Power of three multitrait methods for QTL detection in crossbred populations.
Hélène Gilbert, Pascale Le Roy

To cite this version:
Hélène Gilbert, Pascale Le Roy. Power of three multitrait methods for QTL detection in crossbred populations.. Genetics Selection Evolution, BioMed Central, 2004, 36 (3), pp.347-361. hal-02674188

HAL Id: hal-02674188
https://hal.inrae.fr/hal-02674188
Submitted on 31 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright
Power of three multitrait methods for QTL detection in crossbred populations

Hélène GILBERT, Pascale LE ROY∗

Station de génétique quantitative et appliquée, Institut national de la recherche agronomique, 78352 Jouy-en-Josas Cedex, France

(Received 25 February 2003; accepted 6 January 2004)

Abstract – The multitrait detections of QTL applied to a mixture of full- and half-sib families require specific strategies. Indeed, the number of parameters estimated by the multivariate methods is excessive compared with the size of the population. Thus, only multitrait methods based on a univariate analysis of a linear combination (LC) of the traits can be extensively performed. We compared three strategies to obtain the LC of the traits. Two linear transformations were performed on the overall population. The last one was performed within each half-sib family. Their powers were compared on simulated data depending on the frequency of the two QTL alleles in each of the grand parental populations of an intercross design. The transformations from the whole population did not lead to a large loss of power even though the frequency of the QTL alleles was similar in the two grand parental populations. In these cases, applying the within-sire family transformation improved the detection when the number of progeny per sire was greater than 100.

multitrait / QTL / sib families / simulations

1. INTRODUCTION

The detection of quantitative trait loci (QTL) in animal populations in the last decade has shown much evidence for loci determining quantitative traits of agricultural interest in many species ([3], review). The hypothesis on genetic determinism of the traits, as that resulting from single trait detection, often suggests the existence of loci influencing several traits (e.g. SSC7 in pigs, [1]). But most of the statistical methods used in such frameworks are aimed at detecting QTL trait by trait. However, specific algorithms, called multitrait, have been developed to test the hypothesis of pleiotropy. They exploit information from the residual and/or genetic correlations between the traits [4, 7, 9].

* Corresponding author: Pascale.Leroy@dga.jouy.inra.fr
It has been shown that the multivariate strategy (MV) [6–8] applied to sib families [4] is greatly limited in terms of power and computing time, due to the high number of estimated parameters. Alternative techniques are based on linear transformations of the traits, with the resulting variables being analysed univariately. A first strategy is based on a principal component analysis performed on the phenotypic covariance matrix of the data (PCA). Since the transformation is carried out on the phenotypic variability, it has been shown to perform well in detecting a pleiotropic QTL [4, 9, 12] only if the determinism of the traits is simple, for example with one pleiotropic QTL determining the traits. Another proposition [4] to overcome this limit is to maximise the ratio of the variability due to the putative QTL and the residual variability at each tested position. This latter strategy consists in performing a discriminant analysis at each tested position. It is based on the distribution of the phenotypes of the progeny between groups of haplotypes inherited from their sire. In each group, the performance for each trait is weighted by the probability that the progeny received the corresponding haplotype.

In both cases [4, 14], a unique linear transformation of the traits was performed from the covariance matrix estimated from the whole population. Thus, the effects of the QTL alleles were assumed to be similar in all the families, i.e. all sires, and eventually dams, were assumed to be heterozygous and to share the same two QTL alleles. In animal populations, even coming from crosses of divergent lines such as pig experimental populations [1], QTL alleles might not be fixed, at least for some QTL. The aim of this study was, for one particular case of multitrait QTL detection, to evaluate the consequences on the power and the accuracy of the position estimates of the segregation of the QTL alleles in the grand parental populations in an intercross design.

2. METHODS

Multitrait methods and the single trait method (ST) were compared. These methods applied to a mixture of full and half sib families are extensively described in [4] and [11]. Only the major points are restated here.

