SNP-based genetic diversity among few-branched-1 *(Fbr1)* maize lines and its relationship with heterosis, combining ability and grain yield of testcross hybrids

Shorai Dari[1*], John MacRobert[2], Adré Minnaar-Ontong[3], Maryke.T.Labuschagne[3]

[1]Department of Crop Science, University of Zimbabwe, Harare, Zimbabwe
[2] International Maize and Wheat Improvement Centre (CIMMYT), Harare, Zimbabwe
[3] Department of Plant Sciences, University of the Free State, Bloemfontein, South Africa
*Corresponding author: E-mail: shoraidari@yahoo.co.uk

Abstract

Single nucleotide polymorphism (SNP) markers are regarded as efficient, compared to other marker types in genetic characterization of maize (*Zea mays* L.) germplasm because of their vast coverage of the maize genome. The objectives of this study were to (a) genetically fingerprint 'few-branched-1' (*Fbr1*) and normal tasselled CIMMYT elite lines using SNP markers, to assess their relatedness and level of homozygosity and (b) to determine SNP-based genetic distance among these maize lines and to find association of genetic distances with specific combining ability (SCA), mid-parent heterosis (MPH), high-parent heterosis (HPH) and mean grain yield of the hybrids. Twenty-six CIMMYT maize lines (12 with the *Fbr1* gene, and 14 normal-tasselled) were genotyped using 1074 SNP marker loci. Fifteen of these lines were used in two separate diallel mating designs: a 9x9 and 6x6 crossing set-up, to make hybrids for yield evaluation. Average residual heterozygosity of SNP loci ranged from 0.2-36.1%, with an average of 8.2%, well above the expected ranges for residual heterozygosity found in maize inbred lines. The polymorphic information content (PIC) for the 1074 SNP loci ranged from 0.015-0.50, with an average of 0.25. Mean genetic distance for all pair wise comparisons of lines was lower (0.30) suggesting a high level of relatedness among lines. A number of elite CIMMYT lines were successfully converted to *Fbr1*, and were homozygous for the 1074 SNP loci, thus could be used in breeding programmes involving these new tassel mutants. The unweighted paired group method using arithmetic averages (UPGMA) cluster analysis revealed two discrete clusters for the inbred lines, reflecting heterotic groups used by CIMMYT. In the principal component (PC) analysis, PC1 and PC2 explained 10.87 and 9.08% respectively, of the molecular variance in tassel size for the 1074 SNPs. The results confirmed molecular markers as a powerful complement for use in genetic characterization, in assigning lines into defined heterotic groups and in examining the relationships among inbred lines at deoxyribonucleic acid (DNA) level. Marker-based genetic distances were positively correlated with hybrid performance, SCA and heterosis indicating that they could accurately predict hybrid performance in this set of germplasm. Grain yield for the hybrids ranged from 0.49-2.48 kg/plot, with an average of 1.80 kg/plot. Hybrids from closely related parental lines (according to SNP-based genetic distances) had the lowest mean grain yield, lowest SCA effects for grain yield, and had the lowest heterosis values. Thus, SNP-based genetic distance information would be useful for effective selection by avoiding genetically similar lines when selecting parents for breeding programmes that require genetically diverse lines as parents.

KeyWords *Maize (Zea mays L.); SNP markers; genetic diversity; Fbr1 tassel mutation*

Introduction

The development and application of various DNA marker technologies has contributed significantly to genetic research in maize in the last decades (Yan et al., 2009) and the use of molecular markers for the detection and exploitation of DNA polymorphism has made a significant contribution to the field of molecular genetics (Semagn et al., 2006). DNA-based or molecular markers are tools that can be used effectively for genetic diversity analysis of many crop species. Unlike morphological markers, these markers are not influenced by environmental factors (Smith and Smith, 1992; Westman and Kresovich, 1997); and they are a reflection of the actual level of genetic difference existing between genotypes.

SNP markers have been found to be abundant and evenly distributed throughout the genomes of most plant species. It is considered to be an ideal marker system for genetic research in many crops (Yan et al., 2009). Several high throughput platforms have been developed that allow rapid and simultaneous genotyping of up to a million SNP markers (Yan et al., 2009), and

more than 30 different SNP detection methods have been developed and applied in different crop species (Gupta et al., 2008). Availability of genome sequences of several organisms has allowed the study of sequence variations between individuals, cultivars, and subspecies (Semagn et al., 2006). These studies showed that SNPs and insertions and deletions (InDels) are abundant and distributed throughout the genome in various plant species (Garg et al., 1999; Drenkard et al., 2000; Nasu et al., 2002; Batley et al., 2003a). The abundance of these polymorphisms in plant genomes makes the SNP marker system an attractive tool for mapping, marker-assisted breeding, map-based cloning and in genetic diversity studies (Gupta et al., 2001; Rafalski, 2002; Batley et al., 2003b; Yan et al., 2009).

The identification of parental inbred lines that form superior hybrids is the most costly and time- consuming phase in maize breeding (Betrán et al., 2003). *Per se* performance of maize inbred lines does not predict the performance of maize hybrids for grain yield (Hallauer and Miranda, 1988), thus, predictors of single-cross hybrid value or heterosis between parental inbred lines could therefore increase the efficiency of hybrid breeding programmes. The level of genetic variation between two inbred lines has an influence on the general performance or heterosis in the resulting hybrid (Hinze and Lamkey, 2003). Hence, molecular markers, which reflect such genetic variation can hasten the selection of parental inbred lines (Qi et al., 2010). Previous methods have included diallel crossing, multivariate analyses (Aydin et al., 2007) and several studies have shown that a multifaceted approach which includes morphological, biochemical and intense molecular trait evaluation of candidate inbred lines can be more reliable in heterotic breeding (Rencher, 1995).

The pre-selection of parents is an essential step in the prediction of hybrid performance (Munhoz et al., 2009). The traditionally applied methodology for this purpose is the formation of heterotic groups, based on the evaluation of the pedigree data and its relation with the heterosis values based on morphological traits of interest (Franco et al., 2001; Mohammadi and Prasanna, 2003; Miranda et al., 2008). Molecular markers have been used to detect the variation in the DNA sequence underlying the analysis of existing genetic dissimilarity of the parents (Munhoz et al., 2009), and markers have the advantage of simplifying the screening of parents, which is done through DNA evaluation (Mohammadi and Prasanna, 2003; Crossa and Franco, 2004; Legesse et al., 2008; Balestre et al., 2008; Dandolini et al., 2008; Silva et al, 2009). Several molecular marker platforms have been employed in analysing genetic diversity, quantitative trait loci (QTL) identification and in predicting heterosis in maize, although results on the latter aspect have been inconsistent (Smith et al., 1997; Ajmone-Marsan et al., 1998; Pejic et al., 1998; Melchinger, 1999; Phumichai et al., 2008; Dhliwayo et al., 2009).

The relationship between genetic distance and heterosis was reported before the development of genetic markers (Moll et al., 1965). The theory of quantitative genetics describes a correlation between parental divergence and the heterosis estimates (Falconer and Mackay, 1996). Thus, heterosis is a function of the square of the differences between the allele frequencies in the parents, that is, the genetic divergence and the dominance effect of the alleles controlling the traits in question (Falconer, 1981). However, for maize, the results available for use of molecular markers to predict heterosis cannot be considered conclusive. Dudley et al. (1991) found no significant correlation between these variables in maize. Lanza et al. (1997) obtained a significant correlation between grain yield data and random amplified polymorphic DNA (RAPD) based genetic distances. Amorim et al. (2006) found high correlation between grain yield and genetic divergence for inter-population hybrids, but this correlation was low for intra-population hybrids, showing that markers would be efficient in predicting hybrids derived from different heterotic groups.

