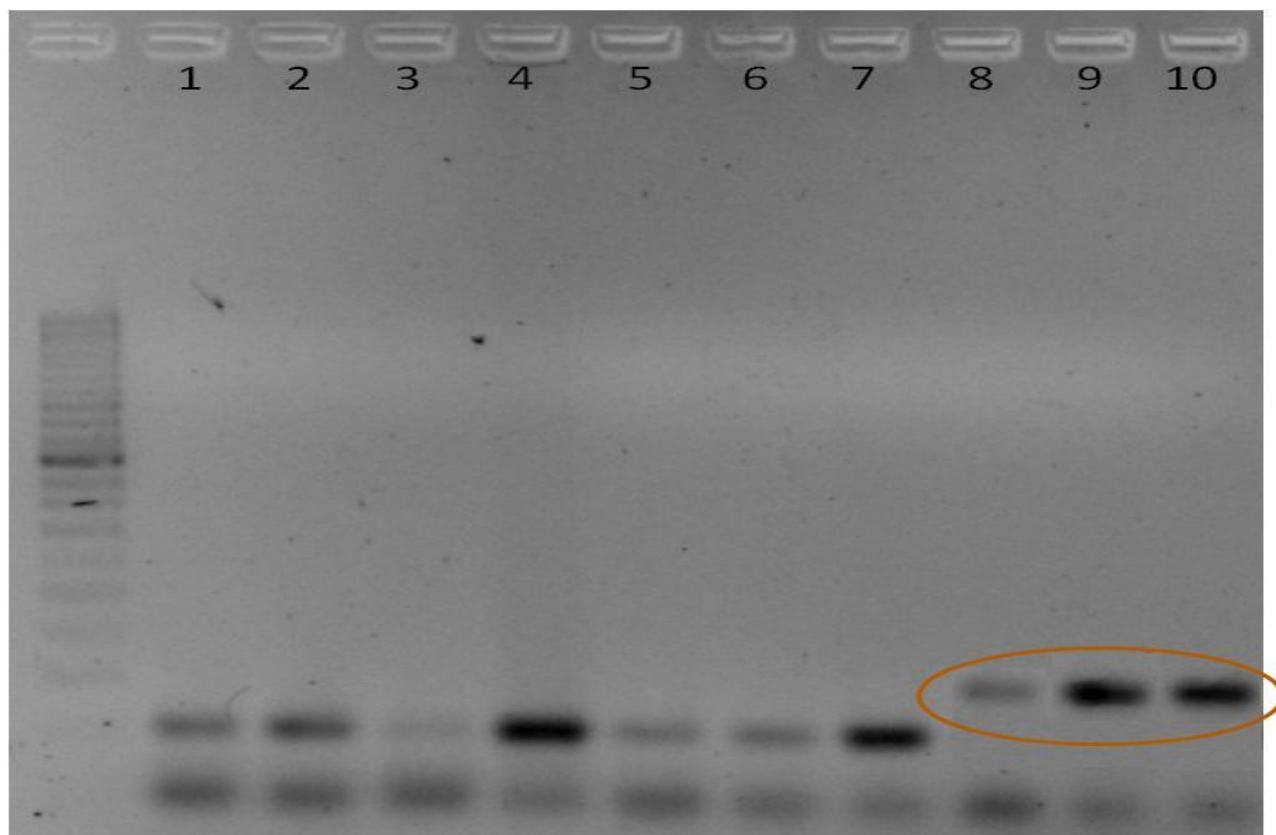
Inbreds	HUZH-53	HUZH-265	HUZH-478	HUZHQM-01	HUZHQM-03	HUZHQM-05	HUZHQM-06	HIGH OIL	DMHOC-09	DMHOC-15
HUZH-53	1.00									
HUZH-265	0.21	1.00								
HUZH-478	0.18	0.27	1.00							
HUZHQM-01	0.44	0.32	0.47	1.00						
HUZHQM-03	0.37	0.32	0.22	0.22	1.00					
HUZHQM-05	0.07	0.17	0.00	0.00	0.21	1.00				
HUZHQM-06	0.35	0.29	0.19	0.19	0.47	0.27	1.00			
HIGHOIL	0.40	0.23	0.41	0.36	0.50	0.06	0.29	1.00		
DMHOC-09	0.42	0.13	0.35	0.45	0.38	0.06	0.24	0.55	1.00	
DMHOC-15	0.33	0.29	0.26	0.43	0.43	0.12	0.35	0.45	0.48	1.00

