Mysore J. Agric. Sci., 57 (4) : 320-332 (2023)

SSR Marker-based Diversity among Elite Maize (*Zea mays* L.) Inbred Lines and Parents of Elite Single Cross Hybrids with Known Heterotic Groups

C. B. Siddu, S. Ramesh, G. Basanagouda and M. P. Kalpana
Department of Genetics and Plant Breeding, College of Agriculture, UAS, GKVK, Bengaluru - 560 065
e-Mail : siddubc16@gmail.com

Authors Contribution

C. B. Siddu : Conceptualization, design, data analysis, curation and original manuscript draft;
S. Ramesh : Conceptualization, editing and supervision;
G. Basanagouda & M. P. Kalpana : Data collection and field experiment

Corresponding Author :
C. B. Siddu
Department of Genetics and Plant Breeding, College of Agriculture, UAS, GKVK, Bangalore

Received : August 2023
Accepted : September 2023

Abstract

Exploiting heterosis through F_1 hybrids is considered one of the most effective methods to increase maize productivity. However, developing successful heterotic hybrids requires the use of the combination of diverse inbred parents to develop breeding populations (BPs) and desirable testers to identify new inbred lines for use in developing heterotic hybrids. As step towards this, the objectives of the present investigation were to (i) assess genetic diversity among elite inbred lines using simple sequence repeat (SSR)-based markers and (ii) to identify most desirable combinations of inbred lines. Sixteen lines which included eleven elite inbred lines with high general combining ability and five parents of three commercialised elite single cross hybrids - Hema, Nithyashree and MAH 14-5 were used in the study. These lines were genotyped using 132 SSR markers distributed across the maize genome. The results revealed polymorphism of hundred markers. The number of alleles per locus ranged from 2 to 8. Narrow difference between average (N_a) (3.31) and effective (N_e) (2.40) number of alleles per locus suggest near-equal frequency of alleles detected at each poly morphic locus. This is amply reflected from high correlation of N_a and N_e with polymorphic information content (PIC). Thirty-three markers with ≥3 alleles (multi-allelic) were more informative than bi-allelic markers as reflected from high magnitude of PIC. A wide range of the estimates of dissimilarity indices between pairs of inbred lines suggested differential frequency of alleles at the polymorphic SSR markers between the inbred lines. The best triplet combinations of two inbred parents (with high dissimilarity indices) and a tester (with high dissimilarity indices between parents and the tester) were identified to maximize the chances of recovering new elite inbred lines for use in heterotic hybrids.

Keywords : SSR Marker based diversity in maize, Dissimilarity indices, Heterotic groups, Triplet combinations

Maize being a C_4 species has highest productivity among cereals (Deepak and Vasudevan, 2023). It is a multipurpose crop as it is being used for food, fodder, fuel and feed purpose. Hence, it is regarded as 4F crop. With ever increasing human population coupled with decreasing land and water resources available for production of crops including maize, necessitates increasing productivity per unit of land and per drop of water. One of the best ways to increase the productivity is to exploit heterosis through F_1 hybrids.

Heterosis has been well exploited in maize by developing and deploying F_1 hybrids in all parts of world including India (Sowmya and Gangappa, 2018). Developing heterotic hybrids requires use of diverse inbred parents with favourable alleles dispersed between them. In most hybrid breeding programmes, development of new inbred lines is based on recycling the elite inbred lines which are either parents of elite commercialized hybrids or those of a large number of non-commercialized test hybrids. As the number of elite inbred lines available increases, the number of

segregating (breeding) populations to be developed for use in developing new inbred lines also increases. In this context, choice of inbred lines to develop breeding populations (BPs) which are likely to results in high frequency of superior inbred lines is the key for the success of hybrid breeding. The frequency of superior inbred lines would be higher if desirable alleles at loci controlling target traits are dispersed among the parent of BPs (Falk, 2010). Dispersion of desirable alleles though results genetic dissimilaritymay results in phenotypic similarity of the parents. Hence, it is difficult to detect the pairs of parents between which favourable alleles are dispersed. It is therefore hypothesized that DNA based-markers are handy for assessing genetic differences among elite inbred lines. The DNA markers especially those based on simple sequence repeats (SSRs) owing to their mono-locus multi-allelic codominant inheritance, selective neutrality, distribution throughout the genome, ease of assay and high reproducibility are considered as ideal ones for assessing diversity among elite inbred lines (Senior *et al.,* 1998; Selvi *et al.,* 2003; Bantte and Prasanna, 2003; Ranatunga *et al.,* 2009; Shehata *et al.,* 2009; Nepolean *et al.,* 2013 and Sserumaga *et al.,* 2014). The twin objectives of the present study were to (i) assess the genetic diversity of elite inbred lines and (ii) identify the triplet combination of pairs of elite inbred lines to develop BPs and tester for selection of desirable new inbred lines derived from BPs using SSR markers.

Material and Methods

Genetic Material

The material for the study consisted of 16 elite inbred lines representing diverse genetic backgrounds (Siddu *et al.,* 2023).While eleven of these 16 lines are elite inbred lines with high general combining ability (Sowmya and Gangappa, 2018) assessed by repeated evaluation and five are parents of three elite single cross hybrids (SCHs) developed by University of Agricultural Sciences (UAS), Bangalore. Thirteen inbred lines belong to known heterotic groups (Table 1).

Table 1
Membership of elite inbred lines to three heterotic groups

Heterotic group 1	Heterotic group 2	Heterotic group 3
NAI 137	MAI 105	CML 395
V 70	40061	CML 568
MAI 194	MAI 21	CML 582
MAI 214		
MAI 345		
MAI 349		

Genomic DNA Isolation and Genotyping

DNA was isolated from the young leaves (18-21 days old) as per modified CTAB (Cetyl Trimethyl Ammonium Bromide) method (Hoisington *et al.,* 1994). Quality and quantity of the sample DNA stock was checked using 0.8 per cent agarose gel with known concentration of uncut lambda DNA of 200 ng/µl. The inbred lines were assayed using 132 genomic SSR markers chosen to represent all the ten chromosomes. The Polymerase Chain Reaction (PCR) mixture contained approximately 2µl of DNA (50 ng per µl), 0.1 µl Taq polymerase (3 units per µl), 2.0 µl 10X TE buffer, 0.4 µl $MgCl_2$ (25 mM), 1.0 µl DNTPs (2 mM) and 2 µl each of forward and reverse primer in a total of 20 µl solution. PCR amplification was carried out in the Applied Biosystems and BIO-RAD thermal cycler which consists of initial denaturation at 95°C for 4 min followed by 35 cycles consisting of denaturation at 94°C for 0.35 min, 0.45 min at respective annealing temperature of primer, extension at 72°C for 1 min. Final extension was at 72°C for 8 min. The denaturation, annealing and extension step is followed by infinite time at 4°C for holding.

