

# Evaluation of Cofactor Markers for Controlling Genetic Background Noise in QTL Mapping\*\*

Chaeyoung Lee\* and Xiaolin Wu<sup>1</sup>

Laboratory of Statistical Genetics, Institute of Environment & Life Science, Hallym University  
Chuncheon, Kangwon-do 200-702, Korea

**ABSTRACT :** In order to control the genetic background noise in QTL mapping, cofactor markers were incorporated in single marker analysis (SMACO) and interval mapping (CIM). A simulation was performed to see how effective the cofactors were by the number of QTL, the number and the type of markers, and the marker spacing. The results of QTL mapping for the simulated data showed that the use of cofactors was slightly effective when detecting a single QTL. On the other hand, a considerable improvement was observed when dealing with more than one QTL. Genetic background noise was efficiently absorbed with linked markers rather than unlinked markers. Furthermore, the efficiency was different in QTL mapping depending on the type of linked markers. Well-chosen markers in both SMACO and CIM made the range of linkage position for a significant QTL narrow and the estimates of QTL effects accurate. Generally, 3 to 5 cofactors offered accurate results. Over-fitting was a problem with many regressor variables when the heritability was small. Various marker spacing from 4 to 20 cM did not change greatly the detection of multiple QTLs, but they were less efficient when the marker spacing exceeded 30 cM. Likelihood ratio increased with a large heritability, and the threshold heritability for QTL detection was between 0.30 and 0.05. (*Asian-Aust. J. Anim. Sci.* 2003, Vol 16, No. 4: 473-480)

**Key Words :** Heritability, Marker Spacing, QTL, Simulation

## INTRODUCTION

Quantitative trait loci (QTLs) have been identified and mapped by associating trait phenotypes with multiple markers. Analytical models in conventional methods often included single marker effect or flanking marker effects. The former method was called the single marker analysis (SMA) and the latter was the interval mapping (IM, Lander and Botstein, 1989). However, these methods were likely to render a low resolution of a QTL due to a QTL effect confounded with the effect of another QTL elsewhere in the linkage group (Wright and Kong, 1997). In order to reduce such bias, Zeng (1993) and Jansen (1993) suggested the use of a marker interval and a few other well-chosen single markers simultaneously, and the method was called composite IM (CIM). Employing additional marker loci made the genetic background noise to be absorbed with a narrow marker region for a significant QTL, resulting in a considerably increased QTL resolution (Zeng, 1994). Although CIM has been increasingly used in practical QTL mapping in animals and plants (Boyle and Gill, 2001; Drake et al., 2001; Robison et al., 2001; Wayne et al., 2001;

Wu et al., 2001, 2002), investigations for optimizing the effects of cofactors were still limited (Piepho and Gauch, 2001; Lee, 2002). The effect of CIM largely depended on the choice of cofactors. In this paper, we evaluated incorporating cofactors in both IM and SMA of QTL mapping through simulation. This improved SMA was referred to as the single-marker analysis with cofactors (SMACO).

## MATERIALS AND METHODS

### A general idea on cofactor markers

Suppose a biallelic QTL was located between markers  $i$  and  $i+1$ . There were also some other QTLs elsewhere in the linkage. Assume the putative QTL of interest was  $r_1$  recombination units apart from the marker  $i$  and  $r_2$  apart from the marker  $i+1$ . The trait value was expressed as below:

$$y_k = \mu + f_i(M_i) + \sum_j f_j(C_j) + e_k \quad (1)$$

where  $y_k$  was the observed value for individual  $k$  in the population,  $\mu$  was the mean of observed values,  $f_i(M_i)$  was a function of genotype for marker  $i$  for SMACO or for marker loci  $i$  and  $i+1$  for CIM, and  $e_k$  was the residual.  $f_j(C_j)$  was a function of genotype for a cofactor. Note that  $j \neq i$  for SMACO, and  $j \neq i, i+1$  for CIM.

Under the assumption of independent and identical normal distributions with variance equaled to  $\sigma^2$  for the trait values, the general form of likelihood function was

\*\* This study was supported by a grant from BioGreen 21 Program, Rural Development Administration, Korea and the Foundation of Hunan Life Science Research Center, China.

\* Reprint request to: Chaeyoung Lee. Tel: +82-31-380-1630, Fax: 82-31-388-3427, E-mail: cleef@sun.hallym.ac.kr

<sup>1</sup> Department of Animal Science and Technology, Hunan Agricultural University, Changsha, Hunan 410128, P. R. China.

Received March 5, 2002; Accepted July 22, 2002

$$L = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \prod_{i=1}^N p(Q_i|M_i) \exp\left[-\frac{(y_i - \mu_i)^2}{2\sigma^2}\right] \quad (2)$$

where  $p(Q_i|M_i)$  was the conditional probability of QTL genotype 1 given marker  $M_i$ , and  $\mu_i$  was the mean of trait values with the QTL genotype 1.

