

UDC 575.633.15

DOI: 10.2298/GENSR1603067N

Original scientific paper

ESTABLISHMENT AND CONFIRMATION OF HETEROTIC GROUPS AND GENETIC DIVERSITY ASSESSMENT OF MAIZE INBRED LINES USING MICROSATELLITE DATA

Ana NIKOLIĆ, Jelena BOGOSAVLJEVIĆ, Zoran ČAMDŽIJA, Milomir FILIPOVIĆ,
Dragan KOVAČEVIĆ, Milan STEVANOVIĆ, Snežana MLADENOVIĆ DRINIĆ

Maize Research Institute "Zemun Polje", Belgrade, Serbia

Nikolić A. J. Bogosavljević, Z. Čamdžija, M. Filipović, D. Kovačević, M. Stevanović, S. Mladenović Drinić (2016): *Establishment and confirmation of heterotic groups and genetic diversity assessment of maize inbred lines using microsatellite data*. - Genetika, Vol 48, No.3, 1067 - 1076.

Twenty-seven maize inbreds (12 commercial and 15 developing lines) from Maize Research Institute breeding program were subjected to microsatellite analysis. The aim was genetic diversity determination, establishing relationships among tested lines and assigning them to heterotic groups according to molecular marker data. Number of alleles detected was 97, with an average of 3.23. Major allele frequency was in a range from 0.33 to 0.82 (average 0.55). The highest value for observed heterozygosity was 10% for several developing lines. Mean values for gene diversity and PIC were 0.56 and 0.48, respectively. Frequency-based distances were calculated using *Roger's* coefficient and average value of 0.57 indicates high genetic diversity in analyzed maize inbreds. Distance matrices were subjected to cluster analysis and PCA. Multivariate analysis methods showed considerable concurrency with pedigree data. Results of analysis with 30 microsatellite markers could be useful for defining/redefining heterotic groups but should be complemented with field testing data.

Keywords: genetic diversity, heterotic groups, maize, SSR

INTRODUCTION

The era of maize hybrids creation and exploitation started more than hundred years ago. Meanwhile, many efforts were invested in gaining the best, considering maize production. Importance of this crop for human population is not declining, although constant endeavors are present in searching for new sources of energy and nutrients.

Corresponding author: Ana Nikolić, Maize Research Institute "Zemun Polje", Slobodana Bajića 1, 11185 Zemun Polje, Serbia, E-mail: anikolic@mrizp.rs, Phone: + 381 64 840 6216

Maize breeding process includes creation of inbred lines through selfing in number of generations. Inbreeding and selection result in formation of genotypes that should contain homozygous alleles at about 95% or more of its loci. Development of maize hybrids with increased hybrid vigor requires crossing of homozygous lines from different heterotic groups. LEE (1995) defines heterotic group as a collection of germplasm which shows higher degree of heterosis (hybrid vigor) when crossed with germplasm from different heterotic group, then when crossed to genotype from its own group. It has been confirmed in many cases that genotypes belonging to different heterotic groups which exhibit high heterosis when crossed are genetically more diverse than combinations of inbreds which show low heterosis. Thus, proper classification of maize lines into heterotic groups plays important role in successful maize breeding and producing maize hybrids with high yielding potential.

Traditionally, conventional breeding methods have been generally used for assigning maize inbred lines into heterotic groups and choosing suitable parental combinations. These methods use information about specific combining ability with some line-pedigree data and/or information about hybrid yield (LIBRANDO and MAGULAMA, 2008; LEGESSE *et al.*, 2009). Besides this approach, establishment and confirmation of heterotic groups could be aided by different DNA marker techniques (HUANG *et al.*, 2001; BARATA and CARENA, 2006; RAJENDRAN *et al.*, 2014). Molecular markers help in assessing genetic diversity, structure and relationship among genetic material under investigation. These data then can be used for defining/redefining heterotic patterns.

Evolution of DNA markers since 1980s led from hybridisation based RFLPs to SNP markers and chip technology creation, in an effort to make these methods more accurate and more highthroughput (HENRY, 2012). Besides certain drawbacks, some types of molecular markers are still frequently used. Microsatellite markers have been widely applied in maize molecular genetic studies in last two decades (WANG *et al.*, 2008; AFAF *et al.*, 2009; PINEDA-HIDALGO, 2013). Codominant nature and known chromosomal position, reproducibility and high level of polymorphism of SSR markers ensure their utilization in the future.

