Theor Appl Genet (2006) 112: 1009–1023
DOI 10.1007/s00122-005-0204-z

ORIGINAL PAPER

Mateo Vargas · Fred A. van Eeuwijk
Jose Crossa · Jean-Marcel Ribaut

Mapping QTLs and QTL × environment interaction for CIMMYT maize drought stress program using factorial regression and partial least squares methods

Received: 15 July 2005 / Accepted: 14 December 2005 / Published online: 15 March 2006

Abstract The study of QTL × environment interaction (QEI) is important for understanding genotype × environment interaction (GEI) in many quantitative traits. For modeling GEI and QEI, factorial regression (FR) models form a powerful class of models. In FR models, covariables (contrasts) defined on the levels of the genotypic and/or environmental factor(s) are used to describe main effects and interactions. In FR models for QTL expression, considerable numbers of genotypic covariables can occur as for each putative QTL an additional covariable needs to be introduced. For large numbers of genotypic and/or environmental covariables, least square estimation breaks down and partial least squares (PLS) estimation procedures become an attractive alternative. In this paper we develop methodology for analyzing QEI by FR for estimating effects and locations of QTLs and QEI and interpreting QEI in terms of environmental variables. A randomization test for the main effects of QTLs and QEI is presented. A population of F_2 derived F_3 families was evaluated in eight environments differing in drought stress and soil nitrogen content and the traits yield and anthesis silking interval (ASI) were measured. For grain yield, chromosomes 1 and 10 showed significant QEI, whereas in chromosomes 3 and 8 only main effect QTLs were observed. For ASI, QTL main effects were observed on chromosomes 1, 2, 6, 8, and 10, whereas QEI was observed only on chromosome 8. The assessment of the QEI at chromosome 1 for grain yield showed that the QTL main effect explained 35.8% of the QTL + QEI variability, while QEI explained 64.2%. Minimum temperature during flowering time explained 77.6% of the QEI. The QEI analysis at chromosome 10 showed that the QTL main effect explained 59.8% of the QTL + QEI variability, while QEI explained 40.2%. Maximum temperature during flowering time explained 23.8% of the QEI. Results of this study show the possibilities of using FR for mapping QTL and for dissecting QEI in terms of environmental variables. PLS regression is efficient in accounting for background noise produced by other QTLs.

Communicated by A. E. Melchinger

M. Vargas
Universidad Autónoma Chapingo,
Chapingo, Edo. Mexico, Mexico

J. Crossa (✉) · M. Vargas
Biometrics and Statistics Unit, International Maize and Wheat Improvement Center (CIMMYT),
6-641, 06600 Mexico D.F., Mexico
E-mail: j.crossa@cgiar.org

F. A. van Eeuwijk
Laboratory of Plant Breeding, Department of Plant Science,
Wageningen University,
386, 6700 AJ Wageningen, The Netherlands

J.-M. Ribaut
Genetic Resources Program, International Maize and Wheat Improvement Center (CIMMYT),
6-641, 06600 Mexico D.F., Mexico

Introduction

Progress in molecular genetics with respect to the creation of ever more polymorphic molecular markers has led to the common application of QTL mapping methodology in genetics and breeding, and has prompted intensive research on biostatistical methods for QTL detection and quantification. Statistical QTL detection and estimation approaches can be roughly grouped into two classes: (1) regression-based methods using least squares or generalized least squares estimation methods and (2) mixture model-based approaches using maximum likelihood (ML) as the estimation method (Lynch and Walsh 1998). Regression methods allow inclusion of different experimental designs, additional treatment

structures and are, generally, less computer-intensive than mixture models. Early attempts to locate QTLs and quantify their effects consisted of analysis of variance at marker positions of phenotypic responses using genotypic marker classes. Equivalently, linear regression approaches were developed that transformed marker information into predictors, after which putative QTL effects were estimated by regressing the phenotypic responses on those genetic predictors. For example, for a codominant marker system, three classes of genotypes can be defined as MM, Mm, and mm, which can be transformed into a genetic predictor for additive genetic effects, x, with the value 1 when the marker genotype is MM, while for marker genotype Mm this genetic predictor takes the value 0, and finally, for marker genotype mm the value of the genetic predictor would be -1. Similarly, genetic predictors for dominance genetic effects can have the value 0 for the marker genotypes MM and mm, while the heterozygote Mm receives the value 1 (Haley and Knott 1992).

With marker regression, the effect of the QTL is typically underestimated (see p. 437 in Lynch and Walsh 1998). An improvement on marker regression constructs additional, virtual genetic predictors in between the observed markers. The values for such virtual genetic predictors are constructed as functions of the genotypes at the left and right flanking markers, and the distance of the virtual genetic predictor to these markers (Haley and Knott 1992; Jiang and Zeng 1997). This approach is known under the name of simple interval mapping (SIM), and was originally introduced as an ML method (Lander and Botstein 1989), until Haley and Knott (1992) presented a regression alternative.

Provided the marker distance is not too large (<20 cM), results from ML-based SIM and regression-based SIM correspond well (Haley and Knott 1992; Martinez and Curnow 1992). With a large number of missing data or wide gaps in the marker linkage map, discrepancies between the two methods can occur. The SIM method represents an advance over analysis of variance or regression at marker positions, but it is based on an often unrealistic genetic model of a single QTL influencing the phenotypic trait, while ignoring the effects of additional QTLs on the same or other chromosomes. The latter issue is addressed by an extension of SIM called composite interval mapping (CIM). CIM combines interval mapping for a single QTL in a given interval with a correction for QTL effects elsewhere in the genome. In the multiple regression approach, QTL effects elsewhere are accounted for by a set of genetic predictors close to or at the positions where QTLs have been found or are suspected. CIM has greater power for QTL detection and higher precision for QTL localization and estimation relative to SIM (Zeng 1994). The idea is that the cofactors reduce, as much as possible, the background noise created by other QTLs in the genome. Haley and Knott (1992), Caliński et al. (2000), and Hackett et al. (2001) presented methods based on univariate and multivariate regression for CIM. Sari-Gorla et al. (1997) proposed a weighted least squares approach coupled with a sequentially rejective Bonferroni method, while a forward selection procedure was used to select a subset of genetic predictors to be used as cofactors.

An important difficulty in using CIM is the appropriate selection of a cofactor set, since the number and chromosome positions of genetic predictors in this set greatly affect the final outcome. When using regression subset selection procedures (forward selection, backward elimination, or stepwise), the number of selected cofactors depends on the chosen significance levels for inclusion or exclusion of cofactors. For example, with $\alpha = 0.05$ for inclusion, a large number of cofactors will be selected. Conversely, for $\alpha = 0.01$, some important cofactors may not be included. Another point to notice in CIM is that as the number of markers increases, the map becomes denser and measures for dealing with multicollinearity among cofactors become increasingly relevant. Typically, QTL model strategies will not survey all possible subsets of cofactors and will provide insufficient protection against collinearity between genetic predictors that are closely linked. When cofactors show high collinearity, interpretation of least squares regression coefficients is complicated, if not erroneous, because the coefficients are estimated very imprecisely and tend to be too large in absolute value (see section 8.3 in Montgomery and Peck 1982). A stepwise procedure for selection of cofactors might alleviate some of the collinearity problems, but will still survey only some of the possible subsets. Alternatively, all subset approaches will fail in the presence of the large number of genetic predictors that need to be included in the regression models, as the number of predictors will exceed the number of observations. Alternative estimation methods like partial least squares (PLS) (this paper) or principal components regression (Hwang and Nettleton 2003) will be useful in such conditions.

When mapping QTLs, the phenotypic evaluation of the same segregating population across multiple environments creates the need for statistical models that allow the modeling and interpretation of QTL by environment interaction (QEI), the differential expression of QTLs in relation to changing environmental conditions. The study of QEI is not only of importance by itself, but also, and maybe even more so, because of the relationship between QEI and the phenotypic phenomenon of genotype $\times$ environment interaction (GEI), the dependence of phenotypic differences among genotypes on the environment. Standard statistical models for GEI implicitly model the GEI as the summed result of all QTLs and QEI involved in the production of a specific phenotypic trait (Crossa 1990; Van Eeuwijk et al. 1996), without considering the possibility that different regions of the genome, i.e., QTLs, can have their specific responses to environmental conditions. Models for QEI and GEI can be synthesized within the context of the factorial regression (FR) framework, where phenotypic responses as observed across a set of environments are modeled on genotypic and environmental covariables.

The use of FR models in GEI problems was described by Denis (1988, 1991) and Van Eeuwijk et al. (1996). In its simplest form, FR is equivalent to the inclusion of genotypic and environmental contrasts (covariables) on the levels of the genotypic and environmental factors for two-way genotype × environment tables. In the presence of large numbers of genotypic and/or environmental covariables, least squares estimation procedures break down. PLS estimation procedures have been shown to provide a good alternative for such situations (Aastveit and Martens 1986; Helland 1988; Talbot and Wheelwright 1989; Vargas et al. 1998, 1999).

