

# Alternative models for QTL detection in livestock. III. Heteroskedastic model and models corresponding to several distributions of the QTL effect

Bruno Goffinet <sup>a</sup>, Pascale Le Roy <sup>b</sup>, Didier Boichard <sup>b</sup>,  
Jean Michel Elsen <sup>c\*</sup>, Brigitte Mangin <sup>a</sup>

<sup>a</sup> Biométrie et intelligence artificielle, Institut national  
de la recherche agronomique, BP27, 31326 Castanet-Tolosan, France

<sup>b</sup> Station de génétique quantitative et appliquée, Institut national  
de la recherche agronomique, 78352 Jouy-en-Josas, France

<sup>c</sup> Station d'amélioration génétique des animaux, Institut national  
de la recherche agronomique, BP27, 31326 Castanet-Tolosan, France

(Received 20 November 1998; accepted 22 April 1999)

**Abstract** – This paper describes two kinds of alternative models for QTL detection in livestock: an heteroskedastic model, and models corresponding to several hypotheses concerning the distribution of the QTL substitution effect among the sires: a fixed and limited number of alleles or an infinite number of alleles. The power of different tests built with these hypotheses were computed under different situations. The genetic variance associated with the QTL was shown in some situations. The results showed small power differences between the different models, but important differences in the quality of the estimations. In addition, a model was built in a simplified situation to investigate the gain in using possible linkage disequilibrium. © Inra/Elsevier, Paris  
**half-sib families / heteroskedastic model / linkage disequilibrium / QTL detection**

**Résumé** – Modèles alternatifs pour la détection de QTL dans les populations animales. III. Modèle hétéroscédastique et modèles correspondant à différentes distributions de l'effet du QTL. Ce papier décrit deux types de modèles alternatifs pour la détection de QTL dans les populations animales : un modèle hétéroscédastique

\* Correspondence and reprints  
E-mail: elsen@toulouse.inra.fr

d'une part, et des modèles correspondants à différentes hypothèses sur la distribution de l'effet de substitution du QTL pour chaque mâle : un nombre fixe et limité d'allèles ou au contraire un nombre infini d'allèles. Les puissances des différents tests construits avec ces hypothèses sont calculées dans différentes situations. L'estimation de la variance génétique liée au QTL est donnée dans certaines situations. Les résultats montrent de faibles différences de puissance entre les différents modèles, mais des différences importantes dans la qualité des estimations. De plus, on construit un modèle dans une situation simplifiée pour étudier le gain que l'on peut obtenir en utilisant un éventuel déséquilibre de liaison. © Inra/Elsevier, Paris

**familles de demi-frères / modèle hétéroscédastique / déséquilibre de liaison / détection de QTL**

## 1. INTRODUCTION

In theoretical papers dealing with QTL detection in livestock, the QTL effects are most often considered to be different across the sires  $i$ , and the residual variance within the QTL genotype as constant among the sires (e.g. [9, 10]). These hypotheses were made in the two previous papers about alternative models for QTL detection in livestock [4, 8]. In this third paper, these two sets of parameters are studied.

First, a heteroskedastic model with residual variance  $\sigma_{e,i}^2$  specific to each sire  $i$  is evaluated. The rationale for this test is that it should be more robust against true heteroskedasticity, for instance when different alleles are segregating at another QTL than the QTL under consideration. However, the power of the tests may be smaller than in the homoskedastic model if the homoskedastic model is correct.

Different possibilities concerning the within sire QTL substitution effect  $\alpha_i^x$  will also be considered: a fixed and limited number of alleles, or an infinite number of alleles. Taking into account these distributions of the QTL effect can increase the power of the tests if the model is correct, and decrease this power if the model is incorrect. Therefore, the behaviour of the tests based on these different models will be compared under different situations concerning the distribution of the QTL effect. More specifically, the case of a biallelic QTL in linkage disequilibrium with the marker, will be explored in greater detail.

Jansen et al. [6] also considered the same kind of model concerning the residual variances and the number of alleles, but did not compare the power of the tests. Coppieters et al. [3] also considered these kinds of models and compared the power of regression analysis and of a non-parametric approach.