The aim of this study was to get detailed information about genetic diversity patterns in commercial as well as in developing maize inbreds from Maize Research Institute using SSR markers. Also, the goal was to compare the data about line grouping revealed using molecular marker analysis and already available pedigree data.

MATERIALS AND METHODS

Twenty-seven inbred lines from Maize Research Institute breeding programmes were genotyped using SSR molecular markers. Among the chosen genotypes, 12 inbreds are commercial inbred lines and the remaining 15 are developing lines (S4) (Table 1a, b).

Genotyping was done using 30 SSR markers chosen on the basis of bin location which provides a uniform coverage of all ten chromosomes in the maize genome. Three markers were selected for each of 10 maize chromosomes.

Genomic DNA isolation, PCR amplification and electrophoresis were done according to NIKOLIC *et al.* (2015). The size of the amplified fragments was determined using 20 bp DNA ladder (Thermo Scientific) and these data were subjected to statistical analysis.

The summary statistics (allele number, major allele frequency, gene diversity (expected heterozygosity), observed heterozygosity, and PIC (*Polymorphism Information Content*) was calculated using PowerMarker V3.25. Frequency of rare and unique alleles was also calculated.

Genetic distances based on allele-frequency were determined according to ROGERS (1972) and dendrogram was constructed using UPGMA method in PowerMarker, 3.5. Visualization of clusters was done in MEGA 6.06. Matrices of genetic distances were subjected to *Principal Component Analysis* (PCA) implemented in NTSYS, 2.11a.

Table 1a – Pedigree of 12 commercial maize inbred lines

Commercial lines	Pedigree
L1	BSSS
L2	BSSS
L3	BSSS
L4	BSSS
L5	Lancaster Sure Crop
L6	Lancaster Sure Crop
L7	Lancaster Sure Crop
L8	Iowa dent
L9	Iowa dent
L10	Unrelated (unknown)
L11	Unrelated (unknown)
L12	Unrelated (unknown)

Table 1b Pedigree of 15 maize developing lines

Developing lines	Pedigree
L13	Iowa dent
L14	Ohio x BSSS
L15	Iowa dent
L16	BSSS
L17	BSSS
L18	BSSS
L19	Iowa dent
L20	Lancaster Sure Crop
L21	BSSS x Iowa dent
L22	Iowa dent
L23	Iowa dent
L24	BSSS
L25	Iowa dent
L26	BSSS
L27	BSSS x Iowa dent

RESULTS AND DISCUSSION

Genetic diversity parameters

Average number of alleles identified was 3.23, while total for all 27 analyzed inbreds was 97. Nearly the same sample size (36 maize lines) and number of SSR markers (25) were analyzed in DEMISSEW *et al.* (2015), and the number of alleles detected was also almost the same (98 alleles), with an average of 3.9 per marker. The same value for average number of alleles (3.9) was presented in LEGESSE *et al.* (2007), detected with the same number of SSR markers (27) as in our study. Number of alleles scored per marker was in a range from 2 (for 6 out of 30 microsatellites studied) to 6 (for two markers with dinucleotide repeats). The results confirmed the highest number of alleles for markers with dinucleotide repeats, as many times stated in different studies (LI *et al.*, 2006; ADETIMIRIN *et al.*, 2008; SSERUMAGA *et al.*, 2014). LANES *et al.* (2014) analysed 90 lines derived from tropical hybrids and populations and with 81 polymorphic SSRs generated 471 alleles with an average of 5.8 alleles per marker. These parameters were also much higher (total number of alleles 1082 and average number of alleles 7.21) in PARK *et al.* (2015) and possible reason is different number and type of SSR markers and sample size used in diverse studies. The minimum value calculated for MAF (major allele frequency) was 0.31, while the highest score was 0.82 and average 0.55. Gene diversity was in a range from 0.32 to 0.71 with a mean of 0.56. SSERUMAGA *et al.* (2014) reported similar value of this parameter - 0.60 for tropical inbreds, while in WANG *et al.* (2008) average gene diversity for maize lines from China was higher - 0.68 probably due to higher average number of alleles per locus. The lowest observed heterozygosity per marker was 0.037 and the highest 0.37 with average of 0.05. The highest value of observed heterozygosity per line was 10%, which is expected for developing lines after three generations of selfing. Some level of heterozygosity is detected in commercial inbreds, too and possible reason could be residual heterozygosity or duplication (the amplification of similar sequences in two different genomic regions). In addition, mean value for PIC was 0.48, and this parameter ranged from 0.29 to 0.69. Molecular markers are considered highly informative when PIC value is greater than 0.50 and one half (15) of SSRs used in this study belongs to that category. The rest of them a reasonably informative with a PIC value between 0.30 and 0.50. DEMISSEW *et al.* (2015) reported similar average value for PIC (0.491), although greater values were presented in other studies (REID *et al.* 2011; PARK *et al.*, 2015). One third of molecular markers used were with dinucleotide repeats but relatively low PIC detected could be caused by high level of genetic similarity among analyzed genotypes. The data about microsatellites used and summary of all discussed genetic diversity parameters are presented in Table 2.