The methods are based on an interval mapping method [10], considering a mixture of full and half sib families. The population is considered as a set of $n$ sire families ($i = 1, \ldots, n$), with $n_i$ mates for the sire $i$ ($j = 1, \ldots, n_i$) and $n_{ij}$ progeny for the dam $ij$ ($k = 1, \ldots, n_{ij}$). Some hypotheses have been made to improve the robustness and the time of computation, based on [2, 5, 13]. In particular, only the most probable sire genotype was retained in the calculation, and the likelihood was linearised within full-sib families. This could not be
totally linearised because of the uncertainty of the dam genotype. According to the notations of Elsen et al. [2], the general form of the likelihood function is then written at the $x$ location as:

$$\Lambda^x = \prod_{i=1}^{n} \prod_{j=1}^{n_i} \sum_{hd_{ij}} p(\text{hd}_{ij}/\widehat{h}s_i, M_i) \prod_{k=1}^{n_{ij}} f_{ijk}$$

where $M_i$ is the marker information of the sire family $i$; $\widehat{h}s_i$ is the most probable sire $i$ genotype for genetic markers; $\text{hd}_{ij}$ is a dam $ij$ genotype having a probability greater than 0.1 for genetic markers; $f_{ijk}$ is the penetrance function, which is specific of the method and test applied.

The test statistic was an approximate likelihood ratio test [2]. Whatever the method, the parameters of the methods were estimated within half- and full-sib families as in [4].

### 2.1. Single trait analysis

For the single trait method, the penetrance function $f_{ijk}$ is assumed to be normal for each trait. For each quantitative trait $l$, the within half-sib and full-sib family phenotypic means and average effects of the QTL substitution, and the variance of the half-sib families were estimated under the hypothesis of the segregation of a QTL at the $x$ location.

All the following methods were aimed at detecting pleiotropic QTL by multitrait analysis. The traits must be standardised before the analysis to avoid confusion between the scales of measure and sizes of the effects.

### 2.2. Multivariate analysis

In the multivariate method, the penetrance function $f_{ijk}$ is assumed to be a $p$-dimension multinormal function, where $p$ is the number of analysed traits. With this method, all the parameters estimated with the single trait analysis are jointly estimated, plus the residual correlation coefficients. They were assumed independent from the sire families to limit the number of estimated parameters (see [4]).

### 2.3. Principal component analysis

Weller et al. [14] first proposed a univariate analysis of linear combinations of the traits. Each linear combination is thus analysed as a trait, as described for the single trait analysis.
A principal component analysis was carried out from the phenotypic covariance matrix, estimated from the overall population, leading to as many phenotypically independent variables as traits. Technically, it is easy to implement. Since it is performed from the phenotypic data, it assumes that the residual variability within QTL genotypes and the phenotypic variability have similar magnitudes. Otherwise, the variability explained by the QTL can be split between different component variables, reducing the power of detection [4, 12].

We also tested the sum of the test statistics of the linear component variables (sPCA) at each position. Such a strategy mimics the multivariate technique if the residual variability within QTL genotypes and the phenotypic variability are of the same magnitude [12].

2.4. Discriminant analysis

More recently, Gilbert and Le Roy [4] proposed to apply the discriminant analysis (DA) technique to the multitrait detection of QTL. As for PCA, a univariate analysis of a linear combination of the traits was conducted, but the calculation of the linear combination was specific to the putative effect of the tested position on the traits.

The linear combination which best discriminates the putative QTL effects was established at each tested position. To calculate the linear combination, groups of progeny were distinguished, based on the sires’ QTL haplotypes they inherited at the tested position. This maximised the ratio of genetic variability (between groups, due to the putative QTL) and the residual variability (within groups, due to any other reason). This was thus the best way to discriminate between the QTL effects on all the traits within the sire families. Moreover, if the sires and dams have the same QTL alleles, as in intercross designs between breeds fixed for the QTL alleles, it is also the best way to estimate QTL effects within full-sib families [4].