Plate 1: Gel images showing banding profile obtained by SSR primer bnlg 2286



Lane 1-10 represents the maize inbreds; M= 200bp DNA size marker. 1-HUZM-53, 2-HUZM-265, 3-HUZM-478, 4-HUZQPM-01, 5-HUZQPM-0, 6-HUZQPM-05, 7-HUZQPM-06, 8-HIGH OIL, 9-DMHOC-09, 10-DMHOC-15.

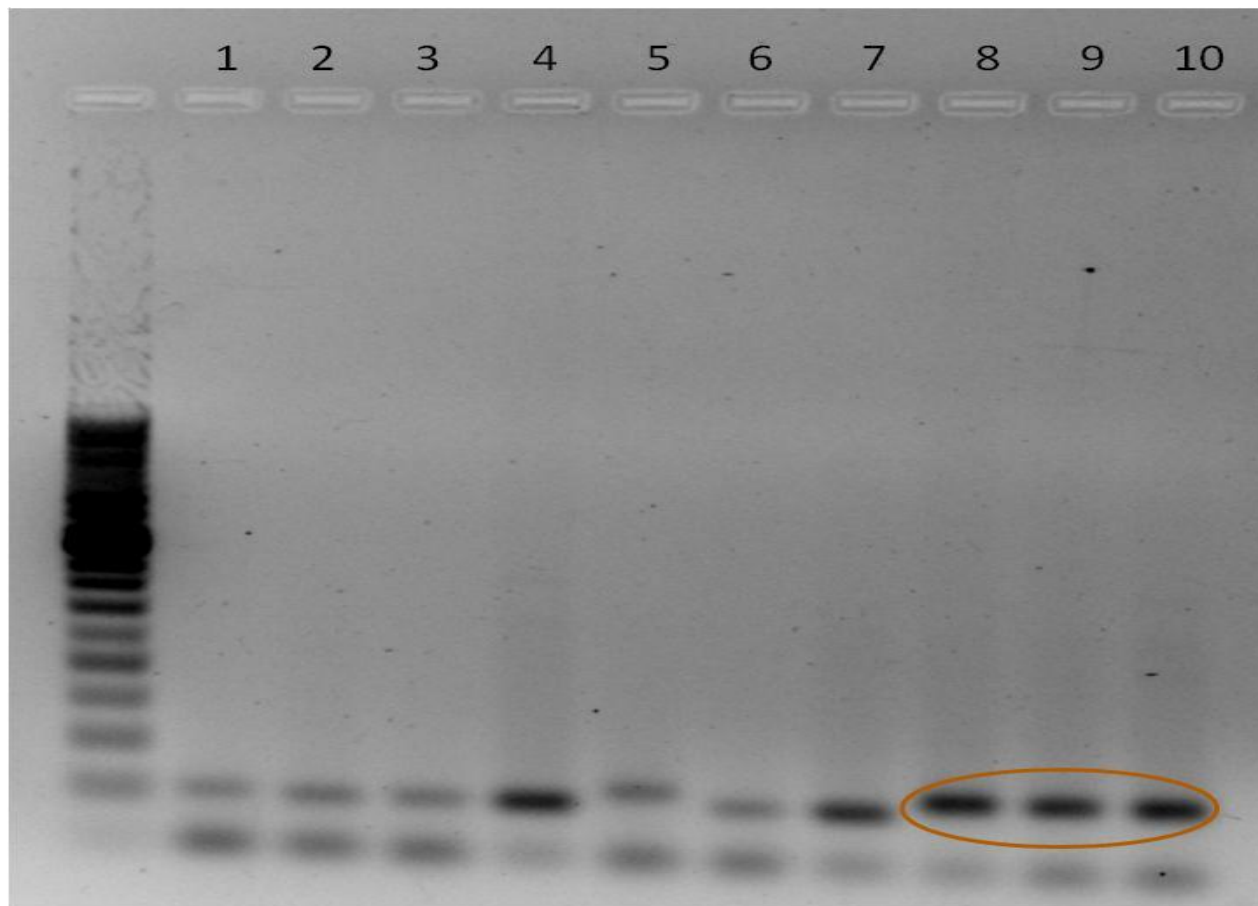
Plate 2: Gel images showing banding profile obtained by SSR primer BNLG 2244