Separation and Scoring of Amplified Products

The PCR products was loaded on 4 per cent agarose gel in 1X TBE buffer stained with 15 µl ethidium

bromide. Amplicons were separated in an electrophoresis unit at 90 V for three hours using 1X TBE buffer. The gel was visualized under UV transillumination and captured in a gel documentation system (Bio Rad). The amplified PCR product for each of the sixteen elite inbred lines were scored manually by using 100bp ladder loaded alongside as reference. The markers which differentiated at least one inbred line were considered as polymorphic ones. At the polymorphic markers, amplicon with distinctly different sizes within the product range in comparison to the ladder were designated as different allele. Based on this criterion, varying numbers of alleles at polymorphic markers were detected. A matrix of SSR markers and their corresponding alleles was prepared in Microsoft Excel spread sheet for use in estimating population genetic parameters.

Estimation of Population Genetic Parameters

SSR marker allelic data matrix of sixteenelite inbred lines was used to estimate various population genetic parameters such as major allele frequency (Mf), total number of alleles (T_a), average number of alleles per locus (N_a), effective number of alleles per locus (N_e), observed heterozygosity (H_o), expected heterozy gosity (H_e)and polymorphic information content (PIC). The estimates of T_a, N_a, N_e, H_o and H_e were obtained using Gen AlEx 6.5 software (Smouse and Peakall, 2012). The estimates of Mfand PIC were obtained using Power Marker 3.25 software (Liu and Muse, 2005). The following formulae were used to estimate these parameters. The N_a was estimated as $N_a = (1/k) \sum^{k}_{i=1} ni$) (Hartl *et al.,* 1997), where, n_i is the total number of detected alleles per SSR locus and 'k' is the number of SSR loci. The N_e was estimated as $N_e = \frac{(1)}{(1-h)} = 1 / \sum^{k}_{i=1} Pi^2$(Hartl *et al.,* 1997), where, p_i is the frequency of the i^{th} allele at a SSR locus and $h = 1-\sum p_i^2$ is the estimate of heterozygosis at a SSR locus. The H_e was estimated as $h = 1 - p_i^2 - q_i^2$, where p_i = frequency of i^{th} allele and q_j = frequency of j^{th} allele. The average H_e across SSR loci was estimated as $H_e = \sum^{k}_{i=1} (1 - p_i^2 - q_i^2) / k,$ where, 'K' is number of SSR loci assuming Hardy-Weinberg equilibrium. The H_o was estimated as the proportion of heterozygote inbred lines at each SSR locus and averaged across all the SSR loci. As sample size (N) is small in the present study, an unbiased estimate of H_e was estimated as, $uH_e = \frac{2N}{(2N-1)} \sum^{k}_{i=1} (1 - p_i^2 - q_i^2)$ (Nei, 1978; Saitou and Nei, 1987), where, N = number of inbred lines. The PIC was estimated as $PIC = 1 - \sum^{k}_{i=1} p_i^2 - \sum^{l-1}_{i=1} \sum^{l}_{j=i+1} 2 p_i^2 p_j^2$ (Botstein *et al.,* 1980), where p_i is the frequency of the i^{th} allele, p_j is the frequency of the j^{th} allele and 'l' is the number of alleles at SSR loci.

Inter-Relationship Among the Estimates of N_a, N_e, PIC, H_e and H_o

Inter-relationship among the estimates of N_a, N_e, PIC, H_e and H_o was explored to obtain insights on the presence of equal / unequal frequency of alleles at SSR loci and presence of rare alleles.

Dissimilarity Between Pairs of Inbred Lines

Dissimilarity indices (Nei, 1978) between pairs of inbred lines were estimated using GenAlEx 6.5 software (Smouse and Peakall, 2012).

Identification of Best Triplet Combination of Two Inbredlines and a Tester

The success of hybrid cultivar development hinges on the choice of inbred lines to develop BPs and selection of a tester to identify superior new inbred lines derived from BPs. The parental lines should be genetically diverse to generate high variability in BPs and tester should be able to discriminate the new inbred lines and produce high test cross performance. This is possible when tester is an elite inbred line that belongs to heterotic group (HG) different from the one to which parents of BPs belong. The pairs of inbred lines that belong to each HG with high dissimilarity indices were selected as parents of BPs and tester from other two HGs were selected such that average dissimilarity indices of crosses between inbred parents and the tester were highest.

Results and Discussion

Polymorphism Among Elite Inbred Lines at SSR Marker Loci

Variability among test genotypes is generally quantified using per cent polymorphism of assayed

markers. A good number (100 of 132) of polymorphic SSR markers (Fig. 1) in the present study suggested considerable variability among 16 elite inbred lines. It is not just per cent polymorphism, but the total number of alleles (T_a) detected at polymorphic markers is considered as one of the good indicators of the extent of variability that exist among the test genotypes. Greater number of alleles (though depends on the size of the test population and resolving system used) imply higher genetic variation among the test genotypes at SSR markers loci (Nei, 1987). In the present study, as high as 331 alleles (Table 2) detected across 100 polymorphic markers suggest substantial variability among 16 elite inbred lines. Arguably, these many numbers of SSR marker alleles especially among elite inbred lines which have experienced repeated selection is significant. High mutability of SSR markers (Li *et al.,* 2004; Kashi and King, 2006) could be attributed to presence of a large number of alleles at the 100 polymorphic markers assayed in the present study. SSRs mutate at rates up to 10 times greater than point mutations (Gemayel *et al.,* 2012). Polymorphism at SSR marker loci results from addition/deletion of the entire repeat units/motifs. The variation in number of repeats of a particular repeat motif across test genotypesis detected as polymorphism. The varied numbers of repeats across test genotypes are detected as different sized amplicons and are considered as different alleles (Vieira *et al.,* 2016). The differences in repeat numbers occur as a result of polymerase strand-slippage during DNA replication and by recombination errors (Li *et. al.,* 2004; Kashi & King, 2006; Gemayel *et al.,* 2012 and Vieira *et al.,* 2016). The multi-allelic and codominant inheritance of alleles at SSR markers and ease of their assay makes them more informative and breeders' friendly ones, as these can be resolved in simple agarose gel even in any basic molecular biology laboratory (Dutta *et al.,* 2013).

