

Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize

G. Blanc · A. Charcosset · B. Mangin ·
A. Gallais · L. Moreau

Received: 19 December 2005 / Accepted: 31 March 2006 / Published online: 20 May 2006
© Springer-Verlag 2006

Abstract Quantitative trait loci (QTL) detection experiments have often been restricted to large bi-allelic populations. Use of connected multiparental crosses has been proposed to increase the genetic variability addressed and to test for epistatic interactions between QTL and the genetic background. We present here the results of a QTL detection performed on six connected F_2 populations of 150 $F_{2,3}$ families each, derived from four maize inbreds and evaluated for three traits of agronomic interest. The QTL detection was carried out by composite interval mapping on each population separately, then on the global design either by taking into account the connections between populations or not. Epistatic interactions between loci and with the genetic background were tested. Taking into account the connections between populations increased the number of QTL detected and the accuracy of QTL position estimates. We detected many epistatic interactions, particularly for grain yield QTL (R^2 increase of 9.6%). Use of connections for the QTL detection also allowed a global ranking of alleles at each QTL. Allelic relationships and epistasis both contribute to the lack of consistency for QTL positions observed among populations, in

addition to the limited power of the tests. The potential benefit of assembling favorable alleles by marker-assisted selection are discussed.

Introduction

Detection of quantitative trait loci (QTL) for agronomic traits has received considerable attention in plant genetics since the late 1980s, to understand the genetic basis of the traits variation and subsequently perform marker-assisted selection (MAS). To perform QTL detection, the general practice has been to develop a few specific biparental mapping populations of large size, in order to guarantee sufficient power of the tests. Analyzing these large specific populations individually has clearly been successful in detecting QTL in plants (Kearsey and Farquhar 1998; Asins 2002; Bernardo 2002) and some QTL could be cloned, in particular not only in rice and tomato (Takahashi et al. 2001; Kojima et al. 2002; Liu et al. 2002, 2003) but also in maize (Doebley et al. 1997; Salvi et al. 2005). However, one can wonder which fraction of the variability available for breeding has been analyzed so far. Indeed, when several populations could be analyzed for the same trait in a given species, inconsistency in QTL positions were generally reported (Beavis et al. 1991; Mihaljevic et al. 2004). A meta-analysis of flowering time and related traits in maize from 55 QTL detection studies concluded that a total of 62 different QTL are likely involved in the variation of these traits, whereas on average 4 to 5 QTL were detected in single-population analyses (Chardon et al. 2004). This suggests that, for complex traits involving several tens

Communicated by J.-L. Jannink

G. Blanc · A. Charcosset · A. Gallais · L. Moreau (✉)
INRA/INA-PG/UPS/CNRS, UMR de Génétique Végétale,
Ferme du Moulon, 91190 Gif sur Yvette, France
e-mail: moreau@moulon.inra.fr

B. Mangin
INRA, Unité de Biométrie et Intelligence Artificielle,
BP 52627, 31326 Castanet Tolosan Cedex, France

of QTL, the genetic diversity in most studies is narrow when compared to that available within the species of interest. There is therefore a growing consensus on the fact that QTL studies should now address diversity more globally. This can be of great interest for applied purposes and for more fundamental research, to understand the nature of allelic variation at QTL.

Considering several populations derived from diverse parental materials increases the probability that a QTL will be polymorphic in at least one population. To go beyond comparison of results between populations, some authors have proposed analyzing jointly the different populations. This can be done first for independent populations (no known pedigree relationship between the parents of the different populations) (Muranty 1996; Xu 1998). In this case, QTL effects are nested (in the statistical sense) within populations and the number of parameters to be estimated increases with the number of populations. Also, the lack of connections between populations does not allow a global comparison of the effects of all QTL alleles segregating in the different populations. An alternative approach is therefore to develop connected populations (common parents among populations). Under the assumption of additivity, considering identical allelic effects over populations rather than nesting effects within populations reduces the total number of parameters and, consequently increases the power of QTL detection (Rebai and Goffinet 1993; Jannink and Jansen 2001). In such an analysis, the effects of alleles segregating are estimated simultaneously, which facilitates a global comparison. This is of particular interest to identify the parental origin(s) of favorable allele(s) at each QTL.

A further interest of connected designs is their potential to address epistatic interactions between QTL and the genetic background, provided the mating design contains “loops” (in the simplest case, three populations derived from three parents $A \times B$, $B \times C$ and $A \times C$). In such designs, epistasis can be tested through the comparison between (1) a “connected additive” model where the allele effects at a QTL are assumed to be identical in the different populations and (2) a “hierarchical” model where allele effects are nested within populations, which accounts for possible interactions with the genetic background. Such an analysis tests for consistency of allelic effects over populations and therefore permits to evaluate the contribution of QTL-by-genetic-background epistatic effects to variation in QTL results observed among populations, relative to that of other factors such as allelic relationships between parental inbreds and statistical noise. Tests for epistasis in connected designs following

this principle have been proposed by several authors (Rebai et al. 1994; Charcosset et al. 1994; Jannink and Jansen 2000, 2001). One of the advantages of these tests, when compared to testing only for digenic interactions, is to enable the detection of epistatic interactions of higher order (Charcosset et al. 1994). The statistical properties of QTL-by-genetic-background interaction tests in the case of a single digenic interaction has been analyzed by means of simulations (Jannink and Jansen 2001). This study showed that it was possible to identify the two QTL involved by using an appropriate statistical test and also proposed guidelines for the interpretation of the sign of the QTL-by-genetic-background interaction effects. For more complex situations, the results are less predictable. Several digenic epistatic interactions that involve a given QTL may add up if similar in sign, yielding a significant interaction with the genetic background whereas none of them were significant. They may also cancel out each other if opposite in signs and lead to no detectable interaction with the genetic background. It is therefore interesting to compare both types of interactions.

Despite these potentialities, very few experimental studies involving connected populations in plants have been reported. Using an interval mapping method based on regression (described by Rebai et al. 1994), significant epistatic effects for silking date QTL in three connected maize recombinant inbred lines were detected (Charcosset et al. 1994). Using the same approach in a diallel among four maize inbreds, silking date QTL were detected but little evidence of interaction with the genetic background was found (Rebai et al. 1997). The lack of power for detecting epistasis might be attributed to the trait considered, indeed silking date in maize is generally assumed to be more additively inherited than more complex traits such as grain yield, and also to the method used, which was simple interval mapping. Composite interval mapping (Zeng 1994), which uses cofactors in QTL detection, increases the power of detection as well as the precision of estimates of QTL position and effects (Jansen 1993; Zeng 1994). To our knowledge only one study (Charcosset et al. 2000) has been published using a composite interval approach in multipopulation designs. But this study was focusing on dominance effects and did not consider epistatic interactions.

We present here the results of a QTL detection carried out for traits of agronomic interest in six connected F_2 populations of maize using the MCQTL software (Jourjon et al. 2005). This software permits the joint analysis of multiple populations using a composite interval mapping method based on a linearized

regression model (Haley and Knott 1992; Charcosset et al. 2000). Our main objective was to compare experimentally the power of different models of QTL detection and to look for epistasis. To do so, we first detected QTL in each population independently, second on the whole design without taking into account connections, then on the global design using connections. Lastly, we tested for digenic interactions and for locus-by-genetic-background interactions, estimated the contribution of epistasis to the variation of the traits studied and checked if epistatic interactions could explain discrepancies among analyses. The joint estimation of the different parental allele effects in a connected model allowed us to identify, for each QTL, the parental inbred line(s) that carried the most interesting allele(s). Based on these results, we will discuss the benefit of assembling favorable alleles coming from the different parental lines, in a more diverse genetic context than simple biparental populations that have been considered so far for MAS in crop breeding.

Materials and methods

Plant material and experimental design

Six F_2 populations, with 150 individuals each, were obtained from a diallel cross between four unrelated maize inbreds, *DE*, *F283*, *F810* and *F9005*. *F9005* is a dent-flint inbred whereas the other three parents are flint inbreds. Each F_2 plant was selfed to obtain a $F_{2,3}$ family. Testcross progenies were produced in isolation plots by crossing the 900 $F_{2,3}$ families and the 4 parental inbreds to a dent inbred tester, *MBS847*. The testcross progenies were evaluated in a total of ten field trials carried out during 2 years (2000 and 2001) at five different locations in Northern France (Dreux, Gif sur Yvette, Lusignan, Mons and Rennes). The testcrosses were grown in two-row plots of 5 m length, spaced 80 cm apart and at plant population densities ranging from 96,000 to 120,000 plants ha^{-1} , depending on the location. In each trial, only 768 testcrosses were evaluated. But considering the whole experimental design, each testcross was evaluated in at least three locations. Within each trial, most testcrosses were not replicated. To assess the precision of the trials, 13 randomly chosen testcrosses were replicated twice and the testcrosses of each parental inbred were replicated between eight and nine times. All the testcross progenies, replicated or not, were arranged in a randomized incomplete block design.

Three agronomic traits were measured. Silking date was the number of days after 1 January when half of

the plants in a plot exhibited silks. Grain moisture at harvest (%) and grain yield adjusted to zero percent grain moisture (tons per hectare, $t\ ha^{-1}$) were measured. These three traits are of major importance in maize breeding. Silking date is known to be involved in the adaptation of maize to environmental conditions. As the interest of breeders is to increase grain yield while keeping low drying costs at harvest, an economic index is commonly considered for variety registration in France. This index, further referred to as index and equal to $10 \times (\text{grain yield}) - 2.5 \times (\text{grain moisture})$, was also included in the analyses.

Genetic map

Genomic DNA was extracted from the four parental inbreds and from a bulk of 20 F_3 plants derived from each of the 900 F_2 . We used 269 microsatellites markers available from Maize GDB (<http://www.maizegdb.org>). Electrophoresis was performed on 4% Metaphor agarose gels. Markers that deviated significantly from expected ratios, based on a Chi-square test at $\alpha=1\%$ were not used. Genetic maps were build using MAPMAKER software (Lander et al. 1987) with the Haldane mapping function (Haldane 1919) and a LOD threshold of 3.0. A linkage map was first built for each population. Then, a consensus map was constructed by considering non-segregating markers in a given population as missing data.

Agronomic data analyses

Plots with less than 70% of the expected stand were excluded from the analysis. Analyses of variance were performed for each trial using SAS GLM procedure (SAS 1990). Models including a random block effect or row and column effect were applied to the field plot performances. When significant, these effects were used to adjust the field data (Moreau et al. 1999). An analysis of variance, with a trial effect, a genotypic effect and the interaction effect between both factors, was then performed on these adjusted data, for each of the six populations. These results were used to estimate broad sense heritability (h^2) within each population, along with a 95% confidence interval (CI) (Knapp et al. 1985). A clustering analysis (Moreau et al. 2004) was performed in order to potentially classify the trials into groups of environments. However, as no clear structure among the trials was found, we calculated adjusted means for each genotype using all trials. Phenotypic correlations between all traits were calculated on these adjusted means over all populations using SAS CORR procedure (SAS 1989b).

