

# Simultaneous fine mapping of closely linked epistatic quantitative trait loci using combined linkage disequilibrium and linkage with a general pedigree

Sang Hong LEE<sup>\*,\*\*</sup>, Julius H.J. VAN DER WERF

The Institute for Genetics and Bioinformatics, School of Rural Science and Agriculture,  
University of New England, Armidale, NSW 2351, Australia

(Received 21 June 2007; accepted 21 September 2007)

**Abstract** – Causal mutations and their intra- and inter-locus interactions play a critical role in complex trait variation. It is often not easy to detect epistatic quantitative trait loci (QTL) due to complicated population structure requirements for detecting epistatic effects in linkage analysis studies and due to main effects often being hidden by interaction effects. Mapping their positions is even harder when they are closely linked. The data structure requirement may be overcome when information on linkage disequilibrium is used. We present an approach using a mixed linear model nested in an empirical Bayesian approach, which simultaneously takes into account additive, dominance and epistatic effects due to multiple QTL. The covariance structure used in the mixed linear model is based on combined linkage disequilibrium and linkage information. In a simulation study where there are complex epistatic interactions between QTL, it is possible to simultaneously map interacting QTL into a small region using the proposed approach. The estimated variance components are accurate and less biased with the proposed approach compared with traditional models.

**fine-mapping / multiple QTL / epistasis / dominance / reversible jump MCMC**

## 1. INTRODUCTION

Phenotypic variation in complex traits may involve the action of many causal genes and their intra- (dominance) and inter-locus interaction (epistasis), in addition to environmental factors. Multiple interacting QTL may play a critical role in quantitative trait variation and epistatic interaction can exist between closely linked quantitative trait loci [5, 11, 15, 28, 33]. Ignoring such non-additive effects due to gene interaction may result in biased estimation

\* Corresponding author: slee38@une.edu.au

\*\* Current address: National Institute of Animal Science, Cheon An, 330-801, Korea.

of QTL position and effects. Interacting QTL showing negligible main effects may not be detected. If the interacting QTL are closely linked, it might be even harder to distinguish their positions.

Within a small region ( $< 10$  cM), closely linked QTL having complicated epistatic interactions may not be detected by conventional linkage QTL mapping methods [23]. Although experimental populations can overcome this by appropriate and complex experimental designs [15,33], natural or outbred populations may present difficulties in the development of such designs. For natural and outbred populations, variance component approaches based on linkage information have been widely used because of their generality and flexibility [2, 3, 26, 30]. In these approaches, additive and non-additive effects including epistasis are treated as random genetic components, which can deal with complex covariance structures based on general pedigrees. However, the use of linkage information alone may limit the power to detect epistasis because estimation of some epistatic components of variance requires many full sibs whose numbers are often limited in outbred and natural populations [30].

Identity by descent (IBD) probabilities between haplotypes of unrelated founders in a recorded pedigree can be estimated on the basis of linkage disequilibrium (LD) using coalescence methods [25, 27, 29]. The LD-based IBD probabilities can give useful extra information about covariance due to additive genetic and dominance effects even without many full sibs as shown by Lee and van der Werf [21]. The use of LD-based IBD probabilities would also be useful for detecting and positioning epistatic QTL. The aim of this study was to investigate how much the mapping resolution improves when considering epistasis in fine mapping of a complex trait and how well epistatic effects can be estimated. The posterior QTL density is estimated with and without considering epistasis using an empirical Bayesian approach based on combined LD and linkage (LDL) information [20].

## 2. MATERIALS AND METHODS

### 2.1. Mixed linear model

A vector of phenotypic observations is written as a linear function of fixed effects, a polygenic term representing the sum of other unidentified additive genetic effects, the additive and dominance effects due to  $n$  QTL, epistatic interaction among the QTL, and residuals. The model can be written as [8],

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \sum_{i=1}^n (\mathbf{Z}\mathbf{a}_i + \mathbf{Z}\mathbf{d}_i) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n (\mathbf{Z}\mathbf{a}_i\mathbf{a}_j + \mathbf{Z}\mathbf{a}_i\mathbf{d}_j + \mathbf{Z}\mathbf{d}_i\mathbf{a}_j + \mathbf{Z}\mathbf{d}_i\mathbf{d}_j) + \mathbf{e} \quad (1)$$

