

## **Analysis of the relationship between coefficients of relatedness and molecular similarity of parental forms with respect to the heterosis effect in maize**

**Agnieszka Tomkowiak<sup>1</sup>, Zbigniew Broda<sup>1</sup>, Krzysztof Moliński<sup>2</sup>,  
Marta Molińska-Glura<sup>3</sup>, Józef Adamczyk<sup>4</sup>**

<sup>1</sup>Department of Genetics and Plant Breeding, Poznań University of Life Sciences, Dojazd 11, 60-632 Poznań, Poland, agatom@interia.pl

<sup>2</sup>Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poznań, Poland

<sup>3</sup>Department and Laboratory of Computer Science and Statistics, Karol Marcinkowski University of Medicine in Poznań, Dąbrowskiego 79, 60-529 Poznań, Poland,

<sup>4</sup>Hodowla Roślin Smolice Sp. z o.o. Grupa IHAR, Smolice 146, 63-740 Kobylin, Poland

### SUMMARY

Within the last twenty years studies have been conducted at many research centers with the aim of dividing breeding materials into heterotic groups based on molecular markers. Molecular techniques make it possible to study the genetic purity of inbred lines, determine their genetic variability and classify breeding materials for which no information is available on their origin. This study aims to investigate relationships between coefficients of relatedness (pedigree analysis) and molecular similarity estimated on the basis of AFLP and RAPD molecular markers, between parental forms of F<sub>1</sub> maize hybrids. Determination of these relationships will make it possible to establish a hierarchy of importance for applied methods concerning selection of parental components for heterotic crossings based on the degree of relatedness and genetic similarity. As a result of the experiment it was shown that in the case of incomplete pedigree information, in the selection of parental components for crossings we may use information concerning molecular similarity determined using AFLP markers and the Jaccard index, which to the least degree differentiates matrices of AFLP and RAPD molecular similarity.

**Key words:** relationship, molecular similarity, maize, AFLP, RAPD

### 1. Introduction

Breeding success in all plant species is determined by access to initial material with possible high genetic divergence. Initial materials, well-identified and

divided into heterotic groups, ensure effectiveness and the reduction of costs for the entire process of hybrid breeding. Next to geographical origin and diallelic crossing, the most frequently applied methods in breeding practice include the analysis of genetic origin (pedigree) and the analysis of genetic similarity determined using methods of molecular genetics. Genetic composition of materials based on their geographical origin is not very reliable due to the considerable international exchange of breeding materials, whereas diallelic crossing, while supplying the greatest amount of information on the material (Dhillon, Singh 1997, Królikowski 1977, Adamczyk 2001), is very costly. In the last twenty years studies have been conducted at many research centers with the aim of dividing breeding materials into heterotic groups based on the analysis of molecular markers. Methods of molecular genetics, in combination with analysis of genetic origin (pedigree), facilitate the selection of parental components for crossings, reducing the number of tested lines in costly field trials to the necessary minimum and thus reducing the costs of the entire breeding process.

The aim of this study was to investigate relationships between relatedness coefficients determined by Henderson methods (1994) and molecular similarity estimated using AFLP and RAPD molecular markers between parental components of  $F_1$  hybrids in maize. The determination of relationships will make it possible to determine a hierarchy of importance for the applied methods concerning the degree of relationship and genetic similarity in the breeding process.

## **2. Plant material and methods**

The plant material used in the analyses comprised parental lines of maize along with  $F_1$  hybrids (Table 1). Parental lines and hybrids were obtained from the company Hodowla Roślin Smolice Spółka z o.o.

**Table 1.** Parental lines of maize

No.	Maternal lines	Paternal lines
1	S160	S336A
2	S41336	S41324A-2
3	S78510	S80660A
4	S54555	S79757
5	S245	S41789
6	S311	Co255
7	S64417	S61328
8	S41796	S41324A-2
9	S41789	S41324A-2
10	S56125A	S41324A-2
11	S63322-3	S61328
12	S64423-2	S61328
13	S68911	S61328

### **RAPD (Random Amplified Polymorphic DNA) molecular markers**

Genomic DNA from the hybrid forms and parental components of maize was isolated using a modified method according to Thompson and Henry (1995).

