MARKER ASSISTED BACKCROSS BREEDING IN MAIZE RESEARCH INSTITUTE ZEMUN POLJE

Marija Kostadinović[*1], Ana Nikolić[1], Danijela Ristić[1], Sofija Božinović[1], Olivera Đorđević Melnik[1], Dragana Ignjatović-Micić[1], Jelena Vančetović[1]

Abstract

The backcross breeding is one of the most extensively used methods in different crop species. With the development of molecular genetics, molecular markers become widely used as a tool aiding conventional breeding. The main goal of marker assisted backcross breeding (MABB), a form of marker assisted selection (MAS), is to develop backcross progenies carrying the gene of interest with maximum recovery of the recurrent parent's genome (RPG). In this research, twelve inbred lines were chosen for the introgression of favourable alleles for several traits. Thirty SSR markers distributed throughout the whole genome were used to determine genetic similarity among these twelve lines and their BC_4 progenies, i.e.to identify the genotypes with the highest proportion of recurrent parent's genome (RPG). Genetic similarity values ranged from 0.79 to 0.99(79-99% RPG).The 48% of progenies had RPG above theoretical value and they were self-pollinated for allele fixation. Also, 41% of them had 99%, what is the value theoretically achieved in BC_6 generation. On the other hand, 52% of individuals had lower RPG content than theory predicts. The results of the research confirmed the advantage of using MABB compared to conventional approach, providing faster achievement of the goal, recovering the recurrent parent's genome in fewer generations of backcrossing and thus enabled enhanced selection efficiency and shortened the breeding process.

Key words: backcross breeding, SSR, marker assisted selection

Introduction

Crop breeding has long been based only on cycles of phenotypic selection and crossing, which resulted in the creation of superior genotypes and increased production (Hu et al., 2018). There are several methods commonly applied in conventional breeding. One of the methods is the backcross breeding and it is extensively used in different crop species (Oliveira et al., 2008; Hasan et al., 2015), including maize (Benchimol et al., 2005). The aim of this approach is to transfer desirable alleles at one or more loci from donors to elite recipient inbred lines. The main drawbacks of this method are: a large number of generations needed for recovering the genome of the recurrent parent and the presence of the parts of the donor's genome eventually linked to non-favourable alleles, surrounding the introgressed allele (Benchimol et al., 2005, Kostadinović et al., 2012).

Molecular genetic tools play an important role in crop breeding as well as in maize breeding for more than three decades. A large number of techniques and approaches have been developed since the beginning of the molecular tool era (Nadeem et al., 2018). Much effort has been invested in improving precision, accuracy, efficiency and cost-cut in this domain. According to available data, continuous and rapid development in this field will be seen in the future.

One of these techniques is SSR markers and since their discovery in the 1980s, they

played an important role in attempts to create superior genotypes more efficient. In spite of many more prominent tools developed later, simplicity of the technique, codominant inheritance, reproducibility, high polymorphism and cost-effectiveness (Sserumaga et al., 2014), qualified microsatellites for wide application in crop breeding till nowadays. These markers can be used for direct selection of the target gene (foreground selection), as well as for fast recovery of recurrent parent's genome (background selection).

The aim of this research was to check the applicability of SSR markers in accelerating the process of maize inbred lines improvement by comparing genetic similarity between recurrent parent and subsequent BC_4 generations of progeny created by crossing elite inbreds and donor lines with favourable alleles.

Material and Methods

Twelve maize inbred lines (L_1-L_{12}), components of the Maize Research Institute leading hybrids, were chosen for the introgression of favourable alleles involved in the expression of the following traits: restorer for CMS-C (RfC), restorer for CMS-S (RfS), blue kernel and red pericarp colour. These twelve lines and their BC_4 progenies were analyzed with molecular markers to determine genetic similarity, i.e. to identify the genotypes with the highest proportion of recurrent parent's genome.

Genomic DNA was isolated from the kernel bulk according to Doyle and Doyle (1987). Bulks were prepared by pooling an equal amount of flour obtained by grounding 10 kernels per sample. The DNA was quantified using bio spectrometer (BioSpetrometer kinetic, Eppendorf) and diluted to a working concentration of 20 ngμL^{-1}.

