Endosperm genotyping as a strategy to differentiate the allele source in maize heterozygous progeny

Francielle Alline Martins[1], Pedro Crescêncio Souza Carneiro[1], Cosme Damião Cruz[1], José Eustáquio de Souza Carneiro[2] and Claudia Teixeira Guimarães[3]

[1]Universidade Federal de Viçosa (UFV), Departamento de Biologia Geral, Avenida P. H. Rolfs, s/nº, CEP 36570-000 Viçosa, MG. E-mail: franufv@yahoo.com.br, carneiro@ufv.br, cdcruz@ufv.br [2]UFV, Departamento de Fitotecnia. E-mail: jesc@ufv.br [3]Embrapa Milho e Sorgo, Caixa Postal 151, CEP 35701-970 Sete Lagoas, MG. E-mail: claudia@cnpms.embrapa.br

Abstract – The objective of this work was to distinguish the parental source of alleles in heterozygous progeny using semiquantitative polymerase chain reaction (PCR) in maize endosperm. Endosperms derived from direct and reciprocal single-cross hybrids between maize inbred lines L3 and L1113-01 were genotyped by semiquantitative PCR methodology (SQ-PCR) using fluorescent microsatellite primers. The amplification products were evaluated by the ratios of fluorescence intensity (RFI), calculated between the peaks corresponding to the alleles derived from each parental line. Based on the statistically significant contrast between RFI mean values of direct and reciprocal single-cross hybrids, it was possible to distinguish the number of alleles received from each parental line and, ultimately, to determine the origin of the alleles of each cross. Thus, endosperm genotyping using SQ-PCR is a promising strategy to map QTL in maize outbred populations.

Index terms: *Zea mays*, allelic origin, heterozygote, natural populations, polygenes, semiquantitative PCR.

Genotipagem do endosperma como estratégia para diferenciar a origem de alelos em progênies heterozigóticas de milho

Resumo – O objetivo deste trabalho foi distinguir a origem de alelos em progênies heterozigóticas usando reação em cadeia da polimerase (PCR) semiquantitativa em endosperma de milho. Endospermas derivados de híbridos simples diretos e recíprocos entre as linhagens de milho L3 e L1113-01 foram genotipados pela metodologia de PCR semiquantitativa (PCR-SQ) com uso de primers microssatélites fluorescentes. Os produtos de amplificação foram avaliados por meio da razão de intensidade de fluorescência (RIF), calculada entre os valores de intensidade dos picos correspondentes aos alelos derivados de cada genitor. Com base no contraste estatisticamente significativo dos valores médios das RIF entre os híbridos simples direto e recíproco, foi possível distinguir o número de alelos recebidos de cada genitor e, finalmente, determinar a origem dos alelos de cada híbrido. Assim, a genotipagem de endosperma utilizando PCR-SQ é uma estratégia promissora no mapeamento de QTLs em populações exogâmicas de milho.

Termos para indexação: *Zea mays*, origem alélica, heterozigoto, populações naturais, poligenes, PCR semiquantitativa.

Introduction

Mapping quantitative trait loci (QTL) is the first step towards identifying specific genes and ultimately understanding gene action (Gilmour, 2007). The simplest procedure to detect QTL is the single-marker analysis, which is based on statistical comparison between phenotypic mean values of the trait on each genotypic class of the marker. A significant result describes the existence of linkage between the marker and the QTL related to the trait (Coelho, 2000).

The segregating population has a critical importance for the success in map construction, thus the importance of parental selection and the crossing-type determination (Carneiro & Vieira, 2002). Lynch & Walsh (1998) distinguished the mapping populations in terms of inbred line crosses and outbred populations. The latter category can be divided into crosses deliberately performed to maximize the efficiency on QTL detection (e.g. half-sib and full-sib populations), and the crosses done in natural populations.

The main difference between inbred and outbred crossings is that the parentals of the latter generally have unknown genotypes (Slate, 2005). This fact has serious consequences, since only a fraction of the offspring is informative in an outbred population (Lynch & Walsh, 1998). An informative genotype must contain at least one informative gamete, which allows the identification of the origin of its alleles (Da & Lewin, 1995). The frequency of informative genotypes is a feature of each population type directly linked to the allelic frequency and crossing type, and relies on previous knowledge of the parental genotype (Guo & Elston, 1999).

