

Molecular characterization of a farmer-preferred maize landrace population from a multiple-stress-prone subtropical lowland environment

FORTUNATE MAKORE¹, EDMORE GASURA¹✉, CALEB SOUTA², UPENYU MAZARURA¹, JOHN DERERA², MELULEKI ZIKHALI², CASPER N. KAMUTANDO¹, COSMOS MAGOROKOSHO³, SHORAI DARI¹

¹Department of Crop Science, Faculty of Agriculture, University of Zimbabwe. P.O Box MP167, Mount Pleasant, Harare, Zimbabwe.

✉email: gasurae@yahoo.com

²Seed Co. Limited, Arnold Research Station. P.O Box CH 142 Chisipite, Harare, Zimbabwe

³International Maize and Wheat Improvement Center (CIMMYT). P.O Box MP163, Harare, Zimbabwe

Manuscript received: 9 December 2020. Revision accepted: 14 January 2021.

Abstract. Makore F, Gasura E, Souta C, Mazamura U, Derera J, Zikhali M, Kamutando CN, Magorokosho C, Dari S. 2021. Molecular characterization of a farmer-preferred maize landrace population from a multiple-stress-prone subtropical lowland environment. *Biodiversitas* 22: 769-777. The study was conducted to assess genetic diversity of 372 maize lines using 116 single nucleotide polymorphism (SNP) markers. Three hundred and forty-seven lines were S₁ lines (coded J lines) from a local maize landrace population and twenty-five were the widely used standard lines. The number of alleles per marker ranged from two to four and the average was three alleles. The average polymorphic information content (PIC) value of 0.405 indicates high genetic diversity for maize lines evaluated in this study. Population structure revealed three distinct sub-populations. Sub-population 1 contained two J lines; sub-population 2 contained five J lines and sub-population 3 contained the rest of the J lines and all the standard lines. Analysis of molecular variance (AMOVA) identified 22% variance among and 78% variance within the three subpopulations, indicating high gene exchange and low genetic differentiation. Hierarchical cluster analysis further divided the lines into nine subgroups placing some of the J lines into known heterotic groups, i.e., J30_3, J393_4, J393_3, and J393_1 in CIMMYT heterotic group B. Allelic variation observed can be a source of allele combination for breeding programs interested in widening their genetic base. The private alleles that were present in the J lines suggest availability of stress-tolerant genes that breeders can incorporate in new hybrids.

Keywords: Genetic diversity, heterotic group, maize landraces, population structure, private alleles

INTRODUCTION

Maize is a very important food crop in sub-Saharan Africa (SSA) but productivity is very low at 1.8 t ha⁻¹ compared to world average of 4.9 t ha⁻¹ (Rezende et al. 2019). Abiotic stresses such as heat, drought, and low nitrogen (low N) are some of factors affecting maize production and it is difficult to find superior genotypes because variability is usually lost due to massive use of uniform commercial varieties. Artificial selection has gradually narrowed the genetic distances in maize germplasm (Whitt et al. 2002). Amount of genetic variability determines the breeding progress in any crop breeding program (Musundire et al. 2019). The US maize breeding program has operated as a closed system that led to loss of genetic variation (Nelson et al. 2016). Unimproved local varieties (landraces) are products of natural selection and are usually adapted to local growing conditions, therefore an important source of favorable traits such as resistance to abiotic and biotic stresses. Miti et al (2010) noted that some landrace varieties tolerated stress caused by low N better than improved maize varieties. Maize landrace populations are a source of genetic variation and comprise large allelic diversity (Holker et al. 2019) that can be used to improve productivity. Maize

landraces can be accessed from gene banks or from farmers and have been widely used by researchers as sources of genetic material to broaden the genetic base of elite maize germplasm (Strigens et al. 2013; Böhm et al. 2017; Brauner et al. 2019). Landraces have the highest genetic variation and best adapted to the environment they evolved (Maxted et al. 1997). The only drawback to their use in hybrid breeding is genetic heterogeneity and a high genetic load and to overcome these limitations breeders can now use new breeding tools such as double haploid technique, genotyping, and marker selection (Böhm et al. 2017). Dividing maize inbred lines into heterotic groups and establishing heterosis using molecular marker breeding has become a very important approach in modern research (Guan et al. 2020). Genetic diversity provides an estimate of average heterozygosity and genetic distances among individuals in a population.

Heterosis at gene level is contributed by variation in the presence of genes or novel beneficial alleles (Springer and Stuper 2007; Zhang et al. 2016). Genetic diversity studies help researchers to make full use of variation and heterosis as inbred lines can be assigned into heterotic group. Conventional breeding can be used to establish diversity but it is difficult to accurately divide heterosis grouping using phenotypic data (Guan et al. 2020). Molecular

markers reveal sites of DNA variation in organisms and are not affected by environmental factors. Co-dominant markers such as single nucleotide polymorphism (SNP) can be used to construct linkage maps, to assess level of genetic diversity within germplasm and in cultivar identity (Pagnotta 2018) because they are abundant and have a high throughput (Collard et al. 2005). The importance of using molecular markers to infer genetic diversity in maize is well documented (Makumbi et al. 2018). In allogamous species like maize, landraces collections represent populations of heterozygous and heterogeneous individuals (Holker et al. 2019) hence the need to self and use S₁ lines to evaluate breeding potential of the germplasm. This study was done to determine the level of diversity in maize S₁ lines derived from a landrace population. The objective of this study was to determine population structure and genetic diversity among the lines and infer potential heterotic groups and breeding potential of the derived lines.

MATERIALS AND METHODS

Germplasm

The maize landrace “Redcore” was obtained from a smallholder farmer in the rural resettlement area, in Chilonga village, in Chiredzi district of Zimbabwe. Predominantly it has a red core and white grain. The landrace has been maintained on-farm, by the farmer for more than 11 years by selecting the large cobs and shelling kernels from the middle of the cobs to constitute 50 kg of seed every season. The attributes of the landrace include large white kernels. On average the farmer harvests at least 2 to 3 tons per hectare of grain. According to the farmer, it is fairly resistant to pests, such as the maize weevils, and has a long taproot which makes it tolerant to drought. The farmer has also observed that it is heat stress-tolerant compared to commercial maize hybrids which are grown in the area. The farmer has been selling seed of this landrace to the neighboring farmers in the area.