Two different strategies were applied to define the groups of progeny. First, two groups were defined from the overall population, with respect to the grand parental origin of the sire’s haplotypes, as in [4]. The transformation was then simply performed from the genetic and residual variabilities estimated from the overall population (PDA), as for PCA. But such a transformation supposed at least that the QTL alleles in the grand parental breeds had different frequencies, to make the between group variability high enough to be detected. Thus, we implemented a second strategy where the groups of progeny were defined within each half-sib family. The linear combination was then specific to the QTL alleles of the sire considered. If the QTL alleles were different between
the sires, the linear combination should thus better characterise their specific effects on the traits. This strategy will be denoted FDA (Familial DA). This assumed that the number of progeny per sire was large enough to accurately estimate the covariance matrix. Then, it was expected to improve the detection when the QTL alleles were not fixed in the grand parental populations. But only the sire’s haplotypes were considered to perform the transformation. Thus, if it is homozygous for the QTL, the linear combination will not discriminate between the groups. Then, the information from heterozygous mates will be lost.

3. A COMPARISON OF POWER AND POSITION ESTIMATIONS

For a particular linkage group, we tested the hypothesis $H_0$ “there is no QTL for any trait on the linkage group” vs. $H_1$ “there is a QTL determining at least one trait”.

3.1. Studied cases

The population came from an intercross design between two grand parental populations, respectively one population of grand sires and one population of grand dams. The three generations of the design were denoted F0, F1, F2. Five hundred F2 progeny were simulated, with 25 progeny per F1 dam. The number of F1 sires varied from 2 to 10. Then, the number of progeny per sire ($nps$) varied from 250 to 50.

Nine genetic markers evenly distributed on a 1 M-long linkage group were simulated. Each had five isofrequent alleles, identical in the two grand parental populations. We considered two traits with equal polygenic heritabilities of 0.2, residual variances of 1 and a residual correlation between the traits equal to $-0.4$. Under the general $H_1$ hypothesis, a QTL influencing the two traits was located at 31 cM on the linkage group. Its two alleles, denoted 1 and 2, had an additive substitution effect of 0.50 on each trait.

This design was similar to one of those reported in [4] with 10 F1 sires and fixed QTL alleles. The multitrait methods PDA, PCA and sPCA were already shown to be powerful to detect pleiotropic QTL in such designs. In the present study, the frequencies of the QTL alleles in the grand parental populations varied and two different situations were studied:

(1) The QTL segregated in the two grand parental populations with $fp(1) = fm(2)$ varying between 1 and 0.5, where $fp(i)$ and $fm(i)$ ($i = 1, 2$) were the frequencies of allele $i$ in the grand sire and in the grand dam populations respectively. Except when $fp(1) = fm(2) = 1$, the 4 QTL genotypes were represented
in the F1 generation. When \( fp(1) = fm(2) = 0.5 \), the groups defined from the grand parental origin of the sire’s haplotypes to perform the discriminant analysis on the overall population were then theoretically identical.

(2) The QTL segregated in only one of the grand parental populations (say the grand sires). \( fp(1) \) varied from 1 to 0.25 (when \( fp(1) = 0 \), all the F1 individuals are homozygous 2/2 for the QTL), and \( fm(2) = 1 \). These simulated designs were more realistic than the case (1). They could be related to experimental crosses between selected (supposed to be the F0 females) and exotic (supposed to be the F0 males) populations.

3.2. Power calculation

We performed 2000 simulations under H0 to estimate chromosome-wise thresholds at the 5% level, and 1000 simulations under H1 to determine the powers. The five methods were successively applied on each replicate: ST, PCA, PDA, FDA and sPCA. For ST and PCA, rejection thresholds were corrected for the number of analysed variables by an approximate Bonferroni correction. This corresponded to the division of the type-I error by the number of variables analysed, considered as being independent.

