Theor Appl Genet (1995) 91:33–37

R. C. Jansen · J. W. Van Ooijen · P. Stam
C. Lister · C. Dean

Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci

Received: 10 October 1994 / Accepted: 24 October 1994

Abstract The interval mapping method is widely used for the genetic mapping of quantitative trait loci (QTLs), though true resolution of quantitative variation into QTLs is hampered with this method. Separation of QTLs is troublesome, because single-QTL is models are fitted. Further, genotype-by-environment interaction, which is of great importance in many quantitative traits, can only be approached by separately analyzing the data collected in multiple environments. Here, we demonstrate for the first time a novel analytic approach (MQM mapping) that accommodates both the mapping of multiple QTLs and genotype-by-environment interaction. MQM mapping is compared to interval mapping in the mapping of QTLs for flowering time in *Arabidopsis thaliana* under various photoperiod and vernalization conditions.

Key words *Arabidopsis thaliana* · Flowering time · Genotype-by-environment interaction · Mapping · Quantitative trait loci

Flowering time in *Arabidopsis*

Arabidopsis thaliana is a model organism for genetic analysis because of its small genome size, short generation time and ease of propagation (Meyerowitz and Pruitt 1985). Transition to flowering is one of the current issues in *Arabidopsis* research (Martinez-Zapater et al. 1994). At least 12 loci for flowering time have been identified by mutational analysis (Koornneef et al. 1991). Also, large differences between ecotypes exist for flowering time; the FRI locus was found to be responsible for some of these differences (Clarke and Dean 1994). The group of early genotypes, which includes the widely used ecotypes Columbia (Col) and Landsberg *erecta* (L*er*), has not been analyzed extensively. Small differences in flowering time within this group have been reported (Koornneef et al. 1991), and it has been suggested that the FLC locus is involved (M. Koornneef, personal communication). Flowering time strongly depends on many environmental factors, amongst which photoperiod and temperature (vernalization treatment) are the most important. Distinct norms of reaction have been reported for several mutants and ecotypes (Martinez-Zapater et al. 1994). In the present article we report the genetic mapping of quantitative trait loci (QTLs) underlying the differences in flowering time between Col and L*er*. Flowering time was recorded under various photoperiod and vernalization conditions in a set of recombinant inbred lines (RILs, Table 1) derived from a cross between Col and L*er*; details of the experimental conditions will be presented elsewhere (Lister and Dean, manuscript in preparation). We used 37 of the restriction fragment length polymorphism (RFLP) markers previously mapped by Lister and Dean (1994).

We successfully applied a novel method of analysis based on multiple-QTL models (MQM mapping; Jansen 1994); QTL-by-environment interaction forms part of the models fitted. For comparison we also applied interval mapping (Lander and Botstein 1989), analyzing the data for each environment separately.

Comparison of interval mapping and MQM mapping

In interval mapping (IM) the likelihood for a single putative QTL is assessed at each map location on the

Communicated by F. Salamini

R. C. Jansen (✉) · J. W. Van Ooijen · P. Stam[1]
Department of Population Biology, Centre for Plant Breeding and Reproduction Research (CPRO-DLO), PO Box 16, 6700 AA Wageningen, The Netherlands

C. Lister · C. Dean
Department of Molecular Genetics, Cambridge Laboratory, JIC, Norwich NR4 7UJ, United Kingdom

Present address:
[1] Department of Plant Breeding, Wageningen Agricultural University, 6709 DB Wageningen, PO Box 386, 6700 AJ Wageningen, The Netherlands

genome and, in the case of multiple environments, a QTL likelihood map is produced for each environment separately (Fig. 1). MQM mapping is an automatic two-stage procedure in the first stage of which "important" markers and marker-by-environment interactions are identified by the backward elimination method in multiple regression on all markers and environments (including interactions). In the second stage the likelihood for a single putative QTL is assessed at each map location (like in IM), but the preselected markers are used as cofactors, except for selected markers flanking the interval under study (Fig. 2). Marker cofactors will hopefully eliminate the major part of the variation induced by QTLs located elsewhere on the genome. A likelihood curve for a single putative QTL with QTL-by-environment interaction is plotted in regions where interaction with environment is still assumed (Fig. 2). The overall 5% significance thresholds for the tests (for the hypothesis of QTL with no QTL-by-environment interaction and for the hypothesis of QTL with QTL-by-environment interaction) were obtained by computer simulation using the actual marker data and analyzing 1000 replicates. Observed flowering times were log-transformed prior to analysis, and distinct variance parameters for each of the environments were included.