A 50 kg bag of grain of the “Redcore” was collected from the farmer in 2018. This was sampled for planting at the Stapleford Research Station, in Harare. A total of 9 000 plants were established in 600 rows of 5 m rows of 15 plants each, spaced at 30 cm within the row and 75 cm between the rows. A random sample of 600 plants was self-pollinated to produce 600 S₁ families. For laboratory genotyping, a random sample of 87 cobs representing different families was collected and total of 347 seed samples was used. From the 87 cobs, 347 seed samples (coded as J lines) were genotyped and 25 samples derived from six standard tester inbred lines (TL) were also genotyped to make a total of 372 inbred lines. The tester lines belong to known regional heterotic groups such as N3, NC, NAW, and the CIMMYT A and B. regional. The genotyping was done in the SeedCo laboratory at Rattray Arnold Research Station, in 2020.

Single nucleotide polymorphic (SNP) marker selection and genotyping

Genomic deoxyribonucleic acid (DNA) isolation was done at Rattray Arnold Research Station (RARS)

laboratory in Harare, Zimbabwe. The DNA from seeds of the 347 inbred lines and 25 seed samples of standard tester inbred lines was extracted using the modified CTAB method (Saghai-Marooof et al. 1984). The DNA was checked for quality using the agarose gel and quantity using a spectrophotometer. A total of 116 single nucleotide polymorphism (SNP) markers from an initial set of 139 SNP markers obtained from the LGC Group Ltd single nucleotide polymorphism (SNP) database were used in genotyping. The 116 remained after removing the markers that had not amplified well in other words the markers that did not cluster into the three KASP groups of FAM (blue), VIC (red), and heterozygous (green) were removed leaving a total of 116 markers which were used in the diversity analysis.

The markers were selected from each of the ten maize chromosomes with 16, 11, 12, 13, 11, 7, 7, 8, 10, and 9 markers each from chromosome one to ten respectively. The markers were selected to be evenly spread from the proximal to the distal end of each chromosome using the centimorgan (cM) distances. The spread across chromosomes enabled the diversity to be determined across the whole genome. This even spread allowed us to use a few markers but still capture the diversity as the markers were not clustered on a few chromosomes.

All KASPTM amplifications were carried out according to Zikhali et al. (2017) with the following changes in 384 well plate instead of 1536-well plates by using 1.8µl instead of 1 µL of KASPTM master mix 1X (LGC Group, UK) and 0.0135 µL of primer mix (12 µL FAM primer at 100 µm + 12 µL of VIC primer at 100 µm + 30 µL of common primer at 100 µm + 46 µL of dH₂O). One microlitre of DNA at 2 ng/µL was previously added on each well of the 384 well plates instead of 1536 plates and dried at 60 °C for 30 min (Zikhali et al. 2017). PCR reactions were carried out by using a touchdown program: 95 °C for 15 min, then 10 cycles of 95 °C for 20 s and 61 °C for 60 s (0.6 °C per cycle), followed by 26 cycles of 95 °C for 20 s and 55 °C for 60 s on a hydrocycler from LGC (Zikhali et al. 2017).

Statistical analysis

Genetic properties of SNP markers

The gene diversity (GD) of locus also known as expected heterozygosity (*He*) describes the proportion of heterozygous genotypes under Hardy-Weinberg equilibrium (Nei 1973). Polymorphic information content (PIC) and expected heterozygosity (*He*) were calculated using GenAlex software version 6.5 (Meirmans 2012; Peakall and Smouse 2012) and RStudio (Team R, 2015). Expected heterozygosity (*He*) = $1 - \sum P_i^2$, where, *P_i* is the frequency of the *i*th allele for the population and sum *P_i²* is the sum of the squared population allele frequencies. Polymorphism information content indicates genetic properties of SNPs in a population and was calculated using the formula of Botstein et al. (1980) as follows:

$$PIC = 1 - \sum_{i=1}^n P_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2P_i^2 P_j^2$$

Where, P_i and P_j are the frequencies of i^{th} and j^{th} alleles for the selected marker, respectively.

Analysis of population structure

Structure software (Pritchard et al. 2000) with a burning length of 5000 and MCMC of 50000 was used to determine the number of groups among the inbred lines. The online genetic software Structure Harvester (Earl and von Holdt 2012) visualized the structure analysis results following the Evanno approach to determine the most probable K-value. RStudio software was then used for cluster analysis to depict the inferred groups using the Gower's distance (Gower 1971) and neighbor-joining algorithm. The silhouette plots using RStudio like structure results also suggested three groups and the dendrogram was then subdivided into three groups using the cutree option in RStudio (Team R 2015).

The number of subpopulations determined with structure was used for analysis of molecular variance (AMOVA) and calculation of Nei's genetic distance using GenAlex v6.5 (Peakall and Smouse 2012). Nei's parameters were calculated according to Nei (1973) using the following formulas; Nei Genetic Identity = $JxJy / (Jx * Jy)^{0.5}$, Nei Unbiased Genetic Identity = $Jxy / (JxUb * JyUb)^{0.5}$, Nei Genetic Distance = $-1 * \ln(\text{Nei Identity})$ and Nei Unbiased Genetic Distance = $-1 * \ln(\text{Nei Unbiased Identity})$, where x represents the allele frequencies in pop 1; y represents the allele frequencies in pop 2; Jx is the sum of x^2 over loci; Jy is the sum of y^2 over loci; $JxJy$ is the sum of xy over loci. $JxUb$ is the sum of xUb over loci; $JyUb$ is the sum of yUb over loci. In addition, genetic indices such as number of loci with private allele, number of different alleles (N_a), Shannon information index (I), and number of effective alleles (N_e) were also calculated using GenAlex v6.5 (Peakall and Smouse 2012) as follows $N_e = 1 / (\text{Sum } P_i^2)$ and $I = -1 * \text{Sum } (P_i * \ln(P_i))$. The duplications were also identified using GenAlex software version 6.5 (Meirmans 2012; Peakall and Smouse 2012).