QTL detection

Quantitative trait loci detection was conducted on the adjusted (for environmental effects) means of testcross families. We did not consider QTL detection for individual environments because precision of performance in individual trials was limited. Also, the clustering approach did not reveal any clear structure among the trials that would have allowed us to define groups of environments to work with. QTL analysis was conducted using MCQTL software (Jourjon et al. 2005), which performs composite interval mapping (CIM) in bi- or multiparental populations, using a linear regression model (Haley and Knott 1992; Charcosset et al. 2000). Several genetic models described below were considered. Since phenotypic evaluation was performed on testcross progenies, statistical additive effects described below include dominance effects between parental alleles and those of the tester (noted T). Indeed, in an F₂ population derived from lines A and B, the testcross progenies of homozygous plants AA (or BB) at a QTL are heterozygous AT (or BT), and the testcross progenies of heterozygous plants, AB, are either AT or BT in equal proportion. So the average testcross value of AB genotypes is intermediate between those of AA and BB genotypes. Dominance effect between parental alleles A and B cannot be tested and the estimated additive QTL effect which corresponds to half the difference between performances of AT and BT progenies is affected by dominance effects of allele T over alleles A and B.

First, we performed an analysis in each individual population (further referred to as single-population analyses) using the following model:

$$y_{pi} = m_p + \alpha_p^q x_{pi}^q + \sum_{c \neq q} \alpha_p^c x_{pi}^c + \varepsilon_{pi} \tag{model 1}$$

where y_{pi} was the adjusted mean performance value of individual i in population p ; m_p was the mean of population p ; α_p^q (α_p^c) was the estimated substitution effect of one parental allele (B) by the other parental allele (A) of population p at the QTL q (or cofactor c); x_{pi}^q (x_{pi}^c) was the expected number of the parental allele (A) given the genotypes at flanking markers; ε_{pi} was the residual error.

In matrix notations, this model can be written as:

$$\mathbf{Y}_p = m_p \mathbf{J}_p + \mathbf{X}_{qp} \mathbf{A}_{qp} + \sum_{c \neq q} \mathbf{X}_{cp} \mathbf{A}_{cp} + \varepsilon \tag{model 1'}$$

where \mathbf{Y}_p was an $I \times 1$ column vector of performances of the I individuals of the population p ; m_p was the

mean of population p ; \mathbf{J}_p was a $I \times 1$ column vector of ones; $\mathbf{X}_{qp}(\mathbf{X}_{cp})$ was a $I \times 2$ matrix containing the expected number (ranging from 0 to 2) of each parental allele of population p at QTL q (cofactor c) given the marker data for each individual i ; \mathbf{A}_{qp} (\mathbf{A}_{cp}) was a 2×1 column vector of the additive effects associated with each parental allele at QTL q (cofactor c) in population p ; and ε was the vector of the residuals of the model. The additive effects were estimated so that their sum equaled zero for each QTL or cofactor.

Second, we detected QTL on the global design by considering the populations as independent (further referred to as multipopulation disconnected analyses), the allelic effects being nested within populations:

$$\mathbf{Y} = \mathbf{J}\mathbf{M} + \mathbf{X}_q \mathbf{A}_q + \sum_{c \neq q} \mathbf{X}_c \mathbf{A}_c + \varepsilon \tag{model 2}$$

where \mathbf{Y} was a $N \times 1$ column vector of performances of N individuals ($N = 900$) coming from P populations ($P = 6$ in our case); \mathbf{J} was a the $N \times P$ matrix whose elements were 0 or 1 according to whether or not individual i belonged to population p and \mathbf{M} was a $P \times I$ vector of population-specific means, m_p . $\mathbf{X}_q(\mathbf{X}_c)$ was an $N \times K$ matrix containing the expected number (ranging from 0 to 2) of allele k at QTL q (cofactor c) given the marker data for each individual i . The total number of allele effects estimated in the global design was $K = 2P$ ($K = 12$ in our case). By definition, on a given line of $\mathbf{X}_q(\mathbf{X}_c)$, only the two elements corresponding to alleles segregating within the population of individual i can be non-null, and their sum equals 2. $\mathbf{A}_q(\mathbf{A}_c)$ was a $K \times 1$ column vector of the within population allele additive effects at QTL q (cofactor c), the sum of the additive effects of the two alleles segregating in a given population was constrained to be zero. ε was the vector of the residuals.

Third, we detected QTL on the complete design while accounting for relationships between parental inbreds (further referred to as multipopulation connected analyses) using the following model:

$$\mathbf{Y} = \mathbf{J}\mathbf{M} + \mathbf{X}_q^* \mathbf{A}_q^* + \sum_{c \neq q} \mathbf{X}_c^* \mathbf{A}_c^* + \varepsilon \tag{model 3}$$

where \mathbf{Y} , \mathbf{J} , \mathbf{M} and ε were as described in model (2). \mathbf{X}_q^* (\mathbf{X}_c^*) was a $N \times K^*$ matrix containing the expected number of parental allele k^* at the QTL q (cofactor c) given the marker data for each individual i , $K^* = 4$ being the number of parental inbreds; \mathbf{A}_q^* (\mathbf{A}_c^*) was a $K^* \times 1$ column vector of the additive effects associated with parental allele k^* at QTL q (cofactor c). In comparison with model (2) this model assumes that the

allelic effects are the same whatever the population considered. The K^* additive effects were estimated so that their sum equaled zero for each QTL q (or cofactors c). Difference in effects among pairs of alleles was tested a posteriori using a t test ($\alpha=5\%$).

Genotypic probabilities used in the models described above were computed every 2 cM, taking into account information from neighbor markers. Cofactor selection and test of QTL effects were performed with F tests. F thresholds were determined by 1,000 permutation tests, to correspond to a global type I risk of 10% (across all populations and total genome). This F threshold was equal to 4.11 in multipopulation disconnected analyses (model (2)) and to 5.96 in multipopulation connected analyses (model (3)). To compare monopopulation analyses (model(1)) to multipopulation ones, using the same genome-wide significance level of 10% over the six populations, we determined the significance level per population by applying the Bonferroni correction. This led us to apply a per population genome-wide level risk of 1.74%. The corresponding F threshold, obtained by means of permutations for each population, was on average equal to 14.93.

The choice of cofactors was based on an iterative process (Charcosset et al. 2000). In the first step, cofactors were selected from all the markers, on a per chromosome basis by backward selection. The cofactors were then used for a CIM analysis on each chromosome to determine the most likely positions of the QTL. On chromosomes where more than one cofactor were found, separate scans were performed to determine the most likely position of each QTL. All the cofactors identified in the first step were taken into account in the model, except the one corresponding to the QTL under study. The analysis stopped in a window of 10 cM around the other cofactors detected on this chromosome. The new QTL positions detected on the different chromosomes were then used as cofactors in a next iteration and a new CIM analysis was performed. The iterative process stopped when the model converged. For the multipopulation disconnected analysis, the markers taken as cofactors for the first step were chosen through a forward selection procedure. It was not possible to apply a backward selection procedure because the large number of parameters during the early steps of the multiple regression procedure created singularities in matrix \mathbf{X}_c . For sake of simplicity we used the same threshold for the selection of markers taken as cofactors in the first step as for the QTL detection, whereas it is often advocated to release on the threshold for choosing cofactors to increase the power of QTL detection. Using the same threshold in

the initial step of detection might have slightly lowered the power of QTL detection.

F values at final QTL position(s) were converted into the corresponding LOD values (Haley and Knott 1992), in order to estimate QTL CI(s) on the basis of a 1.5 LOD unit fall. This led to larger CI than the usually considered 1 LOD unit fall but this value seems to be appropriate for a 95% confidence rate for an F_2 population (van Ooijen 1992; Lynch and Walsh 1998). The allelic effects at each QTL, as well as the phenotypic variance explained, either by each QTL (individual r^2) or by all the detected QTL (R^2), were estimated.

After performing QTL detection, we compared the different analyses in terms of number of QTL detected, size of CI and R^2 . We also sought for congruency of QTL among analyses. For this, two QTL with overlapping CI were considered as corresponding to a same QTL.

Epistasis test

QTL-by-genetic-background interactions

Multipopulation connected analyses (model (3)), assume that one allele has the same effect over populations, whereas multipopulation disconnected analyses (model (2)) account for possible epistatic interactions between QTL and the genetic background. These interactions were tested using the following model:

$$\mathbf{Y} = \mathbf{JM} + \mathbf{X}_q \mathbf{A}_q + \sum_{c \neq q} \mathbf{X}_c^* \mathbf{A}_c^* + \varepsilon \quad (\text{model } 4)$$

where \mathbf{Y} , \mathbf{J} , \mathbf{M} , \mathbf{X}_c^* and \mathbf{A}_c^* were as defined in model (3); \mathbf{X}_q and \mathbf{A}_q were as defined in model (2); and ε was the vector of the residuals. In this model the estimated additive effects of a QTL are nested within populations (using the same constraint as in model (2)). The QTL-by-genetic-background interaction sum of squares, calculated as the difference between the residual sum of squares in model (3) ($\text{RSS}_{(3)}$) and in model (4) ($\text{RSS}_{(4)}$), has $P-(K^*-1)$ degrees of freedom (df). Models (3) and (4), allowed us to perform an F test for QTL-by-genetic-background interaction:

$$F = \frac{(\text{RSS}_{(3)} - \text{RSS}_{(4)})/[P - (K^* - 1)]}{\text{RSS}_{(4)}/[N - 2P - (K^* - 1)(C - 1)]} \quad (5)$$

where N , P , K^* and C were as defined before. The model used for this test corresponded to the final model (3) reached after convergence, with final estimated positions of QTL. The same positions were used in model (4). Note that this test (5) follows an F

distribution under the hypothesis that the estimated QTL positions are the true ones.

