Detection and mapping of quantitative trait loci in farm animals

H. Bovenhuis a,*, J.A.M. van Arendonk a, G. Davis b, J.-M. Elsen c, C.S. Haley d, W.G. Hill e, P.V. Baret e, D.J.S. Hetzel b, F.W. Nicholas f

a Animal Breeding and Genetics group, P.O. Box 338, Wageningen 6700 AH, Netherlands
b CSIRO, Tropical Agriculture, Brisbane, Australia
c INRA, Toulouse, France
d Roslin Institute, Edinburgh, Scotland, UK
e Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh, Scotland, UK
f Dept. Animal Science, University of Sydney, Sydney, Australia

Received 31 January 1997; accepted 18 August 1997

Abstract

Genetic maps of livestock based on molecular genetic markers provide tools for the detection and mapping of genes of economic importance in farm animals. The experimental designs employed for such studies are partly determined by the species characteristics or industry structure. Differences between designs include the number of generations studied, family structure e.g., half-sib or full-sib and the use of outbred lines or line-breed crosses. The appropriate method of analysis will depend on the design of the experiment, information required and assumptions which are valid on the underlying genetic model. An overview is given of the methods which are available for the detection of QTL.

Keywords: QTL; Detection; Statistical analysis; Farm animals

1. Introduction

For many years people have changed the genetic make up of plants and animals through selection without knowledge of the underlying genes. Until recently, the tools to identify the genes responsible for genetic differences between individuals or between populations have not been available. Developments in the area of molecular biology have changed this situation and have allowed genes in humans, plants and farm animals to be identified in association studies. A number of quantitative genetic studies have been undertaken to determine appropriate experimental designs for and to analyze data from association studies. The main aim of this paper is to give an overview of the main characteristics and differences between the methods which have been developed for the analysis of association studies in farm animals. In addition, the issue of hypothesis testing will be discussed. We will start with a brief description of the background of association studies. This paper follows on from a QTL detection workshop which was held in conjunction with the International Society of Animal Genetics conference in Tours in July, 1996.
2. In perspective

In animal breeding, current approaches to estimate the genetic value of an individual depend upon phenotypic observations on the individual itself and/or relatives. For almost all of the traits of interest to animal breeders, differences in phenotypic observations are due to genetic and environmental differences. Further, segregation of genes takes place each time genes are transmitted from parent to offspring. As a result of these two factors, accurate estimation of the breeding value of an animal is possible only if a large number of records on the phenotype of the individual itself or its offspring are available. In general, the requirement of a large number of records postpones the age at which the animal can be selected as a parent and therefore, restricts the attainable annual genetic progress. If genes and their effects on traits of interest are known, typing of animals at the DNA level would make it possible to estimate breeding values independent of phenotypic observations. This should result in a reduction of the generation interval. Instead of having perfect knowledge on genes of interest and their effect, genetic markers might be available which are linked to some of the genes. Several studies have indicated that knowledge about genetic markers linked to genes affecting quantitative traits can increase the selection response of animal breeding programs, especially for traits that are difficult to improve when using traditional selection methods (Smith and Simpson, 1986; Stam, 1986; Kashi et al., 1990; Meuwissen and van Arendonk, 1992; Van der Beek and van Arendonk, 1996; Meuwissen and Goddard, 1996). However, before information from genetic markers can be used in animal breeding programs, genes affecting traits of interest need to be detected and their effects estimated.

Sax (1923) showed that genetic markers can be used to identify genetic factors underlying quantitative traits (now called quantitative trait loci or QTL). It was not until the 1950s and 1960s, however, that the work of Sax was followed up to detect genes affecting quantitative traits in livestock (e.g., Neumann-Sorensen and Robertson, 1961). Most of these studies used enzyme, protein or blood group polymorphisms as genetic markers. In general, these studies were not very successful in detecting QTL. One reason for this failure was the limited number of available genetic markers. Furthermore, associations between marker genotypes and quantitative traits were mostly investigated in a number of more or less randomly selected individuals from an outcross population. When the genetic marker has a direct effect on the trait (i.e., it actually is the QTL) the difference between the marker genotypes will reflect the QTL effect. However, the probability that a randomly selected polymorphism has a direct effect on a trait is small. The locus, however, might very well be linked to a QTL. In the latter situation, marker genotype means are expected to differ only if there is linkage disequilibrium between the marker and the QTL at the population level. Therefore, the probability of detecting a QTL depends upon the amount of linkage disequilibrium between the marker and the QTL. In an outcross population, an appreciable amount of linkage disequilibrium between a marker and QTL is expected only if the marker and the QTL are closely linked and the effective population size is small (Soller, 1991). Note that in this situation, estimated marker effects might differ between populations. These are important reasons why the early studies to detect QTL were not very successful.