Recurrent backcrossing is a traditional breeding method, which is used frequently to transfer alleles at one or more loci from a donor to an elite variety (Allard, 1960; Reyes-Valdés, 2000). Recurrent backcrossing was used to introduce the *Fbr1* tassel mutation from a Mexican donor line into CIMMYT elite maize germplasm and there is great potential of using lines with the *Fbr1* trait in breeding programmes. Alongside the study of effects of the *Fbr1* trait under stress conditions; mapping of quantitative trait loci (QTL) associated with the *Fbr1* trait is work that needs to be done. This will provide valuable information and insights on the usefulness of this new tassel mutation in potential marker-assisted selection (MAS) breeding programmes. The initial step, then, should be to genotype the *Fbr1* lines to assess their homozygosity levels. If the converted lines are not homozygous enough for use in breeding and molecular work, then more backcrossing/selfing may be required before their use. Assessing the relatedness of these maize lines will help in future hybridization programmes involving the lines. More so, allele-based estimates of genetic distances between lines by use of molecular markers will allow the substitution of heterotic grouping based on phenotypic divergence of lines, and will facilitate early identification of contrasting parents for making crosses, hence reducing time required to

conclude breeding programmes. The objective of this study was (i) to fingerprint *Fbr1* and non-*Fbr1* CIMMYT maize lines using SNP markers to assess their relatedness and level of homozygosity; (ii) to determine SNP-based genetic distance estimates among *Fbr1* maize lines and to find correlation of genetic distance with SCA, heterosis and grain yield of the hybrids.

Materials and Methods

Germplasm for SNP and diallel analyses

Twenty six CIMMYT maize inbred lines adapted to the mid-altitude, tropical and/or subtropical environments of southern Africa were used in this study: 12 were *Fbr1* and 14 have normal tassels (Table 1). The 12 lines were an arbitrary sample (within each CIMMYT maize line family, to make sure each family is represented in the sample), of the *Fbr1* genotypes produced after the tassel mutation was introgressed from a Mexican donor line into CIMMYT elite maize lines.

SNP genotyping of maize lines

The maize inbred lines were advanced by selfing during the 2009/2010 summer season at CIMMYT-Harare research station prior to genotyping. After harvesting, 34 seeds per inbred line were packed in envelopes and shipped to BecA hub, Kenya, for the molecular marker analysis.

DNA extraction and SNP genotyping

Seedlings were raised in plastic seed trays for about

Table 1. CIMMYT maize inbred lines characterized by the 1074 known SNP markers

Line	Code	Pedigree	Heterotic Group
L1 L2	CML443 TAS2 CML444 TAS3	[CML443/TAS]BC2-2-9-1-2-B [CML444/TAS]BC2-5Y-3-1-B	A/B B
L3	CML488 TAS	[CML488/TAS]BC2-6-4-2-B	A/B
L4	CML443	CML443	A/B
L5	CML444	CML444	B
L6	CML488	CML488	A/B
		[[CML395/TAS]BC2/[(CML395/CML444)-B-4-1-3-1-	
L7	CML395 TAS	B/CML395//DTPWC8F31-1-1-2-2]-5-1-2-2-B]-8-2-2-B	B
L8	CML443TAS1	[CML443/TAS]BC2-2-5-3-1-B	A/B
		[[CML444/TAS]BC1/[CML444/CML395//DTPWC8F31-4-2-1-6]-2-1-1-1-	
L9	CML444 TAS1	B]-9-3-4-B	B
L10	CML445 TAS1	[[CML445/TAS]BC3/[CML445/ZM621B]-2-1-2-3-1-B]-2-4-2-B	A/B
L11	CML445 TAS2	[CML445/TAS]BC3-1-1-2-1-B	A/B
L12	CML312 TAS	[[CML312/TAS]BC1/MAS[MSR/312]-117-2-2-1-B]-1-3-1-B	A
L13	CML444 TAS2	[CML444/TAS]BC2-6-1-1-B	B
L14	CML442 TAS	[[CML442/TAS]BC1/ZM621A-10-1-1-1-2-BBBBBB]-2-1-B	A
L15	CML445 TAS3	[CML445/TAS]BC3-1-1-2-2-B	A/B
L16	CML445	CML445	A/B
L17	CML395	CML395	B
L18	CML312	CML312	A
L19	CML442	CML442	A
	LaPostaSeqC7-		
L20	F180	LaPostaSeqC7-F180	B
L21	LaPostaSeqC7-F18	LaPostaSeqC7-F18	B
L22	CKL05005	CKL05005	B
L23	G16BNSeqC4	G16BNSeqC4	A
L24	LaPostaSeqC7-F71	LaPostaSeqC7-F71	B
L25	CKL05003	CKL05003	B
L26	CML144	CML144	A

two weeks until three to four-leaf stage in a greenhouse at the BecA hub in Nairobi, Kenya. Equal amounts of leaf tissues were harvested from 10 plants per inbred line, and were bulked, cut into pieces with scissors, and transferred into 1.2 ml strip tubes that contained two 4 mm stainless steel grinding balls. The tissue was freeze-dried (lyophilised) for 3 days using a Labconco freeze dryer (http:// www.labconco.com), as described in the user's manual. The lyophilised leaf samples were ground into fine powder using GenoGrinder 2000 (Spex CertiPrep) at 1500 strokes per minute for 4 minutes at speed = 1x. Genomic DNA was extracted using a modified version of the high throughput mini-prep Cetyl Trimethyl Ammonium Bromide (CTAB) method (Mace et al., 2003).

SNP genotyping and allele calling

Generally, SNP genotyping and allele calling was made by KBiosciences (http://www.KBioscience.co.uk) [2010, November 30] using the KASPar system as described in the user's manual (http://www.kbioscience.co.uk/reagents/KASParSNP_Genotyping_SystemLeafletv6.3.pdf) [2010, November 30].

The design for the KASPar was achieved using the PrimerPicker software found at http://www.kbioscience.co.uk/primer-picker/ [2010, November 30].

Screening of SNP data

SNP markers (a total of 1250) were used for characterising the inbred lines of which 1242 SNPs had data that passed the quality control checks of KBiosciences. Eight SNP markers, BDIBC175, PHM2187_46, PZA01857_1, PZA03012_7, PZA02681_8, PZA00939_1, PHM4757_14 and PHM18705_23, when used to genotype the samples, did not return quality data. One hundred and sixty one of the 1242 successful assays were monomorphic in all the lines, and seven markers had extremely high heterozygosity, therefore 1074 SNPs were used for final evaluation of the maize lines

Diallel analysis and field data collection

Two half diallel crosses were done at Muzarabani (Zimbabwe) to make ($n(n-1)/2$) F_1 crosses (Griffing, 1956), that were evaluated in trials (2010 and 2011) under optimum, low N and drought conditions. The first half diallel crossing set-up involved inbred lines L1-L6 (Table 1), and produced 15 F_1 hybrids that were evaluated in yield trials in 2010 and 2011. The second half diallel crossing set-up involved nine inbred lines: L7-L14 and L3, and produced 36 F_1 hybrids that were, similarly, evaluated for grain yield in 2010 and 2011 under optimum, low N and drought conditions