Percent of rare alleles was 13.4% (frequency <0.05), while intermediate (frequency <0.50) and abundant (frequency > 0.50) comprised 69% and 17.5% of all detected alleles respectively. Artificial selection influences genic and intergenic regions and results in reduced diversity and greater number of rare alleles (JIAO *et al.*, 2012). CHOUKAN *et al.* (2005) found 44 unique alleles in 38 inbred lines with 43 SSR markers and KUMAR *et al.* (2008) detected 9 alleles specific for only one inbred in 16 lines with 24 microsatellites. Eleven unique alleles were found in lines L1, L4, L7, L11, L12 and L18 in this study. Although percent of rare and unique alleles is smaller comparing to other authors, these could identify several lines uniquely.

Table 2 Summary of genetic diversity data

Marker	bin	repeat	AlleleNo	Major.Allele. Frquency	GeneDiversity	Heterozygosity	PIC
umc1269	1.01	(CCT)4	3	0.41	0.66	0	0.59
umc1568	1.02	(TAG)4	3	0.58	0.52	0	0.42
umc2047	1.09	(GACT)4	2	0.59	0.48	0.074	0.37
umc1265	2.02	(TCAC)4	3	0.42	0.64	0	0.56
umc1465	2.04	(ACACA)4	3	0.44	0.63	0.37	0.55
umc1520	2.09	AG(22)	4	0.58	0.54	0.039	0.46
phi036	3.04	AG	3	0.48	0.61	0	0.53
bnlg197	3.06	-	4	0.57	0.6	0.037	0.56
bnlg1350	3.08	AG(13)	3	0.56	0.55	0.074	0.46
umc1288	4.02	(TCCA)4	3	0.82	0.32	0	0.29
umc1651	4.07	(GA)6	2	0.74	0.38	0.148	0.31
umc1109	4.1	(ACG)4	4	0.7	0.45	0.074	0.4
bnlg557	5.03	-	6	0.41	0.73	0	0.69
umc1274	5.03	(TGC)5	3	0.48	0.61	0	0.53
umc1019	5.06	(CT)17	6	0.44	0.68	0	0.63
umc1006	6.02	(GA)19	4	0.72	0.43	0.111	0.37
umc1887	6.03	(CGA)4	2	0.71	0.41	0.039	0.33
phi102	6.05	AT	3	0.48	0.61	0	0.53
umc1695	7	(CGA)4	3	0.62	0.54	0	0.47
umc1324	7.03	(AGC)5	2	0.52	0.5	0	0.38
umc1782	7.04	(GAC)4	2	0.69	0.43	0	0.34
bnlg2235	8.02	AG(23)	5	0.56	0.62	0.222	0.57
phi080	8.08	AGGAG	3	0.52	0.56	0	0.47
umc1638	8.09	(CTCCGG)5	3	0.44	0.6	0.111	0.52
umc1040	9.01	(CT)11	4	0.46	0.65	0.037	0.58
umc1492	9.04	(GCT)4	3	0.37	0.66	0	0.59
umc1310	9.06	(GCG)5	2	0.78	0.35	0	0.29
umc1336	10.03	(ACCAG)4	3	0.48	0.63	0	0.56
umc1506	10.05	(AACAA)4	4	0.33	0.71	0	0.65
umc1645	10.07	(CT)10	2	0.52	0.5	0.148	0.38
Mean			3.23	0.55	0.56	0.05	0.48

