Université catholique de Louvain
Faculté d'ingénierie biologique, agronomique et environnementale

Methodological aspects of the mapping of disease resistance loci in livestock

Pierre Tilquin

September 2003

Thèse présentée en vue de l'obtention
du grade de docteur en sciences agronomiques
et ingénierie biologique

Promoteurs : P. Baret et E. Le Boulengé

Part IV

Conclusion and perspectives

Chapter 10

Conclusions and perspectives

10.1. Context

Researchers who carry out a gene mapping experiment in any species for any quantitative trait of interest have one single objective, i.e. having results. In this context, "having results" means "*to identify the chromosome regions of the genome that are carrying genes (or QTL) involved in the genetical basis of the trait of interest*". This desire is mainly motivated by scientific interests (fundamental or applied), but it is also emphasised by the need to make such experiment economically justified. The housing of animals, the measurement of complex traits, the genotyping of several markers, the statistical analyses, …, these are all components of such experiments that imply very high costs.

From a statistical point of view, the practician's expectation can be expressed as *having the highest power of QTL detection*. As for every statistical test, the statistical power of a QTL detection test is defined as its ability to reject the null hypothesis of no QTL when it is false, i.e. when a true QTL is involved in the determination of the trait of interest. In a more formal language, it corresponds to the probability of rejecting the null hypothesis when it is false, i.e. when the alternative hypothesis is true:

$$P\left(RH_0 \mid H_1\right) = 1 - \beta \tag{10.1}$$

The power is therefore a function of β (the risk of type II error of a statistical test) defined as the probability of accepting the null hypothesis of no QTL when it is false ($P\left(AH_0 \mid H_1\right)$), i.e. missing true QTL effects. Evidently, and as presented in equation (10.1), both concepts are closely linked as they sum to one. A high power of QTL detection will imply that few true QTL will be missed, while a low power will imply that many true QTL will be missed.

Another aspect of the power of QTL detection concerns the size of the effect of the QTL that will be detected. As all QTL do not have the same effects on a specific trait, a high power of QTL detection will imply that QTL with large effects as well as QTL with small effect will be detected, while a low power will imply that only large QTL are identified.

One determinant of the power of a statistical test is the significance level used to reject the null hypothesis. This significance level of a test is the error rate that we accept to face when rejecting a null hypothesis, i.e. α, the type I error rate. In the QTL context, it is defined as the probability of identifying false QTL that we accept. This error rate is generally fixed to 5% or 1% and is based on the expectation that the stronger is the deviation from the null hypothesis of no QTL, the higher is the chance that the QTL will be a true one. While the type I error rate is *a priori* defined, the power of QTL detection is uncontrolled.

However, because of the complexity of the statistical approach used for mapping QTL, there are some relatively controllable factors having an impact on the power of QTL detection. These factors are the following:

(1) the statistical tool used to test for the presence of a QTL, making the assumption or not that the phenotypic trait is normally distributed (*e.g.* least-squares or maximum likelihood-based methods *versus* non-parametric or distribution-free methods);

(2) the size of the experiment, i.e. the number of animals that will be measured as well as the number of animals that will be genotyped (if selective genotyping is performed);

(3) the crossing design of the experiment; either inbred or outbred populations can be used and crossed either in an F_2 cross, or a backcross, or in a half-sib design;

(4) the type, the number and the density of the genotyped markers, evenly spaced or not;

(5) the information content of the markers and therefore their heterozygosity, which depends among other on the degree of divergence between parental strains; the stronger their divergence the higher the heterozygosity of the hybrid.

(6) the accuracy of the phenotypic measurement in terms of reliability, i.e. with the highest proportion of genetic variance, and therefore with the environmental variance reduced as much as possible;

(7) the relevance of the phenotypic measurement, i.e. reflecting at best differences between animals in terms of genetic variability for the phenomenon of interest;

(8) the right transformability of the phenotypic measurement by a mathematical transformation to fulfil the underlying assumption of the chosen statistical tool.

All those factors are known to affect the power of QTL detection. Unfortunately, it was not possible to treat all of them in the present thesis.

However, some of them were explored thoroughly, such as the statistical tool, the transformability of the phenotype and the reliability of the phenotypic measurement. Others were only approached such as the heterozygosity, the marker density and the relevance of the phenotypic measurement.