RESULTS AND DISCUSSION

Characterization and distribution of SNP markers

A total of 372 accessions were genotyped using 116 SNP markers were used in this study. The SNP markers number of alleles ranged from 2 to 3 with an average of two alleles per marker. Genetic diversity (GD) values calculated as expected heterozygosity (H_e) ranged from 0.005 to 0.661 and the mean was 0.45. The polymorphic information content (PIC) values varied from 0.005 to 0.587 with an average of 0.405. The Shannon information index (I) mean was 0.783 and ranged from 0.019 to 1.09 (Table 1).

Duplications

The J lines and tester lines used in the study were replicated four or five times. There were duplications for genotypes TL6, TL1, TL4, TL5, J_60, TL2 and TL3 (Table 2). The tester inbred lines had more duplications compared to the S_1 derived families, where only one out of the 87 was homozygous.

Table 1. Summary statistics for the 116 SNP markers used

Estimates	Average	Min.	Max.
Missing %	0.9	0	5.4
Allele No.	3	2	3
Common allele No.	2.9	1	3
Rare allele No.	0.112	0	2
Expected heterozygosity	0.453	0.005	0.661
Polymorphic information content	0.405	0.005	0.587
Shannon information index	0.783	0.019	1.09

Table 2. Duplications among the lines characterized

Sample no.	Sample	No of duplications	Label of duplication
367	TL6_4	2	A
364	TL6_1	0	A
355	TL1_4	4	B
354	TL1_3	0	B
353	TL1_2	0	B
352	TL1_1	0	B
372	TL4_5	3	C
371	TL4_4	0	C
370	TL4_3	0	C
351	TL5_4	2	D
350	TL5_3	0	D
123	J_60_4	2	E
122	J_60_3	0	E
359	TL2_4	4	F
358	TL2_3	0	F
357	TL2_2	0	F
356	TL2_1	0	F
363	TL3_4	4	G
362	TL3_3	0	G
361	TL3_2	0	G
360	TL3_1	0	G

Note: TL= tester lines, J = S_1 derived lines coded as J lines

Population structure

A huge genetic diversity was observed among 372 maize lines, this was reflected by Gower's genetic distances which ranged from 0 to 0.64 (Figure 1). The Dendrogram from UPGMA cluster analysis clearly shows different groupings of the samples used in the study. The UPGMA cluster method resolved the 372 maize lines into 3 major groups at 0.5 Nei's genetic distance cutoff (Figure 1). Dendrogram show 3 possible groups with the first group having only 2 S_1 lines (J383.3 and J212-2), second group with 5 S_1 lines (J104_4, J104_3, J104_2, J104_1 and J109_3), and third group with a total of 365 lines i.e. the rest of J lines plus all the tester lines. Population structure analysis revealed three groups where the best K (parameter which describes the number of subpopulations that make up the total population) was selected (Figure 2). The model-based quantitative clustering also shows the proportion of genome of individual line, where the duplications are shown with their perfect single color (Figure 3). Silhouette plots computed by R software also show three possible groups with actual numbers of individuals per group; group 1 has 365, group 2 has 5, and group 3 has 2 individuals (Figure 4). The hierarchical

clustering show nine major groups and up to twenty sub-clusters that may be used for breeding purposes (Figure 5). The tester lines used in this study belong to known regional heterotic groups. Some of the S_1 derived lines clustered together with tester lines that belong to known regional heterotic groups. For example, J lines J30_3, J393_4, J393_3, and J393_1 are in the same sub-cluster with tester line 3 which means they all belong to the known regional heterotic group, CIMMYT B.

Genetic diversity

The three subpopulations identified in STRUCTURE were then applied in GenAlEx 6.503 to calculate analysis of molecular variance (AMOVA), Nei genetic distance, and genetic diversity indices. The AMOVA revealed that 22% of the total variation was found among subpopulations while the rest (78%) was within populations (Figure 6). Figure 7 shows AMOVA of two populations that are J lines and tester lines revealing different levels of diversity, 10% of variation is among subpopulations and 90% within populations agreeing with results from the Nei's parameters calculation. Nei genetic identity was 0.91 while Nei unbiased genetic identity was 0.921. Nei genetic distance between these two populations was 0.094 while Nei unbiased genetic distance was 0.083. Genetic differences were observed between J lines and tester inbred lines illustrated by allelic patterns in Figure 8; where population 1 (Pop 1) represents the J lines and population 2 (Pop 2) tester lines. Population 1 had 350 lines with one or more private alleles whilst Pop 2 had no private alleles. The J lines had higher mean number of different alleles

(N_a), number of effective alleles (N_e), Shannon index information (I), N_a frequency, and number of common alleles compared to tester lines. Heterozygosity was lower for population 2 compared to population 1. Allelic patterns for the three possible groups inferred from STRUCTURE show huge diversity (Figure 9). Population 1, 2 and 3 represent group 1, 2 and 3 respectively as shown by the dendrogram from left to right. A total of 366 lines had private alleles with most of these found in population 3. Population 3 (J lines) had high heterozygosity levels of 0.8 compared to other populations.