where  $\mathbf{y}$  is a vector of  $N_r$  observations on the trait of interest,  $\boldsymbol{\beta}$  is a vector of fixed effects,  $\mathbf{u}$  is a vector of  $N$  random polygenic effects for each animal,  $\mathbf{a}_i$  and  $\mathbf{d}_i$  are vectors of  $N$  additive and non-additive random effects due to the  $i$ th putative QTL,  $\mathbf{a}_i\mathbf{a}_j$ ,  $\mathbf{a}_i\mathbf{d}_j$ ,  $\mathbf{d}_i\mathbf{a}_j$  and  $\mathbf{d}_i\mathbf{d}_j$  are vectors of epistatic interactions between the  $i$ th and  $j$ th putative QTL, and  $\mathbf{e}$  are residuals. The random effects ( $\mathbf{u}$ ,  $\mathbf{q}_i$ ,  $\mathbf{d}_i$ ,  $\mathbf{a}_i\mathbf{a}_j$ ,  $\mathbf{a}_i\mathbf{d}_j$ ,  $\mathbf{d}_i\mathbf{a}_j$ ,  $\mathbf{d}_i\mathbf{d}_j$  and  $\mathbf{e}$ ) are assumed to be normally distributed with mean zero and variance  $\mathbf{A}\sigma_u^2$ ,  $\mathbf{G}_i\sigma_{q_i}^2$ ,  $\mathbf{D}_i\sigma_{d_i}^2$ ,  $\mathbf{G}_i\mathbf{G}_j\sigma_{a_i a_j}^2$ ,  $\mathbf{G}_i\mathbf{D}_j\sigma_{a_i d_j}^2$ ,  $\mathbf{D}_i\mathbf{G}_j\sigma_{d_i a_j}^2$ ,  $\mathbf{D}_i\mathbf{D}_j\sigma_{d_i d_j}^2$ , and  $\mathbf{I}\sigma_e^2$ , where  $\mathbf{A}$  is a numerator relationship matrix,  $\mathbf{G}_i$  and  $\mathbf{D}_i$  are additive and dominance genotype relationship matrices at the  $i$ th putative QTL position,  $\mathbf{M}_i\mathbf{M}_j$  is the Hadamard product of the matrices  $\mathbf{M}_i$  and  $\mathbf{M}_j$ , and  $\mathbf{I}$  is a  $N_r$ -order identity matrix.  $\mathbf{X}$  and  $\mathbf{Z}$  are incidence matrices (for the effects  $\boldsymbol{\beta}$  and  $\mathbf{u}$ ,  $\mathbf{q}_i$ ,  $\mathbf{d}_i$ ,  $\mathbf{a}_i\mathbf{a}_j$ ,  $\mathbf{a}_i\mathbf{d}_j$ ,  $\mathbf{d}_i\mathbf{a}_j$ , and  $\mathbf{d}_i\mathbf{d}_j$  respectively). The associated variance covariance matrix ( $\mathbf{V}$ ) of all observations given pedigree and marker genotypes is modelled as

$$\mathbf{V} = \mathbf{ZAZ}'\sigma_u^2 + \sum_{i=1}^n (\mathbf{ZG}_i\mathbf{Z}'\sigma_{a_i}^2 + \mathbf{ZD}_i\mathbf{Z}'\sigma_{d_i}^2) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \mathbf{ZG}_i\mathbf{G}_j\mathbf{Z}'\sigma_{a_i a_j}^2 + \mathbf{ZG}_i\mathbf{D}_j\mathbf{Z}'\sigma_{a_i d_j}^2 + \mathbf{ZD}_i\mathbf{G}_j\mathbf{Z}'\sigma_{d_i a_j}^2 + \mathbf{ZD}_i\mathbf{D}_j\mathbf{Z}'\sigma_{d_i d_j}^2 + \mathbf{I}\sigma_e^2. \quad (2)$$

The LDL-based IBD distribution and covariance structure among chromosome segments or haplotypes are accommodated in the matrices  $\mathbf{G}$  and  $\mathbf{D}$  using an approximate coalescence method [24, 25]. For  $\mathbf{G}$  and  $\mathbf{D}$ , a sampler combining the random walk approach [32] and the meiosis Gibbs sampler [35] was used, which is robust and efficient especially for complex pedigree, many markers and missing genotypes [21]. These  $\mathbf{G}$  and  $\mathbf{D}$  were then incorporated as known quantities into the QTL model selection in an empirical Bayesian approach [6, 19].

## 2.2. Reversible jump Markov chain Monte Carlo for simultaneous mapping of multiple interacting QTL

The number of QTL ( $n$ ), the position of each QTL ( $\rho_i$ ,  $i = 1 \sim n$ ) and the model parameters ( $\Theta = \{\sigma_u^2, (\sigma_{a_i}^2, \sigma_{d_i}^2; i = 1 \sim n), (\sigma_{a_i a_{i+1}}^2, \dots, \sigma_{a_i a_n}^2, \sigma_{a_i d_{i+1}}^2, \dots, \sigma_{a_i d_n}^2, \sigma_{d_i a_{i+1}}^2, \dots, \sigma_{d_i a_n}^2, \sigma_{d_i d_{i+1}}^2, \dots, \sigma_{d_i d_n}^2; i = 1 \sim n - 1), \sigma_e^2\}$ ) are to be estimated for the model (1). Note that  $n$  ranges from 0 to the number of marker brackets since only the middle point of each marker bracket was investigated. The probability of estimated parameters given observed phenotypes is

$$pr(n, \rho, \Theta|y) = \frac{pr(y|n, \rho, \Theta)pr(n, \rho, \Theta)}{\sum pr(y|n, \rho, \Theta)pr(n, \rho, \Theta)} \quad (3)$$