3. Experimental designs for detecting QTL

3.1. Crosses between lines or breeds

Linkage disequilibrium between a genetic marker and a QTL is required to detect the QTL. One way of introducing linkage disequilibrium in a population is by crossing lines that differ with respect to their allele frequencies at marker loci and QTL. Associations between genetic markers and the trait can be studied by comparing the phenotypic performance of F2 or backcross individuals with different marker genotype configurations. A special case of such a design is the use of inbred lines. If it is assumed that the inbred lines are fixed for alternative marker and QTL alleles, then there are a number of important features of this design (Soller, 1991):

- All individuals in the F1 are heterozygous for the marker as well as for the QTL.
- There is complete linkage disequilibrium between the marker and the QTL in the F1.
There are only two marker alleles and two QTL alleles in the population under study, each allele being at a frequency of 0.5.

All individuals in the F1 have the same linkage phase.

This design is frequently used in laboratory animals (e.g., mice) and plants. For farm animals, however, inbred lines are seldom available. In some experiments, crosses between lines with extreme phenotypes have been used instead. In addition, rearing large numbers of F1 and F2 individuals is only possible for some farm animals (e.g., chicken) but not for others (e.g., cattle) because of the long generation interval and the costs of the experiment. As an alternative to a designed experiment involving crosses between lines, existing outcross populations can be used.

3.2. Outbred populations

Linkage disequilibrium in an outcross population between a marker and linked QTL can be found within families. This is due to the co-segregation of the marker and the QTL. For dairy cattle, analysis can be performed within paternal half-sib families using either the ‘daughter design’ or the ‘granddaughter design’ (Weller et al., 1990). The basic idea of the daughter design is to trace marker alleles from the sires to his daughters and to determine whether daughters that inherited alternative sire alleles differ with respect to the quantitative trait. In the daughter design, many daughters of a sire are scored for markers and evaluated for the quantitative trait. However, in the granddaughter design, a number of sons of a proven sire are scored for genetic markers and granddaughters are evaluated for the quantitative trait. In this latter case, the observations on the granddaughters are used to estimate the breeding value of the sons. This breeding value has a lower residual variance compared to a single observation which increases the power of the experiment (Weller et al., 1990; Van der Beek et al., 1995).

3.3. Differences between experimental designs

There are some important differences between analyzing data from a cross between inbred lines and an outcross population:

- Only a fraction of the (grand)sires will be heterozygous for the marker as well as for the QTL.
- (Grand)sires might have different linkage phases and therefore observations cannot simply be pooled across families; marker effects need to be analyzed within families.
- QTL might have more than two alleles and allele frequencies are unknown.
- Linkage phase between marker alleles is unknown.

The differences between a design based on inbred line crosses and a design within an existing outcross population have important consequences for the statistical analysis of the data, both in terms of power and appropriate methodology. The power of different designs has been studied by Weller et al. (1990) and Van der Beek et al. (1995).

4. Methods of analysis in an outcross population

Different approaches can be used to analyze data from an association study. Five different methods of analysis are considered, i.e., methods based on analysis of variance, regression, best linear unbiased prediction (BLUP), mixture models and mixed inheritance models. The main characteristics of these methods are described. A major distinction between the methods is that the first three use only information on mean phenotypic differences between marker genotype classes whereas, the latter two also use information on the distribution of the trait within marker genotype classes. It is generally assumed that the methods are used for the analysis of an experiment with a two or three generation design in which transmission of marker alleles for only one parent is traced. Methods are categorized based on the statistical–genetic model (number of QTL alleles, number of genetic effects) underlying the analysis and the information used in the analysis. For some of the models different parameter estimation procedures can be used e.g., maximum likelihood or Markov Chain Monte Carlo. Categorizing on the basis of parameter estimation procedure might therefore, be misleading.

4.1. Variance analysis

Traditionally, detection of QTL has been performed by contrasting marker genotype effects (e.g.,
Soller et al., 1976). In this type of analysis, markers are analyzed one at a time. The marker genotype effects can be tested using an F-test as is common in analyses of variance. For an outcross population, significant marker effects at the population level are only expected if there is linkage disequilibrium between the QTL and the genetic marker. Neimann-Sorensen and Robertson (1961) noted that for an outcross population, the analyses of marker genotype effects must be performed within sire families. This can be done by contrasting progeny within a family which have inherited alternative parental alleles.

