Aus dem Institut
für Pflanzenzüchtung, Saatgutforschung und Populationsgenetik
der Universität Hohenheim
Fachgebiet: Angewandte Genetik und Pflanzenzüchtung
Prof. Dr. A. E. Melchinger

Biometrical Analyses of Epistasis and the Relationship between Line *per se* and Testcross Performance of Agronomic Traits in Elite Populations of European Maize (*Zea mays* L.)

Dissertation
zur Erlangung des Grades eines Doktors
der Agrarwissenschaften
der Fakultät Agrarwissenschaften
der Universität Hohenheim

von

Dipl.-Ing. sc. agr.
Renata Mihaljević
aus Zagreb

2005

Die vorliegende Arbeit wurde am 02. Mai 2005 von der Fakultät Agrarwissenschaften der Universität Hohenheim als „Dissertation zur Erlangung des Grades eines Doktors der Agrarwissenschaften (Dr.sc.agr.)“ angenommen.

Tag der mündlichen Prüfung:	28. Juli 2005

1. Prodekan:	Prof. Dr. K. Stahr
Berichterstatter, 1. Prüfer:	Prof. Dr. A. E. Melchinger
Mitberichterstatter, 2. Prüfer:	Prof. Dr. R. Blaich
3. Prüfer:	Prof. Dr. N. von Wirén

Contents

[1] Mihaljevic R., H.F. Utz, and A.E. Melchinger. 2004. Crop Sci. 44:114-124.

[2] Mihaljevic R., C.C. Schön, H.F. Utz, and A.E. Melchinger. 2005. Crop Sci. 45:114-122.

[3] Mihaljevic R., H.F. Utz, and A.E. Melchinger. 2005. Crop Sci. 45:2605-2613.

Abbreviations

BC1	backcross of generation F_1 to parent one
BC2	backcross of generation F_1 to parent two
BIC	Bayesian information criterion
CIM	composite interval mapping
cM	centiMorgan
CV	cross validation
DS	data set
ES	estimation set
IV	independent validation
LOD	log odds ratio
LP	line *per se* performance
LR	likelihood ratio
MAS	marker-assisted selection
p	proportion of the genetic variance explained by QTL
P1	parent one
P2	parent two
QTL	quantitative trait locus/loci
RFLP	restriction fragment length polymorphism
TC	testcross
TP	testcross performance
TS	test set

1 General Introduction

Relationship between Line *per se* and Testcross Performance

In hybrid breeding of maize, inbred lines are developed and selected according to both their *per se* performance and their hybrid performance. The latter is evaluated in testcrosses to a tester which is mostly an inbred line unrelated to the germplasm from which lines were developed. Because crossing lines to a tester and conducting yield trials for testcross progenies are expensive and time-consuming, any information on inbred lines that is indicative of their testcross performance is desirable. Relations of yield and other important agronomic traits of inbred lines to the same traits in hybrids have been studied from the time of initiation of hybrid breeding to the present (Hallauer and Miranda, 1981). It has been of great importance to determine whether expression of traits in inbred lines is transmissible to their hybrids.

Experimental estimates of the genotypic correlation between line *per se* (LP) and testcross performance (TP), $\hat{r}_g$ (LP, TP), vary considerably for different crops, traits, and selfing generations. In maize, for traits showing small heterotic effects and high heritability, e.g., grain moisture, ear length or days to flower, estimates of $\hat{r}_g$ (LP, TP) were medium to high. However, they were generally low for the highly heterotic and complex trait grain yield (for review see Hallauer and Miranda, 1981; Seitz, 1989). It was concluded that effective selection based on LP can be made for highly heritable traits, but not for yield and thus the ultimate use of inbred lines in hybrid combinations must be determined from yield evaluations of their testcrosses. Therefore, selection of lines is most commonly based on their general and specific combining ability assessed in topcross tests.

Reasons for the low genotypic correlations between LP and TP reported for grain yield may be that: (i) in advanced selfing generations of unselected materials, recessive genes with detrimental effect occur in homozygous state, (ii) in early selfing generations, LP for heterotic traits like grain yield is affected by different levels of heterozygosity which is not the case for TP, and (iii) overdominance, and/or epistasis are at work.

Smith (1986) demonstrated in theory that low correlations between LP and TP can be fully explained by a simple model with only additive and dominance genetic effects. Accordingly, $\hat{r}_g$ (LP, TP) is a linear function of the proportion of loci at which the inbred tester is homozygous for the favorable allele. As this proportion increases, $\hat{r}_g$ (LP, TP) decreases because the genotypic variance for TP is decreased due to the masking effect of dominant tester alleles over the unfavorable alleles of the lines tested. Thus, the ratio of genotypic variances for LP and TP should be an estimate of the genetic constitution of the tester and indicative of the prevalent type of gene action. While estimates of $\hat{r}_g$ (LP, TP) rely on the summed effects of all genes influencing LP and TP for a given trait, analyses of QTL (quantitative trait locus or loci depending on the context) provide a tool to clarify the basis of this correlation at the molecular level, i.e., for individual genetic factors.

QTL Analyses for Line *per se* and Testcross Performance

Most agronomically important traits such as grain yield, kernel weight, or protein concentration display a continuous distribution of phenotypic values. This is because variation for such traits is influenced by simultaneous segregation of numerous genes and is also affected by a number of environmental effects. Molecular markers have been employed in many species to dissect quantitative traits by estimating the map position and effects of the underlying QTL. Identification of individual genetic factors could lead to several useful applications. First, it could improve the efficacy of breeding in so-called marker-assisted selection (MAS), especially for traits with low heritability or those that can only be measured in one sex (see Soller and Beckmann, 1988; Lande and Thompson, 1990). Second, transgenic technology might be applied to quantitative traits. Third, quantitative genetic theory will be made more realistic when the numbers and properties of the QTL are known (Falconer and Mackay, 1996). A better understanding of the inheritance of quantitative traits may, therefore, lead to the development of improved breeding strategies.

Most QTL studies in maize were conducted with materials obtained by selfing or backcrossing progenies from a cross between two inbred lines. In hybrid breeding of maize, however, performance of inbred lines *per se* does not necessarily provide an

appropriate measure of their yield performance in hybrid combinations as is obvious from the estimates of the genotypic correlation $\hat{r}_g$(LP, TP). Accordingly, it is questionable whether QTL mapped for LP have the same position and/or effect with respect to TP in view of possible dominant or epistatic line × tester interactions. Hence, it may be questioned if MAS for TP based on information from markers flanking the QTL for LP will be efficient. QTL detected for both LP and TP simultaneously represent potential QTL for general combining ability of the lines in the population under study. In the literature, the proportion of common QTL detected for LP and TP was largest for plant and ear height with an unrelated tester, and smallest for grain yield with a related tester (Austin et al., 2000). This was in accordance with the magnitude of genotypic correlations between LP and TP estimated for these traits. For grain yield, therefore, it should be important to map QTL for TP directly using an unrelated tester inbred, which corresponds to the testing situation in a hybrid breeding program.

QTL Congruency across Experimental Populations

The trustworthiness of QTL experiments and the usefulness of their results for MAS depend primarily on the congruency of positions and effects of QTL across different samples of the same cross and among different crosses. Previous studies with populations derived from biparental crosses of elite lines showed only poor to moderate QTL congruency for agronomically important traits in maize and other species. These studies included different samples (Beavis, 1994; Melchinger et al., 1998; Igartua et al., 2000) or different generations of the same cross (Stromberg et al., 1994; Austin and Lee, 1996; Groh et al., 1998) as well as different crosses between related and unrelated parent lines (Abler et al., 1991; Beavis et al., 1991; Bubeck et al., 1993; Stuber, 1995; Thomas et al., 1995; Lübberstedt et al., 1998a,b; Pilet et al., 2001).

In contrast, congruency of QTL between different populations seems to be rather common for crosses of highly divergent parent lines and morphological traits. There are a relatively small number of QTL responsible for morphological divergence and most of the phenotypic variability can be accounted for by one or two QTL with large estimated effects that map to similar regions across comparable studies (Beavis, 1998). Variability exhibited

for quantitative traits of interest to plant breeders is assumed to be either oligogenic or polygenic and due to many more QTL with minor effects than is the case for the morphological traits.

Estimated QTL congruency depends on the sample size employed in QTL mapping as well as the approach used for comparing the QTL detected. With mostly limited sample sizes of mapping populations, the error in estimates of the QTL number, positions, and effects is generally high, especially for polygenic traits (Otto and Jones, 2000; Beavis, 1998; Broman, 2001; Utz and Melchinger, 1994). Therefore, criteria for assessing QTL congruency should allow discrimination between incongruency caused by biological or biometrical reasons.

Three criteria have been proposed in the literature for investigating the congruency of QTL: (i) counting of QTL at congruent genomic sites across the genome as used in numerous studies, (ii) permutation test of correspondence between genome-wide generated log odds ratio (LOD) score profiles described by Keightley and Knott (1999), and (iii) genetic correlation between predicted and observed phenotypic values in an independent sample having a special appeal with regard to MAS (Lande and Thompson, 1990; Melchinger et al., 1998; Utz et al., 2000). Applying so-called independent validation or cross validation (Utz et al., 2000) determines the magnitude of bias influenced by environmental and genotypic sampling, which leads to incongruency of QTL results. Statistical limitations causing incongruency of QTL across samples and populations will be even more manifest in the estimation of the underlying gene action, in particular of epistasis, which is discussed next.

Epistasis

Epistasis is the interaction of alleles at different loci and, thus, a form of non-additive gene action. It may cause a failure of crosses to show expected heterosis, a phenomenon which is the basis of hybrid performance attributed to dominance interactions among alleles at the same locus. Although epistasis may explain for deviations from theoretical expectations of heterotic performance and increasing evidence for its existence has been provided at the molecular level (Cheverud and Routman, 1995), the importance of epistasis

in the performance and heterosis of elite maize hybrids has received surprisingly little attention in research.

Traditional estimation approaches have relied on the analysis of first- and second-degree statistics by using either generation means analysis (Mather and Jinks, 1982) or estimation of epistatic variance components from covariances of relatives generated via special mating designs (Hallauer and Miranda, 1981). Hallauer and Miranda (1981, Chap. 5) reviewed studies that estimated epistatic variance components in maize. They summarized as follows: "It seems that epistasis for a complex trait, such as yield, must exist... but realistic estimates of additive by additive epistasis have not been obtainable. Hence, either the genetic models used are inadequate or epistatic variance is small relative to total genetic variance of maize populations".

Biometrical methods that use mean comparisons (generation means analyses) rather than variance component estimation have regularly indicated that epistatic effects are important for yield in maize. Hence, significant epistatic effects for grain yield in maize are detectable, but not so a significant epistatic variance. A major reason for this is that effects (first-order statistics) are easier to estimate precisely than variances (second-order statistics).

The traditional generation means analysis proposed by Hayman (1958) estimated the *per se* performance of the generations derived from a cross of two pure lines. Herewith, all types of digenic epistatic effects can be estimated. Melchinger (1987) proposed testcrossing the generations from Hayman's analysis to an inbred tester, which removes dominance effects from the model and diminishes competition effects in the experimental design that tended to overwhelm the epistatic effects. With Melchinger's model only the additive × additive type of epistasis can be estimated. Detection of significant epistatic effects, however, is no guarantee for epistasis to be important enough for the breeder. Stuber et al. (1973) and Crow (1999) stated that although epistatic effects are evident, their magnitude would not substantially hinder testcross prediction based on models ignoring epistasis.

Generation means analysis detects only epistatic effects summed over loci, so that positive and negative effects among individual QTL can cancel. QTL analyses do not share this problem, however, in most instances have revealed little or no evidence for epistasis (Stuber et al., 1992; Xiao et al., 1995; Liu et al., 1996). Nevertheless, when individual QTL

were isolated in isogenic backgrounds, epistasis was commonly observed (Doebley et al., 1995; Long et al., 1995; Eshed and Zamir, 1996; Laurie et al., 1997). Also, when genome-wide tests for epistasis were performed, epistatic interactions were detected among marker loci that did not show significant main effects (Damerval et al., 1994; Holland et al., 1997; Li et al., 1997).

Recently, epistatic QTL for yield and its component traits in the autogamous species rice have often been detected (Yu et al., 1997, Li et al., 2001; Luo et al., 2001; Hua et al., 2002; Hua et al., 2003; Mei et al, 2003; etc.). In another autogamous species, *Arabidopsis thaliana*, Kearsey et al. (2003) reported that epistasis of duplicate type which opposes dominance was a common feature of 22 quantitative traits as detected by generation means analysis. For maize, Stuber et al. (1992) reported that dominance was the prevalent gene action underlying hybrid performance.

Objectives

In this study, four and five populations of F_3 to F_6 lines derived from three crosses of elite inbred lines of European flint maize were evaluated for LP and TP, respectively, of five agronomically important quantitative traits: grain yield, grain moisture, kernel weight, protein concentration, and plant height. The population size ranged from 71 to 344. The objectives were to:

(i) estimate phenotypic and genotypic correlations between LP and TP within four populations for all five traits and discuss possible causes for their magnitude,

(ii) determine the positions and gene effects of QTL detected for LP and TP in four and five populations, respectively, for all five traits,

(iii) investigate the influence of the sample and genetic background on QTL congruency among testcross populations,

(iv) determine the gene action of QTL identified for LP and their value for the prediction of TP,

(v) estimate the magnitude of aggregate epistatic effects by generation means analyses of LP and TP in four crosses of European flint lines for grain yield and grain moisture and detect marker pairs with significant genome-wide epistatic effects for

LP and TP of these traits in the four populations previously employed for QTL mapping of LP and TP, and last

(vi) draw conclusions regarding the prospects of MAS.

References

Abler, B.S.B., M.D. Edwards, and C.W. Stuber. 1991. Isoenzymatic identification of quantitative trait loci in crosses of elite maize hybrids. Crop Sci. 31:267-274.

Austin, D.F., and M. Lee. 1996. Comparative mapping in $F_{2:3}$ and $F_{6:7}$ generations of quantitative trait loci for grain yield and yield components in maize. Theor. Appl. Genet. 92:817-826.

Austin, D.F., M. Lee, L.R. Veldboom, and A.R. Hallauer. 2000. Genetic mapping in maize with hybrid progeny across testers and generations: grain yield and grain moisture. Crop. Sci. 40:30-39.

Beavis, W.D., D. Grant, M. Albertsen, and R. Fincher. 1991. Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. Theor. Appl. Genet. 83:141-145.

Beavis, W.D. 1994. The power and deceit of QTL experiments: Lessons from comparative QTL studies. p. 250-266. *In* Proc. of the 49th Annual Corn and Sorghum Industry Research Conf., Chicago, IL. 7-8 Dec. 1994. American Seed Trade Assoc., Washington, DC.

Beavis, W.D. 1998. QTL analyses: power, precision, and accuracy. p.145-162. *In* A.H. Paterson (ed.) Molecular dissection of complex traits. CRC Press, Boca Raton, USA.

Broman, K.W. 2001. Review of statistical methods for QTL mapping in experimental crosses. Lab Animal 30(7):44-52.

Bubeck, D., M. Goodman, W. Beavis, and D. Grant. 1993. Quantitative trait loci controlling resistance to gray leaf spot in maize. Crop Sci. 33:838-847.

Cheverud, J.M., and E.J. Routman. 1995. Epistasis and its contribution to genetic variance components. Genetics 139:1455-1461.

Crow, J.F. 1999. Epistasis. p. 523-524. *In* J.G. Coors and S. Pandey (ed.) The genetics and exploitation of heterosis in crops. ASA, CSSA, and SSSA, Madison WI.

Damerval, C., A. Maurice, J.M. Josse, and D. de Vienne. 1994. Quantitative trait loci underlying gene product variation: a novel perspective for analyzing regulation of gene expression. Genetics 137: 289-301.

Doebley, J., A. Stec, and C. Gustus. 1995. *Teosinte branched 1* and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics 141: 333-346.

Eshed, Y., and D. Zamir. 1996. Less-than-additive epistatic interactions of quantitative trait loci in tomato. Genetics 143: 1807-1817.

Falconer, D.S., and T.F.C. Mackay. 1996. Introduction to quantitative genetics. 4th ed., Longman Ltd. Group, London, UK.

Groh, S., D. Gonzalez-de-Leon, M.M. Khairallah, C. Jiang, D. Bergvinson, M. Bohn, D.A. Hoisington, and A.E. Melchinger. 1998. QTL mapping in tropical maize: III. Genomic regions for resistance to *Diatraea* spp. and associated traits in two RIL populations. Crop Sci. 38:1062-1072.

Hallauer, A.R., and J.B. Miranda. 1981. Quantitative Genetics in Maize Breeding. Iowa State University Press, Ames, IA.

Hayman, B.I. 1958. The separation of epistatic from additive and dominance variation in generation means. Heredity 12:371-390.

Holland, J.B., H.S. Moser, L.S. O'Donoughue, and M. Lee. 1997. QTLs and epistasis associated with vernalization responses in oat. Crop Sci. 37:1306-1316.

Hua, J.P., Y.Z. Xing, C.G. Xu, X.L. Sun, S.B. Yu, and Q. Zhang. 2002. Genetic dissection of an elite rice hybrid revealed that heterozygotes are not always advantageous for performance. Genetics 162:1885-1895.

Hua, J.P., Y.Z. Xing, W. Wu, C. Xu, X.L. Sun, S.B. Yu, and Q. Zhang. 2003. Single-locus heterotic effects and dominance by dominance interactions can adequately explain the genetic basis of heterosis in an elite rice hybrid. Proc. Natl. Acad. Sci. 100:2574-2579.

Igartua, E., M. Edney, B.G. Rossnagel, D. Spaner, W.G. Legge, G.J. Scoles, P.E. Eckstein, G.A. Penner, N.A. Tinker, K.G. Briggs, D.E. Falk, and D.E. Mather. 2000. Marker-based selection of QTL affecting grain and malt quality in two-row barley. Crop Sci. 40:1426-1433.

Kearsey, M.J., H.S. Pooni, and N.H. Syed. 2003. Genetics of quantitative traits in *Arabidopsis thaliana*. Heredity 91:456-464.

Keightley, P.D., and S.A. Knott. 1999. Testing the correspondence between map positions of quantitative trait loci. Genet. Res. Camb. 74:323-328.

Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743-756.

Laurie, C.C., J.R. True, J. Liu, and J.M. Mercer. 1997. An introgression analysis of quantitative trait loci that contribute to the morphological difference between *Drosophila simulans* and *D. mauritiana.* Genetics 145:339-348.

Li, Z., S.R.M. Pinson, W.D. Park, A.H. Paterson, J.W. Stansel. 1997. Epistasis for three grain yield components in rice (*Oryza sativa* L.). Genetics 145:453-465.

Li, Z.K., L.J. Luo, H.W. Mei, D.L. Wang, Q.Y. Shu, R. Tabien, D.B. Zhong, C.S. Ying, J.W. Stansel., G.S. Khush, and A.H. Paterson. 2001. Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield. Genetics 158:1737-1753.

Liu, J., J.M. Mercer, L.F. Stam, G.C. Gibson, Z. Zeng, and C.C. Laurie. 1996. Genetic analysis of a morphological shape difference in the male genitalia of *Drosophila simulans* and *D. mauritiana*. Genetics 142:1129-1145.

Long, A.D., S.L. Mullaney, L. A. Reid, J.D. Fry, C.H. Langley, and T.F.C. Mackay. 1995. High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. Genetics 139:1273-1291.

Lübberstedt, T., D. Klein, and A. Melchinger. 1998a. Comparative QTL mapping of resistance to *Ustilago maydis* across four populations of European flint-maize. Theor. Appl. Genet. 97:1321-1330.

Lübberstedt, T., A. Melchinger, S. Fähr, D. Klein, A. Dally, and P. Westhoff. 1998b. QTL mapping in testcrosses of flint lines of maize: III. Comparison across populations for forage traits. Crop Sci. 38:1278-1289.

Luo, L.J., Z.K. Li, H.W. Mei, Q.Y. Shu, R. Tabien, D.B. Zhong, C.S. Ying, J.W. Stansel, G.S. Khush, and A.H. Paterson. 2001. Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield components. Genetics 158:1755-1771.

Mather, K., and J.L. Jinks. 1982. Biometrical genetics. 3rd ed. Chapman and Hall, London.

Mei, H.W., L.J. Luo, C.S. Ying, Y.P. Wang, X.Q. Yu, L.B. Guo, A.H. Paterson, and Z.K. Li. 2003. Gene actions of QTLs affecting several agronomic traits resolved in a recombinant inbred rice population and two testcross populations. Theor. Appl. Genet. 107:89-101.

Melchinger, A.E. 1987. Expectation of means and variances of testcrosses produced from F_2 and backcross individuals and their selfed progenies. Heredity 59:105-115.

Melchinger, A.E., H.F. Utz, and C.C. Schön. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383-403.

Otto, S. P., and C.D. Jones. 2000. Detecting the undetected: Estimating the total number of loci underlying a quantitative trait. Genetics 156:2093-2107.

Pilet, M.L., G. Duplan, M. Archipiano, P. Barret, C. Baron, R. Horvais, X. Tanguy, M.O. Lucas, M. Renard, and R. Delourme. 2001. Stability of QTL for field resistance to blackleg across two genetic backgrounds in oilseed rape. Crop Sci. 41:197-205.

Seitz, G. 1989. Experimentelle und theoretische Untersuchungen zur Beziehung zwischen Linieneigenleistung und allgemeiner Kombinationsfähigkeit bei Silomais. Ph.D. diss. Univ. of Hohenheim, Stuttgart, Germany.

Smith, O.S. 1986. Covariance between line *per se* and testcross performance. Crop Sci. 26:540-543.

Soller, M., and J.S. Beckmann. 1988. Genomic genetics and the utilization for breeding purposes of genetic variation between populations. p. 161-188. *In* Weir, B.S., E.J.

Eisen, M.M. Goodman, and G. Namkoong (ed.), Proc. Int. Conf. on Quantitative Genetics, 2nd. Sinauer, Sunderland, Mass., USA.

Stromberg, L.D., J.W. Dudley, and G.K. Rufener. 1994. Comparing conventional early generation selection with molecular marker-assisted selection in maize. Crop Sci. 34:1221-1225.

Stuber, C.W., W.P. Williams, and R.H. Moll. 1973. Epistasis in maize (*Zea mays* L.): III. Significance in predictions of hybrid performances. Crop Sci. 13:195-200.

Stuber, C.W., S.E. Lincoln, D.W. Wolff, T. Helentjaris, and E.S. Lander. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:823-839.

Stuber, C.W. 1995. Mapping and manipulating quantitative traits in maize. Trends in Genetics 11:477-481.

Thomas, W.T.B., W. Powell, R. Waugh, K.J. Chalmers, U.M. Barua, P. Jack, V. Lea, B.P. Forster, J.S. Swanston, R.P. Ellis, P.R. Hanson, and R.C.M. Lance. 1995. Detection of quantitative trait loci for agronomic, yield, grain, and disease characters in spring barley (*Hordeum vulgare* L.). Theor. Appl. Genet. 91:1037-1047.

Utz, H.F., and A.E. Melchinger. 1994. Comparison of different approaches to interval mapping of quantitative trait loci. p. 195-204. *In* J. W. van Ooijen, and J. Jansen (ed.) Proc. Meeting of the EUCARPIA Section Biometrics in Plant Breeding, 9th, Wageningen, The Netherlands.

Utz, H.F., A.E. Melchinger, C.C. Schön. 2000. Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. Genetics 154:1839-1849.

Xiao, J., J. Li, L. Yuan, and S.D. Tanksley. 1995. Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. Genetics 140:745-754.

Yu, S.B., J.X. Li, C.G. Xu, Y.F. Tan, Y.J. Gao, X.H. Li, Q. Zhang, and M.A. Saghai Maroof. 1997. Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. Proc. Natl. Acad. Sci. 94:9226-9231.

Congruency of Quantitative Trait Loci Detected for Agronomic Traits in Testcrosses of Five Populations of European Maize

Renata Mihaljevic, H. Friedrich Utz, and Albrecht E. Melchinger*

ABSTRACT

Congruency of estimated positions and effects of QTL in different samples of the same cross or different crosses is an indicator of the reliability of these estimates and their usefulness in marker-assisted selection (MAS). We investigated the influence of the sample and genetic background on QTL congruency among five populations of European maize (*Zea mays* L.). Three samples derived from the same cross comprised 344 ($A \times B^{I}$) and 109 ($A \times B^{II}$) $F_{2:3}$ as well as 71 $F_{4:5}$ ($A \times B^{III}$) lines. Two other crosses comprised 109 ($A \times C$) and 84 ($C \times D$) $F_{3:4}$ lines. All lines were topcrossed to the same inbred tester and evaluated in four or five environments. A combined linkage map of RFLP marker data from all five populations was used in composite interval mapping (CIM). The total number of QTL identified for five agronomically important traits was 42 in $A \times B^{I}$, 18 in $A \times B^{II}$, 20 in $A \times B^{III}$, 28 in $A \times C$, and 23 in $C \times D$. Averaged across traits, the proportion *p* of the genetic variance explained by these QTL varied between 50.4% in the largest population $A \times B^{I}$ and 30.7% in a population of considerably smaller size ($A \times B^{II}$). Cross validation (CV) yielded substantially lower estimates of *p*. Between 10 and 24% of the 42 QTL from $A \times B^{I}$ were also detected within a 20-cM interval in the other four populations. Incongruent QTL among $A \times B$ samples were due to the low power of QTL detection and the large bias in QTL estimates. The genetic correlations between predicted (based on QTL positions from one population) and observed phenotypic values in another population were highest among $A \times B$ samples with a maximum of 0.68 for plant height. Congruency of QTL was found for kernel weight, protein concentration, and plant height and was mainly attributable to one or few QTL of moderate to large size. If more cost-effective than phenotypic selection, MAS will be promising for these traits.

Molecular markers have been employed in numerous species to dissect quantitative traits by estimating the map positions and effects of the underlying quantitative trait loci (QTL). One important aspect concerning efficient use of QTL in MAS is congruency of positions and effects of QTL across different samples of the same cross or different crosses. Previous studies with populations derived from biparental crosses of elite lines showed only poor to moderate QTL congruency for agronomically important traits in maize and other species. These studies included different samples (Beavis, 1994; Melchinger et al., 1998; Igartua et al., 2000) or different generations of the same cross (Stromberg et al., 1994; Austin and Lee, 1996; Groh et al., 1998) as well as different crosses between related and unrelated parent lines (Abler et al., 1991; Beavis et al., 1991; Bubeck et al., 1993; Stuber, 1995; Thomas et al., 1995; Lübberstedt et al., 1998a,b; Pilet et al., 2001).

In contrast, congruency of QTL between different populations seems to be rather common for crosses of highly divergent parent lines and complex but easily classified morphological traits. In interspecific crosses, QTL with mostly drastic effects mapped to the same genomic sites or even syntenic regions (for review see Beavis, 1998). Likewise, Mackay (1995, 1996) and Long et al. (1995) reported for the highly heritable trait bristle number in *Drosophila* a clustering of QTL from different populations in the vicinity of candidate loci.

Important factors influencing QTL congruency are the sample size employed in QTL mapping as well as the approach used for comparing the QTL detected. With mostly limited sample sizes of mapping populations, the error in estimates of QTL number, positions, and effects is generally high, especially for polygenic traits (Otto and Jones, 2000; Beavis, 1998; Broman, 2001; Utz and Melchinger, 1994). Therefore, criteria for assessing QTL congruency should allow discrimination between incongruency caused by biological or biometrical reasons.

Three criteria have been proposed in the literature for investigating the congruency of QTL: (i) counting of QTL at congruent genomic sites across the genome as used in numerous studies; (ii) permutation test of correspondence between genome-wide generated log odds ratio (LOD) score profiles described by Keightley and Knott (1999); (iii) genetic correlation between predicted and observed phenotypic values in an independent sample having special appeal with regard to MAS (Lande and Thompson, 1990; Melchinger et al., 1998; Utz et al., 2000). Determining congruency implies comparisons of at least two samples by use of either an additional independent validation (IV) sample or CV. We applied all three criteria and both validation methods to compare QTL results for traits of presumably different complexity from five populations with both, one, or none of the three elite parents in common.

Our objectives were to (i) determine the positions and gene effects of QTL detected in each of the five populations, (ii) compare QTL congruency across populations by all three criteria, (iii) discuss the influence of the sample and genetic background on QTL congruency for different traits, and (iv) draw conclusions regarding the prospects of MAS in plant breeding.

Inst. of Plant Breeding, Seed Sci., and Population Genetics, Univ. of Hohenheim, 70593 Stuttgart, Germany. Received 24 Mar. 2003. *Corresponding author (melchinger@uni-hohenheim.de).

Published in Crop Sci. 44:114–124 (2004).

677 S. Segoe Rd., Madison, WI 53711 USA

Abbreviations: CIM, composite interval mapping; cM, centiMorgan; CV, cross validation; DS, data set; ES, estimation set; IV, independent validation; LOD, log odds ratio; LR, likelihood ratio; MAS, marker-assisted selection; *p*, proportion of the genetic variance; P1, parent one; P2, parent two; QTL, quantitative trait locus/loci; RFLP, restriction fragment length polymorphism; TC, testcross; TS, test set.

