CHAPTER 3

MULTI-TRAIT MULTI-ENVIRONMENT QTL MODELLING FOR DROUGHT-STRESS ADAPTATION IN MAIZE

M. MALOSETTI[#,###], J.M. RIBAUT[##], M. VARGAS[##], J. CROSSA[##],
M.P. BOER[#,###] AND F.A. VAN EEUWIJK[#,###]

[#] *Laboratory of Plant Breeding, Department of Plant Sciences, Wageningen University, P.O. Box 386, 6700 AJ Wageningen, The Netherlands.*
[##] *International Maize and Wheat Improvement Center (CIMMYT), 6-641, 06600 Mexico D.F., Mexico.*
[###] *Biometris, Wageningen University, P.O. Box 100, 6700 AC Wageningen, The Netherlands.*
E-mail: marcos.malosetti@wur.nl

Abstract. Water shortage is a major cause of yield loss in maize. Thus, breeding for adaptation to water-stressed environments is an important task for breeders. The use of quantitative-trait loci (QTL) models in which the response of complex phenotypes under stressed environments is described in direct relation to molecular information can improve the understanding of the genetic causes underlying stress tolerance. Mixed QTL models are particularly useful for this type of modelling, especially when the data stem from multi-environment evaluations including stressed and non-stressed conditions. The study of complex phenotypic traits such as yield under water-limited conditions can benefit from the analysis of trait components (e.g., yield components) that can be exploited in indirect selection.

Multi-trait multi-environment QTL models help to identify the genome regions responsible for genetic correlations, whether caused by pleiotropy or genetic linkage, and can show how genetic correlations depend on the environmental conditions. With the objective of identifying QTLs for adaptation to drought stress, we present the results of a multi-trait multi-environment QTL-modelling approach using data from the CIMMYT maize-breeding programme.

INTRODUCTION

Water shortage is a major cause of yield loss in maize (*Zea mays* L.). The supply of water by irrigation can alleviate drought stress, but irrigation is costly and not realistic in most of the maize production areas. Yield loss due to water stress can be

J.H.J. Spiertz, P.C. Struik and H.H. van Laar (eds.), Scale and Complexity in Plant Systems Research: Gene-Plant-Crop Relations, 25–36.

tackled by developing varieties better adapted to water-limited conditions. Thus, drought tolerance is a prime objective of many maize-breeding programmes.

Maize is particularly sensitive to drought stress occurring just before and during flowering when the crop's yield potential is defined. When drought stress occurs just before flowering, a delay of silk emergence in relation to male flowering is observed, and this increase of the anthesis-silking interval (ASI) is correlated with lower yield (Bolaños and Edmeades 1993). Therefore, the selection of genotypes that have a short ASI under water-limited conditions has been shown to be useful to improve drought tolerance in maize (Edmeades et al. 1999; 2000; Ribaut et al. 1996).

Quantitative-trait loci (QTLs) associated with drought tolerance can be used in breeding strategies for drought tolerance, especially for selection under stress conditions, where traits typically show low heritability (Ribaut et al. 1996). In QTL mapping, complex phenotypes are modelled in direct relation to molecular information contributing to the understanding of the genetic causes underlying stress tolerance. Mixed models offer a particularly useful statistical framework for QTL analysis (Malosetti et al. 2004), especially when the data stem from multi-environment evaluations including stressed and non-stressed conditions. The study of complex phenotypic responses, such as yield under water-limited conditions, can benefit from the study of their trait components, which can be exploited in indirect selection.

In contrast to single-trait single-environment QTL models, multi-trait multi-environment QTL models simultaneously fit QTLs as affecting several traits in several environments. The attractiveness of such models is that they can help to identify the genome regions responsible for genetic correlations between traits, say yield and its components, whether caused by pleiotropy or genetic linkage, and can show how these genetic correlations depend on the environmental conditions. With the objective of identifying QTLs for adaptation to drought stress, we present the results of a multi-trait multi-environment QTL-modelling approach using data on grain yield (GY) and ASI from the CIMMYT maize-breeding programme.