This method of analysis yields estimates of marker allele substitution effects. However, the analysis does not provide any information about the location of the QTL, i.e., the model cannot distinguish between a tightly linked QTL with a small effect and a loosely linked QTL with a large effect. Another disadvantage of this type of analysis is that some of the progeny cannot be assigned to one of the two parental alleles. These animals have to be excluded from the analysis which results in reduced power. On the other hand, the method does not make any assumptions with respect to the underlying genetic model. If testing is performed using the standard F-statistic, it is assumed that error terms are normally distributed. The analyses can be performed using standard statistical software packages and are not very computationally demanding.

4.2. Regression analysis

Haley and Knott (1992) and Martinez and Curnow (1992) independently of each other introduced at about the same time a regression method of analysis that is capable of utilizing information from flanking markers. In many situations, the probability that an animal has inherited a particular QTL allele from its parent is based only on information from one flanking informative marker. In this case, regression is equivalent to an analysis of variance and consequently, the position and the effect of the QTL cannot be disentangled. However, by utilizing marker bracket information, estimates of both QTL effects and position can be obtained. Instead of regression on the marker genotypes, regression is performed on the probability of an individual having a QTL genotype, given the genotypes for the flanking markers. This probability will depend upon the location of the QTL. By moving a putative QTL along the chromosome, the most likely position of the QTL corresponds to the position with the minimum residual sum of squares. If the analysis is extended beyond the part of the chromosome where flanking markers are available, again it becomes impossible to disentangle the effect of a QTL from its position. In an outcross population, a different regression coefficient is fitted for every family. The method results in an estimate of the QTL position as well as the variance explained by the genotype contrasts within each family. The estimated contrasts can be used to identify the parents which are likely to be heterozygous for the QTL.

Once QTL genotype probabilities are estimated, standard statistical software packages can be used for the regression part of the analysis. Further, regression is not very computationally demanding. The regression method can easily be extended to take into account more complex models, such as including effects of other QTL on the same or different chromosomes (e.g., Spelman et al., 1996).

4.3. BLUP-based analysis

The prediction of an animal’s breeding value is based on additive genetic relationships between individuals. In building the numerator relationship matrix, no knowledge on the actual contribution of a parent to its offspring is used. Fernando and Grossman (1989) and Van Arendonk et al. (1994) showed that information on a single marker can be used in an animal model by fitting additive effects for alleles at a QTL linked to the marker and additive polygenic effects for alleles at the remaining quantitative trait loci. Goddard (1992) extended the model proposed by Fernando and Grossman (1989) to include information from more than one marker. Bink et al. (unpublished) extended the procedure in order to incorporate animals without information on marker genotypes. In assigning QTL genotypes to animals, Bink et al. (unpublished) used information on marker genotypes of the individual or its relatives as well as information on distribution of the trait. Meuwissen and Goddard (1996) presented a method in which the covariance matrix of effects at the marked QTL are approximated. This approximation reduces the com-
Computational requirements and facilitates the inclusion of animals without marker genotype information.

Information on an animal's genotype at a marker locus provides information on transmission of a chromosomal region from parents to offspring. If QTL are located in the chromosomal region, then this information can be used to obtain a more accurate estimate of the breeding value because the inheritance of alleles at the chromosomal region can be traced more precisely than inheritance at an unmarked QTL. In this case, the additive genetic value of an animal can be partitioned into additive genetic value at the marked chromosomal region and the sum of additive genetic effects at all other QTL affecting the trait. The allelic effects at the chromosomal region are represented in the model as a random effect.

For situations with information on a single marker (Van Arendonk et al., 1993) or a marker bracket (Grignola et al., 1994, 1996a,b), this method can be used to estimate variance components using a derivative-free REML approach. With a single marker and information only on two generations, the variance due to allele substitution effects can be estimated, while in the case of marker brackets, position and variance due to QTL can be estimated. With marker and phenotypic data collected over several generations, it becomes possible to separate the position and effect of a QTL and estimate both the variance it causes and its distance from a single marker or from the end of a linkage group. In addition, estimates of the variance due to 'residual' polygenic effects and the residual variance are obtained. The method does not require specification of the number of alleles at the QTL or the allele frequencies.