MATERIALS AND METHODS

Plant Materials

Some of the plant materials used in this study were identical to those employed and described in previous studies on grain traits (Schön et al., 1994; Melchinger et al., 1998; Utz et al., 2000) and forage traits in maize (Lübberstedt et al., 1997). Briefly, four early maturing homozygous European flint lines KW1265, D146, D145, and KW1292, subsequently referred to as A, B, C, and D, respectively, were used as parents. From cross A × B, randomly chosen F_2 plants were selfed to produce 507 F_3 ($F_{2:3}$) lines. These were randomly divided into two samples of 380 and 127 F_3 ($F_{2:3}$) lines designated as A × B^I and A × B^{II}, respectively. The 127 F_3 lines of A × B^{II} were selfed by single-seed descent until generation F_4 to produce 113 F_5 ($F_{4:5}$) lines, designated as A × B^{III}. Furthermore, 131 F_4 ($F_{3:4}$) lines of cross A × C and 140 F_4 ($F_{3:4}$) lines of cross C × D were generated by using bulked seeds of the selfed F_3 plants of each F_3 line. Testcross (TC) seed was produced in isolation plots by mating the unrelated inbred tester (KW5361, [Iodent], referred to as T2 in the notation of Schön et al., 1994), as pollinator to a random sample of 40 plants from each of the F_n lines (F_3 lines in A × B^I and A × B^{II}, F_5 lines in A × B^{III}, F_4 lines in A × C and C × D) as well as to the parent lines A, B, C, and D.

Field Experiments

The TC progenies were evaluated in five experiments. Experiment 1 (A × B^I) was conducted in 1990 and 1991 at two locations in Germany (Gondelsheim and Grucking) as described by Melchinger et al. (1998). The 400 entries consisted of 380 TCs of F_3 lines, TCs of parents A and B included as quintuple entries, and 10 common check hybrids. In addition, data on plant height were taken from forage trials conducted at five environments in Germany as described by Lübberstedt et al. (1997). Experiment 2 (A × B^{II}) was conducted in 1992 and 1993 at two locations in Germany (Eckartsweier and Bad Krozingen). The 150 entries consisted of TCs of the 127 F_3 lines, TCs of the parents A and B included as six and seven entries, respectively, and the same set of 10 check hybrids as in Exp. 1. Because of insufficient quantities of seeds, TC progenies of only 71 F_5 lines of cross A × B^{III} were evaluated in Exp. 3, 109 F_4 lines (A × C) in Exp. 4, and 84 F_4 lines (C × D) in Exp. 5, conducted in 1992 in adjacent trials at five locations with rather diverse agroecological conditions (Chartres in France; Eckartsweier, Grucking, Bad Krozingen, and Gondelsheim in Germany). Experiments 3 to 5 each included 150 entries. Testcrosses of each parent line were included as quintuple entries in each experiment as well as common check hybrids and other lines for completion. The experimental design employed was a 40 × 10 α-design (Patterson and Williams, 1976) for Exp. 1 and a 15 × 10 α-design for the remaining experiments, with two replications each. Two-row plots were overplanted and later thinned to reach a final stand of 80 000 to 110 000 plants ha^{-1} depending on the location. All experiments were machine planted and harvested as grain trials with a combine.

Data were analyzed for the following traits: grain yield (Mg ha^{-1}) adjusted to 155 g kg^{-1} grain moisture, grain moisture (g kg^{-1}) at harvest, kernel weight in mg per kernel determined from four samples of 50 kernels from each plot, protein concentration in grain (g kg^{-1}) estimated by near-infrared reflectance spectroscopy as described by Melchinger et al. (1986), and plant height (cm) on a plot basis as the distance from the soil level to the lowest tassel branch.

RFLP Marker Genotyping and Linkage Map Construction

The procedures for RFLP assays were described by Schön et al. (1994). A subsample of 344 parental F_2 plants of the 380 F_3 lines of A × B^I, and a subsample of 109 parental F_2 plants of the 127 F_3 lines of A × B^{II} were genotyped for a total of 89 RFLP marker loci distributed across the maize genome. A total of 151, 104, and 122 RFLP marker loci were employed to map 113 F_5 lines of A × B^{III}, as well as 131 and 140 F_4 lines of crosses A × C and C × D, respectively. Observed genotype frequencies at each marker locus were tested against expected Mendelian segregation ratios and allele frequency 0.5 by χ^2 tests. Appropriate type I error rates were determined by the sequentially rejective Bonferroni procedure (Holm, 1979). Linkage maps of the individual populations, as well as a joint map combining the molecular data of all populations, were constructed with software JOINMAP Version 3.0 (Van Ooijen and Voorrips, 2001). A LOD threshold of 3.0 was used for declaring linkage in two-point analyses and Haldane's mapping function (Haldane, 1919) was employed for calculating map distances. For the joint map, each linkage group was truncated at both ends. The points of truncation were the most distal markers common to all individual maps.

Agronomic Data Analyses

Analyses of variance were performed for each experiment and environment. Adjusted entry means and effective error mean squares were then used to compute the combined analyses of variance and covariance across environments for each experiment. The sums of squares for entries were subdivided into the variation among TCs of the F_n lines and orthogonal contrasts among the TC means of parent lines P1 and P2 and F_n lines. A corresponding subdivision was conducted on the entry × environment interaction sums of squares. Estimates of variance components $\hat{\sigma}^2_e$ (effective error variance), $\hat{\sigma}^2_{ge}$ (genotype × environment interaction variance) and $\hat{\sigma}^2_g$ (genotypic variance) of F_n TC progenies and their standard errors were calculated as described by Searle (1971, p. 475). Heritabilities (h^2) on a TC progeny mean basis were estimated as described by Hallauer and Miranda (1981, p. 90) and their 95% confidence intervals according to Knapp et al. (1985). Phenotypic ($\hat{r}_p$) and genotypic ($\hat{r}_g$) correlations between the TC performance of F_5 lines of A × B^{III} and F_3 lines of A × B^{II} were calculated for all traits by standard procedures (Mode and Robinson, 1959).

Quantitative Trait Loci Analyses

Quantitative trait loci mapping and estimation of their effects were performed with PLABQTL (Utz and Melchinger, 1996) employing CIM by the regression approach (Haley and Knott, 1992). All QTL analyses were performed with the joint map. An additive genetic model was assumed for the analysis of TC progenies as described in detail by Utz et al. (2000). Cofactors were selected by stepwise regression according to Miller (1990, p. 49) with an "F-to-enter" and "F-to-delete" value of 3.5. Testing for presence of a putative QTL in an interval by a likelihood ratio (LR) test was performed with a 2.5 (= 0.217 LR) LOD threshold in conformity with the foregoing publications on these materials. We also set higher LOD thresholds of 3.5 in A × B^{II} and A × B^{III} as well as 5.0 in A × B^I for certain comparisons across samples. Estimates of QTL positions were obtained at the point where the LOD score assumed its maximum in the region under consideration. For each population, the proportion of the phenotypic variance ($\hat{\sigma}^2_p$) explained by a single QTL was determined as the

square of the partial correlation coefficient (R^2). Estimates of the allele substitution effect (α) of each putative QTL and their partial R^2 were obtained by fitting a model including all significant QTL for the respective trait simultaneously. This model was also used to estimate p_{DS}, the proportion of the genotypic variance ($\hat{\sigma}^2_g$) explained by all QTL detected with the whole data set (DS) for a given trait, by dividing the adjusted total R^2 (R^2_{adj}) by the heritability (h^2) as described by Utz et al. (2000).

Fivefold CV implemented in PLABQTL was used to obtain asymptotically unbiased estimates of p_{DS} (Shao, 1997). For each population, a DS comprising the entry means across environments was divided into five genotypic subsamples. Four of these were combined in an estimation set (ES) for QTL detection and estimation of genetic effects, whereas the remaining subsample was used as a test set (TS) to validate the predictions gained from ES. We call this analysis standard CV. This analysis deviates from CV/G described by Utz et al. (2000), where the ES and TS were defined by omitting one environment of a DS. Here, data from all environments was averaged to obtain phenotypic values, and therefore only five different CV runs are possible by permuting the respective subsamples. A total of 1000 replicated CV runs was performed with 200 randomizations for assigning genotypes to the respective subsamples. Estimates of the proportion of the genotypic variance ($\hat{\sigma}^2_g$) explained by all QTL detected for a given trait were calculated as medians $\tilde{p}_{ES}$ from the 1000 estimates in ES. The validated median $\tilde{p}_{TS.ES}$ was obtained by correlating the observed data in TS with those predicted on the basis of QTL positions and effects estimated in ES. An ad hoc estimate of the bias of p_{DS} was calculated by the difference of medians $\tilde{p}_{ES} - \tilde{p}_{TS.ES}$. The bias of an individual QTL effect in a DS was estimated as the difference of means $\overline{\alpha}_{ES} - \overline{\alpha}_{TS.ES}$ by averaging across all CV runs which contained the individual QTL of a DS within a $\pm$10-cM interval of the QTL position estimated by CIM in a DS. Hereby, $\overline{\alpha}_{ES}$ is the mean estimate in ES, and $\overline{\alpha}_{TS.ES}$ the result of its validation in TS at the QTL position of ES. Within the same interval, the QTL frequency (i.e., the frequency of occurrence of a putative QTL) was determined across the 1000 CV runs.

Three procedures were employed for quantifying the congruency of QTL across populations: (i) number of congruent QTL, whereby individual QTL were considered congruent across two populations if their estimated map position was within a 20-cM distance, irrespective of the sign of estimated α-effects in the two populations; (ii) correlation of LOD score values r (LOD_i, LOD_j) (i, j = A $\times$ B^I, A $\times$ B^{II}, A $\times$ B^{III}, A $\times$ C, and C $\times$ D; $i \neq j$) from populations i and j across the genome (Keightley and Knott, 1999), with significance thresholds for r at the 5% level determined as the 2.5 and 97.5 percentiles of 2000 permutations; (iii) the genetic correlation between predicted and observed TC performance, r_g (M_i, Y_j) (i, j = A $\times$ B^I, A $\times$ B^{II}, A $\times$ B^{III}, A $\times$ C, and C $\times$ D; $i \neq j$). For brevity, a particular r_g (M_i, Y_j) will be denoted as r_g (A $\times$ B^I, A $\times$ B^{II}), for example. Here, M_i is the predicted value based on the QTL positions and effects estimated in the population i (estimation population) and Y_j is the observed value in the population j (validation population). For details, see Utz et al. (2000). The parameter r_g (M_i, Y_j) was estimated for all pairs of populations, except those having no parent in common. The assumption was that in crosses with one parent in common the other parent contributes same allelic effects at the QTL in both crosses. If i and j represent populations of the same cross, r_g (M_i, Y_j) will be comparable with $\sqrt{\tilde{p}_{TS.ES}}$ derived from CV within the population i.

RESULTS

Segregation and Linkage of RFLP Markers

The individual RFLP linkage maps of the five populations generated by JOINMAP corresponded to a large extent with the linear order and marker distances previously determined with mapping software MAPMAKER/EXP (Lander et al., 1987) and GMendel (Holloway and Knapp, 1993), as described by Schön et al. (1994) and Lübberstedt et al. (1997), respectively. A group of four loci (UMC94, BNL8.05a, UMC76, and UMC137) which had previously been mapped on chromosome 1, were not significantly linked to any other markers employed in this analysis. The same was the case with the loci UMC32a and UMC121, as well as UMC109, which had previously been mapped to chromosomes 3 and 9, respectively. We assigned UMC 109 to the linkage group of chromosome 9 in accordance with a widely used reference UMC map (Davis et al., 1999) because it was the only marker common to all populations at the distal portion of the short arm of chromosome 9.

The joint map spanned a total of 1138 cM with an average interval length of 14.4 cM in A $\times$ B^I and A $\times$ B^{II}, 15.0 cM in A $\times$ B^{III}, 12.1 cM in A $\times$ C, and 10.2 cM in C $\times$ D. This map covered approximately 70% of the genome covered by the reference map (Schön et al., 1994) and 84% of the QTL regions detected by Melchinger et al. (1998) in A $\times$ B^I across traits.

In total, six marker loci in populations A $\times$ B^I and A $\times$ B^{II}, and three in A $\times$ C were scored as dominant markers. For markers of the joint map, the observed genotype frequencies generally coincided with the expected Mendelian segregation ratios in A $\times$ B^{II}. Significant deviations were observed once in A $\times$ B^I and A $\times$ B^{III}, twice in A $\times$ C, and in five cases in C $\times$ D. Significant ($P < 0.01$) deviations from 0.5 allele frequency were not found. The joint map is available at http://www.agron.missouri.edu (verified 20 Aug. 2003).

Agronomic Trait Analysis

Herein, only the results for populations A $\times$ B^{III}, A $\times$ C, and C $\times$ D will be presented because agronomic data of populations A $\times$ B^I and A $\times$ B^{II} was reported previously (Schön et al., 1994; Melchinger et al., 1998). Weather conditions were mostly favorable for grain maize production in all five environments, except for noticeable drought stress at Chartres reflected in reduced plant height and kernel weight estimates. The TC progeny means of population A $\times$ B^{III} ($\overline{F}_5$) exceeded TC progeny means of A $\times$ C and C $\times$ D ($\overline{F}_4$) for kernel weight and protein concentration (Table 1). For grain yield and plant height, the highest TC progeny means were obtained in C $\times$ D, whereas for grain moisture, TC mean of A $\times$ C was highest (Table 1). The TC means of P1 and P2 differed significantly ($P < 0.01$) for all traits except grain yield in C $\times$ D and grain moisture in A $\times$ B^{III}. The orthogonal contrast between average TC performance of the parent lines ($\overline{P}$) and the TC mean of the F_n lines ($\overline{F}_n$) was significant ($P < 0.01$) only

Table 1. Estimates of means, variance components, and heritabilities of maize testcross (TC) progenies from parent lines (P1 and P2) and F_5 or F_4 lines from crosses A × B^{III}, A × C, and C × D with inbred tester T2 for five agronomic traits, measured in four (A × B^{III}) and five (A × C and C × D) environments, respectively. For cross A × B, phenotypic and genotypic correlation coefficients between different generations (A × B^{II} and A × B^{III}) are given.

Parameter	Entries	Grain yield	Grain moisture	Kernel weight	Protein concentration	Plant height
	no.	Mg ha^{-1}	g kg^{-1}	g	g kg^{-1}	cm
			Cross A × B^{III}			
TC means†						
P1	5	7.65 ± 0.21‡	283.6 ± 1.4	328.4 ± 1.7	117.4 ± 0.53	233.1 ± 1.0
P2	5	10.31 ± 0.21	285.5 ± 1.4	294.1 ± 1.7	112.2 ± 0.53	224.6 ± 1.0
$\bar{P}$	10	8.98 ± 0.15	284.6 ± 1.0	311.2 ± 1.2	114.8 ± 0.37	228.9 ± 0.7
$\bar{F}_5$	71	8.89 ± 0.10	281.5 ± 1.1	310.4 ± 1.7	116.1 ± 0.49	229.2 ± 0.8
Range of F_5 lines		6.08–10.62	260.3–302.7	281.9–342.0	103.1–123.6	214.2–247.0
Variance components						
$\hat{\sigma}^2_L$		0.492 ± 0.119**	76.8 ± 14.5**	180.7 ± 32.6**	15.39 ± 2.80**	40.54 ± 7.57**
$\hat{\sigma}^2_{ge}$		0.825 ± 0.091**	23.9 ± 4.5**	32.31 ± 6.73**	3.35 ± 0.63**	8.18 ± 2.22**
$\hat{\sigma}^2_\varepsilon$		0.494 ± 0.029	51.8 ± 3.0	84.42 ± 4.89	7.28 ± 0.42	31.63 ± 1.83
Heritability						
$\hat{h}^2$		0.70	0.88	0.92	0.92	0.89
95% C.I. on $\hat{h}^2$§		0.55–0.79	0.83–0.92	0.89–0.95	0.88–0.94	0.84–0.93
Correlation coefficients¶						
$\hat{r}_p$		0.43**	0.39**	0.35**	0.38**	0.33**
$\hat{r}_g$		0.62	0.44	0.40	0.48	0.32
			Cross A × C			
TC means†						
P1	5	7.95 ± 0.37‡	282.6 ± 2.3	333.3 ± 2.5	118.4 ± 0.88	234.4 ± 1.6
P2	5	10.54 ± 0.37	291.6 ± 2.3	257.9 ± 2.5	105.2 ± 0.88	241.4 ± 1.6
$\bar{P}$	10	9.25 ± 0.26	287.1 ± 1.6	295.6 ± 1.8	111.8 ± 0.62	237.9 ± 1.1
$\bar{F}_4$	109	9.65 ± 0.06	287.0 ± 0.8	292.4 ± 1.2	109.3 ± 0.33	236.2 ± 0.7
Range of F_4 lines		8.08–11.11	270.5–307.9	259.3–319.2	101.1–117.5	216.1–255.1
Variance components						
$\hat{\sigma}^2_L$		0.271 ± 0.061**	53.31 ± 8.49**	135.3 ± 20.14**	10.43 ± 1.61**	52.90 ± 7.88**
$\hat{\sigma}^2_{ge}$		0.619 ± 0.061**	20.32 ± 3.59**	22.69 ± 5.42**	2.96 ± 0.56**	8.58 ± 2.15**
$\hat{\sigma}^2_\varepsilon$		0.505 ± 0.029	54.42 ± 3.14	93.88 ± 5.30	8.80 ± 0.50	37.57 ± 2.16
Heritability						
$\hat{h}^2$		0.61	0.85	0.91	0.88	0.91
95% C.I. on $\hat{h}^2$§		0.46–0.71	0.79–0.89	0.87–0.93	0.83–0.91	0.87–0.93
			Cross C × D			
TC means†						
P1	5	10.47 ± 0.21‡	285.0 ± 2.2	259.4 ± 2.1	104.6 ± 0.80	244.6 ± 1.3
P2	5	10.55 ± 0.21	265.6 ± 2.2	285.2 ± 2.1	113.7 ± 0.80	229.3 ± 1.3
$\bar{P}$	10	10.51 ± 0.15	275.3 ± 1.5	272.3 ± 1.5	109.1 ± 0.57	236.9 ± 0.9
$\bar{F}_4$	84	10.45 ± 0.06	276.7 ± 0.8	274.4 ± 1.4	108.9 ± 0.43	238.5 ± 0.7
Range of F_4 lines		9.46–11.85	259.3–294.0	238.8–300.7	97.3–119.6	219.1–257.1
Variance components						
$\hat{\sigma}^2_L$		0.201 ± 0.045**	46.4 ± 8.6**	157.3 ± 25.9**	13.85 ± 2.3**	40.59 ± 7.02**
$\hat{\sigma}^2_{ge}$		0.175 ± 0.038**	18.1 ± 3.9**	12.45 ± 5.3**	2.39 ± 0.56**	4.78 ± 2.30*
$\hat{\sigma}^2_\varepsilon$		0.548 ± 0.032	55.3 ± 3.2	94.21 ± 5.4	8.40 ± 0.48	41.73 ± 1.19
Heritability						
$\hat{h}^2$		0.69	0.84	0.93	0.91	0.89
95% C.I. on $\hat{h}^2$§		0.56–0.78	0.76–0.88	0.90–0.95	0.88–0.94	0.84–0.92

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
† $\bar{P}$ = TC mean of P1 and P2; $\bar{F}_4$, $\bar{F}_5$ = TC means of F_5 and F_4 lines, respectively.
‡ Standard errors are attached.
§ Confidence intervals on $\hat{h}^2$ were calculated according to Knapp et al. (1985).
¶ Phenotypic ($\hat{r}^2_p$) and genotypic ($\hat{r}^2_g$) correlation coefficients among TC progenies of related F_3 (A × B^{II}) and F_5 (A × B^{III}) lines.

for protein concentration in population A × C. For all traits and populations, the range in TC performance of F_n lines considerably exceeded the TC means of the parents.

Genotypic variances ($\hat{\sigma}^2_g$) among TCs of F_n lines were highly significant ($P < 0.01$) for all traits in all populations (Table 1). Genotypic variances among F_5 lines (A × B^{III}) were significantly higher ($P < 0.01$) than those among F_3 lines in A × B^{I} and A × B^{II}. Estimates of genotype × environment interaction variance ($\hat{\sigma}^2_{ge}$) were significantly greater than zero ($P < 0.05$) for all traits in all populations. Except for grain yield, $\hat{\sigma}^2_{ge}$ was consistently smaller than $\hat{\sigma}^2_g$. Heritability was medium for grain yield ($0.61 < \hat{h}^2 < 0.70$), but relatively high for the other traits ($0.84 < \hat{h}^2 < 0.93$) in all three populations. Phenotypic correlations ($\hat{r}_p$) between related TC progenies from F_3 lines (A × B^{II}) and F_5 lines (A × B^{III}) were highly significant ($P < 0.01$) for all traits. Corresponding genotypic correlations ($\hat{r}_g$) ranged from 0.32 to 0.62.

Quantitative Trait Loci Analyses

The QTL results for A × B^{I} and A × B^{II} were reported previously (Schön et al., 1994; Melchinger et al., 1998). Results from QTL analyses of all five populations based on the joint map are presented here for means across environments: the proportion of the genotypic variance explained in Table 2 and the number of QTL detected in Table 3. Detailed information on positions and effects of individual QTL detected can be obtained at http://www.agron.missouri.edu.

Table 2. Proportion of genotypic variance (p) explained by putative QTL determined by three estimation procedures for five agronomic traits; QTL detected in TC progenies of F_3 lines of maize populations A × B^{I} and A × B^{II}, F_5 lines of population A × B^{III}, and F_4 lines of populations A × C and C × D with the inbred tester T2.

Trait	Parameter	Population				
		A × B^{I}	A × B^{II}	A × B^{III}	A × C	C × D
		%				
Grain yield	$\hat{p}_{DS}$†	32.3	25.7	70.1	83.2	78.8
	$\tilde{p}_{ES}$‡	32.1	28.9	55.6	81.6	72.5
	$\tilde{p}_{TS.ES}$§	18.7	6.0	8.2	51.8	35.9
Grain moisture	$\hat{p}_{DS}$	46.0	33.9	26.2	32.4	36.4
	$\tilde{p}_{ES}$	43.6	29.1	22.8	26.3	38.6
	$\tilde{p}_{TS.ES}$	33.0	10.5	3.1	5.2	2.5
Kernel weight	$\hat{p}_{DS}$	51.9	10.5	42.2	34.8	44.2
	$\tilde{p}_{ES}$	53.8	42.2	52.3	41.0	43.8
	$\tilde{p}_{TS.ES}$	42.3	16.4	26.6	13.5	13.5
Protein concentration	$\hat{p}_{DS}$	55.3	53.2	56.0	39.6	50.8
	$\tilde{p}_{ES}$	50.7	45.6	47.0	47.2	52.8
	$\tilde{p}_{TS.ES}$	38.9	17.0	9.8	16.6	19.5
Plant height	$\hat{p}_{DS}$	66.5	30.0	11.2	44.5	35.4
	$\tilde{p}_{ES}$	62.6	27.3	30.2	46.9	36.4
	$\tilde{p}_{TS.ES}$	49.3	16.2	−0.3	22.4	12.8

† Explained by putative QTL detected with composite interval mapping in the entire data set (DS).
‡ Explained by putative QTL detected with standard fivefold cross validation (CV) in the estimation set (ES) given as median ($\tilde{p}_{ES}$) across 1000 replicated CV runs.
§ Explained by putative QTL detected with standard fivefold cross validation (CV) in the test set (TS) given as median ($\tilde{p}_{TS.ES}$) across 1000 replicated CV runs.

Grain Yield

We detected a total of two, three, seven, six, and six putative QTL for grain yield in A × B^{I}, A × B^{II}, A × B^{III}, A × C, and C × D, respectively (Table 3). A simultaneous fit of all putative QTL explained between R^2_{adj} = 15.5 (A × B^{I}) and 54.4% (C × D) of $\hat{\sigma}^2_p$, and between $\hat{p}_{DS}$ = 25.7 (A × B^{II}) and 83.2% (A × C) of $\hat{\sigma}^2_g$ (Table 2). Across populations, the sum of absolute α-effects ranged from 0.92 (A × B^{I}) to 4.07 Mg ha^{-1} (A × B^{III}), corresponding to 8.9 and 45.6% of the TC means of F_3 and F_5 lines, respectively. Cross validation resulted in $\tilde{p}_{TS.ES}$ values ranging from 6.0 (A × B^{II}) to 51.8% (A × C), which were substantially smaller than $\tilde{p}_{ES}$ values (Table 2).

Grain Moisture

We detected nine, four, three, seven, and six QTL for grain moisture in A × B^{I}, A × B^{II}, A × B^{III}, A × C, and C × D, respectively, distributed across the genome (Table 3). Collectively, they accounted for R^2_{adj} = 23.2% of $\hat{\sigma}^2_p$ in A × B^{III} and 37.6% in A × B^{I}, the minimum and maximum obtained for the five populations. The

Table 3. Number of common† QTL for five agronomic traits in populations (A × B^{I}, A × B^{II}, A × B^{III}, A × C, and C × D) (above diagonal), and genetic correlation of predicted and observed testcross performance r_g (M_i, Y_j)‡ (below diagonal). The total number of QTL found in each population is given along the diagonal in italics.

Trait	Validation population	Estimation population				
		A × B^{I}	A × B^{II}	A × B^{III}	A × C	C × D
Grain yield	A × B^{I}	*2*	0	2	1	0
	A × B^{II}	0.26	*3*	1	0	1
	A × B^{III}	0.39	0.21	*7*	2	3
	A × C	0.26	0.15	0.38	*6*	1
	C × D	–	–	–	0.66	*6*
Grain moisture	A × B^{I}	*9*	2	2	3	0
	A × B^{II}	0.46	*4*	0	3	0
	A × B^{III}	0.27	0.20	*3*	1	0
	A × C	0.13	0.05	0.21	*7*	0
	C × D	–	–	–	0.25	*6*
Kernel weight	A × B^{I}	*10*	1	3	1	2
	A × B^{II}	0.63	*2*	1	1	0
	A × B^{III}	0.65	0.44	*3*	1	0
	A × C	0.46	0.45	0.47	*4*	2
	C × D	–	–	–	0.28	*4*
Protein concentration	A × B^{I}	*9*	2	2	3	1
	A × B^{II}	0.47	*6*	3	3	1
	A × B^{III}	0.60	0.43	*6*	3	0
	A × C	0.26	0.37	0.47	*6*	1
	C × D	–	–	–	0.34	*4*
Plant height	A × B^{I}	*12*	2	1	2	1
	A × B^{II}	0.60	*3*	0	1	0
	A × B^{III}	0.68	0.34	*1*	0	0
	A × C	0.20	0.19	0.01	*5*	0
	C × D	–	–	–	0.09	*3*

† QTL with estimated position within a 20-cM distance, irrespective of the sign of the α-effect.
‡ Correlation between the phenotypic means observed in the validation population and predicted genotypic values on the basis of QTL positions and effects derived from the estimation population.

proportion of $\hat{\sigma}^2_g$ explained by all putative QTL ranged from $\hat{p}_{DS}$ = 26.2 (A × B^{III}) to 46.0% (A × B^{I}) (Table 2). The sum of absolute α-effects was between 22.3 g kg^{-1} in A × B^{III} (7.9% of $\overline{F}_5$) and 51.6 g kg^{-1} in C × D (18.6% of $\overline{F}_4$). With CV, $\tilde{p}_{TS,ES}$ values ranged from 2.5 (C × D) to 33.0% (A × B^{I}), which were considerably lower than the corresponding $\tilde{p}_{ES}$ values (Table 2).

Kernel Weight

Ten QTL regions across the genome were significantly associated with kernel weight in population A × B^{I}, two in A × B^{II}, three in A × B^{III}, and four in A × C and C × D (Table 3). A simultaneous fit yielded a minimum R^2_{adj} = 8.3% in A × B^{II} and a maximum R^2_{adj} = 43.9% in A × B^{I}. Simultaneously, all putative QTL explained between 10.5 (A × B^{II}) and 51.9% (A × B^{I}) of $\hat{\sigma}^2_g$ (Table 2). The sum of absolute α-effects varied between 15.3 g in A × B^{II} and 63.8 g in A × B^{I} (4.7 and 20.5% of the TC mean of F_3 lines, respectively). Estimates of $\tilde{p}_{TS,ES}$ ranged from 13.5 (A × C and C × D) to 42.3% (A × B^{I}), and were substantially lower than corresponding estimates of $\tilde{p}_{ES}$ (Table 2).

Protein Concentration

Nine QTL were identified for protein concentration in A × B^{I}, four in C × D, and six QTL in each of the populations A × B^{II}, A × B^{III}, and A × C distributed across the genome (Table 3). Collectively, they explained between R^2_{adj} = 34.7% in A × C and 51.4% in A × B^{III}. Estimates of $\hat{p}_{DS}$ ranged from 39.6 (A × C) to 56.0% (C × D) (Table 2). The sum of absolute α-effects varied from 11.3 g kg^{-1} in C × D (10.3% of $\overline{F}_4$ lines) to 18.0 g kg^{-1} in A × B^{III} (15.5% of $\overline{F}_5$ lines). Cross validation yielded estimates of $\tilde{p}_{TS,ES}$ between 9.8% in A × B^{III} and 38.9% in A × B^{I}, being substantially reduced as compared with corresponding $\tilde{p}_{ES}$ values (Table 2).

Plant Height

A total of 12, 3, 1, 5, and 3 QTL affecting plant height was detected in A × B^{I}, A × B^{II}, A × B^{III}, A × C, and C × D, respectively (Table 3). A simultaneous fit explained between R^2_{adj} = 10.0 (A × B^{III}) and 52.6% (A × B^{I}) of $\hat{\sigma}^2_p$, and between 11.2 (A × B^{III}) and 66.5% (A × B^{I}) of $\hat{\sigma}^2_g$ (Table 2). The largest sum of absolute α-effects was 48.4 cm in A × B^{I} (19.3% of $\overline{F}_3$ lines), the smallest amounted to 6.8 cm in A × B^{III} (2.96% of $\overline{F}_5$ lines). Cross validation yielded estimates of $\tilde{p}_{TS,ES}$ ranging from −0.3 (A × B^{III}) to 49.3% of $\hat{\sigma}^2_g$ (A × B^{I}), which were considerably smaller than their corresponding $\tilde{p}_{ES}$ estimates (Table 2).