Genetic distance

Roger's genetic distance ranged from 0.083 to 0.83. The lowest level of genetic distance was detected between lines L5 and L6 which is expected according to pedigree data. These lines belong to Lancaster Sure Crop (LSC) heterotic group. The most distant inbred lines were L4 and L24, and both genotypes belong to BSSS heterotic group. This result could be caused by errors in pedigree data recording. Besides, almost the same genetic distance (0.82) was found between L7 and L13 and L14 and L25, respectively, and both line pairs are from different heterotic groups according to already available data. Average genetic distance was 0.57, with 76% values

in a range from 0.50 to 0.80 indicating considerable genetic diversity in 27 maize lines. Similar data were presented in DEMISSEW *et al.* (2015).

Multivariate data analysis

Multivariate methods – cluster analysis and PCA were done in order to get better insight in genetic diversity patterns among studied lines.

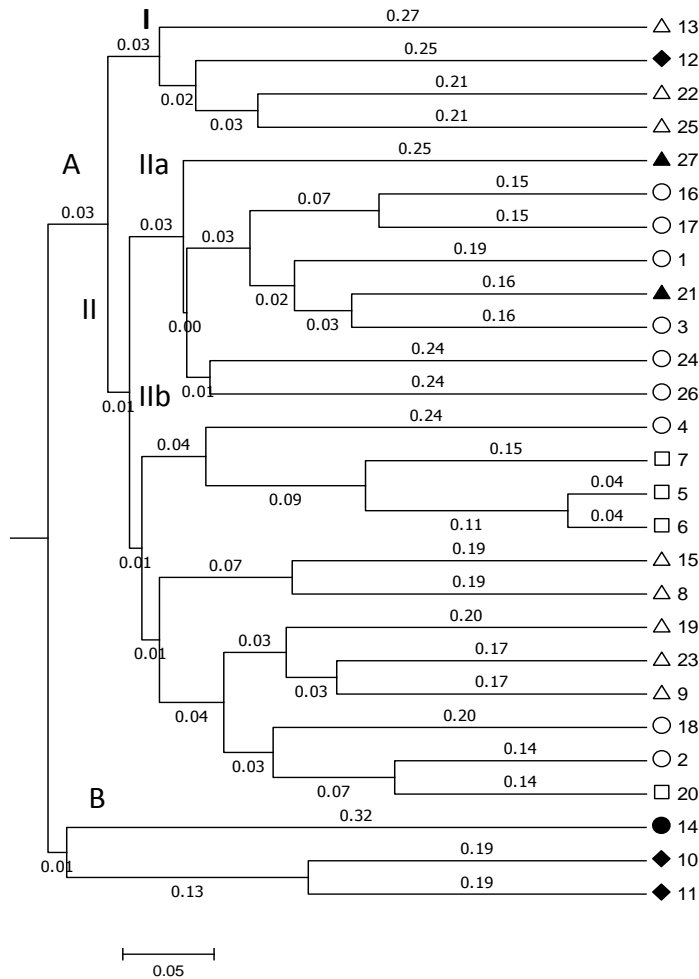


Figure 1 UPGMA dendrogram according to *Roger's* distance. Pedigree is presented with symbols: ○ - BSSS, □ - Lancaster Sure Crop, △ - Iowa dent, ◆ - unknown, ▲ - IowaxBSSS ● - Ohio

Cluster analysis showed moderate correspondence with pedigree data and classified 27 maize genotypes in two main groups (A and B). The larger main group (A) consisted of two subgroups (I and II). Iowa dent (ID) germplasm (lines L13, L22 and L25) constituted group I

with the exception of L12 with unknown pedigree. Group II comprised two subgroups (IIa and IIb). Lines of BSSS pedigree constituted IIa group with the exception of L21 and L27 with BSSSxIowa dent background. These two lines possibly contain more BSSS background. Group IIb was subdivided in group consisted of LSC lines with the exception of L4 (BSSS pedigree) and heterogeneous group comprised of lines belonging to all three heterotic groups under study. The smaller group (B) included only three developing lines. Two inbreds (L10 and L11) belong to germplasm of unknown origin, and the third line (L14) has one different component (Ohio heterotic group) of pedigree comparing to all other 27 genotypes which explains its position in dendrogram. It might be that two lines with unknown pedigree comprise germplasm of Ohio heterotic group origin too, considering their grouping into the same cluster with L14.