Most hypotheses and notations are given in Elsen et al. [4]. To simplify the computations, all the comparisons were made using the most probable sire genotype  $\widehat{hs}_i = \operatorname{argmax}_{hs_i} p(hs_i | M_i)$  and the linearised approximation of the likelihood described in the previous paper. All the simulations were made with 5 000 replications, and the length of the confidence interval for the simulated power was smaller than 1 %. When an analytical solution could not be found, we used a quasi Newton algorithm to compute the maximum likelihood. The chromosome length was 1 Morgan, with 3 or 11 markers, equally spaced, each with two alleles segregating at an equal frequency in the population.

## 2. EVALUATION OF A HETEROSKEDASTIC MODEL

In this section, the power of the  $T^2$  test built under a homoskedastic model [8] will be compared to the power of the  $T^6$  test built under a heteroskedastic model, where  $\sigma_{e,i}^2$  is used in place of  $\sigma_e^2$  in the likelihood  $\tilde{\Lambda}^{x,hs}$ . This comparison will be made for both homoskedastic and heteroskedastic situations. The heteroskedastic situation will be modelled assuming the existence of an independent QTL, i.e. located on another chromosome. This QTL is assumed to be biallelic, with balanced frequencies (0.5) in the sire population and with an additive effect. Dams are homozygous for this QTL. Under this hypothesis, the within offspring residual variance is lower for sires homozygous for this QTL than for the heterozygous sire. Powers were calculated considering an  $H_0$  rejection threshold corresponding to a correct type I error, which is computed in the same situation, homoskedastic or heteroskedastic, with no QTL on the tested chromosome.

Table I concerns true homoskedastic situations, with a residual variance  $\sigma_e^2 = 1$ . In this table, the power of the  $T^2$  and  $T^6$  tests are given for different values of the number of progeny per sire (20 or 50), of the number of markers in the different linkage group (3 or 11), of the position of the QTL (0.05 or 0.35) and of the additive effect of the QTL ( $a = 0.5$  or 1). The two possible QTL alleles thus had the same probability. Note that in this case, the QTL substitution effect equals the QTL additive effect.

**Table I.** Power of the  $T^2$  (homoskedastic) and  $T^6$  (heteroskedastic) tests in situations with  $\sigma_{e,i}^2 = \sigma_e^2$ , for 20 sires, different values of the number of progeny per sire ( $md$ ), of the number of markers ( $L$ ), of the position ( $x$ ) and of the additive effect ( $a$ ) of the QTL under consideration.

| $md$ | $L$ | $x$  | $a$ | $T^2$ | $T^6$ |
|------|-----|------|-----|-------|-------|
| 50   | 11  | 0.05 | 0.5 | 49.8  | 46.1  |
| 50   | 11  | 0.35 | 0.5 | 60.8  | 57.2  |
| 50   | 11  | 0.05 | 1.0 | 96.6  | 95.4  |
| 50   | 11  | 0.35 | 1.0 | 98.7  | 98.1  |
| 20   | 11  | 0.05 | 0.5 | 18.1  | 15.5  |
| 20   | 11  | 0.35 | 0.5 | 19.4  | 18.1  |
| 20   | 11  | 0.05 | 1.0 | 58.2  | 53.0  |
| 20   | 11  | 0.35 | 1.0 | 70.2  | 63.3  |
| 50   | 3   | 0.05 | 0.5 | 26.1  | 25.9  |
| 50   | 3   | 0.35 | 0.5 | 24.3  | 24.2  |
| 50   | 3   | 0.05 | 1.0 | 78.2  | 74.1  |
| 50   | 3   | 0.35 | 1.0 | 71.1  | 67.3  |
| 20   | 3   | 0.05 | 0.5 | 11.9  | 11.1  |
| 20   | 3   | 0.35 | 0.5 | 10.4  | 10.7  |
| 20   | 3   | 0.05 | 1.0 | 29.7  | 27.8  |
| 20   | 3   | 0.35 | 1.0 | 25.9  | 24.8  |