Endosperm is a reserve tissue that guarantees the nutrition of the developing embryo which has several unique genetic properties. This tissue originates from the fusion between a polar male gamete and two female gametophyte nuclei. This fusion generates the triploid endosperm nucleus, which develops by mitotic divisions, carrying two copies of the maternal alleles and one copy of the paternal allele in its genome (Wu et al., 2002; Wang et al., 2008). Therefore, for a locus with two alleles (M_1 and M_2), four genotypes ($M_1M_1M_1$, $M_1M_1M_2$, $M_1M_2M_2$ and $M_2M_2M_2$) are possible, whereas in a diploid plant only three genotypes are possible (Xu et al., 2003). So, the endosperm genotype of direct or reciprocal crosses between two different genotypes will be different (Wu et al., 2002), and the endosperm, as well as the whole plant, represents the next generation (Wen & Wu, 2007).

These attributes make the endosperm a very interesting tissue source, especially for QTL mapping and detection in outbred populations in which heterozygotic genotypes are excluded from the analyses due to the lack of information on the origin of their alleles. A feasible way to include the heterozygotes in this analysis would be genotyping the endosperm, since it is expected that a M_1M_2 heterozygote presents $M_1M_1M_2$ or $M_1M_2M_2$ endosperm genotypes when the female parent donates the M_1 or M_2 allele respectively. Thus, independently of knowing the parental genotypes, the origin of both alleles could be determined in the heterozygotic offsprings, and their phenotypic information could be used to compose the mean values in the contrast to test the association between the *M* locus and the *Q* QTL. Based on endosperm genotyping, two groups could be identified for the *M* locus in the offspring: one group composed by individuals carrying the M_1 allele ($M_1M_1M_1$ and $M_1M_1M_2$), and a second group formed by the $M_2M_2M_2$ and $M_2M_2M_1$ genotypes that received the M_2 allele from the female parent. The association hypothesis between *Q* and *M* can then be evaluated by the contrast between the phenotypic mean values of groups 1 and 2 using t-statistic. By including the heterozygous progenies in the mean values calculation, higher accuracy is expected for QTL detection.

Semiquantitative polymerase chain reaction (SQ-PCR) has results for gene expression analysis highly correlated with Northern blot (Dutta et al., 2007) and in situ hybridization (Tao et al., 2007). Recently, a combination of simple sequence repeat (SSR) markers and semiquantitative PCR in agarose gel has been proposed to evaluate allele dosage in maize endosperm (Martins et al., 2009). The advent of automated DNA sequencers has significantly improved the precision and the resolution in polymorphism detection, mainly considering differences lower than ten base pairs, which is the case of the polymorphisms revealed by SSRs.

The objective of the present work was to implement semiquantitative PCR using fluorescently labeled SSR primers, in order to identify the allelic origin in heterozygous maize progenies by means of endosperm genotyping.

Materials and Methods

Maize seeds of the L3 and L1113-01 inbred lines, as well as their direct and reciprocal crosses, were supplied by the Maize Breeding Program of Embrapa Milho e Sorgo. The hybrids were named after the first line's female parent.

DNA extraction proceeded according to the protocol suggested by Leiva (2006). After removing the pericarp with aid of a razor blade, approximately 100 mg of endosperm (one third of the seed weight) were ground to yield a fine powder. Then, 600 µL of extraction buffer (1% sarcosyl; 100 mmol L^{-1} Tris-HCl pH 8; 100 mmol L^{-1} NaCl; 200 mmol L^{-1} EDTA pH 8.0) were added, vortexed vigorously for 20 s and mixed by continuous inversions at room temperature for 15 min. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added to the mixture, mixed gently and centrifuged at room temperature for 5 min at 15,000 *g*. The aqueous phase was removed to a new tube, an equal volume of chloroform:isoamyl alcohol (24:1) was

added, and the previous procedure was repeated. The DNA was precipitated by adding 450 µL of isopropanol to the supernatant at -80°C for 30 min and centrifuged at 15,000 *g* for 10 min. The pellet was washed with 70% ethanol for 5 min, dried at room temperature for 24 hours and ressuspended in distilled water. DNA was quantified in 0.8% agarose gel with a DNA standard of known concentration. After electrophoresis at 60 V for 30 min, the gel was treated with ethidium bromide (1 mg mL^{-1}) and visualized under ultraviolet light with the Eagle Eye II system (Stratagene Cloning Systems, Inc., La Jolla, CA, USA).