In important maize-growing areas of the world, grain yield reduction is caused by drought during flowering time as well as low nitrogen content of the soil. Drought causes a delay in silking, an increase in the anthesis silking interval (ASI), and therefore a decrease in grain yield. Thus, under drought stress, selection for small ASI in tropical maize should be correlated with grain yield improvement and ASI becomes an important secondary trait with relatively high heritability and more stability than grain yield. Nevertheless, few studies have been conducted on mapping QTLs responsible for the expression of morphological traits under abiotic stresses such as field drought and low-nitrogen conditions. Ribaut et al. (1996) found stable QTLs for ASI across several water regimes (including severe stress and well-watered conditions) on chromosomes 1, 2, 4, 5, 8, 9, and 10, using F_2 derived F_3 families from a cross between a drought-tolerant and -susceptible parent. For the same F_2 derived F_3 families, Ribaut et al. (1997) reported QTLs for grain yield and yield components under well-watered and severe drought stress conditions. Stable QTLs were found on chromosomes 1 and 10.

In this paper, we further develop the statistical approaches described by Crossa et al. (1999) and Van Eeuwijk et al. (2000, 2002) for modeling QTLs and QEI to analyze the population of F_2 derived F_3 families introduced above. The data set combines phenotypic evaluations across eight environments differing in the level of drought stress and soil nitrogen content (Ribaut et al. 1996, 1997). The main objectives of this research were to demonstrate the use of (1) FR for estimating effects and locations of QTL and QEI; (2) FR for modeling and interpreting QEI in terms of products of genetic predictors and environmental variables (the factor environment is characterized and replaced by its related environmental covariables and the factor genotype is characterized and replaced by genetic predictors); (3) a randomization test for the main effects of QTLs and QEI (additive and dominance genetic effects), controlling the genome-wise error rate; (4) PLS for the simultaneous correction for additive and dominance effects of QTLs and QEI in other parts of the genome outside the evaluated chromosome, thereby avoiding the selection of cofactors, while appropriately dealing with collinearity between markers; and (5) interpret the pattern of QTLs and QEI for both yield and ASI in a maize population derived from a cross between a drought-tolerant and -susceptible parent, where this population was evaluated across a range of environments differing in water and nitrogen availability.

Method for mapping QTL and QEI

Background

Crossa et al. (1999) developed a FR model for QEI in tropical maize, generalizing a FR model for GEI. Genetic predictors derived from marker genotypes were introduced in a FR model as genetic covariables, and combined with various environmental covariables that were derived from meteorological records for each of three development stages. The authors found genetic predictors at marker positions associated with biomass that exhibited significant QEI. Using a FR model in combination with a stepwise variable selection procedure, the QEI was partitioned in cross products of genetic predictors (=marker scores) and environmental covariables. The influence of specific environmental covariables (such as maximum or minimum temperature, precipitation, etc.) on QTL expression was quantified. A PLS approach gave similar results to the application of a FR model on a least squares basis. An advantage of PLS was the reduction of the dimensionality of the GEI and QEI, allowing a low dimensional graphical representation of the GEI and QEI.

Van Eeuwijk et al. (2000, 2002) extended the FR models for GEI and QEI developed by Crossa et al. (1999) from the original marker-based regressions to interval mapping and composite interval mapping. The authors presented (1) a randomization test for controlling the genome-wise error rate, following the logic introduced by Churchill and Doerge (1994) and (2) a PLS strategy to deal with the problem of multicollinearity among multiple cofactors. The PLS strategy consisted of (1) taking all the markers outside the chromosome being evaluated as cofactors, (2) regressing the phenotypic responses on this set of markers using multivariate PLS, (3) calculating the fitted values for the phenotypic responses, and (4) using the corrected phenotypic observations, i.e., the residuals from the PLS regression, in a SIM procedure for the chromosome being evaluated. Note that the marker and QTL information on the chromosome under evaluation should be approximately independent of the marker and QTL information on other chromosomes. The situation is comparable to that for an analysis of an experiment laid out in a randomized complete blocks design, where we would first fit the complete blocks, i.e., fit a regression on the qualitative variable "block", and then carry on to work with the residuals of the regression on those complete blocks. Therefore, the PLS correction for background genetic signals on other chromosomes will not induce correlations in the corrected phenotypic observations of step (4), nor will it complicate or even invalidate inference.

The characteristics of PLS regression allows correcting for all cofactors outside the evaluation window, without having to select a subset of cofactors. Furthermore, the PLS strategy allows the use of moderate population sizes, markers with heavy collinearity, and deals with the natural noise in the quantitative trait to be mapped. Van Eeuwijk et al. (2002) illustrated this approach for yield in maize focusing on chromosome 1. The data in that paper covered eight environments and a population of 211 F_2 derived F_3 lines. Recently, Bjornstad et al. (2004a, b) used PLS to identify QTLs and concluded that PLS gave similar results as other QTL mapping procedures, with the advantage of not having to choose a cofactor set.

Two-way fixed effect analysis of variance

As the start for our modeling strategy, we take the conventional fixed effects two-way analysis of variance model with sum to zero constraints running over indices. In this model, the adjusted mean response, $\bar{y}_{ij}$, of the ith genotype ($i = 1, 2, ..., G$) in the jth environment ($j = 1, 2, ..., E$) with n replications for each of the GEI cells is expressed as $\bar{y}_{ij} = \mu + G_i + E_j + (GEI)_{ij} + \bar{\varepsilon}_{ij}$. Here μ is the grand mean over all genotypes and environments, G_i is the additive effect of the ith genotype, E_j is the additive effect of the jth environment, $(GEI)_{ij}$ is the non-additivity or interaction of the ith genotype in the jth environment, and $\bar{\varepsilon}_{ij}$ is the average error, assumed to be NID $(0, \sigma^2/n)$ (where σ^2 is the within-environment error variance, assumed to be constant).

FR model for the analysis and detection of QTL and QEI

The FR model is an extension of the two-way analysis of variance model. In the FR model, the main effects of genotypes (G), environments (E), and the GEI are modeled in relation to genotypic and environmental covariables. For theory and illustrations of possible partitioning of G, E, and GEI effects, see Van Eeuwijk et al. (1996). The FR framework is also suitable for the mapping of QTL main effects and QEI. Van Eeuwijk et al. (2000, 2002) describe how genetic predictors can be constructed from marker information that allows the detection and estimation of additive and dominance QTL effects with or without QEI. Some relevant details of their approach follow.

In FR, genotypic covariables, x_a ($a = 1, 2, ..., A$) with values x_{ia}, can be introduced for the genotypic main effect, G_i: $G_i = x_{ia}\ \rho_a\ +\ (\text{residual})_i$, where ρ_a is the regression coefficient for the regression of G_i on x_a. For more than one genotypic covariable this becomes $G_i = \Sigma_{a=1}^{A} x_{ia}\rho_a + (\text{residual})_i$. When the genotypic covariable x_a is replaced by genetic predictors x_q, the FR framework can also be used for a genome scan for QTL effects. The genetic predictors then should represent linear transformations of the expected QTL genotypes along the genome, as first explained by Haley and Knott (1992) in their seminal paper on a regression approach toward QTL mapping. In a multiple QTL model, the genotypic main effect is replaced by a sum of regression terms $G_i = \Sigma_{q=1}^{Q} x_{iq}\rho_q + (\text{residual})_i$, where ρ_q is the qth QTL main effect. When genetic predictors are calculated at marker positions only, the FR approach reduces to marker regression. For interval mapping and composite interval mapping, genetic predictors in between marker positions need to be constructed as well. Explicit expressions for calculating genetic predictors for different types of segregating populations, and for dominant and codominant markers, as well as for imputing missing markers, can be found in Haley and Knott (1992), Jiang and Zeng (1997), and Lynch and Walsh (1998).

Analogous to the genotypic main effect in FR, the environmental main effect, E_j, also can be regressed on covariables, in this case environmental covariables, z_b with, values z_{jb}. The corresponding partitioning is $E_j = z_{jb}\beta_b + (\text{residual})_j$, for one environmental covariable, or, $E_j = \sum_{b=1}^{B} z_{jb}\beta_b + (\text{residual})_j$, for multiple environmental covariables. The parameters β_b represent the regression coefficients of the regression of the environmental main effect on the environmental covariables.

For the GEI in FR models, three types of partitionings are possible. Firstly, genotypic covariables are measured and environmental coefficients (potentialities) then need to be estimated, $x_{ia}\rho_{ja}$. Secondly, environmental covariables are measured and genotypic coefficients (sensitivities) need to be estimated, $z_{jb}\beta_{ib}$. Finally, both genotypic and environmental covariables are measured and only a scaling constant needs to be estimated, $x_{ia}z_{jb}v_{ab}$. FR models may describe GEI by one or more terms of the above types. An example of a partitioning of GEI in terms of cross products of genotypic and environmental covariables is $(\text{GEI})_{ij} = \Sigma_{a=1}^{A}\Sigma_{b=1}^{B}\ x_{ia}z_{jb}v_{ab} + (\text{residual})_{ij}$, with v_{ab} as a constant that scales the cross product of the genotypic covariables, x_a, with the environmental covariables, z_b. Each cross product represents one degree of freedom in the GEI subspace. A model for GEI consisting of a series of one degree of freedom cross products will be shown below to be very appropriate for the modeling of QEI.