4.4. Mixture model analysis

Weller (1986, 1990) developed mixture model methods to detect QTL using single marker information for crosses between inbred lines and for outcross populations. Solutions are obtained by maximizing the likelihood. For an F2 cross between inbred lines, maximization is usually with respect to five parameters: The mean, the additive and the dominance QTL effect, the recombination fraction between the marker and the QTL and the within-QTL genotype residual variance. Alternatively, it can be assumed that the within-QTL genotype residual variance differs for the three QTL genotypes, resulting in an increase in the number of parameters to be estimated. Besides an estimate of the QTL genotype effect, this method provides an estimate of the QTL location using information from a single marker. In calculating an individual's QTL genotype probability, information from marker genotype as well as information from its phenotype is used. The latter information enables the estimation of the underlying components of the marker allele substitution effect, i.e., recombination rate and genotype effect. Although one would expect this type of analysis to have a slightly higher power than methods that do not make use of phenotypic observations to assign individuals to QTL genotypes, this is not supported by the results of Simpson (1992). The estimated location of the QTL when using single marker information is not very accurate and the information with respect to the location is not complete as it is not known whether the QTL is located to the left or the right of the marker.

For a design in which inbred lines are crossed, a maximum of two QTL alleles will be segregating in the F2, both at frequencies of 0.5. For an outcross population, the QTL allele frequency is not known and therefore, is an additional parameter that needs to be estimated (e.g., Bovenhuis and Weller, 1994). Hypothesis testing is commonly based on the likelihood ratio test statistic.

The mixture model approach suggested by Weller (1986) for crosses between inbred lines has been extended to take into account information from flanking markers. This approach is also known as interval mapping (Lander and Botstein, 1989). The difference in power between single marker analysis and interval mapping depends upon the heterozygosity of the markers and the position of the QTL within the marker bracket (Darvasi et al., 1993, Van der Beek et al., 1995). As compared to the single marker approach, interval mapping has the advantage that a more accurate estimate of the QTL position can be obtained (e.g., Knott and Haley, 1992). Using information from multiple markers will be especially advantageous when analyzing outcross populations, i.e., populations where not all markers will be informative.

Several methods have been suggested for maximizing the likelihood functions, e.g., the EM algo-
rithm, Newton Raphson and derivative-free approaches (see Weller and Ron, 1994). For these types of analysis, no standard statistical software packages are available. Further, maximizing the likelihood is computationally demanding.

Mixture models have also been used to detect genes in cases where there is no information on genetic markers, i.e., segregation analysis (e.g., Knott et al., 1992). In that case, genotype probabilities are based entirely on phenotypic information. As expected, this analysis does not provide any information on the location of the gene.

4.5. Mixed inheritance analysis

When the 'residual' polygenic effects, i.e., genetic effects due to unmarked QTL, are included in the mixture model this results in a mixed inheritance model. To date, mixed inheritance models have mostly been used for segregation analysis, i.e., the detection of major genes based on phenotypic observations only. If information on genetic markers is also available, mixed inheritance models can be used for QTL analysis (Hoeschele, 1994; Uimari et al., 1996a). Obtaining solutions for a mixed inheritance model becomes an almost impossible task especially when applied to livestock populations with a complex pedigree (e.g., Knott et al., 1992; Kinghorn et al., 1993). However, Janss et al. (1995) showed that Markov Chain Monte Carlo methods can be applied successfully to obtain solutions for a mixed inheritance analysis for segregation analysis. In addition to point estimates, the posterior distribution of a parameter can be obtained which provides information on the accuracy of the estimates. These methods are known to be computationally demanding.

5. Hypothesis testing

The threshold for test statistics corresponding to a given level of type I error can be obtained from standard tables. However, critical values are only applicable if a single test is performed. When analyzing experimental data involving a genome-wide search for QTL, many hypotheses are tested. Several of these tests will involve markers or marker brackets located on the same chromosome, and will therefore, not be independent. In order to determine the appropriate significance thresholds, these dependencies have to be taken into account.

Lander and Botstein (1989) developed a formula which calculated an appropriate threshold for the test statistic depending upon several factors including the number of chromosomes and total length of the genome. For human linkage analysis a lod score threshold LOD score of 3.3 (Lander and Schork, 1994) was suggested for a genome wide false positive error rate (Type I) of 5%.