TABLE 2

Estimates of 100 polymorphic SSR-markers assay-based population genetic parameters in 16 maize inbred lines

SSR markers	Number of alleles	Average major allele frequency (MAF)	Average of alleles/locus (N_a)	Average effective alleles/ locus (N_e)	Average observed heterozygosity (H_o)	Average expected heterozygosity (H_e)	Average unbiased heterozygosity (uH_e)	Average Poly morphic Information Content (PIC)
100 Poly morphic	331	0.59	3.31	2.40	0.05	0.52	0.53	0.46
Range	-	0.22 - 0.94	2 - 8	1.13 - 5.88	0 - 0.81	0.12 - 0.83	0.12 - 0.86	0.11 - 0.81

Average (N_a) and Effective (N_e) Number of Alleles Per SSR Locus

Besides T_a, estimates of N_a are also good indicator of polymorphism within the test population and are function of size and kind of population.

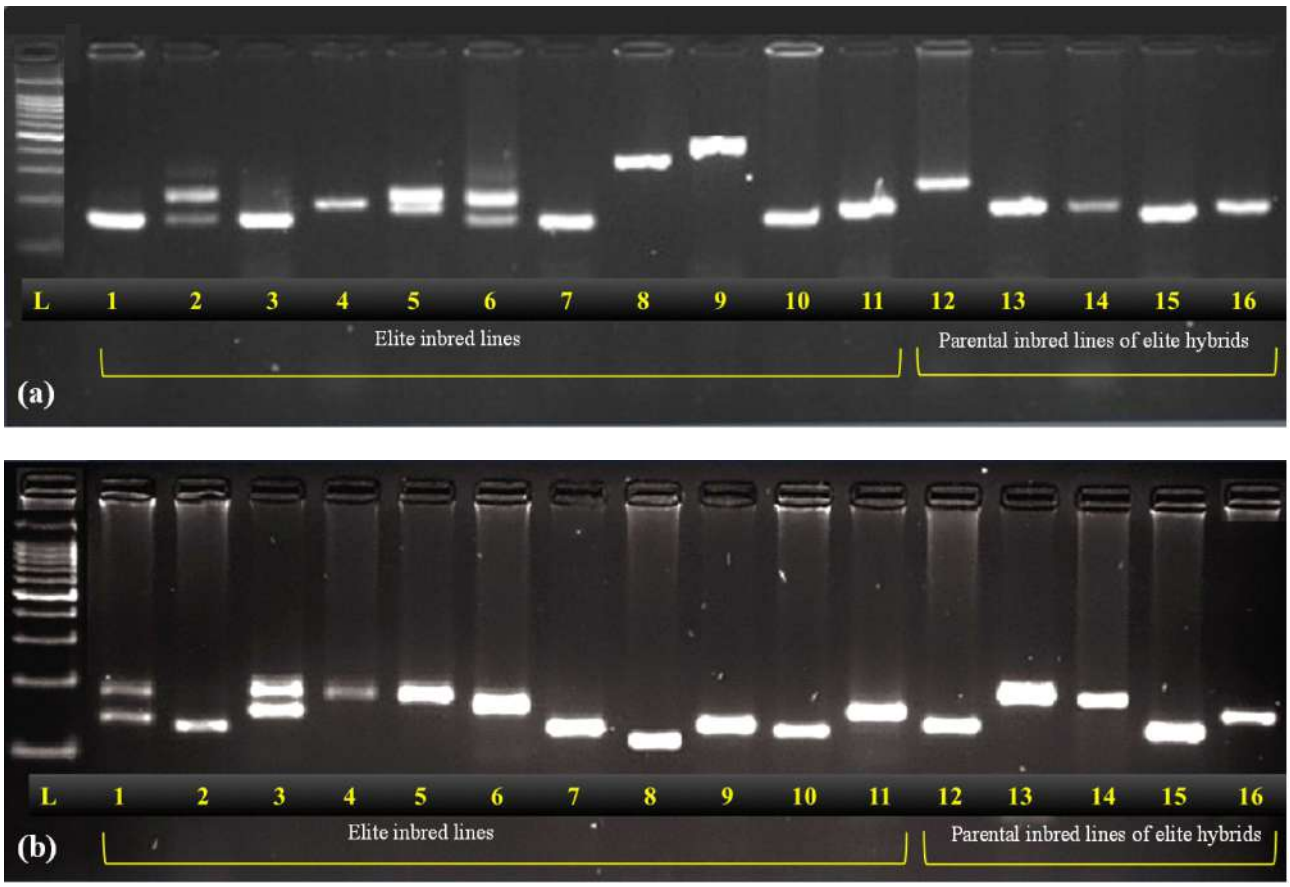

Fig. 1 : A representative agarose gel image showing SSR markers (a) umc1653 and (b) bnlg1175 profiles of 16 maize inbred lines
Lane, L - 100 bp standard DNA ladder; lanes, 1 to 11 – elite inbred lines; lane 12 to 16 parental inbred lines of elite hybrids

Higher estimates of N_a imply greater genetic variation at SSR marker loci (Nei, 1987). In the present study, 2 to 8 alleles across 100 polymorphic loci with N_a of 3.31 alleles could be detected (Table 2). It is possible that the number of alleles among diverse unselected genotypes would be more than those reported across the SSR markers assayed among the 16 elite inbred lines used in the study. Greater the size of the test population, greater is the chance of detection of more number of alleles per SSR locus. However, the estimates of N_a are reported to increase with increase in sample size up to 40 and thereafter stabilize. The correlation coefficient of N_a with sample size is highly significant with $r^2 = 0.97$ (Laurentin, 2009). Researchers who have assayed different kinds and size of populations have, therefore reported varying estimates of N_a. To quote a few, while Kumar *et al.* (2022) detected 2 to 9 alleles, Vathana *et al.* (2019) detected 4 to 17 alleles in a diverse set of Chinese maize inbred lines.