The logarithm of Equation (2) was

$$l = \text{Log } L = -\frac{N}{2} \text{Log}(2\pi\sigma^2) - \sum_{i=1}^N \text{Log} \left\{ \sum_{j=1}^3 p(Q_j|M_i) \exp\left[-\frac{(y_i - \mu_j)^2}{2\sigma^2}\right] \right\} \quad (3)$$

Under the null hypothesis  $H_0: \mu_{QQ} = \mu_{Qq} = \mu_{qq} = \mu$ , the log likelihood was

$$l_{\mu} = -\frac{N}{2} \text{Log}(2\pi\sigma^2) - \sum_{i=1}^N \frac{(y_i - \mu)^2}{2\sigma^2} \quad (4)$$

The following likelihood ratio was defined as a test statistics for the hypothesis that a QTL is not linked to the marker (SMACO) or marker interval (CIM):

$$LR = 2[l - l_{\mu}] \quad (5)$$

The corresponding lod score for the null hypothesis test was

$$\text{lod} = \text{Log}_{(10)} L - \text{Log}_{(10)} L_{H_0} \quad (6)$$

The genetic effect of the putative QTL linked to marker(s) was estimated by maximum likelihood.

#### An example of single marker analysis with cofactors

Since Zeng (1993, 1994) described CIM as a combination of simple IM and multiple linear regression, here we show an example for SMACO using backcross progeny. A progeny population was derived from backcrossing the F1 (Aa) with the parent with AA marker genotype. The analytical regression model for SMACO with backcross progeny was

$$y_{ik} = b_0 + b_1 X_{ik} + \sum_{j=2}^m b_j X_{ijk} + e_{ik} \quad (7)$$

where  $b_0$  was the intercept of the model,  $b_1$  was the slope of regression for the putative QTL linked to marker  $i$ ,  $X_{ik}$  was a dummy variable taking 1 for marker genotype AA and -1 for Aa,  $b_j$  was the partial regression coefficient of the observation on marker  $j$ ,  $X_{ijk}$  was the dummy variable for cofactor marker  $j$  of individual  $k$ , taking 1 for the cofactor marker genotype AA and -1 for Aa, and  $e_{ik}$  was the residual.

Assuming that  $e_{ik}$  was normally distributed with mean zero and variance  $\sigma^2$ , the likelihood function for the

SMACO was

$$L = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \prod_{i=1}^N \left[ (1-r) \exp\left\{-\frac{(y_{ik} - \mu_{QQ})^2}{2\sigma^2}\right\} + r \exp\left\{-\frac{(y_{ik} - \mu_{Qq})^2}{2\sigma^2}\right\} \right] \prod_{i=1}^N \left[ r \exp\left\{-\frac{(y_{ik} - \mu_{Qq})^2}{2\sigma^2}\right\} + (1-r) \exp\left\{-\frac{(y_{ik} - \mu_{QQ})^2}{2\sigma^2}\right\} \right] \quad (8)$$

where  $n_1$  and  $n_2$  were the numbers of individuals having marker genotypes AA and Aa, respectively, at the marker locus  $i$  ( $N = n_1 + n_2$ ). Note that the recessive homozygous genotypes (aa) for the marker was missing in the backcross progeny, and  $p(Q_i|aa) = 0$ .

The logarithm of Equation (8) was

$$l = -\frac{N}{2} \text{Log}(2\pi\sigma^2) + \sum_{i=1}^n \log \left[ (1-r) \exp\left\{-\frac{(y_{ik} - \mu_{QQ})^2}{2\sigma^2}\right\} + r \exp\left\{-\frac{(y_{ik} - \mu_{Qq})^2}{2\sigma^2}\right\} \right] + \sum_{i=1}^n \log \left[ r \exp\left\{-\frac{(y_{ik} - \mu_{Qq})^2}{2\sigma^2}\right\} + (1-r) \exp\left\{-\frac{(y_{ik} - \mu_{QQ})^2}{2\sigma^2}\right\} \right] \quad (9)$$

Under the null hypothesis  $H_0: \mu_{QQ} = \mu_{Qq} = \mu$ , the log likelihood was

$$l_{\mu_0} = -\frac{N}{2} \text{Log}(2\pi\sigma^2) - \sum_{i=1}^n \sum_{k=1}^n \frac{(y_{ik} - \mu)^2}{2\sigma^2} \quad (10)$$

Then the likelihood ratio and the lod score for the null hypothesis were obtained as shown in the Equations (5) and (6).

#### Simulation

A simulation study for diploid organisms was performed to investigate the properties and utilities of cofactors. Typical backcross populations were simulated with QTL Cartographer for Windows Version 1.01 (Basten et al., 1994). From two distinct populations, six males from one population and six females from the other population were randomly selected and mated, producing six full-sib families in the F1 generation. Five daughters from each full-sib family were used to cross back to their corresponding fathers. Ten individuals were born from each pair, and the 300 individuals in the backcross generation were genotyped.

A total of six populations were simulated and two chromosomes that were both 120 cM in linkage length were generated for each population. The input values to simulate these populations were presented in Table 1. In all the populations, 13 markers were evenly located on both chromosomes. For population 1, QTL1 was assigned at 67cM of chromosome 1, and QTL2 at 35 cM and QTL4 at 99 cM on chromosome 2. The additive effects for QTL1.