This test for epistasis can be interpreted as a test for consistency of QTL effects estimated in the different populations. Suppose that, at a given QTL, α_{DE} is the additive effect of allele *DE* and α_{F283} is the additive effect of allele *F283*. The contrast between these two alleles can be directly estimated by analyzing the population *DE* × *F283*. This contrast will be further noted $(\alpha_{DE} - \alpha_{F283})_{DE \times F283}$. If epistasis is absent, each contrast estimated in one population can be written as a linear combination of two or more contrasts estimated in other populations. For instance, $(\alpha_{DE} - \alpha_{F283})_{DE \times F283}$ can be written as the linear combination (d_1) of two other contrasts estimated in related populations i.e. $(\alpha_{DE} - \alpha_{F283})_{DE \times F283} = (\alpha_{DE} - \alpha_{F810})_{DE \times F810} + (\alpha_{F810} - \alpha_{F283})_{F283 \times F810}$. Among all the linear combinations between individual population contrasts that can be defined, only three of them are independent (3 being the number of df associated with the QTL-by-genetic-background interaction test). For instance, we can define four different linear combinations of three population contrasts, all equal to zero when epistasis is absent:

$$\begin{aligned}
 d_1 &= (\alpha_{DE} - \alpha_{F283})_{DE \times F283} \\
 &\quad - [(\alpha_{DE} - \alpha_{F810})_{DE \times F810} \\
 &\quad \quad + (\alpha_{F810} - \alpha_{F283})_{F283 \times F810}]; \\
 d_2 &= (\alpha_{DE} - \alpha_{F283})_{DE \times F283} \\
 &\quad - [(\alpha_{DE} - \alpha_{F9005})_{DE \times F9005} \\
 &\quad \quad + (\alpha_{F9005} - \alpha_{F283})_{F283 \times F9005}]; \\
 d_3 &= (\alpha_{DE} - \alpha_{F810})_{DE \times F810} \\
 &\quad - [(\alpha_{DE} - \alpha_{F9005})_{F9005 \times DE} \\
 &\quad \quad + (\alpha_{F9005} - \alpha_{F810})_{F9005 \times F810}]; \\
 d_4 &= (\alpha_{F9005} - \alpha_{F283})_{F9005 \times F283} \\
 &\quad - [(\alpha_{F9005} - \alpha_{F810})_{F9005 \times F810} \\
 &\quad \quad + (\alpha_{F810} - \alpha_{F283})_{F283 \times F810}].
 \end{aligned}$$

Only three of them are independent as $d_4 = d_1 + d_3 - d_2$.

QTL-by-QTL interactions

Digenic epistasis between the two detected QTL was tested by comparing model (3) to the model:

$$\mathbf{Y} = \mathbf{J}\mathbf{M} + \mathbf{X}_q^* \mathbf{A}_q^* + \mathbf{X}_{q'}^* \mathbf{A}_{q'}^* + \mathbf{X}_{qq'}^* \mathbf{A}_{qq'}^* + \sum_{c \neq q, q'} \mathbf{X}_c^* \mathbf{A}_c^* + \varepsilon \tag{model 6}$$

where elements indexed with q (q') corresponded to the first (second) locus involved in the interaction; $\mathbf{X}_{qq'}^*$

was a $N \times (K^*)^2$ matrix equal to the horizontal direct product of each column of \mathbf{X}_q^* by each column of $\mathbf{X}_{q'}^*$; $\mathbf{A}_{qq'}^*$ was a $(K^*)^3 \times 1$ vector of the effect of the interaction between QTL q and QTL q' ; and the other parameters were as defined in model (3). The interaction has $(K^* - 1)^2$ df. We were able to calculate the right probability matrix $\mathbf{X}_{qq'}^*$ as there was at least one polymorphic marker between QTL q and q' in each population, otherwise the expected number of alleles at both QTL positions would have been correlated. We used the same constraints as in model (3) to make the allelic QTL (or cofactors) main effects estimable. The interaction effects between two loci were estimated so that the sum of the interaction effects between a given allele at one QTL and the K^* alleles at the other QTL equaled zero. This led to include eight different constraints (two QTL times K^* alleles), but only seven of them were independent.

Model combining both type of epistasis

To globally quantify the importance of epistatic interactions for the detected QTL, we used a “combined model” that includes epistatic terms of models (4) and (6):

$$\mathbf{Y} = \mathbf{J}\mathbf{M} + \sum_b \mathbf{X}_b \mathbf{A}_b + \sum_{q, q'} [\mathbf{X}_q^* \mathbf{A}_q^* + \mathbf{X}_{q'}^* \mathbf{A}_{q'}^* + \mathbf{X}_{qq'}^* \mathbf{A}_{qq'}^*] + \sum_{k \neq q, q', b} \mathbf{X}_k^* \mathbf{A}_k^* + \varepsilon \tag{model 7}$$

where all parameters were as defined in models (4) and (6). Elements indexed with b represented QTL interacting with the genetic background; those indexed with q, q' represent QTL involved in QTL-by-QTL interactions; and all other QTL were indexed with k . QTL involved in QTL-by-QTL or QTL-by-genetic-background interactions correspond to those detected with model (3). The epistatic interactions included in this model were selected by a backward selection procedure, among interactions detected individually as significant. The R^2 of this model was compared with R^2 of model (3).

Whole genome scan of epistatic interactions

In addition to the tests presented above, we tested all possible marker-by-marker interactions and marker-by-genetic-background interactions to look for chromosome regions involved in epistatic effects and that may have not been detected based on their individual

effect. The models used were similar to models (3), (4) and (6), except that the indices q and q' corresponded to the marker(s) under study. The cofactors in these models were all the QTL detected in multipopulation connected analysis (model (3)) except those QTL whose CI included the position(s) of the marker(s) under study (to avoid model over-parameterization).

Implementation of epistasis tests and investigation of the proportion of significant tests

As tests of epistasis were not yet implemented in MCQTL software, we developed programs with the SAS/IML (SAS 1989a) language using the genotypic probabilities at QTL (or markers) positions computed by MCQTL to build the incidence matrices of the models. The QTL-by-genetic-background as well as the QTL-by-QTL interactions effects were tested with an F test, considering an individual risk level of 5%. This corresponded to an approximate F threshold of 2.6 and 1.9 for each type of test, respectively. Given the very high number of tests for marker-by-marker and marker-by-genetic-background interactions (about 35,000 for marker-by-marker interactions), we applied the False Discovery Rate (FDR) approach (Storey and Tibshirani 2003). For each trait and each type of epistasis all individual P values were analyzed using the Qvalue software (Storey and Tibshirani 2003) to compute the corresponding q values (i.e. the proportion of false positive tests among all the tests significant at this P value level risk) as well as π_0 , the proportion of true null hypotheses (i.e. no epistasis). We chose to consider a rather stringent FDR value equal to 10% to declare effects as significant at the genome wide level. As the q values computation depends on the observed distribution of the P values, a given FDR may correspond to a different P value level risk for each trait and type of epistasis.

Results

Genetic map

Between 117 and 160 markers (144 on average) were segregating with reliably reading polymorphism in each individual population, yielding a total of 269 mapped markers. Out of these, seven markers were segregating in all the six populations. Markers were most frequently polymorphic in three populations. On average, the map length of individual populations was 1,663 centimorgans (cM) and the mean distance between two markers was 13 cM. The composite map

(Fig. 1) was 1,794 cM long with an average of 7 cM between two markers.

Agronomic results

For silking date, the earliest parental inbred was *DE* and the latest *F810*, with a 3.5 days difference (Table 1). *F810* had the highest grain yield and *F283* had the lowest grain yield, but the difference between these two inbreds was small ($0.45 t ha^{-1}$). *F810* also had the highest grain moisture. This result was consistent with those observed for silking date, as these two traits were positively correlated ($\rho=0.57$). Grain yield was also positively correlated with silking date ($\rho=0.49$) and grain moisture ($\rho=0.41$).

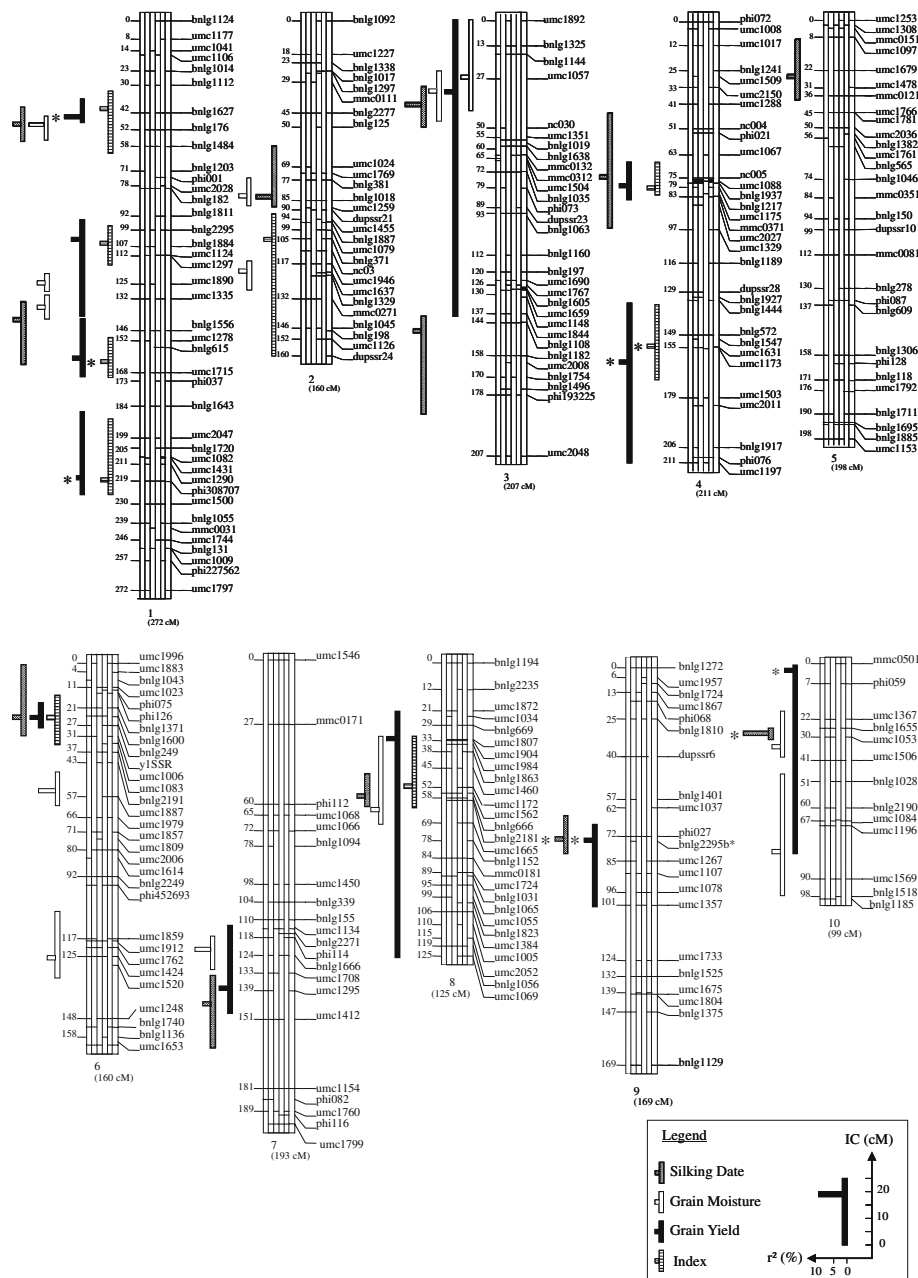
Testcross progenies of the populations derived from *F810* consistently had the latest silking dates and the highest grain yields. Populations with *DE* as a parent flowered earlier. Population testcross means for silking date were intermediate between the parental inbred testcross means, except for the population *F9005* \times *F283* which was earlier than its two parental inbreds. This suggests that silking date mainly behaved additively. This was not true for the other traits. Population testcross means for grain yield were often lower than the average of the two parental means, whereas population means for grain moisture were generally higher than the average of the two parental means. These results suggest the presence of epistasis for these traits (see Discussion), although the differences could also be due to uncontrolled maternal effects during testcross seed production.

