Theor Appl Genet (2001) 103:1243–1253

C. Maliepaard · M.J. Sillanpää · J.W. van Ooijen
R.C. Jansen · E. Arjas

Bayesian versus frequentist analysis of multiple quantitative trait loci with an application to an outbred apple cross

Received: 17 July 2000 / Accepted: 13 January 2001

Abstract Two methods, following different statistical paradigms for mapping multiple quantitative trait loci (QTLs), were compared: the first is a frequentist, the second a Bayesian approach. Both methods were applied to previously published experimental data from an outbred progeny of a single cross between two apple cultivars (*Malus pumila* Mill.). These approaches were compared with respect to (1) the models used, (2) the number of putative QTLs, (3) their estimated map positions and accuracies thereof and (4) the choice of cofactor markers. In general, the strongest evidence for QTLs, provided by both methods, was for the same linkage groups and for similar map positions. However, some differences were found with respect to evidence for QTLs on other linkage groups. The effect of using cofactor markers which were selected differently was also somewhat different.

Keywords QTL mapping methods · Comparison · Bayesian inference · Classical statistics · MQM

Introduction

The analysis of quantitative trait loci (QTLs) using molecular markers has become routine in genetic studies of plant and animal species (see Tanksley 1993; Haley 1995; Doerge et al. 1997; Hoeschele et al. 1997; Kearsey and Farquhar 1998). The detection of QTLs is mostly based on simple interval mapping (SIM), which uses a single-QTL model, or on methods employing an approximate multiple QTL model (MQM) mapping (Lander and Botstein 1989; Haley and Knott 1992; Jansen 1992, 1993, 1994; Zeng 1993, 1994; Jansen and Stam 1994; Xu and Atchley 1995; Kao and Zeng 1997). Haley and Knott (1992) and Martínez and Curnow (1992) applied an exact two-QTL model to map linked QTLs. However, numerical computations tend to become very time consuming as the number of QTLs in the model increases (Jansen 1992). Kao et al. (1999), and Zeng et al. (1999) proposed a stepwise estimation and a model selection framework for an entire genetic architecture covering multiple QTLs, epistasis, G×E and pleiotropy. Carlborg et al. (2000) suggest the utilization of genetic algorithms for a similar purpose.

The use of Markov Chain Monte Carlo (MCMC) methods has greatly facilitated the numerical work arising from considering exact multiple QTL models, often based on incomplete data. These methods seem to open new opportunities for model selection; for example, with regard to the number of QTLs. By applying MCMC sampling in the Bayesian statistical framework, Thaller and Hoeschele (1996), Uimari et al. (1996a) and Uimari and Hoeschele (1997) estimated the number of linked QTLs by using fixed models with a specific linkage-status-variable assigned to each QTL. Satagopan et al. (1996) used the Bayes factor for comparing models that involved different numbers of QTLs. In the Bayesian 'reversible jump' framework (Green 1995; Waagepetersen and Sorensen 2001), parameter estimation and model selection can be considered simultaneously by treating the number of QTLs as an unobserved random variable (see Satagopan and Yandell 1996; Heath 1997; Uimari and Hoeschele 1997; Stephens and Fisch 1998; Sillanpää and Arjas 1998, 1999; George et al. 2000; Yi and Xu 2000a, b; Lee and Thomas 2000; Hurme et al. 2000; Uimari and Sillanpää 2001). Given this wealth of new statistical techniques, of which some are based on the more classical frequentist principles and others on the Bayesian paradigm, it is only fair to enquire about their philosophical and practical differences in a QTL mapping context.

The Bayesian approach to statistical inference is based on an openly subjectivist interpretation of proba-

Communicated by H.C. Becker

C. Maliepaard and M.J. Sillanpää contributed equally to the first authorship

C. Maliepaard · J.W. van Ooijen · R.C. Jansen
Centre for Biometry Wageningen, Plant Research International, P.O. Box 16, 6700 AA Wageningen, the Netherlands

M.J. Sillanpää (✉) · E. Arjas
Rolf Nevanlinna Institute, Research Institute of Mathematics, Statistics and Computer Science, P.O. Box 4, FIN-00014 University of Helsinki, Finland
e-mail: mjs@rolf.helsinki.fi
Tel.: +358 9-191-22764, Fax: +358 9-191-22779

bilities. The probability of an event, or more generally, of a proposition, is viewed as a quantification of "my" uncertainty, or degree of belief, in that this event will happen or that the proposition is true. The frequentist interpretation of probability, on the other hand, views probabilities as limiting frequences of the event concerned, calculated from a (hypothetical) sequence of such events realised under similar conditions (see Royall 1997; Vieland and Hodge 1998; Shoemaker et al. 1999). In classical statistics, probabilistic statements such as confidence intervals or P-values therefore refer to hypothetical repetitions of the experiment, with the randomness corresponding to the variation between the resulting data sets. Probabilities of propositions cannot be defined within this framework because their being true or false cannot be thought of as random outcomes in a sequence of experiments. The Bayesian approach to statistics, on the other hand, is entirely probabilistic, providing also a structured template to incorporate prior information, when available, to the analysis. The result of the Bayesian data analysis can be summarised in the form of the so-called posterior distribution; that is, the (joint) conditional distribution of all unobserved variables in the considered model given the observed data. Key quantities of interest are then conveniently estimated in terms of expectations with respect to such distributions, which are usually carried out in practice by an MCMC procedure. It should be noted that the application of these methods requires substantial understanding and experience of the behaviour of MCMC algorithms under different circumstances (Sillanpää 1999). In contrast to the Bayesian probabilistic way of summarising the results, in the classical statistical analysis the results are mostly presented in the form of point estimates of the model parameters of interest, with additional support coming from statistical significance testing or from finding their approximate confidence regions.

In view of the fact that very different inferential approaches and statistical techniques are being applied, it should come as no surprise that also the results of such analyses can differ. Consequently, two paradigms cannot directly be compared. For a comparative analysis of their performance – for example, in the sense of being able to detect true QTLs, and specifically so as not to raise false alarms – one should obviously use simulated data where the true answers are known. In genuine empirical studies such as the present one, we do not have a certain knowledge of what answers are correct and therefore no clear winner in a comparison between different statistical methods, let alone different paradigms, can be declared. On balance, 'real life' data sets may provide a more challenging environment for testing, in that one or more of the following complications may be present: (1) missing data patterns, (2) uneven distribution of markers on the chromosomes and large variation in their degree of informativeness, (3) deviation from assumed distributions, (4) scoring and measurement errors, (5) outliers and (6) real genetic architecture. An interesting alternative, pointed out to us by a referee, may be found between these two: to simulate new phenotypes (from a known genetic model) to a given 'real life' data set, where marker measurements are real. However, the properties of items (4) and (6) above can then be lost. (In fact, such an approach had been used by us prior to analysing the real data in order to verify proper mixing of the sampling algorithm used.)