**Table 1.** Input values used in the simulation<sup>1</sup>

Population	QTL additive effect <sup>2</sup>				$\sigma_a^2$ <sup>3</sup>	$\sigma_e^2$	Heritability
	QTL1	QTL2	QTL3	QTL4			
1	1.040	0.956	-	1.174	1.000	1.000	0.500
2	-	0.300	0.212	0.321	0.111	1.000	0.100
3	-	0.575	0.344	0.707	0.429	1.000	0.300
4	-	0.851	0.589	1.040	1.000	1.000	0.500
5	-	1.239	0.909	1.624	2.330	1.000	0.700
6	-	2.194	2.057	3.091	9.000	1.000	0.900

<sup>1</sup> Trait mean was 15.8 units. No dominance or epistasis was considered.

<sup>2</sup> QTL1 was located at 67cM on chromosome 1, and QTL2, QTL3 and QTL4 at 35, 90 and 99cM on chromosome 2.

<sup>3</sup>  $\sigma_a^2$  = Additive variance and  $\sigma_e^2$  = Environmental variance.

QTL2, and QTL4 were 1.040, 0.956, and 1.174, respectively, with heritability of 0.5. Dominance and epistasis were not considered for the QTLs. This population was analyzed using both CIM and SMACO. We examined the effect of using various numbers and different types of cofactors to control genetic background noise. QTLs were searched using SMA with 1 (SMACO\_1), 3 (SMACO\_3), 5 (SMACO\_5), 7 (SMACO\_7), and all other loci (SMACO\_all) as cofactors. Note that SMACO\_all was the multiple marker analysis. The CIM was also used with 1 (CIM\_1), 3 (CIM\_3), 5 (CIM\_5), 7 (CIM\_7), and all other loci (CIM\_all) as cofactors. All the markers on chromosomes 1 and 2 were examined in groups such as unlinked markers (UL), markers linked to nuisance QTL (LNQ), markers not closely linked to nuisance QTL (NCLNQ), and flanking markers (FL). The UL were the markers in a linkage group other than the one where the QTL of interest was located. For example, when QTL2 and QTL4 on chromosome 2 were examined, the markers on chromosome 1 were regarded as unlinked markers. The LNQ referred to the ones located within 25 cM from the nuisance QTL, and those located farther than 25 cM from the QTL were regarded as NCLNQ. The FL were flanking the marker interval examined. Populations 2 to 6 were simulated to illustrate the influence of heritability level on QTL detection. The locations of QTL2 and QTL4 on chromosome 2 in these populations were simulated as in population 1, but QTL1 was not simulated. Additionally, QTL3 was simulated at 90 cM on chromosome 2, so it was only 9 cM apart from QTL4. The additive effects of the QTL2, QTL3, and QTL4 varied in the five populations while the same amount of environmental variance was used. The heritability of the five populations ranged from 0.1 to 0.9.

Additionally, various marker densities were generated for population 4 to investigate the influence of marker spacings on QTL mapping. The numbers of markers used per chromosome were 61, 31, 13, 7, 5, and 4, and they were evenly distributed at the two chromosomes. Their corresponding marker spacings were 2 cM, 4 cM, 10 cM, 20 cM, 30 cM, and 40 cM, respectively.

A total of 50 replicates were simulated for each population. The simulated data were analyzed by the proposed methods, and the calculated LR<sub>s</sub> were compared to genome-wide threshold values at a 0.05 significance level. The threshold values were obtained with 1000 replicates by permutation tests (Churchill and Doerge, 1994).

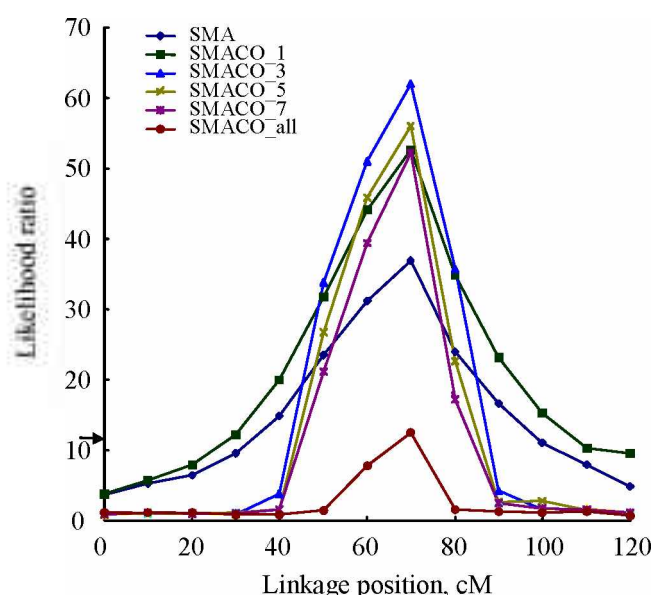
## RESULTS

### Single marker analysis with or without cofactors

In population 1, QTL1 was detected by SMA, regardless of including cofactors (Figure 1). Incorporating cofactors led the likelihood ratio (LR) to be increased at markers closely linked to the QTL except for SMACO\_all, and to be decreased at markers far from the QTL. The LR estimates showed that SMACO\_3 worked most effectively. Table 2 shows that the estimates of QTL1 effects using SMA and SMACO\_all differed ( $p < 0.01$ ) from the input value. Using SMACO\_1, SMACO\_3, SMACO\_5, and SMACO\_7 reduced the difference ( $p < 0.05$ ). The estimation of the QTL effect was improved by introducing cofactors.