**Table I.** Effects of genotypes at 2 QTL with additive model and three interaction models.

	Genotypes for a pair of QTL ( <i>A</i> and <i>B</i> )								
	<i>AA/BB</i>	<i>AA/Bb</i>	<i>AA/bb</i>	<i>Aa/BB</i>	<i>Aa/Bb</i>	<i>Aa/bb</i>	<i>aa/BB</i>	<i>aa/Bb</i>	<i>aa/bb</i>
Additive model	14	12	10	9	7	5	4	2	0
Interaction model 1	20	10	5	10	15	10	5	10	20
Interaction model 2	10	10	20	10	10	20	10	10	0
Interaction model 3	8	8	4	8	20	4	4	4	0

where  $pr(y|n, \rho, \Theta)$  is the likelihood of the observed phenotypes given the estimated parameters,  $pr(n, \rho, \Theta)$  is the joint prior probability of the estimated parameters and the denominator is summed over the probabilities of all possible parameter states. An efficient approach to estimate  $n, \rho$  and  $\Theta$  with additive and dominance QTL effects was shown by [19]. In the process, the QTL model is first defined by the number of QTL and their positions, which are sampled from a proposal distribution. In a second step, residual maximum likelihood (REML) estimates for the model parameters are obtained for a given QTL model. The proposed variables and model parameters are accepted or rejected, according to the acceptance ratio from a reversible jump (RJ) Markov chain Monte Carlo (MCMC) from which the posterior QTL density is derived. Hence, a REML procedure is used nested within a Bayesian RJMCMC [19].

When considering epistatic interaction among QTL, the number of random effects increases, *e.g.* the number of random effects due to  $n$  QTL is  $n$  for the additive model,  $2n$  for the additive and dominance model and  $2n + \{n(n - 1)/2\} \times 4$  for the full model (see (1)). However, not all random effects due to epistasis are significant in the model. Including many non-significant random effects that do not improve the model likelihood causes numerical problems due to over-parameterisation. Therefore, within a RJMCMC step for a given set of QTL, each epistatic component is tested and if the likelihood is not improved, the epistatic component will be removed from the model.

### 2.3. Simulation study

#### 2.3.1. Genetic interaction model

When there are intra-locus interactions within QTL and inter-locus interactions between QTL, each of nine possible genotypes has its own value due to additive and dominance effects and epistatic interaction (Tab. I). In the additive model, allele substitution effects for the favourable allele of the first QTL

and the second QTL were 5 and 2. In the interaction model 1, there were intra- and inter-locus interactions such that some combination of genotypes that were homozygous for both QTL expressed enhanced performance. The interaction model 2 showed complete dominance only when the second QTL was recessive homozygous. In interaction model 3, the combination of genotypes that were heterozygous for both QTL showed enhanced performance. The patterns of interactions for models 1, 2 and 3 have been named the co-adaptive, the dominant, and the dominance-by-dominance epistasis model [4].

### 2.3.2. Simulated data

One hundred generations of a historical population with effective size of 100 were simulated for 26 markers and 4 potential QTL in a 130 cM region. For the region from 10 to 20 cM and 110 to 120 cM as candidate regions, markers were densely positioned at 1 cM intervals. The potential QTL were simulated at 12.5 cM (QTL I), 17.5 cM (QTL II), 112.5 cM (QTL III) and 117.5 cM (QTL IV). Unique numbers were assigned to founder gametes at the QTL in generation 0. In each generation, the number of male and female parents was 50 and their alleles were transmitted to descendants on the basis of Mendelian segregation using the gene-dropping method [22]. Due to genetic drift, the number of alleles at the QTL was reduced. One allele with moderate frequency (0.1 ~ 0.9) was randomly chosen to be the mutation in generation 100 [24]. The number of alleles assumed at each marker locus was four in generation 0 and starting allele frequencies were all at 0.25. The marker allele was mutated at a rate of  $4 \times 10^{-4}$  per generation [9, 10, 37], *i.e.* a new allele was introduced as a mutation. Therefore, this historical population would generate LD among closely linked regions with the random genetic drift and mutation. Note that pedigree and genotype information was deemed not available for these 100 generations. In generation 100 and afterwards, phenotypic values for individuals were simulated as

$$y = \mu + G_{QTL_{II,IV}} + e \quad (\text{for interaction between QTL II and IV}) \quad (4)$$

$$y = \mu + G_{QTL_{I,II}} + G_{QTL_{III,IV}} + e \quad (\text{for interactions between QTL I and II, and QTL III and IV}). \quad (5)$$

The population mean ( $\mu$ ) was 100, values for residuals ( $e$ ) were from  $N(0, \sigma_e^2)$  with  $\sigma_e^2 = 50$ . No polygenic effects were simulated. The QTL genotypes ( $G_{QTL}$ ) for each individual were made up from each pair of QTL. Two interacting QTL that are not linked were investigated with the additive model

and interaction models 1, 2 and 3 using (4). In addition, two pairs of interacting QTL that are closely linked were investigated with the interaction model 1 using (5).