Due to the characteristics of the QTL simulated, the principal component variable which detects the pleiotropic QTL was associated with the lowest eigenvalue of the phenotypic covariance matrix [4]. Thus, only the results from the analysis of this variable will be presented for PCA.

3.3. Accuracy of the position estimates

The accuracy of the position estimates was assessed by considering the mean squared errors (MSE) obtained over the 1000 simulations performed to calculate power. It synthesises the bias and the standard deviation of the estimates (\( MSE = biais^2 + \text{standard deviation}^2 \)).

4. RESULTS AND DISCUSSION

Tables I and II summarise the power of the methods when the QTL alleles segregate in at least one of the grand parental populations. The power of the single-trait method was always improved by the multitrait strategies. This was the consequence of the determinism of the two traits: this great improvement from the multitrait analysis compared with single trait detection when
the product of the effects of the QTL on the traits with the residual correlation
is negative, has been described by many authors [4, 6–9, 12]. The decrease in
power due to the decrease of fp(1) did not modify this statement.

When fp(1) = 1, the QTL alleles were fixed in the grand parental popu-
lations. The more powerful method was PDA, followed by PCA for all nps.
Concerning the MSE of the position estimates (Tabs. III and IV), similar re-
sults were obtained: PDA and PCA had the more accurate estimations, with a
similar magnitude. The accuracy of sPCA and MV were twice those of PCA
and PDA. FDA was intermediate. This order was expected from previous re-
sults [4], since the transformation from the overall population did not miss
information with respect to the allele segregation.

In all these cases, the results obtained for MV and sPCA were very similar,
as stated by [4, 12]. The power was generally the lowest, which might be due
to a huge amount of parameters to estimate together, compared with the trans-
formations performed on the population. They are not described separately
below. Similarly, the results on the power and the MSE of the position esti-
mates showed similar decreases with the decrease of fp(1). The trends of the
MSE are restated separately in the following paragraph only when they were
different from the trends of the power.

4.1. QTL segregating in the two grand parental populations

Table I summarises the power of the methods when fp(1) = fm(2), and
Table III the corresponding MSE of the position estimates.

As the frequencies decreased, PDA and PCA powers tended to reduce faster
than those of FDA and sPCA. The fastest decrease concerned PDA. It turned
to be equivalent to the other methods in terms of power when the QTL alleles 1
and 2 were isofrequent. This was due to the equiprobability of the four QTL
genotypes in the F1 generation. It made the two groups of progeny theoretically
identical. Then, the strategy based on the discriminant analysis performed from
the overall population should become inefficient. Actually, PDA did not lose
that much power as compared with the other methods. Indeed, a low number
of sires in the design (from 2 to 10) implied great sampling variations of the
QTL genotypes of the sires between the simulated designs. This implies that
even when the allele frequencies were assumed equal on average, one of the
QTL alleles was often more represented by chance in the population. Thus,
enough between-group variability was created, making it possible to detect the
pleiotropic QTL without losing much power compared with the other methods.
Table I. Power (%) when the QTL is fixed or the QTL is segregating in the two grand parental populations.