Agronomic management, environments and stress management of trials

The three trial environments were CIMMYT- Harare Maize Research Station (17.80 S, 31.05 E, 1468 masl) (optimum conditions), CIMMYT-Harare Maize Research Station under low N during the summer wet season, and Chiredzi Research Station (21.03 S; 31.57 E, 392 masl) during the winter dry season (under managed drought). Under optimum growing conditions in all sites, a basal application of 400 kg/ha of compound Z fertilizer (8% N: 14% P_2O_5: 7% K_2O: 0.8% Zn) was broadcast and disc–incorporated by a tractor. Ammonium nitrate (33% N) was split applied at 200 kg/ha. The first application of 100kg/ha was done at four weeks after crop emergence and the second split was given at six weeks after emergence. Trials were rain-fed, but a light irrigation was applied immediately after planting to facilitate seed germination and seedling emergence. Irrigation was also applied in the case of a long dry spell. Generally, an irrigation of 7mm/hr for six hours was applied just after planting to facilitate germination. Total water application per irrigation was 42 mm. Thereafter, the irrigation interval varied from 9 to 15 days depending on temperature and crop development stage. Average rainfall was 700-800 mm and 650-700mm potential evapotranspiration was experienced during the growing seasons for Harare in 2010 and 2011.

The experiments under low N were also conducted at Harare using, except for N management, the same crop management practices as under recommended agronomic management. Low N experiments were grown in fields that were depleted of N by continuously cropping maize (main season) or irrigated wheat (winter dry season), removing all stover biomass after harvest and not applying any N fertilizer. No chemical N fertilizer was applied to the low N experiments. For trials under managed drought stress in Chiredzi, three to four irrigations totalling 250 mm of water were applied at the beginning of the season and irrigation stopped at 43 to 57 days after planting (about 50 days before anthesis). The crop completed its life cycle without any further irrigation or rain.

Four sets of trials, the six inbred parents and the 15 hybrids plus five hybrid checks; the nine inbred lines and 36 hybrids plus four hybrid checks, were grown adjacent to each other in three environments in Zimbabwe during 2010 and 2011. The experimental design was an alpha lattice (0,1) (Patterson et al., 1978) with two replications for hybrids and inbreds in each environment. The 15 crosses plus five hybrid checks were grown using one-row plots, two replications and 4 x 5 incomplete lattice designs in all the three environments in 2010 and 2011, while the 36 hybrids plus six hybrid checks for the second trial were laid out in a 6 x 7 alpha lattice design, for grain yield evaluation. Two trials of inbred parents (for

the two hybrid trials) were grown side by side with the hybrids to facilitate estimation of heterosis. Plot size at all locations was a single 4 m row with 0.75 m between rows and 0.25 m between plants within a row, giving final plant populations of ≈ 53 000 plants per hectare at all sites. Grain yield (adjusted to 12.5% moisture content) was obtained considering harvested plot area and counting number of plants and harvested ears per plot.

SNP data analysis

Summary statistics of genetic data such as minor allele frequencies, polymorphic information content (PIC), heterozygosity and number of alleles were computed with Powermarker version
3.25 (Liu and Muse, 2005). Modified Roger's genetic distance (MRD) (Wright, 1978; Goodman and Stuber, 1983) between each pair of inbred lines was computed as:
$MRD_{ij} = ½ [\sum (X_{ai} - X_{aj})^2]^{½}$, where X_{ai} is the frequency of the allele a for individual *i*, and X_{aj} is the frequency of the allele a for the individual *j*.
The PIC for each locus was determined as described by Smith et al. (1997). The genetic relationship among inbred lines was assessed using cluster analysis performed on the MRD distance matrix with UPGMA clustering.
Associations among genotypes were revealed with the Principal Component Analysis (PCA) algorithms (Gower, 1966) implemented in XLSTAT (2010), a statistical and multivariate analysis software (www.xlstat.com), based on MRD estimates between pairs of inbred lines.
Genotypes were grouped into two classes, according to tassel morphology: either normal tasselled (1) or Fbr1 (2). Estimates of missing data were done using nearest neighbour analysis and the PCA type is Pearson (n).

Statistical analysis for agronomic data

Individual analyses of variance were performed for each experiment with the general linear model procedure (PROC GLM) from SAS (SAS, 2003). The adjusted means were used to make subsequent calculations to estimate SCA. SCA was estimated using the Line x tester analysis programme in SAS (SAS, 2003). The fixed-effects model of diallel method 4 was used in the analysis and provided estimates of SCA effects for the hybrids across all environments. Mid parent heterosis was calculated as:

$$MPH = \frac{F1\text{-}MP}{MP} \times 100$$

where, F_1 is the mean of the F_1 hybrid performance and $MP = (P_1 + P_2)/2$ in which P_1 and P_2 are the means of the inbred parents, respectively.
High-parent heterosis was calculated as:

$$HPH = \frac{F1\text{-}HP}{HP} \times 100$$

where HP is the mean of the best parent.
Pearson correlation coefficient (r) between genetic distance (GD) and single cross grain yield (F_1), MP, HP, MPH, HPH, and SCA were calculated from the means across environments. Statistical computations were performed with SAS statistical package (SAS, 2003). Broad sense heritability for grain yield for the hybrid sets was estimated using the formula $(1- 1/F_{value})$. The F value was computed in the ANOVA across sites and years.

Results and discussion

Genetic diversity

Average residual heterozygosity for the maize inbred lines ranged from 0.2 to 36.1%, with an average of 8.2%, which is however, well above the expected ranges for residual heterozygosity found in maize inbreds. This could be because some of the inbred lines are still heterozygous and these need further selfing to reduce the residual heterozygosity. Yan et al., (2009) found heterozygosity ranging from 0 to 9.9%, with an average of 2.5% in a highly diverse global maize collection of 632 inbred lines from temperate, tropical, and subtropical public breeding programmes, which were reported as within expected ranges. Xia et al., (2004) also found an average residual heterozygosity of 4.8% among CIMMYT maize inbred lines investigated with SSR markers, which were in accordance with results reported by Heckenberger et al., (2002).
The PIC values for the 1074 SNP loci ranged from 0.015 to 0.50, with an average of 0.25. Thus, the SNP loci were informative and were able to detect differences among inbred lines based on their genetic relationships. The average PIC value was however lower than that reported previously for tropical and temperate maize lines (Dhliwayo et al., 2009; Betrán et al., 2003; Xia et al., 2004; Senior et al., 1998; Barata and Carena, 2006). There was therefore, relatively little genetic diversity among the germplasm used in this study, which is an indication that most of the inbred lines evaluated were close to fixation. The average inbreeding coefficient of 0.73 for the SNP loci further confirmed the fixation of the maize lines. Since the aim of this study was to identify normal and *Fbr1*-converted lines that are homozygous, and that can be used as parents in breeding programmes involving the *Fbr1* tassel mutation, these homozygous lines are useful in making crosses for test cross evaluations and in making mapping populations in planned marker assisted breeding work.
The minor SNP allele frequency distribution in the CIMMYT maize lines indicated a close-to-uniform distribution in the 0-0.10, 0.11-0.20, 0.21-0.30, 0.31-0.40, and 0.41-0.50 classes. Twenty five percent of the SNP loci fall into a class where the minor allele frequency was

Table 2. Homozygosity levels of the maize inbred lines characterised using the 1074 SNPs