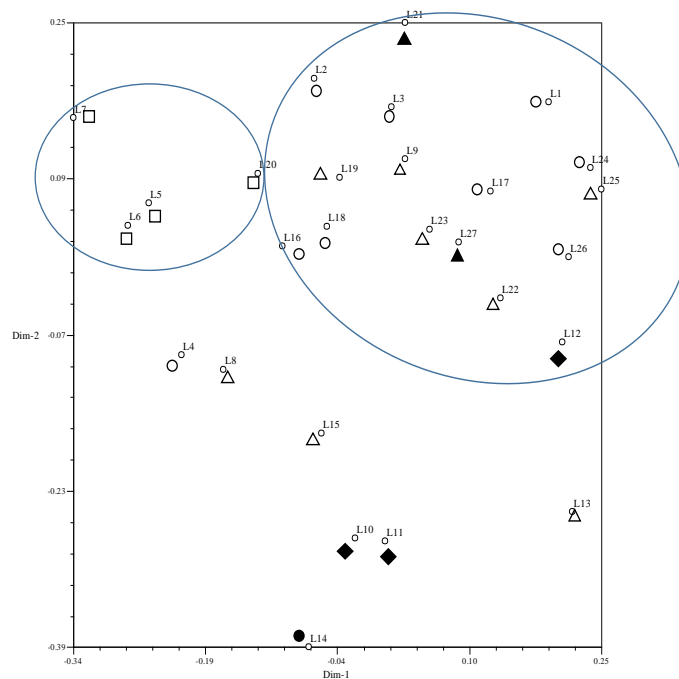


Figure 2 PCA analysis of 27 maize inbreds. Pedigree is presented with symbols:

○ BSSS, □ - Lancaster Sure Crop, △ - Iowa dent ◆ - unknown, ▲ - IowaxBSSS
● - Ohio

Results of PCA showed different grouping of inbreds comparing to cluster analysis. Most of inbreds were arranged in two main groups. One was consisted only of LSC lines, and the other comprised lines of BSSS and ID pedigree. Seven lines of different pedigree were scattered in PCA plot (did not belonged to defined groups). Two lines with unknown germplasm (L10 and

L11) were located closely to each other suggesting common genetic background. Line L14, the only one with Ohio heterotic group germplasm was positioned separately from all other lines.

Inconsistencies in cluster and PCA grouping with pedigree data were observed in many studies with the same object of research (LU *et al.*, 2009; DEMISSEW *et al.*, 2015).

CONCLUSIONS

This study proved that 30 SSR markers were sufficient to clearly separate 27 maize inbred lines, although higher number and different types of microsatellites could give better insight in genetic structure and diversity of analyzed germplasm sample. Heterotic grouping according to microsatellite data should be complemented with the data from field experiments. In spite of disagreements among different types of data (pedigree, combining abilities, molecular genetics data), molecular markers are valuable tool that could help in establishing/confirmation of maize heterotic groups, correcting errors in pedigree data records and elucidating genetic background of tested genotypes and contribute to designing better breeding strategies.

ACKNOWLEDGMENTS

This research is a part of the ongoing project TR-31068 of the Ministry of Education, Science and Technological Development of the Republic of Serbia.

