Euphytica (2017) 213:220
DOI 10.1007/s10681-017-2018-x

Use of genomic and phenotypic selection in lines for prediction of testcrosses in maize I: grain yield and its components

Gustavo Vitti Môro · Mateus Figueiredo Santos · Cláudio Lopes de Souza Júnior

Received: 26 December 2016 / Accepted: 30 August 2017 / Published online: 2 September 2017

Abstract Grain yield (GY) is a direct function of its components and these traits to being less complex and highly correlated with yield. The objectives of this study were to map Quantitative Trait Loci (QTL) for GY and its components in maize lines and GY in their testcrosses, to verify its congruence and the possibility to select testcrosses from the predict means of the lines by using markers information. Two hundred and fifty-six S_1 lines derived from the cross L-08-05F × L-14-04B of tropical germplasm and the testcrosses of these lines with two testers were evaluated in six environments. The traits analysed in the lines were GY, prolificacy, ear height and diameter, number of rows per ear and kernels per row, kernel depth, grain weight, and GY in the testcrosses. QTL were mapped in the lines and in testcrosses and the predicted means of the lines were computed based on QTL effects and in all markers of the genome. Few QTL detected for GY and its components in the lines were coincident with the QTL for yield in testcrosses. The correlations between the predicted means of the lines and the phenotypic means of the testcrosses were not significant or low for most of the components. The coincidence of the selected lines and testcrosses was very low for all traits and the results were similar for both testcrosses and intensity. It is not possible to select testcrosses by using GY or its components information from the lines, even with the aid of molecular markers.

Keywords Correlation · Hybridization · Endogamy · Indirect selection · Marker assisted selection · Yield

Electronic supplementary material The online version of this article (doi:10.1007/s10681-017-2018-x) contains supplementary material, which is available to authorized users.

G. V. Môro (✉)
Department of Plant Production, School of Agricultural and Veterinarian Sciences, São Paulo State University (Unesp), Jaboticabal, São Paulo 14884-900, Brazil
e-mail: gvmoro@fcav.unesp.br

M. F. Santos
Embrapa Beef Cattle, Plant Production Group, Campo Grande, Mato Grosso do Sul 79106-550, Brazil
e-mail: mateus.santos@embrapa.br

C. L. de Souza Júnior
Department of Genetics, University of São Paulo, P.O. Box 83, Piracicaba, São Paulo 13.400-970, Brazil
e-mail: clsouza@usp.br

Introduction

The inbreeding–hybridization system, idealized by Shull (1910), East (1908) and Jones (1918), remains the most important breeding scheme for the commercial production of maize hybrids. In this system, the lines are evaluated in testcrosses, using a tester from a germplasm different from the one used to obtain the lines. Information on lines that may indicate the performance of their testcrosses are desirable because the crossing between the lines and testers, and conducting experiments to evaluate the testcrosses

are expensive and require a great amount of time and human resources (Mihaljevic et al. 2005).

Maize grain yield is the main trait of interest to breeders and farmers. This trait has a quantitative architecture, controlled by a large number of loci (Jugenheimer 1976) that have a small individual effect on the phenotype (Geldermann 1975) and low heritability coefficient estimates (Robinson et al. 1949; Hallauer and Miranda Filho 1988; Malvar et al. 1996; Austin and Lee 1998). Grain yield is a direct function of its components (Jugenheimer 1976), which are the prolificacy or number of ear per plant, average kernel weight, number of rows per ear, number of kernels per row, ear length and diameter and kernel depth. These traits can be used to practice indirect selection for grain yield since they are correlated with yield but less complex than yield itself, less influenced by the environment, and have higher heritability coefficients (Hallauer and Miranda Filho 1988; Arias et al. 1999; Alves et al. 2002).

QTL mapping works are useful for heritability studies on traits of agronomic/economic importance, to understand phenomena such as the correlation between traits and to apply this information to the selection process, known as marker-assisted selection (MAS) (Stuber and Sisco 1992; Berke and Rocheford 1995; Hospital et al. 1997; Bouchez et al. 2002). Meuwissen et al. (2001) proposed a new approach to using the marker-assisted selection, known as Genome-Wide Selection or Genomic Selection (GS). This methodology has been applied intensively in animal breeding, presenting satisfactory results (Kolbehdari et al. 2007; Goddard and Hayes 2007; Legarra and Misztal 2008) but, the majority of studies used mathematical models to evaluate this methodology in plant breeding (Bernardo and Yu 2007; Bernardo 2008, 2009; Liu et al. 2008; Heffner et al. 2009; Massman et al. 2013; Jacobson et al. 2014; Mendes and Souza Júnior 2016).

Few studies have been conducted to verify the coincidence of mapped QTL in lines and their hybrids and/or testcrosses, considering different traits and types of population (Beavis et al. 1994; Mihaljevic et al. 2005). In general, few coincident QTL were found for all traits; however, besides considering the same traits in two generations, they were evaluated in different environments and, therefore, the low QTL congruence may be explained by the high QTL × Environment interaction. Nevertheless, studies on QTL mapping in lines and their testcrosses for grain yield and its components evaluating the two generations in the same environment allow eliminating the QTL × Environment interaction. Therefore, allowing a more accurate comparison of QTL congruence between these two generations while the obtained information can be used to select superior testcrosses based on molecular information of the line components through indirect MAS.

Thus, the objectives of this study were to map QTL for grain yield and its components of S_1 maize lines and testcrosses; to verify the congruence between QTL in these two generations; to estimate correlations between phenotypic and predicted means in lines using molecular markers for grain yield and its components and the phenotypic means of the testcrosses for grain yield; to verify the coincidence of S_1 lines and superior testcrosses selected using molecular markers for grain yield and its components.