Inbred line	% missing data	% homozygosity
CML445/TAS-BC3-Source2	7.7%	63.9%
CML443	4.3%	66.2%
CML312	11.6%	71.2%
CML444-Source1	16.7%	76.6%
CML445/TAS-BC3-Source1	4.1%	77.6%
CML444/TAS-BC2-5Y	3.3%	80.9%
LaPostaSeqC7-F18	11.3%	82.1%
CML444-Source3	20.4%	85.4%
CKL05005	6.9%	87.6%
CML395/TAS	2.8%	90.8%
CML445/TAS	2.2%	91.4%
CML443/TAS-BC2	2.7%	92.1%
CML442/TAS	2.2%	94.2%
CML444/TAS	2.3%	94.5%
CML488/TAS-BC2	1.1%	96.5%
CML443/TASBC2-5Y	2.3%	96.7%
G16BNSeqC4	4.3%	96.7%
LaPostaSeqC7-F71	20.4%	96.8%
DTPWC9-F92	2.5%	96.8%
CML488	1.1%	97.0%
CKL05003	2.4%	97.2%
CML445	3.3%	98.1%
CML444-Source2	1.7%	98.3%
CML312/TAS	2.0%	98.4%
CML444/TAS-BC2	1.7%	98.8%
CML442/CML197/TAS	2.9%	99.1%
CML312/CML445/TAS	1.7%	99.2%
CML395	1.7%	99.2%
LaPostaSeqC7-F180-Source1	3.2%	99.3%
CML144	2.8%	99.3%
CML444-Source4	2.2%	99.3%
CML442	1.1%	99.4%
ZEWAc1F2-134	1.3%	99.5%
LaPostaSeqC7-F180-Source2	1.1%	99.8%

less than 0.10, implying that the genetic characterization done with the 1074 SNP markers was reliable and informative.

Homozygosity of the CIMMYT maize lines

Table 2 show levels of homozygosity of the CIMMYT maize inbred lines characterized using the 1074 SNPs. Because genotyping of lines was done together with a number of other CIMMYT lines on the same SNP genotyping platform, additional inbred lines were added to the analysis to increase the scope of information generated. The total number of lines used for the analysis became 34.

The proportion of homozygous loci for the total markers used for all inbred lines characterized ranged from 63.9% for CML445/TAS-BC3-source2 (the most heterozygous line), to 99.8% for LaPostaSeqC7-F180-source2 (the most homozygous line). Maize inbreds from CML488/TAS- BC2 (in descending order) to LaPostaSeqC7-F180-Source2 were acceptably homozygous (≤ 5% heterozygosity for the 1074 SNPs used). It was surprising that CML443, CML312, CML444-Source1, and CML444-Source3 were heterozygous, as they are expected to be fixed. It was also unexpected that the same lines, though from different sources (for example, the four CML444's) had large differences in terms of homozygosity levels. CML444-Source1 was 76.6% homozygous, while CML444-Source4 had a homozygosity level of 99.3%. The reason could be that, while the greatest care is taken to maintain genetic purity during maintenance of these lines in breeding programmes, there are chances of contamination in the field during pollination, and seed mixes can occur during seed preparation. Consequently, these cause variation within lines that were originally fixed.

Yan et al., (2009) compared 21 CIMMYT maize inbred lines to lines with same name but maintained in different labs for more than 30 years and found that 81% of the lines were still genetically similar while 19% had become different. It is therefore critical to assess homozygosity of CIMMYT lines, especially at molecular level to verify fixation of lines before embarking in critical programmes like making test crosses for QTL analysis (where homozygosity of parental lines is critical), as some of the lines would have become heterozygous during the maintenance course.

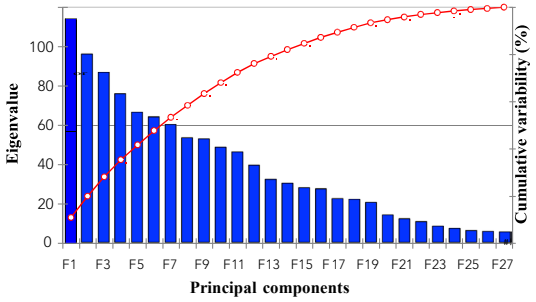

Figure 1 - Scree plot of eigenvalues: corresponding proportion and cumulative variation for all the principal components for tassel size in the maize lines.

Fbr1 lines CML488/TAS-BC2, CML443/TASBC2-5Y, CML312/TAS, CML444/TAS-BC2,

Table 3. Combined analysis of variance across sites and years for grain yield for the two sets of hybrids formed from the two diallel mating designs and parental inbred lines used in F1 hybrid formation

	Hybrid set 1		Hybrid set 2	
Source of variation	**Degrees of freedom**	**Mean square**	**Degrees of freedom**	**Mean square**
Environment	2	214.38***	2	69.023***
Rep (Env)	3	0.59	3	0.38
Entry	34	1.15***	14	2.32***
GCAfemale	7	0.82*	4	2.44***
GCAmale	7	1.86***	5	2.34***
SCA	20	1.06***	5	2.19***
Entry x Env	68	0.45	28	0.53***
GCAfemale x Env	14	0.35	8	0.39
GCAmale x Env	14	0.59	10	0.78***
SCA x Env	40	0.41	10	0.38
Error	319	0.39	120	0.21

	Inbred line set 1		Inbred line set 2	
	Degrees of freedom	**Mean square**	**Degrees of freedom**	**Mean square**
Replication	1	0.071	1	0.11*
Entry	5	0.68***	8	0.51***
Environment	2	4.61***	2	9.69***
Year	1	0.72**	1	1.37***
Entry x Env	10	0.14	16	0.16***
Entry x Year	5	0.18	8	0.14***
Entry x Env x Year	12	0.19*	18	0.31***
Error	34	0.089	52	0.027

* P ≤ 0.05, ** P ≤ 0.05, *** P ≤ 0.001

CML442/CML197/TAS and CML312/CML445/TAS, were also acceptably homozygous (≤ 5% heterozygosity for the 1074 SNPs used), indicating that most alleles from the recurrent parent have been retained after the introduction of the Fbr1 tassel mutation into these genotypes. These Fbr1 converted lines are fixed for the tassel mutation and these lines could be used as parental lines in the development of mapping populations for future marker assisted breeding work.

Principal component analysis

Principal component analysis was done to determine the amount of genotypic variation for tassel size explained by the SNP markers. Principal components (PC) or factors (F) plotted against the cumulative genetic variability explained by the 1074 SNP markers showed that the eigenvalue for PC1 was highest (114.04) and explained 10.87% of genetic variability for tassel size in the maize lines (Figure 1). PC2 and PC3 explained 9.08 and 8.20% of genetic variance for tassel size respectively, and consequently explained 19.95 and 28.15% respectively, of cumulative variability for tassel size. Of the 1074 SNPs, 7.3% did not contribute to the variation observed in PC1 with 69.8% of the SNPs contributing less than 0.1% variation in tassel size. SNP marker PZB00772_7 contributed most to variation observed in PC1 (0.77%). Eleven percent of the 1074 SNPs did not contribute to the variation observed in PC2 while marker PHM4066_11 made the highest contribution (0.57%) to variation observed for tassel size. Of the total SNP markers used, 68.6% contributed less than 0.1% of PC2 variation. These results showed that there was variation in contribution made by different SNP markers to differences observed for tassel size among genotypes.