Leaf discs of 2 mm<sup>2</sup> were treated with 200 µl TPS buffer containing 100 mM Tris HCl (pH 9.5), 1 M KCl and 10 mM EDTA. Incubation took place at 95°C for 15 min.

A polymerase chain reaction (PCR) was run in 12.5 µl reaction mixture containing deionized water, 1M Tris HCl (pH 8.3), 25 mM MgCl<sub>2</sub>, BSA, 2 mM dNTP, primer 5 pmol/µl, Taq polymerase 5 U/µl and DNA extract 25 ng/µl. The Taq polymerase was from MBI-Fermentas, while the other reagents were from SIGMA.

DNA amplification was run using a T3 BIOMETRA thermocycler from POLYGEN. Electrophoresis of the amplification products was run in 1.5% agarose gel containing 1.5 g agarose, 100 ml TBE1x buffer (10.8 g Tris base, 5.5g Boric acid, 4 ml 0.5M EDTA pH 8.0) and 1 µl ethidium bromide.

### **AFLP (Amplified Fragment Length Polymorphism) molecular markers**

Isolation of genomic DNA was performed by the column method with the use of a kit from Qiagen.

DNA was digested using restriction enzymes EcoRI and MseI. Digested DNA was supplemented with 24  $\mu\text{l}$  adapter ligation solution and 1  $\mu\text{l}$  T4DNA ligase, and samples were incubated at 20°C for 3h. After ligation the mixture was supplemented with 90  $\mu\text{l}$  TE buffer at pH 8.0. Pre-amplification was carried out with the use of a T3 BIOMETRA thermocycler from Polygen. After ligation 40  $\mu\text{l}$  pre-amp primer mix, 3.5  $\mu\text{l}$  10x PCR buffer, 15  $\mu\text{l}$   $\text{MgCl}_2$  and 1  $\mu\text{l}$  Taq polymerase (5U/ $\mu\text{l}$ ) were added to the diluted DNA matrix. After pre-amplification MIX1 (primers and dNTP) and MIX2 (Taq polymerase, 10x PCR buffer and  $\text{MgCl}_2$ ) were added to the diluted matrix. Upon completion of PCR the samples were frozen at -20°C. Electrophoresis was run in 5% acrylamide gel for 2.5 h at 60 W, 400 mA and 1400 V. PCR products were detected by silver staining.

### 3. Statistical evaluation and documentation of results

The results of both RAPD and AFLP reactions are documented by electrophoretic images. The molecular masses of the analyzed parental forms were calculated using the UVIMAP computer program. In the next stage the molecular masses of objects were used to determine genetic similarities with the use of the original Mpod300 program. Similarities were determined for three most frequently applied measures:

$$\text{Nei and Li } GS_{NL} = 2n_{xy} / (n_x + n_y),$$

$$\text{Kulczyński } GS_K = n_{xy} (n_x + n_y) / 2n_{xy},$$

$$\text{Jaccard } GS_J = n_{xy} / (n_x + n_y - n_{xy}),$$

where  $n_{xy}$  denotes the number of pairs of common bands in both genotypes, while  $n_x$  and  $n_y$  denote the number of all bands for each genotype respectively. The GS value denotes the index of similarity between the two analyzed genotypes.

The Henderson method was used to construct a matrix of relationships using complete pedigree information for the parental forms. Following the aim of the study – to determine relationships between the matrix of relatedness and matrices of genetic similarity – the reduced matrix of relatedness was used, in which those objects were eliminated which were not subjected to molecular analysis. In order to determine the hierarchy of importance for the applied methods concerning genetic similarity and the degree of relatedness, the Euclidean distance was used, which represents the distance between matrices expressed as the sum of squares of differences between elements of individual matrices.

