J. Appl. Genet. 43(1), 2001, pp. 69-83

# Application of a mixed inheritance model to the detection of quantitative trait loci in swine

Joanna SZYDA[1,2], Zengting LIU[2], Eli GRINDFLEK[3,4], Sigbjørn LIEN[3]

[1]Department of Animal Genetics, Agricultural University of Wrocław, Wrocław, Poland
[2]United Datasystems for Animal Production (VIT), Verden, Germany
[3]Department of Animal Science, Agricultural University of Norway, Ås, Norway
[4]The Norwegian Pig Breeders Association (NORSVIN), Hamar, Norway

**Abstract**. The primary goal of this study was to investigate statistical properties of a mixed inheritance model for the localization of quantitative trait loci (QTL). This is based on the analysis of phenotypic data for the amount of intramuscular fat (IMF) scored on 305 individuals originating from a cross between Duroc and Norwegian Landrace breeds. Marker genotype information is available for F1 and F2 generations. Statistical procedures compared involve i) the interval mapping, ii) the composite interval mapping, iii) a regression method, and iv) a mixed inheritance model accounting for a random animal additive genetic effect and relationships between individuals. The basic statistical properties of the latter approach are then assessed using Monte Carlo simulations showing slight unconservativeness as compared to $\chi^2_{2df}$ and reasonable power to detect QTL of moderate effects. In the analysis of IMF data, the significant evidence for the existing QTL is detected on chromosome 6. A chromosomal region recommended for a second-step fine mapping analysis is identified between markers SW1823 and S0228, based on three types of confidence intervals derived by using: i) the Jackknife algorithm, ii) the numerical variance approximation, and iii) the LOD score approach. The Jackknife algorithm was additionally used to quantify each family's contribution to the test statistic and to the estimate of QTL position.

**Key words**: backcross, meat quality, mixed inheritance model, pig, QTL mapping.

## Introduction

Experimental designs based on crossing inbred lines, which are divergent for the trait of interest, provide the most informative (i.e. powerful) data structure available for mapping quantitative trait loci (QTL) in livestock. Among the most

Received: September 17, 2001. Accepted: October 26, 2001.
Correspondence: Joanna SZYDA, Department of Animal Genetics, Agricultural University of Wrocław, ul. Kożuchowska 7, 51-631 Wrocław, Poland, e-mail: szyda@karnet.ar.wroc.pl

important advantages of such designs are: i) high (or complete) heterozygosity of QTL allowing for the differentiation between the effects of alternative alleles on a quantitative trait, ii) high (or complete) heterozygosity of markers, which enables the identification of recombination events along the chromosome, and iii) strong linkage disequilibrium between markers and QTL (ANDERSSON-EKLUND et al. 1998, SILLANPÄÄ, ARJAS 1998). However in practice, instead of inbred lines, it is common to cross divergent outbred populations or breeds. In such a case, one can still profit from the advantages of a special experimental design, but the degree of heterozygosity and linkage disequilibrium are lower. Furthermore, polygenic effects of individuals from the parental generation belonging to the same breed (population) are no longer identical.

Several statistical approaches have been proposed for the modelling of such experimental data. While the first models mapped QTL utilizing information from one marker locus (SOLLER et al. 1976), the emphasis of a further development was put on using genotypes of two, and later, of many markers simultaneously, utilizing more out of molecular genetic information available (LANDER, BOTSTEIN 1989, HALEY, KNOTT 1992, JANSEN 1993, ZENG 1994, SILLANPÄÄ, ARJAS 1998). Further model enhancements comprise a multivariate analysis (JIANG, ZENG 1995), accounting for genetic imprinting (DE KONING et al. 2000), including family-specific estimates and nonzero covariances between observations (XU 1998), and the separation of polygenic and QTL components (RATHJE et al. 1997, PÉREZ-ENCISO, VARONA 2000, SZYDA et al. 2000, NAGAMINE, HALEY 2001).

In the current study a new model for the analysis of a line cross data is proposed. It is based on the regression approach of HALEY and KNOTT (1992) with modifications comprising the introduction of a random animal additive genetic effect and covariances between related individuals. This method is applied to the detection of a QTL responsible for the amount of intramuscular fat (IMF) using data from a cross between Duroc and Norwegian Landrace breeds. Results are then compared with those based on interval (LANDER, BOTSTEIN 1989, HALEY, KNOTT 1992) and composite interval mapping (ZENG 1994) methods.

## Material and methods

### Data

Phenotypic data on IMF was collected for 305 individuals from the F2 generation. This progeny originated from a cross between Duroc sires and Norwegian Landrace dams in the parental generation, and then from backcrossing each of five F1 sires with eight Landrace-Yorkshire dams. This results in an F2 data structure comprising five large paternal halfsib families, each divided into small (up to eight sibs) fullsib families. For F1 and F2 individuals, marker genotypes at chromosome 4 (11 markers, covering approximately 126 cM), chromosome 6 (9 markers, covering approximately 134 cM), and chromosome 7 (9 markers, covering approximately 137 cM) were available. The detailed description of the experi-

mental design and the molecular information can be found in GRINDFLEK et al. (2000).