Comparison of QTL across Populations

Comparing different samples of the same generation in the same cross, seven out of 18 QTL detected in the smaller population (A × B^{II}) were found to be within a 20-cM distance from the 42 QTL detected in the larger population (A × B^{I}) across all five traits (Table 3). For grain yield, no common QTL was detected. The genome-wide correlation of LOD-score values for A × B^{I} and A × B^{II} was significant ($P < 0.05$) only for kernel weight and plant height (Table 4). The genetic correlation r_g (A × B^{I}, A × B^{II}) ranged from 0.26 for grain yield to 0.63 for kernel weight (Table 3).

Comparing different generations of the same cross originating from the same (A × B^{II} vs. A × B^{III}) or different (A × B^{I} vs. A × B^{III}) samples, out of the 20 QTL detected across all five traits in A × B^{III}, 10 and 5 were in common to A × B^{I} and A × B^{II}, respectively (Table 3). The genome-wide correlation between LOD scores was significant ($P < 0.05$) for kernel weight (above 0.39) in both comparisons and for plant height only when comparing A × B^{I} vs. A × B^{III} (Table 4). The genetic correlation r_g (A × B^{II}, A × B^{III}) reached a maximum of 0.44 for kernel weight and a minimum of 0.20 for grain moisture, whereas the extremes for r_g (A × B^{I}, A × B^{III}) were 0.68 for plant height and 0.27 for grain moisture (Table 3).

In the comparison of populations having one parent in common, out of the 28 QTL detected in A × C across all five traits, only 10, 8, and 7 were common to the QTL detected in A × B^{I}, A × B^{II}, and A × B^{III}, respectively (Table 3). The genome-wide correlation of LOD scores between A × C and A × B^{I} was significant ($P < 0.05$) only for kernel weight (Table 4). This was also the case when A × B^{III} was compared with A × C; however, when comparing A × B^{II} vs. A × C, no significant correlations were obtained (data not shown). For most traits, r_g (A × B^{III}, A × C) was mostly higher than r_g (A × B^{I}, A × C) or r_g (A × B^{II}, A × C). The first correlation refers to populations evaluated in the same environments, which is not the case for the other two correlations. Estimates of r_g (A × B^{I}, A × C) were of medium size (0.46) for kernel weight but considerably lower for other traits. Only four out of 28 QTL identified

Table 4. Genome-wide correlation (r) between log odds ratio (LOD) scores of two populations. The LOD scores are determined by composite interval mapping of putative QTL affecting five agronomic traits in A × B^{I}, A × B^{II}, A × B^{III}, A × C, and C × D.

	Population pair					
	A × B^{I}				A × B^{II}	A × C
Trait	A × B^{II}	A × B^{III}	A × C	C × D	A × B^{III}	C × D
Grain yield	−0.05	0.25	0.07	−0.08	0.16	0.13
Grain moisture	0.27	0.17	0.17	−0.04	−0.16	0.00
Kernel weight	0.39*	0.45*	0.40*	0.03	0.61**	0.16
Protein concentration	−0.00	0.05	0.04	0.02	0.25	−0.00
Plant height	0.63**	0.43*	−0.13	0.04	0.23	0.02

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.

in A × C were in common to the 23 QTL detected in C × D across traits (Table 3). The genome-wide correlations of LOD scores between A × C and C × D were close to zero for all traits (Table 4). The correlations r_g (A × C, C × D) ranged from 0.09 (plant height) to 0.66 (grain yield) despite the fact that for grain yield only one QTL was in common to both populations.

In the comparison of populations having no parent in common, out of the 23 QTL detected across all five traits in C × D, only two to four were in common with A × B^I, A × B^{II}, and A × B^{III} (Table 3). The genome-wide correlation of LOD scores was practically zero for all traits when comparing A × B^I vs. C × D (Table 4). This was also the case when comparing A × B^{II} or A × B^{III} vs. C × D (data not shown).

DISCUSSION

Comparison of Criteria for Assessing QTL Congruency

Assessing the congruency of QTL among populations requires, above all, appropriate criteria and statistical tests. Three criteria were employed in this study. Our first criterion, counting the QTL with congruent positions, has so far predominantly been used in comparisons of QTL from different populations (e.g., Lübberstedt et al., 1998a,b; Pilet et al., 2001; He et al., 2001). Following Melchinger et al. (1998) and Groh et al. (1998), we declared a pair of QTL from two populations as congruent if they were located within a 20-cM distance. This corresponds to the criterion of overlapping bin regions used by Tuberosa et al. (2002) and seems more appropriate than overlapping confidence intervals because CIM does not provide their straightforward calculation (Visscher et al., 1996; Bennewitz et al., 2002). The procedure is useful for determining the number of common QTL in two mapping experiments, but yields no information about the conformity of QTL effects or LOD score profiles.

The second criterion, the correlation coefficient between LOD score profiles overcomes this deficiency. As Keightley and Knott (1999) concluded from simulations and experimental results, however, the correlation coefficients were low and the power to detect congruency decreased already with several QTL underlying the trait. This was corroborated in our study because significant associations were obtained only if one or few large QTL were congruent. Small differences in QTL positions often reduced the correlation substantially. Therefore, we agree with Keightley and Knott on not using this criterion for complex polygenic traits.

Our third criterion, the genetic correlation between predicted and observed phenotypic values, r_g (M_i, Y_j), estimates the QTL congruency quantitatively by taking into account both positions and effects of QTL. It deals adequately with cases of linked QTL (e.g., two linked QTL in a large sample or a *ghost* QTL in a smaller sample) and is best suited for assessing the prospects of MAS because it corresponds to the square root of the proportion of genetic variance explained by QTL. A shortcoming is the large estimation error associated with r_g (M_i, Y_j) if the heritability is low, because the latter occurs in the denominator of the formula. Furthermore, same allelic effects at the QTL must be assumed if populations share one or no parent.

Impact of Shortcomings in QTL Analyses on QTL Congruency across Samples

Lack of QTL congruency across different samples of the same cross reflects the limitations and shortcomings of QTL analyses. They depend on (i) random errors associated with phenotypic and marker data, (ii) sampling of genotypes and environments, and (iii) bias caused by model selection in QTL analyses.

The first factor was presumably of minor importance for explaining the poor QTL congruency between the three populations of A × B, because our phenotypic values referred to means across four or five environments and heritabilities were fairly high for all traits except grain yield (Table 1).

Genotypic sampling influences QTL detection and estimation of their positions and effects to a much higher extent than environmental sampling with more than three environments (Utz et al., 2000). This was corroborated herein also for grain yield, the trait with the highest expected G × E interaction variance. Estimated QTL × E interaction variance components in the PLABQTL analysis were mostly small compared with the QTL variance components across populations, except for A × C, where the two variance components were of similar size. The genetic variance explained by all putative QTL detected in A × C remained high with $\tilde{p}_{TS.ES}$ = 51.8% after standard CV (Table 2). With CV on independent environmental and genotypic samples (i.e., CV/GE in Utz et al. [2000]), however, the above estimate was reduced to $\tilde{p}_{TS.ES}$ = 22.1%. The reason may be the fact that two QTL detected in A × C showed different signs across the five test environments. In such a case, the environmental sample may influence the size of the QTL effect in the mapping population and consequently reduce the QTL congruency with the other populations.

Model selection in QTL mapping can introduce a bias and cause a substantial inflation in QTL estimates (Utz and Melchinger, 1994; Georges et al., 1995; Beavis, 1998; Broman, 2001; Göring et al., 2001). As demonstrated by simulations of these authors, the bias in estimates of individual QTL effects as well as p can be as high as the true parameters, with the bias and sampling error increasing for small sample sizes and small effects of the QTL.

By the same token, the power of QTL detection increases for larger sample sizes and effects of QTL. Assuming a QTL with an estimated $R^2 = 0.10$, which corresponds to the average value across all traits and QTL determined in our study, the power of detecting such a QTL is 0.98 for N = 500 but only 0.65 for N = 100 (Charcosset and Gallais, 1996). The probability of detecting such a QTL simultaneously in two independent samples is obtained by multiplication. Taking bias into account, the true QTL effect is only about half as large

as the estimated QTL effect, which reduces the probability of joint QTL detection in both samples to 0.30. This value is in close agreement with the proportions of congruent QTL detected in A × B^{I} vs. A × B^{II} or A × B^{III}. The QTL congruency is further reduced if a constant Type I error level is chosen because our 2.5 LOD threshold corresponds to a level of 0.14 in A × B^{I}, 0.23 in A × B^{II}, and 0.40 in A × B^{III} with use of the permutation test of Doerge and Churchill (1996).

In conclusion, genotypic sampling and estimation bias can largely explain the low rate of congruency between QTL detected in different samples of the same cross. Consequently, with a low power of QTL detection it remains an open question whether incongruency was due to sampling error or due to genetic causes, as there may be different QTL × environment interactions when populations are grown in different environments or different allelic effects at QTL in the case of different crosses.

Information Gain from Cross Validation

Resampling methods such as CV have been proposed to determine the sampling error and bias of QTL estimates (Utz et al., 2000). By a comparison of CV results from populations A × B^{I}, A × B^{II}, and A × B^{III}, we examined whether CV permits assessment of (i) the power of QTL detection by looking at QTL frequencies, (ii) the bias and standard error of individual QTL effects, and (iii) the bias in p calculated as the difference in corresponding estimates from ES and TS. For a summary across traits, QTL effects were standardized by dividing the estimated substitution effects by the phenotypic standard deviation of entry means.

The fidelity of QTL detection was assessed by QTL frequency, which corresponds to the percentage of the 1000 CV runs, in which the QTL was detected in the ±10-cM interval of the QTL position found by CIM in a DS. As expected, the QTL frequency decreased with decreasing sample size and averaged 0.74 in A × B^{I}, 0.54 in A × B^{II}, and 0.46 in A × B^{III}. Even with N = 344 in A × B^{I}, the QTL frequency exceeded 0.95 only for seven out of the 42 detected QTL. In the smaller samples, the maximum QTL frequency amounted to 0.88. In all three populations, the QTL frequency was significantly correlated with the LOD scores and the absolute standardized QTL effects, which corroborates that it is a good indicator of the power of QTL detection.

The average of the standardized QTL effects across all five traits amounted to 0.34 in A × B^{I}, 0.47 in A × B^{II}, and 0.38 in A × B^{III}. These differences are largely attributable to the increased bias of QTL effects estimated from smaller populations because the CV bias of standardized QTL effects averaged 0.06 in A × B^{I}, but 0.18 in A × B^{II} and A × B^{III}. Large estimated QTL effects generally displayed a smaller bias than the smaller ones. The CV also revealed a large variation in QTL effects estimated from TS in different runs. The variation of estimated bias was also smaller in the group of larger QTL than in the group of smaller QTL, especially in the large population A × B^{I}. Hence, for smaller populations our results corroborate the findings of Göring et al. (2001) that the estimated QTL effects may be virtually independent of the true size of the QTL. Moreover, IV corresponds essentially to a single CV run and shows high standard errors of QTL effects when using small sample sizes unless a QTL is very large.

While individual QTL effects often deviated considerably between CV and IV, estimates of p ($\bar{p}_{TS,ES}$) averaged across traits from CV and r_g^2 (M_i, Y_j) from IV showed good agreement if the large population A × B^{I} was used for QTL mapping (Table 5). This confirms that CV provides asymptotically unbiased estimates of p (Shao, 1997). The LOD thresholds for these comparisons were set higher than 2.5 as we found the congruency to be mostly due to largest QTL.

In conclusion, our findings clearly support the routine use of CV in QTL analyses. With CIM based on the regression approach, the increase in computation time is almost negligible. Moreover, even five to 10 CV runs already allow a fairly robust assessment of the estimation bias of p.

Trait-Specific QTL Congruency

Falconer and Mackay (1996, p. 357) designated QTL explaining >10% of the phenotypic variance or their standardized effects exceeding 0.5, respectively, as "large." The standardized effects averaged across the three populations of the cross A × B were <0.5 as already discussed. However, at least one large QTL was found in each population and for each trait. Although these large QTL were not necessarily detected at congruent positions across populations, for kernel weight, protein concentration, and plant height they could have been detected even with higher LOD thresholds (3.5

Table 5. Mean number of QTL detected with increased log odds ratio (LOD) thresholds in three estimation populations and mean coefficients of genetic correlation between predicted and observed testcross performance r_g (M_i, Y_j)† with M_i derived from estimation population (above and below diagonal) averaged across grain yield, grain moisture, kernel weight, protein concentration, and plant height. The comparable estimates of $\bar{p}_{TS,ES}^{0.5}$ averaged across all traits are given in italics on the diagonal.

Estimation population‡	Number of QTL	Validation population		
		A × B^{I}	A × B^{II}	A × B^{III}
		r_g (M_i, Y_j)		
A × B^{I}	4.2	***0.49***	0.44	0.49
A × B^{II}	2.2	0.34	***0.18***	0.36
A × B^{III}	1.2	0.30	0.39	***0.15***

† Correlation between the phenotypic means observed in validation population and predicted genotypic values on the basis of QTL positions and effects derived from estimation population.

‡ LOD threshold = 5.0 in A × B^{I}, LOD threshold = 3.5 in A × B^{II} and A × B^{III}.

for A × B^{II} and A × B^{III}, 5.0 in A × B^{I}) and contributed substantially to the high genome-wide congruency evidenced by genetic correlations r_g (M_i, Y_j) in Table 3. Large QTL did not act accordingly for grain yield and grain moisture, which may be due to high estimation error or a higher number of small QTL underlying these traits. Moreover, presence of highly integrated epistatic complexes (Stuber et al., 1999) or varied control of these traits via metabolic pathways (Bost et al., 1999) may be other causes for this result.

With sample sizes typically used in QTL mapping experiments, it seems unrealistic to unravel the genetic architecture of polygenic traits. Even with $N = 344$ in A × B^{I}, one can make only cautious inferences concerning the importance and width of a QTL region. Limitations are already manifest in detecting the true number of QTL (Otto and Jones, 2000) and furthermore in estimating the degree of dominance and epistasis of a given trait.

Congruency of QTL from Different Crosses

Owing to the high selection pressure exerted in maize breeding programs, it seems plausible that the same favorable alleles are fixed at a QTL in both parents of a cross within the same heterotic group. Thus, polymorphism at a QTL in one but its absence in the other cross could be a biological cause for incongruency. Furthermore, the divergence of the parental lines of two crosses will be reflected in magnitude and direction of effects found for QTL at congruent positions. Moreover, epistasis can modulate the effect of a QTL depending on the genetic background. Hence, it is not surprising that we found no QTL congruent among all crosses.

Congruency as evidenced by the genetic correlations r_g (M_i, Y_j) was generally diminished if one of the parents varied between crosses. A noticeable higher value of r_g (A × C, C × D) was found for grain yield due to a large congruent QTL on chromosome 1. The higher r_g values of A × B populations with A × C for kernel weight were also mostly attributable to a large congruent QTL on chromosome 8. It is striking that in other QTL studies in maize, QTL for grain yield and its components were reported on the same region of chromosome 1 and on chromosome 8 (Abler et al., 1991; Beavis et al., 1994; Austin and Lee, 1996; Veldboom and Lee, 1996). Each of these QTL may represent either a gene complex or individual genes controlling a specific metabolic pathway or gene network.

Alternative approaches to QTL mapping that do not rely on biparental crosses might provide new tools for investigating the congruency of QTL in different populations. Besides QTL mapping in multiple-line crosses (Rebai and Goffinet, 2000; Xie et al., 1998; Xu, 1998; Liu and Zeng, 2000), the haplotype-based QTL mapping approach recently devised by Jansen et al. (2003) promises progress in this direction, because it can be applied to progeny from multiple related crosses. Furthermore, congruent QTL across different genetic backgrounds can be confirmed by association mapping (Meuwissen and Goddard, 2000; Thornsberry et al., 2001), if candidate genes and/or high density maps are available.

Implications for Marker-Assisted Selection and QTL Mapping

The high estimation error and low power explain why in most published experiments on MAS, only about half of the QTL under selection actually contributed to the realized selection response (Eathington et al., 1997; Mather et al., 1997; Igartua et al., 2000; Bouchez et al., 2002). Obviously, the chances for MAS are substantial if at least a few large QTL are detected, even if some of them are false positives or overestimated.

Marker-assisted selection should be promising in our material for some traits such as kernel weight, protein concentration, and plant height because independent samples of the same cross yielded congruent QTL and explained up to 46% of the genetic variance. For these traits, genetic correlations between A × B^{II} and A × B^{III}, for example, based on the whole genotype (Table 1) corresponded well to the r_g (A × B^{II}, A × B^{III}) based on the QTL genotype (Table 3). Nevertheless, even for these traits we recommend the use of a large population for mapping at least of a size of 300 correspondingly to the one used in this study for A × B^{I} ($N = 380$). The p values estimated from validation were still below the corresponding h^2 estimates; consequently, MAS will be superior to phenotypic selection only if it is more cost-effective (Lande and Thompson, 1990; Knapp, 1998).

In view of the high costs of QTL mapping experiments, it would be advantageous if QTL regions were consistent among crosses and only the most suitable flanking marker and the sign of the QTL allele would have to be determined for each population. Remapping of QTL at regular intervals during MAS is necessary because QTL-marker associations change during several generations of selection (Gimmelfarb and Lande, 1995). A multistage approach with estimation of QTL in one generation and with validation and combined estimation in the next generation would allow for an efficient use of both phenotypic and marker data. An essential prerequisite for this approach is the integration of QTL mapping in ordinary breeding programs with elite germplasm, as suggested by Jannink et al. (2001).

ACKNOWLEDGMENTS

The present study was part of EUREKA project 290 supported by grants from the German Ministry of Research and Technology (BMBF) and KWS Kleinwanzlebener Saatzucht AG, grant 0319233A. The RFLP assays were conducted in the lab of Prof. Dr. R.G. Herrmann, Ludwig-Maximilians-Universität in Munich, by E. Brunklaus-Jung and J. Boppenmaier as well as A. Dally and P. Westhoff at the Heinrich-Heine-Universität in Düsseldorf. The skilled technical assistance of F. Mauch, D. Schilling-Groß, A. Vesting, and the staff at the Plant Breeding Research Station in Eckartsweier in conducting field trials is gratefully acknowledged. This article is dedicated to F.W. Schnell on the occasion of his 90th birthday.

REFERENCES

Abler, B.S.B., M.D. Edwards, and C.W. Stuber. 1991. Isoenzymatic identification of quantitative trait loci in crosses of elite maize hybrids. Crop Sci. 31:267–274.

Austin, D.F., and M. Lee. 1996. Comparative mapping in $F_{2:3}$ and $F_{6:7}$ generations of quantitative trait loci for grain yield and yield components in maize. Theor. Appl. Genet. 92:817–826.

Beavis, W.D. 1994. The power and deceit of QTL experiments: Lessons from comparative QTL studies. p. 250–266. *In* Proc. of the 49th Annual Corn and Sorghum Industry Res. Conf., Chicago, IL. 7–8 Dec. 1994. American Seed Trade Assoc., Washington, DC.

Beavis, W.D. 1998. QTL analyses: Power, precision, and accuracy. p. 145–162. *In* A.H. Paterson (ed.) Molecular dissection of complex traits. CRC Press, Boca Raton, FL.

Beavis, W.D., D. Grant, M. Albertsen, and R. Fincher. 1991. Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. Theor. Appl. Genet. 83: 141–145.

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

Bennewitz, J., N. Reinsch, and E. Kalm. 2002. Improved confidence intervals in quantitative trait loci mapping by permutation bootstrapping. Genetics 160:1673–1686.

Bost, B., C. Dillmann, and D. de Vienne. 1999. Fluxes and metabolic pools as model traits for quantitative genetics. I. The L-shaped distribution of gene effects. Genetics 153:2001–2012.

Bouchez, A., F. Hospital, M. Caussee, A. Gallais, and A. Charcosset. 2002. Marker-assisted introgression of favourable alleles at quantitative trait loci between maize elite lines. Genetics 162:1945–1959.

Broman, K.W. 2001. Review of statistical methods for QTL mapping in experimental crosses. Lab. Anim. 30(7):44–52.

Bubeck, D., M. Goodman, W. Beavis, and D. Grant. 1993. Quantitative trait loci controlling resistance to gray leaf spot in maize. Crop Sci. 33:838–847.

Charcosset, A., and A. Gallais. 1996. Estimation of the contribution of quantitative trait loci (QTL) to the variance of a quantitative trait by means of genetic markers. Theor. Appl. Genet. 93:1193–1201.

Davis, G.L., M.D. McMullen, C. Baysdorfer, T. Musket, D. Grant, M. Staebell, G. Xu, M. Polacco, L. Koster, S. Melia-Hancock, K. Houchins, S. Chao, and E.H. Coe, Jr. 1999. A maize map standard with sequenced core markers, grass genome reference points and 932 expressed sequence tagged sites (ESTs) in a 1736-locus map. Genetics 152:1137–1172.

Doerge, R.W., and G.A. Churchill. 1996. Permutation tests for multiple loci affecting a quantitative character. Genetics 142:285–294.

Eathington, S.R., J.W. Dudley, and G.K. Rufener, II. 1997. Usefulness of marker-QTL associations in early generation selection. Crop Sci. 37:1686–1693.

Falconer, D.S., and T.F.C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Longman Ltd. Group, London.

Georges, M., D. Nielsen, M. Mackinnon, A. Mishra, R. Okimoto, A.T. Pasquino, L.S. Sargeant, A. Sorensen, M.R. Steele, and X. Zhao. 1995. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics 139: 907–920.

Gimmelfarb, A., and R. Lande. 1995. Marker-assisted selection and marker-QTL associations in hybrid populations. Theor. Appl. Genet. 91:522–528.

Göring, H.H.H., J.D. Terwilliger, and J. Blangero. 2001. Large upward bias in estimation of locus-specific effects from genomewide scans. Am. J. Hum. Genet. 69:1357–1369.

Groh, S., D. Gonzalez-de-Leon, M.M. Khairallah, C. Jiang, D. Bergvinson, M. Bohn, D.A. Hoisington, and A.E. Melchinger. 1998. QTL mapping in tropical maize: III. Genomic regions for resistance to *Diatraea* spp. and associated traits in two RIL populations. Crop Sci. 38:1062–1072.

Haldane, J.B.S. 1919. The combination of linkage values, and the calculation of distance between the loci of linked factors. J. Genet. 8:299–309.

Haley, C.S., and S.A. Knott. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69:315–324.

Hallauer, A.R., and J.B. Miranda. 1981. Quantitative Genetics in Maize Breeding. Iowa State Univ. Press, Ames.

He, P., J.Z. Li, X.W. Zheng, L.S. Shen, C.F. Lu, Y. Chen, and L.H. Zhu. 2001. Comparison of molecular linkage maps and agronomic trait loci between DH and RIL populations derived from the same rice cross. Crop Sci. 41:1240–1246.

Holloway, J.L., and S.J. Knapp. 1993. GMendel 3.0: Software for the analysis of genetic markers and maps. p. 1–130. Oregon State Univ., Corvallis.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6:65–70.

Igartua, E., M. Edney, B.G. Rossnagel, D. Spaner, W.G. Legge, G.J. Scoles, P.E. Eckstein, G.A. Penner, N.A. Tinker, K.G. Briggs, D.E. Falk, and D.E. Mather. 2000. Marker-based selection of QTL affecting grain and malt quality in two-row barley. Crop Sci. 40:1426–1433.

Jannink, J.-L., M.C. Bink, and R.C. Jansen. 2001. Using complex plant pedigrees to map valuable genes. Trends Plant Sci. 6(8):337–342.

Jansen, R.C., J.-L. Jannink, and W.D. Beavis. 2003. Mapping quantitative trait loci in plant breeding populations: Use of parental haplotype sharing. Crop Sci. 43:829–834.

Keightley, P.D., and S.A. Knott. 1999. Testing the correspondence between map positions of quantitative trait loci. Genet. Res. (Cambridge) 74:323–328.

Knapp, S.J. 1998. Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. Crop Sci. 38:1164–1174.

Knapp, S.J., W.W. Stroup, and W.M. Ross. 1985. Exact confidence intervals for heritability on a progeny mean basis. Crop Sci. 25: 192–194.

Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124: 743–756.

Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M.J. Daly, S.E. Lincoln, and L. Newburg. 1987. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181.

Liu, Y., and Z.-B. Zeng. 2000. A general mixture model approach for mapping quantitative trait loci from diverse cross designs involving multiple inbred lines. Genet. Res. (Cambridge) 75:345–355.

Long, A.D., S.L. Mullaney, L.A. Reid, J.D. Frey, C.H. Langley, and T.F.C. Mackay. 1995. High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. Genetics 139:1273–1291.

Lübberstedt, T., D. Klein, and A. Melchinger. 1998a. Comparative QTL mapping of resistance to *Ustilago maydis* across four populations of European flint-maize. Theor. Appl. Genet. 97:1321–1330.

Lübberstedt, T., A. Melchinger, S. Fähr, D. Klein, A. Dally, and P. Westhoff. 1998b. QTL mapping in testcrosses of flint lines of maize: III. Comparison across populations for forage traits. Crop Sci. 38:1278–1289.

Lübberstedt, T., A.E. Melchinger, C.C. Schön, H.F. Utz, and D. Klein. 1997. QTL mapping in testcrosses of European flint lines of maize: I. Comparison of different testers for forage yield traits. Crop Sci. 37:921–931.

Mackay, T.F.C. 1995. The genetic basis of quantitative variation: Numbers of sensory bristles of *Drosophila melanogaster* as a model system. Trends Genet. 11:464–470.

Mackay, T.F.C. 1996. The nature of quantitative genetic variation revisited: Lessons from *Drosophila* bristles. Bioessays 18:113–121.

Mather, D.E., N.A. Tinker, D.E. LaBerge, M. Edney, B.L. Jones, B.G. Rossnagel, W.G. Legge, K.G. Briggs, R.B. Irvine, D.E. Falk, and K.J. Kasha. 1997. Regions of the genome that affect grain and malt quality in a North American two-row barley cross. Crop Sci. 37:544–554.

Melchinger, A.E., G.A. Schmidt, and H.H. Geiger. 1986. Evaluation of near infrared reflectance spectroscopy for predicting grain and stover quality traits in maize. Plant Breed. 97:20–29.

Melchinger, A.E., H.F. Utz, and C.C. Schön. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383–403.

Meuwissen, T.H.E., and M.E. Goddard. 2000. Fine mapping of quanti-

tative trait loci using linkage disequilibria with closely linked marker loci. Genetics 155:421–430.
Miller, A.J. 1990. Subset selection in regression. Chapman and Hall, London.
Mode, C.J., and H.F. Robinson. 1959. Pleiotropism and the genetic variance and covariance. Biometrics 15:518–537.
Otto, S.P., and C.D. Jones. 2000. Detecting the undetected: Estimating the total number of loci underlying a quantitative trait. Genetics 156:2093–2107.
Patterson, H.D., and E.R. Williams. 1976. A new class of resolvable incomplete block designs. Biometrika 63:83–92.
Pilet, M.L., G. Duplan, M. Archipiano, P. Barret, C. Baron, R. Horvais, X. Tanguy, M.O. Lucas, M. Renard, and R. Delourme. 2001. Stability of QTL for field resistance to blackleg across two genetic backgrounds in oilseed rape. Crop Sci. 41:197–205.
Rebai, A., and B. Goffinet. 2000. More about quantitative trait locus mapping with diallel designs. Genet. Res. (Cambridge) 75:243–247.
Schön, C.C., A.E. Melchinger, J. Boppenmaier, E. Brunklaus-Jung, R.G. Herrmann, and J.F. Seitzer. 1994. RFLP mapping in maize: Quantitative trait loci affecting testcross performance of elite European flint lines. Crop Sci. 34:378–389.
Searle, S.R. 1971. Linear Models. John Wiley & Sons, New York.
Shao, J. 1997. An asymptotic theory for linear model selection. Statistica Sinica 7:221–264.
Stromberg, L.D., J.W. Dudley, and G.K. Rufener. 1994. Comparing conventional early generation selection with molecular marker-assisted selection in maize. Crop Sci. 34:1221–1225.
Stuber, C.W. 1995. Mapping and manipulating quantitative traits in maize. Trends Genet. 11:477–481.
Stuber, C.W., M. Polacco, and M.L. Senior. 1999. Synergy of empirical breeding, marker-assisted selection, and genomics to increase crop yield potential. Crop Sci. 39:1571–1583.
Thomas, W.T.B., W. Powell, R. Waugh, K.J. Chalmers, U.M. Barua, P. Jack, V. Lea, B.P. Forster, J.S. Swanston, R.P. Ellis, P.R. Hanson, and R.C.M. Lance. 1995. Detection of quantitative trait loci for agronomic, yield, grain, and disease characters in spring barley (*Hordeum vulgare* L.). Theor. Appl. Genet. 91:1037–1047.
Thornsberry, J.M., M.M. Goodman, J. Doebley, S. Kresovich, D. Nielsen, and E.S. Buckler. 2001. *Dwarf8* polymorphisms associated with variation in flowering time. Nat. Genet. 28:286–289.
Tuberosa, R., M.C. Sanguineti, P. Landi, M.M. Giuliani, S. Salvi, and S. Conti. 2002. Identification of QTLs for root characteristics in maize grown in hydroponics and analysis of their overlap with QTLs for grain yield in the field at two water regimes. Plant Mol. Biol. 48:697–712.
Utz, H.F., and A.E. Melchinger. 1994. Comparison of different approaches to interval mapping of quantitative trait loci. p. 195–204. *In* J.W. van Ooijen and J. Jansen (ed.) Proc. of the 9th Meeting of the EUCARPIA Section Biometrics in Plant Breeding, Wageningen, the Netherlands. 6–8 July 1994. Centre for Plant Breeding and Reproduction Research, Wageningen, the Netherlands.
Utz, H.F., and A.E. Melchinger. 1996. PLABQTL: A program for composite interval mapping of QTL. J. Quant. Trait Loci 2(1) [no longer available. See: http://www.uni-hohenheim.de/~ipspwww/soft.html (verified 19 Aug. 2003)].
Utz, H.F., A.E. Melchinger, and C.C. Schön. 2000. Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. Genetics 154:1839–1849.
Van Ooijen, J.W., and R.E. Voorrips. 2001. JoinMap Version 3.0, software for the calculation of genetic linkage maps. Plant Research Int., Wageningen, the Netherlands.
Veldboom, L.R., and M. Lee. 1996. Genetic mapping of quantitative trait loci in maize in stress and nonstress environments: I. Grain yield and yield components. Crop Sci. 36:1310–1319.
Visscher, P.M., R. Thompson, and C.S. Haley. 1996. Confidence intervals in QTL mapping by bootstrapping. Genetics 143:1013–1020.
Xie, C., D.D.G. Gessler, and S. Xu. 1998. Combining different line crosses for mapping quantitative trait loci using the identical by descent-based variance component method. Genetics 149:1139–1146.
Xu, S. 1998. Mapping quantitative trait loci using multiple families of line crosses. Genetics 148:517–524.