All the SMA with or without cofactors were able to discover the two QTLs on chromosome 2 except for SMACO\_all (Figure 2). Using ordinary SMA, the LR estimates at all markers were significant at the 0.05 genome-wide significance level. Introducing 3 to 7 cofactor loci dramatically reduced the genetic background noise.

With multiple QTLs, incorporating cofactors in the analytical models also improved the estimation of QTL



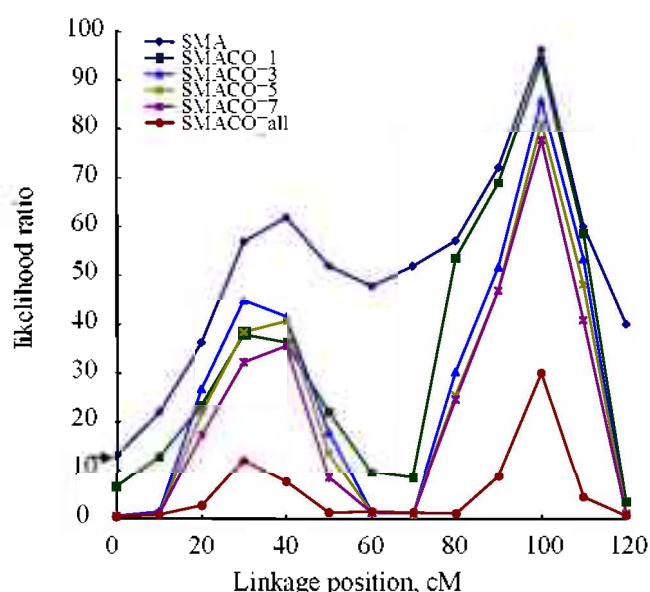
**Figure 1.** Likelihood ratio using SMACO in the search for a single QTL on chromosome 1. SMACO\_n indicated single-marker analysis using n cofactors. The arrow represents the genome-wide threshold value at 0.05 significance level.

**Table 2.** Estimate of QTL genetic effect in population 1 using SMACO<sup>1,2</sup>

	QTL1	QTL2	QTL4
Input value	1.040	0.956	1.174
SMA	0.976** $\pm$ 0.021	1.260** $\pm$ 0.024	1.501** $\pm$ 0.030
SAMCO_1	0.984* $\pm$ 0.020	0.954 $\pm$ 0.022	1.337** $\pm$ 0.029
SMACO_3	0.988* $\pm$ 0.020	0.924 $\pm$ 0.023	1.267** $\pm$ 0.020
SMACO_5	0.988* $\pm$ 0.023	0.910 $\pm$ 0.025	1.213 $\pm$ 0.036
SMACO_7	0.974* $\pm$ 0.025	0.897 $\pm$ 0.033	1.270* $\pm$ 0.038
SMACO_all	0.887** $\pm$ 0.034	0.843* $\pm$ 0.040	1.290** $\pm$ 0.031

<sup>1</sup> SMACO\_1, SMACO\_3, SMACO\_5, SMACO\_7 and SMACO\_all were the single marker analysis with 1, 3, 5, 7, and all other markers as cofactors. The standard errors were empirically obtained from 50 replicates.

<sup>2</sup> \*  $p < 0.05$ , \*\*  $p < 0.01$ .

**Figure 2.** Likelihood ratio using SMACO in the search for multiple QTLs on chromosome 2. SMACO\_n indicated single-marker analysis using n cofactors as controls. The arrow represents the genome-wide threshold value at 0.05 significance level.

effects (Table 2). Estimates of QTL2 and QTL4 effects using ordinary SMA were largely different ( $P < 0.01$ ) from their input values. The overestimation resulted from the genetic background noise due to the segregation of the linked QTL. Incorporating cofactors dramatically absorbed the genetic background noise in estimating QTL effects. The estimates of QTL2 and QTL4 effects using SMACO\_5 did not differ ( $P > 0.01$ ) from their input values. However, using SMACO\_all led to underestimation ( $P < 0.05$ ) of QTL with smaller effect and overestimation ( $P < 0.01$ ) of QTL with larger effect.