Within a QTL analysis by FR, a multiple QEI model follows easily from models for GEI: $(\text{GEI})_{ij} = \Sigma_{q=1}^{Q} x_{iq}\rho_{jq} + (\text{residual})_{ij}$, where ρ_{jq} represents a QEI effect, i.e., differential QTL expression in relation to the main effect QTL expression, for the qth QTL in environment j. QEI for a QTL q' can be further modeled by regressing it on an environmental covariable, z_b: $(\text{GEI})_{ij} = x_{iq'}z_{jb}v_{q'b} + (\text{residual})_{ij}$. For multiple QTLs, this generalizes to $(\text{GEI})_{ij} = \Sigma_{q=1}^{Q}\Sigma_{b=1}^{B} x_{iq}z_{jb}v_{qb} + (\text{residual})_{ij}$.

Testing one or more QTL main effects can be done by comparing the model $\bar{y}_{ij} = \mu + \Sigma_{q=1}^{Q} x_{iq}\rho_q + E_j + (\text{residual})_{ij}$ with the model $\bar{y}_{ij} = \mu + E_j + (\text{residual})_{ij}$. When main effect QTL expression and QEI areconsidered together, this is equivalent to fitting different

QTLs for each environment. Testing multiple QTLs with different effects for each environment, is done by fitting the model $\bar{y}_{ij} = \mu + \Sigma_{q=1}^{Q} x_{iq}\rho_{jq} + E_j +$ (residual)$_{ij}$ and comparing it with the model $\bar{y}_{ij} = \mu + E_j +$ (residual)$_{ij}$. A specific test for QEI compares $\bar{y}_{ij} = \mu + \Sigma_{q=1}^{Q} x_{iq}\rho_q + E_j + \Sigma_{q=1}^{Q} x_{iq}\rho_{jq} +$ (residual)$_{ij}$ with $\bar{y}_{ij} = \mu + \Sigma_{q=1}^{Q} x_{iq}\rho_q + E_j +$(residual)$_{ij}$. F-tests can be constructed from ratios of regression mean squares to independent error terms. In this study the error term used was computed as the median of the individual trial error terms. Since tests are performed at every position in the genome, the genome-wise error rate must be controlled. As the tests at nearby positions will be correlated, it is difficult to control the genome-wise error rate by Bonferroni corrections. Therefore, we elaborated a testing procedure based on randomization.

Randomization test

The strategy proposed by Van Eeuwijk et al. (2000, 2002) was inspired by the Churchill and Doerge (1994) method of randomization of the response vector for controlling Type I error in QTL detection. It also uses ideas from Manly (1997) on the randomization tests for two-way analysis of variance and multiple regression.

The approach consists of computing, first, for all evaluation positions at the chromosome under study, the F statistics (or, equivalently, a LOD score or R^2) for the QTL main effect, QEI, and QTL + QEI (=QTL with different effect in each environment). The distribution of the statistic for testing the null hypothesis of no QTL, no QEI, and no QTL + QEI is obtained from randomizations of the set of vectors of genetic predictors per progeny (F_2, RIL, backcross, etc.) with respect to the set of vectors of phenotypic observations across environments. This means that the genetic predictor values for a progeny family are kept together, just as with the phenotypic observations. However, the coupling of genetic predictor values and phenotypic values is random, so that the genetic predictor values for one family are joined with the phenotypic observations for another family. Per randomization, the maximum value of the test-statistic over the chromosome under study is stored. After 1,000 randomizations, the distribution of the test-statistic can be constructed from the realized values in the randomization. For QTL testing purposes, the 95th and/or 99th quantiles can then be taken to perform tests at a test level of 0.05 or 0.01 genome-wise, respectively. The F-values computed on the original, non-randomized data are then compared with the threshold values obtained from the randomization distribution.

Correcting for QTLs at other chromosomes using PLS

Correction for genetic effects (QTLs) outside the evaluation chromosome can be achieved by application of a multivariate PLS regression (Aastveit and Martens 1986; Helland 1988) of the vectors of phenotypic observations across environments on vectors of genetic predictors, where the latter correspond to the complete set of markers on all chromosomes except the evaluation chromosome. For example, for chromosome 1, the set of genetic predictors in the PLS regression will involve all markers on the chromosomes 2–10. The genetic predictors include both additive and dominance effects, and these effects are allowed to depend on the environment. The residuals from the PLS regressions are subsequently analyzed for QTLs by a SIM procedure, in the expectation that the effects of QTLs on other chromosomes were removed by the PLS regression.

This PLS approach avoids the complex construction of a set of cofactors and it is also supposed to deal adequately with the collinearity problem among genetic predictors (markers). The appropriate rank of the matrix of genetic predictors can be assessed by cross validation (Osten 1988). The randomization tests for QTL detection should be performed on the phenotypic data adjusted for the QTLs elsewhere.

Plant material and genetic mapping

Details of the phenotypic and genetic data analyzed in this study using the approach described above can be found in Ribaut et al. (1996, 1997). Phenotypic data were collected on a set of 211 $F_{2:3}$ families derived by selfing F_2 plants. The two parental lines crossed to obtain this population were P1, Ac7643S_5, derived from Population 43 (La Posta), and P2, Ac7729/TZSRWS_5, from Population 29 (Tuxpeño Caribe). The drought-tolerant inbred line was parent P1, which had a short ASI and performs well under drought, whereas parent P2 was the drought-susceptible. Segregating families were evaluated under eight environments during 1992, 1994, and 1996 under "optimal", water-limited, and low-nitrogen conditions. For each experiment, environmental variables such as solar radiation, minimum and maximum temperatures, precipitation, and sun hours were recorded during three development stages of the crop: vegetative (before flowering), flowering, and grain filling period. In this study, we did the QTL and QEI mapping for grain yield (GY) and ASI (measured as the difference in days between pollen shed and silk emergence). Table 1 gives a brief description of the experiments with their means for the two target traits GY and ASI.

The genetic predictors were calculated according to Jiang and Zeng (1997). A total of 132 RFLP markers distributed along the entire maize genome were used (Ribaut et al. 1996). The FRs were done every 3.33 cM and at the markers. At each position, the amount of variability accounted for by the putative QTL and QEI was calculated and expressed as a percent of the total variation due to G and GEI. We followed the same approach to calculate the percentage of variation accounted for by the main effect of QTL, R^2_{QTL}, the QEI,

Table 1 Trial code, year of the trial, time of sowing, levels of nitrogen and drought, and average grain yield and ASI

Trial code	Year	Sowing	Nitrogen	Drought stress	Mean yield (ton ha^{-1})	ASI (days)
NS92a	1992	Winter	Normal	No	10.5	-1.6
IS92a	1992	Winter	Normal	Intermediate	6.4	-1.0
SS92a	1992	Winter	Normal	Severe	3.7	-0.9
IS94a	1994	Winter	Normal	Intermediate	4.2	1.8
SS94a	1994	Winter	Normal	Severe	4.1	1.9
LN96a	1996	Winter	Low	No	1.8	2.9
LN96b	1996	Summer	Low	No	1.0	3.3
HN96b	1996	Summer	High	No	4.9	-1.1

NS no drought stress, *IS* Intermediate drought stress, *SS* severe drought stress, *LN* low nitrogen, *HN* high nitrogen, *ASI* anthesis silking interval

R^2_{QEI}, and the variation due to fitting QTL in each environment, $R^2_{QTL+QEI}$.

Results

Mapping additive and dominant QTLs for grain yield

To illustrate the new methodological approach presented in this paper, we discuss results of QTL for grain yield and ASI measured across several environments. These traits were selected among a broad set of traits as a grain yield increase represents the final objective of most breeding efforts conducted at CIMMYT, while ASI is a key secondary trait for maize improvement under both water-limited and low-nitrogen stress conditions. A QTL in a chromosome region was considered to be environment specific if both terms QTL + QEI and QEI were significant. On the other hand, when QTL + QEI was significant but the QEI was not, then it was concluded that this QTL has an expression only through its main effect. As confidence intervals for QTL location easily reach 20 cM, we treated QTLs within 20 cM as constituting the same QTL. For grain yield and ASI we followed the idea that a QTL was considered environment-specific when both tests for QTL + QEI and QEI were significant. Alternatively, a QTL had a main effect expression when QTL + QEI was significant, but QEI was not.