Lane 1-10 represents the maize inbreds; M= 200bp DNA size marker.1-HUZM-53, 2-HUZM-265, 3-HUZM-478, 4-HUZQPM-01,

5-HUZQPM-03, 6-HUZQPM-05, 7-HUZQPM-06, 8-HIGH OIL, 9-DMHOC-09, 10-DMHOC-15.

Plate 3: Gel images showing banding profile obtained by SSR primer UMC 1360



Lane 1-10 represents the maize inbreds; M= 200bp DNA size marker. 1-HUZM-53, 2-HUZM-265, 3-HUZM-478, 4-HUZQPM-01, 5-HUZQPM-03, 6-HUZQPM-05, 7-HUZQPM-06, 8-HIGH OIL, 9-DMHOC-09, 10-DMHOC-15

Fig. 1: Dendrogram of ten maize inbreds based on genetic similarity

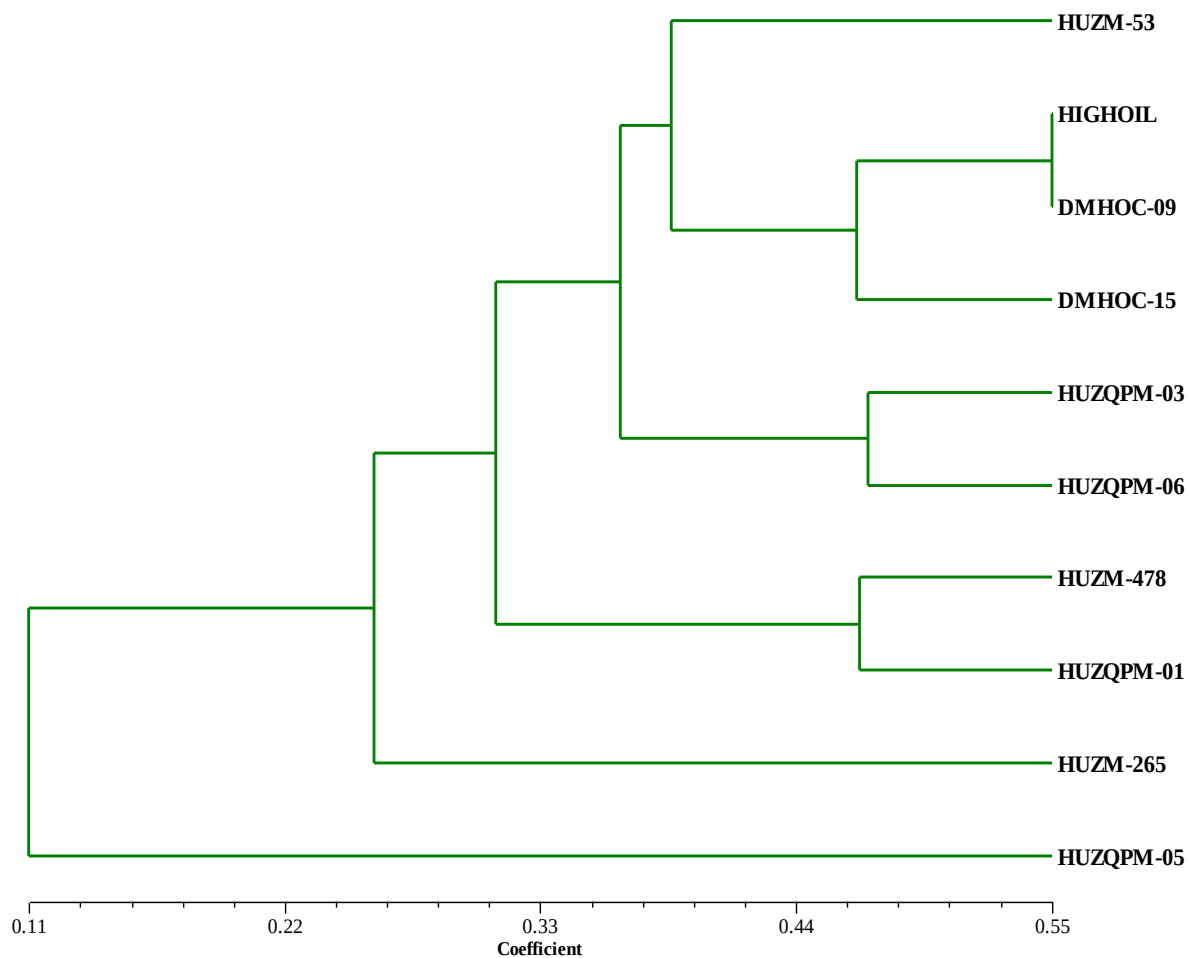
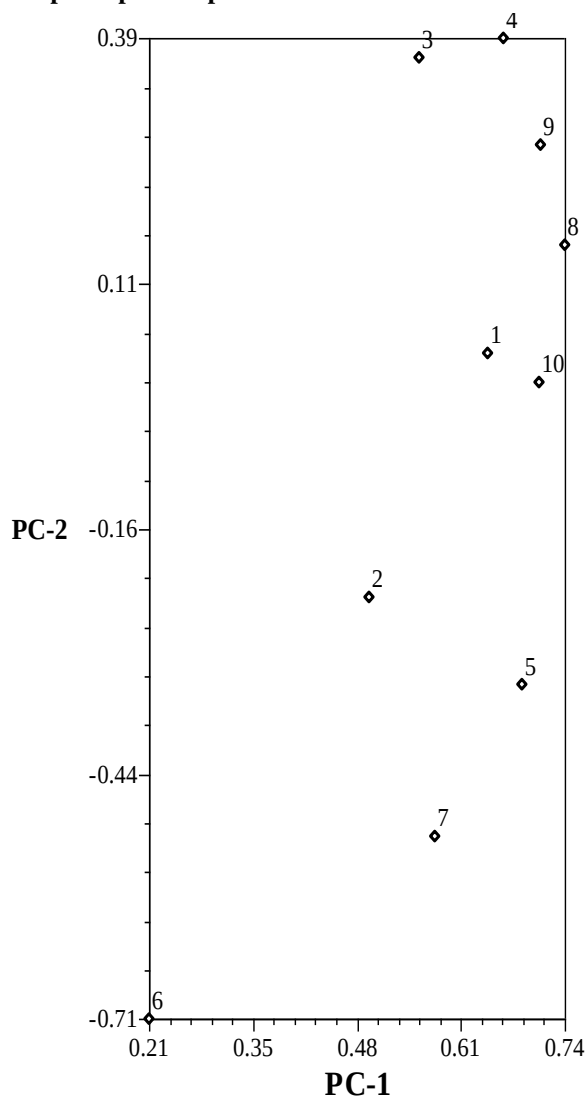


Fig. 2: Two dimensional plot of principal components 1 and 2 based on SSR markers of maize inbreds



1= HUZM-53, 2= HUZM-265, 3= HUZM-478, 4= HUZQPM-01, 5= HUZQPM-03, 6= HUZQPM-05, 7= HUZQPM-06, 8= HIGH OIL, 9= DMHOC-09, 10= DMHOC-15