The relative magnitude of the coefficients (eigenvec-

Table 4. Yield of the *Fbr1* maize hybrids in relation to mid-parent heterosis (MPH), high- parent heterosis (HPH) and specific combining ability (SCA)

(a)

Hybrid†	Yield (kg/plot)	MPH(%)	HPH(%)	SCA	Hybrid	Yield (kg/plot)	MPH(%)	HPH(%)	SCA
1x2	1.54	47.30	20.50	0.037	7x10	1.41	73.09	10.26	-0.106
1x3	1.83	157.07	43.86	0.077	7x11	1.90	141.02	48.71	0.164
1x4	1.87	143.97	46.98	-0.044	7x12	2.16	167.87	69.38	0.095
1x5	1.49	63.74	16.65	0.073	7x13	1.96	131.19	53.69	0.193
1x6	1.95	147.58	53.02	0.348	7x14	1.80	153.93	40.93	0.088
2x3	1.78	88.31	39.64	-0.267	8x10	1.72	133.73	34.79	0.158
2x4	2.19	119.22	71.95	-0.007	8x11	1.21	70.69	-4.98	-0.568
2x5	2.46	115.44	92.67	-0.102	8x12	1.92	163.99	51.00	-0.188
2x6	1.95	91.40	53.11	0.033	8x13	1.99	158.63	56.32	0.320
3x4	1.79	166.52	40.18	0.263	8x14	1.70	169.70	33.41	-0.056
3x5	2.01	148.24	57.85	0.119	9x10	1.38	63.57	8.43	-0.005
3x6	0.49	-28.97	-61.53	-0.748	9x11	2.11	157.96	65.85	0.507
4x5	2.48	186.40	94.38	-0.139	9x12	2.26	169.00	77.06	0.265
4x6	1.83	145.13	43.25	-0.124	9x13	1.11	26.22	-12.82	-0.530
5x6	2.09	135.68	63.63	0.480	9x14	1.84	148.98	44.63	0.202
7x3	1.59	100.58	24.45	-0.220	10x11	1.67	69.60	30.80	0.025
7x8	1.87	247.85	46.97	-0.009	10x12	1.84	83.51	44.33	-0.135
8x3	1.84	157.68	44.33	-0.015	10x13	1.69	62.35	32.95	0.018
8x9	2.41	322.52	89.46	-0.058	10x14	1.69	87.47	32.95	0.009
9x3	1.47	78.58	15.43	-0.211	11x12	2.16	120.97	69.38	-0.062
10x3	1.69	70.70	32.23	-0.032	11x13	2.02	98.41	58.52	0.098
11x3	1.90	98.02	49.43	-0.059	11x14	1.72	95.57	34.79	-0.147
12x3	2.10	113.79	64.61	0.490	12x13	1.68	62.16	32.04	0.116
13x3	1.76	72.46	38.38	-0.027	12x14	1.00	11.77	-21.26	-0.506
3x14	1.80	103.95	41.26	0.086	13x14	2.06	119.46	61.73	0.097

(b) Correlation coefficients between molecular-based genetic distance (MRD), grain yield, mid-parent heterosis (MPH), high-parent heterosis (HPH), and specific combining ability (SCA)

Crosses	Grain Yield (kg/plot)	MPH (%)	HPH (%)	SCA
MPH	0.73***			
HPH	0.99***	0.73***		
SCA	0.63***	0.46***	0.63***	
MRD	0.50***	0.42*	0.50***	0.45**

* P ≤ 0.05, ** P ≤ 0.05, *** P ≤ 0.001

† The pedigree information of the lines used to make the hybrids is shown in Table 1.

tors), which reflects the relative contribution of each genotype to PC scores showed that genotypes CML395/TAS, CML395, and CKL05003 made the highest contribution to PC1 together while genotypes CML445, CML445/TAS-BC3- S1, CML444TAS, CML443/TAS-BC2, CML442/CML197/TAS and CML312/CML445/TAS made high negative contribution to PC1. Inbred lines CML442/TAS and CML312/TAS contributed most to variation in PC2 while lines CML488 and CML488/TAS-BC2 made the highest negative contribution to PC2.

Genetic distance among inbred lines and cluster analysis based on the SNP markers.

The mean genetic distance for all pair wise comparisons was 0.30, which is lower than that reported in previous studies for tropical germplasm (Xia et al., 2004; 2005), and that reported among elite CIMMYT and IITA tropical maize inbred lines (Dhliwayo et al., 2009). Reif et al., (2003), investigating the diversity among seven of CIMMYT's tropical maize populations with molecular markers, also identified low variance between populations. The lower average MRD suggests a high average degree of relatedness among the CIMMYT maize lines. Genetic distance ranged from MRD of 0.02 (between

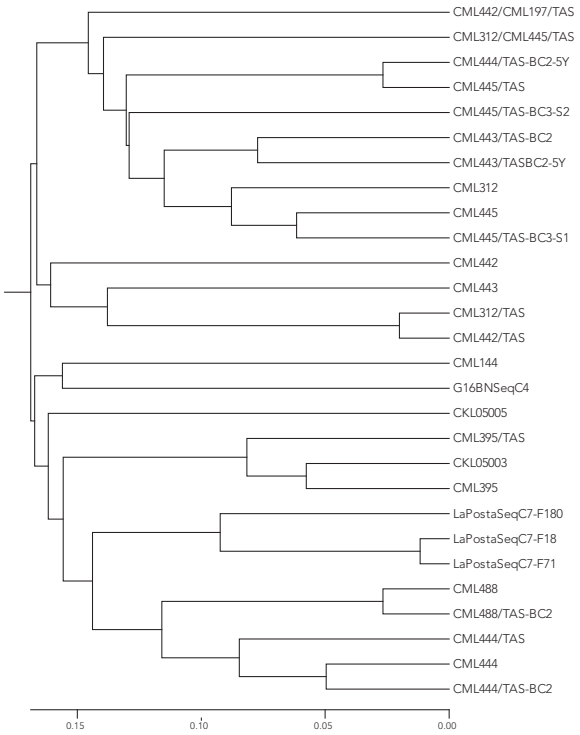

Figure 2 - Dendrogram constructed using unweighted pair group method with arithmetic mean clustering of maize inbred lines from CIMMYT based on 1074 SNPs. The scale bar on the axis is expressed in Modified Roger's distance (MRD) (Wright, 1978; Goodman and Stuber, 1983), which shows percentage dissimilarity between or among genotypes

LaPosta SeqC7-F71 and LaPosta SeqC7-F18) to 0.39 (between La PostaSeqC7-F180 and CML312/CML445/TAS). UPGMA clustering showed two major clusters (Figure 2). One cluster (bottom cluster) consisted of inbred lines in heterotic group B and A/B, while the other cluster (top cluster) consisted of inbred lines in heterotic group A and A/B. Hence, the maize lines were clustered according to pre-defined heterotic grouping, showing efficiency of the SNP markers in characterising these inbred lines. The tight clustering of CML312/TAS and CML442/TAS was surprising (4% dissimilar) since the lines are genetically different and do not have the same ancestry.

Genetic analysis of maize lines and hybrids

The analysis of variance for grain yield of inbred lines and hybrids showed that genotypic and environmental variations were highly significant ($P \leq 0.001$) for both hybrids and parental inbred lines (Table 3). GCA_{female} and GCA_{male} were significant for the two sets of hybrids while SCA was also highly significant in the two hybrid sets. SCA ranged from -0.75 (hybrid 3x6 i.e. CML488TAS x CML488) to 0.507 (hybrid 9x11 i.e. CML444TAS1 x CML445TAS2) (Table 4).