For grain yield, Fig. 1a–f depict the profile of R^2_{QTL}, R^2_{QEI}, and $R^2_{QTL+QEI}$ and the corresponding critical values for $\alpha = 0.01$ based on 1,000 randomizations. Between 105 and 180 cM of chromosome 1 (Fig. 1a) and between 45 and 90 cM of chromosome 10 (Fig. 1f) there are good reasons to believe that there are environment-specific QTLs (the QTL + QEI and the QEI effects were both significant). On the contrary, in the regions between 20 and 50 cM of chromosome 3 (Fig. 1b) and between 60 and 100 cM of chromosome 8 (Fig. 1d) only main effect QTLs were observed. QEI at the end of chromosome 4 (Fig. 1c) and near the end at chromosome 9 (Fig. 1e) was ignored because those QEI peaks did not coincide with the corresponding peaks for QTL + QEI. At chromosome 4 a significant dominance main effect QTL was also found (not shown in Fig. 1).

The results of the QTLs found for grain yield are summarized in Table 2, where the sign of the QTL effect is positive when yield was increased by the allele coming from the drought-tolerant parent, P1, and negative when the yield increasing allele came from the drought-susceptible parent, P2. Thus, for the additive main effect of QTLs on chromosomes 1, 3, 9, and 10, the positive allele came from the susceptible parent, whereas for the QTLs on chromosomes 4 (for both additivity and dominance) and 8, the positive allele came from the drought-tolerant parent (P1). For comparison, Table 2 also shows results of Ribaut et al. (1997) for some low-nitrogen and drought environments.

Mapping additive and dominant QTLs for ASI

Several QTLs have been identified for ASI. Figure 2a–f shows, for ASI, the profile of R^2_{QTL}, R^2_{QEI}, and $R^2_{QTL+QEI}$ and the corresponding critical values for $\alpha = 0.01$ based on 1,000 randomizations. Main effect QTLs with additive genetic effects were found on chromosomes 1, 2, 6, 8, and 10. In addition, on chromosome 1 a main effect QTL effect was detected for dominance. Significant QEI was observed at the end of chromosome 6 (Fig. 2d) and at the end of chromosome 8 (Fig. 2e). The QEI on chromosome 6 will not be considered further, because the peak of the QEI did not coincide with that of the QTL + QEI.

Table 3 summarizes location and effects of the ASI QTLs. The additive main effects of the QTLs on chromosomes 1 (for dominance), 2, 6, 8, and 10 follow the direction of the drought-tolerant parent (P1) that decreases the ASI. Only the QTL on chromosome 1, with an additive genetic effect, made the ASI wider and came from the susceptible parent (P2). Results from Ribaut et al. (1997) are included for comparison.

Interpreting QTL × E for grain yield

In this section we concentrate on positions between 105 and 180 cM of chromosome 1 (Fig. 1a) and between 45 and 90 cM of chromosome 10 (Fig. 1f), where QEI effects for yield were significant. The analysis of variance

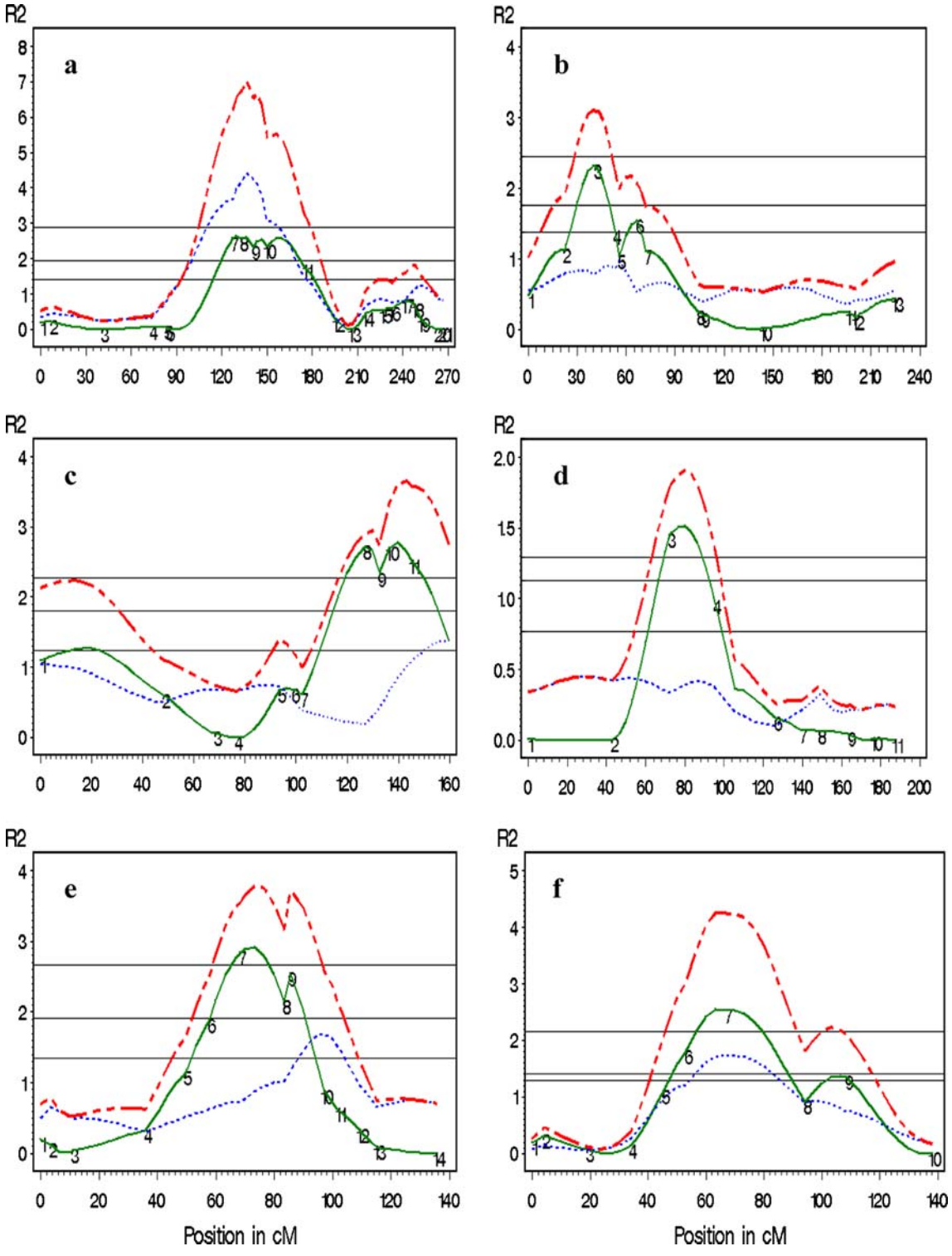

Fig. 1 Profile of R^2 for the additive effects of QTL (*Solid lines*), QEI (*dotted lines*), and QTL + QEI (*broken lines*) on grain yield for **(a)** chromosome 1 (*additive*); **(b)** chromosome 3 (*additive*); **(c)** chromosome 4 (additive); **(d)** chromosome 8 (*additive*); **(e)** chromosome 9 (*additive*); **(f)** chromosome 10 (*additive*). *Horizontal lines* from *top* to *bottom* represent the critical values of QTL + QEI, QTL and QEI, respectively, after 1,000 randomizations ($\alpha = 0.01$). *Numbers* on the QTL profile indicate marker positions

table at position 140 cM of chromosome 1 is given in Table 4. The first part of Table 4 (first five rows) shows the usual analysis of variance for a two-way table of grain yield measured in 211 genotypes (F_2 derived F_3 families) evaluated in eight different environments (E) with the partitioning of the joint effect of G + GEI into G and GEI effects. Most of the variability is due to E, followed by the GEI effects, which is highly significant

Table 2 Number of chromosomes with QTLs showing additive (Add.) or dominance (Dom.) genetic effects for grain yield, position (cM), and R^2 (data adjusted for the effect of two PLS terms)

Chromosome	1 Add.	2 Add.	3 Add.	4 Add.	4 Dom.	6 Add.	7 Add.	8 Add.	9 Add.	10 Add.
Position (cM)	140		40	143	73			80	73	63
R^2	7.02		3.10	3.65	2.92			1.91	3.79	4.26
Additivity[a]	−0.27		−0.27	0.26	0.38			0.23	-0.29	−0.27
Direction[b]	P2		P2	P1	P1			P1	P2	P2
NS92a[c] (cM)	168	86						134		48
IS94a[c] (cM)	154, 229			14			74			59
SS94a[c] (cM)	82			114		57				60
IS94a + SS94a[c] (cM)	82, 156					57		73		61
LNa + LNb + HNb[d] (cM)	104, 234	96	42, 173	59, 119		58, 123			67	60

QTLs reported by Ribaut et al. (1997) are shown for comparison

[a]Additive effects are associated with the allele from the tolerant line (P_1). A positive value means that the P_1 allele increases the numeric value of grain yield

[b]Direction indicates the parental line which contributes to the increase of the numeric value of the trait

[c]Ribaut et al. (1997). *NS* no drought stress, *IS* Intermediate drought stress, *SS* severe drought stress, 1992 cycle a, 1994 cycle a

[d]*LNa* low nitrogen 1996 cycle a, *LNb* low nitrogen 1996 cycle b, *HN* high nitrogen 1996 cycle b

when tested against an error obtained from analyzing the individual trials. The middle part of Table 4 (rows 6–10) shows the variability due to QTL + QEI effects in other parts of the genome than chromosome 1 (i.e., due to QTLs on chromosomes 2–10), the variability due to G + GEI after correction for the QTLs on the other chromosomes, and the corresponding partitioning into G and GEI components. Approximately 28.8% of the original G + GEI was associated with QTLs on other chromosomes. (The degrees of freedom for G and GEI could have been adjusted for the corrections at the other chromosomes, but this was not done as it would not have any observable influence on subsequent parts of the analyses.) The last part of Table 4 shows the partitioning of the G + GEI adjusted for the QTLs on chromomosome 2–10 into variation due to QTL + QEI at position 140 cM of chromosome 1 and deviations from the QTL model. The eight environment-dependent QTLs do not seem to be responsible for a major amount of variation, QTL + QEI variability explained only 6.7% of the G + GEI variability (153.775/2287.0623) in yield. QEI dominated the QTL main effect, when comparing the sums of squares, the QTL main effect explained 35.8% of the QTL + QEI variability, while QEI explained 64.2%.