For QTL mapping results, the posterior QTL density over 20 replicates was estimated in RJMCMC LDL mapping with additive effects only (additive model), additive and dominance effects only (additive and dominance model), or additive, dominance and epistasis effects (full model). In all cases, marker genotypes and phenotypes were available for the last two generations (200 animals) used for analyses.

### 3. RESULTS

#### 3.1. The posterior QTL density

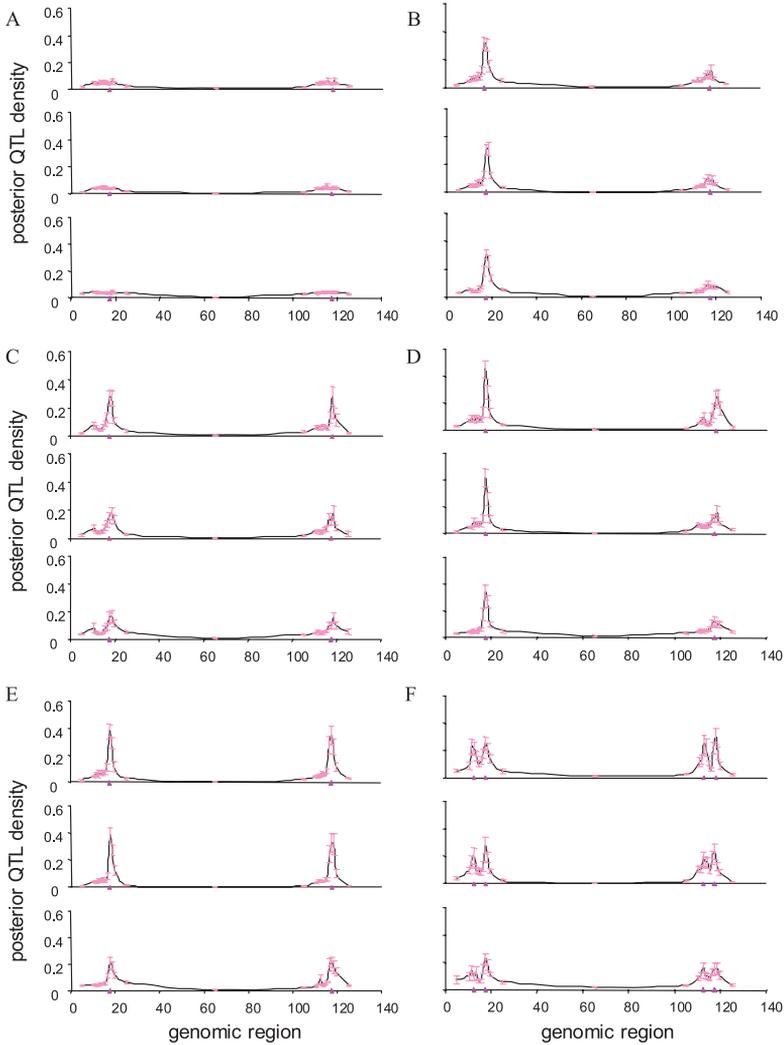
With no effects simulated for additive, dominance and epistatic QTL, the posterior QTL density with the full model is fairly flat, although the regions of densely spaced markers show some slight elevations. The difference of the curve between the full model and other models is negligible (Fig. 1A). With two additive QTL (QTL II and QTL IV), the posterior QTL densities are similar across the three different models of analyses (Fig. 1B). The results show that the QTL density profiles of the model including dominance and epistasis are the same with the additive model when dominance and epistasis are absent (no spurious peaks were found).

When two QTL (QTL II and QTL IV) have a complex interaction (co-adaptive, dominant or dominance-by-dominance epistasis), the full model with additive, dominance and epistasis gives a higher mapping resolution than any other model (Figs. 1 C, D and E). This shows that the full model is more accurate for mapping interacting QTL.