Genotype \times environment interactions were found significant for all the traits (results not shown). However, the clustering approach did not reveal any clear structure among the trials. Across populations, broad sense heritabilities (h^2) ranged from 0.63 to 0.81 for grain yield, from 0.72 to 0.87 for grain moisture, and from 0.63 to 0.80 for silking date (Table 1). These values were high, due to the number of trials performed. Silking date and grain moisture were the most heritable traits on average. Populations derived from *F283*, which was one of the earliest parental inbred, had the highest h^2 for silking date. Two populations, both having *F810* as a parent, had a high h^2 for grain yield: *DE* \times *F810* (0.78) and *F283* \times *F810* (0.80). This last result could be related to the fact that *F283* and *F810* had the most divergent parental means for this trait.

QTL detection and comparison across analyses

A total of 37 QTL were detected for the four traits in single-population analyses (Table 2). Zero to three

Fig. 1 Consensus map and QTL detected for silking date, grain moisture, grain yield and index in the multipopulation connected analyses (model (3)). Each vertical box for a chromosome represents one population, from left to right, *DE* × *F283*, *DE* × *F810*, *F283* × *F810*, *F9005* × *DE*, *F9005* × *F283* and *F9005* × *F810*. A horizontal line in the box shows a polymorphic marker in the corresponding population. The QTL are represented in a different bearing for each trait on the left of each chromosome. The vertical line represents the confidence interval (CI) and the horizontal one is placed at the estimated position of the QTL. Its length is proportional to the individual r^2 . A star on the left of a QTL means this QTL interacted with the genetic background at the 5% significance level



(generally either one or two) QTL were detected within a population for a given trait (Table 2). These QTL explained on average (over populations) 28.9, 29.6 and 25.9% of the phenotypic variation of grain moisture, silking date and grain yield (Table 3), respectively. The average CI of QTL positions were rather large, varying between 30 and 49 cM for the different traits. Several QTL found in different populations for the same trait had overlapping CI (Figs. 2, 3). So considering globally the results found in the different populations, seven regions were detected for grain yield, seven for grain moisture, eight for silking date and five for index (Table 2, Fig. 4). These

numbers corresponded to minimum values. Indeed for a given trait, some QTL were sometimes detected on a same chromosome in different populations but at positions rather distant from each other. In cases where the CI of these QTL were large and overlapped, it was difficult to conclude from single-population analyses whether there were one or more QTL implicated in the variation of the trait. To be conservative and avoid overestimation of the QTL number, we considered in this case a single QTL, even if more might be involved.

A total of 26 QTL were detected with the multipopulation disconnected model (2) (Table 2). The number of QTL detected with model (2) is quite

Table 1 Trait means and heritability (h^2) for the four parental inbreds and the six F_2 populations

Inbred or population	Silking date (days after 1 January)		Grain moisture (%)		Grain yield (t ha ⁻¹)	
	Adjusted means	h^2 (CI)	Adjusted means	h^2 (CI)	Adjusted means	h^2 (CI)
<i>DE</i>	210.54	–	33.33	–	9.42	–
<i>F283</i>	211.75	–	32.87	–	9.20	–
<i>F9005</i>	211.25	–	32.27	–	9.34	–
<i>F810</i>	214.06	–	35.40	–	9.65	–
<i>DE</i> × <i>F283</i>	211.30	0.77 (0.71–0.82)	34.12	0.72 (0.65–0.78)	8.79	0.63 (0.53–0.71)
<i>DE</i> × <i>F810</i>	212.57	0.63 (0.53–0.71)	34.02	0.85 (0.81–0.88)	9.03	0.78 (0.72–0.83)
<i>F283</i> × <i>F810</i>	212.79	0.77 (0.71–0.82)	34.48	0.87 (0.84–0.90)	9.34	0.81 (0.76–0.85)
<i>F9005</i> × <i>DE</i>	211.18	0.73 (0.66–0.79)	33.39	0.84 (0.80–0.88)	8.81	0.70 (0.62–0.77)
<i>F9005</i> × <i>F283</i>	210.66	0.80 (0.75–0.84)	33.27	0.85 (0.81–0.89)	8.68	0.69 (0.61–0.76)
<i>F9005</i> × <i>F810</i>	212.29	0.76 (0.69–0.81)	34.06	0.80 (0.75–0.84)	8.85	0.70 (0.62–0.77)

CI95% confidence interval of heritability

equivalent to the number of regions detected with model (1). Most regions (18) are common between the two analyses (Fig. 4). However, the nine QTL detected in single-population analysis (model(1)) were not detected with model (2), and the seven QTL detected with model (2) were not detected with model (1). Model (2) detected all the QTL that were significant in at least two different populations. CI for QTL detected by model (2) were shorter than or at least equivalent to those of corresponding single-population QTL. For example, the grain moisture QTL detected at position 29.9 cM on chromosome 1 had a 7 cM long CI. This is shorter than the shortest CI of the QTL detected in this region in single-population analyses (15.8 cM long for *DE* × *F810* population; Fig. 2).

When taking into account the connections using model (3), 46 QTL were detected in total. This is nearly twice the number detected with model (2). Eleven QTL were detected for silking date, 12 for grain yield, 13 for grain moisture and 10 for index (Table 2, Fig. 1). Multipopulation connected analyses allowed the detection of 15 additional QTL (Fig. 4) not detected by models (1) or (2). Most of these QTL were detected on chromosomes regions where no QTL had been detected with models (1) or (2). Others

corresponded to chromosome segments where model (3) concluded to two or three different QTL whereas model (1) had detected a QTL in several populations, at different positions and with large overlapping CI, so that we considered a single QTL for the synthesis (Table 3). Only three QTL detected with model (1) were not detected with model (3) (Fig. 4). Two of them, for grain moisture and index, had an *F* test value just below the defined threshold. The average CI length of QTL detected by model (3) ranged from 20 cM (for grain moisture) to 52 cM (for grain yield). On average, the QTL CI estimated with model (3) were shorter than the ones obtained with other models (Table 3). For example, the grain moisture QTL detected at position 55 on chromosome 6 had a CI of 14 cM in the multipopulation connected analysis (model (3)), whereas the shortest CI of the corresponding QTL detected in single-population analyses (model (1)) was equal to 37 cM (Fig. 2). The only exceptions to this general tendency were for grain yield QTL and to a less extent for silking date QTL (Table 3). For grain yield, three QTL specific to the multipopulation connected analysis (model (3)) had very large CI. These large QTL CI were due to very flat *F* curves around the maxima, just above the threshold.

Table 2 Number of quantitative trait loci (QTL) detected for the different traits in single-population and in multipopulation connected or disconnected analyses

Type of analysis	Silking date	Grain moisture	Grain yield	Index	Total
Single populations					
<i>DE</i> × <i>F283</i>	2	2	1	1	6
<i>DE</i> × <i>F810</i>	1	2	3	2	8
<i>F283</i> × <i>F810</i>	2	1	1	0	4
<i>F9005</i> × <i>DE</i>	2	1	1	2	6
<i>F9005</i> × <i>F283</i>	3	3	1	0	7
<i>F9005</i> × <i>F810</i>	3	2	1	0	6
Total	8	7	7	5	27
Multipopulation					
Disconnected	8	8	5	5	26
Connected without epistasis	11	13	12	10	46

Table 3 Comparison of the number of QTL detected, average confidence intervals [CI in centimorgans (cM)], and percentage of variance explained (global R^2) in individual analyses and in multipopulation disconnected or connected analyses

Analyses	Silking date			Grain moisture			Grain yield			Index		
	Number of QTL	CI	R^2	Number of QTL	CI	R^2	Number of QTL	CI	R^2	Number of QTL	CI	R^2
Single-population (model (1))	2.2 ^a (8 ^b)	32	29.6	1.8 ^a (7 ^b)	30	28.9	1.3 ^a (7 ^b)	49	25.9	0.8 ^a (5 ^b)	44	13
Multipopulation disconnected (model (2))	8	25 (23 ^c)	64.3	8	29 (26 ^c)	52.2	5	40 (21 ^c)	18.9	4	31 (39 ^c)	27.7
Multipopulation connected (model (3))	11	28 (16 ^d)	66.0	13	20 (14 ^d)	57.9	12	52 (17 ^d)	46.9	10	33 (23 ^d)	35.6

^aAverage number of QTL detected per population

^bNumber of different regions detected by single-population analyses (model(1))

^cAverage CI of the QTL also detected in the single-population analyses (model(1))

^dAverage CI of the QTL also detected in the multipopulation disconnected analyses (model(2))

When considering only the QTL detected by at least two models, we observed a reduction of the CI length for positions estimated with model (3) compared to those estimated by model (2) and for positions estimated with model (2) when compared to those estimated with model (1).

The QTL detected with model (3) explained 66% of the phenotypic variance for silking date, 46.9% for grain yield, 57.9% for grain moisture and 35.6% for index (Table 3). The QTL for the different traits had generally small effects, with r^2 values of 2–10% for grain moisture, 2–11% for grain yield and 2– 4% for

Fig. 2 Grain moisture QTL on the beginning of chromosome 1 detected with the three different models (*lower part*) and curve of the F value corresponding to the test of interaction between markers and the genetic background (*upper part*) along the chromosome. The QTL are represented in the same way as in Fig. 1