#### Interval mapping with or without cofactors

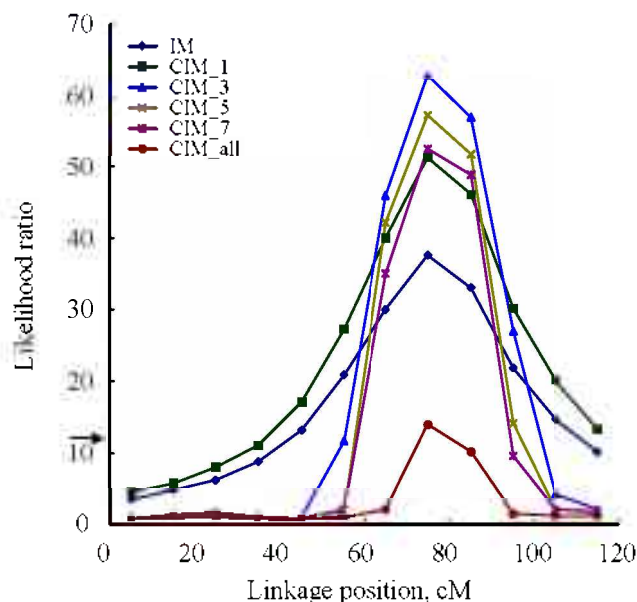
The use of CIM in QTL detection produced results similar to those obtained by SMACO but offered more

accurate estimates of QTL effects than SMACO. When chromosome 1 was analyzed with cofactor(s), LR estimates were increased at marker interval flanking or closely linked to QTL 1 except for CIM\_all, but decreased at other intervals (Figure 3). The LR estimates showed that the use of three cofactors (CIM\_3) worked most effectively as shown with SMACO\_3. The estimation for the QTL 1 effect was improved by introducing cofactor(s) (Table 3).

In the analyses of multiple QTLs on chromosome 2, LR estimates were significant at all marker intervals in IM (Figure 4). Incorporating cofactor markers produced smaller LR estimates at unlinked marker intervals (Figure 4). The effects for QTL 2 and QTL 4 were overestimated ( $p < 0.01$ ) with IM, and the position estimates were different from the input values (Table 3). However, they were greatly improved by introducing cofactor(s), and especially the estimates of the QTL2 and QTL4 effects using CIM\_5 did not differ ( $P > 0.01$ ) from their input values.

#### Types of cofactors

When applying CIM to mapping QTLs on chromosome 2, the estimates of QTL effects and positions using CIM with UL were corresponding to those using IM (Table 4). However, estimation of the genetic effects was improved using linked markers such as NCLNQ, LNQ, and FL. Especially, estimates of genetic effects with LNQ did not differ ( $p > 0.05$ ) from their input values. Furthermore, incorporating such linked markers dramatically reduced the LR estimates at the control marker interval while the LR

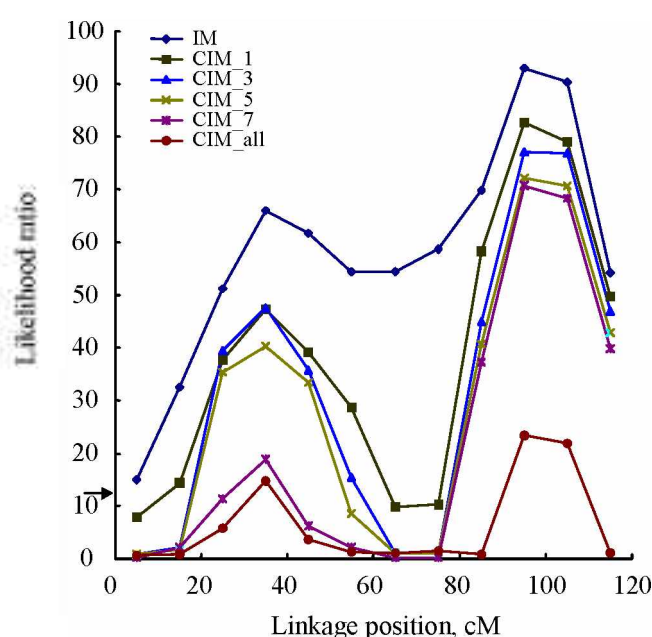
**Figure 3.** Likelihood ratio using CIM in the search for a single QTL on chromosome 1. CIM\_n indicated composite interval mapping using n cofactors as controls. The arrow represents the genome-wide threshold value at 0.05 significance level.

**Table 3.** Estimates of QTL genetic effect (GE) and linkage position (LP, cM) in population 1 using CIM<sup>1</sup>

	QTL 1		QTL 2		QTL 4	
	GE	LP	GE	LP	GE	LP
Input value	1.040	67	0.956	35	1.174	99
IM	1.023±0.020	66.57±0.46	1.127**±0.022	38.29**±0.64	1.488**±0.027	97.86**±0.35
CIM_1	1.036±0.018	37.57±0.43	0.967±0.021	34.29±0.44	1.323**±0.026	98.29*±0.35
CIM_3	1.041±0.018	67.41±0.42	0.954±0.021	34.38±0.44	1.248**±0.020	99.24±0.35
CIM_5	1.044±0.020	67.14±0.44	0.955±0.022	35.48±0.55	1.206±0.027	99.05±0.35
CIM_7	1.042±0.021	67.29±0.46	0.947±0.029	34.81±0.56	1.260**±0.028	99.05±0.34
CIM_all	1.035±0.026	67.33±0.63	0.923±0.031	34.00±0.57	1.278**±0.030	98.38*±0.31

<sup>1</sup> CIM\_1, CIM\_3, CIM\_5, CIM\_7 and CIM\_all were the composite interval mapping with 1, 3, 5, 7, and all other markers as cofactors. The standard errors were empirically obtained from 50 replicates.

