

Research Article

*E. Lamalakshmi Devi and Ngangkham Umakanta would like to share the first authorship together.

Cite this article: Lamalakshmi Devi E *et al.* (2023). Genetic diversity and population structure analysis in early generations maize inbreds derived from local germplasm of Eastern Himalayan regions using microsatellite markers. *Plant Genetic Resources: Characterization and Utilization* **21**, 418–425. <https://doi.org/10.1017/S1479262123000862>

Received: 12 January 2023

Revised: 25 October 2023

Accepted: 30 October 2023

First published online: 29 November 2023



Keywords:

genetic diversity; maize; marker-trait association; population structure; SSR markers

Corresponding authors:

E. Lamalakshmi Devi;
Email: elangbamlama@gmail.com;
Sunil Kumar Chongtham;
Email: sunil.sunil.ch@gmail.com

Genetic diversity and population structure analysis in early generations maize inbreds derived from local germplasm of Eastern Himalayan regions using microsatellite markers

E. Lamalakshmi Devi^{1,*} , Umakanta Ngangkham^{2,*}, Sunil Kumar Chongtham³ , Bhuvaneshwari S⁴, Ingudam Bhupenchandra⁵, Konsam Sarika⁶, Harendra Verma⁷, Akoijam Ratankumar Singh⁸, Amit Kumar⁹, Tensubam Basanta Singh¹⁰, Amit Kumar¹¹, T. L. Bhutia¹², S. K. Dutta¹³, Shaon Kumar Das¹⁴, Ramgopal Devadas¹, Ayam Gangarani Devi¹⁵, S. P. Das¹⁶, Ch. Chinglen Meetei⁶, I. Meghachandra Singh¹⁷ and V. K. Mishra¹⁸

¹Department of Genetics and Plant Breeding, ICAR-RC-NEH Region, Sikkim Centre, Tadong, Sikkim, India;

²Department of Biotechnology, ICAR-RC-NEH Region, Manipur Centre, Lamphelpat, Manipur, India; ³Department of Natural Resource Management, College of Horticulture, Central Agricultural University, Bermoik, Sikkim, India;

⁴Division of Crop Improvement, ICAR-Indian Institute of Vegetable Research, Varanasi, Uttar Pradesh, India;

⁵Department of Soil Science, ICAR-KVK Tamenglong, ICAR RC NEH Region, Manipur Centre, Lamphelpat, Manipur, India; ⁶Department of Genetics and Plant Breeding, ICAR RC NEH Region, Manipur centre, Lamphelpat, Manipur, India; ⁷Department of Genetics and Plant Breeding, ICAR RC NEH Region, Nagaland centre, Medziphema, Nagaland, India; ⁸Department of Plant Pathology, ICAR RC NEH Region, Manipur centre, Lamphelpat, Manipur, India; ⁹Department of Agronomy, ICAR RC NEH Region, Sikkim centre, Tadong, Sikkim, India; ¹⁰Department of Soil Science, ICAR RC NEH Region, Manipur centre, Lamphelpat, Manipur, India; ¹¹Division of Crop Science, ICAR-RC-NEH Region, Umiam, Meghalaya, India; ¹²Department of Vegetable Science, ICAR RC NEH Region, Sikkim centre, Tadong, Sikkim, India; ¹³Department of Horticulture, ICAR RC NEH Region, Sikkim centre, Tadong, Sikkim, India; ¹⁴Department of Soil Science, ICAR RC NEH Region, Sikkim centre, Tadong, Sikkim, India; ¹⁵Department of Plant Physiology, ICAR-RC-NEH Region, Tripura centre, Lembucherra, Tripura, India; ¹⁶Director, ICAR- NRC for Orchid, Pakyong, Sikkim, India; ¹⁷Department of Seed Science, ICAR RC NEH Region, Manipur centre, Lamphelpat, Manipur, India and ¹⁸Director, ICAR-RC-NEH Region, Umiam, Meghalaya, India

Abstract

The North-Eastern region (NER) of India falls under the Eastern Himalayan region and it is a bio-diversity hub. Diverse maize landraces with wide adaptability to extreme climatic and soil scenario like heavy rainfall, drought and acidic soil conditions have been grown in NER since time immemorial. However, maize diversity in NER region has drastically reduced due to introduction of high yielding varieties and hybrids. Modern maize breeding programmes are focused on high yield but other unique traits like stay green trait, prolificacy (more than one fertile ear per plant), self-fertilizing ability are also important and the local germplasm of the NER region can contribute with these unique traits. Prior to the selection of any lines in several breeding programmes, assessment of genetic diversity and population structure are basic requirements. Hence, in the present study assessment of genetic diversity and population structure study in 30 maize inbreds developed from different germplasm of NER was undertaken using SSR markers, selected for their broad distribution throughout the genome, in order to assess the extent of allelic diversity among the lines and whether any population structure could be established. In addition to assessing molecular diversity, the study aims to evaluate the potential for yield and other beneficial and unique alleles that have high potential to contribute in the genetic enhancement programme of maize.

Introduction

Maize is the ‘Queen of Cereals’ and it is used as food, fodder and feed apart from various industrial purposes. The Eastern Himalayan region, mainly the North-Eastern region (NER) of India, is a diverse hub for various crops, including maize. Maize in this region has acquired different unique characteristics like coloured grain having a diverse range from purple, blue, black, white, cream, yellow, orange, red, and variegated; multi-cob bearing (3–6 cobs per plant); mucus-producing roots having the potential to fix atmospheric nitrogen, stay green trait, and waxiness having high amylopectin content. In addition to the above traits, the local maize has high potential as fodder due to its vigorous growth and tillering habit which makes them unique. Local maize has tremendous morphological variability such as



plant height, plant type, duration, kernel colour, texture, nutrient composition, and appearance. (Singh *et al.*, 2019). The plant type of local landraces ranges from wild primitive types to advanced landraces/varieties maintained and selected by farmers from time immemorial.