| fp(1)  | 0.500  | 0.625  | 0.750  | 0.875  | 1.000  |
|--------|--------|--------|--------|--------|--------|
|        | 50     | 100    | 250    | 50     | 100    | 250    | 50     | 100    | 250    | 50     | 100    | 250    |
| PDA    | 31.0   | 40.4   | 46.1   | 37.4   | 45.0   | 52.5   | 51.0   | 59.2   | 61.6   | 68.7   | 76.3   | 74.8   | 89.1   | 92.4   | 95.4   |
| FDA    | 29.9   | 41.5   | 48.4   | 34.3   | 48.5   | 51.6   | 44.5   | 55.0   | 59.1   | 54.7   | 69.0   | 73.9   | 74.5   | 83.5   | 94.3   |
| PCA    | 36.2   | 47.9   | 47.9   | 42.8   | 51.2   | 54.5   | 50.8   | 60.4   | 62.4   | 67.5   | 76.5   | 74.6   | 86.4   | 91.4   | 94.8   |
| sPCA   | 32.1   | 40.7   | 43.2   | 35.6   | 43.2   | 48.2   | 44.9   | 49.3   | 54.1   | 55.8   | 66.6   | 68.4   | 75.1   | 80.6   | 87.1   |
| MV     | 31.9   | 41.0   | 44.1   | 34.9   | 43.2   | 9.0    | 47.3   | 49.5   | 54.0   | 57.2   | 65.7   | 68.2   | 75.7   | 81.1   | 87.7   |
| ST     | 8.7    | 11.9   | 11.6   | 8.1    | 13.5   | 13.0   | 11.4   | 15.7   | 15.6   | 13.9   | 20.2   | 21.5   | 20.3   | 26.1   | 26.6   |

* QTL allele 1 frequency in the grand sire population = QTL allele 2 frequency in the grand dam population.
* number of progeny per sire.
* PDA: discriminant analysis on the overall population; FDA: discriminant analysis within half-sib families; PCA: principal component analysis; sPCA: sum at each position of the likelihood ratio tests of the principal component variables; ST: single trait method; MV: multivariate analysis.
Table II. Power (%) when the QTL is fixed or the QTL is segregating in only one grand parental population.

| $fp(1)^{a}$ | 0.25 | 0.50 | 0.75 | 1.00 |
|------------|------|------|------|------|
| $nps^{b}$  | 50   | 100  | 250  | 50   | 100  | 250  | 50   | 100  | 250  |
| PDA $^{c}$ | 15.1 | 20.5 | 25.7 | 40.1 | 47.4 | 51.2 | 69.4 | 72.6 | 73.8 | 89.1 | 92.4 | 95.4 |
| FDA        | 13.0 | 18.8 | 23.7 | 31.4 | 40.3 | 49.0 | 54.9 | 64.1 | 71.6 | 74.5 | 83.5 | 94.3 |
| PCA        | 14.2 | 20.2 | 22.9 | 38.1 | 46.7 | 50.5 | 68.6 | 71.8 | 72.7 | 86.4 | 91.4 | 94.8 |
| sPCA       | 15.3 | 18.5 | 21.8 | 35.3 | 38.1 | 46.0 | 56.5 | 59.9 | 64.3 | 75.1 | 80.6 | 87.1 |
| MV         | 16.8 | 18.8 | 22.4 | 36.3 | 38.3 | 46.2 | 56.8 | 60.5 | 65.0 | 75.7 | 81.1 | 87.7 |
| ST         | 4.2  | 6.6  | 6.4  | 10.0 | 12.1 | 14.6 | 14.0 | 16.9 | 20.8 | 20.3 | 26.1 | 26.6 |

$^{a}$ QTL allele 1 frequency in the grand sire population (QTL allele 2 is fixed in the grand dam population).

$^{b}$ number of progeny per sire.

$^{c}$ PDA: discriminant analysis on the overall population; FDA: discriminant analysis within half-sib families; PCA: principal component analysis; sPCA: sum at each position of the likelihood ratio tests of the principal component variables; ST: single trait method; MV: multivariate analysis.
Table III. Mean squared errors of position estimates ($10^2$) when the QTL is fixed or the QTL is segregating in the two grand parental populations.