MATERIALS AND METHODS

Field data and molecular-marker data

The data used in this paper were generated at CIMMYT, Mexico, with the objective of detecting QTLs related to yield and other yield-related traits under stressed conditions. A detailed description of field experiments and production of molecular-marker information is given in Ribaut et al. (1996). Briefly, an F_2 population derived from the cross between a drought-resistant parent (Ac7643S_5) and a high-yielding but drought-susceptible parent (Ac7729/TZSRWS_5) was genotyped by RFLP markers. A population of 211 $F_{2:3}$ families derived from that F_2 was subsequently evaluated in three years under different water and nitrogen stress conditions in Mexico (Table 1). Several traits were registered, but in the present chapter we concentrate on GY and ASI.

Table 1. *Description of the environments in which maize genotypes were evaluated: environment codes, location, water regime, nitrogen availability, mean GY and ASI*

Environment	Location	Water regime ¶	Nitrogen	Mean GY ($t\ ha^{-1}$)	Mean ASI (days)
NS92a	Tlaltizapán	WW	normal	10.5	–1.6
IS92a	Tlaltizapán	IS	normal	6.4	–1.0
SS92a	Tlaltizapán	SS §	normal	3.7	–0.9
IS94a	Tlaltizapán	IS	normal	4.2	1.8
SS94a	Tlaltizapán	SS	normal	4.1	1.9
LN96a	Poza Rica	WW	low	1.8	2.9
HN96b	Poza Rica	WW	high	4.9	–1.1
LN96b	Poza Rica	WW	low	1.0	3.3

¶ WW: well watered; IS: intermediate stress; SS: severe stress

§ rainfall around flowering caused only intermediate stress

Multi-trait multi-environment phenotypic model

We first conducted an analysis without introducing molecular-marker information in the model. The multivariate multi-environment mixed model used was (random terms underlined):

$$\underline{y}_{ijt} = \mu_t + E_{jt} + \underline{G}_{it} + \underline{GE}_{ijt} + \underline{\varepsilon}_{ijt} \tag{1}$$

with $\underline{y}_{ijt}$ a vector containing the observations of genotypes (i=1…211), in each of the eight environments (j=1…8), and for the two traits (t=1…2); μ_t an intercept for each trait (overall trait means across genotypes and environments), E_{jt} the environmental effect (fixed), $\underline{G}_{it}$ and $\underline{GE}_{ijt}$ the trait-specific genotypic main effects and genotype-by-environment interaction (GEI) effects, respectively (both random terms), and finally a residual term, which we considered heterogeneous among environments. An unstructured variance–covariance matrix was assumed for the $\underline{G}_{it}$ term thus introducing genetic correlations between traits due to genotypic main effects. For the $\underline{GE}_{ijt}$ term we imposed a factor-analytic model of order 1 (FA1). This model allows parsimonious modelling of genotypic correlations between environments and traits, since it requires fewer parameters than an unstructured model (Smith et al. 2001). In summary, the mixed model as defined above, with an unstructured variance–covariance model for the genotypic main effects of the traits and a FA1 model for the GEI part, allows to consider heterogeneity of genetic variance for the traits across environments, genetic correlations between environments for the same trait, and genetic correlations between traits within and across environments. Residuals were not estimated directly from the analysis since the data consisted of genotypic means per trial, but estimates were available from previous analysis.

Genotypic and GEI variance components per trait were estimated based on Model (1) and used to estimate heritability and genotypic and phenotypic correlations between GY and ASI. The association between GY and ASI was graphically represented by a number of scatter plots. Further investigation of the GEI patterns was performed by plotting the factor loadings for both axes of a factor-analytic variance–covariance model of order 2 (FA2), which can be regarded as an analogue of the AMMI analysis within a mixed-model framework (Smith et al. 2001).