With MQM mapping we found evidence of 12 QTLs; 4 of these display QTL-by-environment interaction (Fig. 2). The QTL-by-environment effects indicate a QTL-by-vernalization interaction, where vernalization decreases the effects of these QTLs (Fig. 2). The result is not unexpected, because vernalization considerably decreases both environmental and genetic variance (Table 1). Further, QTL-by-photoperiod interactions are indicated. For instance, on chromosome 2 the QTL near marker m323 has little effect at LD but a large effect at CL. A full analysis of the identified QTLs and their relationship to previously mapped flowering time loci will be presented elsewhere (Lister and Dean, manuscript in preparation).

We now compare the results of MQM mapping with those of IM and discuss the features of MQM mapping and IM that contribute to the differences. In IM, single-QTL models are used, and independence of residual errors is a basic assumption. In our experiment, however, each RIL is tested in six environments, and the six observations are correlated via genetic identity of the underlying genes. Therefore, the usual assumption of independent residual errors may be seriously violated and a joint analysis accommodating the information from all environments is not possible with IM; QTL likelihood maps can only be produced for each environment separately (Fig.1). Although environment-specific QTLs may be detected this way, the approach is intrinsically weak, because the interaction is not part of the genetic model that is being fitted with IM. In MQM mapping with a complete linkage map however, the major part of this correlation is removed by markers that are used as cofactors in the model. This makes it possible to produce a joint map, including QTL-by-environment interaction, in the univariate regression frame of MQM mapping (Fig. 2).

The IM analysis indicates the presence of at least 4 QTLs in several environments (Fig.1).The fact that IM detects a QTL at a specific map region in one environment but not in another environment may indicate QTL-by-environment interaction (for instance, a QTL is detected near maker g4552 on chromosome 1 in environments SD and LD but not in CL). In the absence of true QTL-by-environment interaction, however, a QTL can also be detected in one environment and not in another environment because the chance of simultaneous detection in both environments is small.Therefore, the IM analysis may be indicative but cannot be conclusive on the presence of QTL-by-environment interaction. However, if pattern of environment-specific QTLs really results from QTL-by-environment interaction, this is readily, and more powerfully detected by MQM mapping.

In MQM mapping, genetic background "noise" is removed by using marker cofactors (Jansen 1994). Therefore, the chance of detecting QTLs is generally higher in MQM mapping than in IM. Further, separation of linked QTLs is much easier in MQM mapping than in interval mapping (Jansen 1994). In IM, linked QTLs of unidirectional effect tend to be mapped as a single "ghost-QTL" at some intermediate position on the marker map (Martinez and Curnow 1992; Jansen 1994); also, linked QTLs of opposite effect may go unnoticed because of their mutually neutralizing effects (Jansen 1994). Both situations and even the more complex configuration of multiple linked QTLs with effects of alternating sign have been encountered in our *Arabidopsis* experiment (Figs. 1 and 2). For instance on chromosome 2, MQM analysis indicates the presence of 2 QTLs. In IM, the QTL near m246 is found in LD and LDV. In the other environments a QTL is mapped at various positions (in the middle of the chromosome near g6842 in CLV, SDV and SD, and near m323 in CL), but support intervals are very large. This illustrates the problems in separating linked QTLs with unidirectional effects by IM. Here, the situation is even more complex due to QTL-by-environment interaction for the QTL near m323. Another example of linked QTLs is indicated on chromosome 3, where MQM mapping detects 2 QTLs. In IM, the QTL near m583 is detected in LDV, SDV and CLV but not in LD, SD and CL. The presence of the second QTL, with opposite effect only in LD, SD and CL, near m457 is one of the reasons for this. The other chromosomes exemplify similar problems in mapping of linked QTLs by IM.

The present study clearly illustrates the advantages of the MQM approach over IM in the detection and mapping of multiple genes underlying quantitative traits, especially when data have been collected in multiple environments. Therefore, we feel that our approach is another step forward towards understanding the genetics of quantitative characters. Our results also