An empirical method for determining threshold values based on the permutation test was suggested by Churchill and Doerge (1994). Their approach is based on removing any association between the quantitative trait and the genetic marker by randomly shuffling the trait values. In this way, each individual retains its marker genotypes but is assigned a different trait value. For each shuffle, an appropriate test statistic is calculated. The distribution of the test statistic corresponds to the distribution under the null hypothesis, i.e., no associations between the marker and the quantitative trait. Because the trait values and the marker genotypes are not altered, the simulated test statistic will automatically take into account characteristics of the data such as its distribution. Churchill and Doerge (1994) also distinguish between a comparisonwise and an experimentwise threshold. The comparisonwise significance level reflects the probability that at a specific locus an extreme deviation of the test statistic is found just by chance. An experimentwise significance level is the probability that somewhere in a whole genome a deviation of the test statistic exceeding the threshold value is found just by chance. Because the distribution of the test statistic is approximated, the accuracy at which the critical value is determined will depend upon the number of permutations.

Spelman et al. (1996) found differences between traits in the distribution of the chromosomewise test statistic which resulted in substantial differences in critical values. In addition, they adjusted their critical values for the fact that they analyzed five correlated traits.

When analyzing genome-wide scans for QTL, stringent significance levels have to be applied. However, applying stringent significance levels will reduce the power of detecting a QTL. Nevertheless,
Lander and Krugylak (1995) emphasized that although initial analysis may not show significant results, they may point to important chromosomal regions which may be the subject of further research.

6. QTL detection workshop

A number of methods have been developed for mapping QTL. As a way to review and compare these new approaches, a workshop was organized for which different research teams performed a QTL analysis on the same data sets. This has been a standard procedure in developing new methods in human genetics (e.g., Hodge, 1995), but is new in animal genetics. Both real and simulated data for a half-sib granddaughter design as well as full-sib design were distributed in the context of the workshop. Most groups concentrated on the analysis of the real half-sib data which we will describe in some detail.

6.1. Data structure

Simulated data were generated under the constraint of the marker data produced by the MII.QTL project of Holland Genetics (The Netherlands) and Livestock Improvement Corporation (New Zealand). QTL were allocated to grandsires depending upon the population frequencies of the QTL alleles. For each son, the inherited QTL allele was sampled given the marker information. Further, breeding values for grandsires and sons and granddaughter yield deviations were simulated. Number of grandsires, sons and grandsires was corresponding the MII.QTL design. In total, 20 grandsire families were genotyped at the University of Liège under the supervision of Wouter Coppieters and Michel Georges. The total number of half-sib sons was 692 (9 to 139 sons per family) and the four largest families accounted for 48% of the sample. The number of daughters tested per son spanned from 1 to 776 (mean = 110). The nine microsatellite markers were situated on chromosome 6. In a previous study, major QTL were located on this chromosome (Georges et al., 1995). The relative positions of markers were 0, 13, 20, 31, 41, 52, 54, 58, 95 cM. For the real data part, three parameters (daughter yield deviation, number of tested daughters and reliability) were provided for five traits: Milk, fat and protein yield, fat and protein percentage. In the simulated data, five different models were defined using in each case a non-QTL heritability of 36% and a phenotypic standard deviation of 1000. The simulated models differed in the number of QTL, their effects, allele frequencies and positions. A polygenic model was also simulated, i.e., 16 QTL with small and additive effects.

6.2. Results

The purpose of this workshop was to draw a parallel between different analytical approaches for the same data set. Any strict comparison of the results is pointless due to the unique and limited data set. Indeed, the originality of the simulation process was constrained to the actual pedigree and to the marker structure which resulted in an unbalanced and relatively poorly informative design. This situation contrasts with simulations usually used in power calculation where the family sizes are generally equal and the markers are evenly spaced.

The workshop in Tours was the first occasion where results from different methods for QTL mapping applied to a common data set were presented. Classical methods of analysis (regression and mixture model methods) were very efficient in the detection of a single QTL but further refinement should be considered in multiple QTL models. For the simulated polygenic model in most of the cases a large unique QTL was detected. Such a result emphasizes the problem of the null hypothesis. The usual practice is testing a major QTL underlying the considered trait against no genetic variation for the considered trait. A more sensible approach could be to test a major QTL hypothesis vs. a polygenic model (Visscher and Haley, 1996).

Some of the results of analyses of the workshop data are already published (Spelman et al., 1996; Uimari et al., 1996b) or are in the process of being published.

The workshop brought together experts from a large number of laboratories and stimulated discussion on the various methods. It also served as a forum for the proposition of new approaches as well. Methods were only applied on one data set with a relatively simple pedigree structure. For a full com-
parison of methods more data sets need to be analyzed and more complex pedigree structures need to be considered. Little attention was given to the full-sib data. This is a reflection of the short amount of time given to groups for analysis and the fact that most groups have experience with the analysis of half-sib data. Currently, more groups have also started to work on data with full-sib families.