**Preliminary analysis without marker information**

In order to get better insight into the characteristics of the data, preliminary analysis of phenotypic observations, ignoring the molecular information, was performed. First, the frequency distribution, reflecting the probability density function of IMF phenotypes, was visually checked for especially extreme departure from normality, as each of the method used for QTL mapping assumes a normal distribution of phenotypes. Furthermore, a sequence of linear regression models was fitted to the data in order to identify which of the available effects: sire, dam, sex, and slaughter weight, have a significant impact on the observed phenotypic variation. Differences in the goodness of fit between particular models were assessed using the likelihood ratio test with a number of degrees of freedom corresponding to the difference in the number of parameters. A final step comprised the estimation of additive genetic variance of the analysed traits. Variance components were estimated via restricted maximum likelihood, using the SAS package (SAS 1996), based on the following model:

$$y_{ijkl} = \mu + sex_i + s_j + d_k + e_{ijkl} ,$$

where: $y_i$ represents a quantitative trait value of *ijkl*-th individual, $\mu$ is the overall mean, $sex_i$ is a fixed effect of sex of individual *i,* $s_j$ and $d_k$ are random effects of sire and dam respectively, and $e_i$ is the random error.

**QTL detection**

**QTL mapping models**

Four statistical models utilizing various amounts of the available information were used for QTL mapping: Lander-Botstein model – LBM (LANDER, BOTSTEIN 1989), Zeng model – ZM (ZENG 1994), Haley-Knott model – HKM (HALEY, KNOTT 1992), and a mixed inheritance model (MIM) based on our modification of the HKM approach introduced by SZYDA et al. (2000). The first two methods were implemented via the QTL-Cartographer package (BASTEN et al. 1994), the maximum likelihood of HKM and MIM was evaluated through the SAS procedure MIXED (SAS 1996).

LBM is the simplest of all the models:

$$y_i = \mu + q_1 x_{q1i} + q_2 x_{q2i} + e_i ,$$

where: $y_i$, $\mu$ and $e_i$ are as above, $q_1$ and $q_2$ denote respectively the fixed effects of a heterozygous (say, Qq), and a homozygous (say, qq) QTL genotype, $x_{q1i}$, and $x_{q2i}$ represent appropriate elements of design vectors for $q_1$, and $q_2$ corresponding to QTL genotype probabilities in F2 offspring. In the LBM approach double recombinations within a marker interval are ignored, so that possible values

of $x_{q1i}$, and $x_{q2i}$ are: 0, 1, $\frac{r_{M_1Q}}{r_{M_1M_2}}$, or $\left(1-\frac{r_{M_1Q}}{r_{M_1M_2}}\right)$, where $r_{M_1Q}$ and $r_{M_1M_2}$ respectively represent recombination rates between a marker ($M_1$) and a putative QTL position ($Q$) within a marker interval, and between flanking markers ($M_1$, $M_2$).

ZM was especially designed for mapping of multiple linked QTL:

$$y_i = \mu + q_{1m} x_{q1mi} + q_{2m} x_{q2mi} + \sum_{k \neq m} (q_{1k} x_{q1ki} + q_{2k} x_{q2ki}) + e_i .$$

As this approach is based on LBM, it models QTL genotype probabilities in the same way as above. Also the first three components of the model, corresponding to a marker interval considered ($m$), have the same meaning, but an extra term is added accounting for additional putative QTL located in the neighbourhood of other available marker intervals ($k$).

HKM was originally formulated in terms of linear regression. To strike its correspondence to LBM, here we describe HKM in the following form:

$$y_i = \mu + q_1 x^*_{q1i} + q_2 x^*_{q2i} + e_1 .$$

One of the most important developments implemented in HKM is to allow for the double recombinations within a marker interval in calculating QTL genotype probabilities, so that $x^*_{q1i}$ and $x^*_{q2i}$ become functions of recombination rates between both markers, as well as between each marker and a putative QTL position.

The MIM approach introduced in this study is related to HKM, but additionally involves the incorporation of $\alpha_i$ = a random additive genetic effect of animal $i$, which accumulates the parental polygenic influence together with the Mendelian sampling effect, and β = additional fixed effects representing nongenetic covariates (in our model – sex):

$$y_i = \mu + \alpha_i + \beta x_{\beta i} + q_1 x^*_{q1i} + q_2 x^*_{q2i} + e_i .$$

As above, $x^*_{q1i}$, and $x^*_{q2i}$ represent appropriate elements of design vectors for $q_1$ and $q_2$ accounting for double recombinants, while $x_{\beta i}$ is a design vector for sex effects in $\beta$.

**Assumptions in QTL mapping**

The assumption underlying all the four above methods is that both paternal lines in a cross are fully homozygous both for markers and a QTL. As a consequence, F1 sires, providing information on putative QTL allele transmissions, are expected to be fully informative (i.e. heterozygous) for all the loci considered. This is an ideal situation so that for the analysis of our real data from a backcross experiment the following approximations are set:

i) possible alleles of a putative QTL are divided in two categories – favourable and unfavourable, so that the practical analysis relies on a biallelic QTL,
ii) all dams mated to F1 sires are homozygous at a putative QTL (say qq),
iii) based on the marker information on dams, offspring, and sires, the marker haplotype phase of F1 sires is known without error,

iv) for a given F1 sire, a favourable QTL allele is assigned to the marker haplotype associated with a higher phenotypic mean value of offspring, which obtained this haplotype, the other haplotype is assigned an unfavourable QTL allele.