10.2. Conclusions

10.2.1. Statistical tools

In the present thesis, we have shown that the choice of the statistical tool can have implications on the power of QTL detection and therefore on the expected results from QTL mapping experiments.

Thanks to the development of an algorithm for the simulation of any type of non-normally distributed traits (Chapter 5), we have shown by means of simulations that the power of least-squares as well as of maximum likelihood-based mapping methods is reduced if non-normally distributed phenotypes (such as bacteria counts) are analysed.

Furthermore, if a non-parametric method is chosen, this thesis has shown that the way of handling ties has an impact on the power to detect QTL for phenotypes presenting a large number of zero values (Chapters 5 and 6). As such phenotypes are encountered when assessing the resistance of livestock animals to bacterial and parasitic diseases, we have modified the non-parametric QTL mapping test proposed by Coppieters *et al.* (1998) for half-

sib designs to allow for the use of midranks instead of random ranks for tied observations. (Chapter 6). The increase in terms of power of QTL detection was significant for proportions of zeros higher than 35%.

Finally, following the concomitant application of both parametric and non-parametric methods on real data sets in chicken and sheep, we confirm the interest of using both approaches when there is evidence of non-normality, as recommended by Kruglyak & Lander (1995a). Indeed, the NP method can be used to confirm QTL identified by the LS method on non-normally distributed phenotypes as it is more robust to non-normality than LS, and inversely the LS method can show a highest power of QTL detection than the NP method on traits that are normally distributed.

10.2.2. Mathematical transformation

The respective efficiencies (i.e. the power of QTL detection and the accuracy of the estimated QTL location) of parametric and non-parametric methods were also compared in an uncommon, but fair and realistic, context. Indeed, as opposed to other simulation studies on non-normally distributed phenotypes (e.g. Coppieters *et al.*, 1998; Zou, 2001), the performance of the non-parametric method on non-normally distributed traits was compared to performance of the least-squares-based approach applied on *mathematically transformed* phenotypes. Indeed, practitians generally use a *log* or square-root transformation to normalise their data before applying the standard least-squares-based approach (Chapter 3).

In this context, and as opposed to other simulation studies in half-sib design by Coppieters *et al.* (1998) and in backcross by Zou (2001), we have demonstrated that the power obtained by the non-parametric interval mapping method is never higher than the power obtained by a least-squares-based method on mathematically transformed phenotypes (see Chapters 5 and 6).

10.2.3. Guidelines

Based on the results of this thesis, and regarding only the choice of the statistical tool in relation to the distribution properties of the phenotypic (or resistance) trait, we propose the following guidelines to guarantee the highest power of QTL detection:

(1) if the distribution of the disease resistance trait is close to normality, a least-squares or maximum-likelihood based method should be chosen.

(2) if the trait is non-normally distributed, time should first be taken to find the most appropriate mathematical transformation; then, if a mathematical transformation is found, a least-squares-based approach should be used.

(3) if the trait is non-normally distributed and if no satisfying mathematical transformation can be identified (let us say if the value of the W test statistic of Shapiro-Wilk is inferior to 0.90), the non-parametric interval mapping approach should be used either alone, or concomitantly to the least-squares-based approach.

Finally, another issue may be the choice of a less stringent significance level (although it implies a higher probability of false positives) when the non-parametric method is chosen, as the power of the non-parametric test is slightly lower than the power of the least-squares-based method.

10.2.4. Reliability of the phenotypic measurement

In the context of faecal egg counts (FEC) that are used to assess the resistance of animals to parasitic diseases, this thesis has shown that a more reliable, and therefore more powerful, phenotypic measurement can be obtained if aggregates of repeated measurements are used instead of single FEC measurements (Chapter 7). This proposal holds if one can reasonably assume that the immune system of animals has no memory, i.e. that susceptible animals didn't get resistant after the first challenge. Therefore, as far as the model of infection allows for it, this conclusion could be extended

to other disease resistance phenotypes for which multiple measurements are usual.

Owing to the extension of our simulation algorithm for non-normally distributed traits used in Chapter 5 to the simulation of two correlated non-normally distributed traits, we have shown that the power of QTL detection for parasitic resistance in livestock would be increased if aggregates of repeated FEC measurements were used. This would especially be the case if the environmental correlation between both FEC measurements is small. As it is a common practice to drench animals (i.e. to apply an anti-parasitic treatment) between different FEC measurements (e.g. Beh *et al.*, 2002; Raadsma *et al.*, 2002), we propose to generalise this practice in all QTL mapping experiments using FEC as resistance trait. This would allow to reduce the environmental correlation between the different FEC and therefore maximise the power (Chapter 7).