Materials and methods

Genetic material

Parental inbred lines L-08-05F and L-14-04B of tropical germplasm were used to develop a reference F_2 population. Both inbreds were developed by the Maize Breeding Program of the Department of Genetics, at the Luiz de Queiroz School of Agriculture, University of São Paulo, Brazil (ESALQ/USP). Inbred L-08-05F, with orange flint kernels, was developed from the population IG-1, and inbred L-14-04B, with yellow dent kernels, was developed from the population BR-106. These populations and their respective inbreds are in different heterotic groups, and they are genetically divergent for several traits (Sibov et al. 2003; Sabadin et al. 2008; Moreira et al. 2009; Môro et al. 2012). The parental inbreds were crossed and three F_1 plants, previously tested against the parental inbreds to check their genetic identity with microsatellite markers, were selfed to develop the F_2 population. Two-hundred and fifty-six plants were randomly taken and selfed to develop $F_{2:3}$ progenies, which corresponded to the S_1 lines used in this work. These were crossed with two tester lines L-04-05F and L-02-03D, also belonging to the Maize Breeding Program of the Department of Genetics of ESALQ/USP. Inbred line L-04-05F was derived from

the IG-1 population, while inbred line L-02-03D was obtained from the IG-2 population. These tester lines belong to different heterotic groups and are divergent for various traits. To obtain the testcrosses, the S_1 lines were used as the female parent (through detasseling) and crossed with each tester line in isolation plots, thereby producing two testcrosses for each of the S_1 lines, totaling 256 testcrosses with L-04-05F (TC1) and 256 testcrosses with L-02-03D (TC2). This study thus used 256 S_1 lines, together with their respective testcrosses.

Experimental procedure

The evaluation of the S_1 lines and the testcrosses was performed in six environments, each combination of location and year being considered as a distinct environment. The 256 lines and their respective testcrosses were evaluated on two experimental stations, Department of Genetics—ESALQ/USP (22°42′S, 47°38′W, 540 m altitude) and Anhembi (22°45′S, 48°00′W, 460 m altitude). The experiments were conducted over three farming years, with two replications per environment; the trials with the S_1 lines and the testcrosses were set up in adjacent areas and were, therefore, in the same environment. For both the trial with the S_1 lines and the two trials with the testcrosses, the experimental design adopted was a simple 16 × 16 lattice. Each plot comprised a 4 m-long row sown with 50 seeds that, after thinning, had 20 plants remaining in each plot. The final spacing between the plants was 0.20 m, while the spacing between the rows was 0.80 m, corresponding to a population of approximately 62,500 plants ha^{-1}.

In the S_1 lines, the following traits were analysed: grain yield (GY, t ha^{-1}), adjusted to 15% of moisture and corrected for average stand; prolificacy (PRO, number of ears $plot^{-1}$), corrected for the average stand; ear length (EL, cm); ear diameter (ED, cm); number of rows per ear (NRE); number of kernels per row (NKR); kernel depth (KD, cm, obtained by [(ED − cob diameter)/2]); and average weight of 500 grains (W500, grams); and, in the testcrosses, grain yield (GY, t ha^{-1}) adjusted to 15% moisture and corrected for average stand.

Analyses of variance and covariance

All analyses of variance and covariance were performed using the PROC GLM procedures of the SAS software package (SAS Institute Inc. 2001). For each environment conducted with the lines and testcrosses, individual analyses of variance were performed, according to the mathematical model for lattice-type experiments (Cochran and Cox 1966). Starting from the adjusted means obtained from these individual analyses, analyses of combined variance of the environment were performed, using a random model. For the testcrosses, in addition to the combined analyses, analyses of grouped combined variance were performed, in which the trials conducted with the two testcrosses were grouped together. Analyses of covariance between the grain yield and its components of the S_1 lines and the grain yield of the testcrosses of each tester were also conducted, using the same procedures as for the analyses of variance, to obtain the coefficients of genetic and phenotypic correlation of the evaluated traits in the lines and testcrosses.

Genetic map

The genetic map used, and the procedures used to develop it, was previously described by Sibov et al. (2003) and Môro et al. (2012). The F_2 plants that gave rise to the $F_{2:3}$ progenies (S_1 lines) were genotyped with microsatellite markers and the genetic map was developed using MAPMAKER/EXP version 3.0b (Lincoln et al. 1992). The map present a total of 177 markers distributed along the 10 linkage groups. The genetic map spanned 2052 cM in length with an average interval of 11.60 cM between adjacent markers.

QTL mapping

The program QTL Cartographer version 1.17 (Basten et al. 2003) was used to perform the QTL mapping, "window size" was set to 10 cM and "walking speed" to 1 cM, and module *Jzmapqtl* was selected. *Jzmapqtl* implements composite interval mapping (CIM) (Zeng 1994) expanded to analysis across multiple environments (Jiang and Zeng 1995). For the mapping in the S_1 lines, the model used was $y_{jm} = b_{0m} + b_m^* x_j^*$

$+ d_m^* z_j^* + \sum_l^t (b_{lm} x_{jl} + d_{lm} z_{jl}) + e_{jm}$ and, in testcrosses, the model was similar but the dominance effects are absent, were y_{jm} is the phenotypic value of the jth genotype evaluated in the mth environment; b_{0m} is the mean effect for environment m; b_m^* corresponds the additive effect of the probable QTL applicable to environment m for the lines and the allelic substitution effect of the probable QTL applicable to environment m for the testcrosses; x_j^* is the identifying variable of the genotype of the probable QTL which takes the values 0, 1 and 2 for the genotypes qq, Qq and QQ, respectively, according to probabilities that depend on the recombination fraction between marker i and the QTL, conditional on the genotypes of the flanking markers i and $i + 1$; d_m^* is the dominance effect of the probable QTL applicable to environment m for the lines; z_j^* is the identifying variable of the genotype of the probable QTL, which assumes values of 0 for the homozygote genotypes (qq or QQ) and 1 of a heterozygote genotype (Qq), according to probabilities that depend on the recombination fraction between marker i and the QTL, conditional on the genotypes of the flanking markers i and $i + 1$; b_{lm} is the partial regression coefficient between the phenotypic values and the values attributed to x_{jl}; x_{jl} are the identifying variables associated with cofactor l and t is the number of markers selected as cofactors; d_{lm} is the partial regression coefficient between the phenotypic values and the values attributed to z_{jl} for the lines; z_{jl} are the identifying variables associated with cofactor l and t is the number of markers selected as cofactors for the lines; and e_{jm} is the residual effect associated with the jth genotype in the mth environment.

The cofactors considered in the analyses were selected per environment, using a "stepwise" (*forward/backward*) regression procedure, with $\alpha = 0.05$ as the critical value for the inclusion or exclusion of a marker in the model. After this first round of cofactor selection, a second selection was applied to the selected cofactors, to leave at most the five most informative cofactors for each environment. This was performed to prevent overparameterization of the model, which can introduce biases in the estimates obtained (Basten et al. 2003).