As agriculture is at a subsistence level in the NER region, farmers still grow traditional landraces for family/community needs. The region is characterized by high rainfall and acidic soils with temperatures varying from temperate to sub-tropical. The maize is grown mainly in hill ecology with slash and burn system of cultivation. The weather conditions in the North Eastern Region (NER) of India are erratic; drought and floods are common phenomena in the region. The local maize germplasm has evolved in balance with the prevalent agroecological conditions of the region and, therefore, they are expected to possess better adaptation to regional stresses.

The advent of molecular markers provides a key breakthrough to surmount the key drawback of classical crop improvement programmes (Devi *et al.*, 2017b; Hasan *et al.*, 2021) as well as provide a platform for an efficient assessment of diversity present within and between plant populations (Boakyewaa *et al.*, 2019; Borah *et al.*, 2021). They are preferred over the morphological or biochemical markers as they are not affected by environmental factors and by the developmental phase of the plant (Badu-Apraku and Fakorede, 2017). Among the various types of molecular markers, microsatellites or simple sequence repeats (SSRs) are considered sufficient for genetic diversity assessment (Laosatit *et al.*, 2022; Islam *et al.*, 2023) due to their high level of polymorphism, abundance, reproducibility, cost effective and amenability to automation (Govindaraj *et al.*, 2015). Studies on the comparative utilization of different molecular markers have publicized that SSRs with moderate density are more effective in assessing the molecular diversity study in maize than SNPs (Jones *et al.*, 2007; Yang *et al.*, 2011; Adu *et al.*, 2019). Likewise, a study by Singh *et al.* (2013) in a series of Indian rice varieties concluded that for population structure study SNP markers showed better genetic relatedness with more population number. However, for the diversity studies SSR showed better grouping samples even at trait level.

Information on the genetic variability and population structure of the base population is a pre-requisite and crucial step for undertaking a successful crop improvement programme (Boakyewaa *et al.*, 2019; Hasan *et al.*, 2021). In the past few years, crop diversity has been intensely reduced due to genetic erosion. Therefore, indigenous germplasm has been given preference for the development of mapping populations, heterotic pools and inbreds (Yadav *et al.*, 2015). Conservation and utilization of diverse germplasm for the introgression of unique and desirable genes into the genetic base is essential.

Population structure is an important tool to understand the evolution of plants and the introduction of molecular techniques and DNA sequencing analysis has led to significant improvements in the insights into the genetics of natural populations (Sbordoni *et al.*, 2012; Boukteb *et al.*, 2021). Therefore, the objective of the present study was to (i) examine the morphological and genetic diversity of maize inbred lines derived from different source germplasm of eastern Himalayas using SSR markers; (ii) genetic relationships analysis among early generation maize inbreds; (iii) population structure analysis and allelic patterns across populations; (iv) studying the correlation and linkage of major QTLs with agronomic traits.

Materials and methods

Plant materials

A set of 30 indigenous early-generation maize inbred lines were selected based on the phenotypic diversity, grain colour and unique traits like mucous producing roots (self-fertilizing ability), prolificacy or multiple cobs bearing and stay green trait to assess the genetic diversity and study the population structure. These early-generation lines were developed through continuous inbreeding accompanied by selection of desirable traits at ICAR-Research Complex for North Eastern Hill Region (ICAR-RC-NEHR), Manipur Centre using local germplasm collected from different places of Manipur and Mizoram states of India.

Genomic DNA isolation and PCR amplification

Genomic DNA was extracted from the leaf of the selected genotypes using a modified CTAB extraction protocol (Porebski *et al.*, 1997). DNA concentration was measured on Genova Nano Analyzer UV/VIS (Mumbai, India). The final concentration of extracted genomic DNA was adjusted to 10 ng/ μ l and used for the PCR amplification. A set of 52 polymorphic Simple Sequence Repeats' (SSR) markers, selected on the basis of different bin locations to maximize the genome coverage were used to study the allelic diversity and population structure of the maize lines. The primer sequences were collected from the maize genomic database (<http://www.maizegdb.org/>) and the oligonucleotide primers were synthesized in purified and lyophilized form from GCC Biotech India Pvt. Ltd. (Kolkata, India).

Polymerase chain reactions were carried out on 96-well thermal cycler, Eppendorf-Mastercycler Nexus Gradient (Hamburg, Germany). All PCRs were performed in a final reaction volume of 20 μ L with the following reagent components: 2 μ l of template DNA (20 ng/ μ l), 10 μ l 2X Hi-Primer Taq PCR Master Mix (GCC Biotech India Pvt. Ltd., India), 1 μ l of each forward and reverse primers and 6 μ l RNA grade nuclease-free water (GCC Biotech India Pvt. Ltd., India). The amplification conditions were set as follows: initial denaturation at 95°C for 5 min, four cycles consisting of denaturation at 95°C for 45 s, primer annealing ranged between 57 and 60°C for 45 s, primer extension at 72°C for 45 s and then 35 cycles consisting of denaturation at 95°C for 45 s, primer annealing 60°C for 45 s, primer extension at 72°C for 45 s and final extension at 72°C for 8 min. The amplified PCR products were resolved using 3.0% ISSR Agarose (GCC Biotech India Pvt. Ltd., India) stained with ethidium bromide, documented under the gel doc system (Gel DocTM XR+, Biorad, India).

Agronomic performance of inbreds

Each of the 30 maize inbreds (details given in online Supplementary Table S1) was grown in two rows of 3 m length each during *spring* 2019 and 2020 at the Research Farm of the ICAR-RC-NEHR, Manipur Centre, Lamphelpat and *spring* 2021 at the Research Farm of the ICAR-RC-NEHR, Sikkim Centre, Tadong following the standard package of practices for maize cultivation. The row-to-row and plant-to-plant distances were maintained at 60 and 20 cm, respectively. The individual plants were carefully selfed and harvested to avoid contamination of inbreds. These early generation inbreds were evaluated for plant height (PH), ear height (EH), number of kernel rows per ear (KRPE),

number of kernels per row (KPR), cob length (CL), cob diameter (CD) and 100 kernel weight (KW).