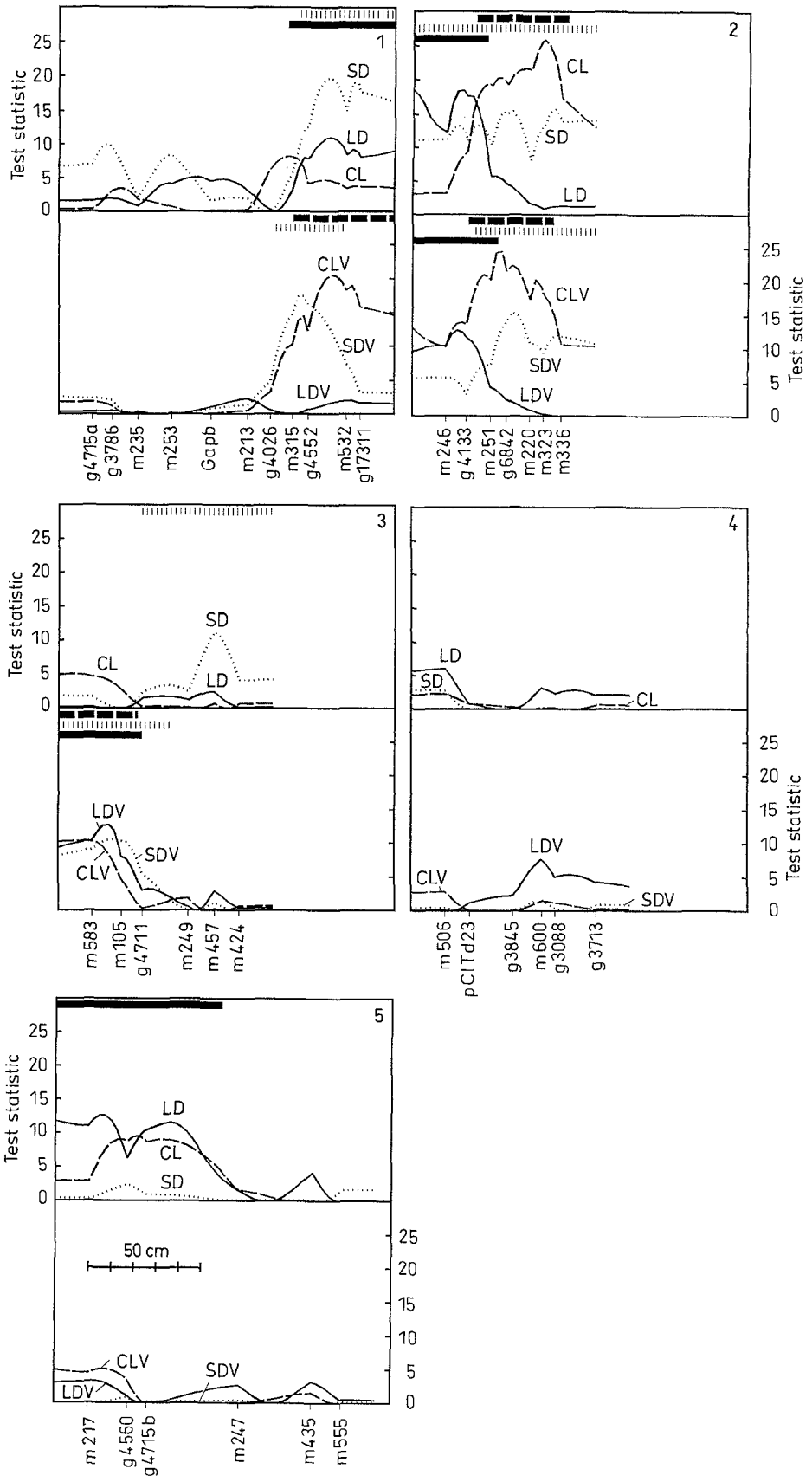

Fig. 1 Genetic mapping of QTLs for flowering time (expressed by leaf number) in *Arabidopsis thaliana*: QTL likelihood maps produced by interval mapping (IM). Chromosome number is indicated at the *right-hand top corner* of each graph; markers are plotted along the abscissa. The *solid*, *dashed* and *dotted curves* represent the test statistic (twice the log of the likelihood ratio) for the hypothesis of a QTL (with no QTL-by-environment interaction) in the environment indicated. The overall 5% significance threshold for the test is 10. *Solid*, *dashed* and *dotted bars* represent two lod (^{10}log of likelihood ratio) support intervals for the map locations of detected QTLs (pattern of curves and bars are corresponding) (*SD* short day (10 h of light), *LD* long day (16 h of light), *CL* continuous light, *LDV*, *SDV*, and *CLV* LD, SD and CL + vernalization, respectively)