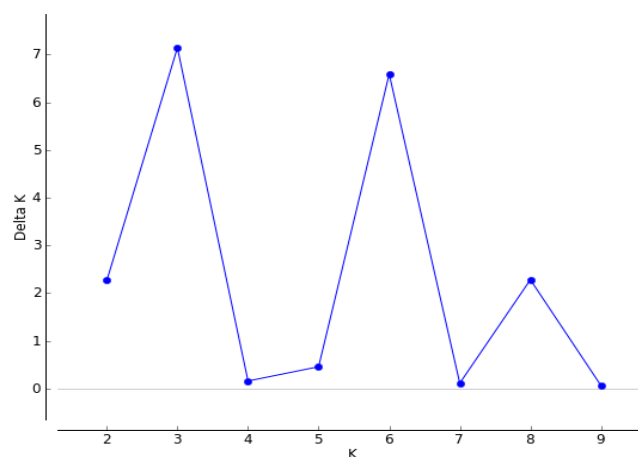


Figure 2. A graph showing the *ad hoc* statistic for Δk computed for k varying from 1 to 10 from which the best k was selected

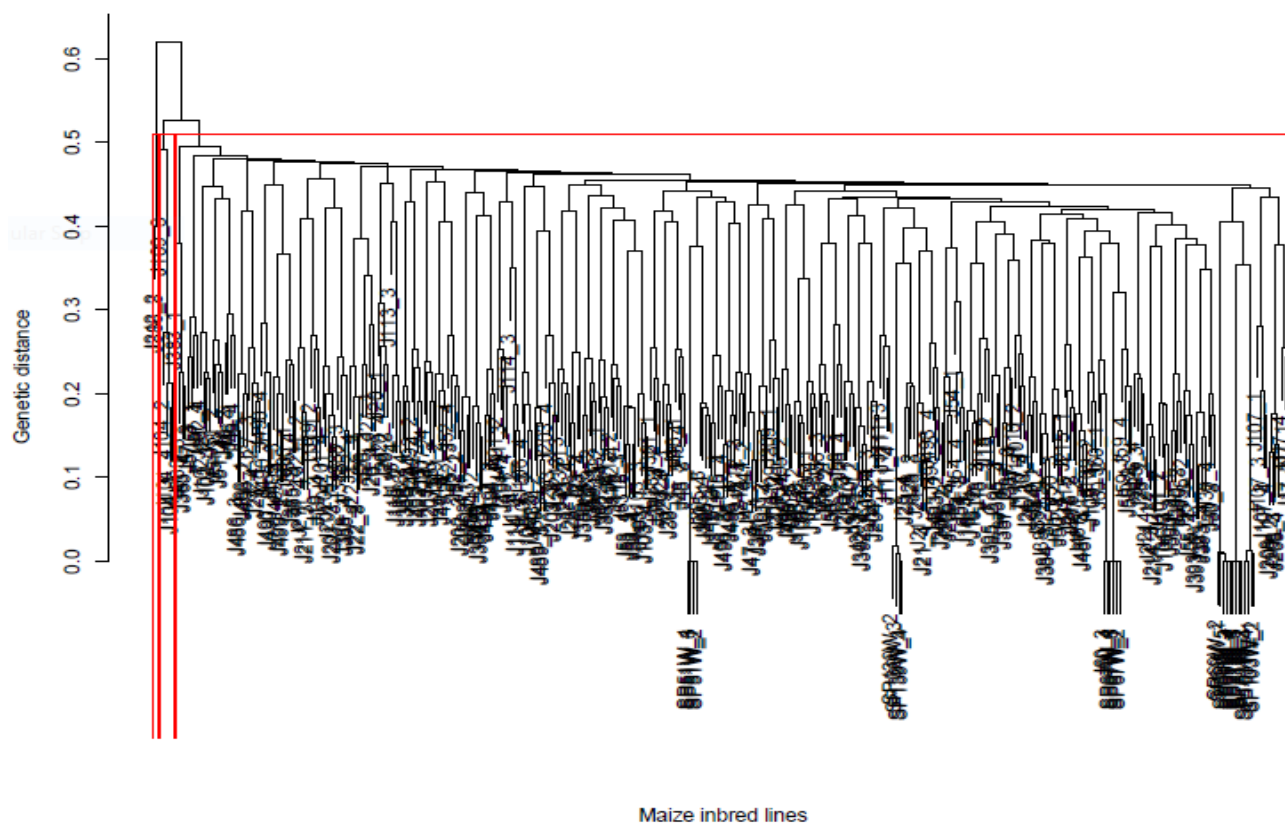


Figure 1. Dendrogram showing relationships among the 372 maize lines

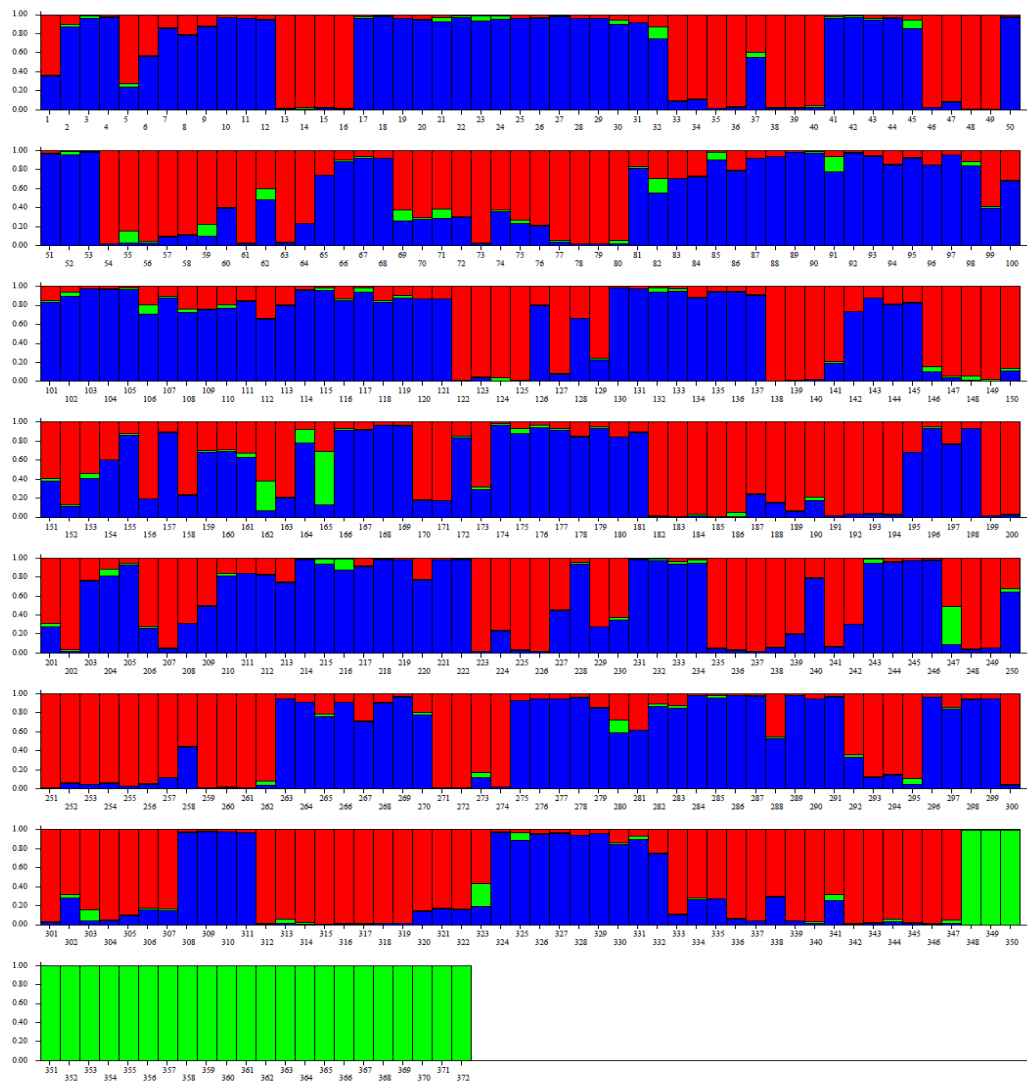


Figure 3. Bar plot of model-based quantitative clustering showing