Methodological comparisons between classical and Bayesian techniques in QTL mapping have so far been made by the following authors. Scheler et al. (1998) compared such methods in an inbred line cross situation with simulated data. However, their 'Bayesian' method is not strictly Bayesian; it is a likelihood method, where all QTL positions are integrated away from the likelihood expression. Uimari et al. (1996b) compared their Bayesian method to frequentist methods using simulated and experimental data in an outbred livestock population with a granddaughter design. They found a good agreement between these methods in their location estimates. Sillanpää and Arjas (1998, 1999) compared their Bayesian method with SIM and MQM using simulated data sets and inbred and outbred experimental designs. Hurme et al. (2000) further compared the Bayesian method of Sillanpää and Arjas (1999) with single-marker-regression using data from Scots pine. Yi and Xu (2000b) compared their Bayesian method with an ML-based method in simulated data. Royall (1997), Vieland (1998) and Vieland and Hodge (1998) compared Bayesian and classical statistics, providing some considerations of a more general nature. Interesting discussions can also be found in Dupuis (1996) and Dupuis and Siegmund (1999). Recently, Shoemaker et al. (1999) emphasised some areas in genetics where Bayesian statistical methods can be particularly useful.

The purpose of this paper is to compare the frequentist methods of SIM and approximate MQM mapping with the Bayesian multiple QTL analysis. For this comparison we used the experimental data from a single large full-sib (FS) family derived from a cross between two apple cultivars. The methods are compared with respect to: (1) the models used, (2) the number of QTLs mapped, (3) the estimated map positions of the QTLs and their accuracies and (4) the choice of marker cofactors. The same data set has been previously analysed using SIM by King et al. (2000).

Materials and methods

Experimental data

A cross between the apple cultivars 'Prima' and 'Fiesta' was carried out at Plant Research International, the Netherlands, in 1988, using Prima as the female parent. A full–sib progeny consisting of 152 genotypes from this cross was vegetatively propagated, and replicate sets or subsets of this progeny plus parents were planted at seven sites in six countries in Europe in 1993 (King et al. 1991; King 1996). In 1995 and 1996 apples from trees from six sites were analysed for fruit firmness using two test methods: (1) acoustic resonance frequency (RF) (Abbott et al. 1992; Chen and De Baerdemaker 1993) and (2) hand penetrometer (PEN)

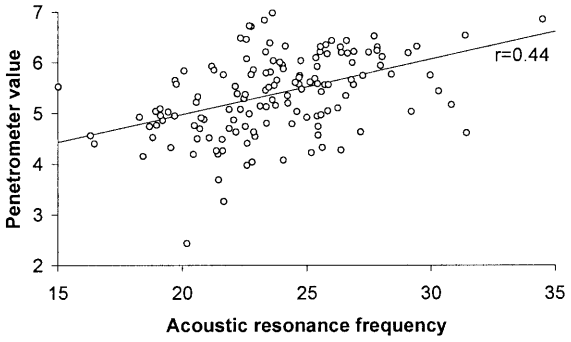


Fig. 1 Scatter plot of penetrometer estimates by resonance frequency estimates of the Prima×Fiesta progeny. Each *circle* indicates an individual in the progeny

(Magness and Taylor 1925; Bourne 1974). For both sets of phenotypic measurements, data over sites and years were analysed using REML in GENSTAT 5, weighting with the number of apples per tree measured and taking as random factors the site/year combination and trees of the same genotype within a site/year; the genotype was taken as fixed. In this way estimates for each genotype were obtained, and these were used for QTL analysis. For the resonance frequency the estimates were over three site/year combinations; for the penetrometer readings the estimates were over nine site/year combinations. Figure 1 shows a scatter plot of the PEN estimates plotted against the RF estimates.

Multipoint information content

The information content of each linkage group was calculated, using the experimental genotype data and taking into account the numbers of missing values and the upgrading of multipoint marker information if markers were only partly informative or if there were missing values. The information content was first defined, for each individual, as the maximum of the four QTL genotype probabilities at a map position. (Here each QTL genotype corresponds to one of four possible pairs of parental alleles whose grandparental origins are known.) These maxima were then averaged over the individuals in the progeny.The information content was calculated with steps of 1 cM and varied from linkage group to linkage group, with a minimum value of 0.58 and an average of 0.86 over the whole genome. The information content of linkage group L01, using a maximum of ten neighbouring markers to upgrade the information, is shown in Fig. 2.

QTL analysis

Linkage maps of Prima and Fiesta were constructed using 290 markers to genotype both parents and the progeny (see Maliepaard et al. 1998). Sixty-seven multi–allelic markers for which both parents were heterozygous allowed the two parental maps to be integrated (17 linkage groups). Linkage phases were first estimated using JOINMAP™ version 2.0 (Stam and Van Ooijen 1995). The marker order and distances on the integrated linkage map and the parental linkage phases were then assumed known in both the frequentist and the Bayesian QTL analysis.

In a full-sib family a QTL or a marker can segregate for four distinct alleles, i.e. parental mating type $Q_1Q_2 \times Q_3Q_4$, producing four different genotypes. Therefore, in both the frequentist and Bayesian approach, three effects (deviances from the effect of the first genotype) are modelled for a QTL and each cofactor. As usual only additive QTL terms (no epistasis) were considered in the models.

Frequentist QTL analysis

In SIM and MQM mapping, the EM-algorithm (Dempster et al. 1977; Van Ooijen 1992) is used for estimation, and standard statistical procedures are used for testing hypotheses. In MQM, the cofactor effects are estimated simultaneously with the QTL effects (Jansen 1994). The LOD profiles over each of the linkage groups were used to determine the map positions of QTLs. These were estimated as the position with the maximum LOD score on a linkage group. Uncertainty of the map position was indicated by a 2–LOD support interval (Conneally et al. 1985; Van Ooijen 1992). The number of QTLs was inferred from the number of LOD peaks exceeding the significance threshold.

In addition to the determination of the 2-LOD support intervals, bootstrapping (Efron 1979, 1982) was also used to obtain approximate central 95% confidence intervals for QTL positions (Visscher et al. 1996), which were expected to be more comparable to the Bayesian credible intervals of Sillanpää and Arjas (1998). The position with the maximum LOD score was retained after each bootstrap analysis.