Differences in QTL expression for the additive genetic effect at position 140 cM of chromosome 1 across the eight environments were clarified by studying the effect for the QTL and QEI (Table 5). Effects of allele substitutions varied between −0.234 and −0.721 ton ha^{-1} for the drought stress trials carried out during 1992 and 1994. The QTL effect in the low-nitrogen experiment of 1996 was effectively zero (−0.019 ton ha^{-1}), while under high-nitrogen conditions in that year it produced a grain yield increase of 0.424 ton ha^{-1}. The average additive effect (QTL main effects) of an allele substitution across all the trials was a yield decrease (−0.273 ton ha^{-1}) (Table 5). QEI effects were negative (or negligible) for drought trials in 1992 and 1994 (−0.310, −0.263, −0.448, −0.258, 0.039 ton ha^{-1}) and positive for nitrogen trials conducted in 1996 (0.289, 0.254, 0.697 ton ha^{-1}).

The environmental covariable that explained the QEI best, 77.6%, was minimum temperature during flowering time (Table 4). The effect of this environmental covariable was highly significant by an *F*-test for the regression mean square over the error obtained from the analysis of individual trials ($F = 76.675/0.75 = 102.23$, $P < 0.001$). The QTL allele coming from parent P1, the drought-tolerant parent, raised yield by 0.065 ton ha^{-1} for each degree Celsius that the minimum temperature at flowering time increased (Table 5). For genotypes that are homozygous at this QTL, the increase doubled. In the fifth column of Table 5, the fitted values of the FR for QEI are expressed as the product of the regression coefficient (0.065) and the value for the minimum temperature during flowering in degree Celsius given as deviation from the overall temperature mean (across the eight environments). Thus, the relative effect of the minimum temperature at flowering of QEI in each environment for grain yield is visualized through the sign and magnitude of the product. The highest negative effect was in the severe water stress in 1992 (SS92a) environment and the highest positive effect was observed in the low-nitrogen environment in 1996 (LN96b) (Table 5). This model did not fit equally well for each environment, as can be seen in the last column of Table 5, where the residual QEI effects are given.

The complete analysis for grain yield QTLs at position 63 cM of chromosome 10 is shown in Table 6. The first part of Table 6 (first five rows) is analogous to that of Table 4. The middle part of Table 6 (rows 6–10) shows the variability of QTL + QEI effects due to genetic predictors at chromosomes 1–9 as well as the remaining G + GEI variability. Approximately 31.4% of the original variability of the grain yield due to G + GEI was associated with the QTLs on the other chromosomes. The QTL + QEI variability explained only 4.3% of the adjusted G + GEI variability (93.87/2203.99). The QTL main effect explained 59.8% of the

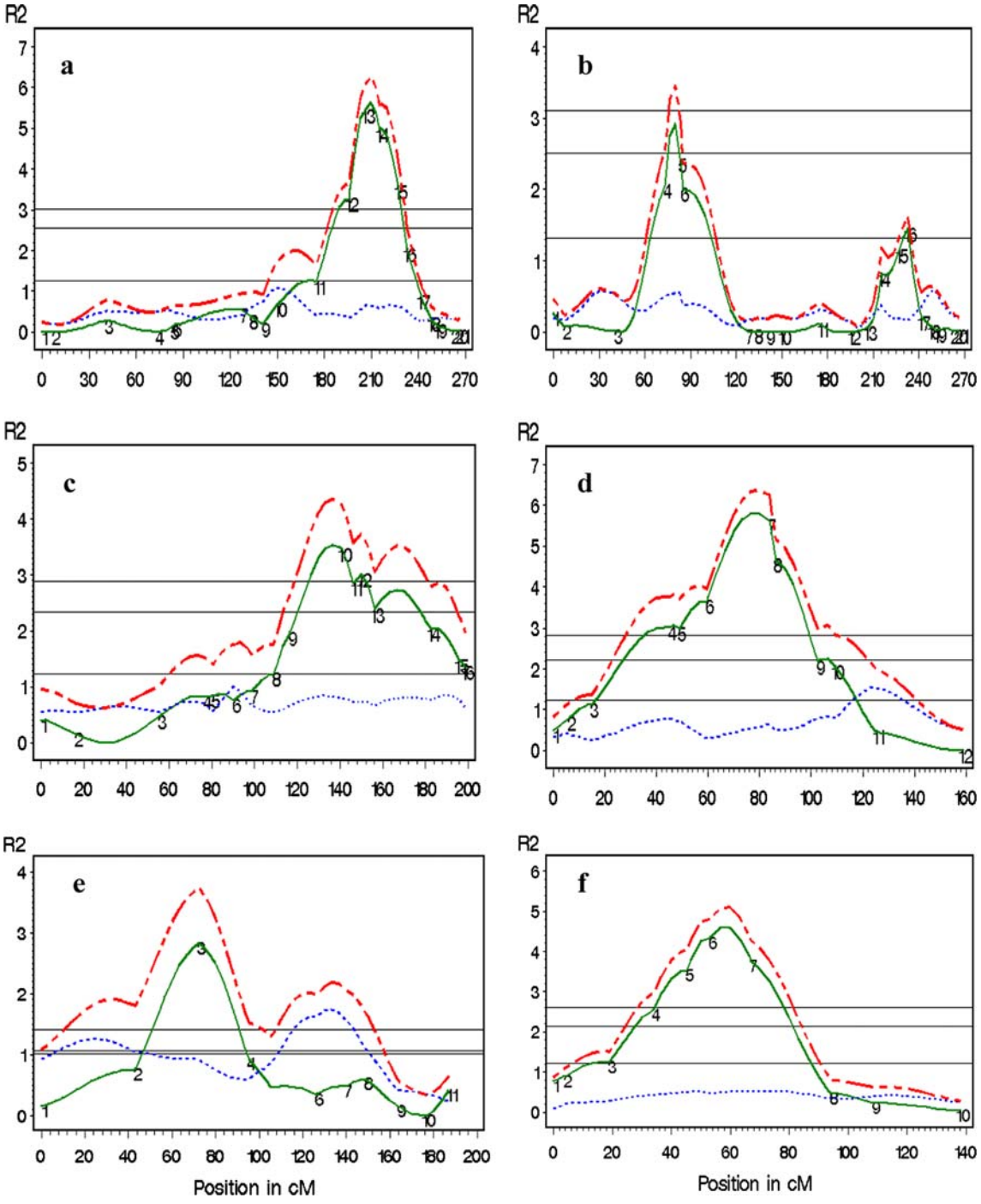

Fig. 2 Profile of R^2 for the effects of QTL (*Solid lines*), QEI (*dotted lines*), and QTL + QEI (*broken lines*) on ASI for **(a)** chromosome 1 (*additive*); **(b)** chromosome 1 (*dominance*); **(c)** chromosome 2 (*additive*); **(d)** chromosome 6 (*additive*); **(e)** chromosome 8 (*additive*); **(f)** chromosome 10 (*additive*). *Horizontal lines* from *top* to *bottom* represent the critical values of QTL + QEI, QTL and QEI, respectively, after 1,000 randomizations ($\alpha = 0.01$). *Numbers* on the QTL profile indicate marker positions

QTL + QEI variability, while QEI explained 40.2% (Table 6).