#### 4. Results

For RAPD markers 10 selected primers generated 93 polymorphic bands; on average one primer generated from 4 to 15 bands differentiating parental lines (Table 2). Primer OPB 07 turned out to be the most effective primer, which differentiated the tested genotypes the most strongly (15 polymorphic bands). Five selected combinations of primers in the case of AFLP molecular markers generated 197 polymorphic bands. One combination of primers generated on average from 5 to 16 differentiating bands (Table 3). The greatest polymorphism was obtained using the combination of E-AGG and M-CAC primers (47 polymorphic bands).

Values of Euclidean distances between matrices of molecular similarity between AFLP and RAPD were determined for individual indexes (Table 4). The following values were obtained: for Nei and Li 4.3, for Kulczyński 4.24, and for Jaccard 1.52. On this basis it may be concluded that the Jaccard index least effectively differentiates matrices of molecular similarity between AFLP and RAPD.

In the case of the relationship between the pedigree matrix and the molecular matrices, the measures of the difference between AFLP and RAPD

matrices for individual indexes are zero. This means that matrices of molecular similarity carry similar information on the structure of these similarities.

Table 2. Characteristics of DNA amplification products obtained as a result of RAPD-PCR

Primer number 5' – 3'	Number of generated DNA fragments	Number of generated polymorphic fragments
OPA 04	14	13
OPA 07	8	5
OPA 09	7	4
OPA 12	11	6
OPB 07	16	15
OPB 10	10	9
OPB 11	11	9
OPB 19	14	11
OPC 06	12	12
OPC 08	9	9

Table 3. Characteristics of DNA amplification products obtained using AFLP markers

Primer number 5' – 3'	Number of generated DNA fragments	Number of generated polymorphic fragments
E-AAC	36	35
M-CAC		
E-ACC	48	41
M-CAG		
E-ACG	51	45
M-CAC		
E-ACG	32	29
M-CAG		
E-AGG	53	47
M-CAC		

Based on the value of the square of Euclidean distance (Table 5) it may be concluded that genetic similarity, determined on the basis of AFLP molecular markers, is closer to the degree of relatedness between components of parental lines of F<sub>1</sub> hybrids in maize than the molecular similarity of RAPD for each of the analyzed measures.

The index of the difference of the pedigree matrix and individual matrices of molecular similarity for all the investigated measures is zero. This means that molecular similarity significantly reflects the degree of relatedness.

**Table 4. Genetic molecular analysis relatedness and similarity determined based on AFLP and RAPD markers between the parental forms of maize**

No.	Maternal lines	Paternal lines	Relatedness	Genetic similarity using AFLP marker (%)			Genetic similarity using RAPD marker (%)		
				Nei and Li	Kulczyński	Jaccard	Nei and Li	Kulczyński	Jaccard
1	S160	S336A	0	16	16	9	19	19	10
2	S41336	S41324A-2	50	18	18	10	10	10	5
3	S78510	S80660A	6	12	12	6	18	18	10
4	S54555	S79757	0	16	16	9	23	23	13
5	S245	S41789	13	12	12	8	20	21	11
6	S311	Co255	50	16	16	9	14	15	8
7	S64417	S61328	0	23	23	13	19	19	10
8	S41796	S41324A-2	50	13	13	7	17	18	9
9	S41789	S41324A-2	50	15	14	7	13	13	6
10	S56125A	S41324A-2	0	21	21	11	5	6	3
11	S63322-3	S61328	0	13	13	7	12	12	6
12	S64423-2	S61328	0	16	16	8	14	14	7
13	S68911	S61328	0	15	16	8	20	16	8

**Table 5. Euclidean distance of the matrix of molecular similarity and the pedigree matrix**

Index	Type of marker	
	AFLP	RAPD
Nei and Li	35.77	38.16
Kulczyński	35.81	38.25
Jaccard	35.13	37.67

## 5. Discussion

The final stage in the long breeding cycle is to obtain hybrid lines. According to Sprague and Eberhardt (1997), we may distinguish three basic stages in the breeding of hybrids. The first involves the generation of two or more populations of breeding lines, which will represent opposite gene pools with the maximum additive variation within each of them. The next stage has the aim of continuous improvement of these populations using methods of cyclic selection and generation of inbred lines. The last stage involves obtaining high-yielding hybrids from each selection cycle using efficient and systematic testing procedures. The objective of line testing is to determine combining ability: general (in topcross or polycross) and specific (in diallelic crossing).