Statistical analyses

The genotypic data of 52 SSR markers were used for the estimation of genetic diversity parameters such as number of alleles per locus (K), observed heterozygosity (H_{Obs}), expected heterozygosity (H_{Exp}) and polymorphic information content (PIC) using the Cervus 3.0 program (Field Genetics Ltd., London, England). GenAlEx 6.502 software (Peakall and Smouse, 2012) was used to compute the principal coordinate analysis (PCoA) by plotting the eigenvector values in a scatter graph taking the first principal component and the second principal component as the axes. A hierarchical analysis of molecular variance (AMOVA) of 30 genotypes between the populations was also conducted using GenAlEx 6.502 software (Peakall and Smouse, 2012). To determine the presence of the genetic structure, STRUCTURE version 2.3.4 (Pritchard *et al.*, 2000) was run with the number of clusters from $K=1$ to $K=5$, with five independent replications per K using the admixture model and correlated allele frequencies, a 50,000 burn-in period and 50,000 MCMC to determine approximations of posterior distributions. The optimum K value was determined through the STRUCTURE HARVESTER programme (Earl and Vonholdt, 2012) by estimating the peak value of ΔK using the method of Evanno *et al.* (2005). The STRUCTURE analyses give a Q-matrix (QST) for each value of K , which lists the estimated membership coefficients for each genotype in each subgroup. An individual has more than 70% of its genome fraction value under a particular K subgroup was assigned to that subgroup. The allele size was evaluated by comparing it with a 100 bp DNA ladder (GCC Biotech India Pvt. Ltd., India) and a phylogenetic tree was constructed using the neighbour-joining method implemented in DARwin-6.0 (Perrier and Jacquemoud-Collet, 2006).

All the statistical analysis of agronomic data was carried out using XLSTAT v2020.4.1 software (www.xlstat.com). The association analysis of seven agronomic traits along with the 52 SSR markers was calculated using GGT 2.0 software (van Berloo, 2008). The association analysis was calculated based on the squared correlation coefficients between the marker data and trait values (R^2 values) and the associated probabilities of the correlation values. The threshold to declare a significant marker-trait association was set to $-\log_{10}(P) = 3$.

Results

Genetic diversity of maize inbreds

In the present study, all 52 markers were successfully amplified in 30 maize-inbred lines (online Supplementary Fig. S1). A total of 189 alleles were generated across the 30 genotypes with a mean of 3.63 alleles per locus (online Supplementary Table S2). The allele size of the 52 SSR markers ranged from 50 bp (*phi 036*) to 295 bp (*p 101049*) which revealed a high level of genetic diversity among the loci. The PIC value of the markers varied from 0.17 (*umc 1622*) to 0.76 (*umc 1153*) with an average value of 0.49 (online Supplementary Table S2). Out of the 52 SSR markers, 24 SSR loci (46.15%) were found to be more informative by having of PIC value of ≥ 0.50 , depicting the higher capability of these loci to differentiate between the genotypes, determine genetic diversity, gene introgression, genetic conservation, etc. The

number of alleles per locus (K) ranged from 2 to 8 with an average of 3.634 ± 1.428 . The *bnlg1031* marker mapped at 8.06 bin showed the highest number of alleles (eight). The value of expected heterozygosity (H_{Exp}) ranged from 0.19 to 0.80, whereas the observed heterozygosity (H_{Obs}) ranged from 0 to 0.89.

The heterozygosity percentage ranged from 0 (SCM 15/9LC) to 27.45 (GP H 6) (online Supplementary Table S3). The graphical representation of heterozygosity in 30 maize inbreds is depicted in Fig. 1. Out of 30 inbred lines, 13 lines showed heterozygosity $\leq 10\%$, which means 43% of the lines attained more than 90% homozygosity. Another three inbred lines *viz.* GP (H) 38, MN 1, and SCM15/9LC showed heterozygosity of less than 5%, which revealed a high level of homozygosity in these lines. These three inbred maize lines may be considered pure lines which will be useful as parents for the development of hybrid, mapping population, genetic mapping, etc. In the present study, the AMOVA result revealed a higher molecular variance within individuals of the population (96%) than among individuals of the population (4%) (online Supplementary Table S4).

Genetic relationships analysis among early generation maize inbreds

The genetic dissimilarity between the genotype pairs ranged from 0.40 to 0.64 with a mean value of 0.57. The cluster diagram grouped the 30 maize genotypes into six distinct major clusters (Fig. 2). The lines developed from the same source population fall under the same cluster/sub-cluster. For instance, lines developed from populations collected from Manipur *viz.* GP (H) 2, GP (H) 28 and SCM 15/9LC were together in the same cluster *i.e.*, A. Likewise, lines like MZ 2, MZ 70, MZ 24, MZ 42, MZ 40, Kolasib 1(2) and MZ 46(B) developed from populations collected from Mizoram clustered together in the same sub-group *viz.*, B1.

The genotypic PCoA revealed that the lines were distributed in all four quadrangles, showing a high degree of genetic variability. A scatter plot generated from the PCoA analysis showed that the first two components accounted for 9.82% and 8.23% of the genetic variation which resulted in a total genetic variation of 18.05% (online Supplementary Fig. S2). Lines derived from the populations from Manipur were found to be closely placed in genotypic PCoA, while that from Mizoram occupied the nearby axes in the PCoA.