Simple sequence repeat (SSR) analysis was done with 30 polymorphic primers spanning over the whole genome (Table 1). Polymerase chain reaction (PCR) was carried out in 25 μL reaction volume containing: DreamTaq™ Green PCR Master Mix (2×), 0.5 μM primers and 20ng DNA. The following touch-down

Table 1. List of primers (name, bin, repeat and sequence) used in SSR analysis
Tabela 1. Lista prajmera (naziv, bin, ponovak i sekvenca) korišćenih u SSR analizi

Name	Bin	Repeat	Sequence (forward and reverse)
umc1282	1.00	(AT)6	5'-TACACTACACGACTCCCAACAGGA-3' 5'-GCGAGGGTTCTTTCCATAGAGAAT-3'
umc2230	1.05	(AGC)5	5'-AACGCGACGACTTCCACAAG-3' 5'-ACACGTAATGTCCCTACGGTCG-3'
bnlg1273	1.06	AG(13)	5'-AAACACCAAACGTCACGTGG-3' 5'-GGCGACGAGATACAGGATGT-3'
umc1013	1.08	(GA)9	5'-TAATGTGTCCATACGGTGGTGG-3' 5'-AGCTGGCTAGTCTCAGGCACTC-3'
umc2047	1.09	(GACT)4	5'-GACAGACATTCCTCGCTACCTGAT-3' 5'-CTGCTAGCTACCAAACATTCCGAT-3'
umc1265	2.02	(TCAC)4	5'-GCCTAGTCGCCTACCCTACCAAT-3' 5'-TGTGTTCTTGATTGGGTGAGACAT-3'
umc2129	2.07	(CGC)5	5'-ACGTGGTCATCACTCACCGC-3' 5'-AAGGAGGAGCGTTCTCGTGG-3'
bnlg1456	3.05	AG(15)	5'-TTCATGAGGACCGTGTTGAA-3' 5'-CTCTAGGTGGTTAAGATTAACTCATT-3'
umc1257	3.09	(CAC)4	5'-CAACGGAAGTGGCTGTAGAGTTTT-3' 5'-ACAGAGCATGTCAGGTATTTGCAG-3'

bnlg1784	4.07	AG(13)	5'-GCAACGATCTGTCAGACGAA-3' 5'-TTGGCATTGGTAATGGGTCT-3'
umc1418	4.08	(GGAAG)4	5'-TCACACACACACTACACTCGCAAT-3' 5'-GAGCCAAGAGCCAGAGCAAAG-3'
bnlg557	5.03	-	5'-TCACGGGCGTAGAGAGAGA-3' 5'-CGAAGAAACAGCAGGAGATGAC-3'
bnlg1046	5.03	AG(39)	5'-TGAGCCGAAGCTAACCTCTC-3' 5'-GATGCAAAGGAGGTTCAGGA-3'
phi087	5.06	ACC	5'-GAGAGGAGGTGTTGTTTGACACAC-3' 5'-ACAACCGGACAAGTCAGCAGATTG-3'
umc1792	5.08	(CGG)5	5'-CATGGGACAGCAAGAGACACAG-3' 5'-ACCTTCATCACCTGCAACTACGAC-3'
umc1006	6.02	(GA)19	5'-AATCGCTTACTTGTAACCCACTTG-3' 5'-AGTTTCCGAGCTGCTTTCTCT-3'
umc1887	6.03	(CGA)4	5'-CTTGCCATTTTAATTTGGACGTTT-3' 5'-CGAAGTTGCCCAAATAGCTACAGT-3'
umc2375	6.06	(GCG)4	5'-GCCGTACTGATGTGATGGTCC-3' 5'-TCTGACATTGTCCTCTTGACCAAA-3'
umc1695	7.00	(CA)8	5'-CAGGTAATAACGACGCAGCAGAA-3' 5'-GTCCTAGGTTACATGCGTTGCTCT-3'
umc1324	7.03	(AGC)5	5'-ATCCATCATCATCATCATTGCTTG-3' 5'-ATGTCATCATGTACCAGGTGTTGG-3'
umc1782	7.04	(GAC)4	5'-CGTCAACTACCTGGCGAAGAA-3' 5'-TCGCATACCATGATCACTAGCTTC-3'
umc2355	8.03	(CCT)5	5'-CTACTCCCCGAAGCCGTCTAAG-3' 5'-CGGGTTGTTGTTGGAGTAGGAC-3'
bnlg1782	8.05	AG(13)	5'-CGATGCTCCGCTAGGAATAG-3' 5'-TGTGTTGGAAATTGACCCAA-3'
phi080	8.08	AGGAG	5'-CACCCGATGCAACTTGCGTAGA-3' 5'-TCGTCACGTTCCACGACATCAC-3'
umc1040	9.01	(CT)11	5'-CATTCACTCTCTTGCCAACTTGA-3' 5'-AGTAAGAGTGGGATATTCTGGGAGTT-3'
umc1492	9.04	(GCT)4	5'-GAGACCCAACCAAAACTAATAATCTCTT-3' 5'-CTGCTGCAGACCATTTGAAATAAC-3'
umc1310	9.06	(GCG)5	5'-GAGGAAGAGTTGGCCAGGATG-3' 5'-AACTCCGAGATCTACGACAACAGC-3'
umc1827	10.04	(GAC)6	5'-GCAAGTCAGGGAGTCCAAGAGAG-3' 5'-CCACCTCACAGGTGTTCTACGAC-3'
umc1506	10.05	(AACA)4	5'-AAAAGAAACATGTTCAGTCGAGCG-3' 5'-ATAAAGGTTGGCAAAACGTAGCCT-3'
bnlg1839	10.07	(AG)24	5'-AGCAGACGGAGGAAACAAGA-3' 5'-TCTCCCTCTCCCTCTTGACA-3'