proportion of genome of individual line for 372 maize lines at k=3 used for genetic diversity studies with 116 SNP markers

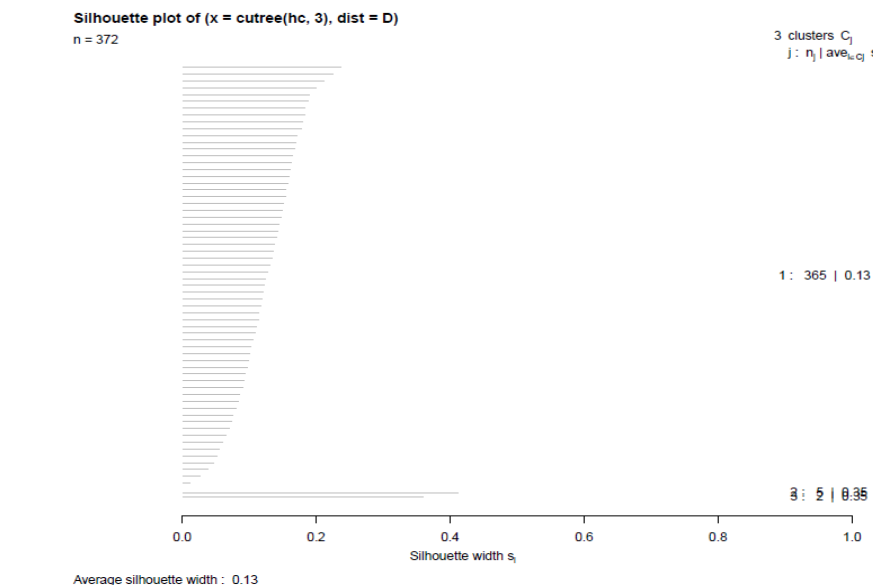


Figure 4. Silhouette plot shows that three groups are possible according to R software

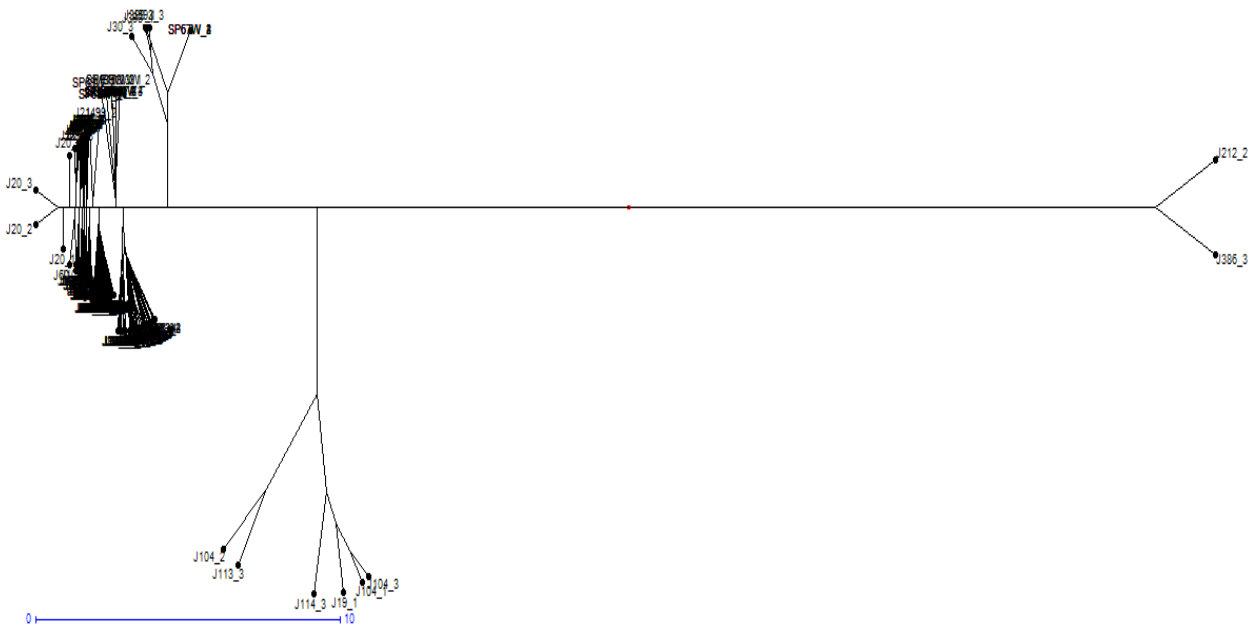


Figure 5. Hierarchical clustering showing nine groups and twenty possible sub-clusters