Population structure analysis and allelic patterns across populations

The peak plateau of ad hoc measure ΔK was detected at $K=3$ (online Supplementary Fig. S3), which indicated that the entire population can be grouped into three possible groups or sub-populations. The results of the population structure analysis were confirmed by the phylogenetic tree which divided the 30 inbred lines into three clusters and each cluster was further subdivided into sub-clusters. Population structures of 30 genotypes based on 52 SSR markers are shown in online Supplementary Fig. S3. Different colour within the group indicates the proportion of shared ancestry with another group that has the same colour as the admixture. The population structure group of 30 genotypes based on inferred ancestry values was also estimated (online Supplementary Table S3). The percentage of individuals in each population belonging to each group was estimated based on Q-matrix (QST) for each value of K and the cut-off probability

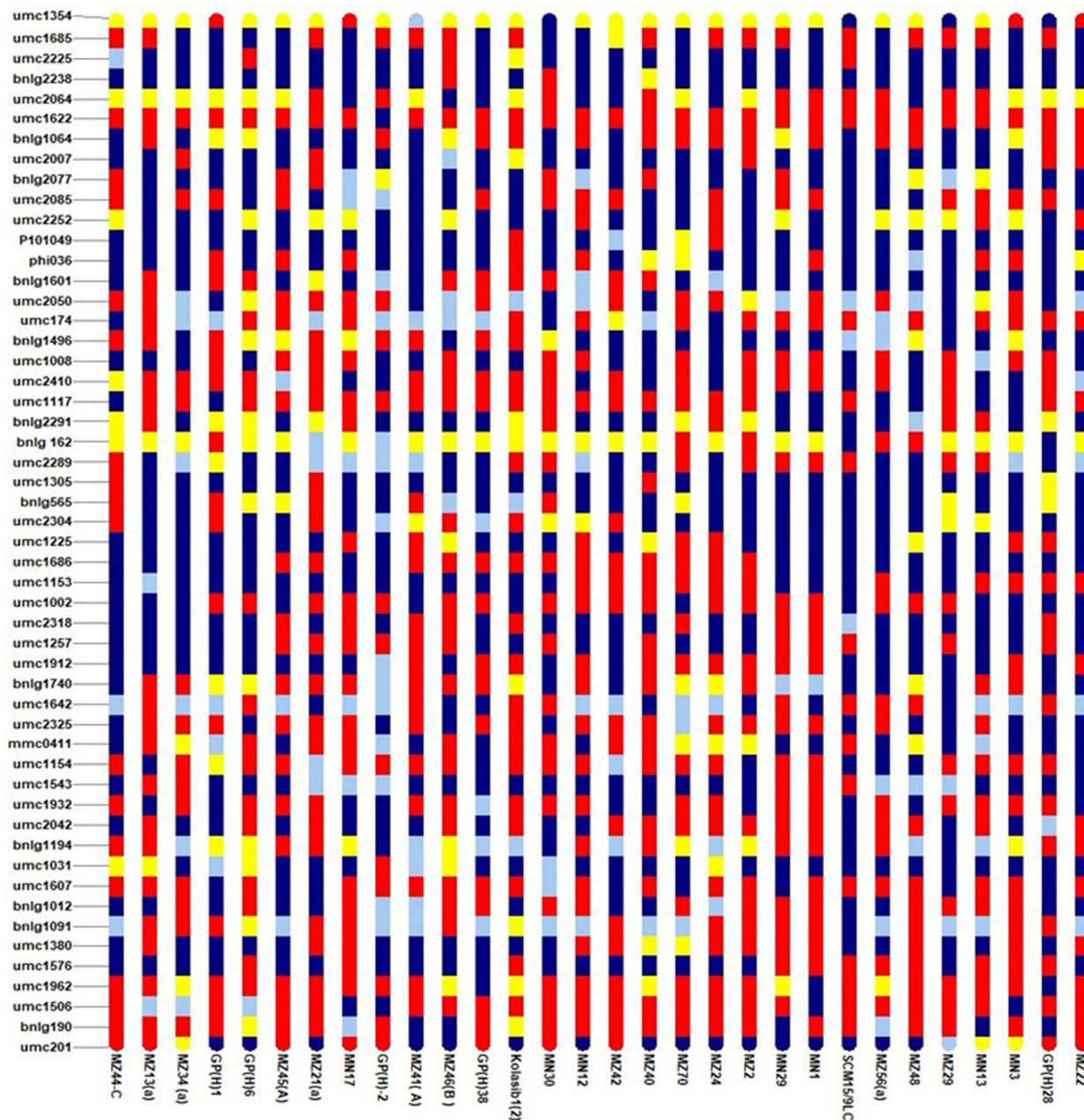


Figure 1. Depiction of allelic patterns of the 52 SSR markers among the 30 early maize inbreds lines through GGT 2.0 software (Van Berloo, 2008). The colour legends such as red, purple, yellow and grey denotes A, B, H (heterozygotic) and missing alleles, respectively.

for assigning to a group was assumed to be 70%. A set of 29 genotypes out of 30 showed the estimated membership coefficients of more than 70% of its genome fraction value and were therefore assigned to a specific group. Whereas, only one genotype was classified as admixture since it showed estimated membership coefficients of less than 70%. To conduct the analysis of molecular variance (AMOVA), the 30 maize inbred lines were grouped into two: Manipur (developed from the Manipur collection) and Mizoram (developed lines using the Mizoram collection). AMOVA revealed highly significant variance among populations and within individuals of the population (online Supplementary Table S4). The overall genetic differentiation indicated that 6% of the total variance is due to differences among populations, while 94% of the total molecular variance is accounted for

within-population molecular variance (online Supplementary Fig. S4).