The critical values for the tests to establish the presence of a QTL and a QTL × Environment interaction given by the number of independent tests, according Vieira et al. (2000). For the present study, these critical values were 25.3 and 21.4, which correspond to LOD scores of 5.5 and 4.7, respectively, larger than those used in most studies in maize (Mihaljevic et al. 2005; Lima et al. 2006).

Congruence of mapped QTL in lines and testcrosses

To identify positional congruence between the QTL detected in the S_1 lines and those in their testcrosses, estimates were made of the confidence intervals for each mapped QTL, expressed as a "one-LOD support interval" (Lander and Botstein 1989). The confidence intervals of the detected QTL in congruent regions were compared in the lines and testcrosses, and the overlapping of these intervals was considered a probable common QTL.

Prediction of line means

Means were predicted for the S_1 lines based on the effects of the QTL. Neglecting epistatic effects between the QTL, these means ($\bar{Y}_{QS_1}$) were obtained from the equation $\hat{\mathbf{y}} = \hat{\boldsymbol{\mu}}\mathbf{1} + \mathbf{X}\hat{\boldsymbol{\beta}}$, in which $\hat{\mathbf{y}}$ is the vector holding the predicted means for the S_1 lines; $\hat{\boldsymbol{\mu}}$ is the overall mean of the S_1 lines; $\mathbf{1}$ is a vector of ones; $\mathbf{X}$ is the matrix of the genetic predictors of the additive and dominance effects of the mapped QTL, with dimensions $N_{S1} \times (2 \times N_{QTL})$, where $N_{S1} = 256$ and N_{QTL} is the number of mapped QTL; and $\hat{\boldsymbol{\beta}}$ is the vector of the genetic values of the mapped QTL, that is, the additive and dominance effects of the mapped QTL. The predictors of the additive effects (a) were obtained from the difference the conditional probabilities of the QTL would present for the genotypes QQ and qq, given the genotypes of the flanking markers of this QTL, while the predictors of the dominance effects (d) were obtained from the conditional probabilities the QTL would display for the genotype Qq, given the genotypes of the flanking markers, that is Predictor (a) = P(QQ|Mi_Mj_) − P(qq|Mi_Mj_) and Predictor (d) = P(Qq|Mi_Mj_). The probabilities were obtained using the R/QTL package of the software R.

The genotypic values (Best Linear Unbiased Predictor—BLUPs) of the molecular markers were obtained using the mixed model $\mathbf{y} = \mu\mathbf{1} + \mathbf{Xg} + \mathbf{e}$, in

which: $\mathbf{y}$ is the vector of dimension $N_{S1} \times 1$ ($N_{S1} = 256$) holding the phenotypic means of the S_1 lines; $\boldsymbol{\mu}$ is the overall phenotypic mean of the S_1 lines; $\mathbf{1}$ is the vector of ones of dimension $N_{S1} \times 1$ which relates the mean to the vector $\mathbf{y}$; $\mathbf{X}$ is the matrix of dimensions $N_{S1} \times N_m$ ($N_m = 177$), whose elements are defined as follows: 1 if the S_1 plant is homozygote for the marker originated from the L14-04B line, -1 if homozygote for the L08-05F originated marker, and 0 if heterozygote; $\mathbf{g}$ is the $N_M \times 1$ vector of the BLUPs of the genetic values of the markers (GVM/BLUP), treated as random effects; and $\mathbf{e}$ is the $N_M \times 1$ vector of the residuals.

The genetic values of the markers were estimated using a mixed-models methodology (Henderson 1984). The BLUPs of the genetic values were obtained assuming that the variance from each marker is equal to σ^2_G/N_m, where σ^2_G is the genetic variance of the S_1 lines and N_m the number of markers (Meuwissen et al. 2001). Starting from the genotypes of the markers of the S_1 lines and the estimates of their genetic values, predictions were obtained of the means for the lines based upon all the markers of the genome ($\bar{Y}_{GS_1}$), through the equation $\hat{\mathbf{y}} = \hat{\boldsymbol{\mu}}\mathbf{1} + \mathbf{X}\hat{\mathbf{g}}$. In this equation, the parameters correspond to those previously described, and have the same dimensions, but with the following differences: the vector $\hat{\mathbf{y}}$ holds the estimates of the predicted means of the S_1 lines; $\hat{\boldsymbol{\mu}}$ is the estimate of the overall mean of the S_1 lines; and $\hat{\mathbf{g}}$ is the vector holding the estimates of the BLUPs of the genetic values of the markers.

Correlations and coincidences for selected superior testcrosses

Correlation coefficients between the phenotypic and predicted means of the lines and the phenotypic means of the testcrosses for the two testers were estimated. Lists of selected genotypes were then generated for the S_1 lines, based in turn on the phenotypic means, the means predicted from the mapped QTL ($\bar{Y}_{QS_1}$), and the means predicted from the genetic values ($\bar{Y}_{GS_1}$). Selection lists were also developed for the two testcrosses, based on their phenotypic means. Selection was then applied to these lists with selection intensities of 10% (best 26 genotypes) and 20% (best 52 genotypes). The number of superior genotype coincidences for the lines and the testcrosses in these selections were then counted.

Results

Analyses of variance and covariance

Highly significant differences ($p \leq 0.01$) were detected for lines and line × environment interactions for all analysed traits. Highly significant differences were also detected for the testcrosses, but no significance was detected for their interaction with the environment for all traits in both testcrosses (data not shown). For grain yield (GY), line means differed from both testcrosses means and, among these, no significant difference was observed between TC1 and TC2. However, GY had a greater variation interval in the lines compared to the testcrosses while TC2 had a greater variation interval between the two testcrosses. Although the experimental coefficient of variation for GY was greater in the lines than in the testcrosses, probably due to their higher average, for the yield components in the lines, the values were reduced (Table 1).