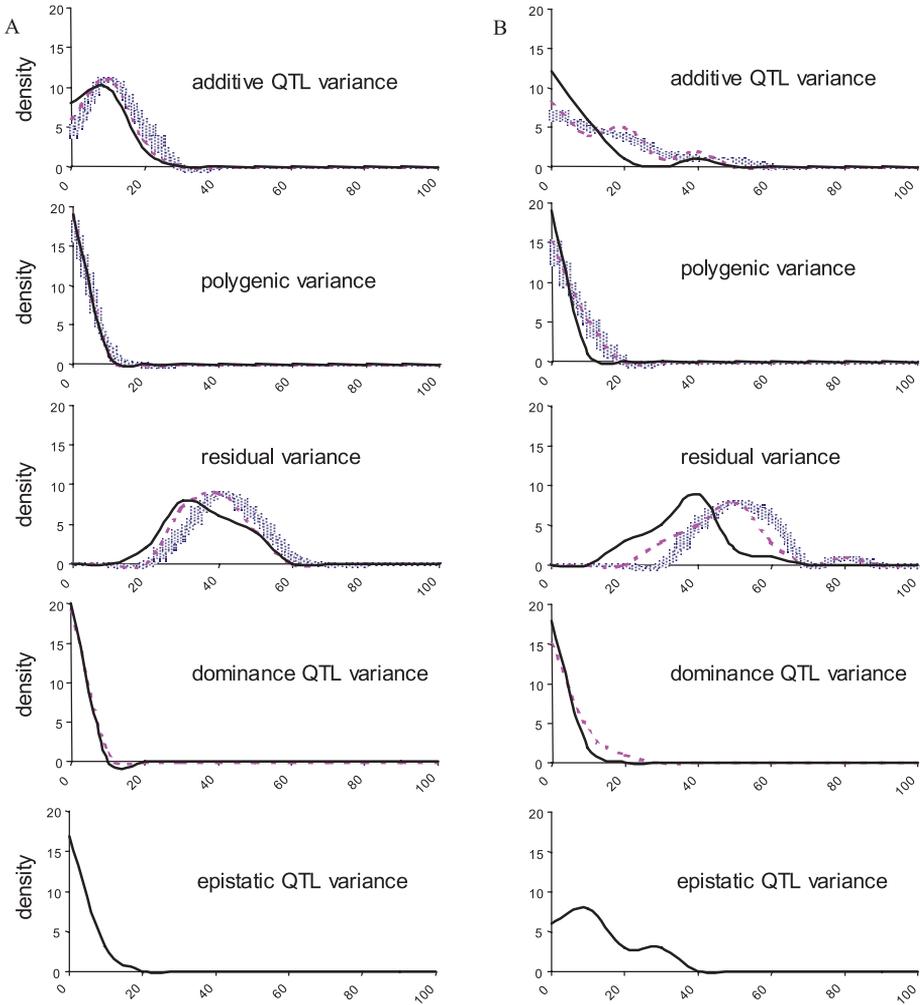
When two sets of two closely linked QTL have a complex interaction (co-adaptive), the posterior QTL density with the full model gives a higher mapping resolution compared to other models (Fig. 1F). The posterior QTL density with the full model is clearly peaked at the true QTL positions, and two closely linked QTL are clearly distinguished. However, the models without epistasis give less accuracy.

#### 3.2. Estimated variance components

Figure 2 shows the histogram of estimated variances from 20 replicates with the full (solid line), additive and dominance (dotted line), and additive model



**Figure 1.** Posterior QTL density with full model (upper), additive and dominance model (middle), and additive model (lower) when additive, dominance and epistasis effects are null (A); QTL II and QTL IV have additive effects without dominance and epistasis effects (B); QTL II and QTL IV have a complex interaction as in the interaction model 1 (co-adaptive epistasis) (C); QTL II and QTL IV have a complex interaction as in the interaction model 2 (dominant epistasis) (D); QTL II and QTL IV have a complex interaction as in the interaction model 3 (dominance-by-dominance epistasis) (E); and the closely linked QTL I and QTL II, and QTL III and QTL IV have a complex interaction of the interaction model 1 (co-adaptive epistasis) (F). Triangle shows the true QTL positions. The shaded vertical bars indicate empirical standard error.



**Figure 2.** Histogram of estimated variances for 20 replicates with full model (solid line), additive and dominance model (dotted line) and additive model (shaded line) when QTL II and QTL IV have additive effects without dominance and epistasis effects (A); and when QTL II and QTL IV have a complex interaction with the interaction model 1 (co-adaptive epistasis) (B). In the case of A, the average of the expected values is 11.36 (standard error = 0.6) for additive QTL variance, 0 (0) for polygenic variance, 50 (0) for residual variance. In the case of B, the average of the expected values is 9.82 (1.38) for additive QTL variance, 0 (0) for polygenic variance, 50 (0) for residual variance, 0.25 (0.05) for dominance variance and 12.91 (0.78) for epistatic variance.