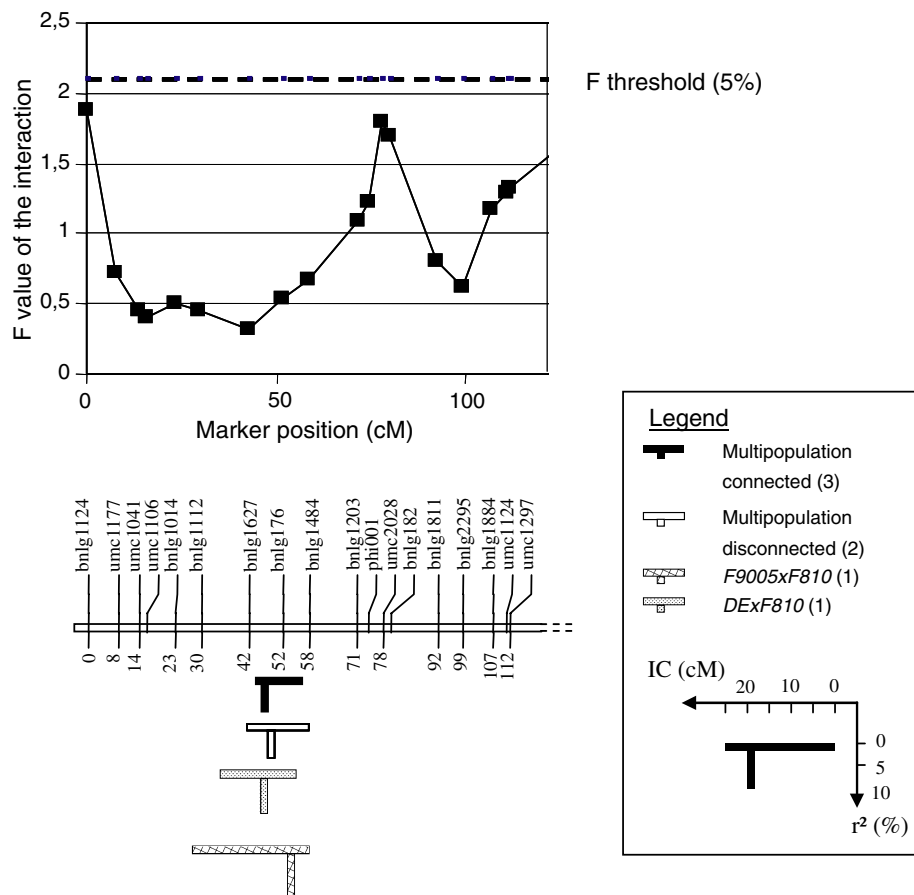
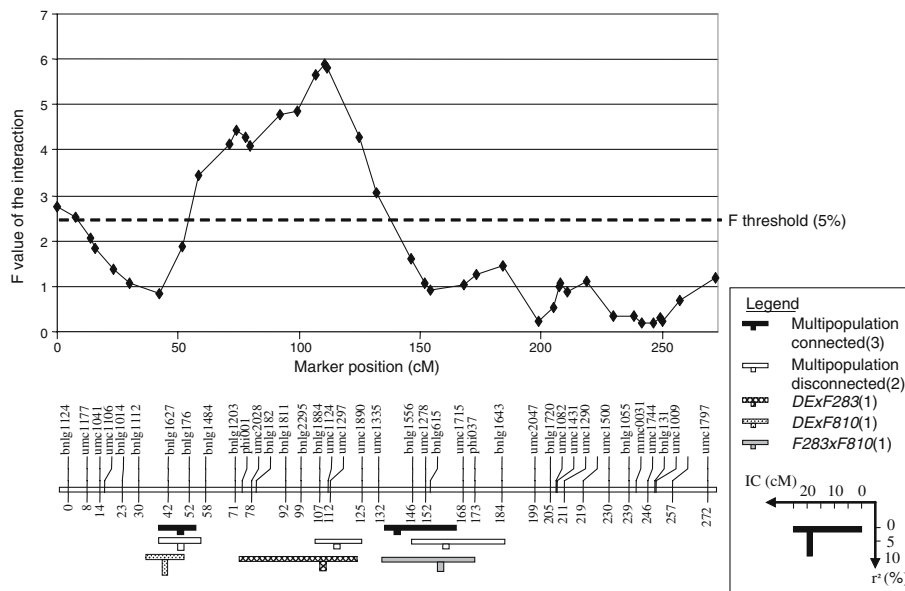


Fig. 3 Silking date QTL detected on chromosome 1 with the three different models (*lower part*) and curve of the *F* value corresponding to the test of interaction between markers and the genetic background (*upper part*) along the chromosome. The QTL are represented in the same way as in Fig. 1

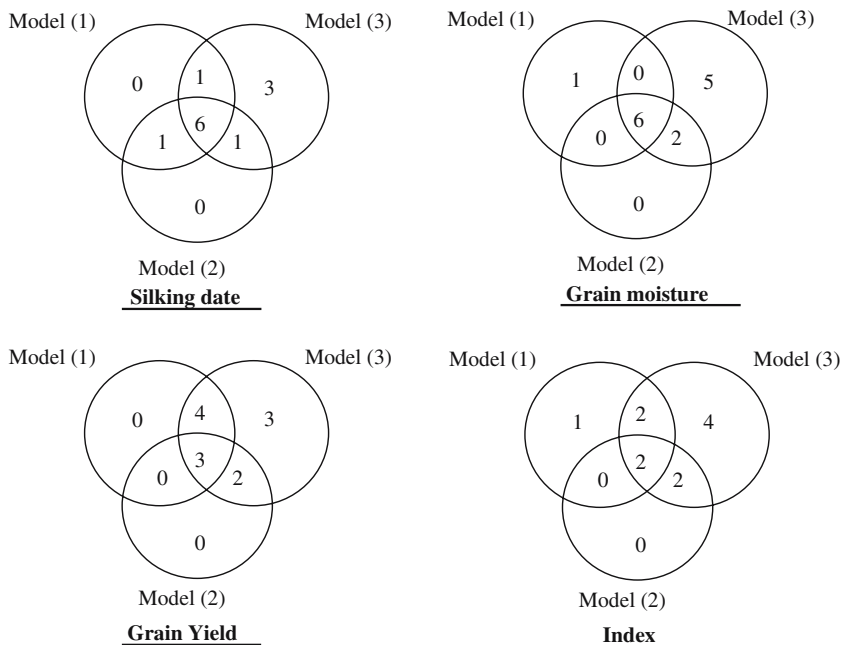


index (Table 4). For silking date, 10 QTL out of the 11 detected had relatively low r^2 of 2–8%, whereas a major QTL was detected on chromosome 10, which explained 18% of the phenotypic variance.

In model (3), the additive effects of the four alleles are estimated simultaneously. For each QTL, tests of the differences between the allele effects allowed us to group them into two to four classes (Table 4). Out of the 46 QTL that were detected, 27 displayed two classes, 18 displayed three classes and only one displayed significant differences between all the four alleles. The major QTL for silking date lo-

cated on chromosome 10 shows two very contrasted allelic classes: a late flowering allele (+ 0.87 days) specific to parental line *F283* and three early flowering alleles with very close effects (–0.25 to –0.34 days). For grain yield, as expected from the parental performances, the favorable alleles were most often contributed by *F810*. For silking date, the early alleles came most often from *DE* or *F283*, as expected from the earliness of these two parental inbreds. It also can be noted that, for all traits, each parental inbred brought both positive and negative allele effects at QTL.

Fig. 4 Number of QTL for each trait detected specifically in single-population analyses (model (1)), multipopulation disconnected analyses (model (2)), multipopulation connected analyses (model (3)), or common to the different models. The number of QTL detected in single-population analyses was determined by considering that two QTL detected in different populations with overlapping CI were the same



As illustrated by Fig. 1, several chromosomal regions displayed significant effects for different traits, suggesting a possible pleiotropic effect of a single QTL. In these cases, additive effects for the different traits were consistent with the correlations observed between

traits. For example, a region on chromosome 1 was involved in the variation of all the traits. In this case, the F810 allele reduced silking date of 0.46 days and grain moisture of 0.34%, but it also decreased grain yield of 0.195 t ha⁻¹ (Table 4).

Table 4 Quantitative trait loci detected in the multipopulation connected design, using model (3), additive effects and epistatic interactions

Trait	QTL number	Chromosome number	Position (in cM)	CI	r^2	Estimated additive effect of allele				Number of significantly different alleles	First-order interaction with QTL (5% significance level)	Genetic-by-background interaction ^c
						DE	F283	F9005	F810			
Silking date	1	1	46	38–56	0.06	0.22 ^a	0.03 ^a	0.21 ^a	-0.46 ^b	2	10, 11	
	2	1	140	134–166	0.06	0.01 ^a	-0.42 ^c	0.06 ^a	0.35 ^b	3	11	
	3	2	85	62–89	0.07	-0.48 ^a	0.19 ^{bc}	0.31 ^c	-0.03 ^b	3	8	
	4	3	41	33–50	0.08	0.27 ^a	-0.51 ^b	-0.19 ^b	0.43 ^a	2	11	
	5	3	150	139–188	0.04	-0.07 ^a	0.11 ^a	-0.3 ^b	0.26 ^c	3	–	
	6	4	75	45–97	0.02	-0.24 ^a	0.09 ^{bc}	0.00 ^b	0.15 ^c	3	–	
	7	5	26	10–38	0.02	0.06 ^a	-0.29 ^b	0.10 ^a	0.12 ^a	2	–	
	8	6	25	2–31	0.04	0.13 ^a	0.08 ^a	-0.37 ^b	0.17 ^a	2	3	
	9	7	145	135–167	0.04	-0.03 ^a	-0.36 ^b	0.15 ^a	0.25 ^c	3	–	
	10	8	58	47–65	0.05	-0.23 ^a	-0.03 ^a	0.40 ^b	-0.15 ^a	2	1	
	11	10	30	28–32	0.18	-0.34 ^a	0.87 ^b	-0.29 ^a	-0.25 ^a	3	1, 2, 4	**
Grain moisture	1	1	46	44–56	0.10	0.13 ^{ab}	0.18 ^a	0.04 ^b	-0.34 ^c	3	–	
	2	1	125	121–127	0.04	-0.44 ^a	0.29 ^b	0.31 ^b	-0.15 ^a	2	–	
	3	1	136	131–142	0.04	0.27 ^a	-0.42 ^b	-0.11 ^b	0.26 ^a	2	–	
	4	2	83	77–87	0.03	-0.18 ^a	0.01 ^{ab}	-0.02 ^b	0.20 ^c	3	11	
	5	2	123	115–129	0.05	0.23 ^a	0.07 ^b	-0.06 ^c	-0.23 ^d	4	–	
	6	3	35	22–47	0.04	0.14 ^a	-0.22 ^b	0.04 ^a	0.04 ^a	2	10, 11	
	7	6	55	47–61	0.07	0.17 ^a	0.14 ^a	-0.20 ^b	-0.11 ^b	2	–	
	8	6	126	107–136	0.05	-0.07 ^a	-0.16 ^a	0.17 ^b	0.05 ^c	3	–	
	9	7	122	118–132	0.09	-0.11 ^a	-0.21 ^a	0.22 ^b	0.10 ^c	3	–	
	10	8	65	31–71	0.07	-0.01 ^a	-0.22 ^b	0.03 ^{ab}	0.20 ^c	3	6	
	11	9	75	64–83	0.06	-0.06 ^a	-0.05 ^a	-0.14 ^a	0.25 ^b	2	4, 6	*
	12	10	34	19–38	0.06	-0.14 ^a	0.27 ^b	-0.01 ^c	-0.13 ^a	3	–	
	13	10	81	46–99	0.02	-0.14 ^a	0.01 ^b	0.03 ^b	0.09 ^b	2	–	
Grain yield	1	1	44	36–48	0.11	0.099 ^a	0.114 ^a	-0.017 ^b	-0.195 ^c	3	3, 11	*
	2	1	105	94–142	0.03	0.102 ^a	-0.086 ^b	0.017 ^c	-0.033 ^{bc}	3	7, 11	
	3	1	160	142–170	0.04	0.067 ^a	-0.082 ^b	-0.083 ^b	0.098 ^a	2	1, 7	
	4	1	217	186–223	0.03	0.049 ^a	0.057 ^a	-0.006 ^a	-0.101 ^b	2	10, 11, 12	*
	5	3	35	0–139	0.02	0.039 ^a	0.001 ^a	-0.094 ^b	0.055 ^a	2	–	
	6	4	79	67–87	0.04	-0.083 ^a	0.015 ^b	-0.028 ^{ab}	0.096 ^c	3	7, 11, 12	
	7	4	164	133–210	0.03	-0.045 ^a	-0.007 ^a	0.103 ^b	-0.052 ^a	2	2, 3, 6, 10, 11	**
	8	6	23	20–29	0.03	-0.021 ^a	0.094 ^b	-0.087 ^c	0.014 ^a	2	11	
	9	7	139	112–149	0.03	-0.057 ^a	-0.057 ^a	0.041 ^b	0.073 ^b	2	–	
	10	8	33	21–125	0.02	-0.032 ^a	-0.040 ^a	0.073 ^b	0.001 ^a	2	4, 7	
	11	9	75	68–101	0.03	-0.020 ^a	-0.025 ^a	-0.054 ^a	0.099 ^b	2	1, 2, 4, 6, 7, 8, 12	*
	12	10	2	0–79	0.03	-0.021 ^a	0.088 ^b	-0.063 ^a	0.003 ^a	3	4, 6, 11	*
Index	1	1	40	32–62	0.04	0.84 ^a	0.56 ^a	-0.36 ^b	-1.04 ^b	2	3	
	2	1	105	94–119	0.04	1.26 ^a	-0.79 ^b	-0.13 ^b	-0.34 ^b	2	–	
	3	1	162	150–172	0.04	0.91 ^a	-0.46 ^b	-1.05 ^b	0.59 ^a	2	1, 5, 8	*
	4	1	219	190–227	0.03	0.21 ^a	0.71 ^a	0.05 ^a	-0.96 ^b	2	8, 9	
	5	2	105	92–160	0.02	-0.51 ^a	-0.31 ^a	0.82 ^b	0.00 ^c	3	3, 7	
	6	3	27	0–45	0.03	-0.02 ^a	0.63 ^a	-0.99 ^b	0.38 ^a	2	–	
	7	4	79	67–82	0.03	-0.61 ^a	0.02 ^b	-0.31 ^{ab}	0.90 ^c	3	5, 8	
	8	4	153	137–170	0.04	-0.50 ^a	0.25 ^b	0.94 ^b	-0.69 ^a	2	3, 4, 7	*
	9	6	23	13–36	0.02	-0.43 ^a	0.73 ^b	-0.55 ^a	0.24 ^b	2	4	
	10	8	52	31–63	0.03	-0.30 ^{ab}	0.22 ^a	0.69 ^c	-0.62 ^b	3	–	

