



# The genetic dissection of immune response using gene-expression studies and genome mapping

Dirk-Jan de Koning\*, Örjan Carlborg<sup>1</sup>, Chris S. Haley

*The Roslin Institute, Roslin Biocentre, Roslin EH25 9PS, UK*

## Abstract

Functional genomics has been applied to the genetic dissection of immune response in different ways: (1) experimental crosses between lines that differ in their (non-) specific immune response have been used to detect quantitative trait loci (QTL) underlying these differences. (2) The measurement of gene expression levels for thousands of genes using microarrays or oligonucleotide chips to identify differential expression with regard to antigen challenge: (a) before and after infection, (b) resistant versus susceptible lines, or (c) combinations of both. Interpretation of QTL results is hampered by the fact that confidence regions of the QTL are large and can contain hundreds of potential candidate genes for the QTL. At the same time, the microarray experiments tend to show large numbers of differentially expressed genes without identifying the relationships between these genes. In the recently proposed 'genetical genomics' framework, members of a segregating population are characterised for genome-wide molecular markers and for gene expression levels. This facilitates the mapping of expression-QTL (eQTL): loci in the genome that control the expression of genes. Initial applications of this approach are critically reviewed and potential applications of this approach with regard to immune response are presented.

© 2005 Elsevier B.V. All rights reserved.

*Keywords:* Gene expression; Genetics; Immune response; Mapping; Quantitative trait locus

## 1. Introduction

The genetics underlying production traits has been thoroughly studied and exploited for genetic improvement of livestock through selective breeding for decades. For many traits, regions of the genome that affect these traits (quantitative trait loci: QTL) have

been identified and in some cases even the molecular polymorphism underlying the QTL has been identified (for reviews on QTL mapping in livestock see Andersson, 2001; Andersson and Georges, 2004). There is very little doubt about the economic and welfare implications of infectious diseases or about the existence of genetic variation in disease susceptibility in livestock populations (Stear et al., 2001). However, the availability of sufficient phenotypic observations on structured populations limits research that is aimed at identifying genetic factors conferring relative susceptibility or resistance against an infectious disease. Research into the genetics underlying

\* Corresponding author. Tel.: +44 131 5274258; fax: +44 131 4400434.

*E-mail address:* [dj.dekoning@bbsrc.ac.uk](mailto:dj.dekoning@bbsrc.ac.uk) (D.-J. de Koning).

<sup>1</sup> Present address: Linnaeus Centre for Bioinformatics, Uppsala University, BMC Box 598, SE-751 24 Uppsala, Sweden.

immune response can now benefit from expertise and infrastructure that has been generated by studying traditional traits (e.g. Sonstegard and Gasbarre, 2001). A combination of proven approaches in QTL detection and emerging technologies in gene transcription analysis can provide a fast track for unravelling the genetic networks underlying differences in immune response. We will briefly review some existing approaches and subsequently elaborate on how they can be merged into a powerful genetical genomics approach.

## 2. QTL approaches

For most infectious diseases, experimental populations have to be custom bred and challenged to study genetic differences in immune response and map genetic loci underlying these differences (Fig. 1A). However, some disease traits are routinely recorded because of their economic importance and the available infrastructure. As a result, QTL have been detected for mastitis (Klungland et al., 2001; Schulman et al., 2004) and somatic cell count (reviewed by Khatkar et al., 2004) using existing family structures and routinely collected data in dairy cattle.

In poultry, experimental crosses of lines that differ in their susceptibility against a specific disease have been used to map QTL affecting Marek's disease (e.g. Vallejo et al., 1998), *E. coli* infection (Yonash et al., 2001), and coccidiosis (Zhu et al., 2003). Siwek et al. (2003a, 2003b) used an F<sub>2</sub> cross between chicken lines that were divergently selected for response against sheep red blood cells (SRBC) to map QTL underlying response to SRBC, Keyhole Limpet Hemocyanin, and *Mycobacterium butyricum* challenge.

For many infectious diseases, model animals like mice have been used to study the genetics of susceptibility (Lee et al., 2003). For transmissible spongiform encephalopathies (TSE), like BSE and scrapie, mice have been used as a model organism. Beside a mutation in the *PRP* gene, other loci affecting susceptibility to TSE have been mapped in mice (e.g. Moreno et al., 2003; including an overview of other studies), cattle (Hernandez-Sanchez et al., 2002; Zhang et al., 2004) and sheep (Moreno et al., 2002). Another example is trypanosomosis, where susceptibility loci have been mapped in mice (Kemp et al., 1997) and subsequently in

cattle (Hanotte et al., 2003). Although studying model species could provide a short cut in theory, it is not clear how well the results translate from, for instance, mice to livestock when the results vary widely between the mouse studies (Moreno et al., 2003). This is also illustrated by the trypanosomosis example where the three major susceptibility loci described in mice, do not correspond very well with the 18 cattle loci with more modest effects (Kemp et al., 1997; Hanotte et al., 2003).