\*  $p < 0.05$ , \*\*  $p < 0.01$ .



**Figure 4.** Likelihood ratio using CIM in the search for multiple QTLs on chromosome 2. CIM<sub>n</sub> indicated a composite interval mapping using *n* cofactors as controls. The arrow represents the genome-wide threshold value at 0.05 significance level.

estimates at the intervals where QTL2 and QTL4 were located remained significant ( $p < 0.05$ ).

#### Heritability level

Populations 2 to 6 with various heritabilities were used to investigate how heritability level influenced QTL detection with cofactors, and the results with chromosome 2 using CIM are shown in Figure 5. All analyses failed to distinguish the two closely linked QTLs, QTL3 and QTL4, and this jointly detected QTL was referred to as QTLJ. Larger LR estimates at the QTL regions were obtained with a larger heritability. QTL2 and QTLJ were detected from all the replicates when the heritability was 0.5 or larger. The results using SMACO showed similar patterns (data not presented). With a heritability of 0.1, SMACO failed to

discover one of the two QTLs in 45 out of 50 replicates, and it failed in 17 out of 50 replicates with a heritability of 0.3. When the heritability was 0.5 or higher, these two QTLs were discovered at the significance level of 0.05 in all replicates.

#### Marker spacing

Population 4 was used to examine how marker spacing influenced QTL detection with cofactors, and the results using CIM are shown in Figure 6. QTL3 and QTL4 were detected jointly when the marker spacings ranged from 4 cM to 40 cM. They were discovered separately only when the marker spacing was 2 cM. In attempts to distinguish two very closely linked QTLs (less than 5 cM apart), both CIM and SMACO failed even when the marker spacing was reduced to 1 cM (data are not presented).

The peak of LR estimates for QTL2 was obviously discovered when the marker spacing ranged from 2 cM to 20 cM. It became hardly detectable when the marker spacing was 30 cM or larger, and the position began to be shifted.

## DISCUSSION

#### Reduction of genetic background noise by cofactors

An ordinary SMA has a theoretical problem of producing underestimated QTL estimates (Liu, 1998), and so does the SMACO (Table 2). However, the genetic background noise attributable to other QTLs was dramatically absorbed with well-chosen markers as cofactors.

Applying single-QTL methods such as SMA and IM to deal with multiple QTLs was challenged by the confounded effect of multiple QTLs because segregation of other QTLs also contributed to the phenotypic variance. Therefore, both SMA and IM could not discern whether significant effects at several linked markers are due to a common QTL or to several linked QTLs. The presence of multiple QTLs introduced serious biases into QTL estimation with these approaches. It often led to a confusing situation where all



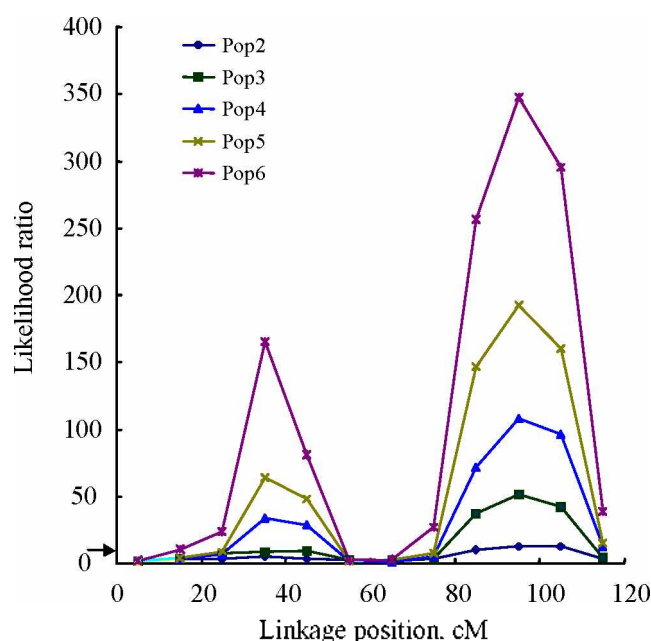
**Table 4.** Estimates of QTL genetic effect (GE) and linkage position (LP, cM) in population 1 using CIM with various types of cofactors: no cofactor (NONE), unlinked markers (UL), markers not closely linked to nuisance QTL (NCLNQ), markers linked to nuisance QTL (LNQ), and flanking markers (FL)<sup>1,2,3</sup>

Types of cofactors	30~40 cM (QTL2)			90~100 cM (QTL4)			60~70 cM
	LR	GE	LP	LR	GE	LP	LR
Input value		0.956	35		1.174	99	
NONE	65.9±2.3	1.127**±0.022	38.29±2.52	92.9±2.1	1.488**±0.027	97.86±2.31	54.3±2.0
UL	70.4±2.1	1.130**±0.019	38.28±2.83	115.6±2.2	1.483**±0.024	97.30±2.31	67.4±1.9
NCLNQ	21.0±1.0	0.956±0.028	26.23±1.87	45.9±1.8	1.254*±0.029	98.75±2.10	5.3±0.6
LNQ	45.2±1.4	0.954±0.018	36.37±1.71	64.4±1.9	1.217±0.026	98.79±2.18	0.4±0.1
FL	7.0±0.8	0.891*±0.031	37.15±1.87	19.3±0.9	1.266*±0.030	99.00±2.53	0.9±0.2

<sup>1</sup> A control non-QTL interval (60~70 cM) was also examined to reflect false positive results.