The mean effective numbers of allele (N_e) were estimated to be 2.31 and 3.99 for Manipur (MAN) and Mizoram (MIZ) populations, respectively (online Supplementary Table S5). The number of alleles (allelic richness) in a population is an important and fundamental measure of genetic variation in a population and is a useful statistic for identifying populations for conservation. The Shannon’s information index (I) was calculated as 0.96 and 1.58 for MAN and MIZ populations, respectively. Among 189 alleles generated across the 30 genotypes, three private alleles (number of alleles unique to a single population) were detected in SSR marker *umc 1354*. These private alleles were revealed in three genotypes originating from Mizoram. A graph depicting

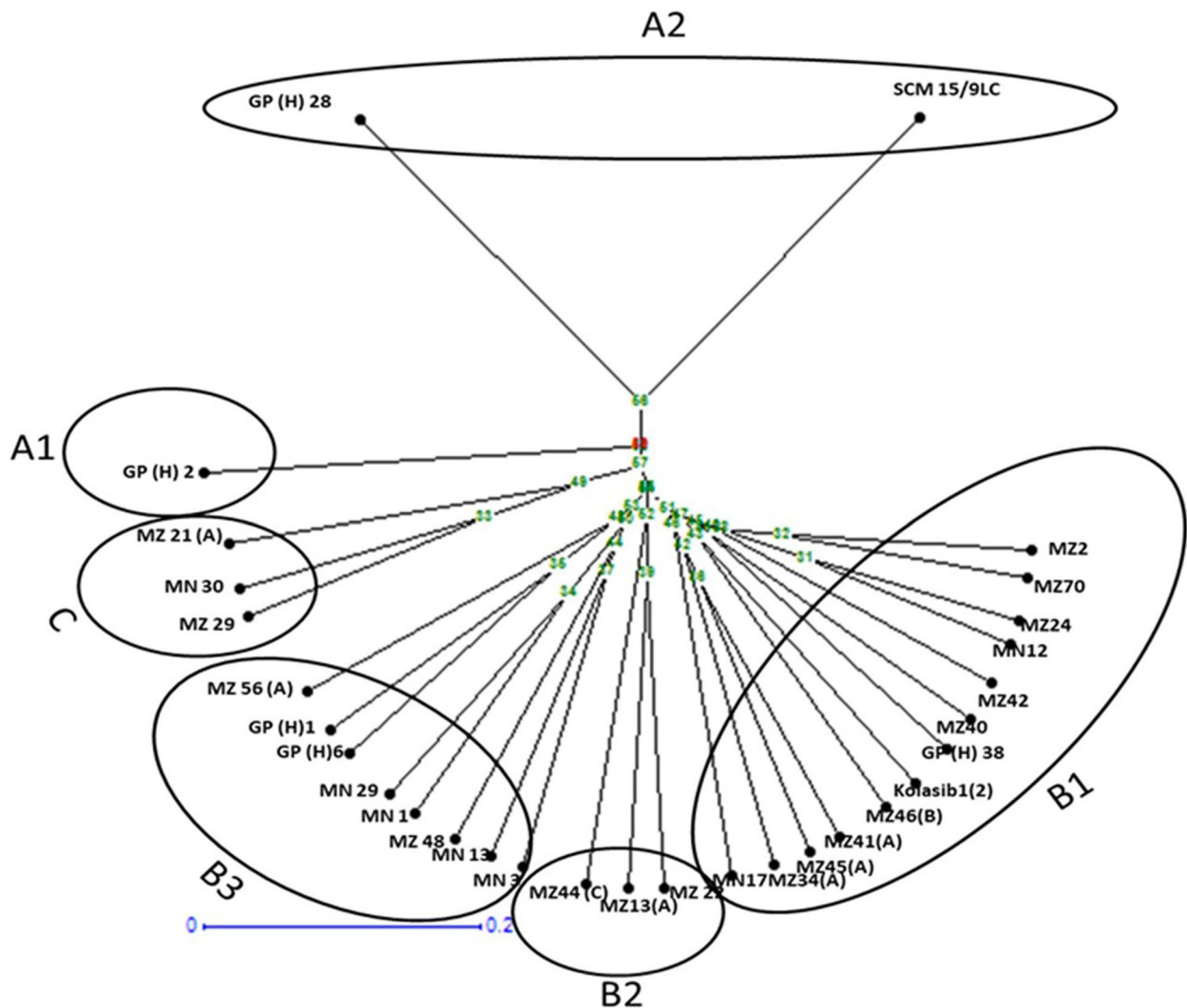


Figure 2. Cluster analysis revealed by 52 SSRs depicting genetic relationship among 30 lines. A, B and C indicate the major cluster, whereas 1, 2 and 3 indicate sub-clusters within each major cluster.

Table 1. Descriptive statistics of seven traits in the 30 early maize inbreds used in the study

Statistic	PH	EH	KRPE	KPR	CL	CD	KW
Minimum	172.00	46.50	9.33	9.75	8.750	1.40	16.23
Maximum	285.00	170.00	15.67	54.72	17.000	3.32	32.15
Mean	226.57	104.25	12.61	20.91	12.91	2.46	22.03
SD	29.34	27.81	1.81	8.18	2.29	0.45	4.58

PH, plant height (cm); EH, ear height (cm); KRPE, number of kernel rows per ear; KPR, number of kernel per row; CL, cob length (cm); CD, cob diameter (cm); KW, 100 kernel weight (g).

allelic patterns across populations is given in online Supplementary Fig. S5. The results of this study provide useful information to maize breeders for the selection of inbred lines and for making promising cross combinations based on genetic distance and clustering.

Agro-morphological diversity and marker-trait association

The statistical parameters related to agro-morphological traits of the lines are presented in Table 1 and huge morphological

diversity was observed. The mean phenotypic data of the 30 lines are provided in online Supplementary Table S6. A positive correlation was found between important agronomic traits (Table 2). The significant and highest positive correlation was observed between PH and CD ($r = 0.22$) followed by KPR and CD (0.21) and EH and KW ($r = 0.16$). We have identified three lines viz. MZ 46 (B), MN 13 and MN 3 producing mucous in the bract roots (online Supplementary Fig. S6). The identified lines have reddish colour aerial roots extending up to 2–3 nodes with high branching of brace roots. These lines

Table 2. Estimation of correlation coefficients at seven agronomic traits in the 30 early maize inbred lines used

Variables	PH	EH	KRPE	KPR	CL	CD	KW
Plant height (cm)	1	0.760	0.128**	-0.028***	0.366	0.217*	0.342
Ear height (cm)		1	0.378	0.063**	0.074**	0.129*	0.162*
Number of kernel rows per ear			1	0.191*	-0.078**	0.226	-0.190*
Number of kernel per row				1	0.278	0.207*	0.018***
Cob length (cm)					1	0.266	0.357
Cob diameter (cm)						1	0.483

*, ** & ***Significance at P value < 0.001, 0.01 and 0.05.