The simulated data did not play a major role in the first workshop. However, it was felt that in future workshops more attention should be paid to simulated data. The number of data sets distributed for analysis, however, should be small in order to keep the workshop focused.

7. Discussion

In this paper, a number of different methods that can be used to analyze QTL detection and mapping experiments have been discussed. Some methods can be applied relatively easily (e.g., regression) and do not put high demands on computing capacities. Other methods require specific knowledge on the nature of QTL effects or put high demands on computing facilities. There is scope for all types of methods to be used: Less computer intensive methods can be used as initial screening, and more computer intensive methods can be used for a detailed analysis of interesting chromosomal regions.

One point of difference between the methods discussed is the assumptions made with respect to the QTL. In the mixture models and the mixed inheritance model it is usually assumed that two alleles at the QTL are segregating. Except for potential computational problems (e.g., number of parameters in the likelihood), there is no basic problem in extending the number of QTL alleles in these models. On the other hand, in the regression method of analysis and the BLUP-based models, no specific assumptions with respect to the number of QTL alleles are made. The assumption of two QTL alleles will be invalid if there are three or more alleles, all at relatively high frequencies and with measurable differences in effects. A situation with two closely linked QTL might be observed as one QTL with multiple alleles. It is expected that a statistical model that explicitly specifies the underlying genetic model (e.g., number of alleles) is more powerful, at least for situations where the model and reality agree. If the model is specified incorrectly, e.g., in reality there are more than two alleles, this may lead to failing to detect true QTL or detecting false QTL. On the other hand, a model where the number of QTL alleles is not explicitly specified is expected to be more robust but less powerful than a correctly specified model.

When flanking markers are used, information is needed on the interference and on the recombination fraction between the flanking markers. Usually, the recombination fraction between markers is assumed to be known and interference is assumed to be absent. The latter assumption in particular greatly simplifies the analysis because recombination events within a bracket become independent of recombination events outside the bracket.

The methods that have been presented make use of different types of information to assign QTL genotypes to individuals. The analysis of variance method does not assign QTL genotypes to individuals and thus, does not make any assumptions with respect to the QTL. Mixture model and the mixed inheritance methods use marker genotype information and phenotypic observations to assign individuals to QTL genotypes whereas, the regression type of analysis and the BLUP-based method only make use of the marker genotype information. However, for these methods, it is possible to make use of information from phenotypic observations to assign QTL genotypes to individuals as was recently shown by Bink et al. (unpublished) for the BLUP-based methods. Obtaining parameter solutions then becomes more cumbersome. Making use of phenotypic observations, in addition to the marker information, should make it possible to extract more information from the data. However, in order to do so, assumptions have to be made about the distribution of the data. Again, if the model is correctly specified, it is expected to be more powerful. Violation of the assumed distribution of observations may lead to false conclusions.

In the future it is expected that QTL genotypes will be assigned to individuals based on full pedigree information. This situation already exists for breeding values where estimates are obtained from data from an ongoing breeding program. For general
pedigrees, the BLUP-based and the mixed inheritance methods are better suited. The regression method can not be extended to include full pedigree information. The permutation test offers a good way of calculating appropriate threshold levels. However, it is not clear how the permutation test can be used in a full pedigree analysis.

Risch and Botstein (1996) suggest that making data publicly available should be a condition for publication. The reason for this is that the statistical power of most studies to detect QTL will be low. This will be especially true if stringent experiment-wise significance levels are applied. Consequently, it is expected that different QTL will be detected in different studies. Only for QTL with large effects will individual studies provide conclusive evidence. By making data publicly available, data can be combined into one analysis.

Multiple QTL mapping methods have been introduced by Jansen (1993) and Zhao-Bang (1993). These methods have been specifically set up for analyzing crosses between inbred lines. By combining interval mapping with multiple regression on a number of pre-selected markers, part of the variation arising from QTL located elsewhere on the genome will be eliminated and thereby, the probability of detecting each QTL will be higher. Multiple QTL mapping methods have been applied successfully to plant data (Jansen and Stam, 1994). Two QTL models have been applied to analyze data from a granddaughter design in cattle (Spelman et al., 1996). However, in this study both QTL were located on the same chromosome. The potential of the multiple QTL mapping methods as suggested by Jansen (1993) for outcross populations requires further study.

Acknowledgements

The authors acknowledge the financial support from the European Human Capital and Mobility fund.

References