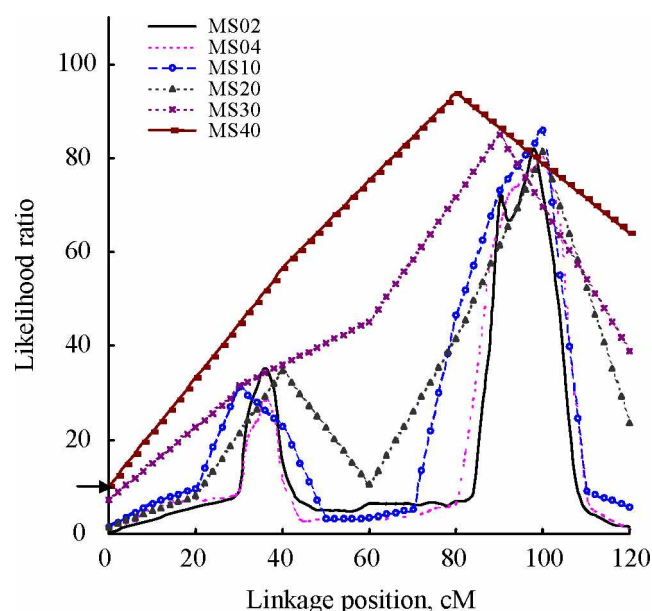
<sup>2</sup> Four cofactors were used in all analyses except for NONE. Since the QTLs on chromosome 2 were being examined, all the markers located on chromosome 1 were regarded as UL. LNQ were those located within 25 cM from the nuisance QTL, and those located greater than 25 cM from the QTL were regarded as NCLNQ. FL were flanking the interval being examined. The standard errors were empirically obtained from 50 replicates.

<sup>3</sup> \*  $p < 0.5$ , \*\*  $p < 0.01$ .



**Figure 5.** Likelihood ratio from mapping multiple QTLs on chromosome 2 simulated with various heritability levels. The input values for heritability ranged from 0.1 to 0.9 with an increment of 0.2. A forward regression using 5 cofactors was used to obtain the results. The arrow represents the genome-wide threshold value at 0.05 significance level.

marker loci were significant. This problem was especially serious with SMA. However, this did not mean that SMA was not capable of searching for multiple QTLs. The current study showed that the accuracy of QTL mapping increased by obtaining the linked QTLs separately. Using SMACO and CIM, the effect of QTL elsewhere in the linkage was efficiently controlled. As a result, LRs increased for markers closely linked to a QTL, but decreased for markers far from the QTL. The results indicated that both CIM and SMACO were powerful in dealing with multiple QTLs.



**Figure 6.** Likelihood ratio from mapping multiple QTLs on chromosome 2 simulated with various marker spacing. The results were obtained by forward regression using 5 cofactors. MSxx indicates marker spacing of xx cM where xx = 2, 4, 10, 20, 30, or 40. The arrow represents the genome-wide threshold value at 0.05 significance level.

### Choice of cofactors

This study indicated that the choice of markers as cofactors was of great importance in mapping multiple QTL. The genetic background noise was dramatically absorbed with linked markers, but not with unlinked markers. This study showed that flanking markers could be desirable cofactors when the heritability was high. Piepho and Gauch (2001) also suggested that the use of flanking markers was the best controls because they were very efficient at absorbing genetic background noise. Note that precaution should be taken with closely linked flanking markers. The

flanking markers tended to over-absorb the QTL effect, especially those with a small heritability. As a result, the effects of weak QTL were underestimated. Another concern in CIM and SMACO was the number of cofactors. This study showed that 3 to 5 cofactors led to good results in mapping multiple QTLs. Note that it was challenged by the over-fitting problem when the heritability was low and when the number of regressor variables was large. No more than 5 markers were recommended for a low heritability.

### Heritability level

This study showed that the heritability level influenced QTL mapping. A higher level of heritability led to more significant LRs and more accurate estimates of QTL effects and positions. The accuracy of marker-trait association was increased with a large heritability, enhancing the ability to detect and map QTLs (Belknap, 1998; Wu, 1999; Kearsey and Farquhar, 1998). The current study suggested that reliable results in QTL mapping were obtained when heritability was 0.3 or larger. This concurred with the findings of Williams and Blangero (1999). They investigated the asymptotic power of LR test for detecting linkage to a QTL and discovered the minimum detectable QTL heritability of 0.35. Knowing such a threshold level of heritability was important for optimizing the experimental design and the sample size for QTL mapping (Weller et al., 1990; Moreno-Gonzalez, 1993).