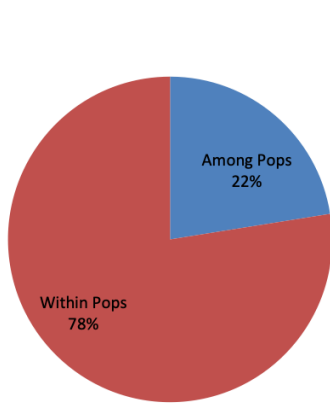


Figure 6. AMOVA is shown as percentages among and within 372 maize lines

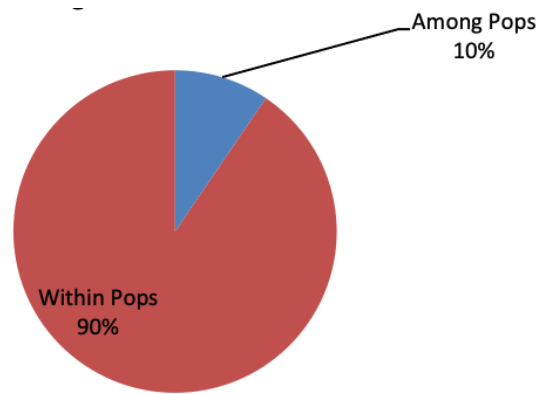


Figure 7. AMOVA is shown as percentages among and within J lines

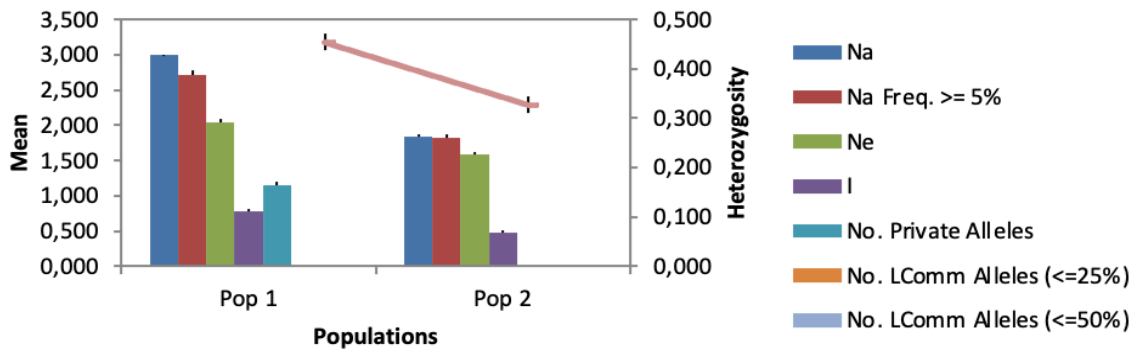


Figure 8. Allelic patterns between J lines and tester line populations

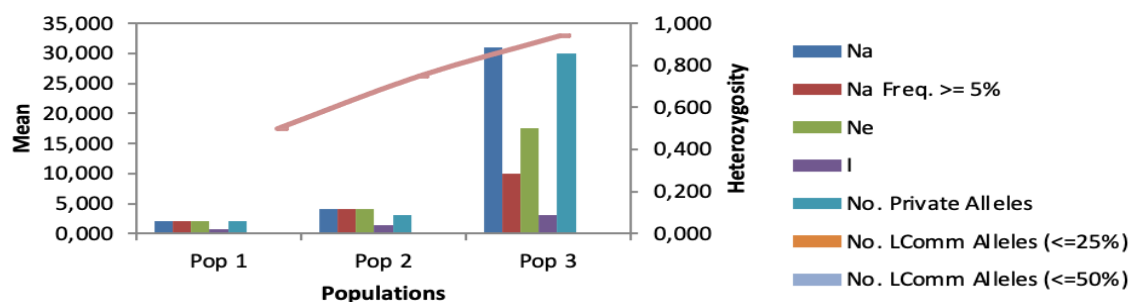


Figure 9. Allelic patterns across three populations

Discussion

In this work, the 116 SNP markers used were effective in discriminating the 372 inbred lines or accessions. The gene diversity (GD) also known as expected heterozygosity (H_e) and PIC values are both measures of genetic diversity among genotypes in breeding populations as they indicate usefulness of markers for linkage analysis. The average expected heterozygosity of 0.453 observed in this study indicated substantial diversity in the population and were similar to GD values of >0.45 recorded by Zhang et al. (2016) where they genotyped 362 maize inbred lines using 56,110 SNP markers. The high level of heterozygosity observed in the current population was expected because these were S_1 or F_3 population which had just gone through one generation of self-pollination. The average PIC of 0.4 recorded in this study was higher than the value of 0.256 recorded by Dao et al. (2014) in the genotypic study of 100 maize inbred lines using 1057 informative SNP markers. The overall H_e values were slightly higher than PIC which is expected since PIC values will always be smaller than H_e values (Shete et al. 2000). According to Botstein et al. (1980) markers with a PIC value >0.5 are considered to be highly informative, 0.25 to 0.5 moderately informative and less than 0.25 slightly informative. Therefore, SNP markers used in the study were informative and effective in discriminating maize genotypes because 80% of the markers had PIC value >0.25 and can be recommended for future genetic diversity studies. The high PIC values can be attributed to high average allele number per marker of three recorded. Presence of duplications indicates level of homozygosity and results from current study were expected as more duplication was observed in standard lines than in J lines where there is a lot of segregation still going on.