The multi-trait multi-environment QTL analysis

The multivariate mixed model previously described was upgraded to include molecular information, leading to the following QTL model:

$$\underline{y}_{ijt} = \mu_t + E_{jt} + x_i^{add}\alpha_{jt} + x_i^{dom}\delta_{jt} + \underline{G}_{it} + \underline{GE}_{ijt} + \underline{\varepsilon}_{ijt} \qquad (2)$$

where the two extra terms in the model (both fixed effects) account for environment-specific additive QTL (α_{jt}) and dominance QTL (δ_{jt}) effects on GY and ASI. The covariables x_i^{add} and x_i^{dom} are called genetic predictors and are a function of the inferred genotypic constitution of the QTL at one particular point on the chromosome (Jiang and Zeng 1997). In short, for a given genotype, the additive genetic predictor (x_i^{add}) had a value –1 when homozygous of the maternal type, 0 when heterozygous, and +1 when homozygous of the paternal type. The dominance genetic predictor (x_i^{dom}) had a value +1 when the genotype was heterozygous at the locus, and 0 otherwise. With the genetic predictors estimated along the chromosomes we fitted the model at the different chromosome positions. The fixed QTL effects were tested by a Wald test (Verbeke and Molenberghs 2000) and the test statistic was plotted along the chromosomes to produce an analogue to the LOD score profile usually presented in QTL-mapping results. A QTL was revealed by a peak value exceeding a threshold value defined to control for multiple testing. Note that the described test is a global test for the presence of a QTL, with an effect on GY, ASI or both. Therefore, at positions where the global test indicated a QTL, we subsequently estimated and tested for the specific effects on GY and ASI being different from zero, the equivalent of a t-test using estimated standard errors. We restrict the analysis here to chromosomes 1 and 10.

When a QTL was found significant for both GY and ASI, a second scan was performed to investigate whether a single pleiotropic QTL or two closely linked QTLs were involved. In the initial scan a pleiotropic model was assumed as the genetic predictors represented the genotypic constitution at the same position for both traits. However, in this later stage, we allowed the genetic predictors to represent different positions on the chromosomes within a window of 20 cM around the initial pleiotropic position. The result of the two-dimensional scan was plotted in a contour plot to identify the region where the maximum for the test statistic was located, with close-linkage detected when that maximum resided far from the diagonal of the plot.

RESULTS AND DISCUSSION

Phenotypic analysis

The genotypic and GEI variance components for both traits are presented in Table 2. As expected, genotypic-variance components were larger for ASI than for GY. Conversely, GEI was more important for GY than for ASI, although the latter showed considerable GEI in some of the environments (particularly the water-stressed trials in 1994 and the low-nitrogen trials in 1996). Estimated heritability per environment was similar for GY and ASI, an observation that was already mentioned in previous studies of this population (Ribaut et al. 1996) and other similar populations used in CIMMYT trials (Bolaños and Edmeades 1996).

Table 2. *Estimates of genotypic variance across environments (V_G), environment-specific genotypic variances ($V_{G(E)}$), and environment-specific error variances (V_E) for grain yield (GY) and anthesis-silking interval (ASI), plus estimates of heritability (h^2), and genotypic (r_G) and phenotypic (r_P) correlations between traits*

	GY (t ha^{-1})				ASI (days)					
	V_G	$V_{G(E)}$	V_E	h^2	V_G	$V_{G(E)}$	V_E	h^2	r_G	r_P
NS92a	0.09	2.71	2.07	0.58	0.96	0.06	1.63	0.38	–0.11	–0.05
IS92a	0.09	1.04	2.03	0.36	0.96	0.21	0.98	0.54	–0.16	–0.07
SS92a	0.09	0.92	1.43	0.41	0.96	0.13	1.27	0.46	–0.12	–0.05
IS94a	0.09	1.33	1.53	0.48	0.96	1.13	2.45	0.46	–0.57	–0.27
SS94a	0.09	1.39	1.52	0.49	0.96	1.36	4.06	0.36	–0.48	–0.20
LN96a	0.09	0.17	0.39	0.41	0.96	1.07	2.04	0.50	–0.01	–0.01
HN96b	0.09	1.56	0.97	0.63	0.96	0.46	0.23	0.86	0.00	0.00
LN96b	0.09	0.20	0.21	0.59	0.96	2.33	2.64	0.56	0.05	0.03