For the QTL at 63 cM on chromosome 10, the allele coming from the drought-tolerant parent, P1, reduced grain yield in all trials carried out during 1992, 1994, and 1996 (Table 7). The average additive effect of an allele substitution (the QTL main effect) produced a yield decrease of 0.278 ton ha^{-1} (Table 7). When the QEI effects were regressed on the set of environmental covariables, maximum temperature at flowering showed the closest relation with these interaction effects, although only 23.8% of the QEI was explained by this regression. The *F*-test produced significance at a test level of 0.05 ($P = 0.0134$). The effect of the environment-specific

Table 3 Number of chromosomes with QTLs showing additive (Add.) or dominance (Dom.) genetic effects for ASI, position (cM), and R^2 (data adjusted for the effect of one PLS term)

Chromosome	1 Add.	1 Dom.	2 Add.	5 Add.	6 Add.	7 Add.	8 Add.	10 Add.
Position (cM)	210	80	137		78		72	60
R^2	6.27	3.45	4.35		6.37		3.73	5.10
Additivity[a] (days)	0.55	−0.55	−0.45		−0.53		−0.35	−0.48
Direction[b]	P2	P1	P1		P1		P1	P1
NS92a[c] (cM)	201		138		90			61
IS94a[c] (cM)	203		134	147	76		76	101
SS94a[c] (cM)	206		130	147	79		75	43
LNa + LNb + HNb[d] (cM)	81, 208		90		71, 116	95	69, 124	45

QTLs reported by Ribaut et al. (1996) are shown for comparison
[a]Additive effects are associated with the allele from the susceptible line (P_2). A positive value means that the P_2 allele increases the numeric value of ASI
[b]Direction indicates the parental line which contributes to the increase of the numeric value of the trait
[c]Ribaut et al. (1996). *NS* no drought stress, *IS* Intermediate drought stress, *SS* severe drought stress, 1992 cycle a, 1994 cycle a
[d]*LNa* low nitrogen 1996 cycle a, *LNb* low nitrogen 1996 cycle b, *HN* high nitrogen 1996 cycle b

Table 4 Partitioning of yield variation at position 140 cM of chromosome 1

Source of variation	Degrees of freedom	Sum of squares	Mean Squares
Environment (E)	7	12777.169	1825.310
G+GEI	1680	3212.868	1.914
F2 family (G)	210	1382.102	6.581
GEI	1470	1829.700	1.245
Total	1687	15988.970	
G+GEI	1680	3212.868	1.914
QTL+QEI Chrom. 2–10	–*	925.806	
G+GEI Chrom. 1 adj.	1680*	2287.062	
F2 family (G) adj.	210	693.358	3.302
GEI adj.	1470	1593.704	1.084
G+GEI Chrom. 1 adj.	1680*	2287.062	
QTL+QEI Chrom. 1 140 cM	8	153.775	19.222
QTL main effect	1	54.986	54.986
QEI	7	98.789	14.113
Min. Temp. Flow.	1	76.675	76.675
Residual QEI	6	22.114	3.686
Deviations	1672	2133.287	1.276

For comparison, an error estimated as the median of the individual trial error was 0.75
*For the correction of the grain yield data due to genetic effects on chromosomes 2–10, degrees of freedom might be discounted (see text)

QTL effects (QEI) is described as a decrease in grain yield of 0.064 ton ha^{-1} for each degree (Celsius) that the maximum temperature at flowering stage increased.

Interpreting QTL × E for ASI

A significant QEI was found for ASI at position 133 cM of chromosome 8. From the variability explained by G + GEI, 28.4% is associated with QTLs at other chromosomes. The QTL + QEI variability explained only 2.2% of the adjusted G + GEI variability (75.69/3458.52). The QEI effect dominated the QTL + QEI effect, when comparing the sums of squares. The QTL main effect explained 20.7% of the QTL + QEI variability, whereas QEI explained 79.3% of the QTL + QEI (data not shown). For the QTL main effect, the allele coming from the drought-tolerant parent (P1) reduced the length of ASI. The average additive effect of an allele substitution (the QTL main effect) produced an average decrease in ASI of 0.157 days. When the QEI effects were regressed on the set of environmental covariables, precipitation at flowering showed the closest relation with the QEI effects and explained 38% of the QEI. The effect of this environmental covariable was significant ($P = 0.0025$). The effect of the environment-specific QTL

Table 5 QTL effects, QTL main effect, QEI effects for grain yield per environment for position 140 cM of chromosome 1

Trial code	QTL per environment	QTL + QEI model		Factorial regression on minimum temperature during flowering for QEI	
	QTL effect	QTL main effect	QEI effect	Fit regression on min. temp. flow.[a]	ResidualQEI
NS92a	−0.583	−0.273	−0.310	0.065×−2.88	-0.123
IS92a	−0.536	−0.273	−0.263	0.065×−3.76	0.019
SS92a	−0.721	−0.273	−0.448	0.065×4.76	-0.138
IS94a	−0.531	−0.273	−0.258	0.065×−3.08	-0.058
SS94a	−0.234	−0.273	0.039	0.065×−2.98	0.232
LN96a	0.158	-0.273	0.289	0.065×1.26	0.207
LN96b	−0.019	−0.273	0.254	0.065×8.12	-0.274
HN96b	0.424	-0.273	0.697	0.065×8.09	0.172
Standard error	0.132	0.037	0.105	0.0073[b]	0.102

Regression of QEI effect on minimum (min.) temperature (temp.) during flowering (flow.) using factorial regression
[a]The fit for the factorial regression model for QEI is expressed as the product of the regression coefficient and the value for the minimum temperature during flowering in degree Celsius, the latter given as the deviation from the mean for that temperature (across the eight environments)
[b]Standard error of slope

Table 6 Partitioning of yield variation at position 63 cM on chromosome 10

Source of variation	Degrees of freedom	Sum of squares	Mean Squares	
Environment (E)	7	12777.169	1825.310	
G+GEI	1680	3212.868	1.914	
F2 family (G)	210	1382.102	6.581	
GEI	1470	1829.700	1.245	
Total	1687	15988.970		
G+GEI	1680	3212.868	1.914	
QTL+QEI Chrom. 1-9	-*	1008.879		
G+GEI Chrom. 10 adj.	1680*	2203.988		
F2 family (G) adj.	210	666.755	3.175	
GEI adj.	1470	1537.234	1.046	
G+GEI Chrom. 10 adj.	1680*	2203.988		
QTL+QEI Chrom. 10 63 cM	8	93.868	11.733	
QTL main effect	1	56.148	56.148	
QEI	7	37.720	5.388	
Max. Temp. Flow.	1	8.986	8.986	
Residual QEI	6	28.720	4.787	
Deviations	1672	2110.121	1.262	

For comparison, an error estimated as the median of the individual trial error was 0.75
*For the correction of the grain yield data due to genetic effects on chromosomes 1 through 9, degrees of freedom might be discounted (see text)

effects (QEI) can be described as an average decrease in ASI of 0.002 days for the increase of 1 mm precipitation during the flowering stage.

Discussion

Comparing the QTLs found in this study for grain yield with those reported by Ribaut et al. (1997) for the NS92a, IS94a, and SS94a environments and for IS94a + SS94a and LN96a + LN96b + HN96b combined analysis, the presence of some common QTLs on almost all of the chromosomes can be observed. Some of them were located practically at the same positions and others were located at different positions, probably due to the different compositions of the sets of environments used in the various studies, in combination with differences in the applied QTL methodology (Table 2). The most stable QTLs were that found on chromosome 10 between 45 and 90 cM (with the peak at 63 cM) and that found by Ribaut et al. (1997) in 48, 59, 60, 61, and 60 cM of chromosome 10 in environments NS92a, IS94a, SS94a, IS94a + SS94a, and LN96a + LN96b + HN96b, respectively (Table 2 and

Table 7 QTL effects, QTL main effect, QEI effects for grain yield per environment for position 63 cM of chromosome 10

Trial code	QTL per environment	QTL + QEI model		Factorial regression on maximum temperature during flowering for QEI	
	QTL effect	QTL main effect	QEI	Fit regression on max. temp. flow.[a]	Residual QEI
NS92a	−0.355	−0.278	−0.077	−0.064×0.608	−0.038
IS92a	−0.393	−0.278	−0.115	−0.064×0.087	−0.109
SS92a	−0.034	−0.278	0.244	-0.064×−0.873	0.188
IS94a	−0.237	−0.278	0.041	-0.064×1.358	0.129
SS94a	−0.339	−0.278	−0.062	−0.064×0.997	0.003
LN96a	−0.068	−0.278	0.210	-0.064×−4.213	−0.061
LN96b	−0.041	−0.278	0.237	-0.064×0.837	0.291
HN96b	−0.757	−0.278	−0.479	−0.064×1.197	−0.402
Standard error		0.037	0.106	0.023[b]	0.110

Regression of QEI effect on maximum (max.) temperature (temp.) during flowering (flow.) using factorial regression
[a]The fit for the factorial regression model for QEI is expressed as the product of the regression coefficient and the value for the maximum temperature during flowering in degree Celsius, the latter given as the deviation from the mean for that temperature (across the eight environments)
[b]Standard error of slope

Fig. 4). For chromosome 1, a significant region for yield was found between 105 and 180 cM (with a peak at 140 cM) (Table 2), whereas Ribaut et al. (1997) found QTLs at different places of chromosome 1 between 82 and 234 cM, depending on the environment where the progenies were evaluated (Table 2 and Fig. 3).