Following these assumptions, QTL genotype effects can be interpreted as $q_1 = E(Y|Qq) = \mu + d$ and $q_2 = E(Y|qq) = \mu - a$, where $d$ and $a$ represent respectively a dominance and an additive QTL effect, whereas corresponding QTL genotype probabilities are equivalent to paternal QTL allele transmission probabilities, so that:

$$x_{q1i} = P(Qq|M_i, Qq_{Si}, qq_{Di}, r) = P(Q|M_i, Qq_{Si}, qq_{Di}, r) = P(Q|M_i, r)P(Qq_{Si})P(qq_{Di}),$$

$$x_{q2i} = P(qq|M_i, Qq_{Si}, qq_{Di}, r) = P(q|M_i, Qq_{Si}, qq_{Di}, r) = P(q|M_i, r)\,P(Qq_{Si})\,P(qq_{Di}),$$

where $M_i$ is the set of marker information for individual i comprising marker genotype of a sire, a dam and individual's own genotype, $Qq_{Si}$ and $qq_{Di}$ are the assumed genotypes at a putative QTL of a sire and a dam of individual i, respectively, $r$ is a set of recombination rates between both markers or between a marker and a putative QTL. For LBM and ZM, $P(Qq_{Si}) = P(qq_{Di}) = 1$. It is however possible to relax the assumption that a given sire is heterozygous at a putative QTL by modelling $P(Qq_{Si})$. We explore this possibility while calculating appropriate transmission probabilities for HKM and MIM by assuming:

$$P(Qq_{Si}) = 1 - \alpha_{TDT},$$

where $\alpha_{TDT}$ is a nominal type I error rate of a TDT test statistic (SZYDA et al. 1998) calculated for each marker interval separately comparing phenotypic means of F2 individuals which obtained one of two possible nonrecombined marker haplotypes from the sire.

**(Co)variance structure**

An additional feature introduced into MIM is the incorporation of information on the relationship between individuals into the (co)variance matrix. While LBM, ZM, and HKM assume no correlation between phenotypic observations $y_i$, our approach models trait values of full- and half-sibs as correlated. This imposed correlation is based on the similarities between polygenic effects of related individuals, as described by a standard additive genetic relationship matrix.

**Estimation**

The estimation of model parameters in MIM is done through solving the mixed model equations (HENDERSON 1984) assuming that the (co)variance matrix for random animal polygenic effects is known. (Co)variance components are estimated prior to the QTL detection, as described above. To test the hypothesis of no QTL (i.e. $H_0$: $q_1 = q_2 = 0$), a likelihood ratio test statistic ($\lambda$), comparing the maximum of likelihood functions obtained under $H_0$ $[L(\hat{\beta}_0)]$ and under the unrestricted model $[L(\hat{\beta}_1)]$, was used:

$$\lambda = -2\left\{\ln[L(\hat{\beta}_0)] - \ln[L(\hat{\beta}_1)]\right\} = \left[y - E_0(y)\right]^T V^{-1}\left[y - E_0(y)\right] - \\ -\left[y - E_1(y)\right]^T V^{-1}\left[y - E_1(y)\right].$$

The components of λ are defined as follows:

$$E_0(y) = \mu + \beta x_\beta, \quad E_1(y) = \mu + \beta x_\beta + q_1 x^*_{q1} + q_2 x^*_{q2},$$

$$L(\hat{\beta}_m) = |V|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left[y - E_m(y)\right]^T V^{-1}\left[y - E_m(y)\right]\right\},$$

where: subscript *m* denotes the particular model (i.e. 0 or 1), *V* represents the phenotypic (co)variances, and the other components are defined as above. The so-called likelihood profile, i.e. the values of λ calculated for fixed map positions along the analysed chromosome region, gives the estimate of the most probable QTL location.

**Confidence intervals for QTL position**

As the aim of the current analysis was not to map a QTL itself, but rather to localize chromosomal regions responsible for a significant proportion of a quantitative variation of IMF, which will be further explored in a second step fine mapping study, we constructed confidence intervals (CI) for a putative QTL position along the chromosome where additional markers should be genotyped. A standard method for estimating 1– α confidence intervals, based on a normal approximation of the asymptotic distribution of maximum likelihood estimates, follows:

$$\left(\hat{\theta} - z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}} < \theta < \hat{\theta} + z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}}\right),$$

where: $\hat{\theta}$ is the estimate of $\theta$, $\alpha$ is the probability of type I error, $z_{\frac{\alpha}{2}}$ is the critical value corresponding to $\frac{\alpha}{2}$ type I error rate based on the standard normal distribution[1], and $\sigma_{\hat{\theta}}$ is the standard deviation of $\hat{\theta}$. In our case $\alpha$ was set to 0.05, $\theta$ represents a QTL position in cM from the leftmost marker, and the unknown $\sigma_{\hat{\theta}}$ is replaced by an estimate.