Two microsatellite primers *bnlg1208* and *umc2025* were fluorescently labeled with 6-FAM and previously selected due to the polymorphism between the inbred lines L3 and L1113-01. The primer *umc2025* amplifies a tetra-nucleotide repeat SSR and is located at bin 1.05, while *bnlg1208* is a di-nucleotide repeat locus mapped at bin 5.03. Additional information about the loci is available in the Maize Genetics and Genomics Database (2006).

Amplification reactions were carried out individually for each primer following the parameters described by Martins et al. (2009). The PCR cycles consisted of an initial denaturation at 95°C for 2 min, nine cycles at 94°C for 20 s, 68°C for 20 s with a decrease of 1°C per cycle, and 72°C for 20 s, followed by 21 amplification cycles at 94°C for 20 s, 60°C for 20 s and 72°C for 20 s. The number of amplification cycles was defined by Martins et al. (2009) as being the best representation of the exponential phase. The PCR products were evaluated in the ABI Prism 377 (Applied Biosystems Incorporation, Foster City, California, USA) genetic analyzer following the manufacturer protocols with a GS500-ROX size standard in each line. The electropherogram was generated by the GeneScan 2.1 software (Applied Biosystems do Brasil, São Paulo, Brazil); each peak represents the relative fluorescence of the detected fragment (y-axis) over the number of scans per hour converted in base pairs according to the size standard (x-axis).

The ratios of fluorescence intensity (RFI) were calculated based on the peak area among the amplified fragments that correspond to the alleles derived from the parental lines L3 and L1113-01. For each L3 x L1113-01 and L1113-01 x L3 hybrid, DNA was extracted from three different samples, which constituted the biological replicates, and three amplification reactions were carried out for each sample. RFI values were subjected to analysis of variance, at 5% probability. When homogeneity of variance was observed, all observations of each primer for each hybrid were jointly evaluated in a completely randomized design, with two fixed treatment effects (L3 x L1113-01 and L1113-01 x L3 hybrids) and nine replicates each. Amplification products of the parental DNA were used as positive or negative controls throughout the analyses.

Results and Discussion

The protocol used for DNA isolation from endosperm yielded DNA with adequate quality to perform the PCR reactions, but for some extractions the amount was not enough to be visualized in agarose gel stained with ethidium bromide. The average DNA yield was 250 ng per seed, which is sufficient for approximately ten amplification reactions.

The low yield of endosperm DNA extraction in comparison to that commonly obtained from leaf tissue can be explained by specific tissue variations, such as cell size and number, rate of cell division and concentration of extracellular DNA, or even by differences in the chemical and structural composition of these tissues (Ramos et al., 2006). This lower efficiency can also be attributed to DNA loss by the endospermatic tissue during the reserve accumulation phase in the seeds. When starch is formed, organelles rupture and the cellular content is modified, with expressive reduction in nucleic acid amount, culminating in cell death and total filling of the starch reserves (McDonald et al., 1994).

Seeds are plant reproduction structures, with reserve tissues rich in carbohydrates, lipids and proteins that are responsible for embryo maintenance during the germination process (Carvalho & Nakagawa, 2000). These molecules hamper DNA isolation, and require more washes with organic solvent (isoamylic chloroform-alcohol) in the extraction protocol. This also reduces the total amount of isolated DNA, since, according to Ramos et al. (2006), part of these molecules cannot be separated from other organic compounds and are thus discarded. However, in a protocol developed by Gao et al. (2008) for maize seed DNA extraction, 30 mg of endosperm fragment yielded enough high-quality DNA for around 200 PCR

reactions. Therefore, DNA extraction from seeds may not be a critical issue for the wide utilization of this genotyping strategy.