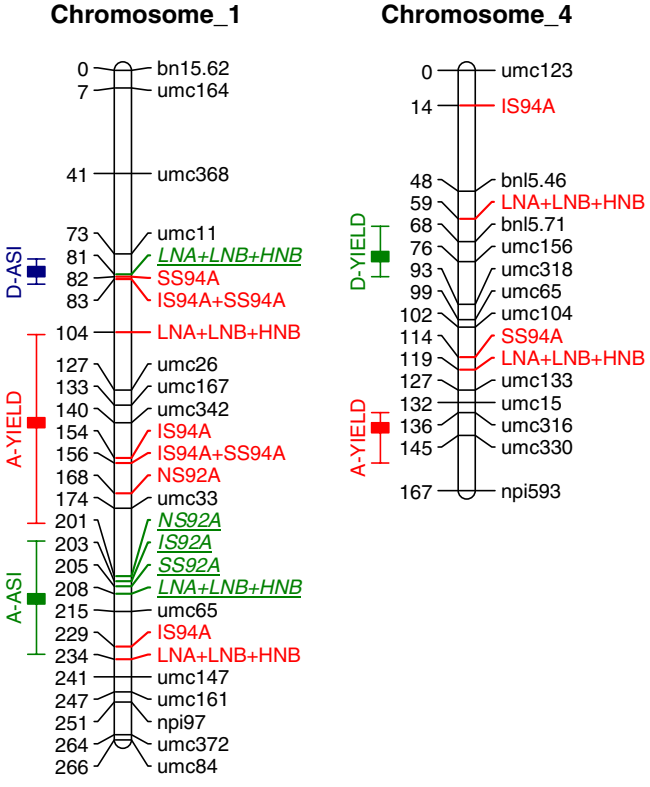

Fig. 3 Positions of additive and dominant QTL + QEI effects for grain yield and ASI for chromosomes 1 and 4. Along the *left-hand side* are the positions of the QTLs found in this study; in *red* are the QTLs for grain yield (A-YIELD, D-YIELD, with A for additive genetic effects and D for dominance) and in *green* the QTLs for ASI (A-ASI and D-ASI). The intervals for the QTLs show the region for which the test statistic was significant. Along the *right-hand side* are the positions of the QTLs found by Ribaut et al. (1996, 1997); *red* denotes the QTLs for grain yield identified by specific trials and the *underlined green italic* identifies the trials where the QTLs for ASI were found

For ASI the number and positions of the QTLs found in this study agreed very well with those identified by Ribaut et al. (1996) (Table 3 and Figs. 3, 4), indicating that QTLs for ASI are stable across the different environments. Contrary to the case of grain yield, for ASI the main effect of QTL at different chromosomes showed a more stable behavior. Only for the additive QTL on chromosome 1, the allele from P2, the susceptible line, produced an increase in ASI, whereas, as expected, P1 alleles reduced ASI for all other significant QTLs (Table 3). The sign of the additivity was consistent for all ASI QTLs across the eight environments. The significant QEI for ASI found in chromosome 8 was well explained by precipitation during flowering stage.

The above results indicate that QTLs for grain yield are less stable than those identified for ASI, as they are more affected by different environmental conditions. The location and the effects of the QTLs for grain yield reported by Ribaut et al. (1997) and the present study vary depending with the nature of the stress used in the trial. The most important consequence of drought stress is a decrease in yield and an increase in GEI (Blum 1988). Thus, as pointed out by Ribaut et al. (1997), inconsistencies in the identification of QTLs for grain yield performance across well-watered and drought environments are expected. However, we also expect to find stable QTLs across varying water regimes due to the spillover effects of yield potential. It is likely to find genotypes performing well under well-watered conditions and under drought, even if the relative yield reduction is large (Edmeades et al. 2001). Beavis (1994) mentioned that the inconsistency of grain yield QTLs across environments can be explained by the fact that yield is under the control of a large number of small-effect QTLs that segregate in the genome.

Concerning the interpretation of QEI for grain yield occurring in chromosomes 1 and 10, in terms of

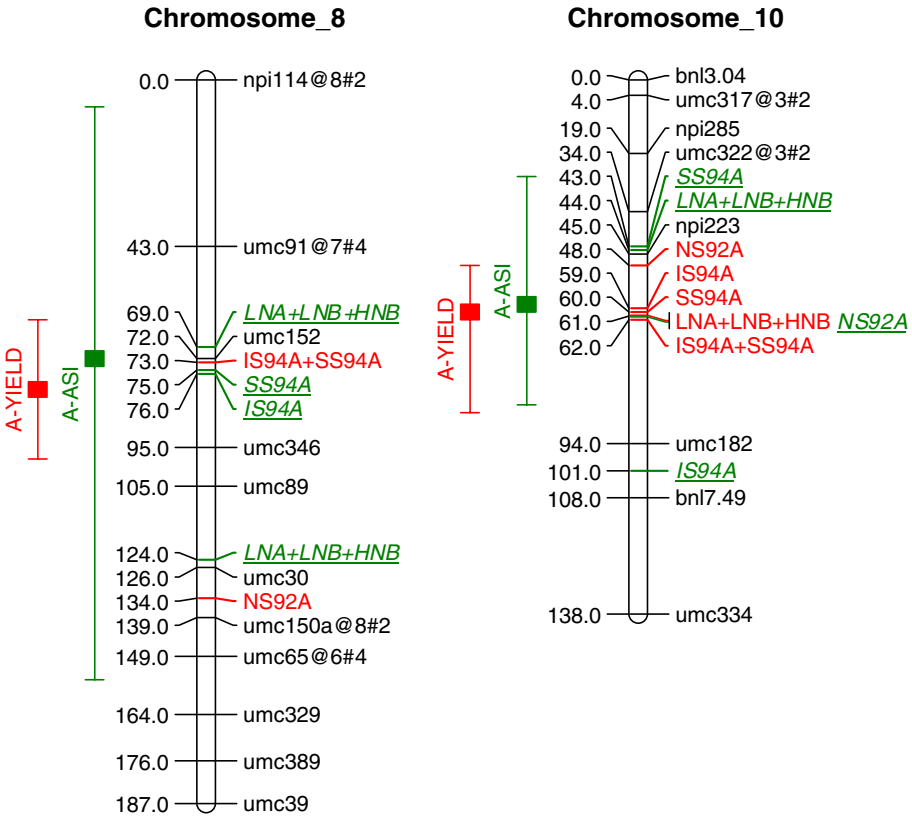

Fig. 4 Positions of additive QTL + QEI effects for grain yield and ASI for chromosomes 8 and 10. Along the *left-hand side* are the positions of the QTLs found in this study; in *red* are the additive QTLs for grain yield (A-YIELD) and in green the additive QTLs for ASI (A-ASI). The intervals for the QTLs show the region for which the test statistic was significant. Along the *right-hand side* are the positions of the QTLs found by Ribaut et al. (1996, 1997); *red* denotes the QTLs for grain yield identified by specific trials and the *underlined green italic* identifies the trials where the QTLs for ASI were found

environmental variables it is clear that temperatures during the flowering stage were the most important factors affecting QEI. Because of the highly polygenic nature of grain yield and large transgressive segregation, the contribution of favorable alleles was relatively balanced between the two parental lines. P2 contributed to an increase of yield at QTLs on chromosomes 1, 3, 9, and 10, while P1 increased yield at QTLs on chromosome 4 (additivity and dominance) and 8 (Table 2). In maize, drought during flowering produces a delay in silking, and thus an increase in the QTL for ASI on chromosome 1 from parent P2 (Ribaut et al. 1996; Hall et al. 1982; Westgate and Bassetti 1990; Bolaños and Edmeades 1993).

It is interesting that there is no evidence of closely linked or pleiotropic additive genetic QTLs for grain yield and ASI on chromosome 1 (Fig. 3). Similarly, there is no evidence for joint additive and dominance genetic effects at the QTLs for ASI on chromosome 1. There is one QTL with an additive effect on grain yield on chromosome 4 (Fig. 3) at 143 cM and another one with a dominance effect at 73 cM, whereas there is no evidence of QTLs for ASI on this chromosome. On chromosomes 8 and 10, QTLs for grain yield and ASI coincide closely with respect to location. Chromosome 8 had an additive QTL for grain yield (at 80 cM) and one for ASI (at 72 cM); chromosome 10 had an additive QTL for grain yield at 63 cM and an additive QTL for ASI at 60 cM (Fig. 4).