Tables II and III concern true heteroskedastic situations. A QTL located on another chromosome was simulated with an  $a_2$  effect. The thresholds of the  $T^2$  and  $T^6$  tests are given in table II for different values of the  $a_2$  effect and for 20 sires, 50 progeny per sire and 11 markers. The results were obtained with 5 000 simulations. The power of the  $T^2$  and  $T^6$  tests are given in table III for different values of the linked QTL additive effect ( $a = 0.5$  or  $1.0$ ), of the position of this linked QTL ( $x = 0.05$  or  $0.35$ ) and of the independent QTL additive effect ( $a_2 = 0, 1, 1.5$  or  $2$ ). For each QTL, the two possible alleles had the same probability.

**Table II.** Thresholds of the  $T^2$  (homoskedastic) and  $T^6$  (heteroskedastic) tests in situations where  $\sigma_{e,i}^2$  depends on  $i$ , owing to segregation of an unlinked QTL with additive effect  $a_2$ . Number of progeny per sire  $md = 50$ , number of markers  $L = 11$ , number of sires = 20 and different values for  $a_2$  and the corresponding extreme values among the sires for  $\sigma_{e,i}^2$ .

| $a_2$ | $\min\sigma_i^2$ | $\max\sigma_i^2$ | $T^2$ | $T^6$ |
|-------|------------------|------------------|-------|-------|
| 0.0   | 1                | 1                | 38.2  | 39.2  |
| 1.0   | 1                | 2                | 38.3  | 39.2  |
| 1.5   | 1                | 3.25             | 38.8  | 39.4  |
| 2.0   | 1                | 5                | 39.6  | 39.1  |

In the true homoskedastic situation, and for a given number of sires and markers, the thresholds of the two tests appear to be very close to each other for all cases (data not shown), which is in agreement with the asymptotic theory in linear models. In a linear model, the asymptotic distribution of Fisher test statistic is the same if the residual variance used in the denominator is replaced by any consistent estimate of this variance. The estimate of the residual variances in the model corresponding to the  $T^6$  test is consistent, as is the estimate in the other model. The thresholds given in table II show that the  $T^6$  test is not sensitive at all to the value of  $a_2$ , whereas  $T^2$  is slightly more sensitive. The use of the threshold corresponding to  $a_2 = 0$  when it is not true can lead to a first type error of 5.5 % instead of 5 %.

The power of the  $T^6$  test appears to be only slightly smaller than the power of the  $T^2$  test in the case of  $\sigma_{e,i} = \sigma_e$ . This very small decrease is in agreement with the difference in power of an analysis of variance test when the number of degrees of freedom of the residual varies from 50 to 1 000, i.e. from the number of progeny per sire to the total number of progeny.

The power of the  $T^6$  test is slightly larger than that of the  $T^2$  test only in cases where the QTL leading to heteroskedasticity has a large effect. Even in these cases, the differences between the power of the two tests remain small and of the same order as for homoskedastic situations, but with the opposite sign.

From these results, and considering that the tests based on the heteroskedastic model take a little less time to compute (about 5 %), the following tests will be based on this model.

**Table III.** Power of the  $T^2$  (homoskedastic) and  $T^6$  (heteroskedastic) tests in situations where  $\sigma_i^2$  depends on  $i$ , owing to the segregation of an unlinked QTL with additive effect  $a_2$ . Number of progeny per sire  $md = 50$ , number of markers  $L = 11$ , different values for  $a_2$  and the corresponding extreme values among the sires for  $\sigma_{e,i}^2$ , and different values for the position ( $x$ ) and the additive effect ( $a$ ) of the QTL under consideration.