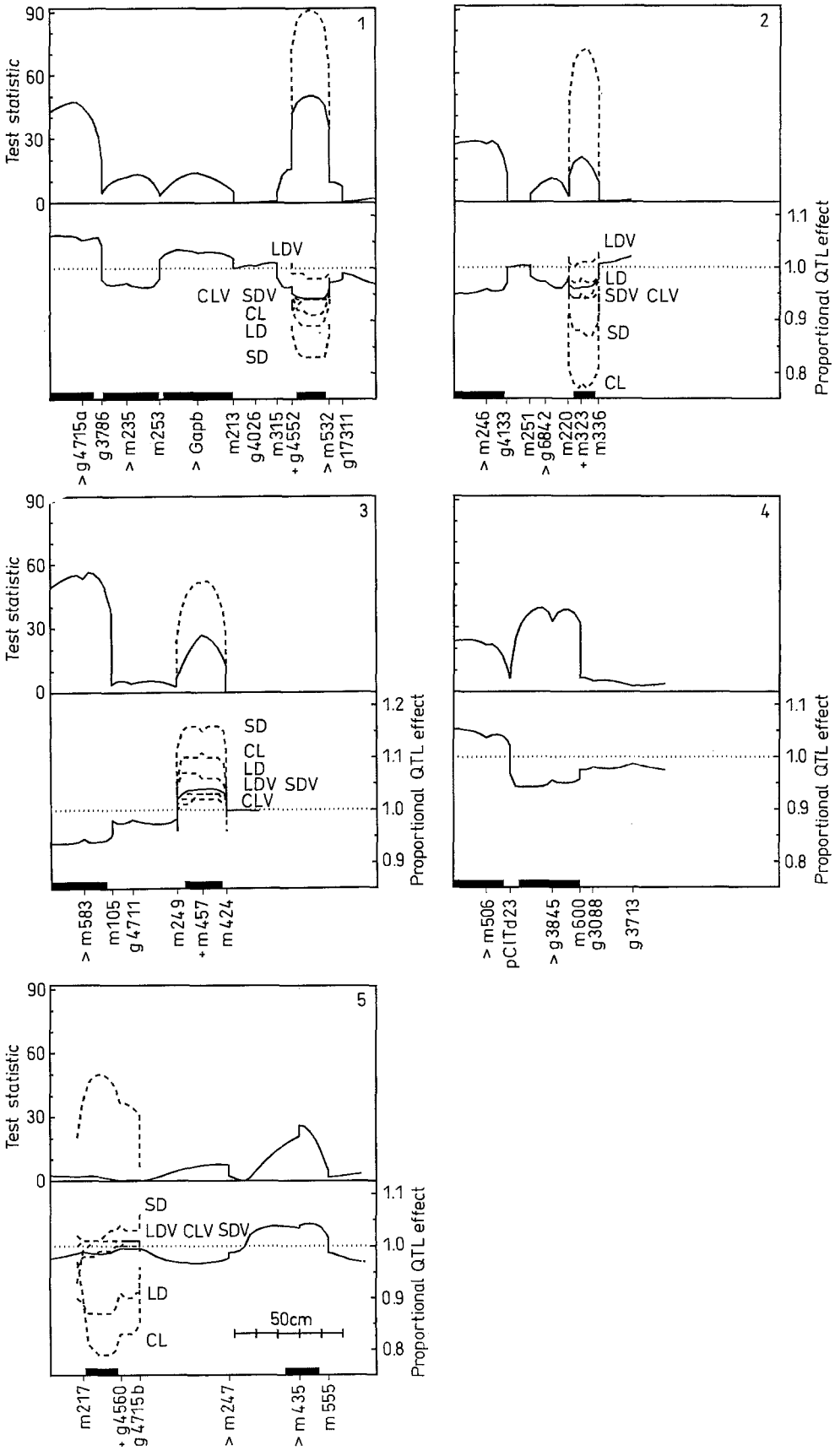

Fig. 2 Genetic mapping of QTLs for flowering time (expressed by leaf number) in *Arabidopsis thaliana*: QTL likelihood maps and QTL effect maps produced by MQM mapping. Chromosome number is indicated at the *right-hand top corner* of each graph, markers are plotted along the abscissa. Selected markers are indicated by + when interaction with environment is still assumed, otherwise by ∧. *Solid curves* indicate the test statistic (twice the log of the likelihood ratio) for the hypothesis of a QTL with no QTL-by-environment interaction assumed (*upper part*) and the estimated QTL effect (*lower part*). The overall 5% significance threshold for this test is 11. *Dashed curves* represent the test statistic for the hypothesis of a QTL with QTL-by-environment interaction (*upper part*) and the estimated QTL effects (*lower part*). The overall 5% significance threshold for the interaction test (the difference between solid and dashed curve) is 22. *Bars* along the *abscissa* indicate two lod (^{10}log of likelihood ratio) support intervals for the map locations of detected QTLs. The QTL effect is expressed proportionally, i.e. the replacement of the putative QTL allele of Col by that of *Ler* (a) has no effect if the QTL effect is equal to 1, (b) proportionally increases the number of leaves, if the QTL effect is larger than 1 and (c) proportionally decreases the number of leaves, if the QTL effect is smaller than 1 (*SD* short day (10 h of light), *LD* long day (16 h of light), *CL* continuous light, *LDV*, *SDV*, and *CLV* LD, SD and CL + vernalization, respectively)

Table 1 Genetic mapping of QTLs for flowering time (expressed by leaf number[a]) in *Arabidopsis thaliana*: some population parameters

Environmental conditions[b]	Phenotypic mean			Phenotypic variance between RILs	Multiple regression of RIL phenotypes on all 37 markers	
	Col[c]	*Ler*[c]	RILs[d]		Residual variance	Variance explained
LD	9.9	7.1	8.6	1.63	0.79	52%
LDV	9.0	7.4	8.6	0.38	0.23	39%
SD	32.9	28.3	29.4	26.69	10.12	66%
SDV	22.2	19.5	21.2	5.51	3.37	39%
CL	18.1	11.5	12.8	9.76	5.55	43%
CLV	11.3	8.3	10.3	0.78	0.37	53%

[a] Leaf number is often taken as a measure of flowering time; leaf numbers in this table represent the total number of rosette and cauline leaves per plant

[b] SD = short day (10 h of light), LD = long day (16 h of light), CL = continuous light, LDV, SDV and CLV = LD, SD and CL + vernalization, respectively, Col = Columbia and *Ler* = Landsberg *erecta*

[c] Two sets of five plants per environment were tested