The components of variance estimates for line traits were all positive and significant ($p \leq 0.05$) and, for the testcrosses, only the Testcrosses × Environment interaction of grain yield for the two testers was null. The magnitudes of the genetic variances of GY between testcrosses were statistically different ($p \leq 0.05$), and TC2 had the highest estimate. The heritability coefficients based on means were significantly different from zero ($p \leq 0.05$) for all traits in lines and testcross yields whereas the highest estimates were obtained for line yield and its components and, the lowest for the grain yield of testcrosses. The GY estimates of 0.90, 0.68, and 0.72 were obtained for S_1, TC1, and TC2, respectively. The estimates of yield components of the lines ranged between 0.83 (PRO) and 0.91 (ED). Furthermore, the heritabilities of the two testcrosses were not significantly different (Table 1).

The genetic correlation coefficient of GY in the lines with TC1 was not significant, but it was significant with TC2, displaying the value of 0.35. For the yield components of the lines and the GY of the testcrosses, these estimates ranged from -0.17 (W500) in TC1 to 0.38 (PRO) in TC2. The phenotypic

Table 1 Mean values, range of variation (R), experimental coefficients of variation (CV %), genetic variances ($\hat{\sigma}^2_G$), interaction genotype × environment variances ($\hat{\sigma}^2_{G\times E}$), heritability coefficients ($\hat{h}^2_{\bar{x}}$), with respective confidence intervals[a] (between brackets), for grain yield and yield components of the S_1 lines and grain yield of the testcrosses (TC)

Generation[b]	Trait[c]	Mean (R)	CV %	$\hat{\sigma}^2_G$	$\hat{\sigma}^2_{G\times E}$	$\hat{h}^2_{\bar{x}}$
S_1	GY	5.27 (2.01; 10.67)	19.91	1.73 [1.44; 2.12]	0.69 [0.61; 0.79]	0.90 [0.88; 0.92]
	EL	14.60 (12.30; 16.83)	5.59	0.61 [0.50; 0.76]	0.25 [0.20; 0.33]	0.85 [0.81; 0.87]
	ED[d]	3.89 (3.41; 4.45)	4.11	0.41 [0.34; 0.50]	0.14 [0.12; 0.16]	0.91 [0.89; 0.92]
	NRE	11.77 (9.11; 15.34)	5.61	0.95 [0.80; 1.16]	0.19 [0.15; 0.23]	0.93 [0.91; 0.94]
	NKR	29.85 (22.37; 36.25)	7.95	6.00 [4.98; 7.46]	2.85 [2.42; 3.42]	0.86 [0.84; 0.89]
	KD[d]	0.83 (0.65; 1.00)	6.87	0.35 [0.29; 0.43]	0.16 [0.13; 0.19]	0.87 [0.84; 0.89]
	W500	137.23 (102.48; 167.45)	8.28	131.16 [108.38; 162.73]	59.92 [49.92; 73.52]	0.86 [0.83; 0.88]
	PRO[d]	0.95 (0.51; 1.37)	14.75	0.16 [0.13; 0.20]	0.09 [0.08; 0.11]	0.83 [0.79; 0.86]
TC1	GY	9.81 (8.44; 11.31)	9.93	0.10 [0.08; 0.13]	–	0.68 [0.61; 0.74]
TC2	GY	10.53 (7.33; 12.55)	9.93	0.18 [0.15; 0.24]	–	0.72 [0.66; 0.77]

[a] Confidence intervals at 0.95 probability level

[b] S_1, TC1 and TC2 refers S_1 lines, testcrosses with the L-04-05F line tester and testcrosses with the L-02-03D line tester, respectively

[c] GY (grain yield, tonnes per hectare), EL (ear length, cm), ED (ear diameter, cm), NRE (number of rows per ear), NKR (number of kernels per row), KD (kernel depth, cm), W500 (average weight of 500 grains, grams), PRO (prolificacy, number of ears per plant)

[d] The genetic variances and the interaction genotype × environment were multiplied by 10 for ED and PRO, and by 100, for KD

correlations between GY of the two generations were also not significant for TC1 and, this estimate was 0.29 for TC2. For the yield components, the phenotypic correlations ranged between −0.13 (W500) in TC1 and 0.30 (PRO) in TC2 (Table 2).

QTL mapping and congruence of mapped QTL

Sixteen QTL were mapped to GY in S_1 lines, of which 88% showed interaction with the environment. QTL numbers varied between 9 (NKR) and 23 (PRO) when mapping yield components in the lines whereas the lowest and the highest percentages of QTL with significant interaction with the environment were observed for NRE (25%) and EL (81%), respectively. For grain yield of testcrosses, 16 and 17 QTL were mapped on TC1 and TC2, respectively. The interaction with the environment was significant in 50 and 65% of the QTL mapped on TC1 and TC2, respectively. Only 1 QTL of GY coincided between the lines and TC1 and TC2 and, considering the yield components in the lines, the number of coincident QTL decreased for all of them. For TC1, this number ranged from 1 (NKR and KD) to 6 (PRO) and, for TC2, ranged from 1 (NRE and W500) to 4 (PRO). When considering the two testcrosses simultaneously, the highest number of matched QTL was 1 for EL, ED, NRE, NKR and PRO (Table 3). The comparison between the two testcrosses of the QTL for grain yield showed only four coincident QTL (data not shown).

Table 2 Phenotypic ($\hat{r}_{\bar{F}}$) and genetic ($\hat{r}_G$) correlations between the yield and yield components of S_1 lines and grain yield of the testcrosses (TC)

Trait[a]	GY—TC1 (L-0405F Tester)		GY—TC2 (L-0203D Tester)	
	$\hat{r}_{\bar{F}}$	$\hat{r}_G$	$\hat{r}_{\bar{F}}$	$\hat{r}_G$
GY	0.11^{ns}	0.12^{ns}	0.29**	0.35**
EL	-0.05^{ns}	-0.08^{ns}	0.02^{ns}	0.02^{ns}
ED	0.06^{ns}	0.06^{ns}	0.17**	0.21**
NRE	0.02^{ns}	0.01^{ns}	0.04^{ns}	0.05^{ns}
NKR	0.07^{ns}	0.08^{ns}	0.18**	0.23**
KD	0.14*	0.17**	0.20**	0.25**
W500	−0.13*	−0.17**	0.01^{ns}	0.01^{ns}
PRO	0.21**	0.25**	0.30**	0.38**

[ns],*,** Non-significant, significantly different from zero at 0.05 and significantly different from zero at the 0.01 probability level, respectively