## 3. Exploiting gene expression technology

The recent development of high throughput gene-expression technologies, such as microarrays, has given rise to a plethora of new research hypotheses and possibilities. Extensive reviews are available about the application (e.g. Butte, 2002), design (e.g. Churchill, 2002), and analysis (e.g. Quackenbush, 2002) of microarray studies. With regard to infectious diseases, microarrays have been proposed to study gene-expression in the parasite (Malaria: Rathod et al., 2002; Trypanosomosis: El Sayed et al., 2000; Hill et al., 2005) as well as host response following infection (e.g. *Mycobacterium paratuberculosis* infection in cattle: Coussens et al., 2004; *Eimeria* infection in poultry: Min et al., 2003). Variation in gene expression between animals/lines following disease challenge is of particular interest to unravel the genetics underlying immune response (Fig. 1B). In mice, comparisons in gene expression between susceptible and 'resistant' lines have been made for several diseases like colon cancer (Guda et al., 2003), corneal infection by *Pseudomonas aeruginosa* (Huang and Hazlett, 2003), and gallstone disease (Dyck et al., 2003).

A first step in the integration of QTL studies with gene expression studies can be the identification of differentially expressed genes as positional candidate genes for the QTL (Fig. 1). A nice example of this strategy is provided by Liu et al. (2001), who created an experimental F<sub>2</sub> cross between susceptible and resistant inbred lines to map QTL underlying resistance against Marek's disease (Vallejo et al., 1998). The founder lines were also used for a microarray study to identify genes that were differentially expressed between the two lines following artificial infection. Fifteen of these genes were subsequently mapped onto the chicken genome and