The exceptional difficulty here consists in the compromise between the huge number of evaluated objects and requirements of scientific experimentation, which is connected with the possible high number of replications and locations. As early as in the 1930s problems appeared with the evaluation of all possible hybrid combinations from bred lines. It is extremely difficult to investigate all possible hybrid formulas from mass bred inbred lines. Along with the rapidly increasing number of new lines, also in hybrid breeding economic barriers have arisen in experiments conducted for all potential formulas.

Occasionally the long-term, complicated and costly breeding cycle may be considerably shortened by conducting selection with the use of DNA markers. This is based on the close linkage between a marker and the locus responsible for inheritance of a production trait. This method is frequently referred to as Marker Assisted Selection (MAS).

According to Adamczyk (1998), the interest of maize breeders in DNA analysis methods is connected with the search for methods other than direct crossing, which would facilitate appropriate selection of parental forms in order to obtain a maximum heterosis effect in  $F_1$  hybrids. The aim is to find a relationship between yields in the  $F_1$  hybrid and the heterogeneity of loci –



markers for its parental forms. The second area of interest in these techniques is related to their use to evaluate genetic purity of breeding materials and their identification for the purpose of copyright protection. In this case the applicability of these techniques has been confirmed and the most commonly used methods are those involving evaluation of length polymorphism of the DNA chain, such as RFLP and AFLP, i.e. restriction and amplification. The issue is not so obvious when these techniques are applied directly in breeding in order to predict hybrid formulas. Results reported by different research teams differ considerably depending on the method used and the selection of experimental material. Due to the relatively low costs of analysis, much attention is focused on techniques based on polymerase chain reaction (PCR), referred to as Random Amplified Polymorphism of DNA (RAPD). As the name indicates, a common characteristic of these techniques is the application of primers with random sequences.

Joanna Sztuba-Solińska (2006) stated that molecular techniques are most effective in the evaluation of genetic distance between parental lines and facilitate the prediction of efficiency for the generation of new hybrid cultivars exhibiting important production traits.

The problem of inbred breeding and crossing of unrelated individuals and individuals characterized by molecular similarity, particularly biological aspects of gene expression in hybrids, are the most current and important in terms of pure science and practical applications in plant breeding. It is assumed that heterosis may be an effect of the metabolism of many enzymes. Quantitative analyses of mRNA content has shown quantitative variation in RNA content and the number of sites connected with genomic replications. Probably the molecular background of heterosis is actually found in the specific interaction of the nuclear genome with the organelle genome. Results of studies on the molecular background of heterosis indicate that heterosis is expressed mainly in an increase of the level of DNA replication and thus the expression of genes being in a unique heterozygous combination, i.e. recombination of relatedness and genetic similarity of parental forms in  $F_1$  hybrids (Barciszewski 1996.).

The hypotheses of dominance and overdominance are based on classical genetics and do not explain the mechanism of heterosis. Polymorphism of DNA markers in combination with epigenetic factors provides a molecular explanation of this phenomenon.

Studies on the dependence between isoenzymes and the heterosis effect, manifested in grain yield in F1 hybrids in maize (*Zea mays* L.) and rice (*Oryza sativa* L.), did not find significant correlations between isoenzymatic markers and the heterosis effect (Peng 1988, Peng 1996).

In 1992, Stubber et al., using molecular markers, identified factors connected with heterosis in a hybrid generated from inbred maize lines.

Becker and Link (2000), who investigated the relationship of genetic distance in inbred maize lines with the heterosis effect, stated that DentxDent maize hybrids coming from parental forms with a considerable genetic distance showed the greatest heterosis effect manifested in grain yield. Similar dependencies were observed by Tomkowiak et al. (2009) and Tomkowiak et al. (2010). Using RAPD and AFLP markers for the analysis of genetic variability in inbred maize lines, they showed that a greater heterosis effect may be expected in F1 hybrids whose parental forms came from the gene pool with a considerable genetic distance.