### Marker spacing

The results from the current study showed that the effects of CIM and SMACO in the search for multiple QTLs were not different with marker spacings from 4 to 20 cM. After a study, Darvasi et al. (1993) reported that the power of detecting a QTL was virtually the same for a marker spacing of 10 cM as that for an infinite number of markers and was slightly decreased for marker spacing of 20 to 50 cM. They also found that reducing marker spacing below the resolving power defined as the 95% confidence interval map position did not improve by narrowing the confidence interval. Similar evidence was presented by Piepho (2000) that the power of QTL detection and the standard errors of genetic effect estimates were affected little by any increase of marker density beyond 10 cM. The present study showed that the location of QTL 2 tended to be shifted with marker spacings from 30 cM to 40 cM. This was in agreement with Martinez and Curnow (1992) and Haley and Knott (1992) who reported spurious QTLs in their search for the presence of multiple QTLs using IM with such large marker spaces. On the other hand, reducing marker space to less than 2 cM allowed the two QTLs (QTL3 and QTL4) that were 9 cM apart to be discovered separately.

## REFERENCES

- Basten, C. J., B. S. Weir and Z. B. Zeng. 1994. Zmap-a QTL cartographer. Proc. 5th World Cong. Genet. Appl. Livst. Prod. 22:65-66. Guelph, Ontario, Canada.
- Belknap, J. K. 1998. Effect of within-strain sample size on QTL detection and mapping using recombinant inbred mouse strains. *Behav. Genet.* 28:29-38.
- Boyle, A. E. and K. Gill. 2001. Sensitivity of AXB/BXA recombinant inbred lines of mice to locomotor activating effects of cocaine: a quantitative trait analysis. *Pharmacogenetics* 11:255-264.
- Churchill, G. A. and R. W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. *Genetics* 138:963-971.
- Darvasi, A., A. Weinreb, V. Mink, J. I. Weller and M. Soller. 1993. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics* 134:943-951.
- Drake, T. A., E. Schadt, K. Hannani, J. M. Kabo, K. Krass, V. Colimayo, L. E. Greaser, J. Goldin and A. J. Lusis. 2001. Genetic loci determining bone density in mice with diet-induced atherosclerosis. *Physiol. Genomics* 5:205-15.
- Haley, C. S. and S. A. Knott. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69:315-324.
- Jansen, R. C. 1993. Interval mapping of multiple quantitative trait loci. *Genetics* 135:205-11.
- Kearsey, M. J. and A. G. Farquhar. 1998. QTL analysis in plants; where are we now? *Heredity* 80:137-142.
- Lander, E. S. and D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199.
- Lee, C. 2002. What holds the future of quantitative genetics? *Asian-Aust. J. Anim. Sci.* 15:303-308.
- Liu, B. H. 1998. *Statistical Genomics: Linkage, Mapping and QTL Analysis*. CRC Press LLC, Boca Raton, FL, USA.
- Martinez, O. and R. N. Curnow. 1992. Estimating the locations and sizes of the effects of quantitative trait loci using flanking markers. *Theor. Appl. Genet.* 85:480-488.
- Moreno-Gonzalez, J. 1993. Efficiency of generations for estimating marker-associated QTL effects by multiple regression. *Genetics* 135:223-231.
- Piepho, H. P. 2000. Optimal marker density for interval mapping in a backcross population. *Heredity* 84:437-440.
- Piepho, H. P. and H. G. Gauch. 2001. Marker pair selection for mapping quantitative loci. *Genetics* 157: 433-444.
- Robison, B. D., P. A. Wheeler, K. Sundin, P. Sikka and G. H. Thorgaard. 2001. Composite interval mapping reveals a major locus influencing embryonic development rate in rainbow trout (*Oncorhynchus mykiss*). *J. Hered.* 92:16-22.
- Wayne, M. L., J. B. Hackett, C. L. Dilda, S. V. Nuzhdin, E. G. Pasyukova and T. F. Mackay. 2001. Quantitative trait locus mapping of fitness-related traits in *Drosophila melanogaster*. *Genet. Res.* 77:107-116.
- Weller, J. I., Y. Kashi and M. Soller. 1990. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *J. Dairy Sci.* 73:2525-2537.

- Williams, J. T. and J. Blangero. 1999. Asymptotic power of likelihood-ratio tests for detecting quantitative trait loci using the COGA data. *Genet. Epidemiol.* 17(Suppl. 1):S397-402.
- Wright, F. A. and A. Kong. 1997. Linkage mapping in experimental crosses: the robustness of single-gene models. *Genetics* 146:417-425.
- Wu, R. L. 1999. Mapping quantitative trait loci by genotyping haploid tissues. *Genetics* 152:1741-1752.
- Wu, X. L., C. Lee, J. Jiang, Y. L. Peng, S. L. Yang, B. N. Xiao, X. C. Liu and Q. S. Shi. 2001. Mapping a quantitative trait locus for growth and backfat on porcine chromosome 18. *Asian-Aust. J. Anim. Sci.* 14:1665-1669.
- Wu, X. L., C. Lee, J. Jiang, Y. L. Peng, H. F. Yan, S. L. Yang, B. N. Xiao, X. C. Liu and Q. S. Shi. 2002. Quantitative trait loci mapping for porcine backfat thickness. *Asian-Aust. J. Anim. Sci.* 15:923-927.
- 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.