Table 3. Association analysis of seven agronomic traits with 52 SSR markers in 30 maize inbred lines

Trait	Marker	P value	Marker_Rsq (R^2 values)	Phenotypic variance (%)	QTL
PH	umc 1642	0.03	0.15	15.30	major QTL
	mmc 0411	0.01	0.20	20.11	major QTL
EH	umc 1642	0.01	0.19	19.68	major QTL
KRPE	mmc 0411	0.05	0.12	12.28	
	umc 1685	0.03	0.15	15.28	major QTL
	umc 2064	0.05	0.12	12.47	
	umc 2252	0.04	0.13	13.75	
KPR	umc 2289	0.01	0.21	21.20	major QTL
	umc 1153	0.00	0.25	25.04	major QTL
CL	umc 1506	0.05	0.12	12.81	
CD	bnlg2190	0.05	0.12	12.21	
KW	bnlg 1601	0.02	0.15	15.91	major QTL
	umc 2021	0.05	0.12	12.90	

can be used in developing mapping populations and further in developing hybrids with the potential to fix atmospheric nitrogen.

In our study, significant marker-trait associations were conducted between the SSR markers and the seven agro-morphological traits (Table 3). The lowest R^2 value (0.12) was detected in *mmc 0411*, *umc 2064*, *umc 1506* and *bnlg 2190*, which was associated with KRPE, CL, CD and KW and the highest (0.25) was in *umc 1153*, which was associated with KPR. It was observed that seven major QTLs were probably linked with different agronomic traits *viz.* PH, EH, KRPE, KPR and KW since the phenotypic variance (%) were estimated to be more than 15%. The SSR markers identified in the present study may serve as useful molecular markers for selecting important agronomic traits in maize breeding.

Discussion

We investigated the extent of genetic diversity, genetic relationship, population structure and allelic pattern using microsatellite markers apart from the study of agro-morphological diversity. The different multivariate techniques revealed the presence of three major possible groups, which was in support with the pedigree information, as most of the lines originated from Mizoram clustered in one group while that originated from Manipur tended

to cluster into the same group. However, it was observed that the clustering patterns did not distinguish the lines based on phenotypes and traits like grain colour, stay green character, mucous producing roots, etc. Our findings are in agreement with previous studies by Xia *et al.*, 2005 and Semagn *et al.*, 2012 who also reported lack of clear clustering patterns in maize germplasm from CIMMYT based on phenotypes, grain colour or type, maturity, environmental adaptation and heterotic groups. The lines we used in our present studies are unique at its own since they have been developed from the landraces collected from the unexplored remote areas of eastern Himalayas. In our studies, the allele size of the 52 SSR markers ranged from 50 bp (*phi 036*) to 295 bp (*p 101049*) with an average of 3.63 alleles per marker locus which revealed a high level of genetic diversity among the loci. Compared to the finding of previous researchers including, Hung *et al.* (2012) (117 alleles, 3.26 average alleles/locus), Yu *et al.* (2012) (60 alleles, 2.73 average alleles/locus), Devi *et al.* (2017b) (203 alleles, 2.69 average alleles/locus), more diversity was observed in our lines since high number of alleles per locus is attributable to the higher genetic diversity of the lines.

In the present study, a large variation was observed among the inbreds (online Supplementary Table S4). Similar findings were also reported by Prasanna (2012), Silva *et al.* (2015), Belalia *et al.* (2019) and Mukri *et al.* (2022). Such inbreds can be used

as parental lines in developing hybrids that are well adapted to the erratic weather conditions of northeast India. In the cluster diagram, inbreds derived from the same source population clustered themselves in the same group. Similar results on genetic relationship analysis in maize were also reported by Devi *et al.* (2017b) and Badu-Apraku *et al.* (2021). Most of the inbreds derived from the landraces of Mizoram fall in the same group (group B) in the cluster diagram, suggesting the presence of high genetic similarity that may be due to the same selection history and the use of closely related populations to derive the inbreds. The presence of a high level of genetic similarity among the inbreds was also reported by Semagn *et al.* (2012), Wu *et al.* (2016) and Wegary *et al.* (2019). The genotypic PCoA revealed that the lines were distributed in all four quadrangles, showing a high degree of genetic variability.

In our study, only one inbred was under the category of admixture, indicating that it is of mixed pedigree, possibly because of the introgression of the gene(s) while developing the inbred line (Yin *et al.*, 2019). Group 3 had the greatest diversity comprising diverse source populations from Manipur and Mizoram. Such findings were also reported by many researchers namely Boakyewaa *et al.* (2019) and Badu-Apraku *et al.* (2021). The clustering pattern of all the genotypes based on marker information had a close correlation with their pedigree information. The lines developed from the same source population fall under the same cluster/sub-cluster. For instance, lines developed from populations collected from Manipur *viz.* GP (H) 2, GP (H) 28 and SCM 15/9LC were together in the same cluster *i.e.*, A. Likewise, lines like MZ 2, MZ 70, MZ 24, MZ 42, MZ 40, Kolasib 1(2) and MZ 46(B) developed from populations collected from Mizoram clustered together in the same sub-group *viz.*, B1. Thus, the presence of distinct subgroups among the inbreds could be attributed to artificial selections made at different locations.

A high range of variation was found for all seven agronomic traits which revealed a high level of variation in these lines. The marker-trait association study showed that many markers are associated with the agronomic traits studied and these markers will be useful for identifying major QTLs linked with desirable traits and also in selecting lines having good agronomic characteristics.