Apart from T_a and N_a, the estimates of N_e are also good indicator of variability among test genotypes. The estimates of N_e are equally frequent alleles which result in frequency of heterozygote genotypes observed in the study (Hartl and Clark, 1997). The estimates of N_e among 16 elite inbred lines ranged from 1.13 to 5.88 with an average of 2.40 alleles per marker (Table 2). The estimates of N_e are most often lower than those of N_a except when detected alleles at each of assayed polymorphic SSR locus are equally frequent. While the former (N_e) are less sensitive, the latter ones (N_a) are more sensitive to population size (Bashalkhanov *et al.,* 2009). The estimates of N_e, therefore allow researchers to compare different test populations for SSR marker polymorphism. Thus, estimates of N_e being less

sensitive to population size are considered as reliable indicators of diversity among sets of test genotypes. Previous researchers such as Nyaligwa *et al.* (2015) detected N_e ranging from 1.00 to 4.70 with a mean of 2.40 alleles per locus. Large differences between N_a and the N_e indicate low frequencies of a few alleles, as they are likely to be present only in a few genotypes. Hence, the estimates of N_e in combination with those of N_a provide useful clues regarding the presence of rare alleles in the test population (Laurentin, 2009). A rather narrow difference between the estimates of N_a (3.31 alleles) and N_e (2.4 alleles) in the present study suggests a relatively equal frequency of alleles detected at each of the assayed polymorphic SSR locus.

Gene Diversity / Heterozygosis

Genetic variability at SSR loci is also quantified based on the estimates of H_o and H_e. The estimates of H_o are the proportion/frequency of genotypes that are heterozygous at a given SSR locus. H_e is the frequency of genotypes that are heterozygous at a locus as expected under the assumption of Hardy-Weinberg Equilibrium (HWE). Though H_e is less sensitive to sample size than H_o, it is suggested to use a minimum of 25 markers and a minimum of 40 genotypes to assess SSR marker-based diversity (Laurentin, 2009). The present study is based on number of SSR markers (100) far greater than the minimum (25), but based on fewer (16) genotypes. Hence, estimate of H_e was corrected for small population size to obtain its unbiased version (μH_e). Marginal differences between the estimates of μH_e (corrected for small population size) and H_e (un-corrected) suggest adequacy of the number of inbred lines used in the present study. This is evident from the range as well as average μH_e and H_e estimates. While the estimates of μH_e ranged from 0.12 to 0.86 with an average of 0.53, those of H_e ranged from 0.12 to 0.83 with an average of 0.52 (Table 2).

Typically, both H_o and H_e range from 0 to 1, where $H_o = 0$ indicates absence of heterozygote genotypes and $H_o = 1$ indicates that all the test inbred lines are heterozygotes. The extent of differences between H_o and H_e could be used to infer about the relative levels of inbreeding/random mating among the inbred lines. $H_o = H_e$ indicater and ommating in the test population; $H_o < H_e$ indicate inbreeding in the test population; $H_o > H_e$ indicate mating system that avoids inbreedingin the test population. Given that the elite inbred lines used in the study are near homozygous at most loci, true to our expectation, the estimates of H_o which ranged from 0.00 to 0.81 with a mean of 0.05 were far lower than those of H_e (Table 2). These results are amply reflected by poor correlation between H_o and H_e (Fig. 2a).

The estimates of H_e are used to quantify the levels of genetic diversity as they depend solely on the number and relative frequencies of alleles. They are also measures of evenness of allelic frequencies. The estimates of H_e in the test population assayed using SSR markers with equally frequent alleles are greater than those in the test population assayed using SSR markers with un-equally frequent alleles. There fore, high correspondence between H_e and N_e is expected. For example, H_e is 0.85 for N_e of 6.67 alleles; $N_e = 1/1 - H_e = 1/1-0.85 = 6.67$. If an SSR locus has '8' alleles (maximum possible $H_e = 0.875$); but if H_e is say only 0.6, the $N_e = 2.5$; $N_e = 1/1 - H_e = 1/1-0.6 = 2.5$. Strong correlation between N_e and H_e (Fig.2b) in the present study indicate equal frequency of the alleles detected at 100 polymorphic SSR marker loci.

Informativeness of SSR Markers

Informativeness / ability of SSR markers to discriminate any set of test genotypes (elite inbred lines in the present study) is quantified by PIC. As is true for H_e, PIC is a function of number of alleles and their relative frequencies (Hilderbrand *et al.,* 1992 and Guo & Elston, 1999). Theoretically, the estimates of PIC can range from 0.00 to 1.00. If a marker has only one allele, PIC will be '0.00' and '1.00' if it has an infinite number of alleles. Markers with PIC values > 0.50 are considered as highly informative; those with PIC values ranging from 0.25 to 0.50 are considered moderately informative and those with PIC values less than 0.25 are considered less informative. In the present study, estimates of PIC ranged from 0.11 to 0.81 (Table 2). Majority (92) of 100 SSR markers were either highly/moderately informative.

The magnitude of PIC may vary for the same set of SSR markers depending on the number and kind of genotypes used in the study. Thus, a set of SSR markers could be more informative for a set of genotypes and it could be moderately or even less informative for a different set of genotypes. This is evident from wide range of PIC reported by previous researchers such as, Nyaligwa *et al.* (2015), Vathana *et al.* (2019) and Kumar *et al.* (2022) in maize.

As is true between N_e and H_e, a good correlation between N_e and PIC is expected, but not between N_a and PIC. Thus, correlation between N_e and PIC could be a useful indicates of occurrence of rare alleles (Laurentin, 2009). A marker with only two equally frequent alleles will have maximum PIC of 0.375. Similarly, a marker with three and four equally frequent alleles will have maximum PIC values of 0.7 and 0.83, respectively. Thus, markers with un-equally frequent 2, 3 and 4 alleles will have PIC values less than 0.375, 0.63 and 0.70, respectively. A marker with greater number of equally frequent will have greater values of PIC. such marker is therefore considered more informative. However, the direct relationship between number of alleles and PIC does not hold good if one or more alleles are more frequent than others (Hildebr and *et al.,* 1992). Thus, mere occurrence of a greater number of alleles will not make a marker more informative if the frequencies of detected alleles are unequal. Hence a combination of high PIC and a greater number of alleles makes a marker more informative. In the present study, the numbers of alleles detected across SSR markers are equally frequent as supported by strong correlation between N_e and PIC (Fig.2c) and between H_e and PIC (Fig.2d).