10.2.5. Other parameters

Beside the choice of the statistical tool and beside the use of a mathematical transformation to normalise the data, some other aspects of the QTL detection power were tackled in this thesis. Indeed, some elements were obtained by the analysis of real data sets (Chapters 8 and 9).

For example, the impact of the **heterozygosity** of the sires in a half-sib design was approached in Chapter 9. In this chapter, a genome scan for QTL involved in salmonellosis resistance in sheep was performed. Due to the low information content of the markers, it can be expected that the significance level of some QTL involved in the determination of bacterial colonisation was underestimated. Indeed, the heterozygosity of the sires selected for such experiments is a critical component of the expected power of QTL detection (e.g. van der Beek *et al.*, 1995; Du *et al.*, 2002). Furthermore, a non-homogenous repartition of the informativity along a chromosome (i.e. mixture of fully and partially informative markers) has the drawback that the estimated position of the QTL is attracted to areas with a high information content (van Arendonk *et al.*, 1998).

In this analysis, the impact of the **marker density** on the power of QTL detection was also encountered. As a small number of markers was genotyped on small chromosome segments rather than entire chromosomes, the peaks in the location scores were not sharp enough to appraise correctly the most likely location of the identified QTL. It was indeed demonstrated that, with all other factors being fixed, the highest power is obtained with a marker density beyond 10 cM (e.g. Darvasi *et al.*, 1993; Piepho, 2000).

Regarding the **transformability** of the phenotypic measurement, by the analysis of a real data set in chicken (F_2 and backcross data sets), we were faced with the fact that the information given by the phenotypic measurements was rather low. Most of the animals had zero values for caecal swab counts, and the non-zero values resembled a discrete trait. In this context, it was relatively difficult to obtain a satisfying mathematical transformation to respect the normality assumption of the least-squares method. The low information of these phenotypes has had an impact on the power of QTL detection.

Another factor related to the transformability of the phenotypic measurement is the **relevance** of the phenotypic measurement. The resistance traits that were measured in the F_2 and backcross data sets analysed in this thesis were not reflecting at best the difference between animals in terms of resistance to salmonellosis. This comes from the difficulty of planning such experiments. The choice of the sampling date (or slaughtering date if traits are measured on slaughtered animals) in resistance studies is indeed a critical factor to guarantee QTL mapping results, i.e. the identification of QTL. Once the artificial inoculation is performed (e.g. with a bacterial strain), the sampling date must be chosen in order to obtain the highest variability between individuals in terms of resistance. If the sampling is performed too late, e.g. if a too high proportion of animals have already cleared the bacteria when the sampling is performed, most of the genetic variation (or information) has disappeared and mapping QTL on such data becomes nearly impossible. Indeed, as stated by Andersen (1990), "*once the data have been ascertained, the most powerful statistical methods will not be able to salvage an inappropriately designed study*".

There are different ways to treat this **sampling** problem. Usually, preliminary experiments are performed on control animals, inoculated with the chosen strain, and sampled regularly to observe the process of the infection. To guarantee the repeatability of the results, the experimental conditions must be as close as possible to the experimental conditions that will be used in the whole experiment. Another way of treating the sampling problem is to maintain beside the whole experiment, control animals reared in the same conditions and to sample them regularly to follow the process of the infection. Although preliminary studies or control animals represent a cost, the investment is weak compared to the global cost of such experiments. Such a strategy is therefore preferable if one wants to avoid to throw away all the results of a QTL mapping experiment.

10.2.6. Actual power

Regarding the power problem in QTL mapping experiments, a very high number of simulation studies have been carried out (e.g. Weller *et al.*, 1990; Knott *et al.*, 1996; Baret *et al.*, 1998; Alfonso & Haley, 1998; Mangin *et al.*, 1999). In these simulation studies, depending on the simulated QTL effect, on the design and on the marker density and informativity, power estimates are ranging from 10 to 100%, but are often as high as 80%.