The self-fertilizing maize was reported for the first time in a landrace of maize in the Sierra Mixe region of Oaxaca, Mexico which was reportedly grown using traditional practices with little or no fertilizer (Van Deynze *et al.*, 2018). Such maize plants have finger-like roots protruding from their stalks that secrete mucus, in which bacteria live. The bacteria fix nitrogen from the air and convert it to a different form of nitrogen and later the plants soak up the fixed nitrogen in the mucus through the finger like roots (Devi *et al.*, 2022). Three of our lines *viz.* MZ 46 (B), MN 13 and MN 3 has secreted mucous in the roots. However, further detailed studies regarding the microbiome association might provide insight of the microbes involved and the related mechanism.

Conclusion

The present study is focused on the estimation of genetic diversity, population structure, marker–trait association as well as agromorphological diversity of the early generation inbreds developed from the local germplasm collected from different places in the eastern Himalayan region. The lines developed from the local germplasm may not be good yielder but it may harbour many

useful unique traits. These lines can be exploited as donor in various maize improvement aspects as well as parents in hybrid development programmes.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S1479262123000862>

Acknowledgements. The research work is a part of the ICAR-funded project (code IXX13387). The authors would like to acknowledge the farmers of Manipur and Mizoram for sharing seeds of maize germplasm and the Director, ICAR-RC- NEH Region, Umiam for the funds and facilities provided.

Competing interests. The authors declare that they have no competing interest.

References

- Adu GB, Awuku FJ, Amegbor IK, Haruna A, Manigben KA and Aboyardana PA (2019) Genetic characterization and population structure of maize populations using SSR markers. *Annals of Agricultural Sciences* **64**, 47–54.
- Badu-Apraku B and Fakorede MAB (2017) *Advances in Genetic Enhancement of Early and Extra-Early Maize for sub-Saharan Africa*. Cham: Springer, 1–604 p.
- Badu-Apraku B, Garcia-Oliveira AL, Petroli CD, Hearne S, Adewale SA and Gedil M (2021) Genetic diversity and population structure of early and extra-early maturing maize germplasm adapted to sub-Saharan Africa. *BMC Plant Biology* **21**, 96. <https://doi.org/10.1186/s12870-021-02829-6>
- Belalia N, Lupini A, Djemel A, Morsli A, Mauceri A, Lotti C, Khelifi-Slaoui M, Khelifi L and Sunseri F (2019) Analysis of genetic diversity and population structure in Saharan maize (*Zea mays* L.) populations using phenotypic traits and SSR markers. *Genetic Resources and Crop Evolution* **66**, 243–257.
- Bharadwaj DN (2018) *Advanced Molecular Plant Breeding: Meeting the Challenge of Food Security*, 1st Edition. Florida, USA: Apple Academic Press, pp. 726.
- Boakyewaa Adu G, B-Apraku B, Akromah R, Garcia-Oliveira AL, Awuku FJ and Gedil M (2019) Genetic diversity and population structure of early-maturing tropical maize inbred lines using SNP markers. *PLoS ONE* **14**, e0214810. <https://doi.org/10.1371/journal.pone.0214810>
- Borah R, Bhattacharjee A, Rao SR, Kumar V, Sharma P, Upadhaya K and Choudhury H (2021) Genetic diversity and population structure assessment using molecular markers and SPAR approach in *Illicium griffithii*, a medicinally important endangered species of Northeast India. *Journal of Genetic Engineering and Biotechnology* **19**, 118. doi: 10.1186/s43141-021-00211-5
- Boukteb A, Sakaguchi S, Ichihashi Y, Kharrat M, Nagano AJ, Shirasu K and Bouhadida M (2021) Analysis of genetic diversity and population structure of *Orobanche foetida* populations from Tunisia Using RADseq. *Frontiers in Plant Science* **12**. <https://www.frontiersin.org/articles/10.3389/fpls.2021.618245>
- Devi EL, Devi CP, Kumar S, Sharma SK, Beemrote A, Chongtham SK, Singh CH, Ch T, Singh TB, Ningombam A, Akoijam R, Singh IM, Singh YR, Monteshori S, Omitya Y, Prakash N and Ngachan SV (2017b) Marker assisted selection (MAS) towards generating stress tolerant crop plants. *Plant Gene* **11**, 205–218.
- Devi EL, Avasthe RK, Meetei CC, Dutta SK, Das SK, Singh AR, Bhutia TL, Kumar A, Saha S and Mishra VK (2022) Miracle maize lines of North-Eastern region that can fix atmospheric nitrogen. *CAU Farm Magazine*, July – September, p. 15.
- Earl DA and Vonholdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetic Resources* **4**, 359–361.
- Evanno G, Regnaut S and Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611–2620.