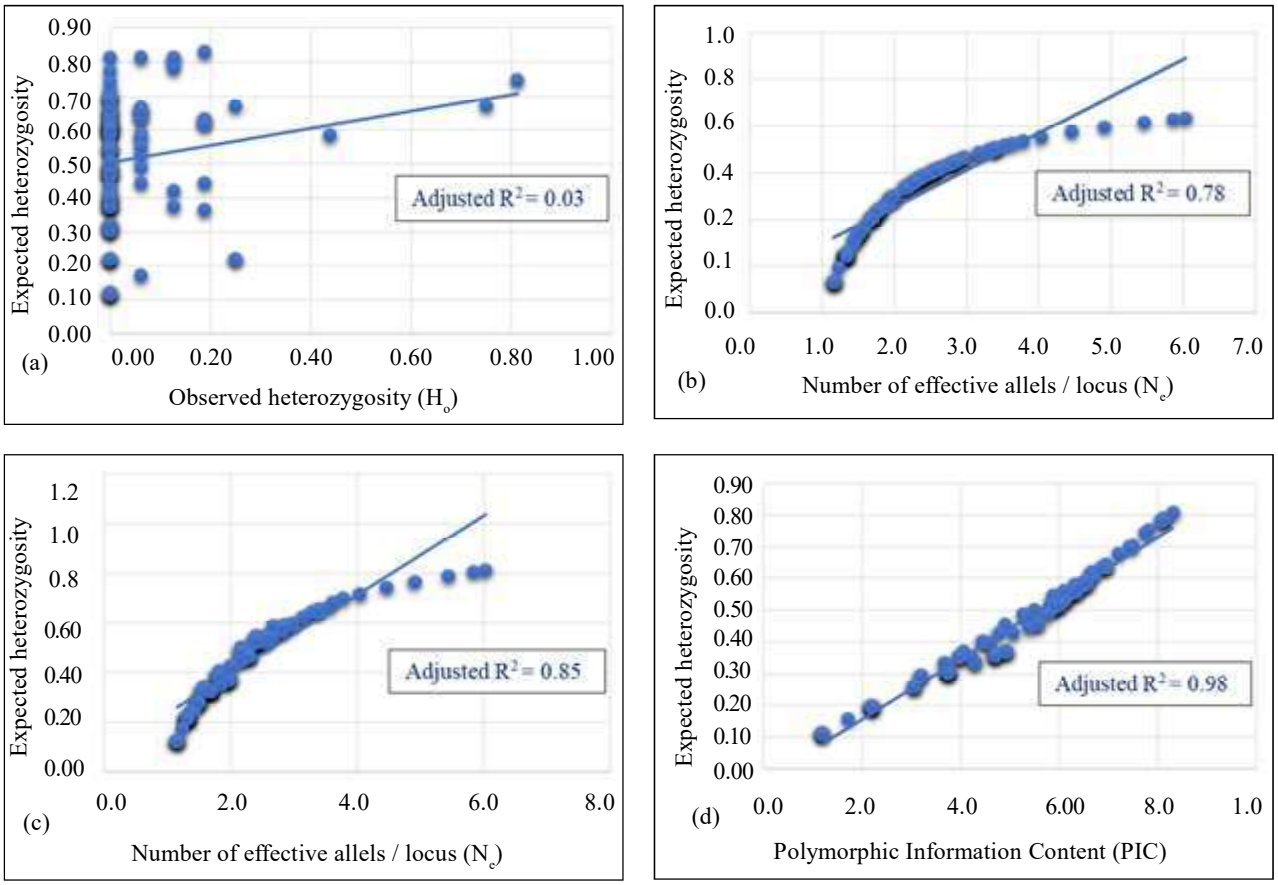

Fig. 2 : Relationships exists between different population genetic parameters. (a) correlation between observed heterozygosity and expected heterozygosity, (b) correlation between effective number of alleles and expected heterozygosity, (c) correlation between effective number of alleles with PIC values and (d) correlation between expected heterozygosity and PIC

TABLE 3

Comparison of population genetic parameters estimated using bi-allelic and multi-allelic SSR markers

SSR markers	Number of markers	Average Effective alleles/ locus (N_e)	Average observed heterozygosity (H_o)	Average expected heterozygosity (H_e)	Average unbiased heterozygosity (uH_e)	Average Polymorphic Information Content (PIC)
Bi-allelic marker	67	1.92	0.03	0.44	0.45	0.38
Multi-allelic marker ($\geq$ 3 allele)	33	3.37	0.10	0.68	0.70	0.63

Core Set of SSR Markers

All population genetic parameters estimated based on multi-allelic markers were higher in magnitude than those estimated based on bi-allelic markers (Table 3). Further, the estimates of PIC and H_e based on multi-allelic markers were almost one and half a times greater than those based on bi-allelic markers. These multi-allelic markers could be regarded a score set of markers. Being multi-allelic, the core set of markers are hyper variable markers and hence exhibit greater ability to discriminate test genotypes. For the same reason, they are easy to assay and resolve even in simple agarose gel (Singh *et al.,* 2012 and Dutta *et al.,* 2013). Fewer core set of markers capture a greater /at least as much genetic diversity as could be captured by a greater number of less informative (low PIC) markers. Hence, these core set of markers could be preferentially used to assess genetic diversity of any set genotypes with high probability of detecting polymorphism.