Figure 1 shows the patterns of GEI in the experiments. For interpretation, the lengths of the vectors representing the environments correspond to the amount of GEI in that environment. The (cosine of the) angle between environmental vectors is proportional to the correlation between the two environments with respect to the GEI. Acute angles represent high positive correlations, obtuse angles indicate high negative correlations, and right angles point to low correlations. For GY, GEI was mainly caused by the contrast between environments in 1992 and 1994 versus those in 1996. This pattern reflects the contrast between two different locations, Tlaltizapán and Poza Rica, that represent rather different growing environments for maize, especially in terms of temperatures and water availability (Tlaltizapán is a drier and cooler location than Poza Rica) (Edmeades et al. 1999). GEI for ASI was mainly caused by the contrast of the trials in 1994 versus the ones in 1992 and 1996. In this case, the contrast seems to reflect the effect of a water-stressed environment (1994) versus those that did not have or had mild water restrictions (though some of the environments in Poza Rica were stressed due to low nitrogen). We emphasize that in those trials of 1992 that were managed in such a way that water stress was supposed to occur, rainfall occurred around flowering and, therefore, the plants did not experience water shortage. This lack of water stress is reflected by the values for average ASI observed in those trials: these values were similar to the ones observed in environments in which water stress was not imposed (Table 1).

Figure 1 also reflects the correlations between GY and ASI, and shows that in most of the environments the association between both traits was rather low (right angle between vectors). The only example of a negative association between GY and ASI was observed in the trials of 1994, where an obtuse angle between GY and ASI vectors indicates a negative correlation (Figure 1). This conclusion is in agreement with the estimated genetic correlations between both traits (Table 2). The lack of association between GY and ASI in most of the environments is also evident from Figure 2, where only in the water-stressed trials of 1994 a moderate association is observed (Figure 2). This observation is consistent with previous results in which the correlation between GY and ASI was mainly observed in water-stressed environments (Bolaños and Edmeades 1996; Chapman and Edmeades 1999).

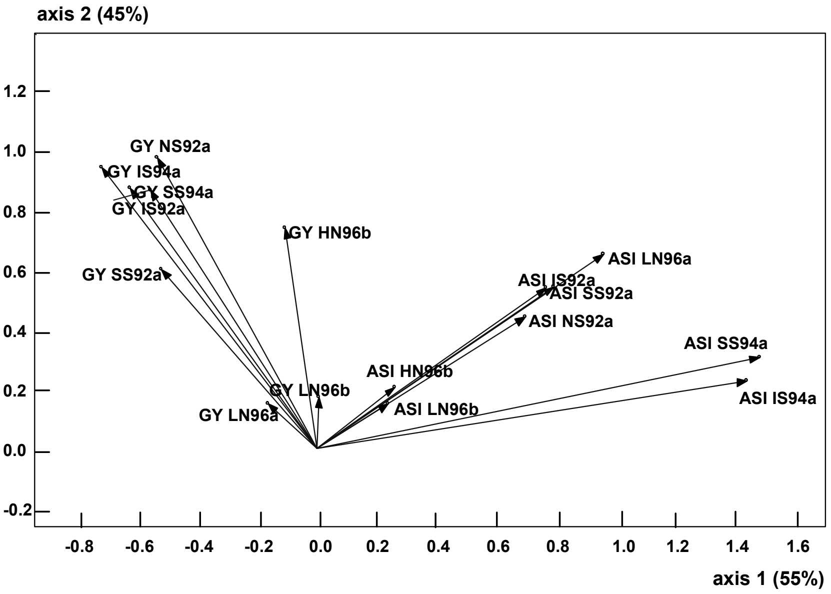

Figure 1. *Plot of loadings from a factor-analytic model of order 2 (FA2) fitted on maize trials carried out in eight environments in Mexico. The labels associated to each vector indicate the observed trait (GY or ASI) and environment*