The reason why hybrid CML488TAS x CML488 was the poorest in yield performance is that both parents constituted the same line (CML488), the difference being that, one of the parental lines has the small tassel (*Fbr1*) gene, while the other parent is normal tasselled. Inbreeding depression may be the major cause of the serious yield reduction. Most hybrids that had high and positive SCA values for grain yield also had high mean grain yield and fall in complementary heterotic groups, for example hybrids 9x11, 1x6, 5x6 and 8x13 had high mean grain yield and positive SCA for grain yield. Lines that constituted these hybrids fall in complementary heterotic groups: A/B and B (Table 1). Thus, CIMMYT's predefined heterotic grouping of lines consistently predicts the performance of hybrids, suggesting that these heterotic groups were well defined.

Correlation of genetic distance with hybrid performance and heterosis

Estimates of grain yield, mid-parent heterosis (MPH), high-parent heterosis (HPH) and specific combining ability (SCA), for the maize hybrids are presented in Table 4a. Grain yield ranged from 0.49 to 2.48 kg/plot for the hybrids: CML488 x CML488/TAS-BC2 and CML443 x CML444 respectively, with average grain yield of 1.80 kg/plot across all hybrid sets. The highest MPH (323%) was recorded for the hybrid 8x9, i.e. CML444/TAS1 x CML443/TAS1, while the lowest MPH (-28.97%) was detected in the combination CML488 x CML488/TAS, which are sister lines. It is worth noting that this particular hybrid also recorded the lowest mean grain yield, and had the smallest SCA for grain yield. HPH recorded an average of 41.51% in the maize hybrids, with hybrid CML488 x CML488/TAS similarly recording the lowest HPH value of -61.53%, while CML443 x CML444 had the highest HPH of 94.38%. Betrán et al., (2003) also recorded the lowest SCA effects for grain yield for hybrids between sister lines LP4 and LP5. They found that SCA across environments was generally negative for hybrids involving inbred lines with the same germplasm origin or related by pedigree, and was greater for hybrids involving inbred lines of different source germplasm origin. Sister lines or lines related by pedigree lack the interaction of the genes, in favour of cumulative dominant alleles, which are useful in the expression of heterosis in F_1 hybrid combination (Qi et al., 2010).

Grain yield for hybrids across environments was positively correlated with MPH, HPH, SCA and genetic distance and the correlation coefficients were highly significant ($P \leq 0.001$, Table 4b). The highest correlation was observed between grain yield and HPH. The correlation between grain yield and MPH was also relatively high ($r = 0.73^{***}$), indicating that heterosis can predict hybrid performance better than SCA among

parental lines or molecular marker-based genetic distance. On the contrary, Betrán et al., (2003) found that SCA among lines was highly correlated with grain yield across stress and non-stress environments and justified prediction of hybrid performance based on SCA. They argued that heterosis is highly dependent on the performance of inbred lines, and there is differential response of inbred lines to stresses and environmental conditions relative to hybrids, rendering predictions based on heterosis erratic and inconsistent across environments. SCA was positively correlated with MPH (r = 0.46***) and HPH (r = 0.63***) across environments.

Parental genetic distance was positively correlated with grain yield, heterosis (MPH and HPH), and SCA. Similarly, Mladenovic-Drinic et al., (2002) found positive correlation between genetic distance and these parameters. Ajmone-Marsan et al., (1998) also obtained highly significant but modest estimates of correlation coefficients between genetic distance and yield within a set of 78 maize hybrids studied, for the two classes of molecular markers, RFLP and AFLP. The correlation coefficient of genetic distance and SCA was moderate and significant (r = 0.45***). Previous experiments with diallel crosses indicated correlation between genetic distance and SCA for grain yield ranging from very low (Dudley et al., 1991), medium (Melchinger et al., 1990; 1992), to rarely, very high (Lee et al., 1989).

The correlation between MPH, and HPH for grain yield with genetic distance were positive and moderate (r = 0.42* and r = 0.50*** respectively). Thus, genetic distance of parental lines, to some extent, determines hybrid vigour expected in hybrid progeny. Boppenmaier et al., (1992), Dhillon et al., (1993) and Ajmone-Marsan et al., (1998) found relatively low values of correlation coefficients between genetic distance and heterosis. Betrán et al., (1997) studied germplasm of tropical white maize using RFLP markers and obtained low values of correlation coefficients between genetic distance and SCA, and between genetic distance with grain yield and heterosis. On the other hand, Smith et al., (1990) obtained very high correlation (r = 0.87) between RFLP- based genetic distance and heterosis in crosses of inbred lines from the same and different heterotic groups.

Dhliwayo et al., (2009), however, found no significant association of genetic distance with grain yield, MPH and SCA. Regarding RAPD markers, Rinaldi et al., (2007) also did not infer a significant correlation between heterosis and productivity in Brazilian popcorn populations. Bernardo (1992) and Melchinger (1999) summarised some theoretical considerations that often lead to poor predictive value of genetic distance for hybrid performance; these include the small role of dominance gene action, low heritability of the trait (as is the case for grain yield), and few trait-relevant QTL linked to the molecular markers. In this study, SCA effects for grain yield were mostly significant, thus dominance gene action played a major role in determining yield of hybrids and broad sense heritability for grain yield as determined by the formula (1- $1/F_{value}$) was relatively high (an average broad sense heritability for the two hybrid sets was H = 0.79). Thus, according to Bernardo (1992) and Melchinger (1999), high predictive value of genetic distance for hybrid performance was expected. Lack or low correlation of genetic distance with hybrid performance, heterosis, and SCA was also suggested to be a result of lack of linkage between genes controlling the traits under analysis, unequal or insufficient genome coverage, random marker distribution and diversified effect of dominance (Melchinger et al., 1990; Charcosset et al., 1991; Kwon et al., 2002). Prediction of heterosis based on marker loci would therefore, be more efficient if the markers are selected *a priori*, for their relationship to the alleles implicated in the heterotic traits.

According to Mladenovic-Drinic et al., (2002), the absence of linkage between molecular markers used to estimate divergence and the genes controlling heterosis for the studied traits could explain low correlation observed between heterosis and genetic distance. Therefore, markers must be in linkage disequilibrium with QTLs to have a predictive value. Charcosset and Essioux (1994) suggested that necessary conditions for prediction efficiency should be fulfilled at the within-group level and at a general level. Linkage disequilibria between markers and QTLs generally differ randomly from one heterotic group to another, thus genetic distance based on neutral marker loci will not be predictive for the performance of between-groups hybrids.

Conlusion and recommendation

The 1074 SNP marker loci used to characterise the maize parental lines indicated that the mean genetic distance for all pairwise comparisons of lines was low (0.30) suggesting a high level of relatedness among lines. Inbred lines can therefore be isolated from this germplasm for future breeding work involving the *Fbr1* tassel mutation. The fact that genetic distances were able to effectively group the maize inbred lines according to heterotic patterns used by CIMMYT highlights the potential value of genetic distances for preliminary classification of poorly characterised germplasm. The results confirm molecular markers as a powerful complement to help assign lines into defined heterotic groups and to examine the relationships among inbred lines at DNA level.

Molecular markers were useful to determine heterotic grouping in a short time.