^{a,b,c,d} Allelic effects followed by a common letter do not differ significantly at the 5% risk level

^cInteraction: *5%, **=5%

Table 5 QTL-by-QTL interactions (QTL × QTL) and QTL by genetic background interactions (QTL × background): number of interactions individually significant at the 5% risk level and number of significant interactions in model (7)

Trait	Test	Number of tests	Individually significant interactions		Model (7)	
			Number	%	Significant interactions	R^2 increase compared to model (3)
Silking date	QTL × QTL	55	5	9.1	1	1.5%
	QTL × background	11	1	9.1	1	
Grain moisture	QTL × QTL	78	3	3.8	1	1.2%
	QTL × background	13	1	7.7	0	
Grain yield	QTL × QTL	66	15	22.7	7	9.6%
	QTL × background	12	5	41.7	1	
Index	QTL × QTL	21	7	23.8	3	6.3%
	QTL × background	10	2	20.0	0	

Epistasis

We performed between 21 and 78 QTL-by-QTL tests for epistasis, depending on the trait (model (6)). Between 3 and up to 15 interactions were detected as significant at the 5% significance level (Table 5). Except for grain moisture, we detected more interactions than expected by chance only at this risk level. Moreover, with 15 significant interactions out of 78 tests, epistasis was detected more frequently for grain yield than for silking date or grain moisture. Many QTL on chromosome 1 were involved in interactions with other loci, especially for silking date and grain yield (Table 4). The grain yield QTL on chromosome 9 interacted with seven other QTL, three of which were located on chromosome 1. The silking date QTL on chromosome 10 (30 cM position), explaining a large part of the phenotypic variance (18%), interacted with three other QTL (Table 4).

For each trait, we detected also more significant QTL-by-genetic-background interactions than expected by

chance at a 5% significance level (model (4); Table 5). Only one QTL interacted with the genetic background for silking date and grain moisture and two for index. Many interactions (5 out of 12 QTL) were detected for grain yield. Significant QTL-by-genetic-background interactions were only found for QTL involved in at least two QTL-by-QTL interactions. Out of the eight QTL interacting with at least three other QTL, all but one (QTL 6 for grain yield) were involved in interactions with the genetic background. Out of the ten QTL interacting with two other QTL, only two exhibited a significant QTL-by-genetic-background interaction, i.e. QTL 11 for grain moisture and QTL 1 for grain yield.

Including significant epistatic interactions in a global model increased the percentage of phenotypic variance explained, especially for grain yield. Indeed, when compared to the percentage of phenotypic variance explained by the QTL detected with model (3), R^2 of model (7) increased of 9.6% for grain yield and only of a few percent for silking date (+ 1.5%) and grain moisture (+ 1.2%) (Table 5).

Table 6 Marker-by-marker and marker-by-genetic background interactions: percentages of significant interactions for individual risk level (P values) of 5% and 1‰ and a false discovery rate (q value) of 10%

Trait	Type of interaction	Percentage of significant tests			Estimated % of true null hypotheses
		$P < 5\%$	$P < 1\text{‰}$	$q < 10\%$	
Silking date	Marker × marker	10.44	0.38	0	69
	Marker × genetic background	12.87	5.15	5.88	49
Grain moisture	Marker × marker	7.72	0.22	0	75
	Marker × genetic background	11.4	3.68	1.47	74
Grain yield	Marker × marker	8.30	0.25	0.03	71
	Marker × genetic background	13.60	2.57	1.47	80
Index	Marker × marker	8.71	0.43	0.36	84
	Marker × genetic background	7.72	4.41	1.84	93

The percentage of true null hypotheses is estimated based on the distribution of the P values

These analyses of epistatic effects for QTL detected with model (3) were complemented by a whole genome scan at marker positions (Table 6). For marker-by-marker interactions, observed P values yielded an estimate of the percentage (π_0) of true null hypotheses (i.e. no epistasis) between 69% for silking date and 84% for index. Epistasis may therefore concern up to 31% of the tests for silking date. For this trait, the lower q value, corresponding to the most significant test, was equal to 11.7%. So, no test could be considered as significant if one wants to limit the FDR at a level of 10%. This indicates that even if there is certainly some epistasis for this trait, it is likely explained by numerous digenic interactions with limited individual effect. On the contrary, for index and to a less extent for grain yield, we observed higher π_0 values but detected several significant digenic interactions when considering an FDR of 10%. The part of the genome involved in epistasis should therefore be smaller for these traits, but with higher contribution of individual effects. As expected significant marker-by-marker interactions were found in regions where QTL-by-QTL interactions were detected. However, it can be noted that the highest marker-by-marker interactions were often found for markers either (1) within the CI of QTL detected by model (3), but not at the markers closest to the estimated QTL positions or (2) at markers located outside the QTL CI but close to their edges. Highly significant digenic interactions were found for grain yield and index between markers located on chromosome 4 and several markers located on chromosomes 1, 7 and 9. For index, another important interaction was found between the beginning of chromosome 8 and the beginning of chromosome 1.

For marker-by-genetic-background interaction tests, π_0 varied between 49% for silking date and 93% for index. As for marker-by-marker interaction tests, this result suggests that epistasis affects a larger proportion of the genome for silking date than for the other traits. Significant marker-by-genetic-background interactions were detected with an FDR of 10% for all the traits. So, for silking date and grain moisture for which no marker-by-marker interactions were found, marker-by-genetic-background interactions mainly result from several digenic interactions of relatively low effect that add up each other. For the other traits, marker-by-genetic-background interactions generally coincided with markers involved in a few highly significant marker-by-marker interactions. Interestingly, investigation of markers-by-genetic-background interactions revealed significant epistatic effects for three chromosome regions where QTL had been detected with

model (1) and not confirmed with model (3): one silking date QTL (chromosome 1), one index QTL (chromosome 9) and one grain moisture QTL (chromosome 3). The QTL for silking date detected in the population $DE \times F283$ at 105 cM on chromosome 1 (Fig. 3) is particularly interesting. In addition to model (1), it was also detected with model (2) but not with model (3). The F values curve (Eq. 5) for marker-by-genetic-background interaction presents a maximum very close to the QTL position estimated with model (2). Estimation of the allelic effects nested within populations underlines inconsistencies in ranking. In the $DE \times F9005$ population, the contrast between the segregating alleles was $\alpha_{DE} - \alpha_{F9005} = -0.33$ and in the population $F9005 \times F283$ the contrast $\alpha_{F9005} - \alpha_{F283} = 0.03$. So, based on these values, we expected to observe in population $DE \times F283$ a contrast $\alpha_{DE} - \alpha_{F283}$ equal to -0.30 (i.e. $\alpha_{F283} > \alpha_{DE}$). The contrast estimated in this population was opposite and equal to $+0.59$ (i.e. $\alpha_{F283} < \alpha_{DE}$). So, the difference between the expected value under the hypothesis of additivity and the observations (d_2) was equal to 0.89. This result is associated with the lack of consistency between the ranking of alleles in the multipopulation connected analyses, (e.g. $\alpha_{F810} > \alpha_{DE} > \alpha_{F9005} > \alpha_{F283}$) and the ranking in individual populations (for instance $\alpha_{DE} < \alpha_{F9005}$ in population $DE \times F9005$). When adding this QTL, with disconnected effects, in model (4), the R^2 was increased by 0.75%.

Discussion

Comparison of models for QTL detection

The multiparental design analyzed here allowed the detection of 46 QTL for four different traits. When performing individual analyses of each of the six populations, only few QTL were detected in two or more populations (based on overlapping QTL position CI). This is in accordance with other studies (Beavis et al. 1991; Mihaljevic et al. 2004) who also reported only poor to moderate QTL congruency for agronomic traits in different biparental populations of maize. Consistent with Mihaljevic et al. (2004), congruent QTL among different populations were often detected in three crosses, that involved a same parental inbred with a large allele effect at this QTL, relative to alleles of the three other parental inbreds.

When compared to single-population analyses, model (2) detected approximately the same number of QTL but the two sets of detected QTL were slightly different, some QTL being detected by model (1) but

not by model (2) and conversely. It has to be noted that the F value curve in these regions was often very close to the threshold, leading to the detection of a QTL with one model but not with the other. The lack of consistency between the results of the two models, which are using the same information, therefore seems to be mainly attributed to the limited power of both models.