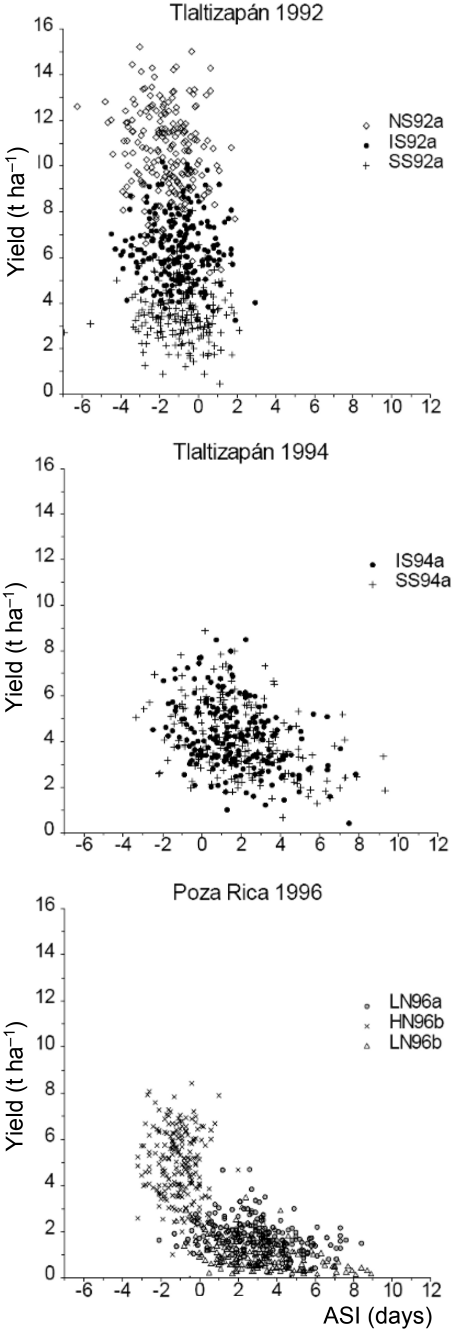

Figure 2. Yield versus ASI in eight environments, with environments grouped by year

QTL analysis

The profiles resulting from the QTL scan of chromosomes 1 and 10 are presented in Figure 3. The profiles show three regions where significant additive QTL effects were found, two on chromosome 1 and one on chromosome 10. The two regions on chromosome 1 were at 137 cM and at 215 cM and the one on chromosome 10 at 62 cM, which agreed with previous studies in the same population, although using different QTL models (Ribaut et al. 1996; Vargas et al. 2006). No significant dominant effects were found on any of the chromosomes.