- data not available

43

program (thermocycler Biometra TProfessional Standard 96) was performed: an initial denaturation at 95ºC for 5min, followed by 15 cycles each of denaturation at 95°C for 30 s, annealing at 63.5ºC for 1min (-0.5ºC per cycle) and extension at 72ºC for 1min; another 22 cycles of 95ºC for 30 s, 56ºC for 1min and 72ºC for 1min with final elongation at 72°C for 4min. The PCR products were separated by electrophoresis on 8% polyacrylamide gel with 20bp molecular weight ladder as a marker. After staining with ethidium bromide, they were photographed under UV light using Biometra BioDocAnalyze gel documentation system.

SSR profiles were converted into a binary matrix based on the presence (1) or the absence (0) of a specific allele. Genetic similarity (GS) was calculated in accordance with Dice (1945): $GS_{ij} = 2a/2a+b+c$; where a is the number of fragments present in both variety i and j (1, 1), b is the number of fragments present in i and absent in j (1,0) and is the number of fragments absent in i and present in j (0,1). Marker data analyses were performed using statistical NTSYSpc2 program package (Rohlf, 2000).

Results and Discussion

Marker assisted backcross breeding (MABB) is widely used for the introgression of target trait(s) from a donor into the genome of a recipient parent. The main goal of this process is to develop backcross progenies carrying the gene of interest with maximum recovery of the recurrent parent's genome (RPG), which can take many generations in conventional breeding. Marker assisted selection for the genetic background of the recurrent parent in the backcrossed progenies (i.e. background selection) accelerates recovery of the recurrent parent genome (Hospital et al.,1992; Ribaut and Hoisington 1998; Wang et al., 2007). As concluded in the study by Hospitalet al. (1992), two generations can be saved by conducting marker assisted background selection.

In the present research, a total of 30 polymorphic SSR markers distributed throughout the genome were efficiently used for identification of the genotypes with the highest proportion of recurrent parent's genome in BC_4 generation (Table 2). Genetic similarity values between the recurrent parents and their

Table 2. Recurrent parent's genome (RPG)percentage for 12 analyzed lines and their corresponding BC4 progenies
Tabela 2. Procenat genoma rekurentnog roditelja (RPG) kod 12 analiziranih linija i njihovih BC4 potomstava

Line	Trait	RPG (%) range	RPG (%) average
L_1-$L_{1/4}$ [1]	Restorer for CMS-C	93-99	95.5
L_2-$L_{2/12}$	Restorer for CMS-C	95-99	97.7
L_3-$L_{3/17}$	Restorer for CMS-C	93-99	97.8
L_4-$L_{4/4}$	Restorer for CMS-S	99	99
L_5-$L_{5/5}$	Restorer for CMS-S	85-93	89
L_6-$L_{6/14}$	Restorer for CMS-S	91-98	94.9
L_7-$L_{7/12}$	Blue kernel	79-89	85.3
L_8-$L_{8/2}$	Blue kernel	92-95	93.5
L_9-$L_{9/5}$	Blue kernel	87-90	88
L_{10}-$L_{10/8}$	Red pericarp	86-96	93.8
L_{11}-$L_{11/3}$	Red pericarp	99	99
L_{12}-$L_{12/5}$	Red pericarp	94-96	95.2

[1] the last number denotes the number of BC4 sub-families (ears) of each line

corresponding BC_4 progenies ranged from 0.79 to 0.99 (79-99% RPG). The highest proportion of the recurrent parent's genome was found in L_4 and L_{11} progenies (99%), followed by L_2 (95-99%), L_1 and L_3 (93-99%). The lowest proportion of the RPG was found in L_7 progenies (79-89%), followed by L_9 (87-90%) and L_5 (85-93%). Average values for the RPG content ranged from 85.3 to 99%.