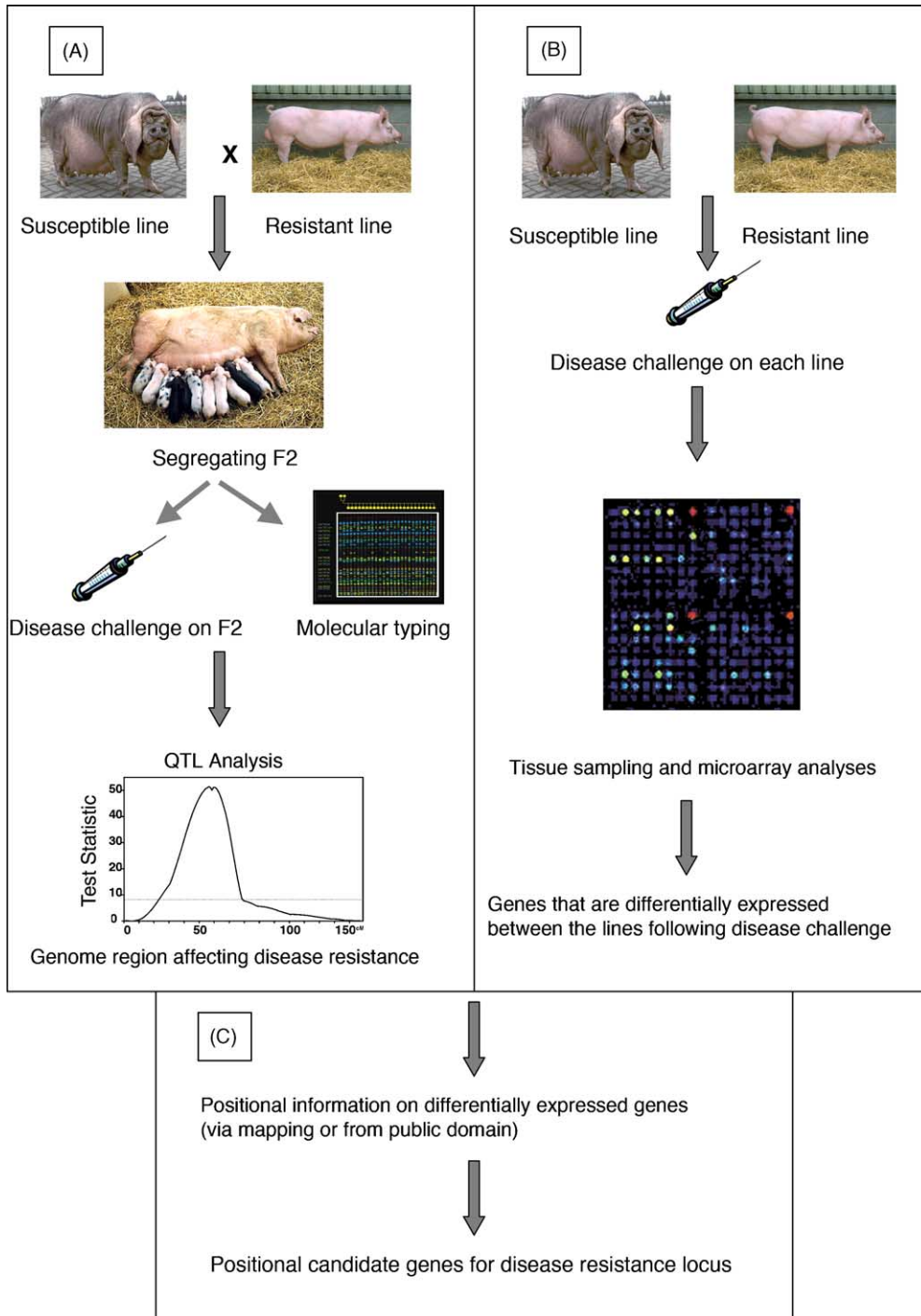


Fig. 1. Graphical representation of applications of functional genomics to disease resistance and their potential integration: QTL mapping (A), gene-expression analysis (B), and combining the information from both (C). The models are based on two lines that differ in their susceptibility for a hypothetical disease.

two of them mapped to a QTL region for Marek's resistance. One of these confirmed a previously identified effect within the growth hormone (*GHI*) gene (Liu et al., 2001). Although the study by Liu et al. (2001) only followed up 15 genes with a mapping effort, the paper illustrates the potential of combining gene expression and QTL data. The resolution of the QTL studies is a potential problem for this approach. The confidence intervals of QTL are usually several tens of centiMorgans wide and harbour hundreds of genes, many of which could be differentially expressed between the two founder lines. Wayne and McIntyre (2002) demonstrate a strategy in *Drosophila*, where they first fine-mapped the QTL to a region with 548 candidate genes, which they subsequently narrowed down to 34 genes that were differentially expressed between the founder lines. It must be noted that the molecular tools for fine mapping and the genome information available for *Drosophila* is much more extensive than for livestock species. A similar project is currently underway for trypanosomiasis (<http://www.genomics.liv.ac.uk/tryps/>). In both cattle and mice, expression differences will be studied between resistant and susceptible breeds, following trypanosome challenge. Subsequently, gene expression will be studied on congenic mouse lines for three QTL regions while in cattle gene-expression will be measured in a backcross design (Steve Kemp, Liverpool *personal communication*).

#### 4. Genetical genomics: mapping genetic loci underlying variation in gene expression

Jansen and Nap (2001) described a strategy to combine genome-wide linkage analysis with expression studies. They propose the use of a segregating population where each member of the population is characterised for molecular markers and gene expression levels. The gene expression values are treated as a quantitative phenotype and the marker genotypes are used to map loci affecting gene expression levels (expression QTL  $\Rightarrow$  eQTL). Expression differences for a given gene can be caused by variation within the gene ('*cis*-acting') or by variation elsewhere in the genome ('*trans*-acting'). Jansen (2003) describe how studying many combinations of gene variants within a segregating population may be more powerful and

informative than studying the effect of a single mutation or knockout in a uniform background. One of the promises of genetical genomics is the unravelling of networks of gene regulation. This is achieved by studying shared locations of eQTL across genes combined with information of (candidate) genes within the eQTL regions and their respective eQTL.

The first genetical genomics study based on experimental data was carried out in budding yeast (Brem et al., 2002). In this study gene expression was measured on two budding yeast lines as well as on 40 haploid segregants (offspring) from a cross between the lines. From the 1500 genes that were differentially expressed, eQTL were detected for 308 gene transcripts. Another 262 QTL were detected for genes that showed no significant difference in expression between the two parental lines. There was a clear distinction between eQTL that were exclusively *cis*-acting and *trans*-acting eQTL that affected between 7 and 94 gene transcripts (Brem et al., 2002). In a more powerful recent study of 86 segregants in the same cross, Yvert et al. (2003) report detection of 2294 eQTL out of 6215 yeast genes. In these crosses, about 25% of the eQTL were *cis*-acting and the other 75% were *trans*-acting. A surprising find was that the *trans*-acting eQTL were not enriched for transcription factors as might have been anticipated (Yvert et al., 2003).