Population structure is informative in understanding genetic diversity and in this study STRUCTURE results (optimal $K=3$) and the UPMGA cluster analysis divided the 372 lines into three sub-populations. Boakyewaa et al. (2019) also noted three subpopulations ($K=3$) within 94 inbred lines using 15,047 SNP markers. Moreover, Silhouette plot from R software also shows three sub-populations, showing average genetic distances (GD) for each sub-population as 0.13, 0.35 and 0.33 for sub-populations 3, 2 and 1 respectively meaning that the lines in sub-population 3 were closely related. Sub-population 1 and 2 contained few J lines, 1.34% and 0.54% respectively, while the standard lines were all clustered together with

most of the J lines in sub-population 3 constituting 98.12% of total accessions. These suggest that the landrace population and standard lines may be coming from the same gene pool considering the background that it was an open-pollinated variety kept by the farmer for more than ten years. The allelic patterns and genetic diversity indices provided insight into diversity within each of the subpopulations. Expected heterozygosity was higher in sub-population 3 compared to 1 and 2 meaning that sub-population 3 was more diverse.

The AMOVA revealed molecular variation of 22% within and 78% among the 372 inbred lines. The magnitude of variation between populations was higher than findings reported by Nyaligwa et al. (2015) in their study of 79 elite lines using SSR markers. Allelic patterns across these three populations show the presence of private alleles and implication for breeding is that the private alleles are an indication of new genes that could be crucial in stress tolerance. It was therefore prudent to do an analysis between the standard lines and the J lines in order to illustrate where the unique genes might be, the AMOVA observed greater variation within (90%) accessions than between (10%) them and this was similar to the findings reported by Nda et al. (2016). Allelic patterns of the populations revealed that the J lines had private alleles meaning they have unique genes. These private alleles can be linked to the 10% molecular variation observed between the two populations. This could be attributed to adaptation of the "Redcore" landrace to drought, heat stress, and low N conditions. High levels of heterozygosity were also observed implying the presence of substantial genetic diversity that breeders can explore. The allelic patterns of the populations revealed private alleles in the J lines suggesting the presence of new genes that may be contributing to drought and heat tolerance. Aci et al (2018) also noted presence of unique alleles in 47 maize landrace populations, thus the landrace populations represent a valuable resource for genetic improvement of elite maize germplasm which lacks alleles for abiotic stress tolerance.

Hierarchical structure further divided sub-population 2 and 3 into 9 sub-clusters indicating a high level of diversity. Zhang et al. (2016) and Giordani et al. (2019) also noted sub-clusters in their studies. Furthermore, there were variations within the sub-clusters, such as the sub-cluster with J104.2, J113.3 and J114.3. Implication for breeding is that a breeder with limited resources may

choose to select just one line from each sub-cluster for testing. The study was able to place some of the J lines into heterotic groups because all the standard lines included in this study belong to known regional heterotic groups. They were intentionally included in genotyping in order to establish if the J lines derived from an open-pollinated variety population would align with any of these known regional heterotic groups. Indeed some of the J lines such as J30_3, J393_4, J393_3, and J393_1 were in the same cluster with CIMMYT's tester lines CML444 and CML312 which belong to CIMMYT heterotic group B. However, the majority of the J lines (343 lines) were not aligned with any of the tester lines as they were placed in different clusters. This implies that the J lines can be crossed with most of the tester lines to determine their hybrid potential.

The results revealed the potential of tropical maize landraces as a worthy source of genetic variation in maize improvement similar to findings by Nelimor et al. (2019) when he investigated diversity among 196 maize landraces. Huge genetic diversity was observed among the J maize inbred lines derived from the "Redcore" landrace population and similar results were observed by Belalia et al (2018) when they genotyped landrace populations, therefore a worthy source of genetic variation. The implication for breeding is that lines from the landrace population can be used in breeding programs and variations between groups and within groups of the J lines can result in heterosis essential in hybrid development. In addition, the lines can be further assessed to determine their breeding value by evaluating their testcross performance in combination with elite single crosses and inbred lines.

In conclusion, the SNP markers used in the study were informative and effective in categorizing the inbred lines. Population structure analysis revealed three subpopulations ($K=3$) within the 372 inbred lines. High genetic distances obtained among the paired inbred lines showed existence of variability of the landrace-derived J lines that can be exploited for development of hybrids. The uniqueness of J lines and their inclination towards standard lines from the breeding programs in the region was also revealed suggesting that the landraces derived lines and standard lines may have come from the same gene pool, this has implications for integration of the J lines in the programs. Duplications were observed in standard lines and not in J lines as expected due to levels of heterozygosity. A few J lines such as J30_3, J393_4, J393_3, and J393_1 belong to CIMMYT heterotic group B. The other J lines can also be used to broaden the existing heterotic groups used in the region.

ACKNOWLEDGEMENTS

Genotyping was supported by the molecular laboratory at Seed Co Ltd, in Zimbabwe. The authors are grateful to the farmer, Aaron Rwasarira of Chilonga Village in Chiredzi District of Zimbabwe for providing the landrace maize seed