In this study, as in King et al. (2000), LOD scores greater than 3.0 were considered as evidence of a QTL. LOD scores greater than

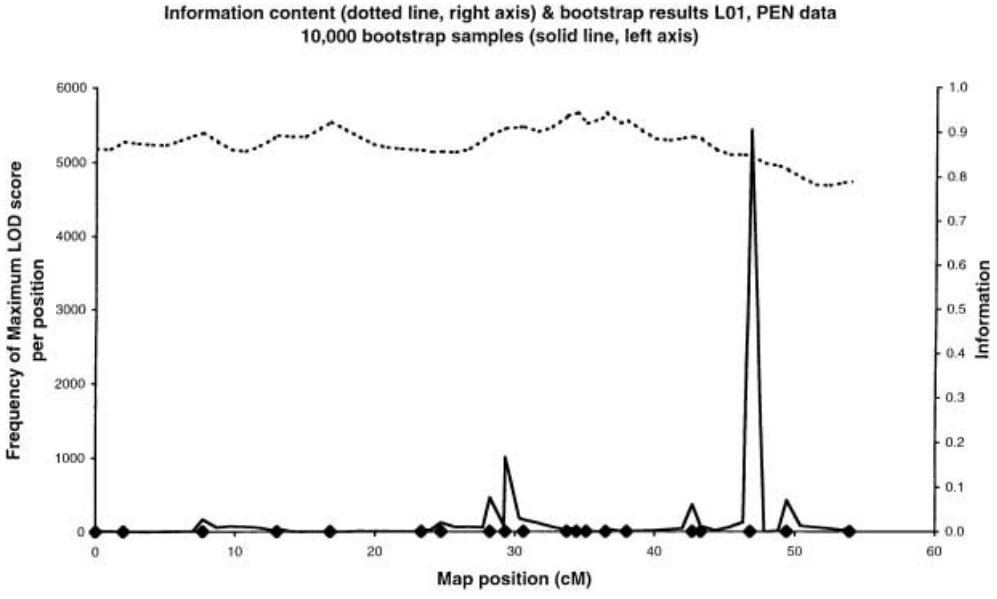


Fig. 2 Information content for linkage group L01, using ten neighbouring marker intervals for upgrading linkage information for missing or partially missing marker genotypes (*right axis*, *upper graph*) and empirical frequency distributions of the estimated positions of QTLs represented as frequency polygons (*left axis*, *lower graph*). Results are from interval mapping of bootstrap samples of the PEN data from 152 individuals in the Prima×Fiesta progeny, linkage group L01. A step size of 1 cM was used. *Symbols* on the *horizontal axis* indicate marker positions. The left *y-axis* corresponds to the frequency of the maximum LOD score per map position

4.5 were considered significant. These values correspond to a per linkage group error rate of 5% for the average linkage group length (63 cM) and a genome–wide error rate of 5% (Van Ooijen 1999).

Flanking markers were used to calculate the probabilities of the four QTL genotypes at a given map position. For missing marker data and for markers that were not completely informative with respect to the four possible parental allelic combinations, marker genotype information from neighbouring markers was used ('all–markers mapping', Knott and Haley 1992; Maliepaard and van Ooijen 1994). Usually up to five neighbouring markers on both the left- and right-hand side of the marker interval were used to upgrade the genotype information. Only for linkage group L01 were ten neighbouring markers used, since there was a group of markers at one end of the linkage group which provided information with respect to one parent only. For cofactors the same procedure was applied in MQM mapping.

The MAPQTL™ 4.0 package (Van Ooijen et al. 2000) was used both for SIM and MQM mapping in this outbred full–sib progeny. Results from SIM were checked against results obtained with the LS approach of Haley et al. (1994) and against the non–parametric test of Kruskal-Wallis, performed on single markers.

Bayesian QTL analysis

Following Sillanpää and Arjas (1998, 1999), an exact multiple QTL model was used for one chromosome while the other chromosomes were controlled by using a preselected set of cofactor markers. In Bayesian analysis, the QTL mapping problem is not formulated in a sequential (hypothesis) testing framework as in frequentist methods. All results of the analysis can be expressed in terms of the posterior distribution of the unknown variables/parameters in the model, given the data. Convenient summary measures for variables of interest, such as the number of QTLs and the QTL positions, can be defined by considering suitable marginals of the posterior distribution or corresponding expectations. Bayesian posterior credible intervals (e.g. Sillanpää and Arjas 1998) for these parameters can be constructed from the marginals of the posterior distribution. In principle, any interval can be taken as a credible interval. The main advantage of credible intervals is in their straightforward interpretation in terms of conditional probabilities of containing the unknown parameters given the data.

In the numerical estimation of the model parameters in the adopted Bayesian hierarchical model (Sillanpää and Arjas 1999), the Metropolis-Hastings-Green algorithm was used (Metropolis et al. 1953; Hastings 1970; Green 1995; Chib and Greenberg 1995; Waagepetersen and Sorensen 2001). In each round of the estimation, the QTL genotype probabilities were determined for each individual in the offspring given current values for completed fully informative markers and/or QTLs. In each round, incomplete marker data were completed for missing genotypes and linkage phases and coded according to their grandparental origin by using all other markers in the same linkage group. With equal probabilities, this block-updating was conducted for the whole family and separately at each marker point or for the entire haplotype and separately for each individual. This modification of the sampling scheme of Sillanpää and Arjas (1999) is described in the Appendix of Hurme et al. (2000). The idea behind it was mentioned in Thompson and Heath (1999). Missing cofactor genotypes were augmented by assuming marker independence and by using M-H where acceptance of the imputed values was always conditional on the current parameter values. Initially in each MCMC run, three QTLs (which was also the maximum allowed) were placed evenly along each linkage group to be analysed. A truncated Poisson distribution was used with mean equal to 2 as the prior for the number of QTLs. The prior for the residual variance was uniform over the range [0.0, 3.35] in the RF data and over [0.0, 0.69] in the PEN data. The right endpoints of these ranges are equal to the variance estimates from the corresponding data. The prior for the regression intercept was taken to be uniform over [–100, 100] in the RF data and over [–13, 13] in PEN data. The prior for QTL genotype effects was N(0,100) in both data sets, and the prior for cofactor effects was uniform over [–13, 13]. The prior for the QTL location was taken to be uniform over the entire length of the particular linkage group. The random walk (see Chib and Greenberg 1995) proposal ranges in the MCMC analyses were chosen to be 2.0 (location), 1.0 (intercept), 0.2 (residual S.D.), 1.5 (QTL coefficients) and 2.0 (cofactor coefficients). The proposal distribution for new QTL effects was N(0.0, 0.5). The burn-in period was not deleted, since high numbers of MCMC cycles (from 2500000 to 5000000) were run in all analyses. In the estimation (Monte Carlo averaging), the MCMC samples were thinned, using only every tenth iteration, because of the limited storage capacity. Credible intervals for the positions of QTLs were constructed from the posterior QTL intensities, as in Sillanpää and Arjas (1998).

The posterior distribution of the number of QTLs in a mapped chromosome can be used as an initial summary measure of the analysis. Based on this measure, chromosomes showing some QTL activity can then be investigated further by looking at their posterior QTL intensities (Sillanpää and Arjas 1998) along the chromosome.

It is important to note here that the whole Bayesian analysis for experimental data was conducted independently from the mapping results obtained by MQM. All MCMC calculations were performed using MULTIMAPPER/OUTBRED software (http:/www.rni.helsinki.fi/~mjs). An overview of packages used in this study for frequentist and Bayesian multiple QTL analysis is described in Manly and Olson (1999).