| $a_2$ | $\min\sigma_i^2$ | $\max\sigma_i^2$ | $x$  | $a$ | $T^2$ | $T^6$ |
|-------|------------------|------------------|------|-----|-------|-------|
| 0.0   | 1                | 1                | 0.05 | 0.5 | 49.8  | 46.1  |
| 0.0   | 1                | 1                | 0.35 | 0.5 | 60.8  | 57.2  |
| 0.0   | 1                | 1                | 0.05 | 1.0 | 96.6  | 95.4  |
| 0.0   | 1                | 1                | 0.35 | 1.0 | 98.7  | 98.1  |
| 1.0   | 1                | 2                | 0.05 | 0.5 | 44.2  | 43.4  |
| 1.0   | 1                | 2                | 0.35 | 0.5 | 54.4  | 52.8  |
| 1.0   | 1                | 2                | 0.05 | 1.0 | 97.6  | 96.8  |
| 1.0   | 1                | 2                | 0.35 | 1.0 | 99.2  | 99.0  |
| 1.5   | 1                | 3.25             | 0.05 | 0.5 | 36.7  | 38.5  |
| 1.5   | 1                | 3.25             | 0.35 | 0.5 | 47.4  | 48.2  |
| 1.5   | 1                | 3.25             | 0.05 | 1.0 | 95.6  | 95.3  |
| 1.5   | 1                | 3.25             | 0.35 | 1.0 | 98.3  | 98.3  |
| 2.0   | 1                | 5                | 0.05 | 0.5 | 28.1  | 34.8  |
| 2.0   | 1                | 5                | 0.35 | 0.5 | 38.2  | 42.5  |
| 2.0   | 1                | 5                | 0.05 | 1.0 | 92.1  | 93.4  |
| 2.0   | 1                | 5                | 0.35 | 1.0 | 96.3  | 96.7  |

### 3. VARIOUS NUMBERS OF ALLELES AT THE QTL LOCUS

In the previous papers [4, 8], QTL substitution effects  $\alpha_i^x$  were defined within with each sire  $i$ . In this paper, two possible alternative situations concerning these effects are considered.

– A limited number of QTL alleles, and therefore a set of only a few possible values for  $\alpha_i^x$ . In this case, the parameters are these values and the probability of QTL genotypes. This is the model used by Knott et al. [7].

– An infinite number of possible values, drawn at random in a normal distribution. This is the model used by Grignola et al. [5].

In these two situations, we will consider that the QTL effects are independently and identically distributed between the sires.

In the two cases, the linearised version of the likelihood can be written as:

$$\tilde{\Lambda}_H^{x, \widehat{hs}}(f) = \prod_i \int_{\alpha_i^x} f(\alpha_i^x) \prod_{j=1}^{n_i} \phi(y p_{ij}; (2p(d_{ij}^x = 1 | \widehat{hs}_i, M_i) - 1) \frac{\alpha_i^x}{2} + \mu_i, \sigma_{e,i}^2) d\alpha_i^x$$

where  $f(\alpha_i^x)$  is the density of the distribution of  $\alpha_i^x$ .

In the situation with two possible alleles at the QTL locus, the likelihood becomes:

$$\begin{aligned} \tilde{\Lambda}_H^{x,\widehat{hs}}(\text{two alleles}) &= \prod_i (p_\alpha^x \prod_{j=1}^{n_i} \phi(y p_{ij}; (2p(d_{ij}^x = 1|\widehat{hs}_i, M_i) - 1) \frac{\alpha}{2} + \mu_i, \sigma_{e,i}^2)) \\ &+ p_\alpha^x \prod_{j=1}^{n_i} \phi(y p_{ij}; (2p(d_{ij}^x = 1|\widehat{hs}_i, M_i) - 1) \frac{-\alpha}{2} + \mu_i, \sigma_{e,i}^2) \\ &+ (1 - 2p_\alpha^x) \prod_{j=1}^{n_i} \phi(y p_{ij}; \mu_i, \sigma_{e,i}^2) \end{aligned}$$

where  $p_\alpha^x = p(\alpha_i^x = \alpha) = p(\alpha_i^x = -\alpha)$  and  $\alpha$  are the two parameters of the distribution.

In the situation with a normal distribution of the QTL effect, the density  $f(\alpha_i^x)$  is the normal density  $\phi(\alpha_i^x; 0, \sigma_\alpha^2)$  and the likelihood is written as  $\tilde{\Lambda}_H^{x,\widehat{hs}}$  (normal).