[a] GY (grain yield, tonnes per hectare), EL (ear length, cm), ED (ear diameter, cm), NRE (number of rows per ear), NKR (number of kernels per row), KD (kernel depth, cm), W500 (average weight of 500 grains, grams), PRO (prolificacy, number of ears per plant)

Table 3 Number of QTL mapped and with significant interaction with the environment, percentage and number (between parenthesis) of QTL for grain yield and its components of the S_1 lines coincident with those of the testcrosses (TC)

Generation[a]	Trait[b]	Number		% QTL coincident (number QTL coincident)[a]		
		QTL	QTL × environment	TC1	TC2	TC
S_1	GY	16	14	6.25 (1)	6.25 (1)	0,00 (0)
	EL	16	13	12.50 (2)	12.50 (2)	6,25 (1)
	ED	16	8	12.50 (2)	18.75 (3)	6,25 (1)
	NRE	12	3	16.67 (2)	8.33 (1)	8,33 (1)
	NKR	9	5	11.11 (1)	33.33 (3)	11,11 (1)
	KD	12	6	8.33 (1)	16.67 (2)	0,00 (0)
	W500	18	6	22.22 (4)	5.56 (1)	0,00 (0)
	PRO	23	17	26.09 (6)	17.39 (4)	4,35 (1)
TC1	GY	16	8	–	–	–
TC2	GY	17	11	–	–	–

[a] S_1, TC1, TC2 and TC refer S_1 lines, testcrosses with L-04-05F line tester, testcrosses with L-02-03D line tester and for both line testers, respectively

[b] GY (grain yield, tonnes per hectare), EL (ear length, cm), ED (ear diameter, cm), NRE (number of rows per ear); NKR (number of kernels per row), KD (kernel depth, cm), W500 (average weight of 500 grains, grams), PRO (prolificacy, number of ears per plant)

Correlations between line and testcross means, and coincidences in selection of superior testcrosses

The correlations between the predicted means based on the QTL effects of the lines ($\bar{Y}_{QS_1}$) and the phenotypic means of the testcrosses ($\bar{Y}_{PT}$) for GY were 0.13 and 0.23 for TC1 and TC2, respectively, while for the components, the estimates varied between 0.14 (EL) in TC2 and 0.19 (PRO) in TC1. Considering the predicted line means based on all markers ($\bar{Y}_{GS_1}$) and $\bar{Y}_{PT}$, the correlations for GY were 0.13 and 0.26 for TC1 and TC2, respectively, and for the components, the lowest estimate was 0.15 (KD) in TC1 and, the highest, 0.29 (PRO) in TC2 (Table 4).

The selection results of the testcrosses based on $\bar{Y}_{PS_1}$,$\bar{Y}_{QS_1}$ and $\bar{Y}_{GS_1}$ for grain yield and their components in the lines did not differ from those obtained for the correlations, and had low values in most situations. The 10% selection intensity for GY, resulted in 2, 2 and 3 coincident testcrosses in TC1 and 6, 5 and 6 coincidences in TC2 for $\bar{Y}_{PS_1}$, $\bar{Y}_{QS_1}$ and $\bar{Y}_{GS_1}$, respectively. For yield components, there was no coincidence for NRE considering $\bar{Y}_{QS_1}$ and $\bar{Y}_{GS_1}$, and for NKR considering $\bar{Y}_{GS_1}$, in TC1, while the greatest coincidence occurred for PRO considering $\bar{Y}_{PS_1}$ and $\bar{Y}_{QS_1}$ in TC2. The selection intensity of 20%, for GY, resulted in 12, 13 and 12 coincident testcrosses in TC1, and 15, 12 and 16 in TC2, for $\bar{Y}_{PS_1}$, $\bar{Y}_{QS_1}$ and $\bar{Y}_{GS_1}$, respectively. In the components, coincidences varied from 7 (NRE) in TC1 considering $\bar{Y}_{QS_1}$, up to 17 (PRO) in TC2 considering $\bar{Y}_{QS_1}$ and $\bar{Y}_{GS_1}$ (Table 5).

Discussions

The analysis of variance indicated the existence of genetic variability for lines and testcrosses, as well as differential performance of the lines in different environments. Also, the coefficients of experimental variation showed a good experimental precision in obtaining the data, with values within the expected for the traits and types of evaluated genotypes (Hallauer and Miranda Filho 1988; Leon and Coors 2002; Bento et al. 2003; Lima et al. 2006). For grain yield, the higher genetic variance estimated for the lines compared to the testcrosses was expected since the lines had a greater variation amplitude of the mean. The genetic variance of the testcrosses is a function of the genetic divergence of the population and the allelic substitution effect of the tester on the loci controlling the trait, depending on the degree of dominance (Bernardo 2002). The variation intervals of the means of the two testcrosses was greater in TC2, resulting in

Table 4 Correlation coefficients between the phenotypic means ($\bar{Y}_{PS_1}$), means predicted based on the effects of the QTL ($\bar{Y}_{QS_1}$) and means predicted based on all the markers ($\bar{Y}_{GS_1}$) of grain yields of S_1 lines and their components with the phenotypic means of grain yield of the testcrosses

Testcross[a]	Trait[b]	$\bar{Y}_{PS_1}$	$\bar{Y}_{QS_1}$	$\bar{Y}_{GS_1}$
TC1	GY	0.11^{ns}	0.13*	0.13*
	EL	-0.05^{ns}	0.10^{ns}	-0.03^{ns}
	ED	0.06^{ns}	0.06^{ns}	0.03^{ns}
	NRE	0.02^{ns}	-0.07^{ns}	0.00^{ns}
	NKR	0.07^{ns}	0.00^{ns}	0.06^{ns}
	KD	0.14*	0.09^{ns}	0.15*
	W500	−0.13*	-0.07^{ns}	-0.08^{ns}
	PRO	0.21**	0.19**	0.23**
TC2	GY	0.29**	0.23**	0.26**
	EL	0.02^{ns}	0.14*	0.01^{ns}
	ED	0.17**	0.11^{ns}	0.11^{ns}
	NRE	0.04^{ns}	-0.07^{ns}	0.02^{ns}
	NKR	0.18**	0.10^{ns}	0.18**
	KD	0.20**	0.11^{ns}	0.19**
	W500	0.01^{ns}	-0.09^{ns}	-0.03^{ns}
	PRO	0.30**	0.16**	0.29**