[d] In total 99 recombinant inbred lines (RILs) were tested, each RIL with five plants per environment

suggest that re-analysis of several QTL experiments reported in literature (cf. Paterson et al. 1988, 1991; Stuber et al. 1992; De Vicente and Tanksley 1993; Hayes et al. 1993; Schön et al. 1993; Laurie et al. 1994) may further lift the veil that covers the link between phenotype and genotype.

Acknowledgement The assistance of Maarten Koornneef in the preparation of the manuscript is gratefully acknowledged.

References

Clarke JH, Dean C (1994) Mapping FRI, a locus controlling flowering time and vernalization response. Mol Gen Genet 242:81–89

De Vicente MC, Tanksley SD (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134:585–596

Hayes PM, Liu BH, Knapp SJ, Chen F, Jones B, Blake T, Franckowiak J, Rasmusson D, Sorrells M, Ullrich SE, Wesenberg D, Kleinhofs A (1993) Quantitative trait locus effects and environmental interaction in a sample of North american barley germ plasm. Theor Appl Genet 87:392–401

Jansen RC (1994) Controlling the type I and type II errors in mapping quantitative trait loci. Genetics 138:871–881

Koornneef M, Hanhart CJ, Van der Veen JH (1991) A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. Mol Gen Genet 229:57–66

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

Laurie DA, Pratchett N, Bezant JH, Snape JW (1994) Genetic analysis of a photoperiod response gene on the short arm of chromosome 2(2H) of *Hordeum vulgare* (barley). Heredity 72:619–627

Lister C, Dean C (1993) Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. Plant J 4:745–750

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

Martinez-Zapater JM, Coupland G, Dean C, Koornneef M (1994) The transition to flowering in *Arabidopsis*. In: Meyerowitz EM, Sommerville CR (eds) *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. pp 403–433

Meyerowitz EM, Pruitt RE (1985) *Arabidopsis thaliana* and plant molecular genetics. Science 229:1214–1218

Paterson AH, Lander ES Hewitt JD, Peterson S, Lincoln SE, Tanksely SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726

Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander ES,Tanksley SD (1991) Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. Genetics 127:181–197

Schön CC, Lee M, Melchinger A, Guthrie WD, Woodman WL (1993) Mapping and characterization of quantitative trait loci affecting resistance against second-generation European corn borer in maize with the aid of RFLPs. Heredity 70:648–659

Stuber CW, Lincoln SE, Wolff DW, Helentjaris T,Lander ES (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:823–839