| fp(1) $^a$ | 0.500  | 0.625  | 0.750  | 0.875  | 1.000  |
|------------|--------|--------|--------|--------|--------|
| $nps$ $^b$ | 50     | 100    | 250    | 50     | 100    | 250    | 50     | 100    | 250    | 50     | 100    | 250    |
| PDA $^c$   | 5.99   | 5.63   | 5.03   | 5.90   | 5.37   | 4.22   | 4.28   | 3.94   | 3.92   | 2.64   | 2.42   | 2.21   | 1.22   | 0.90   | 0.71   |
| FDA        | 5.78   | 5.25   | 4.71   | 5.42   | 4.90   | 4.51   | 4.48   | 3.69   | 4.04   | 3.89   | 2.64   | 2.43   | 1.92   | 1.43   | 0.89   |
| PCA        | 3.91   | 3.56   | 4.04   | 3.63   | 3.76   | 3.23   | 2.90   | 2.82   | 3.04   | 2.24   | 1.53   | 1.44   | 1.07   | 0.85   | 0.74   |
| sPCA       | 5.41   | 4.80   | 5.66   | 5.32   | 4.50   | 4.75   | 4.41   | 3.74   | 4.19   | 3.50   | 2.93   | 2.66   | 2.00   | 1.46   | 1.33   |
| MV         | 5.59   | 4.66   | 5.56   | 5.26   | 4.52   | 4.61   | 4.20   | 3.68   | 4.26   | 3.51   | 3.13   | 2.60   | 2.03   | 1.48   | 1.30   |
| ST         | 8.39   | 8.71   | 8.81   | 8.58   | 8.46   | 7.97   | 8.05   | 7.39   | 7.72   | 7.75   | 6.94   | 6.71   | 6.35   | 5.46   | 5.60   |

$^a$ QTL allele 1 frequency in the grand sire population = QTL allele 2 frequency in the grand dam population.

$^b$ number of progeny per sire.

$^c$ PDA: discriminant analysis on the overall population; FDA: discriminant analysis within half-sib families; PCA: principal component analysis; sPCA: sum at each position of the likelihood ratio tests of the principal component variables; ST: single trait method; MV: multivariate analysis.
Table IV. Mean squared errors of position estimates \(10^2\) when the QTL is fixed or the QTL is segregating in only one grand parental population.

| \(fp(1)^a\) | 0.25  | 0.50  | 0.75  | 1.00  |
|-------------|-------|-------|-------|-------|
| \(nps^b\)  | 50    | 100   | 250   | 50    | 100   | 250   | 50    | 100   | 250   |
| PDA         | 8.20  | 8.13  | 7.59  | 4.68  | 5.25  | 5.15  | 2.09  | 2.57  | 2.34  | 1.22  | 0.90  | 0.71  |
| FDA         | 8.40  | 8.27  | 7.87  | 5.20  | 5.38  | 5.23  | 3.68  | 2.91  | 2.58  | 1.92  | 1.43  | 0.89  |
| PCA         | 7.45  | 7.68  | 7.08  | 3.49  | 3.98  | 4.41  | 1.96  | 1.92  | 2.12  | 1.07  | 0.85  | 0.74  |
| sPCA        | 8.35  | 8.21  | 8.29  | 5.03  | 5.37  | 5.57  | 3.10  | 2.95  | 3.33  | 2.00  | 1.46  | 1.33  |
| MV          | 8.35  | 8.21  | 8.29  | 4.79  | 5.35  | 5.61  | 3.06  | 3.32  | 3.42  | 2.03  | 1.48  | 1.30  |
| ST          | 10.15 | 10.13 | 9.73  | 8.43  | 8.25  | 8.32  | 7.20  | 7.17  | 7.39  | 6.35  | 5.46  | 5.60  |

\(a\) QTL allele 1 frequency in the grand sire population (QTL allele 2 is fixed in the grand dam population).
\(b\) number of progeny per sire.
\(c\) PDA: discriminant analysis on the overall population; FDA: discriminant analysis within half-sib families; PCA: principal component analysis; sPCA: sum at each position of the likelihood ratio tests of the principal component variables; ST: single trait method; MV: multivariate analysis.
For PCA, the detection was based on the phenotypic covariance matrix of the traits in the progeny. Since \( fp(1) = fm(2) \), the general shape of the phenotypic data was not modified. Thus, the decrease of the power was only a consequence of a decrease of the number of informative parents F1. For FDA, the power of detection was clearly improved by the increase of \( nps \). Indeed, the estimation of the within-family covariance matrix was much more accurate for 250 \( nps \) than for 50 \( nps \). This power increased by 30% on average when \( nps \) was between 50 and 250, compared with 17 and 19% respectively for PCA and PDA.