^{ns} ,*,** Non-significant, significantly different from zero at 0.05 probability level and significantly different from zero at 0.01 probability level, respectively

[a] TC1 and TC2 refers testcrosses with L-04-05F line tester and testcrosses with L-02-03D line tester, respectively

[b] GY (grain yield, tonnes per hectare), EL (ear length, cm), ED (ear diameter, cm), NRE (number of rows per ear), NKR (number of kernels per row), KD (kernel depth, cm), W500 (average weight of 500 grains, grams), PRO (prolificacy, number of ears per plant)

greater variability and, also, the L-04-05F tester had a common origin with one of the parents of the population lines, resulting in lower divergence of this tester with the population lines compared to the other tester. The higher heritability coefficients obtained in the lines may result from the higher genetic variability in this generation compared to testcrosses, due to the existence of high inbreeding and a large number of homozygous genotypes in the lines compared to the testcrosses. However, for the yield components, this may be due to the lower complexity and lower environmental influence of these traits compared to yield itself, justifying the use of these components to practice indirect selection for yield (Maita and Coors 1996; Leon and Coors 2002; Bento et al. 2003). The heritabilities of the two testcrosses were not significantly different while for the traits and the number of evaluated environments, the values obtained in this study corroborate those reported by Hallauer and Miranda Filho (1988).

Although some correlation coefficients were significant and even highly significant, the values obtained for these estimates are very low, providing no predictive value for the yield of testcrosses, both from the line yield and its components. Thus, it is not possible to use information on yield components in lines to practice indirect selection for grain yield in their testcrosses, even though indirect selection for grain yield can be practiced efficiently based on information from the yield components for the same generation (Maita and Coors 1996; Leon and Coors 2002; Bento et al. 2003). It is also verified that the tester used influenced these estimates since TC2 had a higher number of significant correlations compared to TC1, although they are so low that, hardly, any response is expected in the yield of testcrosses.

As in other studies in the literature, the present work has mapped a large number of QTL for all traits, and a large number of QTL had significant interaction with the environment (Mihaljevic et al. 2005; Lima et al. 2006). The varying number of detected QTL and the lack of coincidence among them suggest that the population genetic background influences QTL mapping and that the tester influences the mapping analysis directly. Thus, indicating the occurrence of QTL × Tester interaction, probably due to effects of tester specific alleles (Beavis et al. 1994; Groh et al. 1998; Lu et al. 2003; Mihaljevic et al. 2005; Peng et al. 2013).

The low correlation between yield and its components in the lines and their testcrosses can be explained by the low congruence between QTL for yield and the components mapped in the lines with QTL for yield detected in the testcrosses. It is noteworthy that although the correlation was low for PRO in the lines, it was the highest among the components while this trait had the highest QTL congruence with the yields of the testcrosses, suggesting that some of the QTL affecting PRO in the lines also affect grain yield in the testcrosses of these lines. Many congruent QTL were expected to be detected since components and yield are highly correlated (Hallauer and Miranda Filho 1988; Arias et al. 1999; Alves et al. 2002), but these traits were considered in different generations in the

Table 5 Coincidence of superior S_1 lines and testcrosses selected, in percentage and number (in parenthesis), considering selection based on phenotypic means ($\bar{Y}_{PS_1}$), means predicted based on the effects of the QTL ($\bar{Y}_{QS_1}$) and predicted based on all the markers ($\bar{Y}_{GS_1}$) of grain yield of the S_1 lines and their components and the phenotypic mean of grain yield of testcrosses, for 10% (26 selected genotypes) and 20% (52 selected genotypes) selection intensity

Testcross[a]	Trait[b]	IS = 10% (26)			IS = 20% (52)		
		$\bar{Y}_{PS_1}$	$\bar{Y}_{QS_1}$	$\bar{Y}_{GS_1}$	$\bar{Y}_{PS_1}$	$\bar{Y}_{QS_1}$	$\bar{Y}_{GS_1}$
TC1	GY	7.69 (2)	7.69 (2)	11.54 (3)	23.08 (12)	25.00 (13)	23.08 (12)
	EL	7.69 (2)	11.54 (3)	7.69 (2)	17.31 (9)	30.77 (16)	19.23 (10)
	ED	15.38 (4)	7.69 (2)	3.85 (1)	21.15 (11)	21.15 (11)	19.23 (10)
	NRE	11.54 (3)	0.00 (0)	0.00 (0)	25.00 (13)	13.46 (7)	17.31 (9)
	NKR	15.38 (4)	0.00 (0)	15.38 (4)	21.15 (11)	25.00 (13)	26.92 (14)
	KD	19.23 (5)	19.23 (5)	15.38 (4)	23.08 (12)	25.00 (13)	23.08 (12)
	W500	7.69 (2)	15.38 (4)	7.69 (2)	15.38 (8)	23.08 (12)	17.31 (9)
	PRO	7.69 (2)	11.54 (3)	3.85 (1)	23.08 (12)	28.85 (15)	25.00 (13)
TC2	GY	23.08 (6)	19.23 (5)	23.08 (6)	28.85 (15)	23.08 (12)	30.77 (16)
	EL	19.23 (5)	23.08 (6)	19.23 (5)	21.15 (11)	23.08 (12)	17.31 (9)
	ED	15.38 (4)	7.69 (2)	15.38 (4)	15.38 (8)	25.00 (13)	21.15 (11)
	NRE	19.23 (5)	11.54 (3)	15.38 (4)	21.15 (11)	23.08 (12)	19.23 (10)
	NKR	19.23 (5)	15.38 (4)	15.38 (4)	21.15 (11)	26.92 (14)	23.08 (12)
	KD	11.54 (3)	11.54 (3)	15.38 (4)	21.15 (11)	17.31 (9)	21.15 (11)
	W500	11.54 (3)	11.54 (3)	11.54 (3)	19.23 (10)	15.38 (8)	17.31 (9)
	PRO	26.92 (7)	26.92 (7)	19.23 (5)	30.77 (16)	32.69 (17)	32.69 (17)

[a] TC1 and TC2 refers to testcrosses with L-04-05F line tester and testcrosses with L-02-03D line tester, respectively

[b] GY (grain yield, tonnes per hectare), EL (ear length, cm), DE (ear diameter, cm), NRE (number of rows per ear), NKR (number of kernels per row), KD (kernel depth, cm), W500 (average weight of 500 grains, grams), PRO (prolificacy, number of ears per plant)

present study. Therefore, the low QTL congruence observed may have resulted from the dominance or epistatic effects of the QTL and the effects of the specific alleles of the testers. Although there are no studies available in the literature investigating the same traits for these two generations, a few studies on the same traits in the lines and their crosses obtained similar results, with correlation coefficients ranging from low to average. However, few QTL were coincident between the two generations (Beavis et al. 1994; Groh et al. 1998; Austin et al. 2000; Mihaljevic et al. 2005; Peng et al. 2013).