Correlations and QTL Correspondence between Line Per Se and Testcross Performance for Agronomic Traits in Four Populations of European Maize

Renata Mihaljevic, Chris C. Schön, H. Friedrich Utz, and Albrecht E. Melchinger*

ABSTRACT

The magnitude of the genotypic correlation between line per se performance (LP) and testcross performance (TP) is crucial for optimum testing schemes in hybrid breeding as well as simultaneous improvement of commercial hybrids and their inbred parents. The objectives of this study were to (i) obtain estimates of the correlation, and (ii) determine quantitative trait loci (QTL) correspondence between LP and TP within four populations (F_3 to F_6 lines) derived from intrapool crosses of European flint maize (*Zea mays* L.). The number of lines evaluated for both LP and TP ranged from 65 to 280 across the four populations. The LP and TP with a dent inbred tester were evaluated for grain yield, grain moisture, kernel weight, protein concentration, and plant height in four to five environments. Composite interval mapping (CIM) using a joint restriction fragment length polymorphism (RFLP) linkage map of all populations was conducted separately for LP and TP in each population, with sample sizes ranging from 71 to 344. Genotypic correlations between LP and TP, $\hat{r}_g$ (LP, TP), were low to intermediate for grain yield (0.28–0.56) across populations and intermediate to high for the other traits (0.52–0.87). The magnitude of $\hat{r}_g$ (LP, TP) across populations for grain yield was neither associated with the ratio between the genotypic variances for LP and TP nor with the evidence for dominance in LP or epistasis in LP or TP. Genotypic correlations between observed TP and its prediction based on QTL positions and effects for LP were smaller than corresponding values of $\hat{r}_g$ (LP, TP) for all traits. Except for grain yield, more than half of the QTL were in common to LP and TP in the largest population A × B^1. Thus, it seems feasible to apply marker-assisted selection for TP based on QTL detected for LP, for traits with a large proportion of the genotypic variance accounted for by QTL.

Testcross performance of experimental lines is the prime selection criterion in hybrid breeding of maize. An indirect improvement of TP in early selfing generations by selecting for LP is economically advantageous, with a high positive correlation between LP and TP. Experimental estimates of the genotypic correlation between LP and TP, $\hat{r}_g$ (LP, TP), vary considerably for different crops, traits, and selfing generations. In maize, for traits showing small heterotic effects such as grain moisture, ear length, or days to flower, estimates of $\hat{r}_g$ (LP, TP) were medium to high. However, they were generally low for the highly heterotic trait, grain yield (for review see Hallauer and Miranda, 1981; Seitz, 1989).

In early studies, low values of $\hat{r}_g$ (LP, TP) observed for grain yield in advanced selfing generations were most probably due to recessive genes with detrimental effect in homozygous state (Genter and Alexander, 1966). Overdominance, epistasis, and linkage, or the combined action of these factors may also decrease the correlation (Schnell, 1961; Smith, 1986). For heterotic traits, estimates of $\hat{r}_g$ (LP, TP) with experimental lines from early selfing generations may be reduced because different levels of heterozygosity affect LP but not TP.

Assuming absence of linkage and epistasis, Smith (1986) demonstrated theoretically that low correlations between LP and TP can be fully explained by a model with additive and dominance effects. Thus, with biallelism and allele frequencies of 0.5 in a set of lines derived from a population in Hardy-Weinberg equilibrium, $\hat{r}_g$ (LP, TP) is a linear function of the proportion of QTL at which the inbred tester is homozygous for the favorable allele. As the latter increases, $\hat{r}_g$ (LP, TP) decreases due to a reduced genotypic variance (σ_g^2) for TP. Thus, the ratio of σ_g^2 for LP and TP provides a crude estimate of the proportion of dominant favorable alleles fixed in the tester.

The importance of epistatic interactions relative to the masking effect of dominant tester alleles for the reduction of $\hat{r}_g$ (LP, TP) can be assessed from quantitative genetic parameters. Differences among testcross means and changes in the ratios of segregation variances from different testcross generations are expected in the presence of linked epistatic effects (Melchinger, 1987). Such differences are not expected to occur if the masking effect of dominant tester alleles prevails.

While estimates of $\hat{r}_g$ (LP, TP) rely on the net effect of all QTL influencing LP and TP for a given trait, QTL analyses provide a tool to clarify the genetic basis of this correlation at the molecular level. The proportion of common QTL for LP and TP was largest for plant and ear height with an unrelated tester, and smallest for grain yield with a related tester (Austin et al., 2000). This was in accordance with the magnitude of genotypic correlations between LP and TP estimated for these traits. However, comparative QTL studies for LP and TP (Guffy et al., 1988; Beavis et al., 1994; Groh et al.; 1998; Kerns et al., 1999; Austin et al., 2000; Méchin et al., 2001) have so far not targeted the causes of the low genotypic correlations estimated in previous studies.

In this study, we evaluated four populations derived from three crosses of elite inbreds of European flint maize in different selfing generations (F_3 to F_6 lines) for both LP and TP. Our objectives were to (i) obtain reliable estimates of the correlation between LP and TP for five agronomic traits, (ii) examine possible causes for their magnitude by comparing genetic variances as well as the proportion of common QTL for LP and TP

R. Mihaljevic, H.F. Utz, and A.E. Melchinger, Institute of Plant Breeding, Seed Science, and Population Genetics, Univ. of Hohenheim, 70593 Stuttgart, Germany; and C.C. Schön, State Plant Breeding Institute, Univ. of Hohenheim, 70593 Stuttgart, Germany. Received 26 Feb. 2004. *Corresponding author (melchinger@uni-hohenheim.de).

Published in Crop Sci. 45:114–122 (2005).

677 S. Segoe Rd., Madison, WI 53711 USA

Abbreviations: CIM, composite interval mapping; DS, data set; ES, estimation set; LP, line per se performance; P1, parent one; P2, parent two; QTL, quantitative trait locus/loci; RFLP, restriction fragment length polymorphism; TP, testcross performance; TS, test set.

across populations and traits, and (iii) determine the gene action of QTL identified for LP and their value for the prediction of TP.

MATERIALS AND METHODS

Plant Materials

Four early maturing elite European flint lines KW1265, D146, D145, and KW1292, subsequently referred to as A, B, C, and D, were used as parents (P1 and P2) to produce four populations of 380 $F_{2:3}$ lines (A × B^{I}), 120 $F_{4:5}$ lines (A × B^{III}), and 131 (A × C) and 135 (C × D) $F_{3:4}$ lines. Superscripts I and III denote two different samples of the cross A × B according to the notation used in Mihaljevic et al. (2004). Testcross seed was produced in isolation plots by mating the unrelated dent inbred tester (KW5361, subsequently referred to as T2 in the notation of Schön et al., 1994), as pollinator to a random sample of 40 plants from each of the F_n lines ($F_{2:3}$ lines in A × B^{I}; $F_{4:5}$ lines in A × B^{III}; $F_{3:4}$ lines in A × C and C × D) as well as to the parent lines A, B, C, and D. Lines of each population except for A × B^{I} were further selfed, and the resulting $F_{4:6}$ lines of the cross A × B^{III} as well as $F_{3:5}$ lines of the crosses A × C and C × D were evaluated for LP (Table 1). For A × B^{I}, however, seed for evaluation of LP was produced by chain crossing of 20 plants of each $F_{2:3}$ line.

Field Experiments

The lines were evaluated for LP in separate experiments in the Upper Rhine valley. The experimental design employed was a 30 × 10 (A × B^{I}) and 15 × 10 (A × B^{III}, A × C, and C × D) α-design (Patterson and Williams, 1976) with two replications and one-row plots overplanted and later thinned to obtain a final stand of 8.7 plants m^{-2} in all experiments. All trials were conducted at five different environments. For populations A × B^{III}, A × C, and C × D, data from one environment was excluded from the combined analysis across environments due to severe drought stress (Table 1). The corresponding testcrosses evaluated in the same environment were far less affected by the unfavorable weather conditions. Each year-site combination was treated as an environment in subsequent statistical analyses.

The corresponding testcross progenies of the populations A × B^{I}, A × B^{III}, A × C, and C × D were evaluated for TP in separate experiments in the Upper Rhine valley, Lower Bavaria, and France, as described by Melchinger et al. (1998) and Mihaljevic et al. (2004). Population A × B^{I} was grown in four, the remaining three populations (A × B^{III}, A × C, and C × D) in five environments (Table 1). Because of insufficient quantities of seeds, fewer lines were tested for TP than LP in A × B^{III}, A × C, and C × D. The experimental design was a 40 × 10 (A × B^{I}) or a 15 × 10 α-design (A × B^{III}, A × C, and C × D) with two replications and two-row plots overplanted and later thinned to obtain a final stand of 8.7 plants m^{-2} in the Upper Rhine valley (two environments) and 11 plants m^{-2} in the other regions (three environments). All experiments were machine planted and harvested as grain trials with a combine. In the case of A × B^{III}, A × C, and C × D, one test environment was in common for LP and TP but none in the case of A × B^{I} (Table 1).

Data were analyzed for the following traits: grain yield (Mg ha^{-1}) adjusted to 155 g kg^{-1} grain moisture, grain moisture (g kg^{-1}) at harvest, kernel weight expressed as grams per 1000 kernels determined from four samples of 50 kernels from each plot, protein concentration in grain (g kg^{-1}) measured by near-infrared reflectance spectroscopy as described by Melchinger et al. (1986), and plant height (cm) on a plot basis as the distance from the soil level to the lowest tassel branch.

RFLP Marker Genotyping and Linkage Map Construction

The procedures for RFLP assays were described by Schön et al. (1994). A total of 89 RFLP marker loci was used to genotype 344 parental F_2 plants of the 380 $F_{2:3}$ lines from cross A × B^{I}, and 151 RFLPs were used to genotype parental F_4 plants of 120 $F_{4:5}$ or $F_{4:6}$ lines (A × B^{III}) (Table 1). A total of 104 and 122 RFLPs was mapped with 131 and 140 F_3 lines derived from cross A × C and C × D, respectively. The joint linkage map reported by Mihaljevic et al. (2004) comprising data of the four populations plus an additional population (independent sample A × B^{II} of cross A × B), formed the basis of all further analyses. The joint map is available at http://www.maizegdb.org (verified 3 Sept. 2004).

Agronomic Data Analyses

Adjusted entry means and effective error mean squares derived from ANOVAs of each environment (year–site–combination) were used to calculate the combined ANOVAs and ANCOVAs for each experiment. Quantitative genetic parameters, such as variance components and heritabilities, were estimated as described by Melchinger et al. (1998). An approximative F test was used to test whether the genotypic variance ($\hat{\sigma}^2_g$) for LP was larger than $\hat{\sigma}^2_g$ for TP. Degrees of freedom

Table 1. Dimensions of field experiments and of restriction fragment length polymorphism (RFLP) genotyping employed for the evaluation of line per se performance (LP) and testcross performance (TP) in four populations (A × B^{I}, A × B^{III}, A × C, and C × D) of European maize.

	Population							
	A × B^{I}		A × B^{III}		A × C		C × D	
Experiment	LP	TP	LP	TP	LP	TP	LP	TP
Generation	$F_{2:3}$	$F_{2:3}$	$F_{4:6}$	$F_{4:5}$	$F_{3:5}$	$F_{3:4}$	$F_{3:5}$	$F_{3:4}$
				Field experiments				
No. of entries	300	400	150	150	150	150	150	150
Parental lines (P1, P2)	10, 10	5, 5	10, 10	5, 5	7, 7	5, 5	5, 5	5, 5
F_n lines	280	380	120	71	131	109	135	84
Common F_n lines for LP and TP	280		65		109		82	
No. of environments	5	4	4	5	4	5	4	5
Common environments	0		1		1		1	
				RFLP genotyping				
No. of genotypes	344 F_2		120 F_4		131 F_3		140 F_3	
No. of loci	89		151		104		122	

for the one-tailed F test were calculated according to Satterthwaite (1946). Corresponding F tests were also employed to compare $\hat{\sigma}^2_g$ from different generations ($A \times B^I$ and $A \times B^{III}$).

Phenotypic, $\hat{r}_p$ (LP, TP), and genotypic, $\hat{r}_g$ (LP, TP), correlations were calculated between LP and TP using only the common lines (Table 1). Both types of correlation coefficients were calculated using the MANOVA estimators of adjusted entry means described by Liu et al. (1997). Here, the phenotypic covariance was used as an estimator of the genotypic covariance, assuming the covariance of genotype × environment interactions to be negligible. Empirical 95% confidence intervals of the correlation coefficients were estimated by 2000 bootstrap samples according to Liu et al. (1997).

QTL Analyses

QTL mapping and estimation of their effects were performed with PLABQTL (Utz and Melchinger, 1996) employing CIM by the regression approach (Haley and Knott, 1992). The additive genetic model underlying the analysis of TP was described in detail by Utz et al. (2000). For analyses of LP of the F_n lines, the following model was employed:

$$Y_j = m + b_1^* x_{ajl}^* + b_2^* x_{djl}^* + \sum_k b_k x_{jk} + \varepsilon_j,$$

where Y_j denotes the phenotypic trait mean of the jth F_n line averaged across environments; m is the phenotypic trait mean of F_n lines with genotype qq at the lth putative QTL; b_1^* and b_2^* are the additive (a) and the dominance (d, estimated only for $F_{2:3}$ lines of $A \times B^I$) effects as defined by Falconer and Mackay (1996, p. 112) at the putative QTL in the marker interval l with flanking markers l' and l''. x_{ajl}^* and x_{djl}^* are the conditional expectations of the dummy variables Θ_{ajl} and Θ_{djl} given the observed genotypes at the flanking marker loci l' and l'', where Θ_{ajl} assumes values 0, 1, or 2, and Θ_{djl} assumes values 0, 0.5, or 0 if the genotype of the parental F_n individual at the putative QTL is qq, Qq, or QQ, respectively. Θ_{djl} is 0.5 rather than 1 for heterozygotes Qq, because phenotypic traits were evaluated in $A \times B^I$ for $F_{2:3}$ lines and not F_2 plants, which reduces the dominance effect by one half. b_k is the partial regression coefficient of phenotype Y_j on the kth (selected) marker; x_{jk} is a dummy variable (cofactor) taking values 0, 1, or 2, depending on whether the marker genotype of the parental F_n individual j at marker locus k is homozygous qq, heterozygous Qq, or homozygous QQ, respectively. ε_j is a residual variable for the jth F_n line.

Cofactors were selected by stepwise regression according to Miller (1990, p. 49) with an "F-to-enter" and "F-to-delete" value of 3.5. Testing for presence of a putative QTL in an interval by a likelihood ratio (LR) test was performed by using a LOD threshold of 2.5 (= 0.217 LR). Estimates of QTL positions were obtained at the point where the LOD score assumed its maximum value in the region under consideration. For each population, the proportion of the phenotypic variance ($\hat{\sigma}^2_p$) explained by a single QTL was determined as the square of the partial correlation coefficient (R^2). Estimates of the additive effects (and dominance effects for $A \times B^I$) of each putative QTL for LP and their partial R^2 were obtained by fitting a model including all QTL for the respective trait simultaneously. The proportion p of the genotypic variance explained by all detected QTL was also determined from this model for each data set (DS) as $\hat{p}_{DS}$ by dividing the adjusted total R^2 (R^2_{adj}) by the heritability (h^2) as described by Utz et al. (2000).

Five-fold standard cross validation implemented in PLABQTL was used to obtain asymptotically unbiased estimates of p (Utz et al., 2000). For each population, the DS comprising the entry means across environments was divided into five genotypic subsamples. Four of these were combined in an estimation set (ES) for QTL detection and estimation of genetic effects, whereas the remaining fifth subsample was used as a test set (TS) to validate the predictions gained from ES and calculate $\hat{p}_{TS.ES}$ by correlating data predicted on the basis of QTL estimates in ES with those observed in the TS. Five different cross validation runs are possible by permutating the respective subsamples. A total of 1000 replicated cross validation runs was performed with 200 randomizations for assigning genotypes to the respective subsamples. The median $\tilde{p}_{TS.ES}$ was obtained from $\hat{p}_{TS.ES}$ across the 1000 runs.

Congruency of QTL for Line Per Se and Testcross Performance

We assessed congruency of QTL detected for LP and TP of a particular trait in the same population. Two approaches were used for this purpose: (i) counting the number of congruent QTL, whereby individual QTL were considered congruent if their estimated map position was within a 20-cM distance, irrespective of the sign of estimated QTL effects, and (ii) the genotypic correlation between predicted and observed testcross performance, $\hat{r}_g$ (M_{LP}, Y_{TP}), where M_{LP} is the predicted value of a line based on the QTL positions and effects estimated from QTL for LP in a given population, and Y_{TP} is the observed TP of this line (Utz et al., 2000).

RESULTS

Segregation and Linkage of RFLP Markers

The results of the RFLP analyses have been reported previously (Mihaljevic et al., 2004). The joint linkage map of the populations $A \times B^I$, $A \times B^{II}$, $A \times B^{III}$, $A \times C$, and $C \times D$ spanned a total of 1138 cM. This joint map covered about 70% of the genome from the original map of $A \times B^I$ published by Schön et al. (1994).

Agronomic Trait Analysis for Line Per Se Performance

The means of parents P1 and P2 differed significantly ($P < 0.01$) for all traits in all populations except for plant height in $A \times B^I$ and $A \times C$, grain moisture in $A \times B^{III}$ and $C \times D$, and grain yield and kernel weight in $C \times D$ (Table 2). An orthogonal contrast between the mean performance of the parent lines ($\overline{P}$) and the population mean of the F_n lines ($\overline{F}_n$) was highly significant ($P < 0.01$) for all traits in $A \times B^I$, in $A \times C$ for grain yield, and in $C \times D$ for plant height only. For grain yield, kernel weight, and plant height, $\overline{P}$ was significantly smaller than $\overline{F}_n$ in all of these cases. In contrast, $\overline{P}$ vs. $\overline{F}_n$ was not significant for any trait in $A \times B^{III}$.

Genotypic variances for LP were highly significant for all traits in all four populations (Table 3). As expected from quantitative genetic theory, the lines in $A \times B^I$ from an early selfing generation had a significantly ($P < 0.05$) smaller $\hat{\sigma}^2_g$ than lines in $A \times B^{III}$ from an advanced selfing generation for all traits except grain yield. For comparison, $\hat{\sigma}^2_g$ for TP was significantly smaller in $A \times B^I$ than in $A \times B^{III}$ for all traits.

Estimates of genotype × environment interaction variance ($\hat{\sigma}^2_{ge}$) for LP were significantly greater than zero ($P < 0.01$) and consistently smaller than $\hat{\sigma}^2_g$ for all

Table 2. Means of parents P1 and P2, 280 $F_{2:3}$ (A × B^{I}), 120 $F_{4:6}$ (A × B^{III}), 131 $F_{3:5}$ (A × C), and 135 $F_{3:5}$ (C × D) evaluated for line per se performance (LP) for five agronomic traits of European maize estimated in five or four environments.

Generation	Population A × B^{I}	A × B^{III}	A × C	C × D
	Mg ha^{-1}			
Grain yield				
P1	2.77 ± 0.22†	4.03 ± 0.27	3.92 ± 0.23	5.27 ± 0.30
P2	4.88 ± 0.22	5.79 ± 0.27	5.39 ± 0.23	4.48 ± 0.30
$\bar{P}$	3.83 ± 0.16	4.91 ± 0.19	4.65 ± 0.16	4.88 ± 0.21
$\bar{F}_n$	5.70 ± 0.06	4.83 ± 0.10	5.36 ± 0.08	5.35 ± 0.11
	g kg^{-1}			
Grain moisture				
P1	358.3 ± 3.8	318.8 ± 4.6	322.8 ± 3.8	343.4 ± 4.4
P2	342.0 ± 3.8	310.7 ± 4.6	344.4 ± 3.8	329.0 ± 4.4
$\bar{P}$	350.2 ± 2.7	314.7 ± 3.3	333.6 ± 2.7	336.2 ± 3.1
$\bar{F}_n$	333.9 ± 0.8	311.0 ± 1.9	339.4 ± 1.4	330.3 ± 1.5
	g			
Kernel weight				
P1	264.0 ± 3.6	275.8 ± 5.8	269.7 ± 5.0	192.3 ± 3.3
P2	225.3 ± 3.6	228.2 ± 5.8	190.6 ± 5.0	194.1 ± 3.3
$\bar{P}$	244.7 ± 2.6	252.0 ± 4.1	230.1 ± 3.5	193.2 ± 2.4
$\bar{F}_n$	264.2 ± 1.3	258.6 ± 2.2	230.7 ± 1.5	194.6 ± 1.8
	g kg^{-1}			
Protein concentration				
P1	129.7 ± 1.0	126.5 ± 1.6	127.6 ± 1.5	98.39 ± 1.7
P2	115.6 ± 1.0	113.6 ± 1.6	96.6 ± 1.5	120.5 ± 1.7
$\bar{P}$	122.6 ± 0.7	120.1 ± 1.2	112.1 ± 1.1	109.4 ± 1.2
$\bar{F}_n$	118.0 ± 0.3	119.1 ± 0.7	110.8 ± 0.6	108.7 ± 0.8
	cm			
Plant height				
P1	171.4 ± 2.1	180.3 ± 2.3	179.5 ± 2.7	169.3 ± 2.6
P2	167.3 ± 2.1	171.4 ± 2.3	177.0 ± 2.7	130.9 ± 2.6
$\bar{P}$	169.3 ± 1.5	175.9 ± 1.6	178.3 ± 1.9	150.1 ± 1.8
$\bar{F}_n$	184.4 ± 0.6	176.8 ± 1.3	180.1 ± 1.0	159.4 ± 1.1

† Standard errors are attached.

traits in all populations (Table 3). Heritabilities ($\hat{h}^2$) were high for all traits ranging from 0.88 to 0.95 across traits and populations.

Comparison of Line Per Se and Testcross Performance

Mihaljevic et al. (2004) reported results of testcross progeny analysis for A × B^{I}, A × B^{III}, A × C, and C × D. In all four crosses, the population mean $\bar{F}_n$ for LP was lower than $\bar{F}_n$ for TP for all traits except grain moisture and protein concentration (Table 2). The range of F_n lines for LP was larger than for TP in all populations and for all traits (data not shown).

As expected, estimates of σ^2_g for LP were significantly greater than those for TP in all populations and for all traits. Estimates of σ^2_{gc} also were generally greater for LP than for TP, except for grain yield in A × B^{III} and A × C (Table 3).

Phenotypic correlations between LP and TP, $\hat{r}_p$ (LP, TP), were low for grain yield, but significant in all populations (Table 4). For the other traits, $\hat{r}_p$ (LP, TP) values were intermediate ($0.40 < \hat{r}_p < 0.75$). Genotypic correlations between LP and TP, $\hat{r}_g$ (LP, TP), were significant and always greater than $\hat{r}_p$ (LP, TP) across all traits and populations. Estimates of $\hat{r}_g$ (LP, TP) ranged from 0.28 to 0.56 for grain yield and from 0.52 to 0.87 for the other four traits (Table 4).

QTL Analyses of Line Per Se Performance

Results from QTL analyses for LP of all four populations based on the joint map are presented here for means across environments (Table 4). Detailed information on the position and magnitude of effects of individual QTL can be obtained at http://www.maizegdb.org. In the large population A × B^{I}, substantially more QTL were detected than in the smaller populations. The number of congruent QTL detected across the four populations was low. Most QTL found for A × B^{III} were also found in A × B^{I}. Only one QTL with dominant gene action was detected for grain yield in A × B^{I}. The QTL results for TP were reported previously (Mihaljevic et al., 2004).

Comparison of QTL for Line Per Se and Testcross Performance

Across all five traits in A × B^{I}, 21 out of 44 QTL detected for LP were found within a 20-cM distance from QTL detected for TP (Table 4). The relationship between the number of common QTL for LP and TP and the total number of QTL detected for LP was lowest for grain yield. In the advanced generation of cross A × B (A × B^{III}), five out of eight QTL detected for LP were common to QTL detected for TP across all five traits. Out of 24 QTL detected in A × C for LP, 10 QTL were within a 20-cM distance to QTL detected for TP for the same trait. In C × D, six out of 24 QTL

Table 3. Variance components and heritabilities of 280 F_{23} ($A \times B^{I}$), 120 F_{46} ($A \times B^{III}$), 131 F_{35} ($A \times C$), and 135 F_{35} ($C \times D$) lines evaluated for line per se performance (LP) as well as 380 F_{23} ($A \times B^{I}$), 71 F_{45} ($A \times B^{III}$), 109 F_{34} ($A \times C$), and 84 F_{34} ($C \times D$) lines evaluated for testcross performance (TP) with tester T2 for five agronomic traits of European maize estimated in five or four environments.

	Population							
	A × B[I]		A × B[III]		A × C		C × D	
Parameter	LP	TP	LP	TP	LP	TP	LP	TP
	Mg ha^{-1}							
Grain yield								
$\hat{\sigma}^2_g$†	0.877 ± 0.081**†	0.129 ± 0.020**	1.034 ± 0.148**	0.492 ± 0.119**	0.719 ± 0.098**	0.271 ± 0.061**	1.447 ± 0.194**	0.201 ± 0.045**
$\hat{\sigma}^2_{ge}$†	0.170 ± 0.022**	0.155 ± 0.028**	0.274 ± 0.037**	0.825 ± 0.091**	0.183 ± 0.025**	0.619 ± 0.061**	0.306 ± 0.044**	0.175 ± 0.038**
$\hat{\sigma}^2_e$†	0.536 ± 0.022	0.811 ± 0.032	0.383 ± 0.025	0.494 ± 0.029	0.284 ± 0.018	0.505 ± 0.029	0.527 ± 0.034	0.548 ± 0.032
$\hat{h}^2$	0.91	0.48	0.90	0.70	0.90	0.61	0.91	0.69
95% C.I. on $\hat{h}^2$	0.89–0.92	0.38–0.56	0.86–0.92	0.55–0.79	0.86–0.92	0.46–0.71	0.88–0.93	0.56–0.78
	g kg^{-1}							
Grain moisture								
$\hat{\sigma}^2_g$†	135.11 ± 12.95**	59.44 ± 5.30**	378.10 ± 52.64**	76.81 ± 14.49**	246.86 ± 33.2**	53.31 ± 8.49**	268.83 ± 36.29**	46.40 ± 8.55**
$\hat{\sigma}^2_{ge}$†	53.07 ± 4.11**	20.53 ± 2.57**	80.24 ± 9.80**	23.93 ± 4.46**	53.01 ± 6.96**	20.32 ± 3.59**	69.63 ± 8.31**	18.10 ± 3.88**
$\hat{\sigma}^2_e$†	74.57 ± 3.05	65.10 ± 2.56	87.02 ± 5.67	51.76 ± 3.02	76.15 ± 4.90	54.42 ± 3.14	82.72 ± 5.33	55.32 ± 3.18
$\hat{h}^2$	0.88	0.82	0.92	0.88	0.92	0.85	0.91	0.84
95% C.I. on $\hat{h}^2$	0.86–0.90	0.78–0.84	0.90–0.94	0.83–0.92	0.89–0.94	0.79–0.89	0.88–0.93	0.76–0.88
	g							
Kernel weight								
$\hat{\sigma}^2_g$†	415.50 ± 36.92**	76.85 ± 6.62**	555.21 ± 77.64**	180.73 ± 32.63**	269.24 ± 37.49**	135.28 ± 20.14**	394.73 ± 50.92**	157.26 ± 25.97**
$\hat{\sigma}^2_{ge}$†	45.96 ± 5.34**	19.03 ± 2.77**	127.25 ± 15.16**	32.31 ± 6.73**	91.44 ± 10.36**	22.69 ± 5.42**	33.87 ± 7.10**	12.45 ± 5.35**
$\hat{\sigma}^2_e$†	128.20 ± 5.24	74.14 ± 2.93	129.55 ± 8.44	84.42 ± 4.89	94.27 ± 6.06	93.88 ± 5.30	107.24 ± 6.91	94.21 ± 5.43
$\hat{h}^2$	0.95	0.85	0.92	0.92	0.89	0.91	0.95	0.93
95% C.I. on $\hat{h}^2$	0.94–0.96	0.82–0.87	0.89–0.94	0.89–0.95	0.85–0.91	0.87–0.93	0.93–0.96	0.90–0.95
	g kg^{-1}							
Protein concentration								
$\hat{\sigma}^2_g$†	27.14 ± 2.48**	5.10 ± 0.50**	50.63 ± 6.98**	15.39 ± 2.80**	39.03 ± 5.21**	10.43 ± 1.61**	72.96 ± 9.51**	13.85 ± 2.33**
$\hat{\sigma}^2_{ge}$†	3.24 ± 0.57**	2.01 ± 0.28**	10.12 ± 1.13**	3.35 ± 0.63**	8.07 ± 0.98**	2.96 ± 0.56**	9.68 ± 1.50**	2.39 ± 0.56**
$\hat{\sigma}^2_e$†	15.76 ± 0.64	5.78 ± 0.26	8.68 ± 0.54	7.28 ± 0.42	9.71 ± 0.62	8.80 ± 0.50	19.38 ± 1.22	8.40 ± 0.48
$\hat{h}^2$	0.92	0.76	0.93	0.92	0.92	0.88	0.94	0.91
95% C.I. on $\hat{h}^2$	0.91–0.94	0.71–0.80	0.91–0.95	0.88–0.94	0.90–0.94	0.83–0.91	0.92–0.95	0.88–0.94
	cm							
Plant height								
$\hat{\sigma}^2_g$†	102.8 ± 9.42**	33.21 ± 2.64**	202.6 ± 27.40**	40.54 ± 7.57**	113.8 ± 15.54**	52.90 ± 7.88**	152.6 ± 20.10**	40.59 ± 7.02**
$\hat{\sigma}^2_{ge}$†	14.07 ± 2.24**	4.69 ± 0.97**	19.07 ± 3.49**	8.18 ± 2.22**	26.42 ± 3.78**	8.58 ± 2.15**	20.86 ± 3.74**	4.78 ± 2.30*
$\hat{\sigma}^2_e$†	60.39 ± 2.46	47.71 ± 1.26	45.33 ± 2.95	31.63 ± 1.83	44.81 ± 2.88	37.57 ± 2.16	52.40 ± 3.36	41.73 ± 1.19
$\hat{h}^2$	0.92	0.91	0.95	0.89	0.90	0.91	0.93	0.89
95% C.I. on $\hat{h}^2$	0.90–0.93	0.90–0.92	0.93–0.96	0.84–0.93	0.87–0.93	0.87–0.93	0.90–0.95	0.84–0.92

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
† Standard errors are attached.