Genetic Dissimilarity between Pairs of Elite Inbred Lines

The estimates of dissimilarity coefficients between pairs 16 of elite inbred lines ranged from 0.35 to 0.71 with an average of 0.54 (Table 4 & Fig. 3). Josia *et al.* (2021) also reported a wide range of genetic dissimilarity coefficients ranging from 0.05 to 0.56

TABLE 4

Estimates of dissimilarity coefficient matrix of pairs of maize elite inbred lines used in the study

	40061	CAL 1443	CML 395	CML 451	CML 568	CML 581	CML 582	MAI 105	MAI 194	MAI 21	MAI 214	MAI 345	MAI 349	NAI 137	SKV 50	V 70
40061	0.00															
CAL 1443	0.65	0.00														
CML 395	0.58	0.52	0.00													
CML 451	0.65	0.41	0.60	0.00												
CML 568	0.61	0.57	0.47	0.64	0.00											
CML 581	0.59	0.58	0.53	0.60	0.44	0.00										
CML 582	0.53	0.50	0.49	0.53	0.44	0.49	0.00									
MAI 105	0.56	0.51	0.46	0.56	0.49	0.49	0.45	0.00								
MAI 194	0.58	0.57	0.51	0.69	0.53	0.54	0.52	0.53	0.00							
MAI 21	0.47	0.54	0.54	0.56	0.66	0.59	0.58	0.56	0.54	0.00						
MAI 214	0.60	0.60	0.51	0.61	0.58	0.54	0.54	0.44	0.38	0.53	0.00					
MAI 345	0.63	0.57	0.50	0.65	0.54	0.52	0.58	0.57	0.47	0.55	0.45	0.00				
MAI 349	0.62	0.49	0.38	0.63	0.46	0.48	0.50	0.52	0.52	0.52	0.47	0.39	0.00			
NAI 137	0.52	0.54	0.47	0.52	0.49	0.44	0.43	0.35	0.60	0.56	0.52	0.58	0.50	0.00		
SKV 50	0.71	0.42	0.55	0.55	0.59	0.56	0.52	0.51	0.61	0.59	0.58	0.59	0.56	0.49	0.00	
V 70	0.53	0.56	0.53	0.61	0.59	0.61	0.55	0.52	0.46	0.53	0.48	0.44	0.48	0.58	0.67	0.00

in maize. Fairly a good number (17) of pairs of elite inbred lines with dissimilarity coefficients more than 0.60 (Table 5) could be identified. Identification of such a large number of diverse pairs of inbred lines is significant given that the inbred lines used in the present study are elite ones and are likely to share large segments of similar genomic regions and are likely to be fixed for alleles at a large number of quantitative traits as well. Evidences from the literature point out that SSRs are non-randomly distributed across the genome of crops including maize. Their wide-spread distribution in upstream promoter regions is known to affect loci controlling quantitative traits. Thus, SSRs provide a significant source of mutation affecting loci controlling quantitative traits (Li *et. al.,* 2004 and Kashi & King, 2006).

Further, given that the SSR markers chosen for the study are distributed evenly across all the chromosomes, we assume that these markers represent entire genome. This assumption follows that the pairs of elite inbred lines with high magnitudes of marker-based dissimilarity coefficients are likely to differ in the frequency of the alleles at loci controlling quantitative traits as well (Burstin and Charcosset, 1997). If HGs have been well defined on the basis of empirical studies as is the true in our study, marker-based dissimilarity indices help select most potential pairs of inbred lines within each HG for developing BPs and select most appropriate inbred line (s) from opposite HG as tester (s) to identify desirable new inbred lines derived from BPs for use in developing heterotic hybrids. Marker-based dissimilarity indices have been effective for selection of phenotypically diverse inbred lines for developing BPs and testers for use in hybrid cultivar breeding in maize (Melchinger *et al.,* 1992). These reports lend adequate support for selection of inbred lines based on marker-based dissimilarity indices. Eight of the 17 pairs of inbred lines with dissimilarity coefficients more than 0.64 (Table 5) could be considered as most putative parents for developing BPs. In particular, 40061 and CML 451 frequently appeared as one of parents among these eight pairs of inbred lines. Siddu *et al.* (2023) reported that 40061 was one of the ten inbred lines that harbour superior dominant favourable alleles

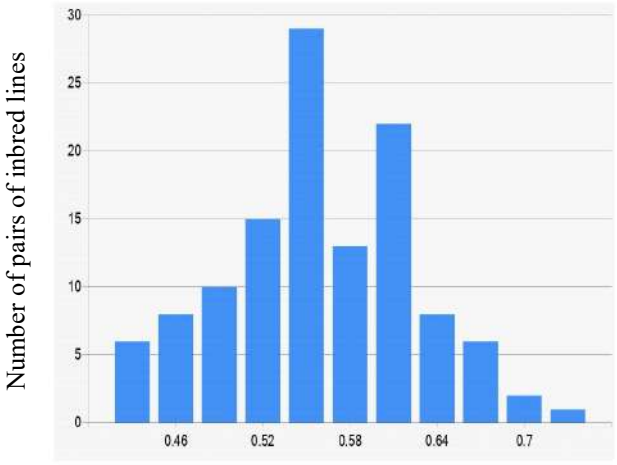

Fig. 3: Histogram depicting frequency of pairs of maize inbred lines with different magnitude of dissimilarity coefficient

TABLE 5

List of maize elite inbred lines with genetic dissimilarity coefficient of ≥ 0.60

Pair of inbred lines		Dissimilarity coefficients
40061	SKV 50	0.71
CML 451	MAI 194	0.69
SKV 50	V 70	0.67
CML 568	MAI 21	0.66
40061	CAL 1443	0.65
40061	CML 451	0.65
CML 451	MAI 345	0.65
CML 451	CML 568	0.64
40061	MAI 345	0.63
CML 451	MAI 349	0.63
40061	MAI 349	0.61
40061	CML 568	0.61
CML 451	V 70	0.61
CML 451	MAI 214	0.61
CML 451	V 70	0.61
SKV 50	MAI 194	0.61
40061	MAI 214	0.60

TABLE 6

List of the best triplet combination of inbred lines identified for each heterotic group (HG)

Pair of inbred lines with high marker dissimilarity from HG 1	Tester parent from HG 2	Tester parent from HG 3
MAI 137 and MAI 194	MAI 21	CML 568
MAI 137 and V 70	MAI 21	CML 568
MAI 137 and MAI 345	40061	CML 568
MAI 137 and MAI 214	40061	CML 568
MAI 137 and MAI 349	40061	CML 568
Pair of inbred lines with high marker dissimilarity from HG 2	Tester parent from HG 1	Tester parent from HG 3
MAI 105 and 40061	MAI 345	CML 568
MAI 105 and MAI 21	MAI 345	CML 568
Pair of inbred lines with high marker dissimilarity from HG 2	Tester parent from HG 1	Tester parent from HG 3
CML 395 and CML 582	V 70	MAI 21
CML 395 and CML 568	V 70	MAI 21

not present in parents of two elite SCHs namely, Hema and Nithyashree. It is therefore advantageous to involve 40061 as one of the parental inbred lines to develop new improved inbred lines for use in developing new heterotic hybrids. 40061 is also suggested for use as donor inbred lines to improve parents of two SCHs and hence to improve them for per se performance (Siddu *et al.,* 2023).