The test built with the likelihood  $\tilde{\Lambda}_H^{x,\widehat{hs}}$  (two alleles) will be  $T^7$  and the test built with the likelihood  $\tilde{\Lambda}_H^{x,\widehat{hs}}$  (normal),  $T^8$ .

In *table IV*,  $T^7$  and  $T^8$  test thresholds are given for different situations concerning the number of markers and the number of progeny per sire. In *table V*, the power of the  $T^6$ ,  $T^7$  and  $T^8$  tests are presented for two kinds of situations. In the first, the QTL had two possible equiprobable ( $p_a = 1/2$ ) alleles with no dominance and an additive effect  $a$ . The QTL substitution effect  $a_i$  for each sire  $i$  is therefore 0 with a probability of 1/2 and  $a$  with a probability of 1/2. We have  $E(a_i^2) = a^2/2$ . The QTL variance due to the sire in the progeny of  $i$  is  $a_i^2/4$ , and globally  $\sigma_s^2 = E(a_i^2/4) = a^2/8$ . In the second, the effect of each value  $a_i$  was drawn at random in a normal distribution,  $\sigma_a^2 = a^2/2$  of null expectation and variance. Therefore,  $E(a_i^2) = a^2/2$  and  $\sigma_s^2 = E(a_i^2/4) = a^2/8$  as in the first case. The results are presented for different values of the parameters.

**Table IV.** Thresholds of the  $T^7$  (two alleles) and  $T^8$  (normal distribution) tests, for different values of the number of progeny per sire ( $md$ ), and the number of markers ( $L$ ).

| $md$ | $L$ | $T^7$ | $T^8$ |
|------|-----|-------|-------|
| 50   | 11  | 6.21  | 4.69  |
| 20   | 11  | 7.54  | 5.92  |
| 50   | 3   | 5.49  | 4.20  |
| 20   | 3   | 6.98  | 4.84  |

It is interesting to note that the thresholds are appreciably smaller than the thresholds presented in *table II*. This is due to the fact that there is only one parameter for the QTL effect in  $T^7$  and  $T^8$ , and 20 in  $T^6$ . The differences between the two kinds of thresholds can be compared with the differences between the  $\chi_{1ddl}^2$  95 % quantile, 3.84, and the  $\chi_{20ddl}^2$  95 % quantile, 31.41.

**Table V.** Powers of the  $T^6$  (heteroskedastic),  $T^7$  (2 alleles) and  $T^8$  (normal distribution) tests, for different values of the number of progeny per sire ( $md$ ), of the number of markers ( $L$ ) and of the additive effect of the QTL ( $a$  or  $\sigma_a^2$ ). The value  $\sigma_a^2 = 1/8$  (resp.  $1/2$ ) is the variance generated by a QTL with an additive effect  $a = 0.5$  (resp. 1).

|                      | $md$ | $L$ | $a$ or $\sigma_a^2$ | $T^6$ | $T^7$ | $T^8$ |
|----------------------|------|-----|---------------------|-------|-------|-------|
| Two alleles          | 50   | 11  | 0.5                 | 57.2  | 55.0  | 56.9  |
|                      | 50   | 3   | 0.5                 | 24.2  | 19.9  | 22.0  |
|                      | 20   | 11  | 0.5                 | 18.1  | 16.5  | 16.5  |
|                      | 20   | 3   | 0.5                 | 10.5  | 8.6   | 9.7   |
|                      | 50   | 11  | 1.0                 | 98.1  | 97.7  | 96.9  |
|                      | 50   | 3   | 1.0                 | 67.3  | 63.3  | 61.2  |
|                      | 20   | 11  | 1.0                 | 63.3  | 57.7  | 55.9  |
|                      | 20   | 3   | 1.0                 | 24.8  | 19.2  | 21.4  |
| Normally distributed | 50   | 11  | 1/8                 | 57.8  | 54.4  | 56.6  |
|                      | 50   | 3   | 1/8                 | 21.3  | 19.6  | 20.2  |
|                      | 20   | 11  | 1/8                 | 18.6  | 17.3  | 16.9  |
|                      | 20   | 3   | 1/8                 | 11.3  | 8.5   | 10.5  |
|                      | 50   | 11  | 1/2                 | 98.0  | 97.0  | 96.5  |
|                      | 50   | 3   | 1/2                 | 63.0  | 58.1  | 59.1  |
|                      | 20   | 11  | 1/2                 | 58.1  | 53.6  | 51.9  |
|                      | 20   | 3   | 1/2                 | 24.0  | 18.7  | 22.2  |

The main and quite strange result was that the power of  $T^6$  is always larger than or equal to the power of the other tests.