The amplified fragments of the *bnlg1208* and *umc2025* primers were plotted on electropherograms (Figure 1), where the inbred lines L1113-01 or L3 presented one peak, and both L1113-01 x L3 and L3 x L1113-01 hybrids showed two peaks with different relative fluorescence intensities, indicating the differences in the original concentration of both alleles in the endosperm, as expected. The *bnlg1208* primer amplified the alleles of 104 e 97 bp (Figure 1 A), while the alleles amplified by the *umc2025* primer

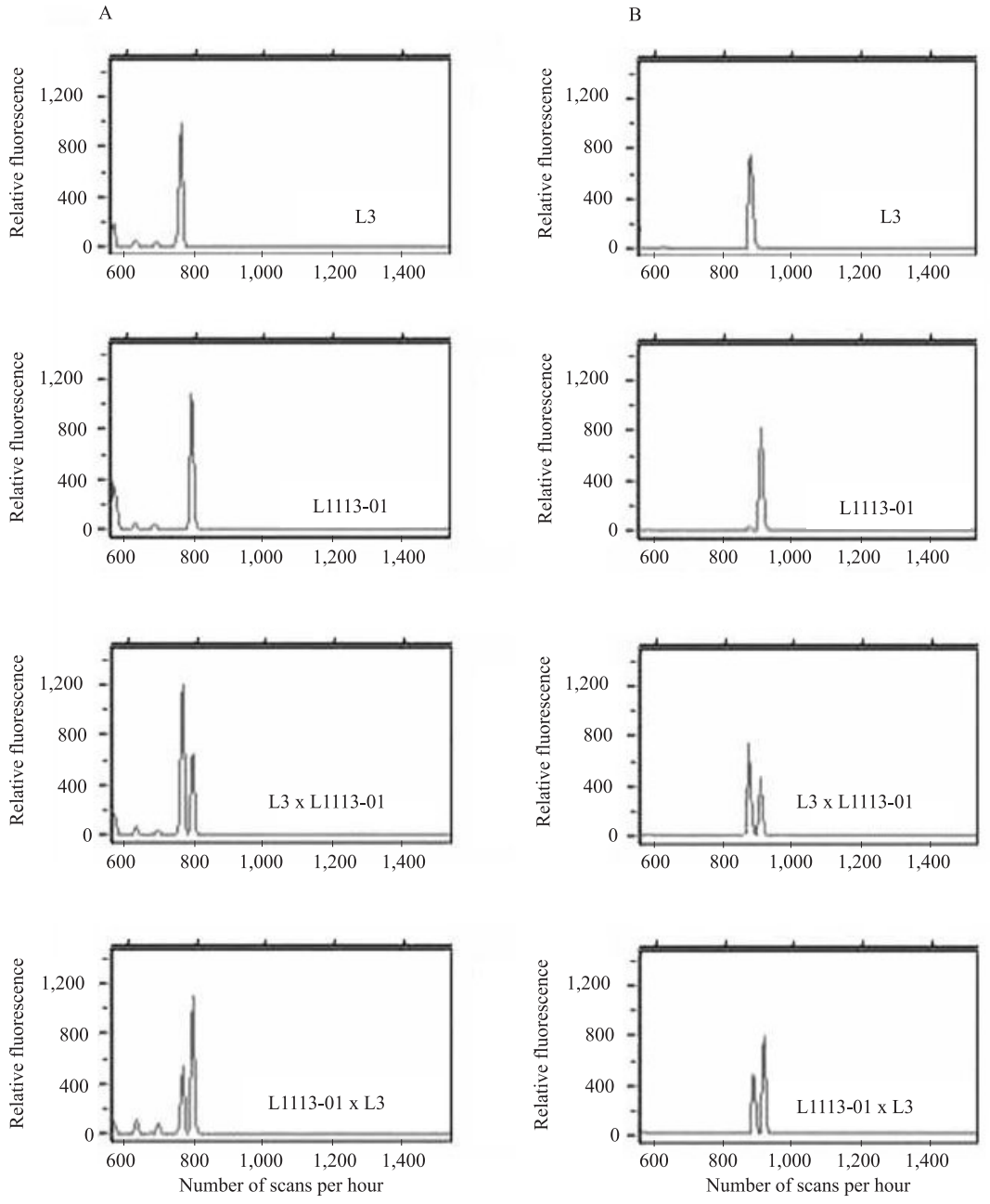

Figure 1. Electropherograms showing alleles amplified using the SSR *bnlg1208* (A) and *umc2025* (B) primers labeled with 6-FAM. The horizontal axis corresponds to the number of scans per hour, where the position of the detected fragments is converted in base pairs according to the GS500-ROX size standard. Inbred lines L3 and L1113-01 show one allele, and both L1113-01 x L3 and L3 x L1113-01 hybrids showed two alleles.

were of 130 and 123 bp for the L1113-01 and L3 lines (Figure 1 B), respectively.

The amplification conditions defined by Martins et al. (2009) and applied in the present work were also suitable for allelic dosage discrimination using fluorescently labeled SSR primers, and allowed the discrimination of the two reciprocal hybrids. The significant contrast ($p<0,05$) between the mean values of RFI for both direct and reciprocal hybrids using the fragments amplified by the *bnlg1208* and *umc2025* primers enabled assuming the different contributions of the two parental alleles in the endosperm of the direct and reciprocal hybrids (Table 1). The differentiation of the initial number of DNA molecules by PCR has been demonstrated since the 1990's, assessing the amplified products during the exponential phase (Ferre, 1992).