In the present thesis, simulations were carried out to mimic an experimental design used by Frédéric Lantier and co-workers (e.g. Moreno *et al.*, 2003) to assess the sheep resistance to salmonellosis. According to these simulations, the chance of detecting a QTL (α=0.05, chromosome-wide) accounting for 8% of the total phenotypic variance and affecting bacterial colonisation was estimated to 57% with the least-squares-based method and 55% with the non-parametric approach. Although this estimation was performed in a highly informative situation, this expected power of QTL detection is rather low. Indeed, a 57% power of QTL detection means that only 57% of the QTL involved in the trait will be caught, and that 43% of the true QTL will be missed.

Beside this low expected power, the adequacy of the additive genetic model could also be questioned. Indeed, as all other simulation studies, the simulations performed in this thesis were based on the assumption that the QTL effect was additive, which means that having inherited one or the other QTL allele from its sire, implied for an individual that a given quantity was added or subtracted (+*a* or -*a*) from its genetic value. In the context of bacterial colonisation, this model might not be appropriate. Indeed, the action of resistance or susceptibility genes on bacterial (or parasitic) colonisation in organs is still unknown. For example, having a susceptibility allele might have a multiplicative effect on the bacterial growth rather than an additive effect.

Based on the above power estimates and on the questionable adequacy of the additive model for modelling QTL effects in phenotypes such as bacteria or parasite counts, we point out that there is an urgent need for the development of new methodologies able to take into account the distributional properties of phenotypes such as bacteria counts. Such a method would perhaps allow to achieve a higher power of QTL detection than the power obtained by the classical least-squares-based method.

10.3. Perspectives

An interesting perspective to deal with this power problem of phenotypes based on a counting process (such as bacteria counts or faecal egg counts), is the use of Poisson or negative-binomial-regression models (see point 2.3, Chapter 2).

In candidate gene studies on the resistance of sheep to parasitic infections, Paterson *et al.* (1998) and Colman *et al.* (2001a) used a generalised linear model approach with a negative binomial structure of errors to model faecal egg counts (FEC), as suggested by preliminary studies on FEC (Wilson *et al.*, 1996; Wilson & Grenfell, 1997). The generalised linear model approach allowed them to include in the model factors such as the allele at the candidate gene, the cohort (year) or the sex of animals.

In another study in mice, Nagase *et al.* (1999) used a single-marker negative binomial multiple regression analysis to screen for predisposition loci associated with tumour multiplicity. In the same context, but in rat, Lan *et al.* (2001) proposed an interval mapping approach based on a negative binomial structure of errors to account for overdispersion of tumour counts. No details were given in this study about the way conditional probabilities of inheriting QTL alleles were introduced.

However, in these sheep and mouse studies, the results obtained with the negative-binomial approach were not compared to the results of a least-squares-based approach.

More recently, Thomson (2003) proposed a so-called *generalised estimating equations* approach for the interval mapping of QTL for non-normal traits. This approach, based on a Poisson regression model, was developed in the context of a backcross experiment to locate QTL for litter size in mice. As the Poisson distribution constrains the variance of the distribution to be equal to the mean, a negative binomial model to allow for overdispersion could be more appropriate for the analysis of count data (Tempelman & Gianola, 1999).

The approaches developed in the above mentioned studies are promising, but were developed in the context of inbred line crosses (F_2 or backcross). For such designs, one can indeed assume that all phenotypic measurements belong to the same distribution. In this context, one needs to estimate only the parameters of one assumed distribution at the position being tested.

In half-sib structures as observed in sheep or cattle breeding, this assumption might not be realistic because of the family-based heterogeneity. Indeed, in half-sib (or outbred) designs, different distributions (and therefore parameters) should be assumed for each half-sib families as overdispersion can be assumed to differ between sire-families.

The development of such a model may offer a very interesting perspective for the mapping of QTL involved in the resistance to bacterial or parasitic

diseases in livestock. However, the way towards such an adapted QTL mapping methods could be long and tricky.

10.4. General conclusion

From the present thesis, we can conclude that, beside the choice of the QTL mapping method, the success obtained from a QTL mapping experiment is also a matter of quality of information. The quality of phenotypic measurements as well as of marker data is of great importance if one wants to get the best results. Researchers in the field of QTL mapping should focus on having a good quality for both types of information (phenotype and marker), as one cannot give good results without the other.

At the end of this work, we hope that this call will be heard and, that this thesis will contribute to a better understanding of the mechanisms underlying the disease resistance in livestock animals.