Selection of cofactors, frequentist analysis

After a first round of SIM, cofactors were selected from regions where the LOD was greater than 3.0. MQM analysis was then performed using these markers as cofactors. The low threshold was chosen since in SIM the error variance still comprises genetic variance from other segregating QTLs and therefore the full power of the test is not yet used. In subsequent rounds of MQM mapping, marker cofactors were added or dropped according to this 3.0 threshold. Because of the 'all–markers mapping' approach Maliepaard and van Ooijen 1994 any marker could be chosen as a cofactor, regardless of the segregation type (informativeness) or the number of missing values. Initially, no marker cofactors were used on the linkage group where a QTL was fitted. However, for linkage groups with evidence of a QTL, MQM mapping was also done using cofactors on those same linkage groups in order to check for the possibility of having detected a ghost QTL.

Selection of cofactors, Bayesian analysis

In the Bayesian approach a single preliminary analysis was performed (without marker cofactors) using the multiple QTL model, allowing for up to three QTLs on each linkage group under investigation. Based on this multiple-QTL analysis, cofactors were then chosen from linkage groups in the regions showing higher than 0.2 posterior probability for single or multiple QTLs. In fact, a rather sharp distinction between linkage groups with and without evidence of a QTL was observed. For the selected linkage groups, cofactors were chosen from the regions showing high and condensed posterior QTL intensities.

Note that in the Bayesian approach no cofactors were chosen from the linkage group to be analysed because a multiple QTL model was used. Following Sillanpää and Arjas (1999) cofactor genotypes were augmented without using genotype information from neighbouring markers, although in principle this can be done the same way as for the QTL genotypes. Instead, the most informative marker in a region (distinguishing four genotypes in the progeny) was chosen as a cofactor, or a set of two or more cofactor markers, in order to maximise the information content.

We also remark that from an inferential perspective a more coherent Bayesian procedure would have been to consider all linkage groups jointly in a single variable dimensional QTL analysis, always using the "current" QTLs as cofactors. The reason for performing separate analyses for each linkage group, each being conditional on cofactors determined in a preliminary analysis, was essentially that this made the computations much less demanding.

Results

Preliminary analysis (no cofactors used)

Frequentist QTL analysis (SIM)

For the PEN data the threshold was exceeded on five linkage groups, including L01 and L10. For the RF data, the LOD score threshold of 3.0 was exceeded on linkage groups L01 and L10 (Table 1). The LOD score graph for PEN for L01 clearly showed a double peak; the graph of RF was rather irregularly shaped, but also showed multiple peaks (Fig. 3). The 2-LOD support intervals for the PEN and RF data for L01 and L10 are presented in Table 2 and indicated in Fig. 3.

Table 1 Preliminary analysis (without cofactors). Maximum LOD score and the posterior distribution of the number of QTLs in linkage groups with evidence of a QTL in simple interval mapping and/or Bayesian analysis with penetrometer and acoustic resonance frequency data

Linkage group	*Frequentist*:	*Bayesian*:		
	Maximum LOD	Posterior distribution		
		Number of QTLs		
		0	1	2 or 3
Penetrometer				
L01	6.5	0.56	0.43	0.00
L03	3.6	1.00	0.00	0.00
L08	4.7	0.96	0.03	0.00
L10	7.4	0.09	0.91	0.00
L15	3.5	0.98	0.01	0.01
Others	<3	>0.98	<0.01	<0.02
Resonance frequency				
L01	4.4	0.26	0.73	0.01
L10	4.5	0.31	0.68	0.00
L11	1.5	0.25	0.73	0.03
L15	2.6	0.73	0.26	0.01
Others	<3	>0.92	<0.08	<0.03

Bayesian QTL analysis

The preliminary Bayesian analysis of the PEN data provided evidence of a single QTL on linkage groups L01 and L10. On all other linkage groups the posterior probability of one or more QTLs was less than 0.04. Analysis of the RF data resulted in three linkage groups where the posterior probability of at least one QTL exceeded 0.5. For linkage group L15 the posterior probability for one QTL was 0.26, on all other linkage groups the posterior probability of one or more QTLs was less than 0.08 (Table 1). The data did not support the existence of more than a single QTL on any of the linkage groups, although for linkage group L01 a strong bimodality was observed, both for the PEN and the RF data (Fig. 3). The credible regions for the PEN and RF data are indicated in Table 2, together with the posterior probabilities of containing at least one QTL in these respective regions.