REFERENCES

- Aci MM, Lupini A, Mauceri A, Morsli A, Khelifi L, Sunseri F. 2019. Genetic variation and structure of maize populations from Saoura and Gourara oasis in Algerian Sahara. *BMC Genetics* 19 (1). DOI: 10.1186/s12863-018-0655-2.
- Belalia N, Lupini A, Sunseri F. 2018. Analysis of genetic diversity and population structure in Saharan maize (*Zea mays* L.) populations using phenotypic traits and SSR markers. *Genet Resour Crop Ev* 66: 243-257.
- Brauner PC, Schipprack W, Melchinger AE. 2019. Testcross performance of doubled haploid lines from flint maize landraces is promising for broadening the genetic base of elite germplasm. *Theor. Appl Genet* 132: 1897-1908.
- Böhm J, Schipprack W, et al. 2017. Tapping the genetic diversity of landraces in allogamous crops with doubled haploid lines: a case study from European flint maize. *Theor. Appl Genet* 130: 861-873.
- Botstein D, White RL, Skolnick M, Davis RW. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 2: 314-331.
- Boakyewaa Adu G, Badu-Apraku B, Akromah R, Garcia-Oliveira AL, Awuku FJ, Gedil M. 2019. Genetic diversity and population structure of early-maturing tropical maize inbred lines using SNP markers. *PLoS ONE* 14 (4): e0214810. DOI: 10.1371/journal.pone.0214810.
- Collard BCY, Jahufer MMZ, Brouwer JB, Pang ECK. 2005. An Introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica* 142:169-196
- Dao, A., J. Sanou, V. Gracen, E.Y. Danquanh. 2014. Heterotic relationship between INERA, CIMMYT and IITA maize inbred lines under drought and well-watered conditions. *Maydica* 59: 201-210
- Earl DA, von Holdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4: 359-361
- Giordani W, Scapim CA, Ruas PL, Ruas CF, Contreras R, Coan M, Fonseca ICB, Goncalves LS. 2019. Genetic diversity, population structure and AFLP markers associated with maize reaction to southern rust. *Bragantia* 7 (2): 183-196.
- Gower JC. 1971. A general coefficient of similarity and some of its properties. *Biometrics* 27 (4): 857-871.
- Guan H, Wang S, Zhang R, Meng Q, Wen H. 2020. Maize genetic diversity and heterosis group. *Rev Fac Agron (LUZ)* 37 (2): 1093-1101.
- Holker AC, Mayer M, Presterl T, et al. 2019. European maize landraces made accessible for plant breeding and genome-based studies. *Theor Appl Genet* 132: 3333-3345 .
- Makumbi D, Assanga S, Diallo A, Magorokosho C, Asea G, Worku M, Banziger M. 2018. Genetic analysis of tropical Midaltitude adapted maize populations under stress and non-stress conditions. *Crop Sci* 58: 1492-1507.
- Maxted N, Ford-Lloyd B, Hawkes JG. 1997. Complementary conservation strategies In: Maxted N, Ford-Lloyd B, Hawkes JG (eds.). *Plant Genetic Conservation: The in situ Approach*. Chapman & Hall, London.
- Meirmans PG. 2012. AMOVA-Based clustering of population genetic data. *J. Hered* 103 (5): 744-750.
- Miti F, Tongoona P, Derera J. 2010. S1 selection of local maize landraces for low soil nitrogen tolerance in Zambia. *Afr J Plant Sci* 4 (3): 67-81
- Musundire L, Derera J, Dari S, Tongoona P. 2019. Genetic variation and path analysis of introgressed maize inbred lines for economic traits. *J Agric Sci* 11: 17. DOI:10.5539/jas.v11n17p135.
- Nda HA, Akanvou L, Pokou ND, Akanza KP, Kouakou CK, Zoro Bi IA. 2016. Genetic diversity and population structure of maize landraces from Côte d'Ivoire. *Afr J Biotechnol* 15 (44): 2507-2516.
- Nei M. 1973. Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* 70: 3321-3323.
- Nelimor C, Badu-Aprakub B, Nguetta SPA, Tetteh AY, Garcia-Oliveira AL. 2019. Phenotypic characterization of maize landraces from Sahel and Coastal West Africa reveals marked diversity and potential for genetic improvement. *J Crop Improv* 34 (1): 122-138.
- Nelson PT, Krakowsky MD, Coles ND, Holland JB et al. 2016. Genetic characterization of the North Carolina State University maize lines. *Crop Sci* 56: 259-276.

- Nyaligwa L, Hussein S, Amelwork B, Ghebrehiwot H. 2015. Genetic diversity analysis of elite maize inbred lines of diverse sources using SSR markers. *Maydica* 60-2015: M29.
- Pagnotta MA. 2018. Comparison among methods and statistical software packages to analyze germplasm genetic diversity by means of codominant markers. *J Multidiscip Res* 1: 197-215 .
- Peakall R, Smouse PE. 2012. GenAlex 6.5: genetic analysis in Excel. *Populations genetic Software for Teaching and Research. Bioinformatics* 28 (19): 2537-2539.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Rezende WS, Beyene Y, Mugo S, et al. 2019. Performance and yield stability of maize hybrids in stress-prone environments in eastern Africa. *Crop J* 8 (1): 107-118 DOI: 10.1016/j.cj.2019.08.001.
- R Development Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- Saghai-Marouf MA, Soliman KM, Jorgensen RA, Allard RW. 1984. Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. *Proc Natl Acad Sci USA* 81: 8014-8018.
- Shete S, Tiwari H, Elston RC. 2000. On estimating the heterozygosity and polymorphism information content value. *Theor Popul Biol* 57: 265-271 .
- Springer NM, Stupar RM. 2007. Allelic variation and heterosis in maize how do two halves make more than a whole? *Genome Res* 17 (3): 264-275.
- Strigens A, Schipprack W, Reif JC, Melchinger AE. 2013. Unlocking the genetic diversity of maize landraces with double haploid opens new avenues for breeding. *PLoS ONE* 8 (2): e57234. DOI: 10.1371/journal.pone.0057234.
- Whitt SR, Wilson LM, Tenailon MI, Gaut BS, Buckler ES. 2002. Genetic diversity and selection in the maize starch pathway. *Proc Natl Acad Sci USA* 99 (20): 12959-12962.
- Zhang X, Zhang H, Li L et al. 2016. Characterizing the population structure and genetic diversity of maize breeding germplasm in Southwest China using genome-wide SNP markers. *BMC Genomics* 17: 697. DOI: 10.1186/s12864-016-3041-3.
- Zikhali M, Wingen LU, Leverington-Waite M, Specel S, Griffiths S. 2017. The identification of new candidate genes *Triticum aestivum* FLOWERING LOCUS T3-B1 (TaFT3-B1) and TARGET OF EAT1 (TaTOE1-B1) controlling the short-day photoperiod response in bread wheat. *Plant Cell Environ* 40 (11): 2678-2690.