Table 4. Phenotypic ($\hat{r}_p$) and genotypic ($\hat{r}_g$) correlations between line per se performance (LP) and testcross performance (TP), the number of quantitative trait loci (QTL) detected for LP and TP as well as the number of common QTL, the proportion of the genotypic variance ($\tilde{p}_{TS,ES}$) explained by these QTL for five agronomic traits of European maize, and the genotypic correlation between LP and TP based on estimated QTL $\hat{r}_g$ (M_{LP}, Y_{TP}).

Parameter	Population			
	A × B[I]	A × B[III]	A × C	C × D
		Grain yield		
$\hat{r}_p$ (LP, TP)	0.19 (0.09; 0.30)†	0.33 (0.07; 0.58)	0.38 (0.21; 0.54)	0.42 (0.23; 0.57)
$\hat{r}_g$ (LP, TP)	0.28 (0.13; 0.44)	0.45 (0.07; 0.87)	0.54 (0.30; 0.78)	0.56 (0.33; 0.75)
$\hat{r}_g$ (M_{LP}, Y_{TP})‡	0.23 (0.08; 0.36)	0.37 (0.00; 0.50)	0.35 (0.03; 0.55)	0.47 (0.00; 0.67)
No. of QTL (LP)	9	2	3	3
No. of QTL (TP)	2	7	6	6
No. of common QTL	1	1	1	1
$\tilde{p}_{TS,ES}$ (%) (LP)§	27.4	3.5	12.3	3.8
$\tilde{p}_{TS,ES}$ (%) (TP)§	18.7	8.2	51.8	35.9
		Grain moisture		
$\hat{r}_p$ (LP, TP)	0.62 (0.55; 0.69)	0.68 (0.54; 0.79)	0.61 (0.49; 0.72)	0.40 (0.20; 0.56)
$\hat{r}_g$ (LP, TP)	0.73 (0.65; 0.81)	0.84 (0.70; 0.98)	0.74 (0.60; 0.89)	0.52 (0.27; 0.74)
$\hat{r}_g$ (M_{LP}, Y_{TP})‡	0.40 (0.29; 0.47)	0.15 (0.02; 0.30)	0.34 (0.19; 0.45)	0.43 (0.21; 0.55)
No. of QTL (LP)	5	1	7	9
No. of QTL (TP)	9	3	7	6
No. of common QTL	3	0	1	3
$\tilde{p}_{TS,ES}$ (%) (LP)§	13.5	2.1	22.2	28.5
$\tilde{p}_{TS,ES}$ (%) (TP)§	33.0	3.1	5.2	2.5
		Kernel weight		
$\hat{r}_p$ (LP, TP)	0.59 (0.50; 0.67)	0.72 (0.59; 0.82)	0.64 (0.52; 0.73)	0.67 (0.52; 0.77)
$\hat{r}_g$ (LP, TP)	0.66 (0.57; 0.75)	0.79 (0.67; 0.90)	0.72 (0.61; 0.83)	0.71 (0.56; 0.82)
$\hat{r}_g$ (M_{LP}, Y_{TP})‡	0.53 (0.39; 0.63)	0.46 (0.11; 0.64)	0.49 (0.39; 0.64)	0.26 (0.14; 0.48)
No. of QTL (LP)	10	2	3	2
No. of QTL (TP)	10	3	4	4
No. of common QTL	6	2	1	0
$\tilde{p}_{TS,ES}$ (%) (LP)§	21.6	9.4	14.9	12.2
$\tilde{p}_{TS,ES}$ (%) (TP)§	42.3	26.6	13.5	13.5
		Protein concentration		
$\hat{r}_p$ (LP, TP)	0.62 (0.53; 0.69)	0.73 (0.58; 0.84)	0.69 (0.54; 0.80)	0.72 (0.60; 0.81)
$\hat{r}_g$ (LP, TP)	0.74 (0.64; 0.84)	0.82 (0.67; 0.92)	0.78 (0.62; 0.90)	0.79 (0.66; 0.89)
$\hat{r}_g$ (M_{LP}, Y_{TP})‡	0.39 (0.27; 0.51)	0.55 (0.30; 0.65)	0.49 (0.22; 0.66)	0.55 (0.45; 0.66)
No. of QTL (LP)	7	2	7	5
No. of QTL (TP)	9	6	6	4
No. of common QTL	4	2	5	1
$\tilde{p}_{TS,ES}$ (%) (LP)§	22.6	7.6	7.6	15.9
$\tilde{p}_{TS,ES}$ (%) (TP)§	38.9	9.8	16.6	19.5
		Plant height		
$\hat{r}_p$ (LP, TP)	0.68 (0.61; 0.74)	0.70 (0.46; 0.86)	0.75 (0.61; 0.85)	0.52 (0.36; 0.65)
$\hat{r}_g$ (LP, TP)	0.81 (0.74; 0.87)	0.80 (0.51; 1.00)	0.87 (0.72; 0.99)	0.60 (0.42; 0.74)
$\hat{r}_g$ (M_{LP}, Y_{TP})‡	0.65 (0.57; 0.72)	0.34 (0.31; 0.55)	0.58 (0.12; 0.75)	0.55 (0.35; 0.64)
No. of QTL (LP)	13	1	4	5
No. of QTL (TP)	12	1	5	3
No. of common QTL	7	0	2	1
$\tilde{p}_{TS,ES}$ (%) (LP)§	35.2	16.4	5.0	19.3
$\tilde{p}_{TS,ES}$ (%) (TP)§	49.3	−0.3	22.4	12.8

† Empirical 95% confidence interval.
‡ Correlation between the observed TP and predicted genotypic values on the basis of QTL positions and effects derived from LP, divided by the heritability.
§ Proportion of genotypic variance (p) explained in the test set (TS) by all QTL detected with five-fold cross validation in the estimation set (ES) given as median ($\tilde{p}_{TS,ES}$) across 1000 replicated cross validation runs.

detected for LP across all five traits were common to QTL detected for TP of the same traits.

Estimates of the genotypic correlation between predicted and observed testcross performance, $\hat{r}_g$ (M_{LP}, Y_{TP}), varied considerably across populations for all traits (Table 4). For grain yield, $\hat{r}_g$ (M_{LP}, Y_{TP}) was highest in C × D and lowest in A × B[I], which was unexpected considering the difference in population size. For the other traits, $\hat{r}_g$ (M_{LP}, Y_{TP}) was highest (0.61) for plant height in A × B[I], and lowest (0.15) for grain moisture in A × B[III].

The number of common QTL generally was not reflected in the magnitude of $\hat{r}_g$ (M_{LP}, Y_{TP}) (Table 4). For grain moisture and plant height in A × B[III] and kernel weight in C × D, significant correlations $\hat{r}_g$ (M_{LP}, Y_{TP}) were detected in spite of zero common QTL between LP and TP. The correlations $\hat{r}_g$ (LP, TP) and $\hat{r}_g$ (M_{LP}, Y_{TP}) corresponded well for grain yield. This was not the case for the other four traits, where $\hat{r}_g$ (LP, TP) was substantially higher than $\hat{r}_g$ (M_{LP}, Y_{TP}) except for grain moisture and plant height in C × D.

DISCUSSION

Correlations between Line Per Se and Testcross Performance

The magnitude of the genotypic correlation between LP and TP is an indicator of the prospects of simultaneously improving commercial hybrids as well as their inbred parents. In maize, a wide range of estimates for phenotypic and genotypic correlations between LP and TP

was reported in the literature, depending on the trait investigated (for review see Hallauer and Miranda, 1981). In our study, genotypic correlations estimated for LP and TP across four populations derived from crosses within the European flint pool were comparable with those obtained for U.S. dent material. Lowest estimates were found for grain yield [$\hat{r}_g$ (LP, TP) = 0.28–0.56]. As expected for traits with higher heritability and presumably mainly additive gene action, such as grain moisture, kernel weight, protein concentration, and plant height, estimates of the respective correlation were generally high [$\hat{r}_g$ (LP, TP) > 0.7] across all four populations with only a few exceptions.

Genotypic correlations were higher than phenotypic correlations for all traits and populations. As expected from theory, when LP and TP are evaluated in different environments, the difference between the genotypic and the phenotypic correlations is a function of the heritability for LP and TP for the respective cross. For grain yield, heritability estimates for TP were smaller compared with LP mainly due to the reduced genotypic variance, slightly lower testing intensity (A × B^{I}), or a higher $\hat{\sigma}^2_{ge}$ for TP than LP. The $\hat{\sigma}^2_{ge}$ of TP in A × B^{III} and A × C were larger than in C × D for grain yield, although all three populations were tested in the same five environments. Thus, TP of lines from C × D seems to be more robust against environmental changes than TP of lines from A × B and A × C. For the other four traits, $\hat{\sigma}^2_{ge}$ was consistently larger for LP than for TP, resulting in similar heritability estimates despite a significant decrease in $\hat{\sigma}^2_g$ for TP.

The decrease in $\hat{\sigma}^2_g$ for TP compared with LP can be used as an indication of the strength (performance level or gene frequency) of the tester and of the expected genotypic correlation between LP and TP. For a tester, which carries dominant alleles masking the effect of the segregating alleles at many loci, σ^2_g for TP is decreased and correlations are expected to be lower. Smith (1986) showed that with complete dominance and a gene frequency of 0.5 in the population under study, the genotypic correlation between LP and TP is inversely proportional to the ratio of σ^2_g for LP and TP. For the biallelic case and an above average inbred tester from the same population, the genotypic correlation between LP and TP would be 0.5 or lower (Smith, 1986).

Considering all four populations and all traits, no significant association was found between the ratio of the two variances and $\hat{r}_g$ (LP, TP). The ratio of $\hat{\sigma}^2_g$ for LP vs. TP varied from 2.0 (kernel weight in A × C) to 7.2 (grain yield in C × D). Highest variance ratios were obtained for grain yield, as expected for a trait presumably controlled by many genes with large dominance effects, but only in A × B^{I} and C × D. Despite surprisingly low ratios for grain yield in A × B^{III} and A × C (2.1 and 2.6, respectively), genotypic correlations in these two crosses were intermediate.

Reasons can be given for the difficulties in predicting genotypic correlations from this ratio. First, Smith (1986) had assumed the biallelic case with the tester originating from the same population as the test units. In our study, however, the inbred tester originated from the opposite dent pool and was known for its excellent combining ability for yield with the flint pool.

Second, lines in all four populations had different levels of inbreeding. Different from TP, LP of an F_3 line for a heterotic trait like grain yield is affected by the heterozygosity level of its parental F_2 plant. However, despite a wide range in heterozygosity at marker loci (28.3 to 75.4%) in the F_2 plants of population A × B^{I}, this parameter showed only a weak correlation ($\hat{r}_g$ = 0.13, $P < 0.05$) with LP for grain yield (data not shown). These results were in accordance with the detection of only one out of nine QTL with dominant gene action for LP of grain yield.

Third, the low precision in estimating genotypic correlations (see large confidence intervals of the estimates presented in Table 4) could be a further explanation for the lack of association between the magnitude of genotypic correlations and the reduction in genotypic variance in the testcrosses. For grain yield and population A × B^{III} for example, the 95% confidence interval for the estimate of r_g (LP, TP) ranged from 0.07 to 0.87. Highest precision, that is, smallest confidence intervals, was obtained for plant height and grain moisture in population A × B^{I}, with the highest number of common lines tested for both LP and TP (N = 280). This is in agreement with results from Liu et al. (1997), who found that the heritability of the trait and sample size had a strong effect on the precision of estimates of genotypic correlations.

QTL Detected for Line Per Se and Testcross Performance

When comparing QTL mapping results for LP and TP across populations, with the exception of grain yield, generally fewer QTL were detected in populations A × B^{III}, A × C, and C × D than in A × B^{I}, reflecting the decreased power of QTL detection with smaller sample sizes. The same was true for the proportion of $\hat{\sigma}^2_g$ explained by QTL estimated from cross validation. For TP and LP similar numbers of QTL were detected in a given population for all traits except grain yield. The higher heritabilities and the slightly larger sample sizes in LP trials as compared with TP trials did not have a significant effect on the number of QTL detected. For grain yield, however, substantially fewer QTL were detected for TP of population A × B^{I} than in the other populations and for LP. In addition to genetic factors, sampling could be a reason for these results. With cross validation, Utz et al. (2000) showed for TP of population A × B^{I} that the number of detected QTL for grain yield can vary from zero to eight, depending on the genotypic sample used for QTL detection. In cross validation of LP data from A × B^{I}, the number of QTL detected for grain yield varied from 3 to 11.

Evidence for genetic factors, such as dominance and epistasis, which influence both heterosis and the correlation between LP and TP, should have been provided by the QTL analysis. It was surprising, however, that in the LP of population A × B^{I}, only one of the nine QTL exhibited dominant gene action for grain yield, and only

one pair of marker loci had a significant additive × additive epistatic effect. In the smaller populations and across all traits, epistatic effects were rarely detected. One reason for these results could be that the level of dominance for LP detected in the segregating intrapool population may not be a valid estimate for the importance of dominant allelic interactions with the tester from the opposite gene pool. Moreover, the estimation error is high for the level of dominance of QTL effects (Falconer and Mackay, 1996) especially in $F_{2:3}$ lines with only half the dominance effect assessed compared with F_2 plants. These statistical limitations apply even more to the estimation of additive × dominance or dominance × dominance type of interaction effects. It is, therefore, not surprising that controversial results can arise from the same data depending on the statistical model used for analysis (Cockerham and Zeng, 1996). Furthermore, choosing the correct model for estimation of epistatic effects is complicated, because additive × dominance epistatic effects frequently become significant only if their corresponding main effects are dropped from the model but not if they are included.

Thus, convincing evidence for allelic or nonallelic interactions at the QTL level could not be detected in our study, neither for LP nor for TP. The investigation of epistatic effects seems promising only if few genes regulate the trait under study and pairs of candidate loci are chosen a priori.

QTL Regions Common to Line Per Se and Testcross Performance

Analogous to a high genotypic correlation between LP and TP, a high congruency of QTL identified in both types of progenies is desirable. Beavis et al. (1994) and Austin et al. (2000) found little congruency of yield QTL detected for LP and TP. In this study, more than half of the QTL regions detected were in common for LP and TP in $A \times B^I$ for all traits except grain yield. The number of detectable common QTL may have been reduced in this study because our joint map covered only 70% of the genome covered by the reference map (Schön et al., 1994). Furthermore, considering that the power of QTL detection was smaller than 100% in both samples, and that the probability of simultaneous detection of a QTL in both progeny types is obtained by multiplication, these results meet expectations. Melchinger et al. (1998) found similar results for the congruency of QTL between two testcross series derived from $A \times B^I$. With the exception of grain yield, their QTL mapping results agreed between testers for a number of traits and more than half of the QTL detected with one tester were also found with the other tester. Thus, we conclude that for traits with mainly additive gene action, such as grain moisture, kernel weight, protein concentration, and plant height, QTL detected for LP should be predictive for TP.

To assess the value of QTL identified for LP in predicting TP, we calculated the genotypic correlation $\hat{r}_g$ (M_{LP}, Y_{TP}). Except for grain moisture and plant height in $A \times B^{III}$ and kernel weight in C × D, at least one common QTL could be detected for the two types of progenies for all traits and all four populations. However, the number of detected QTL was not indicative of the magnitude of $\hat{r}_g$ (M_{LP}, Y_{TP}). For example, for grain yield, one common QTL was detected in all four populations, but $\hat{r}_g$ (M_{LP}, Y_{TP}) ranged from 0.23 to 0.47 due to the differences in partial R^2 explained by the respective QTL. On the other hand, even with zero common QTL, a correlation significantly different from zero could be observed for plant height in $A \times B^{III}$ and kernel weight in C × D. This must be attributed to (i) QTL detected for LP but with effects below the detection threshold for TP or (ii) QTL linked to those detected for LP. Whether the choice of LOD threshold in QTL mapping for LP has an effect on the magnitude of $\hat{r}_g$ (M_{LP}, Y_{TP}) needs to be investigated. Using cross validation, Schön et al. (2004) showed that with a less conservative threshold in QTL estimation, on average, a larger proportion of the genotypic variance could be predicted in test sets.

Estimates of $\hat{r}_g$ (M_{LP}, Y_{TP}) were smaller than those of $\hat{r}_g$ (LP, TP) for all traits in all populations, because $\hat{r}_g$ (M_{LP}, Y_{TP}) can only be predictive for the proportion of genotypic variance explained by the QTL for LP ($\tilde{p}_{TS:ES}$), which was generally smaller than 50%. The magnitude of $\hat{r}_g$ (M_{LP}, Y_{TP}) should vary for the different traits under study and be a function of the validated genotypic variance explained by the QTL for LP. However, the experimental data only partially confirmed these expectations. A major reason could be the lack of precision in estimates of r_g (M_{LP}, Y_{TP}) shown by the large confidence intervals especially for the three smaller populations, which was most pronounced for grain yield.

Implications for Hybrid Maize Breeding

The magnitude of the genotypic correlation estimated for LP and TP of four different crosses were in accordance with earlier published results on U.S. dent material. Results for traits with mainly additive gene action, such as grain moisture, kernel weight, protein concentration, and plant height, were encouraging with respect to early selection for LP and indirect improvement of TP. For these traits, more than half the QTL detected for LP and TP were in common. Probably because of the limited power of QTL detection especially in the smaller populations, the proportion of $\hat{\sigma}^2_g$ explained by QTL for LP was medium to low, and thus resulted in a relatively low correlation between the marker-predicted and the observed TP. With sufficiently large sample sizes for QTL estimation and independent validation, it seems feasible, however, to apply marker-assisted selection based on QTL detected for LP if a substantial proportion of $\hat{\sigma}^2_g$ can be accounted for. For grain yield, r_g (LP, TP) were low, though always greater than the prediction based on markers. Therefore, the application of marker-assisted selection and/or phenotypic selection for LP to improve TP must be evaluated economically. Because of statistical limitations, it was not possible to separate genetic effects such as dominance or epistatic interactions to obtain an unambiguous explanation for the low correlations between LP and TP, neither from

the analysis of the phenotypic data nor from the results of the QTL analyses. Therefore, the expansion of the theoretical and simulation study performed by Smith (1986) to the multi-allelic case with different levels of dominance warrants further research.

ACKNOWLEDGMENTS

The present study was part of EUREKA project 290 supported by grants from the German Ministry of Research and Technology (BMBF) and KWS Kleinwanzlebener Saatzucht AG, grant 0319233A. It was also supported by a grant from the Deutsche Forschungsgemeinschaft, Grant No. ME 931/3-1. The RFLP assays were conducted in the lab of Prof. Dr. R.G. Herrmann, Ludwig-Maximilians-Universität in Munich, by E. Brunklaus-Jung and J. Boppenmaier as well as A. Dally in the lab of Prof. Dr. P. Westhoff at the Heinrich-Heine-Universität in Düsseldorf. The skilled technical assistance of F. Mauch, D. Schilling-Groß, A. Vesting, and the staff at the Plant Breeding Research Station in Eckartsweier in conducting field trials is gratefully acknowledged.

REFERENCES

Austin, D.F., M. Lee, L.R. Veldboom, and A.R. Hallauer. 2000. Genetic mapping in maize with hybrid progeny across testers and generations: Grain yield and grain moisture. Crop Sci. 40:30–39.

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

Cockerham, C.C., and Z.-B. Zeng. 1996. Design III with marker loci. Genetics 143:1437–1456.

Falconer, D.S., and T.F.C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Longman Group, London.

Genter, C.F., and M.W. Alexander. 1966. Development and selection of productive S_1 inbred lines of corn (*Zea mays* L.). Crop Sci. 6:429–431.

Groh, S., M.M. Khairallah, D. Gonzalez-de-León, M. Willcox, C. Jiang, D.A. Hoisington, and A.E. Melchinger. 1998. Comparison of QTLs mapped in RILs and their test-cross progenies of tropical maize for insect resistance and agronomic traits. Plant Breed. 117: 193–202.

Guffy, R.D., C.W. Stuber, and T. Helentjaris. 1988. Molecular markers for evaluating quantitative traits across varying genetic backgrounds and environments in maize. p. 82. *In* 1988 Agronomy abstracts. ASA, Madison, WI.

Haley, C.S., and S.A. Knott. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69:315–324.

Hallauer, A.R., and J.B. Miranda, Fo. 1981. Quantitative genetics in maize breeding. Iowa State Univ. Press, Ames, IA.

Kerns, M.R., J.W. Dudley, and G.K. Rufener II. 1999. Tester and type of progeny affect QTL detection in maize. Maydica 44:69–83.

Liu, B.H., S.J. Knapp, D. Birkes. 1997. Sampling distributions, biases, variances, and confidence intervals for genetic correlations. Theor. Appl. Genet. 94:8–19.

Méchin, V., O. Argillier, Y. Hébert, E. Guingo, L. Moreau, A. Charcosset, and Y. Barrière. 2001. Genetic analysis and QTL mapping of cell wall digestibility and lignification in silage maize. Crop Sci. 41:690–697.

Melchinger, A.E. 1987. Expectation of means and variances of testcrosses produced from F_2 and backcross individuals and their selfed progenies. Heredity 59:105–115.

Melchinger, A.E., G.A. Schmidt, and H.H. Geiger. 1986. Evaluation of near infrared reflectance spectroscopy for predicting grain and stover quality traits in maize. Plant Breed. 97:20–29.

Melchinger, A.E., H.F. Utz, and C.C. Schön. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383–403.

Mihaljevic, R., H.F. Utz, and A.E. Melchinger. 2004. Congruency of quantitative trait loci detected for agronomic traits in testcrosses of five populations of European maize. Crop Sci. 44:114–124.

Miller, A.J. 1990. Subset selection in regression. Chapman and Hall, London.

Patterson, H.D., and E.R. Williams. 1976. A new class of resolvable incomplete block designs. Biometrika 63:83–92.

Satterthwaite, F.E. 1946. An approximate distribution of estimates of variance components. Biometrics 2:110–114.

Schnell, F.W. 1961. On some aspects of reciprocal recurrent selection. Euphytica 10:24–30.

Schön, C.C., A.E. Melchinger, J. Boppenmaier, E. Brunklaus-Jung, R.G. Herrmann, and J.F. Seitzer. 1994. RFLP mapping in maize: Quantitative trait loci affecting testcross performance of elite European flint lines. Crop Sci. 34:378–389.

Schön, C.C., H.F. Utz, S. Groh, B. Truberg, S. Openshaw, and A.E. Melchinger. 2004. QTL mapping based on resampling in a vast maize testcross experiment confirms the infinitesimal model of quantitative genetics for complex traits. Genetics 167(1):485–498.

Seitz, G. 1989. Experimentelle und theoretische Untersuchungen zur Beziehung zwischen Linieneigenleistung und allgemeiner Kombinationsfähigkeit bei Silomais. Ph.D. diss. Univ. Hohenheim, Stuttgart, Germany.

Smith, O.S. 1986. Covariance between line per se and testcross performance. Crop Sci. 26:540–543.

Utz, H.F., and A.E. Melchinger. 1996. PLABQTL: A program for composite interval mapping of QTL. J. Quant. Trait Loci 2(1). Available at http://www.uni-hohenheim.de/~ipspwww/soft.html (verified 10 Sept. 2004). Univ. Hohenheim, Germany.

Utz, H.F., A.E. Melchinger, and C.C. Schön. 2000. Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. Genetics 154:1839–1849.

Published online October 27, 2005

No Evidence for Epistasis in Hybrid and Per Se Performance of Elite European Flint Maize Inbreds from Generation Means and QTL Analyses

Renata Mihaljevic, H. Friedrich Utz, and Albrecht E. Melchinger*

ABSTRACT

Favorable epistatic gene complexes may be important for hybrid performance of maize (*Zea mays* L.). This study was conducted to assess the importance of epistasis in per se and testcross performance for grain yield and grain moisture in four crosses among four elite European flint maize lines by generation means analyses as well as genome-wide tests for significant digenic epistatic effects between marker loci. For each cross, six generations (P1, P2, F_1, F_2, BC1, BC2) and testcrosses of these generations plus the F_2–Syn1, F_2–Syn2, and F_2–Syn3 generations in combination with an unrelated dent tester were evaluated in four environments. Testcross generation means of $\overline{P}$, $\overline{BC}$, F_1, F_2, F_2–Syn1, F_2–Syn2, and F_2–Syn3 did not significantly differ from each other for grain yield and grain moisture, indicating that epistasis between unlinked and moderately linked loci was negligible in its net effect. Depending on the cross, QTL mapping for per se and testcross performance with the dent tester was conducted with 71 to 344 lines (F_3 to F_6) grown in four environments. In genome-wide two-way ANOVAs, significant epistatic interactions were found with only a few marker pairs that did not improve the fit of the model after including main-effect QTLs previously detected by composite interval mapping. Poor correspondence of the results from per se and testcross analyses reflects dominance and epistatic interactions between parental and tester alleles. Our results suggest that epistasis is of minor importance for both traits with regard to the optimum type of population (F_2 vs. BC) in recycling breeding of elite maize inbreds. Estimates of digenic epistasis detected with genome-wide tests must be treated with caution because of the problems associated with model selection in QTL mapping with the sample sizes commonly used.

Epistasis is regarded as one possible cause of heterosis. Although increasing evidence for the existence of epistasis has been provided at the molecular level (Cheverud and Routman, 1995), its importance for heterosis and performance of elite maize hybrids has received surprisingly little attention. One reason for this might be the limited power of biometric methods of quantitative genetics, which test for the net effect of genes or gene combinations summed over all loci (Holland, 2001).

Traditional approaches to assess the importance of epistasis have relied on the analysis of first- and second-degree statistics by using either generation means analysis (Mather and Jinks, 1982) or estimation of variance components from covariances of relatives generated via special mating designs (Hallauer and Miranda, 1981). Nevertheless, the underlying reference populations were in most studies not representative of elite hybrids because crosses within heterotic groups were mainly employed.

To overcome this problem, Melchinger (1987) proposed the testcross generation means analysis. Hereby, the basic generations are not evaluated for their per se performance but for their performance in testcross to a tester from the opposite heterotic pool. This formally eliminates dominance effects from the model, which otherwise tend to override estimates of epistatic effects. Furthermore, by testing interpool hybrids, the results are of direct relevance for hybrid breeding.

First experimental results from a testcross generation means analysis were reported by Melchinger et al. (1988) on a cross of European dent lines. Epistasis was generally of minor importance but significant for grain and forage dry matter content as well as root lodging resistance. In U.S. dent germplasm, Lamkey et al. (1995) found significant epistatic effects for grain yield and grain moisture explaining 21 and 18% of the variation among testcross generation means, respectively. In a follow-up study with 40 hybrid combinations, only five crosses yielded significant additive × additive epistatic effects for grain yield (Hinze and Lamkey, 2003). Hitherto, no study is available on the importance of epistasis in elite lines of European flint maize germplasm.