Received February 22th, 2016

Accepted August 25th, 2016

REFERENCES

- ADETIMIRIN, V.O., I. VROH-BI, C. THE, A. MENKIR, S.E. MITCHELL, S. KRESOVICH (2008): diversity analysis of elite maize inbred lines adapted to west and central Africa using SSR markers. *Maydica*, 53: 143-149.
- AFAF, I. SHEHATA, H. A. AL-GHETHAR, A. A. AL-HOMAIDAN (2009): Application of simple sequence repeat (SSR) markers for molecular diversity and heterozygosity analysis in maize inbred lines. *Saudi Journal of Biological Sciences* Volume 16, Issue 2, October, Pages 57–62.
- BARATA, C., M.J. CARENA (2006): Classification of North Dakota maize inbred lines into heterotic groups based on molecular and testcross data. *Euphytica*, 151: 339-349.
- DEMISSEW, A., H. SHIMELIS, J. DERERA, S. KASSA (2015): Genetic purity and patterns of relationships among tropical highland adapted quality protein and normal maize inbred lines using microsatellite markers. *Euphytica*, 204:49–61.
- HENRY, R.J. (2012): Evolution of DNA Marker Technology in Plants. *Molecular Markers in Plants*: Blackwell Publishing Ltd.; p. 1–19.
- HUANG, Y.Q., J.S. LI (2001): Classification of heterotic groups with RFLPs among 45 maize inbred lines. *Agric. Sin. China*, 34(3):244–250.
- JIAO Y., H. ZHAO, L. REN, W. SONG, B. ZENG, J. GUO, B. WANG, *et al.* (2012): Genome-wide genetic changes during modern breeding of maize. *Nature Genetics*, 44: 812 – 815.
- KUMAR, B., S. RAKSHIT, R.D. SINGH, R.N. GADAG, R. NATH, A.K. PAUL (2008): Genetic Diversity of Early Maturing Indian Maize (*Zea mays* L) Inbred Lines Revealed by SSR Markers *J. Plant Biochemistry & Biotechnology*, Vol. 17(2): 133-140.
- LANES, E.C.M., J.M.S. VIANA, G.P. PAES, M.F.B. PAULA, C. MAIA, E.T. CAIXETA, G.V. MIRANDA (2014): Population structure and genetic diversity of maize inbreds derived from tropical hybrids *Genet. Mol. Res.*, 13 (3): 7365-7376.
- LEE, M. (1995): DNA markers and plant breeding programs. *Adv. Agron.*, 55:265-344.