SSRs are usually good choices of markers for mapping studies due to their multiallelic nature, which maximizes the proportion of informative parents in outbred crosses (Slate, 2005). However, the methodology proposed in the present work can be applied irrespective of the chosen marker. While using dominant markers such as Amplified Fragment Length Polymorphism (AFLP), the ratios of fluorescence intensity must be calculated between the intensity value of the band amplified in the sample and the intensity value of the band of a standard sample containing defined amounts of the allele. Thus, the semiquantitative PCR using fluorescent SSR markers in maize endosperm genotyping allows inferences regarding the origin of each parental allele in a heterozygous progeny. This establishment will permit the inclusion of the heterozygous individuals in the QTL analyses, significantly increasing the accuracy in outbred-population mapping studies.

Table 1. Summary of analysis of variance of contrasts between RFI mean values of the L3 x L1113-01 and L3 x L1113-01 hybrids for the amplification products using the SSR *bnlg1208* and *umc2025* primers.

Source of variation	df	Mean square	
		bnlg1208	*umc2025*
Contrast	1	6.9364*	6.8508*
Residual	16	0.0151	0.0326
L3 x L1113-01		1.795	1.765
L1113-01 x L3		0.554	0.531
CV (%)		15.73	

*Significant at 5 % probability.

Conclusions

1. Endosperm genotyping by semiquantitative PCR methodology using fluorescent microsatellite markers permits the identification of different allelic origins in maize heterozygotes.

2. The proposed methodology is applicable to any population type and can be tested for other marker systems.

Acknowledgements

To Conselho Nacional de Desenvolvimento Científico e Tecnológico and to Fundação de Amparo à Pesquisa do Estado de Minas Gerais, for financial support; to Dr. Sidney Netto Parentoni, for providing the seeds of the lines and of the direct and reciprocal hybrids; to Embrapa Milho e Sorgo, for technical and financial support.

References

CARNEIRO, M.S.; VIEIRA, M.L.C. Mapas genéticos em plantas. **Bragantia**, v.61, p.89-100, 2002.

CARVALHO, N.M. de; NAKAGAWA, J. **Sementes**: ciência, tecnologia e produção. 4.ed. Jaboticabal: FUNEP, 2000. 588p.

COELHO, A.S.G. Considerações gerais sobre a análise de QTL's. In: PINHEIRO, J.B.; CARNEIRO, I.F. (Ed.). **Análise de QTL no melhoramento de plantas**. Goiânia: Fundação de Apoio à Pesquisa, 2000. p.1-36.