Our results revealed the efficiency of backcrossing in converting elite normal-tasselled CIMMYT maize lines to few-branched-lines since most of *Fbr1* lines were homozygous for the SNP loci used. The fact that many homozygous elite lines with the *Fbr1* trait were identified could open a new window in potential marker assisted selection (MAS) for the trait. Furthermore, more homozygous lines with the *Fbr1* trait could be used in breeding programmes aimed at unveiling the untapped potential of these new mutants in maize production.

Grain yield for the hybrids ranged from 0.49-2.48 kg/plot, with an average of 1.80 kg/plot and hybrids constituted from closely related parental lines (according to SNP-based genetic distances) had the lowest mean grain yield, lowest SCA effects for grain yield, and low heterosis. Although determination of the genetic basis of hybrid performance and measuring the relationship between marker-based genetic distance and complex agronomic traits like yield are reported to be quite complex, significant and positive correlations of genetic distance with grain yield, heterosis and SCA were found in this study. Thus, SNP- based genetic distances could be used as efficient predictors of hybrid performance in this set of maize germplasm. Results of this study suggest that SNP-based genetic distance information would aid in the selection of genetically wide lines to include in breeding programmes where inclusion of diverse lines as parents is critical, for example, in synthetic variety formation. Although few breeding programmes rely less on recurrent selection schemes, DNA-based genetic distance could be useful in guiding the introgresion of exotic germplasm into existing local heterotic germplasm, or in initial grouping of uncharacterised germplasm. Our results showed that SNP- based genetic distances were effective in grouping CIMMYT maize lines into predefined heterotic groups. However, it would also be important to test the utility of these SNP-based genetic distances for selecting lines for use in formation of synthetic varieties or in defining a new pair of complementary heterotic populations for subsequent exploitation.

Acknowledgement

Financial support of this work by the International Maize and Wheat Improvement Centre (CIMMYT) is gratefully acknowledged.

References

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: 219-227

Allard RW, 1960. Principles of plant breeding. Wiley, New York

Amorim EP, Amorim UBO, Santos JB, de Souza AP, de Souza JC, 2006. Genetic distance based on SSR and grain yield of inter and intra-populational maize single cross hybrids. Maydica 51: 507-513

Aydin N, Sabri G, Ahmet Y, Ahmet O, Giovanni F, Hikmet B, 2007. Estimating genetic variation among dent corn inbred lines and topcrosses using multivariate analysis. Journal of Applied Biological Science 1: 63-70

Balestre M, Machado JC, Lima JL, Souza JC, Nobrega Filho L, 2008. Genetic distance estimates among single cross hybrids and correlation with specific combining ability and yield in corn double cross hybrids. Genetics and Molecular Research 7: 65-73

Barata C, Carena MJ, 2006. Classification of North Dakota maize inbred lines into heterotic groups based on molecular and testcross data. Euphytica 151: 339-349

Batley J, Barker G, O'Sullivan H, Edwards KJ, Edwards D, 2003a. Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. Plant Physiology 132: 84-91

Batley J, Mogg R, Edwards D, O'Sullivan H, Edwards KJ, 2003b. A high-throughput SNuPE assay for genotyping SNPs in the flanking regions of *Zea mays* sequence tagged simple sequence repeats. Molecular Breeding 11: 111-120

Bernardo R, 1992. Relationship between single-cross performance and molecular marker heterozygosity. Theoretical and Applied Genetics 83: 628-634

Betrán FJ, Beck DL, Edmeades GO, Ribaut JM, Bänziger M, Sánchez C, 1997.

Genetic analysis of abiotic stress tolerance in tropical maize hybrids. In: CIMMYT ed. The Genetics and Exploitation of Heterosis in Crops. An International Symposium, Mexico City, 17-22 Aug 1997. Mexico, DF, CIMMYT. pp. 28-29

Betrán FJ, Ribaut JM, Beck D, Gonzalez de Leon D, 2003. Genetic diversity, specific combining ability, and heterosis in tropical maize under stress and nonstress environments. Crop Sci 43: 797-806

Boppenmaier J, Melchinger AE, Brunklaus-Jung E, Geiger HH, Herman RG, 1992.

Genetic diversity for RFLPs in European maize inbreds: I. Relation to performance of flint x dent crosses for

forage traits. Crop Sci 32: 895-902
Charcosset A, Lefort-Buson M, Gallais A, 1991. Relationship between heterosis and heterozygosity at marker loci. Theor Appl Genet 81: 571-575
Charcosset A, Essioux L, 1994. The effect of population structure on the relationship between heterosis and heterozygosity at marker loci. Theor Appl Genet 89: 336-346
Crossa J, Franco J, 2004. Statistical methods for classifying genotypes. Euphytica 137: 19- 37.
Dandolini TS, Scapim CA, Amaral Junior AT, Mangolin CA, da Silva Machado MFP, de Souza Mott A, Lopes AD, 2008. Genetic divergence in popcorn lines detected by microsatellite markers. Crop Breeding and Applied Biotechnology 8: 313-320
Dhliwayo T, Pixley KV, Menkir A, Warbuton M, 2009. Combining ability, genetic distances, and heterosis among elite CIMMYT and IITA tropical maize inbred lines. Crop Sci 49: 1301-1310
Dhillon BS, Boppenmaier J, Pollmer WG, Herrmann RG, Melchinger AE, 1993.
Relationship of restriction fragment length polymorphism among European maize inbreds with ear dry matter yield of their hybrids. Maydica 38: 245-248
Drenkard E, Richter BG, Rozen S, Stutius LM, Angell NA, Mindrinos M, Cho RJ, Oefner PJ, Davis RW, Ausubel FM, 2000. A simple procedure for the analysis of single nucleotide polymorphisms facilitates map-based cloning in Arabidopsis. Plant Physiology 124: 1483-1492
Dudley JW, Saghai Maroof MA, Rufener GK, 1991. Molecular markers and grouping of parents in maize breeding programs. Crop Sci 31: 718-723
Falconer DS, 1981. Introduction to quantitative genetics,2nd edition, New York: John Wiley and
Sons.
Falconer DS, Mackay TF, 1996. Introduction to quantitative genetics. 3rd edition Longman, New York, USA.
Franco J, Crossa J, Ribaut JM, 2001. A method for combining molecular markers and phenotypic attribites for classifying plant genotypes. Theor Appl Genet 103: 944-952
Garg K, Green P, Nickerson DA, 1999. Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags. Genome Research 9: 1087-1092
Goodman MM, Stuber CW, 1983. Race of maize: VI. Isozyme variation among races of maize in Bolivia. Maydica 28: 169-187
Gower JC, 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325-338
Griffing B, 1956. Concept of general and specific combining ability in relation to diallel crossing systems. Australian J Biol Sci 9: 463-493
Gupta PK, Roy JK, Prasad M, 2001. Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Current Science 80: 524-535
Gupta PK, Rustgi S, Mir RR, 2008. Array-based high-throughput DNA markers for crop improvement. Heredity 101: 5-18
Hallauer AR, Miranda Fo JB, 1988. Quantitative genetics in maize breeding. Iowa State University Press, Ames, IA, USA
Heckenberger M, Bohn M, Ziegle JS, Joe LK, Hauser JD, Hutton M, Melchinger AE,
2002. Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties. I. Genetic and technical sources of variation in SSR data. Molecular Breeding 10: 181-191
Hinze LL, Lamkey KR, 2003. Absence of epistasis for grain yield in elite maize hybrids.
Crop Sci 43: 46-56
Kwon SJ, Ahn SN, Jeong EG, Hwang HG, Choi HC, Moon HP, 2002. Relationship between genetic divergence and hybrid performance in Japonica rice grown in a cold water irrigated field. Euphytica 128: 389-396
Lanza LLB, Souza Junior CL, Ottoboni LMM, Vieira MLC, de Souza AP, 1997.
Genetic distance of inbred lines and prediction of maize single-cross performance using RAPD markers. Theor Appl Genet 94: 1023-1030
Lee M, Godshalk EB, Lamkey KR, Woodman WW, 1989. Association of restriction fragment length polymorphisms among maize inbreds with agronomic performance of their crosses. Crop Sci 29: 1067-1071
Legesse BW, Myburg AA, Pixley KV, Twumasi-Afriyie S, Botha AM, 2008.
Relationship between hybrid performance and AFLP based genetic distance in highland maize inbred lines. Euphytica 162: 313-323
Liu K, Muse SV, 2005. POWERMARKER: An integrated analysis environment for genetic marker analysis. Bioinformatics 21: 2128-2129
Mace ES, Buhariwalla HK, Crouch JH, 2003. A high-throughput DNA extraction protocol for tropical molecular breeding programs. Plant Molecular Biology Reporter 21: 459a-459h
Melchinger AE, Lee M, Lamkey KR, Woodman WL,