**Numerical approximation of $\sigma_{\hat{\theta}}$**

Following MEYER and HILL (1992), the variance of the estimate of QTL position can be approximated numerically based on the values of the likelihood profile (*lnL*) in the vicinity of the estimate using:

[1] Assuming that $\theta$ follows a normal distribution.

$$\sigma^2_{\hat{\theta}A} = -\frac{\Delta^2}{\ln L(\hat{\theta}-\Delta) + \ln L(\hat{\theta}+\Delta) - 2\ln L(\hat{\theta})},$$

with Δ set to 1cM.

**Estimation $\sigma_{\hat{\theta}}$ of based on the Jackknife algorithm**

Another approach towards estimating $\sigma_{\hat{\theta}}$ used in our study is the application of the Jackknife algorithm, considering a paternal-halfsib family as a resampling unit. The Jackknife variance of $\hat{\theta}$ is defined as:

$$\sigma^2_{\hat{\theta}J} = \frac{N-1}{N}\sum_{j=1}^{N}(\hat{\theta}-\hat{\theta}_j)^2 ,$$

where: $N$ is the number of paternal-halfsib families (i.e. 5) and $\hat{\theta}_j$ is the estimate of $\hat{\theta}$ based on data without halfsib family $j$ (EFRON, TIBSHIRANI 1993).

**Support interval**

Support interval utilizes information on the whole shape of the available likelihood profile, by comparing the difference between a maximal likelihood $[\ln L(\hat{\theta})]$ and likelihood values for other QTL positions along the chromosome $[\ln L(\hat{\theta}_i)]$. The 0.95 probability of a true position being within a support interval is approximated by the loglikelihood difference of 2 (LIU 1998):

$$\ln L(\hat{\theta}) - \ln L(\hat{\theta}_i) \le 2 .$$

## Jackknife analysis

In order to get more insight into the influence of a particular sire on QTL mapping results, Jackknife plots were constructed of a maximal value of λ and of the corresponding estimate of a QTL position along the chromosome. Practically, it means that an analysis was performed five times, but each time one of five paternal halfsib families was removed from the data set. Corresponding values of λ and a QTL position were then compared on a plot, but without using any formal test.

## Significance levels for QTL detection

As pointed out by ZENG (1994), VISSCHER et al. (1996a), SCHELER et al. (1998), and DUPUIS and SIEGMUND (1999), the $H_0$ distribution of λ depends on the sample size, the number of markers, intermarker distances and QTL effects. We express the significance of QTL analysis results using both nominal and chromosomewise rates of type I error. A nominal type I error assessment is based on $\chi^2_{1df}$ for LBM and ZM and $\chi^2_{2df}$ for HKM and MIM. A chromosomewise type I error is obtained by permuting IMF phenotypes across individuals and re-

calculating a likelihood ratio test profile (CHURCHILL, DOERGE 1994). After each permutation the highest value of $\lambda$ found along the chromosome is scored. In order to construct an empirical distribution of $\lambda$ such procedure was repeated 500 times for LBM, ZM and HKM, and 100 times for MIM.

**Statistical properties of MIM**

Power and type I error rate of MIM were assessed using 100 and 150 simulated data sets, respectively. The simulated data was intended to resemble the real data set in terms of: i) the experimental design (five sires each mated to eight dams, eight piglets per mating), ii) trait variation (phenotypic variance $\sigma_p^2 = 0.46$, and polygenic variance $\sigma_a^2 = 0.18$), and iii) molecular information (6 marker loci equally spaced every 15 cM). The same polygenic mean and variance were assumed for sire and dam lines, so that for power calculations core differences between parental lines were based on their QTL effect, which were fixed for the alternative alleles. Three different additive genetic effects of the favourable QTL allele $0.5\sigma_a$ , $0.3\sigma_a$ , or $0.1\sigma_a$, and 5% type I error rate were considered. The computed power can be regarded as a maximal power, because the ideal situation was assumed that all marker genotypes in $F_2$ are informative.

## Results

The frequency distribution of IMF did not show especially extreme departures from normality. The additive genetic and phenotypic variance estimates of 0.18 and 0.46 respectively, result in the heritability of 0.39. This estimate, however, corresponds with polygenic heritability without considering a QTL variation, and due to the small sample size it is subjected to a large standard error. The preliminary analysis of phenotypic records showed that sire, dam, and sex significantly influence the phenotypic variation of the trait. Consequently, in HKM and MIM sex was used as a covariate, while parental influence was modelled through introducing (co)variances between related individuals in MIM.

Likelihood profiles for QTL position resulting from four different methods are shown in Figure 1. All the methods point at the same marker interval SW1823-S0003 as the most probable QTL location. The detailed results comprising QTL position estimates in cM from the leftmost marker with corresponding 95% CI nominal and chromosomewise type I error rates are given in Table 1. According to the increasing amount of information utilized by four QTL mapping models we observe an increase in the significance, so that the lowest nominal type I error rate for LBM is 0.00134, whereas a corresponding value obtained for MIM is much lower being equal to 0.000004. As the main goal of the current study is the localization of chromosomal regions of interest for fine mapping, genomewise