- Govindaraj M, Vetriventhan M and Srinivasan M (2015) Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. *Genetics Research International* **2015**, 431487.
- Hasan N, Choudhary S, Naaz N, Sharma N and Laskar RA (2021) Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. *Journal of Genetic Engineering and Biotechnology* **19**, 128. doi: 10.1186/s43141-021-00231-1
- Hung TN, Huyen TN, Loc NV and Cuong BM (2012) The application of SSR molecular indicator to assess the purity and genetic diversity of waxy corn inbred lines. *Journal of ISSASS (International Society for Southeast Asian Agricultural Sciences)* **18**, 45–54.
- Islam MA, Alam MS, Maniruzzaman M and Haque MS (2023) Microsatellite marker-based genetic diversity assessment among exotic and native maize inbred lines of Bangladesh. *Saudi Journal of Biological Sciences* **30**, 103715.
- Jones ES, Sullivan H, Bhatramakki D and Smith JSC (2007) A comparison of simple sequence repeat and single nucleotide polymorphism marker technologies for the genotypic analysis of maize (*Zea mays* L. *Theoretical and Applied Genetics* **115**, 361–371. <https://doi.org/10.1007/s00122-007-0570-9>
- Laosatit K, Amkul K, Somta P, u ma Tanadul O, Kerdri C, Mongkol W, Jitlaka C, Suriharn K and Jompuk C (2022) Genetic diversity of sweet corn inbred lines of public sectors in Thailand revealed by SSR markers. *Crop Breeding and Applied Biotechnology* **22**(4), 1–8. <https://doi.org/10.1590/1984-70332022v22n4a45>
- Mukri G, Gadag RN, Bhat JS, Nepolean T, Gupta NC, Mittal S, ML N, Kumar R and Pal D (2022) Characterization of sub-tropical maize (*Zea Mays* L.) inbred lines for the variation in kernel row numbers (KRN). bioRxiv preprint. <https://doi.org/10.1101/2022.09.23.509129>
- Peakall RO and Smouse PE (2012) GenAlex 6.5: genetic analysis in excel. Population genetic software for teaching and research. *Bioinformatics (Oxford, England)* **28**, 2537–2539.
- Perrier X and Jacquemoud-Collet JP (2006) *DARwin Software*. Available at <http://darwi.n.cirad.fr/darwin>
- Porebski S, Bailey LG and Baum BR (1997) Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter* **15**, 8–15.
- Prasanna BM (2012) Diversity in global maize germplasm: characterization and utilization. *Journal of Biosciences* **37**, 843–855.
- Pritchard JK, Stephens M and Donnelly P (2000) Inference of population-structure using multilocus genotype data. *Genetics* **155**, 945–959.
- Sbordoni V, Allegrucci G and Cesaroni D (2012) Population structure. In White WB and Culver DC (eds), *Encyclopedia of Caves*, 2nd Edn. USA: Academic Press, pp. 608–618. <https://doi.org/10.1016/B978-0-12-383832-2.00090-6>
- Semagn K, Magorokosho C, Vivek BS, Makumbi D, Beyene Y, Mugo S, Prasanna BM and Warburton ML (2012) Molecular characterization of diverse CIMMYT maize inbred lines from eastern and Southern Africa using single nucleotide polymorphic markers. *BMC Genomics* **13**, 113. <https://doi.org/10.1186/1471-2164-13-113>
- Silva TA, Cantagalli LB, Saavedra J, Lopes AD, Mangolin CA, Pires MF, Machado S and Scapim CA (2015) Population structure and genetic diversity of Brazilian popcorn germplasm inferred by microsatellite markers. *Electronic Journal of Biotechnology* **18**, 181–187.
- Singh N, Choudhury DR, Singh AK, Kumar S, Srinivasan K, Tyagi RK, Singh NK and Singh R (2013) Comparison of SSR and SNP markers in estimation of genetic diversity and population structure of Indian rice varieties. *PLoS One* **8**, e84136. doi: 10.1371/journal.pone.0084136
- Singh AR, Devi EL, Dayal V, Saha S, Lungmauna, Dutta SK, Boopathi T and Singh SB (2019) Diversity of landraces maize in Mizoram: prospects, challenges and opportunities. In *Souvenir of National Workshop on Scientific Maize Cultivation in North East India*. Aizawl, India: ICAR Research Complex for NEH Region, Umiam, pp. 98–104.
- van Berloo R (2008) GGT 2.0: versatile software for visualization and analysis of genetic data. *Journal of Heredity* **99**, 232–236.
- Van Deynze A, Zamora P, Delaux PM, Heitmann C, Jayaraman D, Rajasekar S, Graham D, Maeda J, Gibson D, Schwartz KD, Berry AM, Bhatnagar S, Jospin G, Darling A, Jeannotte R, Lopez J, Weimer BC, Eisen JA, Shapiro HY, Ane JM and Bennet AB (2018) Nitrogen fixation in a landrace of maize is supported by a mucilage-associated diazotrophic microbiota. *PLoS Biology* **16**, e2006352. <https://doi.org/10.1371/journal.pbio.2006352>
- Wegary D, Teklewold A, Prasanna BM, Ertiro BT, Alachiotis N, Negera D, Awas G, Abakemal D, Ogugo V, Gowda M and Semagn K (2019) Molecular diversity and selective sweeps in maize inbred lines adapted to African highlands. *Scientific Reports* **9**, 13490. <https://doi.org/10.1038/s41598-019-49861-z>
- Wu Y, Vicente FS, Huang K, Dhlwayo T, Costich DE, Semagn K, Sudha N, Olsen M, Prasanna BM, Zhang X and Babu R (2016) Molecular characterization of CIMMYT maize inbred lines with genotyping-by-sequencing SNPs. *Theoretical and Applied Genetics* **129**, 753–765.
- Xia XC, Reif JC, Melchinger AE, Frisch M, Hoisington DA, Beck D, Pixley K and Warburton ML (2005) Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers: II. Subtropical, tropical mid-altitude, and highland maize inbred lines and their relationships with elite U.S. and European maize. *Crop Science* **45**, 2573–2582. doi: 10.2135/cropsci2005.0246
- Yadav OP, Hossain F, Karjagi CG, Kumar B, Zaidi PH, Jat SL, Chawla JS, Kaul J, Hooda KS, Kumar P, Yadava P and Dhillon BS (2015) Genetic improvement of maize in India: retrospect and prospects. *Agricultural Research* **4**, 325–338.
- Yang X, Xu Y, Shah T, Li H, Han Z, Li J and Yan J (2011) Comparison of SSRs and SNPs in assessment of genetic relatedness in maize. *Genetica* **139**, 1045–1054.
- Yin V, Sa KJ, Lim SE and Lee JK (2019) Genetic diversity and association analyses of Chinese maize inbred lines using SSR markers. *Plant Breeding and Biotechnology* **7**, 186–199.
- Yu RH, Wang YL, Sun Y and Liu B (2012) Analysis of genetic distance by SSR in waxy maize. *Genetics and Molecular Research* **11**, 254–260.