In order to compare the  $T^6$  and  $T^7$  tests more thoroughly when the model really has two alleles, a very large number of simulations were performed in a simplified situation. A very informative marker, linked totally to the QTL was assumed to exist, and the residual variance was assumed to be known (20 sires and 50 progeny per sire). The  $T^6$  and  $T^7$  tests were simplified accordingly. The  $T^6$  test was found to be more powerful (with a difference of 3–4 %) than the  $T^7$  test for  $0.1 < p_a^x < 0.9$ , and  $T^7$  was more powerful (with the same differences) than  $T^6$  for the other values of  $p_a^x$ . This confirms that the loglikelihood ratio test is not the more powerful test in mixture situations, for all values of the alternative parameters. Andrews and Ploberger [1, 2] showed that the loglikelihood ratio test is admissible but not optimal in cases, such as mixture models, where a parameter disappears under the null hypothesis (here the probability of having one of the two alleles). We tried a value  $p_a = 0.05$  in the general framework with  $md = 50$ ,  $L = 11$ ,  $a = 0.5$ , but unfortunately the  $T^6$  test remains more powerful (with a difference of 2 %) than the  $T^7$  test.

Concerning the comparison between  $T^6$  and  $T^8$  in situations where the QTL effect is normally distributed, it is clear in such simple and balanced situations that both  $T^6$  and  $T^8$  are asymptotically equivalent to the test based on the value of  $\sum_i \hat{\alpha}_i^2$  where the  $\hat{\alpha}_i$  are the maximum likelihood estimators of the QTL substitution effect. Therefore, their power should have been quite

the same. The relatively poor performance of  $T^8$  is perhaps partially due to numerical problems, because in some cases (2 %), the algorithm had difficulties in converging and the corresponding simulations were excluded from the results.

The estimation of the QTL variance due to the sire  $\sigma_s^2$  obtained with the different models is shown in *table VI*. With the models used in  $T^6$  and  $T^7$ , this estimation is obtained as a function of the estimates of the  $a_i$  or  $a$ ; with  $T^8$ , it is estimated directly. The value 0.03125 (resp. 0.125) of  $\sigma_s^2$  corresponds to values  $a = 0.5$  and  $\sigma_a^2 = 0.125$  (resp. 1.0 and 0.5).

**Table VI.** Estimation of the sire QTL variance  $\sigma_s^2$  using the  $T^6$  (heteroskedastic),  $T^7$  (two alleles) and  $T^8$  (normal distribution) models, for different values of the number of progeny per sire ( $md$ ), of the number of markers ( $L$ ) and of  $\sigma_a^2$ .

|                      | $md$ | $L$ | $\sigma_s^2$ | $T^6$ | $T^7$ | $T^8$ |
|----------------------|------|-----|--------------|-------|-------|-------|
| Two alleles          | 50   | 11  | 0.03125      | 0.081 | 0.057 | 0.031 |
|                      | 50   | 3   | 0.03125      | 0.37  | 0.069 | 0.032 |
|                      | 20   | 11  | 0.03125      | 0.169 | 0.097 | 0.038 |
|                      | 20   | 3   | 0.03125      | 0.96  | 0.127 | 0.05  |
|                      | 50   | 11  | 0.125        | 0.174 | 0.141 | 0.122 |
|                      | 50   | 3   | 0.125        | 0.49  | 0.161 | 0.103 |
|                      | 20   | 11  | 0.125        | 0.27  | 0.194 | 0.118 |
|                      | 20   | 3   | 0.125        | 1.24  | 0.226 | 0.115 |
| Normally distributed | 50   | 11  | 0.03125      | 0.079 | 0.062 | 0.032 |
|                      | 50   | 3   | 0.03125      | 0.034 | 0.070 | 0.032 |
|                      | 20   | 11  | 0.03125      | 0.167 | 0.097 | 0.039 |
|                      | 20   | 3   | 0.03125      | 0.93  | 0.133 | 0.053 |
|                      | 50   | 11  | 0.125        | 0.98  | 0.157 | 0.120 |
|                      | 50   | 3   | 0.125        | 0.479 | 0.168 | 0.100 |
|                      | 20   | 11  | 0.125        | 0.264 | 0.201 | 0.109 |
|                      | 20   | 3   | 0.125        | 1.24  | 0.23  | 0.111 |