1990. Genetic diversity for restriction fragment length polymorphisms: relation to genetic effects in maize inbreds. Crop Sci 30: 1033-1040

Melchinger AE, Boppenmaier J, Dhillon WG, Pollmer WG, Herrmann RG, 1992.

Genetic diversity for RFLP in European maize inbreds: II Relation to performance of

hybrids within vs. between heterotic groups for forage traits. Theoretical and Applied Genetics 84: 672-681

Melchinger AE, 1999. Genetic diversity and heterosis. In: Coors JG, Pandey S, (Eds.).

The genetics and exploitation of heterosis in crops. ASA,CSSA, and SSSA, Madison, WI, pp. 99-118

Miranda GV, Souza LV, Galvao JCC, Guimaraes LJM, de Melo AV, dos Santos IC, 2008. Genetic variability and heterotic groups of Brazilian popcorn populations. Euphytica 162: 431-440

Mladenovic-Drinic S, Trifunovic S, Drinic G, Konstantinov K, 2002. Genetic divergence and its correlation to heterosis in maize as revealed by SSR-based markers. Maydica 47: 1-8

Mohammadi SA, Prasanna BM, 2003. Analysis of genetic diversity in crop plants a salient statistical tools and considerations. Crop Sci 43: 1235-1248

Moll RH, Lonnquist JH, Fortuna JV, Johnson EC, 1965. The relationship of heterosis and genetic divergence in maize. Genetics 52: 139-144

Munhoz REF, Prioli AJ, Amaral Junior AT, Scapim CA, Simon GA, 2009. Genetic distances between popcorn populations based on molecular markers and correlations with heterosis estimates made by diallel analysis of hybrids. Genetics and Molecular Research 8: 951-962

Nasu S, Suzuki J, Ohta R, Hasegawa K, Yui R, Kitazawa N, Monna L, Minobe Y,

2002. Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa*, *Oryza rufipogon*) and establishment of SNP markers. DNA Research 9: 163-171

Patterson HD, Williams ER, Hunter EA, 1978. Block designs for variety trials. Journal of Agricultural Science (Cambridge) 90: 395-400

Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Castiglioni P, Taramino G, Motto M, 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theor Appl Genet 97: 1248-1255

Phumichai C, Doungchan W, Puddhanon P, Jampatong S, Grudloyma P, Kirdsri C, Chunwongse J, Pulam T, 2008. SSR-based and grain yield-based diversity of hybrid maize in Thailand. Field Crops Res 108: 157-162

Qi X, Kimatu JN, Li Z, Jiang L, Cui Y, Liu B, 2010. Heterotic analysis using AFLP markers reveals moderate correlations between specific combining ability and genetic distance in maize inbred lines. African Journal of Biotechnology 9: 1568-1572

Rafalski A, 2002. Applications of single nucleotide polymorphisms in crop genetics. Current Opinion in Plant Biology 5: 94-100

Reif JC, Melchinger AE, Xia XC, Warburton ML, Hoisington DA, Vasal SK, Beck D,

Bohn M, Frisch M. 2003. Use of SSRs for establishing heterotic groups in subtropical maize. Theor Appl Genet 107: 947-957

Rencher AC, 1995. Methods of multivariate analysis. John Wiley and Sons Inc. pp. 627

Reyes-Valde's MH, 2000. A model for marker-based selection in gene introgression breeding programs. Crop Sci 40: 91-98

Rinaldi AR, Carpentieri-Pipolo V, Gerage AC, Ruas CF, Fonseca Junior NS, de Souza A,

de Souza SGH, Garbuglio DD, 2007. Correlation between heterosis and genetic divergence estimated of diallel crosses and RAPD molecular markers in populations of popcorn. Bragantia 66: 183-192

SAS Institute, 2003. SAS System for Windows. Version 9.1. SAS Inst., Inc. Cary. NC

Semagn K, Bjornstad A, Ndjiondjop MN, 2006. An overview of molecular marker methods for plants. African Journal of Biotechnology 5: 2540-2568

Senior ML, Murphy JP, Goodman MM, Stuber CW, 1998. Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. Crop Sci 38: 1088-1098

Silva TA, Pinto RJB, Scapim CA, Mangolin CA, da Silva Machado MFP, Carvalho MSN, 2009. Genetic divergence in popcorn genotypes using microsatellites in bulk genomic DNA. Crop Breeding and Applied Biotechnology 9: 31-36

Smith OS, Smith JSC, Bowen SL, Tenborg RA, Wall SJ, 1990. Similarities among a group of elite maize inbreds as measured by pedigree, F1 grain yield, grain yield, heterosis, and RFLPs. Theor Appl Genet 80: 833-840

Smith JSC, Smith OS, 1992. Fingerprinting crop varieties. Review of Agriculture Academy Press, USA. pp. 85-140

Smith JSC, Chin ECL, Shu H, Smith OS, Wall SJ, Senior ML, Mitchell SE, Kresovich S, Ziegle J, 1997. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. Theor Appl Genet 95: 163-173

Westman AL, Kresovich S, 1997. Use of molecular marker techniques for description of plant genetic

variation. In: Callow JL, Ford-Lloyd BV, Newburry HJ eds. Biotechnology and plant genetic resources. CAB International. pp. 9-45

Wright S, 1978. Evolution and genetics of populations. Vol. IV. The University of Chicago Press, Chicago

Xia XC, Reif JC, Hoisington DA, Melchinger AE, Frisch M, Warburton M, 2004.

Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers: I. Lowland tropical maize. Crop Sci 44: 2230-2237

Xia XC, Reif JC, Melchinger AE, Frisch M, Hoisington DA, Beck D, Pixley KV, Warburton ML, 2005. Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers: II. Subtropical, tropical midaltitude, and highland maize inbred lines and their relationships with elite U.S. and European maize. Crop Sci 45: 2573-2582

XLSTAT, 2010. XLSTAT (Win) 2010 by Addinsoft. (http://www.xlstat.com) [last accessed 10 Jan. 2011]

Yan J, Shah T, Warburton ML, Buckler ES, McMullen MD, Crouch J, 2009. Genetic characterisation and linkage disequilibrium estimation of a global maize collection using SNP markers. PLoS ONE 4(12): e8451. Doi: 10.1371/journal.pone.0008451. pp. 1-14