Apart from the understanding of the genetic complexity of a target trait, the major output of a QTL analysis is the identification of suitable genomic regions to be included in a marker-assisted selection (MAS) breeding program. Currently, the major limitation of a MAS program is the "instability" of QTL expression across environments and across genetic backgrounds (Ribaut and Hoisington 1998). As demonstrated through studies, there is real gain in accelerating breeding activities using molecular markers as a complementary tool (Morris et al. 2003), but few practical examples have been reported so far. The likely reason for the latter observation is that QTLs of complex traits such as grain yield explain only small amounts of the phenotypic variation, so that it can be expected that predictions based on those models do not perform very well. Therefore, QEI estimation, as presented in this paper, represents a key step forward toward the identification of target genomic regions for MAS experiments. As discussed, QTL identified for ASI were consistent across the eight environment trials. This result is very relevant for MAS purposes because it indicates that by pyramiding favorable alleles for ASI at significant loci identified in this study, we can expect to achieve genetic gain within the P1 × P2 genetic background across a broad set of environments including optimal and limited-water conditions and low-nitrogen conditions. For grain yield, a much larger QEI than for ASI has been observed. However, a couple of QTLs have been

consistently identified across environments that might be considered for inclusion in a MAS experiment.

The challenge for a MAS program is to predict which genotypes have alleles that will rank them as elite, and take advantage of existing genetic information generated through different genomics approaches without having to map QTLs in new crosses (a very time-consuming and expensive activity). This type of MAS, based on consensus regions, is already being tested at CIMMYT (Ribaut et al. 2004). In this context, the identification of QTL that are "stable" across environments represents a key element of success, and thus the QEI characterization, by assessing which environmental factors affect the QTL stability (such as maximum and minimum temperature during flowering found in this study) allows us, to a certain extent, to make predictions of genetic effects in new environments with comparable climatological patterns.

In conclusion, results of this study show the possibilities of using FR for mapping QTL and for dissecting QEI in terms of environmental variables affecting certain QTLs. Furthermore, PLS regression proved to be a convenient tool for canceling background noise produced by other QTLs. The FR model framework with a randomization test for controlling the genome-wise error rate in conjunction with the use of the PLS for dealing with the problem of multicollinearity among cofactors and adjusting for all the markers outside the chromosome being evaluated, seems to be a useful strategy for mapping QTLs with additive and dominance effects and studying QEI in terms of external environmental variables. At present, we are working at extending the FR framework for GEI and QEI in the direction of mixed models in order to have more flexibility in modeling heterogeneity of genetic variance across environments and variations in genetic correlations between environments, which we currently had to ignore. A publication on a mixed model analysis of the QEI for the maize data used in this paper is in preparation.

References

Aastveit H, Martens H (1986) ANOVA interactions interpreted by partial least squares regression. Biometrics 42:829–844

Beavis WD, Smith OS, Grant D, Fincher R (1994) Identification of quantitative trait loci using a small sample of topcrossed and F_4 progeny from maize. Crop Sci 34:882–896

Bjornstad A, Westad F, Martens H (2004a) Analysis of genetic marker–phenotype relationships by jacknifed partial least squares regression (PLSR). Hereditas 141:149–165

Bjornstad A, Gronnerod S, Mac Key J, Tekauz A, Crossa J, Martens H (2004b) Resistance to barley scald (*Rhynchosporium secalis*) in the Ethiopian donor lines 'Steudelli' and 'Jet', analyzed by partial least squares regression and interval mapping. Hereditas 141:166–179

Blum A (1988) Plant breeding for stress environments. CRC Press, Boca Raton, FL

Bolaños J, Edmeades GO (1993) Eight cycles of selection for drought tolerance in lowland tropical maize. II. Responses in reproductive behavior. Field Crops Res 31:253–268

Caliński T, Kaczmarek Z, Krajewski P, Frova C, Sari-Gorla M (2000) A multivariate approach to the problem of QTL localization. Heredity 84:303–310

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Crossa J (1990) Statistical analyses of multilocation trials. Adv Agron 44:55–85

Crossa J, Vargas M, Van Eeuwijk FA, Jiang C, Edmeades GO, Hoisington D (1999) Interpreting genotype × environment interaction in tropical maize using linked molecular markers and environmental covariables. Theor Appl Genet 99:611–625

Denis J-B (1988) Two-way analysis using covariates. Statistics 19:123–132

Denis J-B (1991) Ajustement de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. Revue de Statistique Appliquée 39:5–24

Edmeades GO, Cooper M, Lafitte R, Zinselmeier C, Ribaut J-M, Habben JE, Löffler C and Bänziger M (2001) Abiotic stresses and staple crops. In: Nosberger J, Geiger HH, Struik PC (eds) Crop science: progress and prospects. Proceedings of the Third International Crops Science Congress, August 17–21, 2000, Germany. CABI, Wallingford, UK, pp 137–155

Hackett CA, Meyer RC, Thomas WTB (2001) Multi-trait QTL mapping in barley using multivariate regression. Genet Res 77:95–106

Haley CS, Knott SA (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69:315–324

Hall AJ, Vilella F, Trapani N, Chimenti C (1982) The effects of water stress and genotype on the dynamics of pollen-shedding and silking in maize. Field Crops Res 5:349–363

Hwang JTG, Nettleton D (2003) Principal components regression with data-chosen components and related methods. Technometrics 45:70–79

Helland IS (1988) On the structure of partial least squares. Comm Statist Simulation Comput 17:581–607

Jiang C, Zeng Z-B (1997) Mapping quantitative trait loci with dominant and missing markers in various crosses for two inbred lines. Genetica 101:47–58

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

Lynch M, Walsh B (1998) Genetics and the analysis of quantitative traits. Sinauer Assoc., Inc., Sunderland, pp 980

Manly BFJ (1997) Randomization, bootstrap and Monte Carlo methods in biology, 2nd edn. Chapman and Hall, London

Martínez O, Curnow RN (1992) Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theor Appl Genet 85:480–488

Montgomery DC, Peck EA (1982) Introduction to linear regression analysis. Wiley, New York

Morris M, Dreher K, Ribaut J-M, Khairallah M (2003) Money matters (II): costs of maize inbred line conversion schemes at CIMMYT using conventional and marker-assisted selection. Mol Breed 11:235–247

Osten DW (1988) Selection of optimal regression models via cross-validation. J Chemometrics 2:39–88

Ribaut J-M, Hoisington DA (1998) Marker-assisted selection: new tools and strategies. Trends Plant Sci 3:236–239

Ribaut J-M, Hoisington DA, Deutsch JA, Jiang C, Gonzalez-de-Leon D (1996) Identification of quantitative trait loci under drought conditions in tropical maize 1. Flowering parameters and the anthesis-silking-interval. Theor Appl Genet 92:905–914

Ribaut J-M, Jiang C, Gonzalez-de-Leon D, Edmeades GO, Hoisington DA (1997) Identification of quantitative trait loci under drought conditions in tropical maize 2. Yield components and marker-assisted selection strategies. Theor Appl Genet 94:887–896

Ribaut J-M, Bänziger M, Setter T, Edmeades G, Hoisington D (2004) Genetic dissection of drought tolerance in maize: a case study. In: Nguyen H, Blum A (eds) Physiology and biotechnology

integration for plant breeding. Marcel Dekker Inc., New York, pp 571–611

Sari-Gorla M, Caliński T, Kaczmarek Z, Krajewski P (1997) Detection of QTL × environment interaction in maize by a least squares interval mapping method. Heredity 78:146–157

Talbot M, Wheelwright AV (1989) The analysis of genotype × environment interactions by partial least squares regression. Biuletyn Oceny Odmian Zeszyt 21/22:19–25

Van Eeuwijk FA, Denis J-B, Kang MS (1996) Incorporating additional information on genotypes and environments in models for two-way genotype by environment tables. In: Kang MS, Gauch HG Jr (eds) Genotype-by-environment interaction: new perspectives. CRC Press, Boca Raton, FL, pp 15–49

Van Eeuwijk FA, Crossa J, Vargas M, Ribaut J-M (2000) Variants of factorial regression for analysing QTL by environment interaction In: Gallais A, Dillmann C, Goldringer I (eds) Eucarpia, quantitative genetics and breeding methods: the way ahead. INRA edn. Versailles. Les colloques 96:107–116

Van Eeuwijk FA, Crossa J, Vargas M, Ribaut J-M (2002) Analyzing QTL-environment interaction by factorial regression, with an application to the CIMMYT drought and low-nitrogen stress programme in maize. In: Kang MS (ed) Quantitative genetics, genomics and plant breeding. CABI International, New York

Vargas M, Crossa J, Sayre K, Reynolds M, Ramírez ME, Talbot M (1998) Interpreting genotype × environment interaction in wheat using partial least squares regression. Crop Sci 38:679–689

Vargas M, Crossa J, Van Eeuwijk FA, Ramírez ME, Sayre K (1999) Using AMMI, factorial regression, and partial least squares regression models for interpreting genotype × environment interaction. Crop Sci 39:955–967

Westgate ME, Bassetti P (1990) Heat and drought stress in corn: what really happens to the corn plant at pollination?. In: Wilkinson D (ed) Proceedings of the 45th Annual Corn and Sorghum Research Conference, ASTA, Washington, pp12–28

Zeng Z-B (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1468