With traditional generation means analysis, significant epistatic effects have been detected for important agronomic traits of maize (Hayman, 1958; Gamble 1962a, 1962b; Melchinger et al., 1986). Positive additive × additive and negative dominance × dominance epistatic effects were small compared with additive and dominance effect (Melchinger et al., 1986).

Both testcross generation means analysis and ordinary generation means analysis estimate only net effects of genes or gene combinations summed over loci. Thus, positive and negative epistatic effects among individual quantitative trait loci (QTL) may cancel each other. QTL analyses allow dissecting quantitative traits into the effects of individual factors. In most instances, they revealed little or no evidence for epistasis (Stuber et al., 1992; Xiao et al., 1995; Liu et al., 1996; Lu et al., 2004). However, when individual QTL were isolated in isogenic backgrounds, epistasis was commonly observed (Doebley et al., 1995; Long et al., 1995; Eshed and Zamir, 1996; Laurie et al., 1997).

With composite interval mapping, we rarely found significant digenic epistatic effects among the detected QTL for testcross and per se performance of lines derived from three crosses of European flint maize (Mihaljevic et al., 2004, 2005). However, with genome-wide

Institute of Plant Breeding, Seed Science, and Population Genetics, Univ. of Hohenheim, 70593 Stuttgart, Germany. Received 25 Dec. 2004. *Corresponding author (melchinger@uni-hohenheim.de).

Published in Crop Sci. 45:2605–2613 (2005).
Crop Breeding, Genetics & Cytology
doi:10.2135/cropsci2004.0760

677 S. Segoe Rd., Madison, WI 53711 USA

Abbreviations: ANOVA, analysis of variance; BC1, BC2, first backcrosses of generation F_1 to parents 1 and 2, respectively; BIC, Bayesian information criterion; P1, parent one; P2, parent two; QTL, quantitative trait locus/loci.

tests for epistasis, many important epistatic interactions were detected even among marker loci that did not show significant main effects (Damerval et al., 1994; Li et al., 1997; Holland et al., 1997).

The major goal of the present study was to assess the importance of epistasis for grain yield and grain moisture in four crosses of elite European flint maize with different approaches. Our objectives were to (i) estimate the relative importance of aggregate epistatic effects by generation means analyses of per se and testcross performance, (ii) perform genome-wide tests for significant epistatic effects between individual marker loci, and (iii) compare the results of each analysis and previous QTL analyses for both per se and testcross performance.

MATERIALS AND METHODS

Plant Materials

Four early-maturing elite European flint inbreds KW1265, D146, D145, and KW1292, subsequently referred to as A, B, C, and D, respectively, were used as parental lines in this experiment. Lines A and D are private inbreds developed by KWS SAAT AG; lines B and C are public inbreds proprietary to the University of Hohenheim. The generations P1 and P2 (parents), F_1, F_2, F_2–Syn1, F_2–Syn2, F_2–Syn3, and first backcrosses BC1 and BC2 of the F_1 to P1 and P2, respectively, were developed from each of the following four crosses: A×B, A×C, A×D, and C×D. The F_2–Syn1 to F_2–Syn3 generations were produced by paired plant crosses using a minimum of 250 pairs per generation starting in the F_2 generation. For each cross, testcross seed was produced by mating each generation to the unrelated dent tester inbred T2 (KW5361, serving as pollen parent) previously used for QTL mapping of testcross performance (Schön et al., 1994; Melchinger et al., 1998; Mihaljevic et al., 2004).

Field Experiments

Testcross Generation Means Analysis

Testcross progenies of generations P1, P2, F_1, F_2, F_2–Syn1, F_2–Syn2, F_2–Syn3, BC1, and BC2 were evaluated in a 5×10 α-design (Patterson and Williams, 1976) at four environments (Eckartsweier, Bad Krozingen, Zell, and Stuttgart-Hohenheim) in Germany with three replications. Testcrosses of P1 and P2 were included as duplicate entries.

Generation Means Analysis

The generations P1, P2, F_1, F_2, BC1, and BC2 derived from each of the four crosses were evaluated for per se performance in a split-plot design with generations comprising the main plots and crosses comprising the subplots. The trials were grown at four environments (Eckartsweier, Bad Krozingen, Zell, and Hochburg) in Germany with four replications.

For all experiments, plots consisted of two rows, 4.0 m long and 1.5 m wide with 0.7 m between rows. Two-row plots were overplanted and later thinned to reach a final stand of 90 000 plants ha^{-1}. All experiments were machine planted and harvested as grain trials with a combine. Data were analyzed for grain moisture (g kg^{-1}) at harvest and grain yield (Mg ha^{-1}) adjusted to 155 g kg^{-1} grain moisture.

Agronomic Data Analyses

Lattice and split-plot analyses of variance for testcross and per se data, respectively, were performed for each environment. Adjusted entry means and effective error mean squares from the lattice analyses as well as means and error mean squares from the split-plot analyses were then used to compute the combined analyses of variance across environments (Cochran and Cox, 1957). Generation means across environments were further used in the quantitative genetic analyses.

Testcross Generation Means Analysis

Two genetic models were fitted to the testcross generation means (Melchinger, 1987). Model 1^T accounts for additive effects only. Model 2^T allows for epistatic effects between unlinked pairs of loci but ignores linked epistatic pairs. The superscript T in the following models indicates that these values pertain to testcross effects.

$$\text{Model } 1^T\text{: } Y^T = m^T + x\,(\alpha^T)$$
$$\text{Model } 2^T\text{: } Y^T = m^T + x\,(\alpha^T) + x^2\,(\alpha\alpha^T),$$

where Y^T = testcross mean of the generation considered; m^T = testcross mean of the gene-orthogonal F_2 reference population in linkage equilibrium derived from the cross P1 × P2 (Schnell, 1965); x = coefficient that is generation-dependent and a linear function of the proportion of germplasm from the two parent lines ($x = -1, 1, 0, 0, 0, 0, -0.5, 0.5$ for generations P1, P2, F_1, F_2, F_2–Syn1, F_2–Syn2, F_2–Syn3, BC1, and BC2, respectively); (α^T) = additive effect summed over loci (equivalent to one-half the average effect of a gene substitution (α^T) at a single locus with a positive sign if P2 contains the favorable allele); $(\alpha\alpha^T)$ = additive × additive digenic epistatic effect summed over locus pairs.

Generation Means Analysis

Two genetic models were fitted to the per se performance data of the six generations. Model 1 includes only additive and dominance effects. Model 2 allows for epistatic effects between unlinked pairs of loci but ignores linked epistatic pairs. All effects were defined according to the F_2 metric (Hayman, 1958).

$$\text{Model 1: } Y = m + x\,(a) + z\,(d)$$
$$\text{Model 2: } Y = m + x\,(a) + z\,(d) + x^2\,(aa),$$

where Y = mean of the per se performance of the generation considered; m = mean of all inbred lines derived from the cross P1 × P2; (a) and (d) = summed additive and dominance effects, respectively (a single locus effect will have a positive sign if P2 harbors the favorable or dominant allele at the respective locus); (aa) = summed additive × additive digenic epistatic effects. The parameter notation follows Kearsey and Pooni (1996).

The formulas for the genotypic means of the various generations are

$$\text{P1: } Y = m - (a) - 0.5\,(d) + (aa),$$
$$\text{P2: } Y = m + (a) - 0.5\,(d) + (aa),$$
$$\text{F}_1\text{: } Y = m + 0.5\,(d),$$
$$\text{F}_2\text{: } Y = m,$$
$$\text{BC1 } (\text{F}_1 \times \text{P1})\text{: } Y = m - 0.5\,(a) + 0.25\,(aa),$$
$$\text{BC2 } (\text{F}_1 \times \text{P2})\text{: } Y = m + 0.5\,(a) + 0.25\,(aa).$$

Estimation of Effects and Model Fit

The genetic parameters for all four models were estimated using weighted least squares:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{W}\mathbf{y}),$$

where $\hat{\boldsymbol{\beta}}$ denotes the column vector of estimated genetic effects; $\mathbf{X}$ the matrix with elements that are a function of the generation; $\mathbf{W}$ the weight matrix with the inverse of the variances of the generation means on the diagonal and zero on the off-diagonal; and $\mathbf{y}$ the column vector Y or Y^T, respectively. Weighted estimates were calculated because the parental generations were tested as duplicate entries. Standard errors for the genetic parameters were estimated as the square root of the diagonal of the $(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}$ matrix. The coefficient of determination (R^2) was calculated to estimate the proportion of the variation among generation means accounted for by each model.

For both testcross and per se performance data, the goodness-of-fit of a model was tested with a weighted Chi-square (Mather and Jinks, 1982), $\chi^2 = \Sigma\,[(O - E)^2 \times W]$, where O = the observed generation mean, E = the expected generation mean, and W = the inverse of the variance of the generation mean.

QTL Experiments

QTL analyses for testcross and per se performance of the crosses A×B, A×C, and C×D were published previously (Schön et al., 1994; Melchinger et al., 1998; Mihaljevic et al., 2004, 2005). No QTL analysis was performed for the cross A×D because the population size was too small ($N = 42$) to obtain meaningful results. Briefly, four populations, $A\times B^{I}$ (344 $F_{2:3}$ lines for testcross and 280 $F_{2:3}$ lines for per se performance), $A\times B^{III}$ (71 $F_{4:5}$ for testcross and 120 $F_{4:6}$ for per se performance), A×C (109 $F_{3:4}$ lines for testcross and 131 $F_{3:4}$ lines for per se performance), and C×D (84 $F_{3:4}$ lines for testcross and 135 $F_{3:4}$ lines for per se performance) were employed in QTL analyses. Here, $A\times B^{I}$ and $A\times B^{III}$ represent different samples of the same cross, the notation being in accordance with Mihaljevic et al. (2004, 2005). All these populations were reanalyzed here with a genome-wide test for epistatic effects to detect interactions among QTL which do not necessarily have a significant main effect. The number of markers employed ranged from 73 to 95 depending on the population. Only those markers used for constructing the joint map across populations described by Mihaljevic et al. (2004, 2005) were employed herein for further analyses. The average marker density on the joint map ranged from 10.2 cM in C×D to 15.0 cM in $A\times B^{III}$.

Digenic epistatic effects, (*aa*) for per se performance and ($\alpha\alpha^T$) for testcross performance, between all pairs of marker loci were tested by EPISTACY, a two-way ANOVA routine in SAS based on the F_2 metric (Holland, 1998). Epistatic interactions were declared significant if they exceeded the threshold of $P < 0.001$. This threshold was determined because 45 independent combinations exist among the ten linkage groups of maize. A comparison-wise error rate of 10^{-3} would correspond approximately to an experiment-wise error rate of 0.05. This seems a liberal estimate of the genome-wise error rate for epistatic interactions (Holland et al., 1997).

The Bayesian information criterion (BIC; Piepho and Gauch, 2001) implemented in software PLABQTL (Utz and Melchinger, 1996) was used to compare the model including only positions of main-effect QTL estimated by standard composite interval mapping with an extended model, which included the position of the main-effect QTL plus those marker pairs with significant epistatic effects detected by EPISTACY.

RESULTS

Testcross Generation Means

Testcross means of parents P1 and P2 differed significantly ($P < 0.05$) for both grain yield and grain moisture in all crosses except C×D, where both parents had similar means for both traits (Table 1). No significant ($P < 0.05$) differences existed between the parental mean $\overline{P}$, backcross mean $\overline{BC}$, and F_1 and F_2 generations in any cross for both traits. Likewise, no significant changes were observed between testcrosses of generations F_1, F_2, F_2–Syn1, F_2–Syn2, and F_2–Syn3 for all crosses and both traits. Model 1^T explained over 78% of the variation among generation means for grain yield in all crosses except C×D (Table 2). The χ^2 goodness-of-fit test for Model 1^T was not significant in any of the four crosses. Inclusion of epistatic effects in Model 2^T resulted in a substantial increase of R^2 values for A×B and A×C, with estimates of ($\alpha\alpha^T$) being significant. For grain moisture, the χ^2 goodness-of-fit test for Model 1^T was significant ($P < 0.05$) in A×D. R^2 values of Model 1^T varied between 57.5 and 73.0% for grain moisture and increased substantially for Model 2^T in A×B, A×C, and A×D. In all three crosses, estimates of additive

Table 1. Means and their standard errors of testcross progenies with dent tester T2 of nine generations from four crosses of European flint maize lines evaluated in four environments for grain yield and grain moisture.

Generation	Cross A×B	Cross A×C	Cross A×D	Cross C×D	SE†	Cross A×B	Cross A×C	Cross A×D	Cross C×D	SE†
	Grain yield (Mg ha^{-1})					Grain moisture (g kg^{-1})				
T × P1‡	7.35	7.35	7.35	9.15	0.21	369.8	369.8	369.8	383.1	2.1
T × P2‡	8.60	9.15	8.88	8.88	0.21	381.8	383.1	382.6	382.6	2.1
T × $\overline{P}$	7.98	8.25	8.12	9.02	0.15	375.8	376.4	376.2	382.8	1.5
T × F_1	8.17	8.63	8.40	9.26	0.25	370.5	377.5	371.8	383.7	2.4
T × F_2	8.32	8.81	7.95	9.12	0.25	375.1	377.5	376.4	384.7	2.4
T × F_2–Syn1	8.23	8.57	7.82	8.73	0.25	373.5	382.1	370.6	386.1	2.4
T × F_2–Syn2	8.32	8.65	8.44	9.18	0.25	370.8	380.6	369.6	383.2	2.4
T × F_2–Syn3	8.39	8.71	8.53	9.46	0.25	369.6	376.1	367.7	377.3	2.4
T × BC1	7.76	7.83	7.83	9.04	0.25	369.3	378.2	368.6	384.1	2.4
T × BC2	8.44	8.87	8.32	8.85	0.25	373.2	382.3	374.0	381.8	2.4
T × $\overline{BC}$	8.10	8.35	8.08	8.94	0.18	371.2	380.2	371.3	382.9	1.7

† Standard error for respective generation mean.
‡ Testcrosses of P1 and P2 were included as duplicate entries in each replication.

Table 2. Genetic effects and their standard errors estimated from testcross progeny means of four crosses (A×B, A×C, A×D, C×D) for grain yield and grain moisture. Regression estimates and their standard errors were determined by fitting Model 1[T] and Model 2[T] to testcross generation means across four environments.

Generation	Cross A×B	Cross A×C	Cross A×D	Cross C×D	Cross A×B	Cross A×C	Cross A×D	Cross C×D
	Grain yield (Mg ha^{-1})				Grain moisture (g kg^{-1})			
	Model 1[T]† (Fit for additive effects)							
m^T	8.16 ± 0.06§**	8.48 ± 0.08**	8.16 ± 0.08**	9.07 ± 0.07**	372.8 ± 0.91**	378.4 ± 0.82**	372.6 ± 1.15**	382.9 ± 0.84**
(α^T)	0.63 ± 0.10**	0.92 ± 0.13**	0.72 ± 0.14**	−0.14 ± 0.12	5.66 ± 1.60**	6.24 ± 1.44**	6.24 ± 2.03*	−0.58 ± 1.47
$\chi^2_{(7)}$‡	3.46	6.68	7.67	5.80	9.63	7.80	15.58*	8.22
R^2 (%)	86.1	87.2	78.5	15.9	64.3	73.0	57.5	2.2
	Model 2[T]† (Fit for additive and additive × additive effects)							
m^T	8.26 ± 0.04**	8.62 ± 0.06**	8.20 ± 0.11**	9.11 ± 0.10**	371.5 ± 0.86**	379.2 ± 0.97**	370.9 ± 1.11**	383.0 ± 1.14**
(α^T)	0.63 ± 0.05**	0.92 ± 0.08**	0.72 ± 0.15**	−0.14 ± 0.13	5.66 ± 1.21**	6.24 ± 1.35**	6.24 ± 1.55**	−0.58 ± 1.59
($\alpha\alpha^T$)	−0.30 ± 0.07**	−0.40 ± 0.11**	−0.10 ± 0.21	−0.12 ± 0.18	4.04 ± 1.62*	−2.49 ± 1.82	5.08 ± 2.09*	−0.15 ± 2.13
$\chi^2_{(6)}$‡	0.83	1.94	7.38	5.41	4.73	5.95	7.84	8.21
R^2 (%)	96.7	96.3	79.3	21.5	82.4	79.4	78.6	2.2

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
† For definition of genetic effects, see Materials and Methods.
‡ Chi-square degrees of freedom in parentheses.
§ Standard error is attached.

effects (α^T) were highly significant ($P < 0.01$) for both traits. Estimates of epistatic effects were negative in all crosses for grain yield. For grain moisture, estimates of ($\alpha\alpha^T$) were significant in two crosses and of positive sign. C×D deviated from the other three crosses in that R^2 values were low (≤21.5%) for both models and estimates of (α^T) and ($\alpha\alpha^T$) were nonsignificant for both traits.

Generation Means

Means of parents P1 and P2 differed significantly ($P < 0.05$) for grain yield in all crosses but not for grain moisture (Table 3). For all crosses, the F_1 generation outyielded ($P < 0.05$) the F_2 and $\overline{BC}$; the F_2 generation means were significantly smaller than the $\overline{BC}$ means in A×D and C×D for grain yield. For grain moisture, no significant differences existed among these generations in three of the four crosses (Table 3).

For grain yield, R^2 values for Model 1 exceeded 94% for all crosses, despite significant χ^2 values for crosses A×D and C×D (Table 4). Estimates of epistatic effects (*aa*) were significant only for C×D. The χ^2 goodness-of-fit test of Model 2 was nonsignificant for all crosses except A×D. For grain moisture, R^2 values for Model 1 were lower and ranged between 63.1 and 90.5%. Inclusion of epistatic effects in Model 2 improved the fit, but estimates of epistatic effects (*aa*) were not significant for either cross.

Additive effects were smaller than dominance effects for grain yield in all crosses, but for grain moisture only in A×B and A×D. Both types of effects were highly significant ($P < 0.01$) in most instances for grain yield, but only in two instances for grain moisture. Dominance effects were consistently negative for grain moisture.

Digenic Epistatic Interactions

Testcross Performance [($\alpha\alpha^T$) Type of Epistasis]

The number of marker pairs with significant ($P < 0.001$) epistatic interactions for grain yield was two for $A\times B^I$, zero for $A\times B^{III}$, and one for A×C and C×D (Table 5). The absolute size of the ($\alpha\alpha^T$) effects for grain yield ranged from 0.21 to 0.31 Mg ha^{-1} across populations. The sum of absolute values of the two ($\alpha\alpha^T$) effects in $A\times B^I$ of opposite sign was about half the sum of absolute additive QTL effects estimated in the same population. In A×C and C×D, the absolute

Table 3. Means and their standard errors of six generations from four crosses (A×B, A×C, A×D, C×D) of European flint maize lines evaluated in four environments for grain yield and grain moisture.

Generation	Cross A×B	Cross A×C	Cross A×D	Cross C×D	SE†	Cross A×B	Cross A×C	Cross A×D	Cross C×D	SE†
	Grain yield (Mg ha^{-1})					Grain moisture (g kg^{-1})				
P1§	2.91	2.91	2.91	4.99	0.33	351.6	351.6	351.6	365.5	7.1
P2§	4.65	4.99	3.82	3.82	0.33	352.8	365.5	353.1	353.1	7.1
$\overline{P}$	3.78	3.95	3.37	4.40	0.23	352.2	358.6	352.4	359.3	5.0
F_1	7.80	9.01	7.68	10.72	0.17	332.7	356.2	306.1	353.9	3.2
F_2	6.05	6.63	5.07	6.93	0.17	335.9	361.9	318.3	354.7	3.2
BC1	5.68	6.19	5.44	7.41	0.17	328.4	348.4	318.8	359.6	3.2
BC2	6.76	7.36	6.21	7.07	0.17	342.6	363.7	321.6	349.3	3.2
$\overline{BC}$	6.22	6.78	5.83	7.24	0.12	335.5	356.0	320.2	354.4	2.2
Heterosis (%)‡	106.3	128.1	127.9	143.2	–	−5.5	−0.7	−13.1	−1.5	–

† Standard error for respective generation mean.
‡ Heterosis is measured as 100 (F_1 − P)/P.
§ A given line was evaluated once as duplicate entry for different crosses.

Table 4. Genetic effects and their standard errors estimated from generation means of four crosses (A×B, A×C, A×D, C×D) for grain yield and grain moisture. Regression estimates and their standard errors were determined by fitting Model 1 and Model 2 to generation means across four environments.

Generation	Cross A×B	Cross A×C	Cross A×D	Cross C×D	Cross A×B	Cross A×C	Cross A×D	Cross C×D
	Grain yield (Mg ha^{-1})				Grain moisture (g kg^{-1})			
	Model 1† (Fit for additive and dominance effects)							
m	6.04 ± 0.11§ **	6.65 ± 0.08**	5.55 ± 0.18**	7.27 ± 0.14**	337.5 ± 2.54**	357.8 ± 1.76**	322.3 ± 2.55**	355.1 ± 0.77**
(*a*)	0.97 ± 0.23*	1.10 ± 0.16**	0.61 ± 0.36	−0.46 ± 0.28	8.10 ± 5.44	11.56 ± 3.78	1.88 ± 5.46	−8.46 ± 1.65*
(*d*)	3.85 ± 0.39**	4.96 ± 0.26**	4.28 ± 0.60**	6.50 ± 0.47**	−15.34 ± 8.99	−2.73 ± 6.24	−40.36 ± 9.03*	−4.14 ± 2.72
χ^2_D‡	5.76	2.68	14.01**	8.61*	7.84*	3.79	7.93*	0.72
R^2 (%)	97.5	99.3	94.7	98.5	63.1	76.1	87.0	90.5
	Model 2† (Fit for additive, dominance, and additive × additive effects)							
m	6.21 ± 0.20**	6.75 ± 0.14**	5.35 ± 0.34**	6.93 ± 0.07**	333.9 ± 4.53**	360.0 ± 3.28**	316.5 ± 2.28**	354.1 ± 1.40**
(*a*)	0.97 ± 0.23*	1.10 ± 0.17*	0.61 ± 0.40	−0.46 ± 0.08*	8.10 ± 5.53	11.56 ± 4.00	1.88 ± 2.79	−8.46 ± 1.70*
(*d*)	3.32 ± 0.65*	4.64 ± 0.47**	4.92 ± 1.12*	7.58 ± 0.24**	−4.50 ± 14.7	−9.54 ± 10.6	−22.6 ± 7.38	−0.98 ± 4.51
(*aa*)	−0.62 ± 0.61	−0.37 ± 0.44	0.74 ± 1.05	1.26 ± 0.22*	13.6 ± 14.4	−8.55 ± 10.4	22.4 ± 7.24	3.97 ± 4.42
χ^2_D‡	3.80	1.98	11.24**	0.51	5.41	2.83	1.37	0.51
R^2 (%)	98.4	99.5	95.7	99.9	74.6	82.2	97.8	93.3

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
† For definition of genetic effects, see Materials and Methods.
‡ Chi-square degrees of freedom in parentheses.
§ Standard error is attached.

($\alpha\alpha^T$) effect was by far less than half of the sum of the absolute additive QTL effects.

For grain moisture, no significant ($\alpha\alpha^T$) effects were detected in the largest population A×B^{I}. Three epistatic marker pairs were detected in A×C and C×D, respectively. All three had a positive sign in A×C, but one showed a negative sign in C×D. The absolute size of ($\alpha\alpha^T$) effects ranged from 1.2 to 4.4 g kg^{-1} across populations. For grain moisture, the sum of absolute ($\alpha\alpha^T$) effects was about one-third of the sum of absolute additive QTL effects in A×C and about one-fourth in C×D but comparatively small in A×B^{III} (Table 5). Of the 11 epistatic marker pairs detected across populations and traits, no marker was flanking a QTL with main effects.

The effects estimated with PLABQTL by the model including only epistatic marker pairs from EPISTACY were reduced in size only, apart from two changes in sign when the model included both the epistatic marker pairs and the main-effects QTL previously detected by composite interval mapping (Table 5). According to the BIC, the model extended for epistatic marker pairs was not superior to the basic model, including only main-effect QTL in each population except for grain moisture in C×D.

Per Se Performance [(*aa*) Type of Epistasis]

Between one and three marker pairs per population showed significant ($P < 0.001$) epistatic interactions for grain yield (Table 6). The absolute size of the (*aa*) effects

Table 5. Marker pairs showing significant additive × additive epistasis for testcross performance of grain yield and grain moisture in populations A×B^{I}, A×B^{III}, A×C, and C×D.

Cross	Detected in EPISTACY: Marker 1†	Marker 2†	P	aa^T effect estimated in PLABQTL: Pair only	All pairs and QTL‡	Sum of absolute a^T–effects§
			Grain yield (Mg ha^{-1})			
A×B^{I}	BNL3.04(10)	UMC132(6)	0.000 601	−0.21	−0.18	–
	UMC44(10)	UMC53(2)	0.000 171	0.24	0.23	–
Sum¶	–	–	–	0.45	–	0.92
A×B^{III}	–	–	–	–	–	–
A×C	BNL8.15(5)	UMC159(6)	0.000 165	−0.31	−0.19	2.68
C×D	BNL6.22(5)	UMC64(10)	0.000 522	0.28	0.10	2.42
			Grain moisture (g kg^{-1})			
A×B^{I}	–	–	–	–	–	–
A×B^{III}	BNL10.17(4)	UMC36(2)	0.000 774	1.20	0.89	22.3
A×C	BNL5.71(5)	UMC60(3)	0.000 784	3.23	−0.36	–
	BNL9.11(8)	UMC127(4)	0.000 072	3.70	2.25	–
	UMC1(5)	UMC60(3)	0.000 607	3.64	2.35	–
Sum¶	–	–	–	10.6	–	34.2
C×D	BNL10.13(10)	UMC35(7)	0.000 231	4.44	−7.21	–
	BNL7.71(5)	UMC138(6)	0.000 220	4.40	2.17	–
	UMC28(6)	UMC6(2)	0.000 101	−4.13	−5.62	–
Sum¶	–	–	–	13.0	–	51.6

† The number in parentheses indicates the linkage group of the marker.
‡ For details, see Materials and Methods.
§ a^T effects from QTL analyses of testcross performance (Mihaljevic et al., 2004).
¶ Sum of absolute effects.

Table 6. Marker pairs showing significant additive × additive epistasis for per se performance of grain yield and grain moisture in populations A×B[I], A×B[III], A×C, and C×D.

Cross	Detected in EPISTACY			*aa* effect estimated in PLABQTL		
	Marker 1†	Marker 2†	P	Pair only	All pairs and QTL‡	Sum of absolute *a* effects§
			Grain yield (Mg ha^{-1})			
A×B[I]	BNL10.17(4)	BNL14.28(9)	0.000 407	−0.35	−0.33	2.74
A×B[III]	BNL15.18(1)	BNL9.11(8)	0.000 006	−0.52	−0.43	–
	UMC109(9)	UMC96(3)	0.000 725	−0.34	−0.26	–
Sum¶	–	–	–	0.86	–	0.81
A×C	BNL10.24(3)	UMC138(6)	0.000 630	−0.29	−0.28	–
	UMC12(8)	UMC15(4)#	0.000 686	0.34	0.28	–
Sum¶	–	–	–	0.63	–	0.95
C×D	BNL15.21(7)	BNL6.06(3)	0.000 632	−0.47	−0.53	–
	UMC103(8)	UMC166(5)	0.000 002	0.69	0.48	–
	UMC127(4)	UMC153 (9)	0.000 660	0.48	0.65	–
Sum¶	–	–	–	1.64	–	1.74
			Grain moisture (g kg^{-1})			
A×B[I]	UMC37(1)	UMC3(3)	0.000 889	−5.24	−4.66	28.9
A×B[III]	BNL3.06(9)	BNL9.44(8)	0.000 825	6.93	8.27	–
	UMC140(9)	UMC159(6)	0.000 972	−7.04	−6.84	–
Sum¶	–	–	–	14.0	–	10.2
A×C	UMC120(8)	UMC130(10)	0.000 486	5.76	5.63	–
	UMC155(10)	UMC51(5)	0.000 811	−7.17	−2.23	–
Sum¶	–	–	–	12.9	–	40.2
C×D	BNL8.39(7)	UMC37(1)	0.000 965	−6.41	−3.35	53.1

† The number in parentheses indicates the chromosome location of the marker.
‡ For details, see Materials and Methods.
§ *a* effects from QTL analyses of per se performance (Mihaljevic et al., 2005).
¶ Sum of absolute effects.
# Underlined markers are flanking the detected main-effect QTL.

for this trait ranged from 0.29 to 0.69 Mg ha^{-1} across populations. The sum of absolute (*aa)* effects was comparable with the sum of absolute additive QTL effects in A×B[III], A×C, and C×D (Table 6). The (*aa)* effects were negative in cross A×B, but of opposite sign in A×C and C×D.

For grain moisture, two marker pairs with opposite sign of (*aa)* effects were detected in A×B[III] and A×C, and one marker pair with negative (*aa)* effect was detected in A×B[I] and C×D (Table 6). The absolute size of (*aa)* effect ranged from 5.24 to 7.17 g kg^{-1} across populations. Only in A×B[III] was the sum of absolute (*aa)* effects comparable to the sum of absolute additive QTL effects. In the other populations, the sum of absolute additive QTL effects was a multiple of the sum of absolute (*aa)* effects. Of all 14 epistatic marker pairs detected across all populations and traits, only one marker for grain yield and two markers for grain moisture were flanking QTL with main effects.