Identification of Triplet Combination of Inbred Lines for use in Hybrid Cultivar Breeding

The old adage in plant breeding is to cross elite × elite parents. This adage underscores the importance of selection of parents to develop BPs. The desirable BPs are the ones which have high trait mean and variance. The chance of developing superior inbred lines for use as parents of heterotic hybrids is enhanced by starting with such desirable BPs. It is assumed that desirable alleles at loci controlling major traits are fixed, while those at loci controlling quantitative traits still segregate and are randomly distributed among the elite inbred lines (Falk, 2010 and Bernardo, 2020). Crossing such elite × elite inbred lines are likely to result in new inbred lines with superior combination of alleles better than the parents. The BPs derived from elite inbred lines with fewer defects have been the major sources of high performing hybrid cultivars and constitute score genetic resources for developing new improved inbred lines in hybrid breeding programmes (Falk, 2010 and Bernardo, 2020).

Quantitative genetic theory (Bernardo, 2023) suggests that in addition to two inbred parents to develop BPs to derive new inbred lines, tester is the third inbred parent required in hybrid breeding. As F_1 hybrid are the cultivar types used for commercial production, the new inbred lines derived from BPs need to be evaluated for their performance in hybrid combinations. The tester is therefore required as a third inbred line partner (along with the two inbred parents) to evaluate new inbred lines in their test cross (TC) hybrid combinations. While both the inbred parents of BPs should be selected from the same HG, the tester should be selected from an opposite HG to maintain heterotic patterns. An ideal tester is the one which maximise both mean and

variance of TC hybrid combinations. The ideal triplet combination of two inbred parents + tester inbred lines is the one whose average TC hybrids of the two inbred parents with the tester is highest. Considering these quantitative genetics concepts of hybrid breeding, five pairs of inbred linesin HG 1 and two each in other two HGs with relatively high dissimilarity coefficients between and high average dissimilarity coefficients between the parental inbred linesand thetester wereidentified as the most desirable triplet combination of inbred lines (Table 6) for use in developing heterotic hybrids.

The 16 elite inbred lines differed at 100 out of 132 SSR markers. Thus, 75 per cent of the assayed markers were polymorphic. Thirty-three of these 100 polymorphic markers with more than threealleles were highly informative. These 33 multi-allelic markers showed greater ability to discriminate even the16 elite inbred lines which are likely to share large segments of similar genomic regions. These 33 multi-allelic markers are suggested for use as core set of markers for their preferentially use in assessing diversity among any chosen set of test genotypes. Dissimilarity indices of the order of more than 0.64 between a good number of pairs of elite inbred lines were identified. Using these dissimilarity indices, triplet combinations of elite pairs inbred lines + tester for use in developing heterotic maize hybrids were identified.

Acknowledgements : The senior author gratefully acknowledges Council of Scientific and Industrial Research (CSIR), New Delhi, India for providing Junior Research Fellowship (JRF) vide No. 09/ 0271(12086)/2021-EMR-1 dated 01-05-2021 for pursuing Ph.D. degree program at University of Agricultural Sciences, Bangalore, India.

References

BANTTE, K. AND PRASANNA, B, M., 2003, Simple sequences repeat polymorphism in quality protein maize (QTL) lines. *Euphytica,* **129** : 337 - 344.

BASHALKHANOV, S., PANDEY, M. AND RAJORA, O. P., 2009, A simple method for estimating genetic diversity in large populations from finite sample sizes. *BMC Genet.*, **10** : 1 - 10.

BERNARDO, R., 2020, Breeding for quantitative traits in plants. Third Edition. Stemma Press. Woodbury, Minnesota, USA.

BERNARDO, R., 2023, Essentials of plant breeding. Second edition. Stemma Press. Woodbury, Minnesota, USA.

BOTSTEIN, D., WHITE, R. L., SKOLNICK, M. AND DAVIS, R. W., 1980, Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.,* **32** (3) : 314.

BURSTIN, J. AND CHARCOSSET, A., 1997, Relationship between phenotypic and marker distances: theoretical and experimental investigations. *Hered.,* **79** (5) : 477 - 483.

DEEPAK, T. AND VASUDEVAN, S., 2023, Phenological studies on parental lines of promising maize hybrids for their seed production potentiality. *Mysore J. Agric. Sci.,* **57** (1) : 82 - 88.

DUTTA, S., MAHATO, A. K., SHARMA, P., RANJEET, S. R., SHARMA, T. R. AND SINGH, N. K., 2013, Highly variable Arhar simple sequence repeat markers for molecular diversity and phylogenetic studies in pigeon pea (*Cajanus cajan* (L.) Milli sp.). *Plant Breed.,* **132** : 1439 - 0523. http://doi: 10.1111/pbr.12014.

FALK, D. E., 2010, Generating and maintaining diversity at the elite level in crop breeding. *Genome,* **53** : 982 - 991.

GEMAYEL, R., CHO, J., BOEYNAEMS, S. AND VERSTREPEN, K. J., 2012, Beyond junk-variable tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. Genes, **3** (3) : 461 - 480.

GUO, X. AND ELSTON, R., 1999, Linkage information content of polymorphic genetic markers. *Hum. Hered.,* **49** (2) : 112 - 118. https://doi.org/10.1159/000022855.

HARTL, D. L. AND CLARK, A. G., 1997, Principles of population genetics. Sunderland: Sinauer associates, Massachusetts.

HILDEBRAND, C. E., DAVID, C., TORNEY, C. AND WAGNER, P., 1994, Informativeness of polymorphic DNA markers. *Los Alamos Sci.,* **20** : 100 - 102.

HOISINGTON, D., KHAIRALLAH, M. AND GONZALEZ-DE-LEON, D., 1994, Laboratory Protocols. 2nd ed., CIMMYT applied molecular genetics laboratory. Mexico, D. F., CIMMYT.

JOSIA, C., MASHINGAIDZE, K., AMELEWORK, A. B., KONDWAKWENDA, A., MUSVOSVI, C. AND SIBIYA, J., 2021, SNP-based assessment of genetic purity and diversity in maize hybrid breeding. PLoS One, 16(8), e0249505.