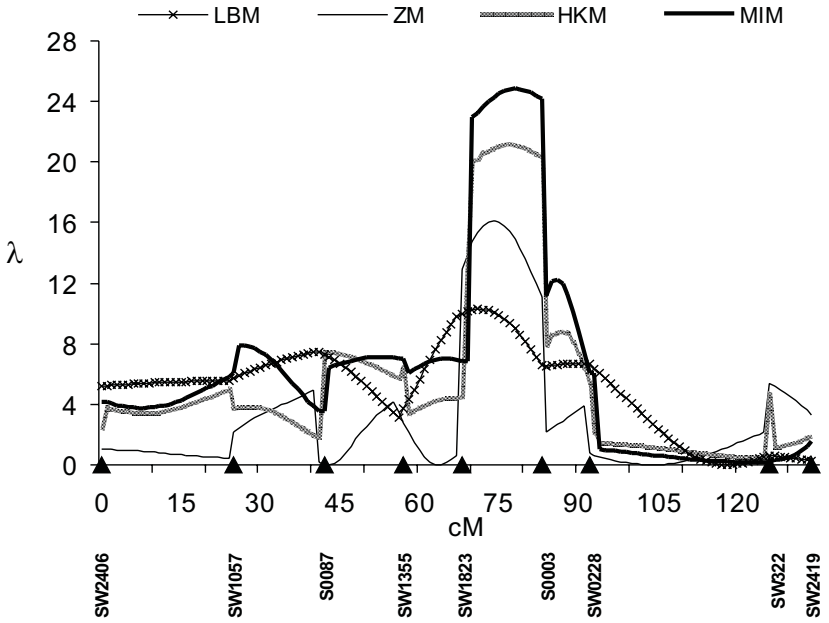


Figure 1. Likelihood profiles for IMF based on LBM, ZM, HKM, and MIM. Triangles represent marker locations.

**Table 1**. QTL mapping results for IMF based on LBM, ZM, HKM, and MIM. Estimates of QTL position in cM from the leftmost marker. 95% confidence intervals for $\hat{\theta}$ based on i) a numerical approximation of $\sigma_{\hat{\theta}}$ (superscript A), ii) the Jackknife estimate of $\sigma_{\hat{\theta}}$ (superscript J), iii) and the support interval (superscript S). Likelihood ratio test values with nominal ($\alpha_N$) and chromosomewise ($\alpha_C$) type I error rates

| Method | QTL position estimate [cM] | Confidence intervals for QTL position | Likelihood ratio test value | $\alpha_N$ | $\alpha_C$ |
|---|---|---|---|---|---|
| LBM | 71 | 62-80 $^A$<br>60-82 $^J$<br>65-79 $^S$ | 10.28 | 0.00134 | 0.009 |
| ZM | 74 | 69-79 $^A$<br>64-84 $^J$<br>70-79 $^S$ | 16.12 | 0.00006 | 0.003 |
| HKM | 77 | 70-84 $^A$<br>61-93 $^J$<br>70-83 $^S$ | 21.20 | 0.00002 | 0.001 |
| MIM | 78 | 71-85 $^A$<br>66-90 $^J$<br>70-83 $^S$ | 24.86 | 0.000004 | < 0.01 |

Note on symbols – as above.