Taking into account the connections for QTL detection, (model (3)) allowed us to detect all QTL detected in single-population analyses (model(1)) but 3, all QTL detected with model (2) but 1 and to detect 15 additional QTL detected neither by model (1) nor by (2). Model (3), therefore, better valorizes the information than model (1) and leads to relatively a larger gain in power of detection than model (2). As a consequence, model (3) explained a higher proportion of phenotypic variance than model (2). These results fall in line with the experimental results obtained by Rebai et al. (1997) on silking date QTL in maize by means of simple interval mapping. Moreover, model (3) led to a reduction of the QTL position CI for the different traits, when compared to models (2) and (1).

Based on theoretical results, the power of model (3) was expected to be higher than the power of model (2) for an additive determinism of the trait, due to a reduced number of parameters (Rebai and Goffinet 1993, 2000). On the other hand, Jannink and Jansen (2001) showed by simulation that considering allele effects as identical over populations may have a lower power in case of epistatic effects and some conditions on the repartition of alleles among parents. They also concluded that nesting QTL effects within population only lead to a small reduction of the power of QTL detection in case of additivity. However, they considered the case of three connected populations derived from the cross between three different inbred lines. In this situation, considering allele effects as identical over populations reduced the df for the QTL effect only by one. In our case, the reduction is more substantial since considering model (3) instead of model (2) led to remove 3 df per QTL fitted. In this experimental study, out of a total of 46 detected QTL, we found a single QTL detected by model (2) and not by model (3), versus 22 detected by model (3) and not by model (2) (Fig. 4). Only a very small minority of QTL, therefore, present conditions that lead to an advantage to model (2) for QTL detection, despite traits analyzed include grain yield, known to display relatively strong epistatic effects (see below). This suggests that model (3) should be preferred to model (2) for QTL detection in multiparental designs. The advantage of model (3) over model (2) is expected to increase as the difference

between the number of populations and the number of inbred lines from which these populations are derived increases, that is to say as the degree of connectedness between populations increases.

Besides a clear gain in power, a major interest of model (3) relative to model (2) is to permit a global evaluation of allelic effects at each detected QTL. This allows ranking parental alleles at each QTL, testing for differences in effects and therefore evaluating the number of allele classes with equivalent effects. The balanced number of favorable alleles found for the parents used in this study illustrates the necessity to address QTL mapping and breeding within broad genetic pools. The design presented here attempts to do so but still involved a limited number of parents when compared to usual breeding practices (e.g. 100 populations started each year from tens of parents) or the size of core-collections representing a large fraction of diversity within one species (e.g. 24 accessions for 96% of the molecular diversity of a worldwide collection of *Arabidopsis thaliana* (McKhann et al. 2004)). The statistical approach that we used here can in principle be extended to more complex designs. We assumed fully informative QTL (one allele per parent), which led here to a relatively limited number of parameters (three per QTL using model (3)), and yet needed a quite long calculation time. If more parents were considered, the number of parameters would have increased considerably, leading to higher computation time and/or a loss in power (if a same total population size was considered). An improvement of the method in case of more parents could be to reduce the number of parameters by grouping alleles with equal effects and testing for class effects. This idea is supported by our results on a posteriori pairwise comparisons of additive effects at the QTL (advocated by Rebai and Goffinet 2000): out of 46 QTL, 26 displayed only two significant different allelic effects. Jannink and Wu (2003) proposed a solution to estimate the number of allelic effects with an MCMC algorithm to infer the probability that QTL alleles are identical in state. They evaluated the impact of such a procedure on allelic effect estimates in comparison to models assuming parents that carry unique alleles. They showed it was difficult to accurately assess the number of alleles, either when the population sizes were low or when the parental alleles had close effects. Another solution is to infer the number of “real” alleles from the genotypes at markers located around QTL position. Using the assumption that marker haplotypes are a good indicator of QTL allele sharing, Jansen et al. (2003) showed by simulation that this a priori information could be used to reduce the number of parameters in the model

and to increase power. But using only few markers to define a haplotype for a QTL might not be sufficient. So this approach requires a quite dense genetic map with highly informative markers. Recently, Li et al. (2005) compared a combined cross QTL detection on four connected inbred line crosses of mice issued from five parents to single-population analyses. To reduce the total number of parameters in the combined analyses, they proposed to encode a priori the five parental alleles in a biallelic manner. They could do so because common strains of inbred laboratory mice are mainly derived from two original subspecies and present a biallelic segregation for most loci. This may not be applicable to other species. Contrary to our results, they did not observe a clear gain in power conferred by the combined analysis, what may be due to a “false” biallelic encoding. However, the combined cross analysis allowed them to reduce the QTL position CI when compared to single-population QTL CI. Other multi-allelic approaches have been investigated by considering allelic effects as random (Xu 1998; Yi and Xu 2002; Crepieux et al. 2004). In these approaches, the allelic effects at each QTL are assumed to be normally distributed and one parameter per QTL (the variance explained by the QTL) is estimated. So, the number of parameters in the model does not increase with the number of parental lines as in a fixed model. These approaches are all the more appropriate as the number of segregating alleles is large and the experimental design complex.

Epistatic effects

Our results also illustrate that global analysis of connected multiparental designs makes it possible to test for epistatic interactions of individual QTL with the genetic background. The approach we used was slightly different from the one proposed by Jannink and Jansen (2001). They performed the detection using maximum likelihood, and first used a “full” model (where allelic effects were nested within populations) to detect QTL. Then to test for QTL-by-genetic-background interactions, they kept a “full” model for the cofactors, and considered a “reduced” model (i.e. considering allele effects as identical over populations) only for the QTL under study. The approach we followed is somehow the reverse one. We performed the QTL detection with a “reduced” model (by taking into account the connections, model (3)) and then used a “full” model (without connections, model (4)) only for the QTL under study. The model we used is more parsimonious and takes advantage of the gain of power conferred by model (3).

Compared to former comparable analyses (Charcosset et al. 1994; Rebai et al. 1997), our results underline more significant QTL-by-genetic-background epistatic effects, especially for grain yield (5 significant tests out of 12). This is consistent with the complexity of the trait (Butron et al. 2004) and phenotypic evidence for epistatic effects. Consistent with results of Lamkey et al. (1995), Melchinger et al. (1988) and Moreno-Gonzalez and Dudley (1981), we found a difference between the mean testcross performance of the six F_2 populations and the average testcross performance of the corresponding parental lines, which is an indication of epistatic effects. Note, however, that other studies conducted with the same approach did not report epistasis (Eta-Ndu and Openshaw 1999; Hinze and Lamkey 2003), suggesting that these effects depend on the genetic material which is considered and possibly environmental factors. All QTL presenting a significant epistatic interaction with the genetic background also displayed at least two significant pairwise interactions with other QTL. Conversely, some significant interactions were detected between two QTL in cases where none of them displayed significant interaction with the genetic background. This may be attributed to both (1) a relatively high proportion of false positives in pairwise tests and (2) a repartition of alleles among parents so that several digenic epistatic effects cancel out each other and result in no significant QTL-by-genetic-background interactions. The QTL-by-genetic background interaction test did not allow the detection of additional QTL showing epistatic effects compared to QTL-by-QTL interaction test. This was not expected and might be specific of this experiment.

Some regions were involved in epistatic interactions for different traits. For example, around position 150 on chromosome 1 (Fig. 1), QTL were detected for all traits and displayed numerous significant interactions for grain yield, silking date and index. This region was reported as presenting significant interactions with the genetic background (Charcosset et al. 1994). A region around position 50 on chromosome 9, which had significant additive effect for grain moisture and grain yield, displayed significant epistatic effects for all the traits. In a study comparing maize growth and developmental QTL positions to gene localization, Khavkin and Coe (1997) showed that numerous QTL have been reported in this region (referred to as bin 9.03), which corresponds to a cluster of developmental genes. Developmental genes are known to interact in complex regulatory patterns which may explain why we detected many epistatic interactions in this region. Lastly, many digenic interactions were significant for silking

date and grain yield on chromosome 10 for a region around position 30. This region corresponds to a major silking date QTL ($r^2 = 18\%$ for model (3)), which was detected in the three populations involving *F283* and presents a significant QTL-by-genetic-background effect. Interestingly, this region corresponded to a QTL “hot spot” reported by Chardon et al. (2004), in a large study about flowering time in maize, including 313 QTL from the literature. The numerous interactions we detected in this region supports the assumption that it is involved in a complex interaction pattern controlling floral initiation.

Besides interactions between QTL that displayed significant additive effects, some new “epistatic regions” were detected using a whole genome scan at marker positions. Interestingly, some of these regions, that present significant background interactions, correspond to QTL detected in single-population analysis (model (1)) but not in the multipopulation connected analysis (model (3)). So the test of interaction with the genetic background helped us to conclude that epistasis at least partly explains differences between results obtained with the different models. In this study, epistasis tests were performed in a second step after QTL detection. The occurrence of chromosome regions showing epistasis without significant additive effect suggests that it could be beneficial to include epistasis effects in the QTL detection step. Wang et al. (1999), showed that including pairwise epistatic interactions as random effects in the model of QTL detection may improve significantly the power of QTL detection and the accuracy of the estimated QTL parameters. A new Bayesian approach also including epistasis was more recently proposed (Yi et al. 2005). The authors also concluded that including epistasis increases the accuracy of the QTL parameter estimates. These methods are not yet implemented in softwares adapted to multipopulation experimental design but our results clearly illustrate that these models deserve consideration for the development of new versions.

Marker-assisted selection in multiparental populations

In this experiment, the connections between populations allowed us to identify the parental origin of favorable allele(s) at each QTL. An appealing perspective to this work from an applied point of view would therefore be to assemble these favorable alleles in a single line. This can be achieved thanks to marker-assisted breeding schemes in which individuals are selected based on information at markers in regions of

interest (see Servin et al. 2004, for a method adapted to a multiallelic context). Application of this approach to multiparental designs has to account for (1) the lack of polymorphism of markers and (2) the size of CI around QTL positions. In such context, knowing the genotypes at markers flanking the QTL position is not sufficient to control that a given candidate carries the allele of interest. An haplotype-based approach combining pedigree information and genotypes at several markers within the CI is required.

One can anticipate that the efficiency of MAS in terms of phenotypes of the end product material will depend on the stability of QTL effects. Discrepancies have indeed been observed between the predicted genetic gain and the progress obtained after quantitative trait allele transfer in biparental populations (Zhu et al. 1999; Bouchez et al. 2002). These could be interpreted as a lack of stability of QTL effects due to statistical errors in initial QTL detection, epistasis and/or interactions with the environment. Consequences of epistasis might be even more important in a multiallelic context where the genetic background is highly variable. To prevent from consequences of epistasis one can imagine selecting only on QTL not involved in epistatic interactions. In our case, only 2 QTL among the 12 detected for grain yield with model (3) were not interacting neither with another QTL nor with the genetic background. However, the gain of power that we observed using model (3) versus model (2) indicates the additive effects of QTL were preponderant. Moreover, in model (3) the estimated allelic effects at QTL can be interpreted as the ‘stable’ part of the QTL effects over the different genetic backgrounds. So one can expect that selecting on these effects will minimize the impact of epistasis on the genetic gain.