It has been demonstrated through simulation studies (Hospital et al., 1992) that increasing the number of markers to more than three per chromosome was not efficient in the early generations of backcrossing. On the other hand, Ribaut and Hoisington (2002) indicated that an increased number of markers should be considered as a way of optimizing the selection, due to the higher probability of crossover in later generations. In the study described in this paper, 60% of progenies had RPG content about the average value with their number decreased to the upper and lower values. Such an almost symmetrical distribution of progenies with low, moderate and high RPG content indicated a sufficient number of marker data points. Also, Ribaut and Hoisington (2002) have pointed out that experiments should be on a case-by-case basis considering the nature of the germplasm involved (e.g. agronomic quality and a number of lines to be converted, genetic similarity between donor and recurrent parent) and the technical options available at the marker level.

Using conventional breeding methods, it takes 6-8 backcrosses to fully recover there current parent's genome. Theoretically, the proportion of the RPG after n generations of backcrossing is given by $(2^{n+1} - 1)/2^{n+1}$ (Collard et al., 2005). These percentages are only achieved with large populations. They are usually lower with smaller population sizes that are typically used in actual plant breeding programs. Although the average percentage of the recurrent parent's genome is 96.9% for the BC_4 generation, someindividuals may possess more of the RPG than others. The recovery efficiency of recurrent parent's genome is depending on the choice of donor parent, concerning its origin and genetic similarity with RP. In our case, 48% of

progenies had RPG above this theoretical value and they were self-pollinated for allele fixation. Also, 41% of them had 99%, what is the value theoretically achieved in BC_6 generation, which shows that MABB made a genetic gain in RPG recovery. This result is similar to previous works (Benchimol et al., 2005). On the other hand, 52% of individuals had lower RPG content, probably due to the linkage between a target gene and nearby genes from the donor parent and/or chance (stochastic or non-random positions of chiasmata) (Semagn, 2006).

The study presented in this paper confirmed the benefit of the background selection in BC_4 generation compared to BC_2 and BC_3 because it reduces the time required for release of the improved line. As stated in Thakur et al. (2014), BC_2 plants with more than 87% RPG may also be used for selfing. However, BC_4 progenies with 99% genome of the recurrent parent may directly be used for hybrid creation. It can be concluded that this study represents the practical application of molecular markers in backcross breeding which can be routinely used to enhance selection efficiency and shorten the breeding process.

Conclusions

Application of 30 polymorphic SSR markers distributed throughout the genome was efficient for identification of the genotypes with the highest proportion of recurrent parent's genome in BC_4 generation. This analysis identified 48% of BC_4 progenies with RPG above the theoretical value, with41% of them having 99%, what is the value theoretically achieved in BC_6 generation. Our results confirmed the advantage of MABB by reducing the number of generations required to recover a recurrent parent's genome.

References

Benchimol LL, de Souza CL, de SouzaAP(2005): Microsatellite-assisted backcross selection in maize. Genet. Mol. Biol. 28(4): 789-797.

Collard BCY, Jahufer MZZ, 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.

Dice LR (1945): Measures of the amount of ecologic association between species. Ecology, 26: 297-302.

Doyle JJ, Doyle JL (1987): A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19: 11-15.

Frisch M, Bohn M, Melchinger AE (1999): Minimum sample size and optimum positioning of flanking markers in marker-assisted backcrossing for transfer of a target gene. Crop Sci. 39: 967-975.

Hasan MM, Rafii MY, Ismail MR, Mahmood M, Rahim HA, Alam MA, et al. (2015): Marker-assisted backcrossing: a useful method for rice improvement. Biotechnol. Biotechnol. Equip. 29: 237–254.

Hospital F, Chevalet C, Mulsant P (1992): Using markers in gene introgression breeding programs. Genetics, 132:1199-1210.

Hu H, Scheben A, Edwards D (2018): Advances in Integrating Genomics and Bioinformatics in the Plant Breeding Pipeline. Agriculture, 8: 75.

Kostadinović M, Ignjatović-Micić D, Stanković G, Mladenović Drinić S (2012): Biohemijska analiza F2 zrna genotipova kukuruza visokog kvaliteta proteina. Selekcija i semenarstvo, 18(1): 33-39.