KASHI, Y. AND KING, D. G., 2006, Simple sequence repeats as advantageous mutators in evolution. Trends Genet., **22** (5) : 253 - 259.

KUMAR, B., CHOUDHARY, M., KUMAR, P., KUMAR, K., KUMAR, S., SINGH, B. K., LAHKAR, C., MEENAKSHI, KUMAR, P., DAR, Z. A., DEVLASH, R., HOODA, K. S., GULERIA, S. K. AND RAKSHIT, S., 2022, Population structure analysis and association mapping for turcicum leaf blight resistance in tropical maize using SSR markers. Genes, **13** (4) : 618.

LAURENTIN, H., 2009, Data analysis for molecular characterization of plant genetic resources. Genet. Resour. Crop Evol., **56** (2): 277 - 292. https://doi.org/10.1007/s10722-008-9397-8

LI, Y. C., KOROL, A. B., FAHIMA, T. AND NEVO, E., 2004, Microsatellites within genes: structure, function, and evolution. Mol. Biol. Evol., **21** (6) : 991 - 1007.

LIU, K. AND MUSE, S. V., 2005, Power Marker: an integrated analysis environment for genetic marker analysis. Bioinformatics, **21** (9) : 2128 - 2129. https://doi.org/10.1093/bioinformatics/bti282

MELCHINGER, A. E., BOPPENMAIER, J., DHILLON, B. S., POLLMER, W. G. AND HERRMANN, R. G., 1992, Genetic diversity for RFLPs in European maize inbreds: II. Relation to performance of hybrids within versus between heterotic groups for forage traits. *Theor. Appl. Genet.,* **84** : 672 - 681.

NEI, M., 1978, Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genet.,* **89** (3) : 583 - 590.

NEI, M.,1987, Molecular evolutionary genetics. Columbia university press.

NEPOLEAN, T., SINGH, I., HOSSAIN, F., PANDEY, N. AND GUPTA, H. S., 2003, Molecular characterization and assessment of genetic diversity of inbred lines showing variability for drought tolerance in maize. *J. Plant Biochem. Biotechnol.,* **22** (1) : 71 - 79.

NYALIGWA, L., HUSSEIN, S., AMELEWORK, B. AND GHEBREHIWOT, H., 2015, Genetic diversity analysis of elite maize inbred lines of diverse sources using SSR markers. Maydica, **60** (3) : 29 - 36.

RANATUNGA, M. A. B., MEENAKSHISUNDARAM, P., ARUMUGACHAMY, S. AND MAHESWARAN, M., 2009, Genetic diversity analysis of maize (*Zea mays* L.) inbreds determined with morphometric traits and simple sequence repeat markers. Maydica; **54** : 113 - 123.

SAITOU, N. AND NEI, M., 1987, The neighbor-joining method: a new method for reconstructing phylo genetic trees. *Mol. Biol. Evol.,* **4** (4): 406 - 425. https://doi.org/10.1093/oxfordjournals.molbev.a040454.

SELVI, A., NAIR, N. V., BALASUNDARAM, N. AND MOHAPATRA, T., 2003, Evaluation of maize microsatellite markers for genetic diversity analysis and fingerprinting in sugar cane. Genome, **46** : 394 - 403.

SENIOR, M. L., MURPHY, J. P., GOODMAN, M. M. AND STUBER, C. W., 1998, Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. *Crop Sci.,* 1998; **38** :1088 - 1098.

SHEHATA, A. I., AL-GHETHAR, H. A. AND AL-HOMAIDAN, A. A., 2009, Application of simple sequence repeat (SSR) markers for molecular diversity and heterozygosity analysis in maize inbred lines. *Saudi J. Biol. Sci.,* **16** : 57 - 62.

SIDDU, C. B., RAMESH, S., KALPANA, M. P., BASANAGOUDA, G., SATHISH, H., ANIKETH, S., KARTHIK, N., SINDHU, D., KEMPARAJU, M., KIRANKUMAR, R., CHANDANA, B. R., SINCHANA, B. V., MAHADEVU, P., BALASUNDARA, D. C. G. AND MALLIKARJUNA, N., 2023, Identification of inbred lines harbouring favourable dominant alleles not present in the parents of three elite maize (*Zea mays* L.) single cross hybrids. Genet. Resour. Crop. Evol. https://doi.org/10.1007/s10722-023-01666-w

SINGH, N. K., GUPTA, D. K., JAYASWAL, P. K., MAHATO, A. K., DUTTA, S., SINGH, S., BHUTANI, S., DOGRA, V., SINGH, B. P., KUMAWAT, G., PAL, J. K., PANDIT, A., SINGH, A., RAWAL, H., KUMAR, A., PRASHAT, G. R., KHARE, A., YADAV, R., RAJE, R.S., SINGH, M.N., DATTA, S., FAKRUDIN, B., WANJARI, K.B., KANSAL, R., DASH, P. K., JAIN, P. K., BHATTACHARYA, R., GAIKWAD, K., MOHAPATRA, T., SRINIVASAN, R. AND SHARMA, T. R., 2012, The first draft of the pigeonpea genome sequence. *J. Plant Biochem. Biotechnol.* **21** : 98 - 112. doi: 10.1007/s13562-011-0088-8.

SMOUSE, R. P. AND PEAKALL, R., 2012, GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics,* **28** (19) : 2537 - 2539. https://doi.org/10.1111/j.1471-8286.2005.01155.x.

SSERUMAGA, J. P., MAKUMBI, D., JI, H., NJOROGE, K., MUTHOMI, J. W., CHEMININGWA, G. N., SI-MYUNG, L., ASEA, G. AND KIM, H., 2013, Molecular characterization of tropical maize inbred lines using microsatellite DNA markers. *Maydica*, **59** : 267 - 274.

VATHANA, Y., SA, K. J., LIM, S. E. AND LEE, J. K., 2019, Genetic diversity and association analyses of Chinese maize inbred lines using SSR markers. *Plant Breed. Biotech.,* **7** (3) : 186 - 199.

VIEIRA, M. L., SANTINI, L., DINIZ, A. L. AND MUNHOZ, C. D., 2016, Microsatellite markers: what they mean and why they are so useful. *Genet. Mol. Biol.,* **39** (3) : 312 - 328. https://doi.org/10.1590/1678-4685-GMB-2016-0027.