As observed for phenotypic means, when using molecular markers information, most of the correlation coefficients were not significant and, when the estimates were significant, the values were very low. Nevertheless, when the correlations between the predicted means in the lines and testcross phenotypic means were significant, the values based on QTL information/effects were, generally, lower, than those based on all markers. This result suggests that, possibly, some loci affecting the traits were not detected in the QTL mapping, confirming the higher efficiency for predicting means based on information from all markers of the genome, instead of using only those markers associated with the trait through the QTL analysis. These results are consistent with the literature (Meuwissen et al. 2001; Bernardo and Yu 2007; Lorenzana and Bernardo 2009; Massman et al. 2013; Jacobson et al. 2014), which shows that predicted means based on all markers are more accurate than those obtained using QTL information since the values are closer to the phenotypic means. Although the used testers were of different origins, no significant effect was observed on the correlation coefficients, for means predicted based on both QTL and all markers. Even though indirect selection can bring satisfactory results in the same generation (Maita and Coors 1996; Leon and Coors 2002; Bento et al. 2003), it is not possible to predict the yield of a

testcross from yield or yield components of the lines used to obtain the testcross even when using information os molecular markers.

The selection of higher yield testcrosses based on yield and yield components of the parental lines is inefficient, regardless of the information used, phenotypic or molecular marker, since the number of S_1 lines and coincident testcrosses was very low for all studied traits. As observed for the correlation coefficients, there was no marked effect of the tester on the coincidence of selected testcrosses, that is, the coincidence was low in the two testcrosses. For the two selection intensities and all the traits studied, the coincidence of selected testcrosses was very low and, the comparison of these results with the obtained correlation coefficients seem to suggest that this coincidence was random in most cases. Comparing the three types of selection, it can be verified that, for the two intensities of selection considered and in both testcrosses, the number of coincident superior testcrosses did not differ markedly for the phenotypic data, QTL and all markers of lines, and varied according to the trait considered in the lines. Therefore, it is not possible to use Genomic Selection to select more productive testers based on information on yield or yield components in the lines even though molecular markers are an excellent tool for genetic breeding programs in some situations (Bernardo 2009; Massman et al. 2013; Mendes and Souza Júnior 2016). Furthermore, when marker information is used to perform indirect assisted selection in testcrosses, the information on the markers must be obtained, necessarily, in testcross populations and not in line populations.

Concluding remarks

The results show that, in the population used in this research, derived of parents L-08-05F × L-14-04B of tropical germplasm, the correlation between grain yield and its components in the lines and their testcrosses are not significant for most of the components and when significant, their value is very small, independent of the tester used. For these traits, the coincident QTL mapped in the lines and testcrosses are also very low and, therefore, most of the QTL responsible for the expression of line yield and its components do not affect the yield of their testcrosses, probably because the genetic background of the population influences QTL expression. Due to the homozygosity of the lines, which is not normal in maize, the line phenotype can be "masked" by one or a few pairs of deleterious genes with homozygotes and, by crossing these homozygous lines in testcrosses, the heterozygosity of these loci is restored and the natural condition for this species, thus eliminating the expression of the most deleterious genes in homozygosity. Therefore, it is concluded that it is not possible to perform selection for testcross yield from the line yield and, also, indirect selection for the testcross yield from the information on the line components, even when using molecular marker information.

Acknowledgements This research was supported by "Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-140964/2006-1)" and by the Department of Genetics of the Agriculture College "Luiz de Queiroz"-University of São Paulo. C. L. Souza Jr. and G. V. Môro are recipient of a research fellowship from CNPq. The authors are grateful to Dr. Anete Pereira de Souza, from the University of Campinas, for the genetic mapping of the population, and to A. J. Desidério, A. S. Oliveira, C. R. Segatelli, and for their assistance with the field experiments.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Alves GF, Ramalho MAP, Souza JC (2002) Alterações nas propriedades genéticas da população CMS-39 submetida à seleção massal para a prolificidade. Revista Brasileira de Milho e Sorgo 1:89–101

Arias CAA, de Souza Júnior CL, Takeda C (1999) Path coefficient analyses of ear weight in different types of progeny in maize. Maydica 44:251–262

Austin DF, Lee M (1998) Detection of quantitative loci for grain yield and yield components in maize across generations in stress and nonstress environments. Crop Sci 38:1296–1308

Austin DF, Lee M, Veldboom LR, Hallauer AR (2000) Genetic mapping in maize with hybrid progeny across testers and generations: grain yield and grain moisture. Crop Sci 40:30–39

Basten CJ, Weir BS, Zeng Z-B (2003) QTL cartographer: version 1.17

Beavis WD, Smith OS, Grant D, Fincher RR (1994) Identification of quantitative trait loci using a small sample of topcrossed and F4 progeny from maize. Crop Sci 34:882–896

Bento DAV, Ramalho MAP, Souza JC (2003) Seleção massal para prolificidade em milho na época normal e na "safrinha". Revista Brasileira de Milho e Sorgo 2:78–87

Berke TG, Rocheford TR (1995) Quantitative trait loci for flowering, plant and ear height, and kernel traits in maize. Crop Sci 35:1542–1549

Bernardo R (2002) Breeding for quantitative traits in plants. Stemma Press, Woodbury

Bernardo R (2008) Molecular markers and selection for complex traits in plants: learning from the last 20 years. Crop Sci 48:1649–1664

Bernardo R (2009) Genome wide selection for rapid introgression of exotic germplasm in maize. Crop Sci 49:419–425