- LEGESSE B.W., A.A. MYBURG, K.V. PIXLEY, A.M., BOTHA (2007): Genetic diversity of African maize inbred lines revealed by SSR markers. *Hereditas*, *144*: 10-17.
- LEGESSE, B.W., K.V. PIXLEY, A.M. BOTHA (2009): Combining ability and heterotic grouping of highland transition maize inbred lines *Maydica*, *54 (1)*: 9-14.
- LI, M.S., X.H. LI, S. SALVI, R. TUBEOSA, L.X., YUAN, F. ROTONDO, L. BAI, S.H. ZHANG (2006): Genetic relationships among CIMMYT subtropical QPM and Chinese maize inbred lines based on SSRs. *Maydica*, *51*: 543-549.
- LIBRANDO, R.P., E.E. MAGULAMA (2008): Classifying white inbred lines into heterotic groups using yield combining ability effects. *USM R & D J*, *16(1)*: 99-103.
- LIU K., M. GOODMA, S. MUSE, J.S. SMITH, E. BUCKLER, J. DOEBLEY (2003): Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics*, *165*: 2117-2128.
- LU Y., J. YAN, C. GUIMARAES, S. TABA, Z. HAO, S. GAO, S. CHEN, J. LI, S. ZHANG, B. VIVEK, C. MAGOROKOSHO, S. MUGO, D. MAKUMBI, S. PARENTONI, T. SHAH, T. RONG, J. CROUCH, Y. XU (2009): Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. *Theor. Appl. Genet.*, *120*:93–115.
- NIKOLIĆ, A., D. IGNJATOVIĆ-MIČIĆ, D. KOVAČEVIĆ, Z. ČAMDŽIJA, M. FILIPOVIĆ, S. MLADENOVIĆ DRINIĆ (2015): Genetic diversity of maize inbred lines as inferred from SSR markers. *Genetika*, *47 (2)*: 489-498.
- ONTIVEROS, J., P. SÁNCHEZ-PEÑA, J.A. GARZÓN-TIZNADO, M.O. VEGA-GARCÍA, J.A. LÓPEZ-VALENZUELA (2013): Microsatellite-based genetic diversity among accessions of maize landraces from Sinaloa in México. *Hereditas*, *150*: 53 – 5.
- PARK, J.Y., R.V. RAMEKAR, K. J. SA, J. K. LEE (2015): Genetic diversity, population structure, and association mapping of biomass traits in maize with simple sequence repeat markers. *Genes Genom.*, *37*:725–735.
- PINEDA-HIDALGO, K.V., K.P. MÉNDEZ-MARROQUÍN, E.V. ALVAREZ, J. CHÁVEZ-ONTIVEROS, P. SÁNCHEZ-PEÑA, J.A. GARZÓN-TIZNADO, M.O. VEGA-GARCÍA, J.A. LÓPEZ-VALENZUELA (2013): Microsatellite-based genetic diversity among accessions of maize landraces from Sinaloa in México *Hereditas. Dec*, *150(4)*:53-9.
- RAJENDRAN, A., A. MUTHIAH, J. JOEL, P. SHANMUGASUNDARAM, *et al.* (2014): Heterotic grouping and patterning of quality protein maize inbreds based on genetic and molecular marker studies. *Turk. J. Biol.*, *38*: 10-20.
- REID ,L.M., K. XIANG, X. ZHU, B. R. BAUM, S. J. MOLNAR (2011): Genetic diversity analysis of 119 Canadian maize inbred lines based on pedigree and simple sequence repeat markers. *Can. J. Plant Sci.*, *91*: 651-661
- ROGERS, J.S. (1972): Measures of genetic similarity and genetic distances. *Studies in Genetics*, Univ. Texas Publ., *7213*: 145–153.
- SSERUMAGA, J. P., D. MAKUMBI, J. HYEONSO, K. NJOROGI, J. W. MUTHOMI, G. N. CHEMINING'WA, L. SI-MYUNG, G. ASEA, H. KIM (2014): Molecular characterization of tropical maize inbred lines using microsatellite DNA markers. *Maydica*, *59*:267-274.
- WANG, R., Y. YU, J. ZHAO, Y. SHI *et al.* (2008): Population structure and linkage disequilibrium of a mini core set of maize inbred lines in China. *Theor. Appl.Genet.*, *117*: 1141-1153.

**DEFINISANJE HETEROTIČNIH GRUPA INBRED LINIJA KUKURUZA I
ISPITIVANJE GENETIČKOG DIVERZITETA PRIMENOM MIKROSATELITA**

Ana NIKOLIĆ, Jelena BOGOSAVLJEVIĆ, Zoran ČAMDŽIJA, Milomir FILIPOVIĆ,
Dragan KOVAČEVIĆ, Milan STEVANOVIĆ, Snežana MLADENOVIĆ DRINIĆ

Institut za kukuruz „Zemun Polje“

Izvod

Dvadeset sedam samooplodnih linija (12 komercijalnih i 15 linija u razvoju) iz programa oplemenjivanja Instituta za kukuruz analizirano je mikrosatelitima. Cilj ovog istraživanja je bilo utvrđivanje genetičkog diverziteta testiranih linija i svrstavanje u heterotične grupe na osnovu podataka dobijenih molekularnim markerima. Detektovano je ukupno 97 alela, a prosečan broj po markeru iznosio je 3.23. Frekvencija najučestalijeg alela je bila u opsegu od 0.33 do 0.82 (prosečna vrednost 0.55). Najveći procenat heterozigotnosti od 10% uočen je kod nekoliko linija u razvoju. Srednje vrednosti genetičkog diverziteta i PIC parametra bile su 0.56 i 0.48 respektivno. Genetičke distance su izračunate na osnovu frekvencija korišćenjem koeficijenta po Rogers-u i srednja vrednost od 0.57 ukazuje na značajan diverzitet ispitivanih linija. Matrice distanci korišćene su za klaster analizu i PCA. Rezultati ove analize ukazuju na to da podaci dobijeni pomoću mikrosatelita mogu biti korisni u definisanju/redefinisanju heterotičnih grupa ali je potrebno da budu dopunjeni rezultatima testova urađenih u polju.

Primljeno 22. II. 2016.

Odobreno 25. VIII. 2016.