It appears that the estimator obtained using  $T^8$  is the only quite unbiased estimator of  $\sigma_s^2$ . The bias is very large when using the other tests. A practical solution would be to use the simple  $T^6$  test to detect a QTL and to use the estimate associated with  $T^8$  when a QTL is detected.

#### 4. BETWEEN SIRES LINKAGE DISEQUILIBRIUM

To investigate the usefulness of using a model including a linkage disequilibrium between markers and QTL alleles at the between sires level, a simplified situation, which mimics the real situation, but which is considerably easier to compute, was considered.

The QTL is supposed to be located on a marker locus, with all the 20 sires considered  $A, B$  heterozygous for this marker. The dams are considered as carrying other alleles and therefore all the progeny are informative. We denote  $Y_A(i)$  (resp.  $Y_B(i)$ ) the mean of the  $n_A(i)$  (resp.  $n_B(i)$ ) progeny of sire  $i$  carrying allele  $A$  (resp.  $B$ ). The two possible alleles at the QTL are denoted  $Q$ , with an

additive effect of  $a/2$  and  $q$ , with an additive effect  $-a/2$ . The model for the expectation of  $Y_A(i)$  and  $Y_B(i)$  is:

$$E(Y_A(i)) = \mu + a/2 \text{ with a probability } p_1 = p(\text{QTL} = Q/A)$$

$$E(Y_A(i)) = \mu - a/2 \text{ with a probability } 1 - p_1 = p(\text{QTL} = q/A)$$

$$E(Y_B(i)) = \mu + a/2 \text{ with a probability } p_2 = p(\text{QTL} = Q/B)$$

$$E(Y_B(i)) = \mu - a/2 \text{ with a probability } 1 - p_2 = p(\text{QTL} = q/B)$$

The variability around this expectation will be considered as normally distributed, with mean 0 and variance  $\sigma^2/n_A(i)$  (resp.  $\sigma^2/n_B(i)$ ) assumed to be known. We will consider two tests: the analysis of variance test which corresponds to the model  $E(Y_A(i)) - E(Y_B(i)) = a_i$ , without an assumption concerning the distribution of the  $a_i$ , and the likelihood ratio test corresponding to the mixture model concerning the sire allele. The first test is analogous to test  $T^6$  and will be denoted  $T^{6'}$  and the second, analogous to test  $T^7$  will be denoted  $T^{7'}$ . This is only an analogy because the residual variance is assumed to be known, all the progeny are informative and the tests are computed only on the marker.

**Table VII.** Powers of the  $T^{6'}$  and  $T^{7'}$  tests, for an additive QTL effect of  $a = 0.5$ , different values of the linkage disequilibrium, and different values of the number of progeny  $md$ .

| $p_1$     | $p_2$ | $T^{6'}$ | $T^{7'}$ |
|-----------|-------|----------|----------|
| $md = 25$ |       |          |          |
| 0.5       | 0.5   | 58.1     | 51.3     |
| 0.6       | 0.4   | 64.8     | 58.0     |
| 0.7       | 0.3   | 68.5     | 74.0     |
| 0.8       | 0.2   | 76.9     | 90.6     |
| 0.9       | 0.1   | 86.9     | 98.2     |
| $md = 50$ |       |          |          |
| 0.5       | 0.5   | 92.3     | 88.4     |
| 0.6       | 0.4   | 94.3     | 91.5     |
| 0.7       | 0.3   | 96.7     | 97.1     |
| 0.8       | 0.2   | 99.1     | 99.6     |
| 0.9       | 0.1   | 99.7     | 100.0    |

The powers of these two tests for  $\sigma^2 = 1$ ,  $a = 0.5$ , with different numbers of informative progeny  $n_A(i) + n_B(i) = \text{constant}$  across the sires, and different values of the parameters  $p_1$  and  $p_2$ , are given in *table VII*. Note that the 25 informative progeny would correspond to the mean number of informative progeny for 50 dams and a single biallelic marker.