Schadt et al. (2003) describe three projects to elucidate the genetic regulation of gene expression in maize, mice, and man. In the mouse experiment, 111 F<sub>2</sub> mice and the founder lines were characterised for 23,574 gene transcripts. For 3701 genes, at least a single eQTL was detected even though only 2123 of these genes showed a significant expression difference in the founder lines. The F<sub>2</sub> were also segregating for obesity related traits and by combining the subcutaneous fat pad mass (FPM) measurements with the expression data it was demonstrated that the animals with high FPM could be divided into two groups, each with specific gene expression patterns and different QTL underlying the high FPM phenotype (Schadt et al., 2003). For the maize experiment, 76 F<sub>2</sub>-derived F<sub>3</sub> plants were arrayed for 18,805 genes, which revealed 6481 genes with at least a single eQTL. The authors describe a novel type of epistatic interaction in the expression levels for two genes that is only evident within the different genotype classes of the eQTL for

these genes (Schadt et al., 2003). However, the methods used to detect this new type of gene interaction are not formally described, let alone validated. In an attempt to include the genetic regulation of gene transcription in humans, gene expression was measured on lymphoid cell lines of 56 individuals from four families (Schadt et al., 2003). While 2726 genes were differentially expressed between the family founders, about 29% of these had a significant genetic component. It was stated that given the complexity of the family data, the 56 individuals were insufficient to do a full-blown eQTL analyses on this data.

A public resource for eQTL mapping had been made available through WebQTL (<http://www.webqtl.org/>; Chesler et al., 2004). The database contains expression data from brain tissues and 'traditional' phenotypes' as well as genome-wide molecular marker information for about 30 recombinant inbred lines (RILs) of mice (BXD) that were derived from a cross between C57BL/6J (B) and DBA/2J (D). Traits or gene transcripts can be selected and subsequently used for a number of analyses from simple correlation analysis to genome-wide interval mapping. An additional advantage is that researchers can purchase the BXD mice, add their phenotypes to the database and perform QTL analysis without any genotyping efforts as well as exploit all the information on traits that are already in the database.

Table 1 shows a comparison of the statistical power to detect additive eQTL for the various eQTL designs to date. The statistical power has been calculated deterministically following Lynch and Walsh (1998).

By imposing the same arbitrary threshold for every experiment ( $P < 0.001$ ), we ignored different genome lengths and expected recombination levels between designs, thus simplifying the results in Table 1. From the published experiments, the experiment by Yvert et al. (2003) has the highest statistical power to detect eQTL (Table 1). Compared to the experimental designs commonly encountered in QTL detection, all eQTL experiments to date are very small, and this is clearly reflected in their statistical power (Table 1). This is easily attributed to the high cost of gene expression analysis. The reduced power can be partly overcome by the use of haploid lines (yeast) or recombinant inbred lines (RIL), which greatly simplifies the genetic model and increases the power of detecting eQTL. An experiment with  $n$  RILs has the same statistical power to detect eQTL as an  $F_2$  population of size  $2n$  (Table 1). RILs have a further advantage that replicate tissue samples can be taken on identical genotypes, allowing a reduction in the environmental variance of the gene expression levels when taking multiple measurements for the same genotype. This is clearly illustrated in Table 1 where, for the RI mice, the power to detect QTL doubles when using three replicate observations if the repeatability of gene-expression is 0.50.

Even though the initial results show some of the potential of eQTL mapping, there are many issues that have not yet been thoroughly investigated. A major problem is that of multiple testing. In eQTL mapping this occurs on two levels. First, multiple correlated tests are carried out during the genome scan for eQTL.

Table 1

Comparison of statistical power to detect QTL at  $P < 0.001$  for the different eQTL designs to date

| Reference             | Population                                 | $N$ | Statistical power for different QTL effects<br>(QTL effects in phenotypic SD) |      |      |      |      |
|-----------------------|--------------------------------------------|-----|-------------------------------------------------------------------------------|------|------|------|------|
|                       |                                            |     | 0.25                                                                          | 0.40 | 0.5  | 0.6  | 0.75 |
| Brem et al. (2002)    | Haploid yeast                              | 40  | 0.05                                                                          | 0.2  | 0.51 | 0.73 | 0.99 |
| Yvert et al. (2003)   | Haploid yeast                              | 86  | 0.16                                                                          | 0.67 | 0.94 | 0.99 | 0.99 |
| Schadt et al. (2003)  | $F_2$ mice                                 | 111 | 0.07                                                                          | 0.37 | 0.68 | 0.9  | 0.99 |
| Schadt et al. (2003)  | $F_3$ maize                                | 76  | 0.04                                                                          | 0.19 | 0.41 | 0.67 | 0.94 |
| Chesler et al. (2004) | RI mice                                    | 33  | 0.03                                                                          | 0.14 | 0.31 | 0.57 | 0.93 |
|                       | One replicate                              |     |                                                                               |      |      |      |      |
| Chesler et al. (2004) | RI mice 3<br>Three replicates <sup>a</sup> | 33  | 0.05                                                                          | 0.29 | 0.62 | 0.91 | 0.99 |

All models compared for a point-wise threshold of  $P < 0.001$ , which corresponds to a LOD score of 3.0 for an  $F_2$  design.