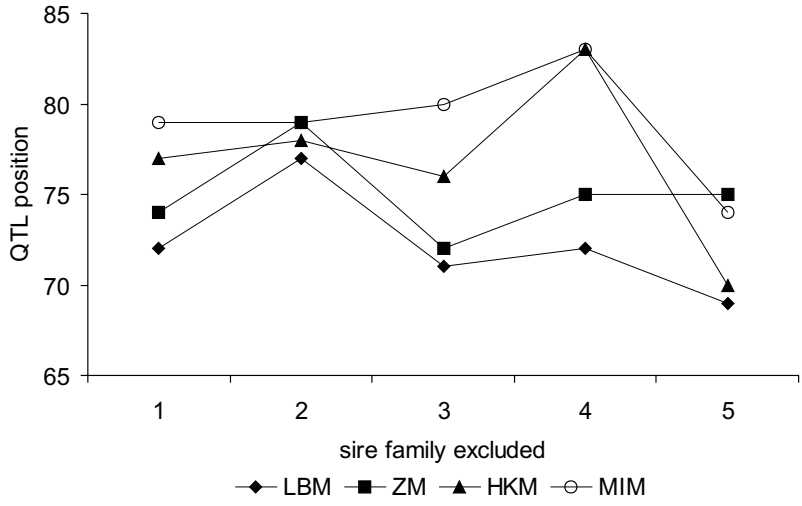


Figure 2. Jackknife plots of the estimated QTL position

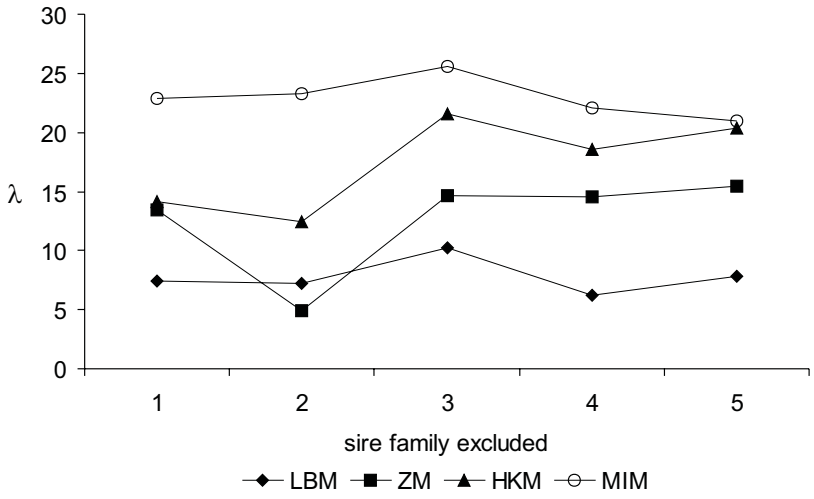


Figure 3. Jackknife plots of the test statistics

significance levels are not reported here, since the precise assessment of significance is not crucial on this stage. The QTL position from the leftmost marker estimated by the four methods varies from 71 cM for LBM to 78 cM for MIM, but remains within the same marker interval. There is a considerable difference in the length of CI between the three approaches applied. For all the four QTL map-

ping methods, using the Jackknife estimate of $\sigma_{\hat{\theta}}$ resulted in the longest CI, covering on average 24.5 cM. The support intervals and CI based on the numerical approximation of $\sigma_{\hat{\theta}}$ were of similar length with the average of 12.3 and 14.0 cM, respectively. Although there were no striking differences in CI length between QTL mapping methods, ZM always corresponded to the shortest estimate. In general, SW1823-S0003-S0228 are the marker intervals included in most of the CI.

Jackknife plots (Figure 2) show variation in the estimates of QTL positions as influenced by the sire family. All the methods show a similar pattern of changes in the QTL position between Jackknife data sets. The differences between the leftmost and the rightmost estimates of a QTL position are large, amounting to 3, 7, 13, and 9 cM, respectively for LBM, ZM, HKM, and MIM. Jackknife plots of the test statistic (Figure 3) visualize differences in the information content between halfsib families. Regardless of the applied statistical model, family 2 was the most informative as its exclusion from the data set resulted in the largest drop of $\lambda$. The opposite pattern is found for family 3. Still, the results remain highly significant for almost each Jackknife data set. LBM and MIM are the most robust, showing less variation in $\lambda$ as compared to ZM and HKM.

The distribution of $\lambda$ under the null hypothesis shows a slight unconservative tendency as compared to the asymptotic $\chi^2_{2df}$ distribution. For the assumed theoretical type I error probability of 0.05 the empirical type I error of 0.075 was obtained. Even for a relatively small data set of 305 individuals, MIM possesses 0.71 power to detect a QTL of 0.5 $\sigma_a$. Whereas power of 0.44 for mapping a moderate-sized QTL is still satisfactory, the identification of effects under $0.3\sigma_a$ has much lower probability (e.g. 0.12 for $0.1\sigma_a$). However, more simulation repetitions are needed to provide estimates of type I error and power with greater accuracy.

## Discussion

In the statistical analysis of line cross experiments the influence of polygenic effects was either ignored (e.g. LANDER, BOTSTEIN 1989, HALEY, KNOTT 1992) or corrected prior to QTL mapping (e.g. DE KONING et al. 1999). Other models attempt to incorporate the effects of other loci by using additional available markers as cofactors (ZENG, 1994, SILLANPÄÄ, ARJAS 1998). Further extensions are proposed by XU (1998), who incorporates differences between families by introducing a family-specific QTL effect and models the error (co)variances accordingly. Also the polygenic component is partially accounted for in this model through a fixed family effect. Recently, PÉREZ-ENCISO and VARONA (2000) proposed a mixed model conceptually similar to MIM, as both models introduce (co)variances between individuals based on the polygenic component. While

PÉREZ-Enciso and VARONA model these (co)variances conditionally on the available marker information, separately for each of arbitrarily defined chromosome segments, MIM uses values averaged over all polygenes, assuming that the polygenic variance in $F_2$ individuals is $\frac{\sigma^2_{\text{sire}} + \sigma^2_{\text{dam}}}{2}$. However, instead of fixing the polygenic variance components at values estimated prior to QTL mapping it would be more accurate in MIM to reestimate them at each step along the chromosome. Preliminary analysis showed that both procedures give almost identical likelihood values, but the methods differ considerably in computing time. Contrary to PÉREZ-ENCISO and VARONA and similarly to the approach presented by SZYDA et al. (2000), the estimation procedure used by NAGAMINE and HALEY (2001) is based on the assumption that the polygenic (co)variance is independent on the QTL position tested and thus the same estimates are used in the model across the whole chromosome. The core difference between both approaches is in terms of hypothesis testing, since NAGAMINE and HALEY (2001) use the F statistics and SZYDA et al. (2000) the likelihood ratio test. The F test involving only one of the model parameters ($\sigma^2_e$) is likely to provide biased results when $\hat{\sigma}^2_e$ is biased, as shown in Table 5 of NAGAMINE and HALEY (2001). Thus, using $\lambda$ involving loglikelihood values related to all the model parameters seems to be more robust towards inaccurate or biased estimation.