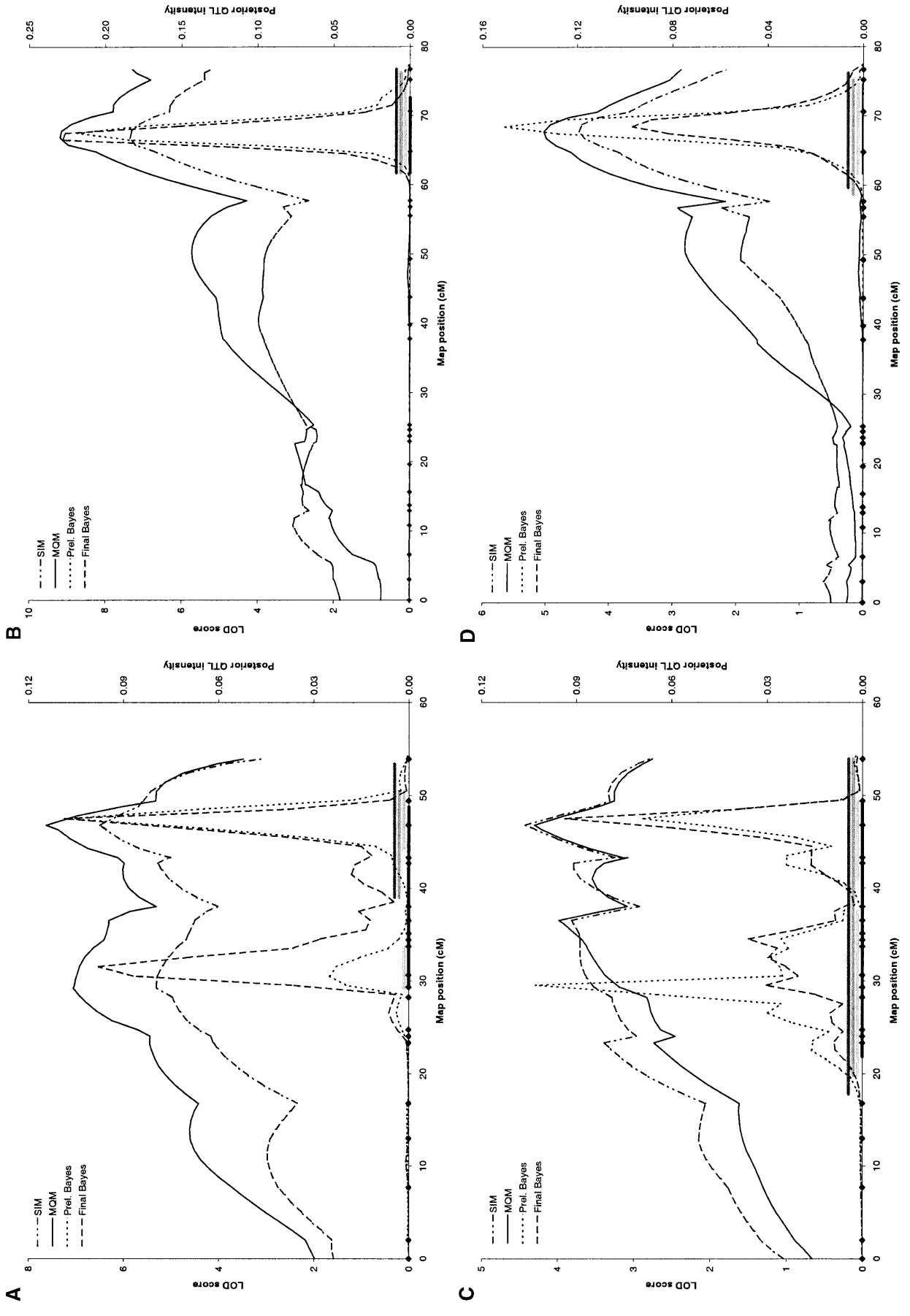
Choice of cofactors and final analysis

Frequentist QTL analysis

Based on the results of the preliminary analysis, cofactor markers were chosen on L01, L03, L08, L10 and L15 for the PEN data. Using these cofactors in MQM mapping, the maximum LOD score of linkage group L03 decreased to a value below 3.0, so that the marker on L03 was dropped as a cofactor. No new regions with LOD scores greater than 3.0 were obtained in the next round. The final model for the PEN data included cofactors on linkage groups L01, L10 and L15. On these linkage groups the LOD significance threshold of 4.5 was ex-

Table 2 2-LOD support intervals in the frequentist QTL (SIM and MQM) analyses and credible regions in the Bayesian (preliminary and final) analyses for linkage groups L01, L10 and L15. The posterior probabilities of containing at least one QTL in the corresponding credible regions were estimated directly from the MCMC runs, as the proportion of iteration cycles in which there was at least one putative QTL within the region

Linkage group	Preliminary analysis (without cofactors)			Final analysis (including cofactors)		
	Frequentist:	*Bayesian*:	Posterior	*Frequentist*:	*Bayesian*:	Posterior
	SIM	Preliminary	Probability	MQM	Final	Probability
Penetrometer						
L01	[39.0, 53.4]	[29.1, 33.5]∪[44.2, 49.9]	0.39	[39.0, 50.4]	[29.9, 49.3]	0.72
L10	[61.7, 76.2]	[63.1, 74.3]	0.89	[61.7, 76.2]	[62.3, 72.7]	0.92
L15				[0.0, 17.4]	[0.0, 9.1]	0.19
Resonance frequency						
L01	[17.8, 53.9]	[20.1, 37.5]∪[40.8, 49.9]	0.71	[23.3, 53.9]	[22.1, 49.3]	0.57
L10	[59.7, 76.2]	[62.2, 73.4]	0.66	[58.7, 76.2]	[61.6, 74.6]	0.52
L15		[0.0, 2.1] ∪ [7.3, 9.5] ∪ [25.4, 37.0]	0.18	[0.0, 32.9]	[0.0, 17.4]∪ [23.8, 30.9]	0.57


A
B
C
D
LOD score
Posterior QTL intensity
Map position (cM)
SIM
MQM
Prel. Bayes
Final Bayes

ceeded. The LOD score for linkage group L08 (4.3) was just below this threshold (Table 3). Compared to the SIM results, the 2–LOD support interval on L01 was just a little bit smaller, on L10 the interval was identical (Table 2).

For the RF data, two cofactor markers were chosen, on L01 and L10. In the analysis of the RF data with these cofactors, a LOD score of 4.0 was obtained on linkage group L15. LOD scores for the markers on L01 and L10 remained above 3.0. A cofactor marker on linkage group L15 was added and in the next round no new genome regions with LODs over 3.0 were found. The final model for the RF data included only the cofactor marker on linkage group L10. LOD scores for the other linkage groups were below the significance threshold of 4.5 (Table 3). For the RF data also, the 2–LOD support interval for L01 was a bit smaller than in SIM but much larger than the intervals estimated for the PEN data. The L10 interval was practically identical to the situation in SIM and also to the intervals estimated for the PEN data (Table 2). The effect on linkage group L01 was mainly a contrast between the alleles from Prima; on L10 mainly from Fiesta (King et al. 2000).

Table 3 Final analysis (including cofactors). Maximum LOD score and the posterior distribution of the number of QTLs in linkage groups with evidence of a QTL in MQM mapping and/or Bayesian analysis with penetrometer and acoustic resonance frequency data

Linkage group	*Frequentist*:	*Bayesian*:		
	Maximum LOD	Posterior distribution		
		Number of QTLs		
		0	1	2 or 3
Penetrometer				
L01	7.6	0.22	0.78	0.00
L03	2.6	1.00	0.00	0.00
L08	4.3	0.99	0.01	0.00
L10	9.2	0.05	0.95	0.00
L15	5.8	0.79	0.21	0.00
Others	<3	>0.98	<0.01	<0.03
Resonance frequency				
L01	4.2	0.41	0.59	0.01
L10	5.0	0.43	0.56	0.01
L11	1.5	0.94	0.06	0.00
L15	4.0	0.37	0.62	0.01
Others	<3	>0.91	<0.09	<0.01

Bayesian QTL analysis

For the PEN data three cofactor markers were chosen: two on linkage group L01 and one on linkage group L10. For the RF data a total of seven cofactors were chosen on linkage groups L01, L10, L11 and L15. These consisted of two markers on L01 to cover a larger area of this linkage group (to account for the two peaks), two on each of linkage groups L11 and L15 to cover for all possible allelic combinations and also a larger area, and one informative marker on L10.

For the PEN data there was strong evidence for a QTL on linkage groups L01 and L10, some evidence of a QTL on L15 and hardly any evidence for a QTL on the other linkage groups. In the analysis of the RF data with cofactors, most linkage groups did not present any evidence of the presence of one or more QTLs. There was evidence of a single QTL on linkage groups L01, L10 and L15 (Table 3). There was no support to more than a single QTL on any of these linkage groups. The posterior probability of a QTL on linkage group L11 decreased considerably when cofactors were used. Figure 3 shows some of the results from both analyses with and without cofactors. Credible regions for the PEN and RF data are given in Table 2, together with the posterior probabilities for the presence of at least one QTL in these respective regions.