Bernardo R, Yu J (2007) Prospects for genome wide selection for quantitative traits in maize. Crop Sci 47:1082–1090

Bouchez A, Hospital F, Causse M, Gallais A, Charcosset A (2002) Marker-assisted introgression of favorable alelles at quantitative trait loci between maize elite lines. Genetics 162:1945–1959

Cochran WG, Cox GM (1966) Experimental design. Wiley, New York

East EM (1908) Inbreeding in corn. Conn Agric Expt Sta Rept for 1907:419–428

Geldermann H (1975) Investigations of inheritance of quantitative characters in animals by gene markers. I. Methods. Theor Appl Genet 46:319–330

Goddard ME, Hayes BJ (2007) Genomic selection. J Anim Breed Genet 124:323–330

Groh S, Khairallah MM, González-de-Leon D, Willcox M, Jiang C, Hoisington DA, Melchinger EH (1998) Comparison of QTL mapped in RILs and their test-cross progenies of tropical maize for insect resistance and agronomic traits. Plant Breed 117:193–202

Hallauer AR, Miranda Filho JB (1988) Quantitative genetics in maize breeding. Iowa State University Press, Ames

Heffner EL, Sorrells ME, Jannink JL (2009) Genomic selection for crop improvement. Crop Sci 49:1–12

Henderson CR (1984) Applications of linear models in animal breeding. University of Guelph, Ontario

Hospital F, Moreau L, Charcosset A, Gallais A (1997) More the efficiency of marker assisted selection. Theor Appl Genet 95:1181–1189

Jacobson A, Lian L, Zhong S, Bernardo R (2014) General combining ability model for genomewide selection in a biparental cross. Crop Sci 54:895–905

Jiang C, Zeng Z (1995) Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics 140:1111–1127

Jones DF (1918) The effects of inbreeding and crossbreeding upon development. Conn Agric Expt Sta Bull 207:5–100

Jugenheimer RW (1976) Corn improvement, seed production and uses. Wiley, New York

Kolbehdari D, Schaeffer LR, Robinson JAB (2007) Estimation of genome-wide haplotype effect in half-sib designs. J Anim Breed Genet 124:356–361

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Legarra A, Misztal I (2008) Computing strategies in genome-wide selection. J Dairy Sci 91:360–366

Leon N, Coors JG (2002) Twenty-four cycles of mass selection for prolificacy in the Golden Glow maize population. Crop Sci 42:325–333

Lima MLA, Souza CL Jr, Vieira DA, Souza PH, Garcia LC (2006) Mapping QTL for grain yield and plant traits in a tropical maize population. Mol Breed 17:227–239

Lincoln SE, Daly MJ, Lander ES (1992) Constructing genetic maps with Mapmaker Exp 3.0. Whitehead Institute for Biometrical Research, Cambridge

Liu X, Fu J, Gu D, Liu W, Liu T, Peng Y, Wang J, Wang G (2008) Genome-wide analysis of gene expression profiles during the kernel development of maize (*Zea mays* L.). Genomics 91:378–387

Lorenzana RE, Bernardo R (2009) Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor Appl Genet 120:151–161

Lu H, Romero-Severson J, Bernardo R (2003) Genetic basis of heterosis explored by simple sequence repeat markers in a random-mated maize population. Theor Appl Genet 107:494–502

Maita R, Coors JG (1996) Twenty cycles of biparental mass selection for prolificacy in the open pollinated maize population Golden-Glow. Crop Sci 36:1527–1532

Malvar RA, Ordás A, Revilla P, Cartea ME (1996) Estimates of genetic variance in two Spanish populations of maize. Crop Sci 36:291–295

Massman JM, Jung H-JG, Bernardo R (2013) Genomewide selection versus marker-assisted recurrent selection to improve grain yield and stover-quality traits for cellulosic ethanol in maize. Crop Sci 53:58–66

Mendes MP, Souza Júnior CL (2016) Genomewide prediction of tropical maize single-crosses. Euphytica 209:651–663

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829

Mihaljevic R, Schon CC, Utz HF, Melchinger EH (2005) Correlations and QTL correspondence between line per se and testcross performance for agronomic traits in four populations of European maize. Crop Sci 45:114–122

Moreira JUV, Bento DAV, Souza PH, Souza CL Jr (2009) QTL mapping for reaction to Phaeosphaeria leaf spot in a tropical maize population. Theor Appl Genet 119:1361–1369

Môro GV, Santos MF, Bento DAV, Aguiar AM, Souza CL Jr (2012) Genetic analysis of kernel oil content in tropical maize with design III and QTL mapping. Euphytica 185:419–428

Peng B, Li Y, Wang Y, Liu C, Liu Z, Zhang Y, Tan W, Wang D, Shi Y, Sun B, Song Y, Wang T, Li Y (2013) Correlations and comparisons of quantitative trait loci with family per se and testcross performance for grain yield and related traits in maize. Theor Appl Genet 126:773–789

Robinson HF, Comstock RE, Harvey PH (1949) Estimates of heritability and the degree of dominance in corn. Agron J 41:353–359

Sabadin PK, Souza CL Jr, Souza PH, Garcia AAF (2008) QTL mapping for yield components in a tropical maize population using microsatellite markers. Hereditas 145:194–203

SAS Institute Inc (2001) SAS/STAT User's guide, v.6.03. SAS Institute, Cary

Shull GH (1910) Hybridization methods in corn breeding. Am Breed Mag 6:63–72

Sibov ST, Souza CL Jr, Garcia AAF, Silva AR, Mangolin CA, Benchimol LL, Souza PH (2003) Molecular mapping in tropical maize using microsatellite markers. 2. Quantitative trait loci (QTL) for grain yield, ear height, and grain moisture. Hereditas 139:107–115

Stuber CW, Sisco P (1992) Marker-facilitated transfer of QTL alleles between inbred lines and responses in hybrids. In: Proceedings of 46th Ann Corn Sorghum Res. Conference. ASTA, Washington pp 104–113

Vieira C, Pasyukova EG, Zeng ZB, Hackette JB, Lyman RF, Mackay TFC (2000) Genotype-environment interaction for quantitative trait loci affecting life span in *Drosophila melanogaster*. Genetics 154:213–227

Zeng Z-B (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1466