In the analysis of real data the most notable approximation underlying all the models used is the assumption that both parental lines are fixed for alternative QTL alleles. As already mentioned by DE KONING et al. (1999), departure from this assumption in a real data set results in a lower power and biased estimates of QTL effects. One should be aware that it might also affect the accuracy of the estimated QTL position, as already indicated by Jackknife plots in Figure 2. On the other hand, BIDANEL et al. (2001), while analysing growth and fatness data from 1103 F2 individuals from outbred line cross, found that a line cross model remains robust towards departures from the above assumption. In the current application of HKM and MIM we try to account for such an unrealistic assumption by the appropriate modelling $P(Qq_{Si})$, as described above.

The power of the mixed model studied by PÉREZ-ENCISO and VARONA (2000) for a variety of inheritance modes was practically 1 (over 30 repetitions). This is higher than the power of MIM obtained in the current study, but we assumed fewer $F_2$ individuals and much lower QTL effects. The advantages of the incorporation of a polygenic component are well demonstrated by VISSCHER and HALEY (1996b) who report elevated type I errors by fitting a major gene model to data with a polygenic variation and reduced power by fitting a multiple QTL model using markers as cofactors to the data with a single QTL.

All the four methods used in our study are in agreement by indicating an interval between markers SW1823 and S0228 as the most probable region for locating a QTL for IMF. The difference in significance between the methods reflects dif-

ferent amounts of information incorporated in the underlying models. QTL for IMF on chromosome 6 have been reported in other crosses of outbred populations in pigs (DE KONING et al. 1999, DE KONING et al. 2000, GERBENS et al. 2000, ÓVILO et al. 2000). A genome scan in a cross between Chinese Meishan × commercial Dutch pigs detected a putative QTL for IMF between markers S0003 and SW2419, close to the telomere of the chromosome (DE KONING et al. 1999, GERBENS et al. 2000). Another analysis of data from the same cross, including more markers and using a different statistical model, indicates that the QTL is paternally expressed and has its most likely position in marker interval S0220-S0121, i.e. closer to the centromere (DE KONING et al. 2000). The most likely QTL position of DE KONING et al. (2000) is almost identical to the result of our study. It is noteworthy that based on the data set available in our study it is only possible to investigate the effect of putative QTL alleles transmitted from sires, i.e. paternal QTL expression. A QTL for IMF on porcine chromosome 6 has also been detected in a cross between Iberian × Landrace pigs (ÓVILO et al. 2000). The most likely QTL position in their study was in marker interval S0228-SW1881, which is telomeric in our results, but similar to the positions reported by DE KONING et al. (1999) and GERBENS et al. (2000).

The accurate identification of the QTL position is not possible using the available marker map. However, our main goal was rather to identify CI marking the area of SSC6, which is going to be subjected to fine mapping in a follow-up study. Several approaches exist to determine CI for the QTL position. These can be roughly categorized as: i) methods based on resampling techniques, i.e. the bootstrap and the Jackknife, ii) methods extracting information from the curvature of the likelihood profile in the neighbourhood of the most probable QTL position, and iii) methods based on asymptotic properties of $\lambda$. Based on the asymptotic properties of the test statistic, DUPUIS and SIEGMUND (1999) proposed CI which are independent of QTL effects, but used an unrealistic assumption that the QTL is located at the marker. A similar approach of CIERCO and MANGIN (1996) incorporates the information on the distance between markers. Here we construct CI based on the Jackknife, the numerical approximation and the support interval approaches. We found that the CI based on Jackknife are the longest. This may result from a poor approximation of the variance of a QTL position estimate because of a halfsib family chosen as a resampling component, which is a rather rough unit. Choosing a finer unit, e.g. a fullsib family, would result in a better approximation, but at the same time will be more computationally intensive. The numerical approximation approach and support interval provide similar, shorter CI and are very easy to compute. Comparing statistical properties of various CI, DUPUIS and SIEGMUND (1999) also found support interval as the most exact approach, robust towards changes in sample size and a QTL position relative to flanking markers.

As pointed out by DE KONING et al. (1999), the practical application of currently available methods based on the mixed inheritance model is very limited.

We show here that MIM is well feasible to evaluate multimarker data sets, even by using commonly available software such as the SAS and, at the same time, to utilize more of the genetic information than, the classical interval mapping models underlying LBM and HKM.

## REFERENCES

ANDERSSON-EKLUND L., MARKLUND L., LUNDSTRÖM K., HALEY C.S., ANDERSSON K., HANSSON I., MOLLER M., ANDERSSON L. (1998). Mapping quantitative trait loci for carcass and meat quality traits in a Wild sire × Large White intercorss. J. Anim. Sci. 76: 694-700.

BASTEN C.J., WEIR B.S., ZENG Z.-B. (1994). Computing strategies and software. In: Proceedings of the 5th World Congress on Genetics Applied to Livestock Production (C. Smith, J.S. Gavora, B. Benkel, J. Chesnais, W. Fairfull, J.P. Gibson, B.W. Kennedy and E.B. Burnside, eds.). Organizing Commitee, 5th World Congress on Genetics Applied to Livestock Production, Guelph, Ontario, Canada 22: 65-66.