(shaded line) when using two additive QTL (Fig. 2A) or when using two interacting QTL with the interaction model 1 (co-adaptive epistasis) (Fig. 2B). Each estimated value is the average of sampled values of all RJMCMC rounds in each replicate. When there are only additive effects without dominance and epistasis, not much difference is found across three models. The distribution of estimated variance with full, additive and dominance, or additive model shows the high density around the expected value ( $11.36 \pm 0.6$  (standard error) for additive QTL variance and 50 (standard error = 0) for residual variance) (Fig. 2A). Since polygenic, dominance and epistatic effects are null, the distribution of estimated variances for them are close to zero with all models including the full model. When epistatic interactions are involved in addition to additive effects, the distributions of estimated variances are different across the models. When using the full model, the distribution of estimated additive QTL variance coincides with the expected value ( $9.82 \pm 1.38$ ). However, the estimation is upwardly biased with the additive and dominance model, and more biased with the additive model. Since polygenic variance was simulated as zero, the estimated variance components were close to zero with all models although they were overestimated more frequently with the additive and dominance model, or the additive model, compared with the full model. The distribution of estimated residual variance coincides with the expected value of 50 although the estimation is underestimated with the full model and overestimated with the other models. The expected value for dominance variance is negligible ( $0.25 \pm 0.05$ ). Therefore, the estimated variance components are distributed around zero although overestimation is shown more frequently with the additive and dominance model than the full model. For estimated epistasis variance with the full model, the mode of distribution coincides with the expected value ( $12.91 \pm 0.78$ ). The results with other scenarios for dominant epistasis or dominance-by-dominance epistasis are similar in that estimated QTL variances with the full model are more accurate than those with other reduced models (result not shown).

#### 4. DISCUSSION

When there were complex epistatic interactions between QTL, the mapping resolution with the full model considering epistasis considerably increased, compared to that with the reduced models. We investigated four different scenarios for epistatic QTL with 20 replicates each, and in 49% (29% or 23%) of these cases, the full model (additive and dominance model, or additive model) gave a higher posterior QTL density than the other models. It was also shown

that considering epistasis helped to simultaneously map closely linked interacting QTL into a fine region. Moreover, the estimated variance components were shown to be biased and less accurate with the additive model, or the additive and dominance model when epistatic interactions exist. This could be remedied when using the full model including epistasis components.

Purcell and Sham [30] advocated that the power and accuracy of detecting epistatic QTL are lower than those for additive QTL. This is probably due to the fact that information used to estimate all four epistatic components of variance in natural (human) populations is not sufficient (*e.g.* limited number of full sibs). Moreover, when using linkage information only, there is little information about segregation patterns in a small region (few recombination events). Therefore, the analysis based on linkage information only cannot identify any QTL even with additive effects [19]. This situation would be worse for QTL with non-additive effects. It is noted that the dominance relationship matrix based on linkage information only is often non-positive definite due to lack of information when using a general pedigree of two generations. However, when using additional LD information, the situation is much improved as was already shown for dominance [20]. The present study shows that the accuracy to detect and identify epistatic QTL is increased because of using LD information in addition to linkage information.

When a large number of parameters are included in a statistical model (*e.g.* the full model), it is usually a concern that spurious signals for ghost QTL can be generated. We tested if spurious QTL peaks were generated when all QTL effects (additive, dominance and epistatic effects) were zero (Fig. 1A), or when there were additive effects but dominance and epistatic effects were zero (Fig. 1B). It was shown that analysis with the full model would be reasonably robust to false positives.

The prior probability for the QTL number in the RJMCMC was drawn from a Poisson distribution with a mean of 1. This would decrease the frequency of including less significant QTL in the model especially when using a whole genome approach (many polygenic terms having small effects). Although several studies [13, 31, 38] showed that estimates of QTL number are robust against prior assumptions, it may be possible and useful to obtain more informative prior information from the previous studies for the genome (*e.g.* meta analyses).

In using LD information, the levels of LD and its distribution in the population are important issues. Several studies have shown that power and precision of fine mapping of QTL are closely related to the levels of LD [17, 19]. Sved [34] showed that the smaller segments have stronger LD which is more

useful in LD mapping [16]. Goddard *et al.* [12] reported that even between different breeds, LD information can be utilised when marker spacing is less than 50 kb. As more markers become available (*e.g.* > 50 000 SNP), the present method based on LD information can be a powerful tool for gene mapping.

Hadamard products of relationship matrices are widely used for fitting epistatic components in variance component approaches [26, 36] based on Cockerham's model [8]. However, Cockerham's model [8] assumes linkage equilibrium between interacting loci and does not account for LD between closely linked loci which may cause some bias in the estimation of variance components. We explicitly used LD information in our method and observed little evidence of bias. Furthermore, it may be worth considering to implement orthogonal models for epistasis (*e.g.* [1, 14]) in the approach which may further improve the mapping resolution and the accuracy of estimated variances.