◀ **Fig. 3A–D** LOD score graphs for interval mapping (SIM) and MQM mapping and Bayesian posterior QTL intensity graphs of preliminary and final analyses of linkage groups which showed evidence of a QTL. Posterior QTL intensities with bin length 1 cM are represented as frequency polygons. Explanations of the different line types are given in the *upper left corner* of each panel. *Bars* at the bottom of each graph indicate 2-LOD support intervals and Bayesian credible intervals for the four situations, in the same order (*top* to *bottom*) as the legend. *Symbols* on the *horizontal axes* indicate marker positions. The *left* (*right*) *y-axis* corresponds to the LOD score (posterior QTL intensity). **A** Linkage group L01, PEN data, **B** linkage group L10, PEN data, **C** linkage group L01, RF data; **D** linkage group L10, RF data

Discussion

Comparison of the frequentist with the Bayesian approach: correspondence

The frequentist and Bayesian approach for multiple QTL analysis were both applied to two data sets collected from a single full-sib family of apple. The data sets, PEN and RF, were different in the way the fruit firmness phenotypes were measured. Phenotypes in both data sets were genotype means over different sites and years. King et al. (2000) analysed the individual data sets per site and year previously, together with additional sensory measurements of fruit firmness. These individual data sets indicated possible QTLs on linkage groups L01 and L10 and some evidence for QTLs on L15. In the present study evidence was also found for QTLs on these linkage groups, and for these there was a good agreement between the results of the frequentist and the Bayesian method. Both methods indicated a single QTL for fruit firmness (both for the PEN data and the RF data) on linkage groups L01 and L10, and perhaps L15 (in the frequentist approach stronger evidence for L15 came from the PEN data, whereas in the Bayesian approach it

was from the RF data). For these linkage groups there also was a very good correspondence of the estimated QTL positions, as indicated by the LOD score and Bayesian QTL intensity peaks (Fig. 3).

There was no indication of a second QTL on either of these linkage groups, even though the graphs for linkage group L01 showed bimodality for the Bayesian intensity for both the PEN and the RF data. This was also visible for the LOD score but less extreme than in the Bayesian analysis. This bimodality could not be explained by the variation in the information content, since the information content is very high throughout this linkage group and there is no visible decrease in information in the region where the posterior density or the LOD score drops (Fig. 2). Upon inspection of the marker data, it was found that there were three double recombinants in the region between 36 cM and 43 cM. However, these could not explain the decrease in the LOD score curve as was verified by more detailed inspection (results not shown). When studying the single recombinants in the region between 29 cM and 43 cM, we observed that the phenotypic values of these individuals indeed could explain a decrease in significance going from 29 to 38 cM, and a subsequent increase in significance from 38 to 43 cM. It seems likely that the sampling bias among the recombinants is responsible for the observed bimodality of the curves. It cannot be excluded that there may also be some errors in the marker data or in the map order and/or the map distances.

Comparison of the frequentist with the Bayesian approach: differences

Although the results of the Bayesian and the frequentist approach agreed very well on QTLs on linkage groups L01 and L10, still some differences between the methods were observed. In the preliminary frequentist analysis of the PEN data, LOD scores larger than 3.0 were also found on linkage groups L03, L08 and L15. The preliminary Bayesian analysis did not provide evidence for QTLs on these linkage groups. On the other hand, the preliminary Bayesian analysis resulted in an elevated posterior density for a QTL on linkage group L11 (RF data), whereas with the frequentist approach a maximum LOD score of only 1.5 was obtained (Table 1). Additionally, the non–parametric test of Kruskal–Wallis was used to verify the results without being required to assume normality of the data. This test also indicated the possible presence of QTLs on linkage groups L03 and L08 and slight evidence of a single marker on linkage group L11.

In this study rather diffuse priors for parameters in the model were used, and therefore we do not believe that posterior inferences were noticeably influenced by the priors. It is also unlikely that the use of the ML procedure gave rise to the high LOD scores in linkage groups L03, L08 and L15, since the LS method (Haley et al. 1994) resulted in almost equal LOD scores. Apparently, the difference between the Bayesian and frequentist method is also not just a matter of a difference in power: evidence for a QTL is found with one method and not with the other, as well as *vice versa*.

Cofactor choice and its effect

In both methods the choice of cofactors was based on the results from the preliminary analysis, so marker cofactors were chosen only in those regions with elevated LOD scores or high QTL intensities. Note that this does not necessarily provide us with the optimal set of cofactors. For example, in MQM when two QTLs are present on a linkage group, the highest LOD score may be found in between these two (as a ghost QTL). Choosing a cofactor at that position may absorb most of the genetic effects generated by the two QTLs, so that these will not be detected in subsequent rounds. This could be prevented by performing a backward elimination procedure to select cofactors on a linkage group of interest (Jansen 1993). In this study possible ‘ghost QTLs’ were checked for on hindsight by using different pairs of cofactors on those linkage groups with evidence of a QTL. In the Bayesian multiple QTL model, this ‘ghost QTL’ behaviour is not expected. The choice of an incorrect cofactor may also occur when the information content is rather variable across the linkage group. In smaller data sets there is also the danger that, due to chance, a major QTL has distorted segregation within the marker classes of a marker on a different linkage group or partial cosegregation with an unlinked marker. When this occurs, another type of ‘ghost QTL’ may be detected on the latter through association with the real QTL. In fact, partial cosegregation was observed between a set of marker pairs on linkage groups L01 and L08, and this may explain the decrease in the LOD score for L08 when a cofactor on L01 was used. However, currently it is not yet feasible to compare and evaluate efficiently different possible sets of cofactors. The logical solution to this problem would be to consider the entire genome in a single multiple–QTL analysis.

Although the choice of cofactors was based on the same principles in both methods, the cofactors were used differently in the models, and also the choices were different. Consequently, the effect of using the marker cofactors seemed different in the two methods. In the frequentist method the effect was generally an increase of the LOD scores, while the shape of the LOD score graph remained very similar. In the Bayesian analysis the differences were more notable. For L11 the change was rather drastic, since the posterior probability for a single QTL on this linkage group decreased from 0.73 down to 0.06. The increase for L15 (from 0.26 to 0.62) was also rather large. For the Bayesian analysis, the shape also changed. For example, for the PEN data the bimodality on L01 became stronger when cofactors were used, and the intensity in the region around 30 cM increased so that the two peaks became almost equally high.

Position estimates and their accuracies

The estimated positions of QTLs were very similar for both methods and data sets and irrespective of the cofactors chosen. The correspondence is striking especially for linkage group L10. The QTL position estimates obtained with the Bayesian approach visually appear to be more accurate than the results with the frequentist approach. This is because they show sharper peaks, whereas the LOD curve is rather flat. The chosen credible regions, however, are not very different from the 2-LOD support intervals. Note that the results cannot be compared directly. With respect to the visual appearance of the peaks, the results from the Bayesian analysis were expected to be more comparable to results from bootstrapping. Indeed, these were more similar (Fig. 2). A part of the difference is explained by the logarithmic scale of LOD scores, whereas results from the Bayesian analysis and from bootstrapping are based on frequency distributions. In the bootstrapping results we observed a similar bias as Walling et al. (1998), resulting in higher frequencies of the maximum LOD score at the marker positions, especially near the estimated QTL position. Unlike in the examples considered in Sillanpää and Arjas (1998, 1999), the posterior probabilities that a given credible region contains at least one QTL were here calculated directly from the MCMC output, as explained in Table 2. Such a direct approximation seems to give more accurate numerical estimates that the earlier method which was based on a Poisson approximation. Where the LOD support intervals are concerned, these are defined differently than confidence intervals. Van Ooijen (1992) demonstrated, both for BC and F_2 populations, that 2-LOD support intervals may be conservative only if the QTL effect is large. Dupuis and Siegmund (1999) showed with simulations that 1–LOD and 1.5–LOD support intervals provided a QTL coverage probability of approximately 90% and 95%, respectively, for dense maps (markers at every 1 cM) and an even greater percentage for sparse maps. These authors also compared confidence regions in simulations with a single QTL and concluded that the coverage probabilities of LOD support regions and Bayesian credible intervals were roughly comparable in large samples.