DA, Y.; LEWIN, H.A. Linkage information content and efficiency of full-sib and half-sib designs for gene mapping. **Theoretical and Applied Genetics**, v.90, p.699-706, 1995.

DUTTA, A.; SINGH, D.; KUMAR, S.; SEN, J. Transcript profiling of terpenoid indole alkaloid pathway genes and regulators reveals strong expression of repressors in *Catharanthus roseus* cell cultures. **Plant Cell Reports**, v.26, p.907-915, 2007.

FERRE, F. Quantitative or semi-quantitative PCR: reality versus myth. **PCR Methods and Applications**, v.2, p.1-9, 1992.

GAO, S.; MARTINEZ, C.; SKINNER, D.J.; KRIVANEK, A.F.; CROUCH, J.H.; XU, Y.B. Development of a seed DNA-based genotyping system for marker-assisted selection in maize. **Molecular Breeding**, v.22, p.477-494, 2008.

GILMOUR, A.R. Mixed model regression mapping for QTL detection in experimental crosses. **Computational Statistics & Data Analysis**, v.51, p.3749-3764, 2007.

GUO, X.Q.; ELSTON, R.C. Linkage information content of polymorphic genetic markers. **Human Heredity**, v.49, p.112-118, 1999.

LEIVA, J. **DNA isolation from mature endosperm**. Available at: <http://ag.arizona.edu/research/larkinslab/>. Accessed on: 30 Sept. 2006.

LYNCH, M.; WALSH, B. **Genetics and analysis of quantitative traits**. Sunderland: Sinauer Associates, 1998. 980p.

MAIZE GENETICS AND GENOMICS DATABASE. **SSR data in MaizeGDB**. Available at: <http://www.maizegdb.org/ssr.php>. Accessed on: 20 Sept. 2006.

MARTINS, F.A.; CARNEIRO, P.C.S.; GUIMARÃES, C.T.; MAGALHÃES, J.V.; CARNEIRO, J.E.S.; CRUZ, C.D. Distinction between plant samples according to allele dosage by semiquantitative polymerase chain reaction. **Genetics and Molecular Research**, v.8, p.319-327, 2009.

MCDONALD, M.B.; ELLIOT, L.J.; SWEENEY, P.M. DNA extraction from dry seeds for RAPD analyses in varietal identification studies. **Seed Science and Technology**, v.22, p.171-176, 1994.

RAMOS, N.P.; BRUNELLI, K.R.; CAMARGO, L.E.A.; MARCOS FILHO, J. Sensibilidade dos microssatélites para determinar a pureza varietal em sementes de milho. **Revista Brasileira de Sementes**, v.28, p.99-105, 2006.

SLATE, J. Quantitative trait locus mapping in natural populations: progress, caveats and future directions. **Molecular Ecology**, v.14, p.363-379, 2005.

TAO, N.G.; HU, Z.Y.; LIU, Q.; XU, J.; CHENG, Y.J.; GUO, L.; GUO, W.; DENG, X.X. Expression of phytoene synthase gene (*Psy*) is enhanced during fruit ripening of Cara Cara navel orange (*Citrus sinensis* Osbeck). **Plant Cell Reports**, v.26, p.837-843, 2007.

WANG, X.F.; HU, Z.Q.; WANG, W.; LI, Y.; ZHANG, Y-M.; XU, C. A mixture model approach to the mapping of QTL controlling endosperm traits with bulked samples. **Genetica**, v.132, p.59-70, 2008.

WEN, Y.G.; WU, W. Interval mapping of quantitative trait loci underlying triploid endosperm traits using F3 seeds. **Journal of Genetics and Genomics**, v.34, p.429-436, 2007.

WU, R.; MA, C.X.; GALLO-MEAGHER, M.; LITTELL, R.C.; CASELLA, G. Statistical methods for dissecting triploid endosperm traits using molecular markers: an autogamous model. **Genetics**, v.162, p.875-892, 2002.

XU, C.; HE, X.; XU, S. Mapping quantitative trait loci underlying triploid endosperm traits. **Heredity**, v.90, p.228-235, 2003.

Received on March 25, 2009 and accepted on September 13, 2009