The computational effort of running the RJMCMC with the full model is relatively large because all epistatic components have to be considered in the model. Genome wide scans would be difficult with a very large pedigree and a very large number of markers. In the case of dense markers (with tens of thousands of SNP), simple methods such as single marker regression with RJMCMC can be used because there is a high chance that a single marker can be completely associated with a causal mutation. It will also be possible to detect interactions between QTL using simple regression methods. However, with the regression method, simplifying assumptions have to be made when there are missing genotypes, and covariance among haplotypes and useful LD information in complex pedigree is likely ignored. Therefore, it seems useful to use our proposed method for analysis of putative QTL regions arising from whole genome scans.

The number of records used in this study was relatively small ( $N_r = 200$ ). This might cause inaccurate estimation of variance components with the full model especially for residuals although the mode of distribution coincided with the expected value (Fig. 2). Although we did not formally assess the power of detecting epistatic QTL, we observed that the mapping resolution was reasonably high. The posterior QTL density profile was highly peaked on each true position, and the peak was higher with the full model. When using QTL II and QTL IV with the interaction model 1, the average ratio of additive variance over phenotypic variance was  $\sim 0.07$  for each QTL, and epistatic variance over phenotypic variance was  $\sim 0.18$ . When using two pairs of QTL with the interaction model 1 (interaction between QTL I and QTL II, and interaction between QTL III and QTL IV), the average ratio of additive variance over phenotypic variance was  $\sim 0.06$  for each QTL, and epistatic variance over

phenotypic variance was  $\sim 0.13$  for each pair of QTL. Hence, the QTL effects in our simulation study may have been relatively large. With a higher number of records, it would be possible to identify QTL with smaller effects and obtain more accurate estimated variances.

Generally, variance component approaches are known to be more robust than regression-based methods or maximum likelihood methods [7]. They can handle a large number of missing genotypes, complex relationships and multi-allelic markers. Moreover, the present method combined with LDL information can efficiently detect QTL and their intra- and inter-locus interactions, making it possible to simultaneously fine-map main and epistatic QTL. With increasing density of genetic markers, the present approach would be an efficient tool to explore more complex genetic models for quantitative traits.

## ACKNOWLEDGEMENTS

This study was supported by PIC International and Sheep Genomics Australia.

## REFERENCES

- [1] Alvarez-Castro J.M., Carlborg O., A unified model for functional and statistical epistasis and its application in quantitative trait loci analysis, *Genetics* 176 (2007) 1151–1167.
- [2] Blangero J., Williams J.T., Almasy L., Quantitative trait locus mapping using human pedigrees, *Hum. Biol.* 72 (2000) 35–62.
- [3] Blangero J., Williams J.T., Almasy L., Variance component methods for detecting complex trait loci, *Adv. Genet.* 42 (2001) 151–181.
- [4] Carlborg O., Haley C.S., Epistasis: Too often neglected in complex trait studies? *Nat. Rev. Genet.* 5 (2004) 618–625.
- [5] Caron H., van Schaik B., van der Mee M., Baas F., Riggins G., van Sluis P., Hermus M.-C., van Asperen R., Boon K., Voûte P.A., Heisterkamp S., van Kampen A., Versteeg R., The human transcriptome map: Clustering of highly expressed genes in chromosomal domains, *Science* 291 (2001) 1289–1292.
- [6] Casella G., Empirical Bayes Gibbs sampling, *Biostatistics* 2 (2001) 485–500.
- [7] Cherny S.S., Sham P.C., Cardon L.R., Introduction to the special issue on variance components methods for mapping quantitative trait loci, *Behav. Genet.* 34 (2004) 125–126.
- [8] Cockerham C.C., An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present, *Genetics* 39 (1954) 859–882.