Multiple linked loci

In this study we found no evidence for more than a single QTL on any linkage group, so that we were not able to compare the performance of the two methods when more QTLs are present. In general, the Bayesian method seems to be well suited to detect multiple QTLs on a linkage group since these are modelled explicitly. This is supported by simulation studies (Sillanpää and Arjas 1998, 1999). Although MQM mapping can also be used to detect multiple QTLs on a single linkage group, the necessary computation time may turn out to be long if there are also cofactors on other linkage groups and if the 'all-markers mapping' approach is applied to upgrade marker information for all cofactors and the fitted QTL. This may be solved by using more multi-allelic markers and by omitting those QTL genotype combinations which have a probability close to zero (Jansen 1995).

Environmental cofactors

In neither the frequentist nor the Bayesian method were environmental cofactors such as the site/year combinations included in the model, although this is certainly possible and has been done (i.e. Jansen et al. 1995; Tinker and Mather 1995; Korol et al. 1998; Hurme et al. 2000). This would also be more in agreement with the Bayesian paradigm of using all prior information and of including uncertainties rather than using point estimates. The use of estimated means over sites and years for QTL mapping may have undesirable effects since some genetic effects may be lost in an adjustment for environmental cofactors. It would be preferable to also include these environmental cofactors into the analysis and estimate all effects simultaneously. However, this would be computationally more demanding, and the sample size in this case would not allow for a reliable estimation of all main and interaction effects.

Conclusion

Both methods provided evidence for the main QTLs on the same linkage groups, and with similar map positions. However, there were also some differences with respect to evidence for QTLs on other linkage groups. The response to adding cofactor markers was also somewhat different. The shape of the graphs of the LOD score and Bayesian posterior intensity were found to differ as well. Neither method provided evidence for more than a single QTL on any linkage group.

Acknowledgements We are grateful to Dr. Marco Bink for his helpful comments to the manuscript. The experimental data were kindly provided by the partners in the CEC project CT920473, 'Development of the European Apple Crop', which was funded by the CEC Framework III research and technical development programme in Agriculture and Agricultural Industries and the national governments of the EU.

References

Abbott JA, Affeldt HA, Liljedahl LA (1992) Firmness measurement of stored 'Delicious' apples by sensory methods, Magness-taylor, and sonic transmission. J Am Soc Hortic Sci 117:590–594

Bourne MC (1974) Comparison of results from the use of the Magness-Taylor pressure tip in hand- and machine-operation. J Text Stud 5:105–108

Carlborg Ö, Andersson L, Kinghorn B (2000) The use of a genetic algorithm for simultaneous mapping of multiple interacting quantitative trait loci. Genetics 155:2003–2010

Chen H, De Baerdemaeker J (1993) Effect of apple shape on acoustic measurements of firmness. J Agric Eng Res 56:259–266

Chib S, Greenberg E (1995) Understanding the Metropolis-Hastings algorithm. Am Stat 49:327–335

Conneally PM, Edwards JH, Kidd KK, Lalouel J-M, Morton N, Ott J, White R (1985) Report of the committee on methods of linkage analysis and reporting. Cytogenet Cell Genet 40:356–359

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc Ser B 39:1–38

Doerge RW, Zeng Z-B, Weir B (1997) Statistical issues in the search for genes affecting quantitative traits in experimental populations. Stat Sci 12:195–219

Dupuis J (1996) Statistical aspect of trait mapping using a dense set of markers: a partial review. In: Speed T, Waterman MS (eds) Genetic mapping and DNA sequencing. Springer, Berlin Heidelberg New York, pp 111–131

Dupuis J, Siegmund D (1999) Statistical methods for mapping quantitative trait loci from a dense set of markers. Genetics 151:373–386

Efron B (1979) Bootstrap methods: another look at the jackknife. Ann Stat 7:1–26

Efron B (1982) The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia

George AW, Mengersen KL, Davis GP (2000) Localization of a quantitative trait locus via a Bayesian approach. Biometrics 56:40–51

Green PJ (1995) Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination. Biometrika 82:711–732

Haley CS (1995) Livestock QTLs – bringing home the bacon? Trends in Genet 11:463–524

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

Haley CS, Knott SA, Elsen J-M (1994) Mapping quantitative trait loci in crosses between outbred lines using least squares. Genetics 136:1195–1207

Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97–109

Heath SC (1997) Markov Chain Monte Carlo segregation and linkage analysis for oligogenic models. Am J Hum Genet 61: 748–760

Hoeschele I, Uimari P, Grignola FE, Zhang Q, Gage KM (1997) Advances in statistical methods to map quantitative trait loci in outbred populations. Genetics 147:1445–1457

Hurme P, Sillanpää MJ, Arjas E, Repo T, Savolainen O (2000) Genetic basis of climatic adaptation in Scots pine by Bayesian quantitative trait locus analysis. Genetics 156:1309–1322

Jansen RC (1992) A general mixture model for mapping quantitative trait loci by using molecular markers. Theor Appl Genet 85:252–260

Jansen RC (1993) Interval mapping of multiple quantitative trait loci. Genetics 135:205–211

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

Jansen RC (1995) Genetic mapping of quantitative trait loci in plants – a novel statistical approach. PhD thesis, Wageningen, the Netherlands. ISBN 90-73771-14-5

Jansen RC, Stam P (1994) High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136:1447–1455

Jansen RC, Van Ooijen JW, Stam P, Lister C, Dean C (1995) Genotype by environment interaction in genetic mapping of multiple quantitative trait loci. Theor Appl Genet 91:33–37

Kao C-H, Zeng Z-B (1997) General formulas for obtaining the MLEs and the asymptotic variance-covariance matrix in mapping quantitative trait loci when using the EM algorithm. Biometrics 53:653–665

Kao C-H, Zeng Z–B, Teasdale RD (1999) Multiple interval mapping for quantitative trait loci. Genetics 152:1203–1216

Kearsey JJ, Farquhar GL (1998) Short review: QTL analysis in plants; where are we now? Heredity 80:137–142

King GJ (1996) Progress of apple genetic mapping in Europe. HortScience 31:1108–1111

King GJ, Alston FH, Batlle I, Chevreau E, Gessler C, Janse J, Lindhout P, Manganaris AG, Sansavini S, Schmidt H, Tobutt KR (1991) The 'European apple genome mapping project' – developing a strategy for mapping genes coding for agronomic characters in tree species. Euphytica 56: 89–94

King GJ, Maliepaard C, Lynn JR, Alston FH, Durel CE, Evans KM, Griffon B, Laurens F, Manganaris AG, Schrevens E, Tartarini S (2000) Quantitative genetic analysis and comparison of physical and sensory descriptors relating to fruit flesh firmness in apple (*Malus pumila* Mill.). Theor Appl Genet 100:1074–1084