Beside epistasis, QTL-by-environment interactions can also generate QTL effects instability. As advocated by Wang et al. (1999), it is therefore interesting to include both epistasis and QTL-by-environment interaction effects in QTL detection models, first to better anticipate sources of QTL effect instability but also to further increase the power of detection and the accuracy of QTL parameter estimates. Our populations were phenotyped in a large range of environments (10 field trials) and genotype \times environment interactions were found significant for all the traits (results not shown). However, contrary to what was observed in another study by Moreau et al. (2004), a clustering approach did not reveal any clear structure among these trials that would allow us to define groups of environments to work with. Besides, as no software was available to combine QTL-by-environment interactions and a multiparental analysis, we decided to

work on the average performances over the whole design and did not look for QTL-by-environment interactions. Further efforts are needed to improve this approach, but our results already clearly illustrate the benefits of using connected populations for QTL detection and MAS. Multipopulation QTL analyses are certainly a good way to better valorize the numerous connected populations produced each year in conventional breeding programs to (1) detect QTL in elite germplasm and identify sources of allelic variation directly useful for selection and (2) to better understand the genetic architecture of trait, and in particular evaluate the contribution of epistasis. This should contribute to strengthen the link between selection and more fundamental research.

Acknowledgments This research program was funded by INRA and Agriobtentation. We are grateful to our colleagues involved in marker analyses at le Moulon (D. Madur, V. Combes, F. Dumas), and colleagues involved in material production and field testing at INRA le Moulon (P. Bertin, P. Jamin, D. Coubriche, S. Jouanne), at INRA Dreux, at INRA Rennes, at INRA Lusignan and at INRA Mons. We also thank the BIA unit at INRA Toulouse for the maintenance of MCQTL software (B. Mangin, B. Ngom, J. Marcel). We are grateful to Guy Decoux for the informatic program used for the graphical display of the maps. We are grateful to Rex Bernardo for very helpful advices on the writing of this manuscript.

References

- Asins MJ (2002) Present and future of quantitative trait locus analysis in plant breeding. *Plant Breed* 121:281–291
- Beavis WD, Grant D, Albertsen M, Fincher R (1991) Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. *Theor Appl Genet* 83:141–145
- Bernardo R (2002) Quantitative traits in plants. Stemma, Woodbury
- Bouchez A, Hospital F, Causse M, Gallais A, Charcosset A (2002) Marker-assisted introgression of favorable alleles at quantitative trait loci between maize elite inbred lines. *Genetics* 162:1945–1959
- Butron A, Velasco P, Ordas A, Malvar RA (2004) Yield evaluation of maize cultivars across environments with different levels of pink stem borer infestation. *Crop Sci* 44:741–747
- Charcosset A, Causse M, Moreau L, Gallais A (1994) Investigation into the effect of genetic background on QTL expression using three recombinant inbred lines (RIL) populations. In: van Ooijen JW, Jansen J (eds) *Biometrics in plant breeding: applications of molecular markers*. CRPO-DLO, Wageningen, The Netherlands, pp 75–84
- Charcosset A, Mangin B, Moreau L, Combes L, Jourjon MF et al. (2000) Heterosis in maize investigated using connected RIL populations. In: *Quantitative genetics and breeding methods: the way ahead*. INRA, Paris, France, pp 89–98
- Chardon F, Virilon B, Moreau L, Falque M, Joets J et al (2004) Genetic architecture of flowering time in maize as inferred from QTL meta-analysis and synteny conservation with the rice genome. *Genetics* 168:2169–2185
- Crepieux S, Lebreton C, Servin B, Charmet G (2004) IBD-based QTL detection in inbred pedigrees: a case study of cereal breeding programs—IBD-based multi-cross QTL mapping. *Euphytica* 137:101–109
- Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. *Nature* 386:485–488
- Eta-Ndu JT, Openshaw SJ (1999) Epistasis for grain yield in two F2 populations of maize. *Crop Sci* 39:346–352
- Haldane JBS (1919) The combination of linkage values and the calculation of distances between the loci of linked factors. *J Genet* 8:299–309
- Haley CS, Knott SA (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69:315–324
- Hinze LL, Lamkey KR (2003) Absence of epistasis for grain yield in elite maize hybrids. *Crop Sci* 43:46–56
- Jannink JL, Jansen RC (2000) The diallel mating design for mapping interacting QTLs. In: *Quantitative genetics and breeding methods: the way ahead*. INRA, Paris, France, pp 81–88
- Jannink JL, Jansen R (2001) Mapping epistatic quantitative trait loci with one-dimensional genome searches. *Genetics* 157:445–454
- Jannink JL, Wu XL (2003) Estimating allelic number and identity in state of QTLs in interconnected families. *Genet Res* 81:133–144
- Jansen RC (1993) Interval mapping of multiple quantitative trait loci. *Genetics* 135:205–211
- Jansen RC, Jannink JL, Beavis WD (2003) Mapping quantitative trait loci in plant breeding populations: use of parental haplotype sharing. *Crop Sci* 43:829–834
- Jourjon MF, Jasson S, Marcel J, Ngom B, Mangin B (2005) MCQTL: multi-allelic QTL mapping in multi-cross design. *Bioinformatics* 21:128–130
- Kearsey MJ, Farquhar AG (1998) QTL analysis in plants; where are we now? *Heredity* 80:137–142
- Khavkin E, Coe E (1997) Mapped genomic locations for developmental functions and QTLs reflect concerted groups in maize (*Zea mays* L.). *Theor Appl Genet* 95:343–352
- Knapp SJ, Stroup WW, Ross WM (1985) Exact confidence intervals for heritability on a progeny mean basis. *Crop Sci* 25:192–194
- Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T et al. (2002) Hd3a, a rice ortholog of the *Arabidopsis* FT gene, promotes transition to flowering downstream of Hd1 under short-day conditions. *Plant Cell Physiol* 43:1096–1105
- Lamkey KR, Schnicker BJ, Melchinger AE (1995) Epistasis in an elite maize hybrid and choice of generation for inbred line development. *Crop Sci* 35:1272–1281
- Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ et al. (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174–181
- Li RM, Lyons A, Wittenburg H, Paigen B, Churchill GA (2005) Combining data from multiple inbred line crosses improves the power and resolution of QTL mapping. *Genetics* 104:033993
- Liu J, Van Eck J, Cong B, Tanksley SD (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proc Natl Acad Sci USA* 99:13302–13306
- Liu JP, Cong B, Tanksley SD (2003) Generation and analysis of an artificial gene dosage series in tomato to study the mechanisms by which the cloned quantitative trait locus fw2.2 controls fruit size. *Plant Physiol* 132:292–299
- Lynch M, Walsh B (1998) Mapping QTLs: inbred line crosses—precision of ML estimates of QTL position. In: *Sinauer*

- Associates (ed) Genetics and analysis of quantitative traits. Sinauer, Sunderland, pp 448–450
- McKhann HI, Camilleri C, Berard A, Bataillon T, David JL et al. (2004) Nested core collections maximizing genetic diversity in *Arabidopsis thaliana*. *Plant J* 38:193–202
- Melchinger AE, Schmidt WGHH (1988) Comparison of test-crosses produced from F2 and first backcross populations in maize. *Crop Sci* 28:743–749
- Mihaljevic R, Utz HF, Melchinger AE (2004) Congruency of quantitative trait loci detected for agronomic traits in test-crosses of five populations of European maize. *Crop Sci* 44:114–124
- Moreau L, Monod H, Charcosset A, Gallais A (1999) Marker-assisted selection with spatial analysis of unreplicated field trials. *Theor Appl Genet* 98:234–242
- Moreau L, Charcosset A, Gallais A (2004) Use of trial clustering to study QTL*environment effects for grain yield and related traits in maize. *Theor Appl Genet* 110(1):92–105
- Moreno-Gonzalez J, Dudley JW (1981) Epistasis in unrelated maize hybrids determined by three methods. *Crop Sci* 21:644–651
- Muranty H (1996) Power of tests for quantitative trait loci detection using full-sib families in different schemes. *Heredity* 76:156–165
- van Ooijen JW (1992) Accuracy of mapping quantitative trait loci in autogamous species. *Theor Appl Genet* 84:803–811
- Rebai A, Goffinet B (1993) Power of tests for QTL detection using replicated progenies derived from a diallel cross. *Theor Appl Genet* 86:1014–1022
- Rebai A, Goffinet B (2000) More about quantitative trait locus mapping with diallel designs. *Genet Res* 75:243–247
- Rebai A, Goffinet B, Mangin B, Perret D (1994) QTL detection with diallel schemes. In: van Ooijen JW, Jansen J (eds) *Biometrics in plant breeding: applications of molecular markers*. CRPO-DLO, Wageningen, The Netherlands, pp 170–177
- Rebai A, Blanchard P, Perret D, Vincourt P (1997) Mapping quantitative trait loci controlling silking date in a diallel cross among four lines of maize. *Theor Appl Genet* 95:451–459
- Salvi S, Sponza G, Morgante MFK, Meeley R, Ananiev E et al. (2005) The maize QTL Vgt1 corresponds to a ca. 2.7-kb non-coding region upstream of an Ap2-like gene. In: *Maize meeting*, Lake Geneva, Wisconsin, pp 195
- SAS (1989a) SAS IML user's guide, Version 6.03, 4th edition. SAS Institute, Cary
- SAS (1989b) SAS procedures guide, Version 6, 3rd edition. SAS Institute, Cary
- SAS (1990) SAS/STAT user's guide, Version 6, 4th edition. SAS Institute, Cary
- Servin B, Martin OC, Mezard M, Hospital F (2004) Toward a theory of marker-assisted gene pyramiding. *Genetics* 168:513–523
- Storey JD, Tibshirani R (2003) Statistical significance for genome wide studies. *PNAS* 100:9440–9445
- Takahashi Y, Shomura A, Sasaki T, Yano M (2001) Hd6, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the alpha subunit of protein kinase CK2. *Proc Natl Acad Sci USA* 98:7922–7927
- Wang DL, Zhu J, Li ZK, Paterson AH (1999) Mapping QTLs with epistatic effects and QTLxenvironment interactions by mixed linear model approaches. *Theor Appl Genet* 99:1255–1264
- Xu SZ (1998) Mapping quantitative trait loci using multiple families of line crosses. *Genetics* 148:517–524
- Yi N, Xu S (2002) Mapping quantitative trait loci with epistatic effects. *Genet Res* 79:185–198
- Yi N, Yandell BS, Churchill GA, Allison DB, Eisen EJ et al. (2005) Bayesian model selection for genome-wide epistatic quantitative trait loci analysis. *Genetics* 170:1333–1344
- Zeng ZB (1994) Precision mapping of quantitative trait loci. *Genetics* 136:1457–1468
- Zhu H, Briceño G, Dovel R, Hayes PM, Liu BH et al. (1999) Molecular breeding for grain yield in Barley: an evaluation of QTL effects in a spring barley cross. *Theor Appl Genet* 98:772–779