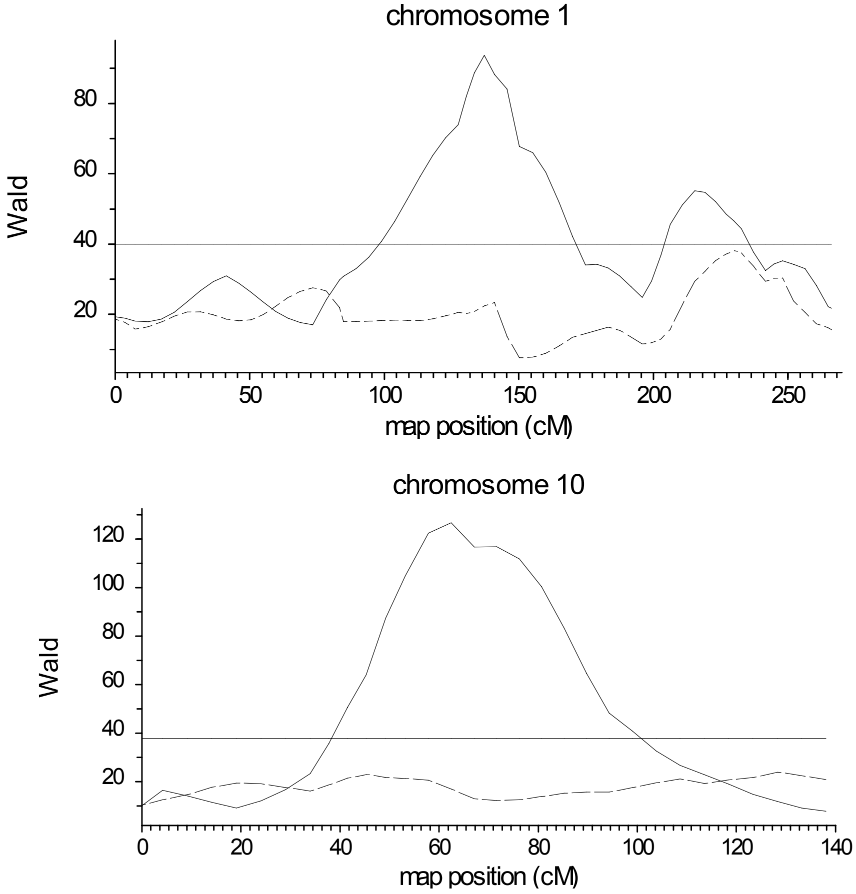

Figure 3. *Result of a multi-trait multi-environment QTL-mapping scan (simple interval mapping) on chromosomes 1 and 10 of maize. The profile represents the test statistic under the null hypothesis of no additive (solid line) or dominance (broken line) QTL effect on GY or ASI in any environment. The horizontal line represents a threshold above which the null hypothesis is rejected*

Two linked QTLs for GY and ASI on chromosome 1

The estimates of the QTL effects revealed that while the QTL at 137 cM on chromosome 1 had an effect only on GY, the one at 215 cM on the same chromosome had an effect exclusively on ASI (Table 3). The magnitudes and signs of the effects for GY reflect the higher GEI observed in this trait, as for example the allele coming from the high-yielding parent (the father) resulted in higher yields in four environments, but lower yield in the high-nitrogen environment of Poza Rica (Table 3), and no superiority in the rest of the environments. The effects for ASI were more consistent across the environments, with some variation in the magnitude, but not in sign (Table 3). Considering the signs of the effects of these two QTLs, the genetic correlation that those two QTLs induce is in the expected direction (except in HN96b), that is, a negative correlation between GY and ASI. However, the impact of this correlation might be low in view of the weak linkage between the QTLs (137 and 215 cM).

Table 3. *Environment-specific QTL effects for GY (t ha^{-1}) and ASI (days). A negative sign indicates that the high-value allele is coming from the maternal line (drought-resistant) and a positive sign indicates that the high-value allele is from the paternal line (high-yielding line)*

	chr 1, 137 cM		chr 1, 215 cM		chr 10, 62 cM	
Environment	GY	ASI	GY	ASI	GY	ASI
NS92a	0.65 *	–0.1	–0.26	–0.4 *	0.50 *	0.4 *
IS92a	0.63 *	–0.1	–0.12	–0.5 *	0.60 *	0.5 *
SS92a	0.82 *	–0.1	0.09	–0.6 *	0.19	0.5 *
IS94a	0.61 *	–0.3	–0.21	–0.7 *	0.47 *	0.5 *
SS94a	0.31	–0.3	0.00	–1.1 *	0.62 *	0.7 *
LN96a	0.01	0.3	0.01	–0.7 *	0.11	0.8 *
HN96b	–0.37 *	0.0	0.07	–0.4 *	0.97 *	0.2
LN96b	0.05	0.1	0.00	–0.5 *	0.10	0.5 *