Theoretical foundations for the relationship between genetic distance and heterosis were presented by Bernardo (1992). He stated that molecular markers may be useful in the prediction of heterosis, on condition that they exhibit a strong dominance effect, the frequency of alleles is negatively correlated with the parents, their heredity is high and there is a connection between Quantitative Trait Loci (QTL).

In the extensive available literature concerning the dependence of the heterosis effect in F1 hybrids on the genetic distance between parental components and their degree of relatedness, some authors are of the opinion that molecular markers are the most effective tools in the determination of genetic distance between parental lines. They claim that the analysis of genetic distance between parental components facilitates the prediction of efficiency of

generation of new hybrid cultivars, exhibiting production traits important from the point of view of breeding. By contrast, opponents of these theories believe that detection of genetic similarity between parental components is insufficient for the generation of hybrid cultivars, and that in the selection of parental components in crossings we need to focus rather on their degree of relatedness.

DNA marker technology has been used as a tool for plant genetic investigation and estimating specific combining ability for predicting hybrid grain yield (Melhinger et al 1990). Also Dudley et al (1992) carried out investigations in which the major goals were to evaluate methods of using molecular data to identify parents useful for improving a single cross and to identify parents useful for improving a hybrid with genotypic marker of inbred lines (Phillips R.L. et al. 2001).

The aim of this paper is to determine relationships occurring between indexes of relatedness determined based on the methods proposed by Henderson and molecular similarity estimated on the basis of AFLP and RAPD molecular markers between parental forms of  $F_1$  hybrids in maize. Determination of these relationships made it possible to identify the hierarchy of importance for the applied methods concerning both the degree of relationship and genetic similarity in the breeding process. It was shown in this study that molecular similarity, irrespective of the type of marker, significantly reflects the degree of relationship; however, genetic similarity determined on the basis of AFLP molecular markers is closer to the degree of relatedness between parental components of  $F_1$  maize hybrids than molecular similarity expressed on the basis of RAPD markers, irrespective of the applied index. It was also found that the Jaccard index to the least degree differentiates the matrix of molecular similarities between AFLP and RAPD. This means that in this case matrices of molecular similarity carry the same information on the structure of similarity.

Along with the broadening knowledge on genome structure it has become possible in the polymerase chain reaction to apply primers, which are complementary or partly complementary to specific, conservative sequences found at the intron-exon interface. A practically unlimited number of primers

may be designed, in which conserved sequences are supplemented with oligonucleotides with random sequences. It is in the abundance of possible variants that we may see the cause of divergent results being obtained, and at the same time also difficulties in the identification of primers with an optimal sequence.

Results of many studies, e.g. Rafalski et al. (1998) and Adamczyk (2001), indicate the possibility of application of PCR techniques with partly complementary primers in line clustering, which considerably facilitates breeding. In turn, when using the RFLP and AFLP techniques, only in one case to date has a result been obtained indicating their suitability for prediction of hybrid formulas, with the AFLP technique being more informative than RFLP (Ajmone-Marsan et al. 1998). Obviously these techniques generate more polymorphic DNA images than RAPD, but they are much more costly and more laborious.

## 6. Conclusions

- Molecular similarity, irrespective of marker type, significantly reflects the degree of relatedness, and for this reason the analysis of genetic distance between parental components may turn out to be sufficient to predict the efficiency of generation of new hybrid cultivars exhibiting production traits of importance from the point of view of breeding.
- Genetic similarity determined with the use of AFLP molecular markers is closer to the degree of relatedness between parental components of F1 maize hybrids than molecular similarity determined based on RAPD markers, irrespective of the applied index.
- It results from our experiment that in the case of incomplete pedigree information, in the selection of parental components for crossings we may use the information concerning molecular similarity determined using AFLP markers and the Jaccard index, which to the least degree differentiates matrices of molecular similarity between AFLP and RAPD.

- The results from this investigation corroborated results from many earlier attempts to correlate marker allele diversity of parental lines with hybrid performance.