It appears that the use of a model with a linkage disequilibrium can increase the power if there is really a linkage disequilibrium (that is a large difference between  $p_1$  and  $p_2$ ) but can lose power when there is a small linkage disequilibrium. These results depend heavily however on the hypothesis made in this simplified situation.

– QTL location knowledge; this knowledge increases the power of the two tests but perhaps does not change the difference between the two tests.

- The females do not carry either of the sire's alleles; it is not a very realistic situation, but it leads to easier computations and one can think that it does not change the power difference between the two tests.
- The use of a completely linked marker; it is considerably more difficult to build a model with one or several partially linked markers and the gain in using this information would be smaller than the gain presented in *table VII*.

## 5. CONCLUSIONS

In many situations, the power of the simple  $T^6$  test, which is easier and faster to compute, is equal to or a little bit better than the power of the other tests. This result could be specific to QTLs of little effect. In the present study, we focused on QTL effects of such a relatively small magnitude because, with QTLs with larger effects, all the tests would have had the same power, one. For QTLs with large effects, the comparison should rely upon other criteria than power, such as the length of the QTL location confidence interval. Nevertheless, the  $T^8$  test is appreciably better than the other test in estimating QTL variance.

The model using a linkage disequilibrium can lead to more power in some situations. Nevertheless, it is of interest only if one can be sure that there is really a linkage disequilibrium. The other problem for the use of this model is the extension to a general situation where the QTL is not located on a marker.

## REFERENCES

- [1] Andrews D.W.K., Ploberger W., Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica* 62 (1994) 1383–1414.
- [2] Andrews D.W.K., Ploberger W., Admissibility of the likelihood ratio test when a nuisance parameter is present only under the alternative, *Ann. Stat.* 23 (1995) 1609–1629.
- [3] Coppieters W., Kvasz A., Farnir F., Arranz J.-J., Grisart B., Mackinnon M., Georges M., A rank-based nonparametric method for mapping quantitative trait loci in outbred half-sib pedigrees: application to milk production in a granddaughter design, *Genetics* 149 (1998) 1547–1555.
- [4] Elsen J.M., Mangin B., Goffinet B., Le Roy P., Boichard D., Alternative models for QTL detection in livestock. I. General introduction, *Genet. Sel. Evol.* 31 (1999) 213–224.
- [5] Grignola F.E., Zhang Q., Hoeschele I., Mapping linked quantitative trait loci via residual maximum likelihood, *Genet. Sel. Evol.* 29 (1997) 529–544.
- [6] Jansen R.C., Johnson D.L., Van Arendonk J.A.M., A mixture model approach to the mapping of quantitative trait loci in complex populations with an application to multiple cattle families, *Genetics* 148 (1988) 391–400.
- [7] Knott S.A., Elsen J.M., Haley C., Methods for multiple-marker mapping of quantitative trait loci in half-sibs populations, *Theor. Appl. Genet.* 93(1996) 71–80.
- [8] Mangin B., Goffinet B., Le Roy P., Boichard D., Elsen J.M., Alternative models for QTL detection in livestock. II. Likelihood approximations and sire marker genotype estimations, *Genet. Sel. Evol.* 31 (1999) 225–237.
- [9] Soller M., Genizi A., The efficiency of experimental designs for the detection of linkage between a marker locus and a locus affecting a quantitative trait in segregating populations, *Biometrics* 34 (1978) 47–55.
- [10] Weller J.L., Kashi Y., Soller M., Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle, *J. Dairy Sci.* 73 (1990) 2525–2537.