* $P<0.05$

The QTL on chromosome 10

In contrast to chromosome 1, the QTL on chromosome 10 had a significant effect on both GY and ASI, in four of the eight environments (Table 3). Another remarkable difference was that the induced correlation was positive rather than negative. From a physiological point of view, a short ASI is an indicator of a better crop status (higher crop and ear growth rates), which relates to a higher yield (Edmeades et al. 2000; Westgate 2000). However, in our example, the allele coming from the high-yielding parent also caused a higher ASI value (Table 3), inducing a positive correlation between GY and ASI. On the one hand, and since GY is a complex trait determined by many processes during development, it is possible that the disadvantage of a longer ASI determined by this QTL is compensated by an advantage given by the same QTL at a later developmental stage, e.g., grain filling. On the other hand, this result may point to less-explored physiological mechanisms, which determine the increase of both GY and ASI. From a breeder's point of view, our result suggests that phenotypic selection for short ASI will not necessarily retain all positive alleles for GY, which highlights the potential of marker-assisted selection as a complement to conventional phenotypic selection.

A relevant question that follows from the results found for chromosome 10 is whether pleiotropy or genetic linkage is present. We addressed this question by refitting the model allowing for changing positions of the putative QTLs for GY and ASI. The results are presented in Figure 4 where the pleiotropic model (indicated in the figure by a dotted diagonal line) can be compared with alternative linkage models. Our result indicates that the area where the maximum of the test statistics was found (white area) included the pleiotropic model, though a close-linkage model cannot be excluded either (Figure 4). Whichever of the two models is the real underlying genetic model, the region would be considered as 'functionally pleiotropic' as breaking this association will always be difficult in practice.

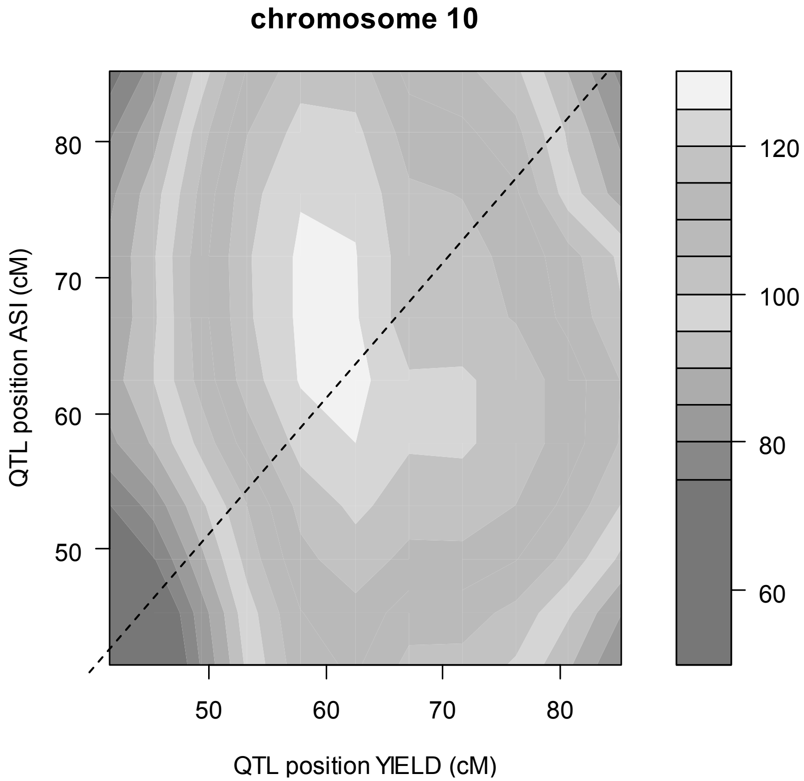

__Figure 4.__ Contour plot of the Wald statistic for QTL effects on chromosome 10 with varying positions for GY (horizontal axis) and for ASI (vertical axis). The results of pleiotropic models are represented on the diagonal (dotted line)