## REFERENCES

- Adamczyk J. (1998): A review of maize breeding methods and their practical applicability. *Biul. IHAR* 208: 123-130.
- Adamczyk J. (2001): The role of new hybrids in enhancing efficiency of different utilisation systems of maize. Appendix to "Agro Serwis". *Biznes-Press*. Warsaw: 7-12.
- Ajmone Marsan P., Castiglioni P., Fusari F., Kuiper M., Motto M. (1998): Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. *Theor. Appl. Genet.* 96: 61-64.
- Barciszewski J., Gunhil E., Siboska, Bent O., Pedersen Brian F.C., Clark Suresh I.S. Rattan. (1996): Evidence for the presence of kinetin in DNA and cell extracts. *FEBS Letters* 393: 197-200.
- Becker H.C., Link W. (2000): Heterosis and hybrid breeding. *MCC 2000 Mendel*: 1-2.
- Bernardo R. (1992): Relationship between single-cross performance and molecular marker heterozygosity. *Theor. Appl. Genet.* 83: 628-634. Centenary Congress: 319-327.
- Dudley J.W., Saghai Maroot M.N., Rufener S.K. (1992): Molecular marker information and selection of parents in corn breeding programs *Crop. Sci.* 32: 301-304.
- Dhillon B.S., Singh J. (1997): Combining ability and heterosis in diallele crosses in maize. *Theor. Appl. Genet.* 3: 117-122.
- Henderson C. R. (1994): Estimation of changes in herd environment. *J. Dairy Sci.* 32: 706.
- Jaccard P. (1901): Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin de la Société Vaudoise des Sciences Naturelles* 37: 547-579.
- Królikowski Z. (1977): Inheritance of some quantitative characters of maize under diallelic crossing. *Genet. Pol.* 1: 15-26.
- Melehinger A.E., Lee M., Lamkey K.R., Woodman A. (1990): Genetic diversity for restriction fragment length polymorphism. Relation to estimated effects in maize inbreds. *Crop Sci.* 30: 1033-1040
- Nei M., and Li W. H. (1979): Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* 76: 5269-5273.
- Peng S., Yang J., Garcia F. V., Laza R.C., Visperas R.M., Sanico A. L., Chavez A. Q., Virmani S.S. (1998): Physiology-based crop management for yield maximization of hybrid rice. *Advances in Hybrid Rice Technology, Proceedings of the 3rd International Symposium on Hybrid Rice 14-16 November 1996 Hyderabad India*: 12-13.
- Peng J. Y., Glaszmann J. C., Virmani S.S. (1988): Heterosis and isozyme divergence in Indica. *Crop Science* 28(3): 561-563.

- Phillips R.L., Visil J.K. (2001): DNA-Based Markers in Plants. *Advances in Cellular and Molecular Biology of Plants* 6: 125–131.
- Rafalski A., Gidzińska M., Wiśniewska I. (1998): Systemy PCR w badaniach pokrewieństwa genetycznego linii kukurydzy [PCR systems in studies on genetic relationship in maize lines]. *Biul. IHAR* 208: 131–140.
- Sprague G.F., Eberhardt S.A. (1997): Corn Breeding. In: *Corn and Corn Improvement*. Sprague F.G. (ed.), Agron. Monog. 18., ASA, CSSA and SSSA, Madison, Wisconsin: 14.
- Stuber C., Lincoln S.E., Wolff T., Helentjaris T. (1992): Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics* 132: 823–839.
- Sztuba-Solińska J. (2006): KOSMOS – problemy nauk biologicznych *PTP* 54: 227-239.
- Thompson D., Henry R. (1995): Single step protocol for preparation of plant tissue for analysis by PCR. *Biotechniques*, 19: 394–400.
- Tomkowiak A., Broda Z., Moliński K. (2010): Attempt to adapt a mathematical model for the heterosis effect in maize F1 hybrids depending on the genetic distance of parental forms. *Plant Breeding and Seed Science IHAR* 62: 72–89.
- Tomkowiak A., Broda Z., Adamczyk J. (2009): Assessment of genetic diversity of corn lines suitable for breeding of heterosis hybrids, based on molecular markers AFLP and RAPD. *ACTA Scientiarum Polonorum, Agricultura* 8 (1): 69–82.