- [9] Dallas J.F., Estimation of microsatellite mutation rates in recombinant inbred strains of mouse, *Mamm. Genome* 3 (1992) 452–456.
- [10] Ellegren H., Mutation rates at porcine microsatellite loci, *Mamm. Genome* 6 (1995) 376–377.
- [11] Georges M., Coppieters W., Charlier C., Polymorphic miRNA-mediated gene regulation: Contribution to phenotypic variation and disease, *Curr. Opin. Genet. Dev.* 17 (2007) 166–176.
- [12] Goddard M.E., Hayes B.J., McPartlan H.C., Chamberlain A.J., Can the same genetic markers be used in multiple breeds? in: *Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, 2006, Belo Horizonte, Brazil*, pp. (paper number) 22–16.
- [13] Jannink J.-L., Fernando R.L., On the Metropolis-Hastings acceptance probability to add or drop a quantitative trait locus in Markov chain Monte Carlo-based Bayesian analyses, *Genetics* 166 (2004) 641–643.
- [14] Kao C.-H., Zeng Z.-B., Modeling epistasis of quantitative trait loci using Cockerham's model, *Genetics* 160 (2002) 1243–1261.
- [15] Kroymann J., Mitchell-Olds T., Epistasis and balanced polymorphism influencing complex trait variation, *Nature* 435 (2005) 95–98.
- [16] Kruglyak L., Prospects for whole-genome linkage disequilibrium mapping of common disease genes, *Nat. Genet.* 22 (1999) 139–144.
- [17] Lee S.H., van der Werf J.H.J., The efficiency of designs for fine-mapping of quantitative trait loci using combined linkage disequilibrium and linkage, *Genet. Sel. Evol.* 36 (2004) 145–161.
- [18] Lee S.H., van der Werf J.H.J., The role of pedigree information in combined linkage disequilibrium and linkage mapping of quantitative trait loci in a general complex pedigree, *Genetics* 169 (2005) 455–466.
- [19] Lee S.H., van der Werf J.H.J., Simultaneous fine mapping of multiple closely linked quantitative trait loci using combined linkage disequilibrium and linkage with a general pedigree, *Genetics* 173 (2006) 2329–2337.
- [20] Lee S.H., van der Werf J.H.J., Using dominance relationship coefficients based on linkage disequilibrium and linkage with a general complex pedigree to increase mapping resolution, *Genetics* 174 (2006) 1009–1016.
- [21] Lee S.H., van der Werf J.H., Tier B., Combining the meiosis Gibbs sampler with the random walk approach for linkage and association studies with a general complex pedigree and multi marker loci, *Genetics* 171 (2005) 2063–2072.
- [22] MacCluer J.W., VandeBerg J.L., Read B., Ryder O.A., Pedigree analysis by computer simulation, *Zoo Biology* 5 (1986) 147–160.
- [23] Maloof J.N., Quantitative genetics: Small but not forgotten, *Heredity* 96 (2006) 1–2.
- [24] Meuwissen T.H.E., Goddard M.E., Fine mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci, *Genetics* 155 (2000) 421–430.
- [25] Meuwissen T.H.E., Goddard M.E., Prediction of identity by descent probabilities from marker-haplotypes, *Genet. Sel. Evol.* 33 (2001) 605–634.

- [26] Mitchell B.D., Ghosh S., Schneider J.L., Birznieks G., Blangero J., Power of variance component linkage analysis to detect epistasis, *Genet. Epidemiol.* 14 (1997) 1017–1022.
- [27] Morris A.P., Whittaker J.C., Balding D.J., Little loss of information due to unknown phase for fine-scale linkage disequilibrium mapping with single-nucleotide-polymorphism genotype data, *Am. J. Hum. Genet.* 74 (2004) 945–953.
- [28] Orgogozo V., Broman K.W., Stern D.L., High-resolution quantitative trait locus mapping reveals sign epistasis controlling ovariole number between two drosophila species, *Genetics* 173 (2006) 197–205.
- [29] Perez-Enciso M., Fine mapping of complex trait genes combining pedigree and linkage disequilibrium information: A Bayesian unified framework, *Genetics* 163 (2003) 1497–1510.
- [30] Purcell S., Sham P.C., Epistasis in quantitative trait locus linkage analysis, *Behav. Genet.* 34 (2004) 143–152.
- [31] Sillanpää M.J., Gasbarra D., Arjas E., Comment on "On the Metropolis-Hastings acceptance probability to add or drop a quantitative trait locus in Markov chain Monte Carlo-based Bayesian analyses", *Genetics* 167 (2004) 1037.
- [32] Sobel E., Lange K., Descent graphs in pedigree analysis: Applications to haplotyping, location scores, and marker-sharing statistics, *Am. J. Hum. Genet.* 58 (1996) 1323–1337.
- [33] Steinmetz L.M., Sinha H., Richards D.R., Spiegelman J.I., Oefner P.J., McCusker J.H., Davis R.W., Dissecting the architecture of a quantitative trait locus in yeast, *Nature* 416 (2002) 326–330.
- [34] Sved J.A., Linkage disequilibrium and homozygosity of chromosome segments in finite populations, *Theor. Popul. Biol.* 2 (1971) 125–141.
- [35] Thompson E.A., Heath S.C., Estimation of conditional multilocus gene identity among relatives, in: Seillier-Moiseiwitsch F. (Ed.), *Statistics in molecular biology and genetics*, IMS lecture notes, Institute of Mathematical Statistics, American Mathematical Society, Providence, RI, 1999, pp. 95–113.
- [36] Wang K., Score tests for epistasis models on quantitative traits using general pedigree data, *Genet. Epidemiol.* 25 (2003) 314–326.
- [37] Weber J.L., Wong C., Mutation of human short tandem repeats, *Hum. Mol. Genet.* 2 (1993) 1123–1128.
- [38] Yi N., Xu S., Allison D.B., Bayesian model choice and search strategies for mapping interacting quantitative trait loci, *Genetics* 165 (2003) 867–883.