For the two latter methods, the improvement of the power was greater when \( nps \) increased from 50 to 100 (22 and 14% respectively) than when it increased from 100 to 250 (2 and 9% respectively). These percentages tended to be higher when the frequencies of the QTL alleles decreased, due to a greater impact of the different QTL allele segregation between the families.

On the one hand, these trends led to an inversion in the order of the multitrait methods between extreme frequencies of QTL alleles in the grand parental populations. On the other hand, FDA was as powerful as PDA when \( nps \) was greater than 100 or 250, depending on \( fp(1) \). As \( fp(1) \) decreased, the advantage of FDA compared with PDA increased. FDA can then be employed for the analysis of designs in outbred populations with large half-sib families, such as (grand)daughter designs.

One important specificity of the MSE of the position estimates, compared with the power, concerned PCA. It was always the more accurate method to estimate the position in the cases we simulated. Otherwise, the evolution of the MSE with respect to the decrease of \( fp(1) \) or the number of progeny per sire followed trends similar to that of the power.

### 4.2. QTL segregating in only one of the grand parental populations

Compared with the previous situation, two points are to be highlighted (Tab. II for the power, Tab. IV for the MSE of the position estimates). In this case some of the F1 individuals were homozygous 2/2 for the QTL. The F2 progeny QTL genotypes then presented an excess of allele 2. Thus, the general shape of the phenotypic data was modified compared with the cases with equal proportions. This potentially decreased the power of PCA and sPCA, since the phenotypic component variables were modified.

First, PDA was the more powerful method in most cases. FDA was then never more powerful than PDA. Indeed, in this situation, only one class of QTL heterozygous F1 existed in the design, with allele 1 coming from the grand sire
and allele 2 from the grand dam. But when fp(1) \neq 0, F1 homozygous 2/2 led
to an excess of the QTL genotype 2/2 and a lack of 1/1 in the progeny. Thus,
the phenotypic covariance matrix no longer represented the variability due to
the QTL. On the contrary, the two groups of haplotypes determined from the
overall population by PDA were always different. But actually, the power of
PCA did not decrease more than those of the other methods.

Second, the improvement of the power of PDA expected from FDA was
null in this case. The powers of PDA and FDA were similar for all fp(1) only
if nps = 250. Then, FDA was powerful only when the number of progeny
per sire was enough (greater than 200) to accurately estimate the covariance
matrix.

Finally, very similar trends were observed for the accuracy of the QTL po-
sition estimates, but as previously noted, the accuracy was always better for
PCA, in a lower magnitude than when fp(1) = fm(2).

5. CONCLUSIONS

We compared two multitrait methods based on linear transformations of the
traits from the overall population with one multitrait method based on within
half-sib family transformations. When both QTL alleles were represented in
the grand parental populations, the discriminant analysis performed from the
overall population was still powerful, even if the groups of progeny were theo-
retically identical, due to the small and finite number of sires simulated. When
the QTL alleles were not fixed in one grand parental population, however, the
excess of progeny of the homozygous QTL genotypes did not hugely penalise
the power of PCA. Thus, the multitrait strategies based on transformations
from the overall population are robust when the QTL alleles are not fixed in
F0 and the number of half sib families is relatively low (less than 10).