CONCLUSIONS

With an example in maize we showed how information stemming from managed-stress trials can be exploited to investigate the genetic causes of drought-stress adaptation. Mixed models are particularly suitable to model complex phenotypic responses across environments (stressed and not stressed), including the commonly observed GEI. In addition, multivariate mixed model approaches allow to model the association between traits in their dependence on the environmental conditions. One step further, molecular marker information can be incorporated to identify the genome regions underlying variation and co-variation between traits, thereby providing relevant information for practical plant breeding. Questions on the relevant regions to select for and on pleiotropy versus genetic linkage determining correlations between traits can be addressed. This information can be advantageously integrated in breeding procedures for direct and indirect selection of better adapted genotypes.

ACKNOWLEDGEMENTS

This chapter was written as part of the project 'An eco-physiological–statistical framework for the analysis of G×E and QTL×E as occurring in abiotic-stress trials, with applications to the CIMMYT drought-stress programs in tropical maize and bread wheat'. Generation Challenge Program, Competitive Grant 4.

REFERENCES

Bolaños, J. and Edmeades, G.O., 1993. Eight cycles of selection for drought tolerance in lowland tropical maize. II. Responses in reproductive behavior. *Field Crops Research,* 31 (3/4), 253-268.

Bolaños, J. and Edmeades, G.O., 1996. The importance of the anthesis-silking interval in breeding for drought tolerance in tropical maize. *Field Crops Research,* 48 (1), 65-80.

Chapman, S.C. and Edmeades, G.O., 1999. Selection improves drought tolerance in tropical maize populations: II. Direct and correlated responses among secondary traits. *Crop Science,* 39 (5), 1315-1324.

Edmeades, G.O., Bolaños, J., Chapman, S.C., et al., 1999. Selection improves drought tolerance in tropical maize populations. I. Gains in biomass, grain yield, and harvest index. *Crop Science,* 39 (5), 1306-1315.

Edmeades, G.O., Bänziger, M. and Ribaut, J.M., 2000. Maize improvement for drought-limited environments. *In:* Otegui, M.E. and Slafer, G.A. eds. *Physiological bases for maize improvement.* Food Products Press, New York, 75-111.

Jiang, C.J. and Zeng, Z.B., 1997. Mapping quantitative trait loci with dominant and missing markers in various crosses from two inbred lines. *Genetica,* 101 (1), 47-58.

Malosetti, M., Voltas, J., Romagosa, I., et al., 2004. Mixed models including environmental covariables for studying QTL by environment interaction. *Euphytica,* 137 (1), 139-145.

Ribaut, J.M., Hoisington, D.A., Deutsch, J.A., et al., 1996. Identification of quantitative trait loci under drought conditions in tropical maize. 1. Flowering parameters and the anthesis-silking interval. *Theoretical and Applied Genetics,* 92 (7), 905-914.

Smith, A., Cullis, B. and Thompson, R., 2001. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics,* 57 (4), 1138-1147.

Vargas, M., Van Eeuwijk, F.A., Crossa, J., et al., 2006. Mapping QTLs and QTL × environment interaction for CIMMYT maize drought stress program using factorial regression and partial least squares methods. *Theoretical and Applied Genetics,* 112 (6), 1009-1023.

Verbeke, G. and Molenberghs, G., 2000. *Linear mixed models for longitudinal data*. Springer Verlag, New York.

Westgate, M.E., 2000. Strategies to maintain ovary and kernel growth during drought. *In:* Otegui, M.E. and Slafer, G.A. eds. *Physiological bases for maize improvement*. Food Products Press, New York, 113-137.