The effect size estimated with PLABQTL by the model including only the marker pairs detected with EPISTACY was mostly larger compared with the model that included these marker pairs plus the positions of main-effect QTL detected previously by composite interval mapping (Table 6). According to the BIC, the latter model was consistently not superior to the basic model including only main-effect QTL.

DISCUSSION

Favorable epistatic gene action between tightly linked genetic loci has been suggested as a major cause of grain yield heterosis and hybrid vigor in maize (Cockerham and Zeng, 1996). The lack of success in recycling breeding with certain elite lines provides further indirect evidence for the presence of epistasis. With this breeding approach, tightly linked positive epistatic combinations of genes can be accumulated by selection over several generations. Conversely, if lines are extracted from populations undergoing recurrent selection, epistasis between linked loci is expected to be of lower importance because recurrent intermating promotes disruption of linked genes.

In the testcross generation mean analysis, epistasis between unlinked loci can alter only the means of generations prior to the F_2 (i.e., $\overline{P}$, $\overline{BC}$, and F_1) because the gametic array produced by the F_1 (or any generation derived from it by random mating) is expected to be in linkage equilibrium. The contribution of positive epistasis between linked loci should therefore decline monotonically in the order $\overline{P} > \overline{BC} > F_2 > F_2$-Syn1 $> F_2$-Syn2 $> F_2$-Syn3 as a function of the recombination frequency (Melchinger, 1987). Since the parental lines of this study were developed by recycling breeding of elite lines, we expected to find epistasis in generation means analyses for both per se and testcross performance.

Epistasis in Testcross and Per Se Generation Means

In the testcross generation means analysis, contrasts $\overline{P}$ vs. F_1, $\overline{P}$ vs. $\overline{BC}$, or $\overline{BC}$ vs. F_1 were not significant for grain yield and grain moisture and, thus, provided no evidence for epistasis among unlinked loci (Table 2). Likewise, generations F_2 to F_2-Syn3 were not significantly different in their testcross means. In contrast, four of the eight estimated additive × additive epistatic effects ($\alpha\alpha^T$) were significant. Following the common

procedure in the literature (e.g., Hinze and Lamkey, 2003), the latter were tested against the deviation means squares, which are often smaller than the estimated error variance of generation means corresponding to SE in Table 1. With the latter error term in this cases, no significant ($\alpha\alpha^T$) effects were detected. Given the low number of degrees of freedom for the residual error or test environments in the generation means analyses, the power of these tests is relatively poor. Therefore, the relative importance of epistasis for testcross performance will be briefly assessed by considering the ratio $AA^T\% = (\alpha\alpha^T)/m^T \times 100$.

Averaged across all four crosses in our study, $AA^T\%$ amounted to −2.7% for grain yield and 0.4% in grain moisture. Thus, for grain yield there was no indication for positive epistasis. By comparison, Lamkey et al. (1995) observed a high reduction in testcross performance for grain yield after eight generations of random mating in cross B73×B84, corresponding to $AA^T\%$ of 6.3% for grain yield and 1.8% for grain moisture. In a more extensive study with 40 crosses of current U.S. elite lines, Hinze and Lamkey (2003) found epistasis to be unimportant for grain yield with an average estimate of $AA^T\%$ of −0.8% and a range between −5.2% and 5.2%, depending on the cross. Thus, the net effect of epistasis on testcross generation means seems to be generally of minor importance, but higher values for individual crosses and environments cannot be ruled out.

A clear distinction of the contribution of unlinked vs. linked locus pairs or epistasis of higher order than ($\alpha\alpha^T$) is complicated by the fact that (i) the effects of epistasis cannot be completely separated from those of linkage (Melchinger, 1987), (ii) epistatic effects are partly contributing to additive effects and higher-order epistatic effects are contributing to estimates of lower order effects (Cheverud and Routman, 1995), (iii) ($\alpha\alpha^T$) effects are confounded with additive × dominance and dominance × dominance interactions between parental and tester alleles (Eta-Ndu and Openshaw, 1999), and/or (iv) maternal effects are confounded with epistatic effects. In the present study, maternal effects cannot be ruled out, because the testcross seed was produced in an isolation plot with the dent tester line as pollen parent and the various generations from the four crosses as seed parents. In reciprocal crosses of three-way hybrids, Schnell and Singh (1978) reported an average yield advantage of 3.1% for hybrids produced on a vigorous F_1 seed parent as compared to those produced on an inbred line seed parent, which have poorer early vigor owing to their smaller seed weight. Obviously, this type of maternal effect would be present in the comparison of $\overline{P}$ with other generations.

In cross C×D, it was striking that the estimate of (α^T) was fairly small (Table 2) even though the two parents showed pronounced differences in their per se performance (Table 3). Thus, the weak line D expressed strong dominance with the tester. The influence of dominance with the tester on estimates of (α^T) were discussed in detail by Melchinger et al. (1998). Hence, it seems plausible that in crosses A×D and C×D the correspondence of estimates of (α^T) with (*a*) and ($\alpha\alpha^T$) with (*aa*) was relatively poor. The agreement between both types of estimates was much better in crosses A×B and A×C. In all instances, the large dominance effect for per se performance reflected the substantial heterosis for grain yield in maize even in crosses within heterotic groups (Table 3, last line).

In conclusion, our study confirms the limitations of generation means analysis for an assessment of the importance of epistasis for quantitative traits. Recognizing these difficulties, marker-based analyses of epistatic effects have been suggested to be more powerful (Damerval et al., 1994; Li et al., 1997; Holland et al., 1997).

Mapping of Epistatic QTL

We found several marker pairs showing significant two-locus epistasis in addition to main-effects QTL. In general, these marker pairs were not flanking main-effect QTL. The sum of the absolute epistatic effects was often half or more of the sum of the absolute additive QTL effects. Thus, at first glance epistasis seems to be important in the analysis of QTL, which is in agreement with experimental results from other plants (Li et al., 1997; Holland et al., 1997; Kearsey et al., 2003). However, when the position of main-effect QTL previously identified in each cross by Mihaljevic et al. (2004, 2005) were included in the model, epistatic effects did not improve the model fit measured by the Bayesian information criterion. As for the generation means analyses, we therefore discuss the limitations of estimation of two-locus epistasis.

The first problem is that the true number and position of QTL, which correspond to the correct statistical model for estimating the gene effects, are unknown and must be determined by model selection (Zeng et al., 1999). The general procedure is to identify among a large number of regressor variables (markers) those that account for the largest proportion in the variance of the response variable (phenotypic values). Subsequently, these genome positions are used for estimation of QTL effects and the proportion *p* of the genotypic variance explained by the detected QTL. With a limited sample size, model selection leads to an overestimation of QTL effects and *p* because of sampling effects and, consequently, to a biased assessment of the prospects of marker-assisted selection (Melchinger et al., 1998; Utz et al., 2000). A genome-wide search for epistatic effects among QTL aggravates the problems associated with model selection, because the number of regressor variables (marker pairs) increases tremendously (for two-locus epistasis in quadratic progression, for three-locus epistasis in cubic progression, etc.). Furthermore, collinearity of dummy marker variables in the selected model disturbs the estimation of additive and epistatic QTL effects, especially with dense marker maps. Moreover, epistatic pairs of QTL are fit directly at marker locus positions rather than in intervals, which may reduce the power of QTL epistasis tests compared to the additive effect tests. Determining the appropriate experiment-wise error rate is therefore of crucial importance (Holland, 1998; Holland et al., 1997).

Similar to mapping of a QTL, in which the effect of

other QTL should be taken into account for example by including cofactors in the model, the same principle applies to the search for epistatically interacting pairs of QTL (Zeng et al., 1999). In our study, the size of epistatic effects for line per se and testcross performance was reduced, when all previously detected main-effect QTL were added to the model (Tables 5 and 6). Bogdan et al. (2004) reported similar results from simulations. They recommended a larger penalty in the BIC for epistatic terms than for main effects. Even with the ordinary BIC, no epistatic terms remained in the model in our study.

The need for validation with an independent sample or cross validation (Utz et al., 2000) is even more compelling for epistatic than for main effects of QTL. An ultimate proof for the presence of an epistatic pair of QTL and an unbiased estimation of its gene effects requires the isolation of the pair of QTL in a homogeneous background by means of near isogenic lines (NILs) or similar approaches (Doebley et al., 1995). Moreover, we strongly recommend using larger populations at least of the sample size of our biggest experiment ($N = 344$) for detection and mapping of epistatic QTL for complex traits such as grain yield and grain moisture. The presence of minor biological epistasis, however, cannot be ruled out at least for the cross $A \times B^1$, where evidence for weak epistasis was detected with a relatively large number of progenies.

Conclusions and Consequences for Breeding

In agreement with the findings of Hinze and Lamkey (2003) for U.S. dent germplasm, our results indicate that epistasis hardly influences the testcross means of F_2 or BC populations produced from elite European flint lines. Consequently, epistasis can be ignored with regard to the choice of the type of base population to be preferably used in recycling breeding (Melchinger et al., 1988). Moreover, we conclude that epistasis generally does not benefit single crosses over other types of hybrids and can safely be ignored in predicting the performance of three-way or double-cross hybrids from the means of their nonparental single crosses (Melchinger et al., 1987).

Our QTL analyses demonstrate that for complex traits such as grain yield, it is extremely difficult to map epistatic QTL with high fidelity and separate their effects from those of main-effects QTL. This is due to the problems associated with model selection, even when relatively large sample sizes are used. A promising way out of this deadlock could be novel approaches such as "genetical genomics" (Jansen and Nap, 2001), in which genome-wide expression data obtained from genomics, proteomics, and metabolomics are analyzed in parallel with phenotypic and marker data to unravel the basis of metabolic, regulatory and developmental pathways. Altogether, the novel approaches currently invented and used in systems biology to study the function and control of genetic networks promise to contribute to a better understanding of epistasis at the molecular level, and in turn also at the level of the entire genotype (Carlborg and Haley, 2004).

It is anticipated that the relative importance of epistatic effects in hybrid maize breeding may strongly increase with the current paradigm shift in line development from recurrent selfing to the production of doubled haploids (Seitz, 2004, pers. communication). This is because the variance of epistatic effects of order m among unlinked loci contributes to the genetic variance among S_n lines ($n \geq 1$) only with a coefficient $(1 - 0.5^n)^m$ (Cockerham, 1963). Hence, with early generation testing in traditional line development, digenic epistasis and even more so higher-order epistasis contribute only marginally to the genetic variance among S_1 or S_2 lines compared with additive effects. In contrast, with doubled haploid lines (corresponding to S_∞ lines), the coefficients of all epistatic variance components are equal to one and, hence, epistasis contributes fully to the genetic variance from the very beginning of the selection process. Moreover, because recombination is limited to a single meiosis for each breeding cycle, doubled haploids minimize recombination between linked loci and, thus, should be very effective in conserving tightly linked complexes of genes with positive epistasis. The drawback of restricted recombination is, however, the low chances to identify positive complexes of genes if these occur in repulsion phase in the parents. This requires either extremely large population sizes or several generations of intermating before producing the doubled haploid lines. It will therefore be of interest to investigate the importance of epistasis after several cycles of recycling breeding with doubled haploid lines have been completed.

ACKNOWLEDGMENTS

The present study was supported by a grant from the Deutsche Forschungsgemeinschaft, Grant No. ME 931/3-2. The RFLP assays were conducted in the lab of Prof. Dr. R.G. Herrmann, Ludwig-Maximilians-Universität in Munich, by E. Brunklaus-Jung and J. Boppenmaier as well as A. Dally in the lab of Prof. Dr. P. Westhoff at the Heinrich-Heine-Universität in Düsseldorf as part of EUREKA project 290. The skilled technical assistance of F. Mauch and the staff at the Plant Breeding Research Station in Eckartsweier in conducting field trials is gratefully acknowledged.

REFERENCES

Bogdan, M., J.K. Ghosh, and R.W. Doerge. 2004. Modifying the Schwarz Bayesian information criterion to locate multiple interacting quantitative trait loci. Genetics 167:989–999.

Carlborg, Ö., and C.S. Haley. 2004. Epistasis: Too often neglected in complex trait studies? Nat. Rev. Genet. 5:618–624.

Cheverud, J.M., and E.J. Routman. 1995. Epistasis and its contribution to genetic variance components. Genetics 139:1455–1461.

Cochran, W.G., and G.M. Cox. 1957. Experimental designs. 2nd ed., John Wiley & Sons, New York.

Cockerham, C.C. 1963. Estimation of genetic variances. p. 53–94. *In* Statistical genetics and plant breeding. NAS-NRC 982, Washington, DC.

Cockerham, C.C., and Z.-B. Zeng. 1996. Design III with marker loci. Genetics 143:1437–1456.

Damerval, C., A. Maurice, J.M. Josse, and D. de Vienne. 1994. Quantitative trait loci underlying gene product variation: A novel perspective for analyzing regulation of gene expression. Genetics 137: 289–301.

Doebley, J., A. Stec, and C. Gustus. 1995. *Teosinte branched 1* and

the origin of maize: Evidence for epistasis and the evolution of dominance. Genetics 141:333–346.
Eshed, Y., and D. Zamir. 1996. Less-than-additive epistatic interactions of quantitative trait loci in tomato. Genetics 143:1807–1817.
Eta-Ndu, J.T., and S.J. Openshaw. 1999. Epistasis for grain yield in two F_2 populations of maize. Crop Sci. 39:346–352.
Gamble, E.E. 1962a. Gene effects in corn (*Zea mays* L.). Separation and relative importance of gene effects for yield. Can. J. Plant Sci. 42:339–348.
Gamble, E.E. 1962b. Gene effects in corn (*Zea mays* L.). Relative importance of gene effects for plant height and certain component attributes of yield. Can. J. Plant Sci. 42:349–358.
Hallauer, A.R., and J.B. Miranda. Fo. 1981. Quantitative genetics in maize breeding. Iowa State Univ. Press, Ames, IA.
Hayman, B.I. 1958. The separation of epistatic from additive and dominance variation in generation means. Heredity 12:371–390.
Hinze, L.L., and K.R. Lamkey. 2003. Absence of epistasis for grain yield in elite maize hybrids. Crop Sci. 43:46–56.
Holland, J.B. 1998. EPISTACY: A SAS program for detecting two-locus epistatic interactions using genetic marker information. J. Hered. 89:374–375.
Holland, J.B. 2001. Epistasis and plant breeding. Plant Breed. Rev. 21:27–92.
Holland, J.B., H.S. Moser, L.S. O'Donoughue, and M. Lee. 1997. QTLs and epistasis associated with vernalization responses in oat. Crop Sci. 37:1306–1316.
Jansen, R.C., and J.-P. Nap. 2001. Genetical genomics: The added value from segregation. Trends Genet. 17:388–391.
Kearsey, M.J., and H.S. Pooni. 1996. The genetical analysis of quantitative traits. Chapman & Hall, London.
Kearsey, M.J., H.S. Pooni, and N.H. Syed. 2003. Genetics of quantitative traits in *Arabidopsis thaliana*. Heredity 91:456–464.
Lamkey, K.R., B.J. Schnicker, and A.E. Melchinger. 1995. Epistasis in an elite maize hybrid and choice of generation for inbred line development. Crop Sci. 35:1272–1281.
Laurie, C.C., J.R. True, J. Liu, and J.M. Mercer. 1997. An introgression analysis of quantitative trait loci that contribute to the morphological difference between *Drosophila simulans* and *D. mauritiana*. Genetics 145:339–348.
Li, Z., S.R.M. Pinson, W.D. Park, A.H. Paterson, and J.W. Stansel. 1997. Epistasis for three grain yield components in rice (*Oryza sativa* L.). Genetics 145:453–465.
Liu, J., J.M. Mercer, L.F. Stam, G.C. Gibson, Z. Zeng, and C.C. Laurie. 1996. Genetic analysis of a morphological shape difference in the male genitalia of *Drosophila simulans* and *D. mauritiana*. Genetics 142:1129–1145.
Long, A.D., S.L. Mullaney, L.A. Reid, J.D. Fry, C.H. Langley, and T.F.C. Mackay. 1995. High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. Genetics 139:1273–1291.
Lu, H., J. Romero-Severson, and R. Bernardo. 2004. Genetic basis of heterosis explored by simple sequence repeat markers in a random-mated maize population. Theor. Appl. Genet. 107:494–502.
Mather, K., and J.L. Jinks. 1982. Biometrical genetics. 3rd ed. Chapman and Hall, London.
Melchinger, A.E. 1987. Expectation of means and variances of test-crosses produced from F_2 and backcross individuals and their selfed progenies. Heredity 59:105–115.
Melchinger, A.E., H.H. Geiger, G. Seitz, and G.A. Schmidt. 1987. Optimum prediction of three-way crosses from single crosses in forage maize (*Zea mays* L.). Theor. Appl. Genet. 74:339–345.
Melchinger, A.E., H.H. Geiger, and F.W. Schnell. 1986. Epistasis in maize (*Zea mays* L.) 2. Genetic effects in crosses among early flint and dent inbred lines determined by three methods. Theor. Appl. Genet. 72:231–239.
Melchinger, A.E., W. Schmidt, and H.H. Geiger. 1988. Comparison of testcrosses produced from F_2 and first backcross populations in maize. Crop Sci. 28:743–749.
Melchinger, A.E., H.F. Utz, and C.C. Schön. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383–403.
Mihaljevic, R., C.C. Schön, H.F. Utz, and A.E. Melchinger. 2005. Correlations and QTL correspondence between line per se and testcross performance for agronomic traits in four populations of European maize. Crop Sci. 45:114–122.
Mihaljevic, R., H.F. Utz, and A.E. Melchinger. 2004. Congruency of quantitative trait loci detected for agronomic traits in testcrosses of five populations of European maize. Crop Sci. 44:114–124.
Patterson, H.D., and E.R. Williams. 1976. A new class of resolvable incomplete block designs. Biometrika 63:83–92.
Piepho, H.-P., and H.G. Gauch. 2001. Marker pair selection for mapping quantitative trait loci. Genetics 157:433–444.
Schnell, F.W. 1965. Die Kovarianz zwischen Verwandten in einer genorthogonalen Population. I. Allgemeine Theorie. Biometr Zeitschrift 7:1–49.
Schnell, F.W., and I.S. Singh. 1978. Epistasis in three-way crosses involving early flint and dent inbred lines of maize. Maydica 23: 233–238.
Schön, C.C., A.E. Melchinger, J. Boppenmaier, E. Brunklaus-Jung, R.G. Herrmann, and J.F. Seitzer. 1994. RFLP mapping in maize: Quantitative trait loci affecting testcross performance of elite European flint lines. Crop Sci. 34:378–389.
Stuber, C.W., S.E. Lincoln, D.W. Wolff, T. Helentjaris, and E.S. Lander. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:823–839.
Utz, H.F., and A.E. Melchinger. 1996. PLABQTL: A program for composite interval mapping of QTL. J. Agric. Genom. 2:1–5. http://www.cabi-publishing.org/gateways/jag/papers96/paper196/indexp196.html; verified 14 July 2005.
Utz, H.F., A.E. Melchinger, and C.C. Schön. 2000. Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. Genetics 154:1839–1849.
Xiao, J., J. Li, L. Yuan, and S.D. Tanksley. 1995. Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. Genetics 140:745–754.
Zeng, Z.-B., C.-H. Kao, and C.J. Basten. 1999. Estimating the genetic architecture of quantitative traits. Genet. Res. (Cambridge) 74:279–289.

5 General Discussion

Correlations between Line *per se* and Testcross Performance

Genotypic correlations between LP and TP, $\hat{r}_g$(LP, TP), estimated herein for five agronomic traits in four populations derived from elite European flint line crosses, were comparable with those obtained for U.S. dent material. The magnitude of $\hat{r}_g$(LP, TP) was trait-specific: for traits of high heritability, such as grain moisture, kernel weight, protein concentration, and plant height, estimates were generally larger than 0.7 across all four populations. For grain yield, estimates were constantly lower and did not exceed an intermediate level of 0.5.

Low estimates of the correlation between LP and TP can be explained by linkage and/or epistasis. But even if linkage and/or epistasis are absent, low estimates $\hat{r}_g$(LP, TP) can result from the masking effects of favorable dominant alleles from a high-performance tester. Testcrosses therefore identify those lines with a high frequency of favorable alleles that are in low frequency (or absent) in the tester. However, lines identified as having high testcross performance by a strong tester may not contain enough favorable alleles or the right favorable alleles to be lines with high LP (Smith, 1986). For an above average inbred tester originating from the same line population and for the biallelic case, the genotypic correlation between LP and TP would be 0.5 or lower as shown by simulations (Smith, 1986).

According to Smith (1986), the genotypic correlation between LP and TP is inversely proportional to the ratio of the genotypic variances for LP and TP when complete dominance and a gene frequency of 0.5 is assumed. Across populations and traits, we found no evidence for this association in our study. The first reason for this is that our experiments did not fulfill one assumption on which Smith's theory was based: our tester was not related to the population, so triallelism rather than biallelism may apply. Second, lines in all four populations differed in their level of heterozygosity, which affected their LP, and third, estimated genotypic correlations showed large confidence intervals, i.e., low precision.

For grain yield, low $\hat{r}_g$(LP, TP) had poor precision (largest confidence intervals compared with the other four traits). This requires testing for both LP and TP and/or combining the data in selection index to improve the performance of both, i.e., ensure sufficient seed yield and yield improvement. In the literature, however, combined selection for LP and TP proved less efficient than selection for TP unless unadapted material without preselection for LP was employed (Gallais, 1997). For grain yield, therefore, it seems more important to map QTL for TP than LP if they are to be used efficiently for MAS.

QTL Mapping across Samples and Populations

One important aspect concerning efficient use of QTL in MAS is congruency of positions and effects of QTL across different samples of the same cross or different crosses. However, QTL analyses are subject to: (i) random errors associated with phenotypic and marker data, (ii) genotypic and environmental sampling, and (iii) bias caused by model selection in multiple regression. Simulations (Utz and Melchinger, 1994; Beavis, 1994) demonstrated that for experiments with small sample size and small QTL effects typical of complex traits like grain yield, the bias in estimates of individual QTL effects as well as the proportion p of the total genotypic variance explained by the QTL detected can be of the same order of magnitude as the true parameters. Because of the resulting low power of QTL detection, only a few QTL will be identified in such an experiment. It is also unlikely that QTL detected with one progeny sample will be re-detected with another independent sample. However, higher congruency is expected for traits of higher heritability like grain moisture, kernel weight, protein concentration, and plant height.

With a QTL of an estimated proportion of phenotypic variance of $R^2 = 0.10$, which corresponds to the average value across all five traits and QTL detected for TP in our study, the power of detecting such a QTL is 0.98 for N = 500 but only 0.65 for N = 100 (Charcosset and Gallais, 1996). The probability of detecting such a QTL simultaneously in two independent samples is obtained by multiplication. Considering bias, the true QTL effect may on average be only about half as large as the estimated QTL effect. This reduces the probability of its simultaneous detection in both samples of the size N = 500

and N = 100 to 0.30. This value is in close agreement with the proportions of congruent QTL detected across samples in this study. Therefore, bias and sampling can well explain the QTL incongruency across samples. Hereby, genotypic sampling generally influences QTL detection and estimation of their positions and effects to a much higher extent than environmental sampling, if more than three test environments are employed (Utz et al., 2000).

Considerable incongruency of QTL also indicates that QTL analyses, as currently performed, can only give limited information on the true number of genes underlying complex quantitative traits. The power of detection is generally too low to provide evidence for the infinitesimal model (Schön et al., 2004).

Even if QTL are detected at congruent positions (within 20-cM distance), this is no guarantee for their usefulness in MAS because no information on the conformity of their effects is given. The latter is provided by two approaches which estimate QTL congruency quantitatively by taking into account both positions and effects of QTL: independent validation and cross validation (Utz et al., 2000). Cross validation is performed without the need for an additional independent sample and yields asymptotically unbiased estimates of *p* (Shao, 1997). Population size of at least 300 employed for mapping QTL of TP in this study and cross validation are recommended if prospects of MAS based on the given QTL results are to be assessed.

Apart from bias and sampling error, incongruency of QTL from different crosses within the same heterotic group can be due to biological reasons. Owing to the high selection pressure exerted in maize breeding programs, equal favorable alleles may be fixed in both parents of a cross of lines from the same heterotic group. Thus, polymorphism at a QTL in one, but its absence in the other cross could be a biological cause for incongruency. Moreover, epistasis can modify the effect of a QTL depending on the genetic background.

In our study, congruency was diminished if one of the parents varied between crosses and was least for unrelated crosses. Exceptions were attributable to large congruent QTL for TP on chromosomes 1 and 8 detected for grain yield and kernel weight, respectively. In these regions, QTL for grain yield and its components have been reported previously (Abler et al., 1991; Beavis et al., 1994; Austin and Lee, 1996; Veldboom and Lee, 1996).

These QTL, which were congruent even among crosses with only one parent in common, may represent a gene cluster or single genes controlling a specific metabolic pathway.

For the investigation of QTL congruency among populations of different genetic background, QTL mapping in multiple-line crosses (Rebai and Goffinet, 2000; Xie et al., 1998; Xu, 1998; Liu and Zeng, 2000), and haplotype-based approach (Jansen et al., 2003) may be more powerful than QTL mapping with biparental crosses.

Whereas for kernel weight, protein concentration, and plant height "large" (Falconer and Mackay, 1996) QTL contributed substantially to the quantitative congruency, this was not the case for grain moisture and yield, probably due to high estimation error of position and heterotic effects of detected QTL or a larger number of small QTL underlying the genetic architecture of these traits. Therefore, MAS across testcross populations seems promising only for the highly heritable traits kernel weight, protein concentration, and plant height. However, its efficacy will primarily depend on its cost efficiency relative to conventional phenotypic selection.

Congruency of QTL for Line *per se* and Testcross Performance

Testcross progenies carry only one allele per locus from either parent in combination with the tester allele. A QTL is detected when the substitution effect of replacing the allele of one parent with the allele of the other parent is significant. The possible interaction of parental alleles with the tester alleles has to be kept in mind when comparing QTL for LP and TP.

For LP and TP similar numbers of QTL were detected in a given population for all traits except grain yield. More than half of the QTL regions detected were in common for LP and TP in the largest population for all traits but grain yield, which suited the $\hat{r}_g$(LP, TP) estimates. The number of detected common QTL may have been reduced due to statistical limitations of QTL analysis as discussed in the previous paragraph. The proportion of common QTL detected for LP and TP of grain moisture, kernel weight, protein concentration, and plant height (i.e., traits with presumably predominant additive gene action) was similar to that found by Melchinger et al. (1998) between testcross

progenies of two different testers. QTL detected with both testers are potential QTL affecting general combining ability (GCA). QTL detected for LP which are common to QTL detected for TP across testers should be predictive for TP, in particular the GCA part of it. For grain yield, substantially fewer QTL were detected for TP in the largest population than in the other populations and for LP. Apart from genetic factors, sampling could be a reason for this.

The quantitative assessment of QTL congruency having a special appeal to MAS was provided by the genotypic correlation between predicted TP based on QTL for LP and observed TP, $\hat{r}_g$ (M_{LP}, Y_{TP}). This estimate should vary accordingly among traits and be a function of the validated genotypic variance explained by the QTL detected for LP. However, the experimental data only partially confirmed these expectations because of the low precision in estimates of $\hat{r}_g$(M_{LP}, Y_{TP}) evidenced by the large confidence intervals especially for grain yield in the smaller populations.

For grain yield, estimated gene action of QTL detected for LP was primarily additive, and evidence for dominance and/or epistasis which influence both heterosis and the correlation between LP and TP was hardly found. Even in the largest mapping population with N = 280, only one of the nine QTL for LP showed significant dominance effect and only one QTL with additive effect showed significant additive × additive epistatic effect as well. It is likely that with $F_{2:3}$ lines rather than F_2 plants, dominance effects are detected on a reduced level. Moreover, the level of dominance for LP detected in a segregating flint population may not be the same as in testcrosses with an unrelated dent tester. Therefore, estimation error seems to be the major reason for the failure of detecting dominance and/or epistatic effects for QTL of LP. Thus, due to statistical limitations the causal analysis of the low correlation between LP and TP for grain yield remains unsatisfactory. For this reason, we performed generation means analyses and pursued genome-wide search for epistasis.

Epistasis between Unlinked and Linked Loci in Testcross and *per se* Generation Means

In the testcross generation means analysis, epistasis between unlinked loci can alter only the means of generations prior to the F_2 (i.e., $\overline{P}$, $\overline{BC}$, and F_1) because the gametic array produced by the F_1 (or any generation derived from it by outbreeding) is already in linkage equilibrium. In this study, contrasts $\overline{P}$ *vs.* F_1, $\overline{P}$ *vs.* $\overline{BC}$, or $\overline{BC}$ *vs.* F_1 were not significant for grain yield and grain moisture and, thus, provided no evidence for epistasis among unlinked loci.

It is, however, the presence of epistasis between tightly linked loci which has been suggested as a major cause for grain yield heterosis and hybrid vigor in maize (Cockerham and Zeng, 1996). Such favorable epistatic gene combinations may get accumulated by selection over several generations if the breeders' practice prefers developing new lines by recycling of elite lines. However, if lines are developed from advanced populations undergoing recurrent selection, random mating, which follows each cycle of selection, will provide enough opportunity for recombination events which disrupt favorable epistatic complexes.