Knott SA, Haley CS (1992) Maximum-likelihood mapping of quantitative trait loci using full-sib families. Genetics 132: 1211–1222

Korol AB, Ronin YI, Nevo E (1998) Approximate analysis of QTL–environment interaction with no limits on the number of environments. Genetics 148:2015–2028

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

Lee JK, Thomas DC (2000) Performance of Markov Chain-Monte Carlo approaches for mapping genes in oligogenic models with an unknown number of loci. Am J Hum Genet 67: 1232–1250

Magness JR, Taylor GF (1925) An improved type of pressure tester for the determination of fruit maturity. USDA Dept Circ No. 350. USA, Washington D.C. US

Maliepaard C, Alston FH, Van Arkel G, Brown LM, Chevreau E, Dunemann F, Evans KM, Gardiner S, Guilford P, Van Heusden AW, Janse J, Laurens F, Lynn JR, Manganaris AG, Den Nijs APM, Periam N, Rikkerink E, Roche P, Ryder C, Sansavini S, Schmidt H, Tartarini S, Verhaegh, JJ, Vrielink-Van Ginkel M, King GJ (1998) Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. Theor Appl Genet 97:60–73

Maliepaard C, Van Ooijen JW (1994) QTL mapping in a full-sib family of an outcrossing species. In: van Ooijen JW, Jansen J (eds.) Biometrics in plant breeding: applications of molecular markers. Proc of 9th Meet EUCARPIA Section Biometrics Plant Breed. PUDOC, Wageningen, the Netherlands, pp 140–146

Manly KF, Olson JM (1999) Overview of QTL mapping software and introduction to Map Manager QTL. Mammalian Genome 10:327–334

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

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. J Chem Phys 21:1087–1092

Royall R (1997) Statistical evidence: a likelihood paradigm. Chapman & Hall, London

Satagopan JM, Yandell BS (1996) Estimating the number of quantitative trait loci via Bayesian model determination. (Special Contributed Paper Session Genet Anal of Quantit Traits Complex Dis. Biometric Sect, Joint Stat Meet, Chicago, Ill.) (available at ftp://ftp.stat.wisc.edu/pub/yandell/revjump.html)

Satagopan JM, Yandell BS, Newton MA, Osborn TC (1996) A Bayesian approach to detect quantitative trait loci using Markov Chain Monte Carlo. Genetics 144:805–816

Scheler P, Mangin B, Goffinet B, Le Roy P, Boichard D, Elsen JM (1998) Properties of a Bayesian approach to detect QTL compared to the flanking markers regression method. J Anim Breed Genet 115:87–95

Shoemaker JS, Painter IS, Weir BS (1999) Bayesian statistics in genetics. A guide for the uninitiated. Trends in Genetics 15:354–358

Sillanpää MJ (1999) Bayesian QTL mapping in inbred and outbred experimental designs. PhD thesis, Helsinki Finland. ISBN 952-9528-56-6 (electronically available at http://www.rni.helsinki.fi/~mjs)

Sillanpää MJ, Arjas E (1998) Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. Genetics 148:1373–1388

Sillanpää MJ, Arjas E (1999) Bayesian mapping of multiple quantitative trait loci from incomplete outbred offspring data. Genetics 151:1605–1619

Stam P, Van Ooijen JW (1995) JoinMap™ version 2.0: software for the calculation of genetic linkage maps. Plant Research International, Wageningen

Stephens DA, Fisch RD (1998) Bayesian analysis of quantitative trait locus data using Reversible Jump Markov Chain Monte Carlo. Biometrics 54:1334–1347

Tanksley SD (1993) Mapping polygenes. Annu Rev Genet 27:205–233

Thaller G, Hoeschele I (1996) A Monte Carlo method for Bayesian analysis of linkage between single markers and quantitative trait loci: I. Methodology. Theor Appl Genet 93: 1161–1166

Thompson EA, Heath SC (1999) Estimation of conditional multilocus gene identity among relatives. In: Sellier-Moiseiwitsch F (ed) Statistics in molecular biology and genetics. IMS Lecture Notes, Institute of Mathematical Statistics, American Mathematical Society, Hayward, CA, pp 95–113

Tinker NA, Mather DE (1995) Methods for QTL analysis with progeny replicated in multiple environments. J Agric Genom 1:1 (http://www.ncgr.org/jag/)

Uimari P, Thaller G, Hoeschele I (1996a) The use of multiple markers in a Bayesian method for mapping quantitative trait loci. Genetics 143:1831–1842

Uimari P, Zhang Q, Grignola F, Hoeschele I, Thaller G (1996b) Analysis of QTL workshop. I. Granddaughter design data using least-squares, residual maximum likelihood and Bayesian methods. J Agric Genom 2:7 (http://www.ncgr.org/jag/)

Uimari P, Hoeschele I (1997) Mapping linked quantitative trait loci using Bayesian analysis and Markov chain Monte Carlo algorithms. Genetics 146:735–743

Uimari P, Sillanpää MJ (2001) A Bayesian oligogenic analysis of quantitative and qualitative traits in general pedigrees. Genet Epidemiol (in press)

Van Ooijen JW (1992) Accuracy of mapping quantitative trait loci in autogamous species. Theor Appl Genet 84:803–811

Van Ooijen JW (1999) LOD significance thresholds for QTL analysis in experimental populations of diploid species. Heredity 83:613–624

Van Ooijen JW, Boer MP, Jansen RC, Maliepaard C (2000) MapQTL™ version 4.0: software for the calculation of QTL positions on genetic maps. Plant Research International, Wageningen

Vieland V (1998) Bayesian linkage analysis, or: How I learned to stop worrying and love the posterior probability of linkage. Am J Hum Genet 63:947–954

Vieland VJ, Hodge SE (1998) Book Reviews. Am J Hum Genet 63:283–289

Visscher PM, Thompson R, Haley CS (1996) Confidence intervals in QTL mapping by bootstrapping. Genetics 143:1013–1020

Waagepetersen R, Sorensen D (2001) A tutorial on reversible jump MCMC with a view toward applications in QTL-mapping. Int Stat Rev 69:49–62

Walling GA, Visscher PM, Haley CS (1998) A comparison of bootstrap methods to construct confidence intervals in QTL mapping. Genet Res 71:171–180

Xu S, Atchley WR (1995) A random model approach to interval mapping of quantitative trait loci. Genetics 141: 1189–1197.

Yi N, Xu S (2000a) Bayesian mapping of quantitative trait loci for complex binary traits. Genetics 155:1391–1403

Yi N, Xu S (2000b) Bayesian mapping of quantitative trait loci under the identity-by-descent-based variance component model. Genetics 156:411–422

Zeng Z-B (1993) Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proc Natl Acad Sci USA 90:10972–10976

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

Zeng Z-B, Kao C-H, Basten CJ (1999) Estimating the genetic architecture of quantitative traits. Genet Res 74:279–289