BIDANEL J.-P., MILAN D., IANNUCCELLI N., AMIGUES Y., BOSCHER M.-Y., BOURGEOIS F., CARITEZ J.-C., GRUAND J., LE ROY P., LAGANT H., QUINTANILLA R., RENARD C., GELLIN J., OLLIVIER L., CHEVALET C. (2001). Detection of quantitative trait loci for growth and fatness in pigs. Genet. Sel. Evol. 33: 289-309.

CHURCHILL G.A., DOERGE R.W. (1994). Empirical threshold values for quantitative traits mapping. Genetics 138: 963-971.

CIERCO C., MANGIN B. (1996). Construction of confidence intervals for quantitative trait loci location in $F_2$ populations. Biometrics 52: 1268-1282.

DE KONING D.J., JANSS L.L.G., RATTINK A.P., VAN OERS P.A.M., DE VRIES B.J., GROENEN M.A.M., VAN DER POEL J.J., DE GROOT P.N., BRASCAMP E.W., VAN ARENDONK J.A.M. (1999). Detection of quantitative trait loci for backfat thickness and intramuscular fat content in pigs (*Sus scrofa*). Genetics 152: 1679-1690.

DE KONING D.J., RATTINK A.P., HARLIZIUS B., VAN ARENDONK J.A.M., BRASCAMP E.W., GROENEN M.A.M. (2000). Genome-wide scan for body composition in pigs reveals important role of imprinting. Proc. Natl. Acad. Sci. USA 97: 7947-7950.

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

EFRON B., TIBSHIRANI R.J. (1993). An introduction to the Bootstrap. Chapman&Hall, New York.

GERBENS F., DE KONING D.J., HARDERS F.L., MEUWISSEN T.H.E., JANSS L.L.G., GROENEN M.A.M., VEERKAMP J.H., VAN ARENDONK J.A.M., TE PAS M.F.W. (2000). The effect of adipocyte and heart fatty acid-binding protein genes on intramuscular fat and backfat content in Meishan crossbred pigs. J. Anim. Sci. 78: 552-559.

GRINDFLEK E., SZYDA J., LIU Z., LIEN S. (2000). Detection of quantitative trait loci for meat quality in a commercial slaughter pig cross. submitted to Mamm. Genome 12: 299-304.

HALEY C.S., KNOTT S.A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69: 315-324.

HENDERSON C.R. (1984). Applications of Linear Models in Animal Breeding. University of Guelph Press, Guelph.

JANSEN R.C. (1993). Interval mapping of multiple quantitative trait loci. Genetics 135: 205-211.

JIANG C., ZENG Z.-B. (1995). Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics 140: 1111-1127.

LANDER E.S., BOTSTEIN D. (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185-199.

LIU B.H. (1998). Statistical Genomics. CRC Press. Boca Raton, New York.

MEYER K., HILL W.G. (1992). Approximation of sampling variances and confidence intervals for maximum likelihood estimates of variance components. J. Anim. Breed. Genet. 109: 264-280.

NAGAMINE Y., HALEY C.S. (2001). Using the mixed model for interval mapping of quantitative trait loci in outbred line crosses. Genet. Res. 77: 199-207.

ÓVILO C., PÉREZ-ENCISO M., BARRAGÁN C., CLOP A., RODRÍGUEZ C., OLIVER M.A., TORO M.A., NOGUERA J.L. (2000). A QTL for intramuscular fat and backfat thickness is located on porcine Chromosome 6. Mamm. Genome 11: 344-346.

PÉREZ-ENCISO M., VARONA L. (2000). Quantitative trait loci mapping in $F_2$ crosses between outbred lines. Genetics 155: 391-405.

RATHJE T.A., ROHRER G.A., JOHNSON R.K. (1997). Evidence for quantitative trait loci affecting ovulation rate in pigs. J. Anim. Sci. 75: 1486-1494.

SAS Institute Inc. (1996). SAS/STAT Software: Changes and Enhancements through Release 6.11, SAS Institute Inc., Cary, NC.

SCHELER P., MANGIN B., GOFFINET B., LEROY P., BOICHARD D., ELSEN J.M. (1998). Properties of a Bayesian approach to detect QTL compared to the flanking markers regression method. J. Anim. Breed. Genet. 115: 87-95.

SILLANPÄÄ M.J., ARJAS E. (1998). Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. Genetics 148: 1373-1388.

SOLLER M., BORDY T., GENIZI A. (1976). On the power of experimental design for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theor. Appl. Genet. 47: 35-39.

SZYDA J., LIU Z., WILD V. (1998). Properties of the transmission-disequilibrium test to the detection of quantitative trait loci in livestock. J. Appl. Genet. 39: 309-319.

SZYDA J., GRINDFLEK E., LIU Z., LIEN S. (2000). Dissection of Genetic Background Underlying Meat Quality Traits in Swine. Quality of Meat and Fat in Pigs as Affected by Genetics and Nutrition, Wageningen Pers, Wageningen.

VISSCHER P.M., HALEY C.S. (1996a). Detection of putative quantitative trait loci in line crosses under infinitesimal genetic models. Theor. Appl. Genet. 93: 691-702.

VISSCHER P.M., THOMPSON R., HALEY C.S. (1996b). Confidence intervals in QTL mapping by bootstrapping. Genetics 143: 1013-1020.

XU S. (1998). Mapping quantitative trait loci using multiple families of line crosses. Genetics 148: 517-524.

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