Nadeem MA, Nawaz MA, Shahid MQ, Doğan Y, Comertpay G, Yıldız M, Hatipoğlu R, Ahmad F, Alsaleh A, Labhane N, Özkan H, Chung G, Baloch FSh (2018): DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing, Biotechnology & Biotechnological Equipment, 32(2): 261-285.

Oliveira LK, Melo LC, Brondani C, Peloso MJ, et al. (2008): Backcross assisted by microsatellite markers in common bean. Genet. Mol. Res. l7: 1000-1010.

Ribaut JM, Hoisington D (1998): Marker-assisted selection: new tools and strategies. Trends Plant Sci. 3(6): 236-239.

Ribaut JM, Hoisington D (2002): Simulation experiments onefficiencies of gene introgression by backcrossing. Crop Sci. 42: 557-565.

Rohlf FJ (2000): NTSYS-pc Numerical Taxonomy and Multivariate Analysis System version 2.1. Owner's manual.

Semagn K, Bjørnstad Å, Ndjiondjop MN (2006): Principles, requirements and prospects of genetic mapping in plants. Afr. J. Biotechnol. 25: 2569-2587.

Sserumaga JP, Makumbi D, Ji H, Njoroge K, Muthomi JW, Chemining'wa GN, Simyung L, Asea G, Kim H (2014): Molecular characterization of tropical maize inbred lines using microsatellite DNA markers. Maydica, 59: 267-274.

Thakur VV, Tiwari S, Tripathi N (2014): Studies on Genome Recovery Percentage for *Opaque2* Introgressed in Backcross Population for Marker Assisted Selection of Quality Protein Maize.Trends in Biosciences, 7(15): 1897-1901.

Wang J, Chapman SC, Bonnett DG, Rebetzke GJ, Crouch J(2007): Application of population genetic theory and simulation models to efficiently pyramid multiple genes via marker-assisted selection. Crop Sci. 47: 582-588.

PRIMENA MOLEKULARNIH MARKERA KOD POVRATNIH UKRŠTANJA U INSTITUTU ZA KUKURUZ ZEMUN POLJE

Marija Kostadinović, Ana Nikolić, Danijela Ristić, Sofija Božinović, Olivera Đorđević Melnik, Dragana Ignjatović-Micić, Jelena Vančetović

Sažetak

Povratno ukrštanje je pristup u klasičnom oplemenjivanju koji se koristi za unošenje jednog ili više poželjnih alela, na jedan ili više lokusa, iz donora u adaptiranu ili komercijalnu liniju (rekurentni roditelj). U poslednjih nekoliko decenija, zahvaljujući razvoju metoda molekularne genetike razvija se oplemenjivanje potpomognuto molekularnim markerima. U cilju bržeg dobijanja željenih genotipova (potomstvo koje nosi poželjni gen sa maksimalnim povraćajem genoma rekurentnog roditelja) primenjuje se kombinacija povratnog ukrštanja i selekcije pomoću SSR molekularnih markera. Za ovo istraživanje je odabrano dvanaest samooplodnih linija kukuruza (L_1-L_{12}) u cilju unošenja poželjnih alela uključenih u ekspresiju svojstava: restorer za CMS-C (RfC), restorer za CMS-S (RfS), za plavu boju zrna i za crvenu boju perikarpa. Samooplodne linije i njihovo BC_4 potomstvo su analizirani SSR molekularnim markerima u cilju identifikacije genotipova sa najvećim procentom genoma rekurentnog roditelja. Za analizu je odabrano 30 SSR markera raspoređenih na svih 10 hromozoma kukuruza. Vrednosti genetičke sličnosti između rekurentnih roditelja i njihovih BC_4 potomstava su bile u opsegu od 0.79 do 0.99 (79-99% genoma rekurentnog roditelja). Rezultati su pokazali da je 48% potomstva imalo povraćaj genoma rekurentnog roditelja iznad teoretske vrednosti i ono je podvrgnuto procesu samooplodnje u cilju fiksacije alela. Takođe, kod 41% potomstva povraćaj je iznosio 99%, što predstavlja vrednost koja se teoretski dostiže u BC_6 generaciji. Sa druge strane, 52% potomstva je imalo manji povraćaj genoma rekurentnog roditelja od teorijskog. U ovom istraživanju prikazana je praktična primena molekularnih markera u povratnom ukrštanju i ovaj pristup se može koristiti za postizanje bolje efikasnosti i skraćivanje procesa selekcije.

Ključne reči: povratno ukrštanje, SSR, selekcija pomoću molekularnih markera