The within half sib family discriminant analysis was at least as powerful
as PDA only when the number of progeny per sire was greater than 200 in
our designs. It was a consequence of the lack of accuracy of the covariance
matrix estimation by FDA when nps was low, and again the small number
of sires which maintained the power of PDA. This could also come from an
increase in the number of parameters to be estimated due to an increase in the
number of sires in the design, the (co)variance matrix being estimated within
half-sib families. For example, in the cases simulated here, with two sires in
the pedigree, there were 52, 58 and 98 parameters to estimate under H1 re-
spectively for PDA, FDA and MV. With 10 sires, it became 76, 130 and 146
respectively for PDA, FDA and MV. The relative advantage of the discriminant
analysis compared with MV (or SPCA) in terms of the number of parameters estimated might be reduced in such circumstances, but there is still a great advantage in terms of computing. In practice, in experimental crosses for QTL detection between divergent breeds such as in pigs, which correspond to the simulated designs, PDA does not need to be replaced by FDA when the QTL alleles are not supposed to be fixed. But analysing large half-sib families, FDA represents an attractive alternative, where sire family contribution to the test statistic might be better estimated than with PDA.

These results were obtained for a very simple determinism of the traits. In particular, it would be interesting to test the ability of these methods to detect pleiotropic QTL when the phenotypes are also distributed differently between families due to the segregation of other QTL. On the contrary, the power of the methods when the QTL has more than two alleles in at least one of the F0 population should be tested.

REFERENCES

[1] Bidanel J.P., Rothschild M.F., Current status of quantitative trait locus mapping in pigs, Pig News Inf. 23 (2002) 39N–53N.
[2] Elsen J.M., Mangin B., Goffinet B., Boichard D., Le Roy P., Alternative models for QTL detection in livestock. I. General introduction, Genet. Sel. Evol. 31 (1999) 213–224.
[3] Georges M., Andersson L., Livestock genomics comes over ages, Genome Res. 6 (1996) 907–921.
[4] Gilbert H., Le Roy P., Comparison of three multitrait methods for QTL detection, Genet. Sel. Evol. 33 (2003) 281–304.
[5] Goffinet B., Le Roy P., Boichard D., Elsen J.M., Mangin B., Alternative models for QTL detection in livestock. III. Heteroskedastic model and models corresponding to several distributions of the QTL effects, Genet. Sel. Evol. 31 (1999) 341–350.
[6] Jiang C., Zeng Z.B., Multiple trait analysis of genetic mapping for quantitative trait loci, Genetics 140 (1995) 1111–1127.
[7] Knott S.A., Haley C.S., Multiple least squares for quantitative trait loci detection, Genetics 156 (2000) 899–911.
[8] Korol A.B., Ronin Y.I., Kirzhner V.M., Interval mapping of quantitative trait loci employing correlated trait complexes, Genetics 140 (1995) 273–284.
[9] Korol A.B., Ronin Y.I., Itskovich A.M., Peng J., Nevo E., Enhanced efficiency of quantitative trait loci mapping analysis based on multivariate complexes of quantitative traits, Genetics 175 (2001) 1789–1803.
[10] Lander E.S., Botstein D., Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps, Genetics 121 (1989) 185–199.
[11] Le Roy P., Elsen J.M., Boichard D., Mangin M., Bidanel J.P., Goffinet B., An algorithm for QTL detection in mixture of full and half sib families, in: 6th World Cong. Genet. Appl. Livest. Prod., 11-16 January 1998, Armidale, Australia, 26 (1998) pp. 257–260.

[12] Mangin B., Thoquet P., Grimsley N.H., Pleiotropic QTL analysis, Biometrics 54 (1998) 88–99.

[13] 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.

[14] Weller J.I., Wiggans G.R., VanRaden P.M., Ron M., Application of a canonical transformation to detection of quantitative trait loci with the aid of genetic markers in a multi-trait experiment, Theor. Appl. Genet. 92 (1996) 998–1002.

To access this journal online: www.edpsciences.org