If lines are developed by recycling, the contribution of positive epistasis between linked loci should decline monotonically in the order $\overline{P} > \overline{BC} > F_2 > F_2$-Syn1 > F_2-Syn2 > F_2-Syn3. Although the parents in this study were developed by recycling breeding, we found no significant decline in our testcross generation means analysis and, thus, no evidence for epistasis among linked loci. Theoretically, intermating the F_2 generation for several generations before producing testcrosses alike production of Syn generations is a recommended approach to detect epistasis between linked loci (Lamkey et al., 1995). Herein, it is nevertheless likely that two or three generations of recombination were not sufficient for disrupting tightly linked epistatic complexes of genes. Another reason for our failure to detect epistasis among linked loci may be that positive and negative epistatic effects cancelled each other in sum.

In contrast, half of the estimated additive × additive epistatic effects ($\alpha\alpha^T$) were significant. The latter were tested against the deviation means squares, which are often smaller than the estimated error variance of generation means. Furthermore, interactions of (α^T) and ($\alpha\alpha^T$) with environments were ignored in these tests. When these factors were

taken into account in appropriate F-tests, no significant ($\alpha\alpha^T$) effect was detected. Given the low number of degrees of freedom for the residual error or test environments in the generation means analyses, the power of these tests is relatively poor. A clear distinction of the contribution of unlinked vs. linked locus pairs or epistasis of higher order than ($\alpha\alpha^T$) is complicated by the fact that (i) the effects of epistasis cannot be completely separated from those of linkage (Melchinger, 1987), (ii) epistatic effects are partly contributing to additive effects and higher-order epistatic effects are contributing to estimates of lower order effects (Cheverud and Routmann, 1995), (iii) ($\alpha\alpha^T$) effects are confounded with additive × dominance and dominance × dominance interactions between parental and tester alleles (Eta-Ndu and Openshaw, 1999), and/or (iv) maternal effects are confounded with epistatic effects. In this study, maternal effects must be taken into consideration because the testcross seed was produced in an isolation plot with the dent tester line as pollen parent and the various generations from the four crosses as seed parents.

It was primarily additive × additive type of epistasis we were interested in because our lines were selected for general combining ability with a number of testers. Thus, dominance types of epistasis should be less important in our material. Furthermore, we mainly discussed the results of testcross generation means analyses because these are of direct relevance to hybrid breeding. Comparing analogous effects from the *per se* and testcross generation means analyses, however, is an indicator for the presence of interactions between alleles of parental line and tester alleles.

For example, the additive effect from the testcross generation means analysis (α^T) is confounded by the dominant types of effects between alleles of P1 and P2 and the alleles of the tester. Thus, the additive effect from testcross generation means analysis will be equal to the additive effect from the generation means analysis of LP, i.e. (a), only in the absence of dominance effects between the parental and tester alleles (Melchinger et al., 1998). The same applies to additive × additive types of epistasis. Indications for line × tester interactions on the basis of a disagreement of analogous types of gene effects for LP and TP were particularly evident in one (C×D) of the four crosses investigated. In cross C×D, it was striking that the estimate of (α^T) was fairly small even though the two parents showed pronounced differences in their LP. Thus, the weak line D expressed strong dominance with the tester.

In conclusion, our study clearly confirms the confinements of generation means analysis for an appropriate assessment of the importance of epistasis for quantitative traits. We therefore proceed with the marker-based approach to search for epistasis.

Mapping of Epistatic QTL

Using the marker-based approach, epistasis seemed important at first sight similar to the results reported with other plants (Li et al.,1997; Holland et al., 1997; Kearsey et al. 2003). We detected several marker locus pairs which showed significant epistasis. In general, those marker pairs were not flanking main-effect QTL previously identified by Mihaljevic et al. (2004, 2005). However, when the position of these main-effect QTL was included in the model, the epistatic effects between pairs of marker loci did no longer improve the model fit measured by the Bayesian information criterion.

Thus, estimation of two-locus epistasis is also subject to a number of statistical limitations. First, the true number and position of QTL and, hence, the correct statistical model for estimating their genetic effects, are unknown and must be determined by model selection (Zeng et al., 1999). With multiple regression approach, the general procedure is to identify amongst a large number of regressor variables (markers) those that account for the largest proportion in the variance of the response variable (phenotypic values). Subsequently, these genome positions are used for estimation of QTL effects and the proportion p of genotypic variance explained by the QTL detected. With a limited sample size, model selection leads to an overestimation of QTL effects and p due to sampling effects and consequently to a biased assessment of the prospects of MAS (Melchinger et al., 1998; Utz et al., 2000). The genome-wide search for epistatic effects among QTL is expected to aggravate the problems associated with model selection because the number of regressor variables (marker pairs) and multicollinearity among them increase tremendously. It is therefore highly important to determine the appropriate experiment-wise error rate (Holland, 1998; Holland et al., 1997).

Bogdan et al. (2004) showed by simulations that epistatic terms appearing in a model without the related main effects cause the standard model selection criteria to have a strong tendency to overestimate the number of interactions. Accordingly, the effect size of

epistatic effects estimated herein for LP and TP was reduced when main effect QTL were added to the model. Furthermore, it was not not possible to distinguish between markers with a tight linkage to a QTL pair with small epistatic effect and markers with a loose linkage to a QTL pair with large epistatic effect.

The need for validation with an independent sample or cross validation (Utz et al., 2000) is even more compelling for epistatic than for main effects of QTL. The certainty in the existence of epistatic interactions, however, will require their isolation in a homogenous background by using near-isogenic lines (NILs) (Doebley et al., 1995). This allows to measure single epistatic QTL effects in the absence of the confounding influence of other segregating QTL. The identification of epistatic interactions between QTL is a valuable starting point for a more thorough understanding of genetic networks underlying the inheritance of complex traits (Carlborg and Haley, 2004). Development of high-throughput techniques and bioinformatic tools in the framework of genomics and proteomics provides a new source for the identification of candidate loci that underlie pairs of interacting QTL.

Concluding Remarks

Our results indicate that epistasis hardly influences the testcross means of F_2 and BC populations produced from elite European flint lines. Epistasis can therefore be ignored with regard to the choice of the type of base population to be preferably used in recyling breeding (Melchinger et al., 1988). Nevertheless, it is anticipated that the relative importance of epistatic effects in hybrid maize breeding may strongly increase with the current shift in line development from recurrent selfing to the production of doubled haploids (Seitz, 2004, pers. communication). This is because, with early generation testing in traditional line development, digenic epistasis and even more so higher-order epistasis contribute only marginally to the genetic variance among S_1 or S_2 lines compared to additive effects (Cockerham, 1963). In contrast, with doubled haploid lines (corresponding to S_∞ lines), the coefficients of all epistatic variance components are equal to one and, hence, epistasis contributes fully to the genetic variance from the very beginning of the selection process. Moreover, because recombination is limited to a single meiosis for each

breeding cycle, doubled haploids minimize recombination between linked loci and, thus, should be very effective in preserving tightly linked complexes of genes with positive epistasis. A disadvantage of restricted recombination is, however, the reduced chance of identifying positive complexes of genes if these occur in repulsion phase in the parents. This requires either extremely large population sizes or several generations of intermating before producing the doubled haploid lines.

Genome-wide mapping of epistatic QTL does not show sufficient precision and cannot separate estimated epistatic effects from those of main-effect QTL. This is due to the problem of model selection, even when relatively large sample sizes are used for mapping. On the other hand, we can say about QTL detected for their additive effects that the chances of MAS are substantial if at least a few large QTL are detected, even if some of them are false positives or overestimated. MAS across different samples should be promising in our material for some traits such as kernel weight, protein concentration, and plant height because congruent QTL yielded up to 46% of the genetic variance. For these traits, genetic correlations based on the whole genotype corresponded well to the genetic correlation based on the QTL genotype. Nonetheless, even for these traits we recommend the use of a large population of at least N = 300 and cross validation. As the proportion of the validated variance explained by the QTL detected was still below the estimated heritability for these traits, MAS will have to be more cost-efficient than phenotypic selection to be applied (Lande and Thompson, 1990; Knapp, 1998).

For all traits across populations, estimates of the correlation between the QTL-predicted and observed TP, $\hat{r}_g(M_{LP}, Y_{TP})$, were smaller than those of $\hat{r}_g$(LP, TP) for the whole genotype, because $\hat{r}_g(M_{LP}, Y_{TP})$ is only predictive for the validated proportion of genotypic variance explained by the QTL for LP, which was generally below 50%. Only if a substantial proportion of genotypic variance can be explained by the detected QTL, MAS based on the QTL detected for LP can be applied, provided it is more cost-efficient than the indirect phenotypic selection for TP based on LP selection.

References

Abler, B.S.B., M.D. Edwards, and C.W. Stuber. 1991. Isoenzymatic identification of quantitative trait loci in crosses of elite maize hybrids. Crop Sci. 31:267-274.

Austin, D.F., and M. Lee. 1996. Comparative mapping in $F_{2:3}$ and $F_{6:7}$ generations of quantitative trait loci for grain yield and yield components in maize. Theor. Appl. Genet. 92:817-826.

Beavis, W.D., O.S. Smith, D. Grant, and R. Fincher. 1994. Identification of quantitative trait loci using a small sample of topcrossed and F_4 progeny from maize. Crop Sci. 34:882-896.

Beavis, W.D. 1994. The power and deceit of QTL experiments: Lessons from comparative QTL studies. p. 250-266. *In*: Proceedings of the 49th Annual Corn and Sorghum Industry Research Conference, ASTA, Washington DC, USA.

Bogdan, M., J.K. Ghosh, and R.W. Doerge. 2004. Modifying the Schwarz Bayesian information criterion to locate multiple interacting quantitative trait loci. Genetics 167:989-999.

Carlborg, O. and C.S. Haley. 2004. Epistasis: too often neglected in complex trait studies? Nat. Rev. Genet. 5:618-625.

Charcosset, A., and A. Gallais. 1996. Estimation of the contribution of quantitative trait loci (QTL) to the variance of a quantitative trait by means of genetic markers. Theor. Appl. Genet. 93:1193-1201.

Cheverud, J.M., and E.J. Routman. 1995. Epistasis and its contribution to genetic variance components. Genetics 139:1455-1461.

Cockerham, C.C. 1963. Estimation of genetic variances. p. 53-94. *In* Statistical Genetics and Plant breeding. NAS-NRC 982, Washington D.C.

Cockerham, C.C., and Z-B. Zeng. 1996. Design III with marker loci. Genetics 143:1437-1456.

Doebley, J., A. Stec, and C. Gustus. 1995. *Teosinte branched 1* and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics 141: 333-346.

Eta-Ndu, J.T., and S.J. Openshaw. 1999. Epistasis for grain yield in two F_2 populations of maize. Crop Sci. 39:346-352.

Falconer, D.S., and T.F.C. Mackay. 1996. Introduction to quantitative genetics. 4th ed., Longman Ltd. Group, London, UK.

Gallais, 1997. Combined testcross and S1 selection for the improvement of testcross and inbred performances. Crop Sci. 37:1126-1133.

Holland, J.B., H.S. Moser, L.S. O'Donoughue, and M. Lee. 1997. QTLs and epistasis associated with vernalization responses in oat. Crop Sci. 37:1306-1316.

Holland, J.B. 1998. EPISTACY: a SAS program for detecting two-locus epistatic interactions using genetic marker information. J. Hered. 89:374-375.

Jansen, R.C., J.-L. Jannink, and W.D. Beavis. 2003. Mapping quantitative trait loci in plant breeding populations: use of parental haplotype sharing. Crop Sci. 43:829-834.

Kearsey, M.J., H.S. Pooni, and N.H. Syed. 2003. Genetics of quantitative traits in *Arabidopsis thaliana*. Heredity 91:456-464.

Knapp, S.J. 1998. Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. Crop Sci. 38:1164-1174.

Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743-756.

Li, Z., S.R.M. Pinson, W.D. Park, A.H. Paterson, J.W. Stansel. 1997. Epistasis for three grain yield components in rice (*Oryza sativa* L.). Genetics 145:453-465.

Liu, Y., and Z.-B. Zeng. 2000. A general mixture model approach for mapping quantitative trait loci from diverse cross designs involving multiple inbred lines. Genet. Res. Camb. 75:345-355.

Lamkey, K.R., B.J. Schnicker, and A.E. Melchinger. 1995. Epistasis in an elite maize hybrid and choice of generation for inbred line development. Crop Sci. 35:1272-1281.

Melchinger, A.E. 1987. Expectation of means and variances of testcrosses produced from F2 and backcross individuals and their selfed progenies. Heredity 59:105-115.

Melchinger, A.E., W. Schmidt, and H.H. Geiger. 1988. Comparison of testcrosses produced from F2 and first backcross populations in maize. Crop Sci. 28:743-749.

Melchinger, A.E., H.F. Utz, and C.C. Schön. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383-403.

Mihaljevic, R., C.C. Schön, H.F. Utz, and A.E. Melchinger. 2005. Correlations and QTL correspondence between line *per se* and testcross performance for agronomic traits in four populations of European maize. Crop Sci. 45:114-122.

Mihaljevic, R., H.F. Utz, and A.E. Melchinger. 2004. Congruency of quantitative trait loci detected for agronomic traits in testcrosses of five populations of European maize. Crop Sci. 44:114-124.

Rebai, A., and B. Goffinet. 2000. More about quantitative trait locus mapping with diallel designs. Genet. Res. Camb. 75:243-247.

Schön, C.C., H.F. Utz, S. Groh, B. Truberg, S. Openshaw, and A.E. Melchinger. 2004. Quantitative trait locus mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. Genetics 167:485-498.

Shao, J. 1997. An asymptotic theory for linear model selection. Statistica sinica 7:221-264.

Smith, O.S. 1986. Covariance between line *per se* and testcross performance. Crop Sci. 26:540-543.

Utz, H.F., and A.E. Melchinger. 1994. Comparison of different approaches to interval mapping of quantitative trait loci. p. 195-204. *In* J. W. van Ooijen, and J. Jansen (ed.) Proc. Meeting of the EUCARPIA Section Biometrics in Plant Breeding, 9th, Wageningen, The Netherlands.

Utz, H.F., A.E. Melchinger, C.C. Schön. 2000. Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. Genetics 154:1839-1849.

Veldboom, L.R., and M. Lee. 1996. Genetic mapping of quantitative trait loci in maize in stress and nonstress environments: I. Grain yield and yield components. Crop Sci. 36:1310-1319.

Xie, C., D.D.G. Gessler, and S. Xu. 1998. Combining different line crosses for mapping quantitative trait loci using the identical by descent-based variance component method. Genetics 149:1139-1146.

Xu, S. 1998. Mapping quantitative trait loci using multiple families of line crosses. Genetics 148:517-524.

Zeng, Z-B., C.-H. Kao, and C.J. Basten. 1999. Estimating the genetic architecture of quantitative traits. Genet. Res. Camb. 74: 279-289.

6 Summary

Relations of yield and other important agronomic traits of inbred lines to the same traits in hybrids have been studied from the time of initiation of hybrid breeding to the present. Because crossing lines to a tester and conducting yield trials are expensive and time-consuming, reliable information on inbred lines that is indicative of their testcross performance is crucial for optimum testing schemes in hybrid breeding as well as simultaneous improvement of commercial hybrids and their inbred parents.

It has therefore been of great importance to determine the magnitude of correlation between line *per s*e performance (LP) and testcross performance (TP) and investigate if epistasis influences this correlation. The comprehensive study on hand was performed with five populations (F_3 to F_6 lines) differing in size (ranging from 71 to 344), level of inbreeding, and the number of common parents. The populations employed were derived from three biparental crosses within the heterotic pool of European elite flint maize (*Zea mays* L.). All five populations were evaluated for TP (using an unrelated dent tester inbred) of five agronomically important quantitative traits: grain yield, grain moisture, kernel weight, protein concentration, and plant height. Four of these populations were also evaluated for LP of the same five traits.

The objectives were to (i) estimate phenotypic and genotypic correlations between LP and TP within four populations for all five traits, (ii) map quantitative trait loci (QTL) for LP and TP in four and five populations, respectively, for all five traits, (iii) validate estimated QTL effects and positions for TP by assessing QTL congruency among testcross populations differing in size and genetic background, (iv) determine the value of LP-QTL for the prediction of TP, (v) estimate the importance of epistatic effects for LP and TP of grain yield and grain moisture by generation means analysis as well as genome-wide testing for epistatic marker pairs, and (vi) draw conclusions regarding the prospects of marker-assisted selection (MAS).

Genotypic correlations between LP and TP, $\hat{r}_g$(LP, TP), estimated herein were comparable with those obtained for European flint or U.S. dent material. The magnitude of $\hat{r}_g$(LP, TP) was trait-specific: for traits of high heritability, i.e. grain moisture, kernel weight, protein concentration, and plant height, estimates were generally larger than 0.7 across all four populations, whereas for grain yield, estimates were consistently lower and

did not exceed the intermediate level of 0.5. For grain yield, lowest $\hat{r}_g$(LP, TP) were estimated with lowest precision (largest confidence intervals). This requires testing for both LP and TP and/or combining the data in a selection index to ensure sufficient inbred performance (seed production) and yield improvement. However, combined selection for LP and TP proved less efficient than sole selection for TP unless unadapted material was employed.

For kernel weight, protein concentration, and plant height, we detected "large" congruent QTL across testcross populations derived from the same cross, which individually explained up to 46% of the validated genotypic variance *p*. However, as the *p* values estimated from validation were still below the corresponding heritability estimates, MAS will be superior to phenotypic selection only if it is more cost-efficient.

For the above traits, similar numbers of QTL for LP and TP were detected across populations. More than half of the QTL regions detected for LP were in common for LP and TP in the largest population (N = 280). To assess the value of QTL identified for LP in predicting TP, we calculated the genotypic correlation $\hat{r}_g(M_{LP}, Y_{TP})$. This parameter assesses QTL congruency for LP and TP quantitatively and is thus the key parameter for assessing the prospects of MAS. The number of common QTL for LP and TP (qualitative QTL congruency) was generally not indicative of the magnitude of $\hat{r}_g(M_{LP}, Y_{TP})$ due to the differences in the effect size of the respective QTL detected for LP and used for the prediction of TP.

For all traits, $\hat{r}_g(M_{LP}, Y_{TP})$ were smaller than $\hat{r}_g$(LP, TP). This is because $\hat{r}_g(M_{LP}, Y_{TP})$ is only predictive for the validated proportion of genotypic variance explained by the QTL for LP, which was generally below 50% because of the limited power of QTL detection, in particular with small sample sizes below 100. Only if QTL detected for LP explain a substantial proportion of the genotypic variance, MAS based on these QTL can be applied, provided it is more cost-efficient than an indirect phenotypic selection for TP based on LP.

QTL detection power was drastically reduced for the complex trait grain yield with a presumably large number of small QTL underlying its genetic architecture. Thus, the number of common QTL for LP and TP as well as the QTL congruency across testcross populations was much lower for grain yield than the other four traits. Estimated gene

action of QTL detected for LP was primarily additive for grain yield. Evidence for dominance and/or epistasis, which may be a reason for the low $\hat{r}_g$ (LP, TP) and the low number of common QTL for LP and TP was generally weak.

Both generation means analysis for LP and TP and genome-wide search for epistatic marker pairs yielded no evidence for epistasis. This is not only because the detected epistatic effects could not be validated, but also because there is low chance to find epistasis unless the generation examined displays the full epistatic variance such as expected from doubled haploids produced from an F_1 cross. Thus, it is anticipated that the relative importance of epistatic effects in hybrid maize breeding may strongly increase with the currently happening shift in line development from recurrent selfing towards the production of doubled haploids.

7 Zusammenfassung

Zentrales Ziel in Hybridzüchtungsprogrammen von Mais (*Zea mays* L.) ist die Selektion von Linien mit hoher Kreuzungsleistung. Da die Herstellung und Prüfung der Testkreuzungen in Hybridzüchtungsprogrammen sehr zeit- und kostenaufwendig sind, wurde schon früh in der Geschichte der Maiszüchtung versucht, die Eigenleistung der Linien (EL) als Selektionskriterium für eine Vorauswahl der Linien heranzuziehen. Zudem ist die EL der Linien für eine ökonomische Saatgutproduktion relevant, insbesondere bei der Herstellung von Einfachhybriden.

Die Aussichten einer simultanen Verbesserung der EL- und Testkreuzungsleistung (TL) sowie einer indirekten Verbesserung der TL durch Selektion auf EL werden von der genotypischen Korrelation r_g (EL, TL) zwischen den beiden Selektionskriterien bestimmt. Die Höhe dieser Korrelation wird von einer Reihe genetischer Faktoren bestimmt, unter anderem möglicherweise vom epistatischen Zusammenwirken beteiligter Gene, das ebenfalls Gegenstand dieser Studie war. Die vorliegende Arbeit wurde an fünf Populationen durchgeführt (F_3 bis F_6 Linien), die aus drei biparentalen Kreuzungen zwischen vier Elitelinien des europäischen Flint-Formenkreises hervorgegangen waren. Diese unterschieden sich in ihrem Umfang (zwischen 71 und 344 Linien) und Inzuchtgrad sowie der Anzahl gemeinsamer Eltern. Alle fünf Populationen wurden auf ihre TL mit einer aus dem Dent-Formenkreis stammenden Inzuchtlinie (Tester) evaluiert. Insgesamt wurden fünf agronomisch wichtige quantitative Merkmale erfaßt: Kornertrag, Kornfeuchte, Tausendkorngewicht, Proteingehalt und Wuchshöhe. Vier dieser Populationen wurden gleichzeitig auf ihre EL in diesen Merkmalen geprüft.

Anhand dieses Materials wurden folgende Fragestellungen untersucht: (i) Wie hoch ist die phänotypische und genotypische Korrelation zwischen EL und TL bei wichtigen Merkmalen von Körnermais? (ii) Wie konsistent sind die gefundenen QTL (quantitative trait locus/loci) für ein gegebenes Merkmal in verschiedenen auf TL geprüften Populationen sowie beim Vergleich von EL und TL in verschiedenen auf EL und TL zugleich geprüften Populationen? (iii) Inwiefern liefern die Ergebnisse aus QTL-Analysen für EL und TL eine Erklärung für die geschätzten genotypischen Korrelationen zwischen diesen beiden Kriterien? (iv) Welche Bedeutung haben epistatische Effekte auf der Ebene von Generationsmittelwertanalysen für EL und TL sowie auf der Ebene einzelner QTL?

Die geschätzten genotypischen Korrelationen in unseren Populationen des europäischen Flint-Formenkreises stimmten größenmäßig mit publizierten Schätzwerten aus den US-amerikanischen Studien mit den Linien des Dent-Formenkreises überein. Generell ergaben sich für Merkmale mit höherer Heritabilität und hauptsächlich additiver Genwirkung wie Kornfeuchte, Tausendkorngewicht, Proteingehalt und Wuchshöhe höhere Schätzwerte der r_g (EL, TL) (> 0.7) als für den Kornertrag, für den die niedrigsten Werte mit geringster Präzision ermittelt wurden. Daraus folgt, dass für die Merkmale Kornfeuchte, Tausendkorngewicht, Proteingehalt und Wuchshöhe eine relativ verläßliche Vorhersage der TL aufgrund der EL der Linien möglich ist. Beim Kornertrag hingegen ist eine direkte Bewertung der TL notwendig.

Für Tausendkorngewicht, Proteingehalt und Wuchshöhe wurden für TL in den Populationsvergleichen derselben Kreuzung übereinstimmende QTL gefunden, die einzeln bis zu 46% der validierten genotypischen Varianz erklärten. Da dieser Anteil allerdings unter der Heritabilität einer Prüfung an vier Umwelten liegt, ist die marker-gestützte Selektion (MAS) nur dann effizienter als eine direkte Auslese auf TL, wenn die Beobachtungswerte sehr viel aufwendiger bzw. teurer zu erheben sind als die Markerdaten.

Für diese Merkmale wurden in der größten Population über die Hälfte der für EL detektierten QTL auch für TL detektiert. Die Anzahl der für EL und TL gemeinsamen QTL war über die Populationen allerdings nicht proportional zu der Größe von r_g (M_{EL}, Y_{TL}). Letzteres ist die Korrelation zwischen der vorhergesagten TL aufgrund der QTL-Ergebnisse für EL und der tatsächlich beobachteten TL und somit eine quantitative Erfassung der Übereinstimmung von QTL über EL und TL. Sie stellt den Schlüsselparameter für die Erfolgsaussichten der MAS dar. Die Schätzwerte von r_g (M_{EL}, Y_{TL}) waren bei allen Merkmalen kleiner als r_g (EL, TL) weil die r_g (M_{EL}, Y_{TL}) nur denjenigen Anteil der genotypischen Varianz vorhersagen kann, welcher auch tatsächlich durch die detektierten QTL für EL erklärt wird. Dieser war jedoch generell kleiner als 50% aufgrund der limitierten QTL-Detektionsgüte (Power) bei Populationsgrößen unter 100. Insofern ist auch hier der ökonomische Aspekt bei der Bewertung der Erfolgsaussichten von MAS maßgebend.

Da die Güte der QTL-Detektion bei Populationsgrößen kleiner 100 und insbesondere bei kleinen QTL komplexer Merkmale wie Kornertrag stark abnimmt, reduzierte sich für dieses Merkmal entsprechend die Wahrscheinlichkeit einer gleichzeitigen Detektion für EL

und TL bzw. die konsistente Detektion von QTL in verschiedenen auf TL geprüften Populationen. Große Populationsumfänge sind notwendig, um die Übereinstimmung zwischen QTL-Experimenten und die Aussichten von MAS auch bei mittleren bzw. kleinen QTL beurteilen zu können. Aus den QTL-Analysen für EL ergaben sich weiterhin nur schwache Hinweise auf dominante und epistatische Geneffekte als Ursache für die beobachteten niedrigen Schätzwerte r_g (EL, TL) für Kornertrag.

Generationsmittelwertanalysen für EL und TL sowie genomweite Tests auf Epistasie lieferten ebenfalls keine eindeutigen Hinweise auf Epistasie. Dies ist nicht zuletzt eine Folge der Implementierung von statistischen Validierungsverfahren in dieser Studie, welche die starke Überschätzung der genetischen Effekte in den zurzeit angewandten statistischen QTL-Verfahren aufdecken und zur Vorsicht im Umgang mit den Ergebnissen hinsichtlich ihrer Nutzung für MAS mahnen. Der Nachweis einer Genwirkungsweise ist aber nichtsdestoweniger vom züchterischen Verfahren zur Entwicklung des im Experiment verwendeten Materials abhängig. Insofern stellt der sich vollziehende Wandel in der züchterischen Praxis bei der Entwicklung von Linien in Richtung Produktion von Doppelhaploiden, bei welchen die epistatische Varianz der gekoppelten Loci erhalten bleibt, mit Sicherheit eine Verbesserung für künftige Epistasieuntersuchungen dar.

8 Acknowledgments

At the very end of this treatise I wish to thank everyone who contributed to my research in many different ways. Thanks to all of you!

I am deeply indebted to Prof. Dr. A.E. Melchinger for giving me the opportunity to obtain a Ph.D. on this topic, his continuous scientific advice, creative suggestions, and support in writing this treatise.

I also wish to thank Prof. Dr. H.F. Utz for numerous discussions on science and life. I highly appreciated his comments on my drafts and his excellent statistical advice. He was always available when I needed him.

Special thanks to Ms. C.C. Schön for reviewing a draft of the first publication and her scientific contribution to the second publication.

My appreciation is extended to the people who helped me with the organization of the field trials and together with Prof. Dr. Melchinger gave me the opportunity to work in the winter nursery in South Africa: Dr. D. Klein and his team in Eckartsweier, and F. Mauch and his team in Hohenheim.

Thanks to all my colleagues and friends in the institute, especially Martin, Željko, Elisabeth, Beate, Susanne, and Astrid. I also gratefully appreciate the kindness and assistance of Ms. B. Boesig.

My special gratitude to my dearest Stanislava, Nevenka, Duro, Suzana, Martin, Oliver, Julian, and Steffen for their continuous support, patience, and love.

This study was funded by the Deutsche Forschungsgemeinschaft (DFG).

9 Curriculum Vitae

Personal Details

Name:	Renata Mihaljević
Date of Birth:	21 July 1974
Place of Birth:	Zagreb, Croatia
Nationality:	Croatian
Marital Status:	unmarried

Elementary and High School Education

1980 – 1988	Elementary School in Croatia
1988 – 1993	Friedrich-Schiller-Gymnasium in Fellbach
1993 – 1994	Private School for Business English – **staatl. anerkannte Wirtschaftskorrespondentin**

University Education

10/1994 – 08/2000	Agricultural sciences at the University of Hohenheim – **Dipl-Ing. sc. agr**. Diploma thesis at the Institute of Plant Breeding, Seed Science and Population Genetics at the University of Hohenheim (Prof. Dr. A.E. Melchinger): "Characterization of resistance gene analogs from the *Scmv1*- and *Scmv2* region of maize" published in Theoretical and Applied Genetics 2002. Vol. 105: 355-363.
10/2000 – 11/2004	Institute of Plant Breeding, Seed Science and Population Genetics at the University of Hohenheim (Prof. Dr. A.E. Melchinger) – **Doctorate candidate for a degree Ph.D. in Agricultural Sciences**

Work Experience

04/1997 – 09/1997	Internship on the farm Sohn in Esslingen-Mettingen

01/1998 – 04/2000	Research assistant at the Institute of Plant Breeding, Seed Science and Population Genetics at the University of Hohenheim (Prof. Dr. H.H. Geiger and Prof. Dr. A.E. Melchinger)
12/2004 – to date	Employee with Pioneer Hi-Bred Northern Europe Service Division GmbH at the Research and Technology Center in Eschbach, Germany