<sup>a</sup> Assuming a repeatability of 0.50 for gene transcripts and three replicates for every RI line.

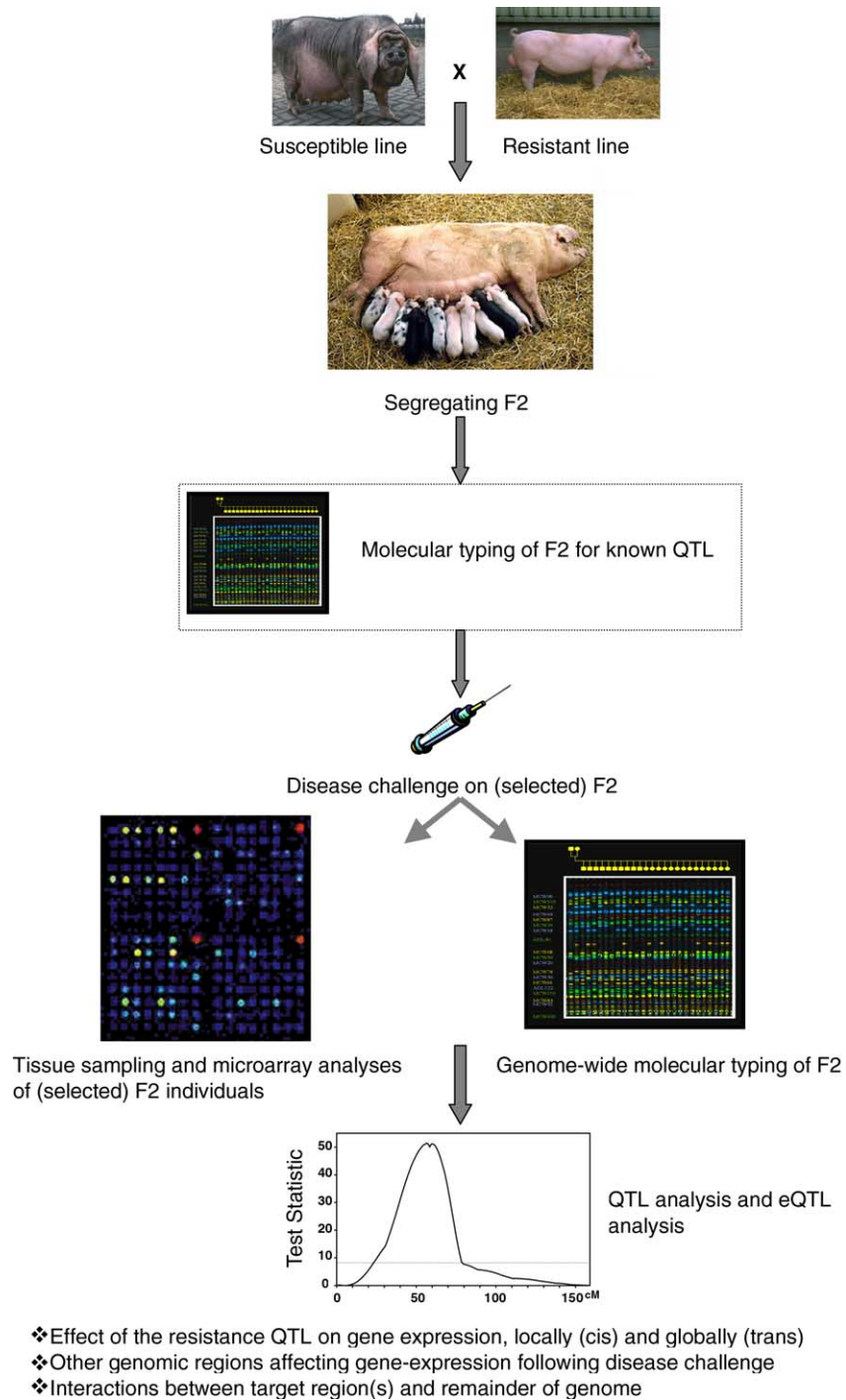


Fig. 2. Graphical representation of the application of genetical genomics to disease resistance. The models are based on two lines that differ in their susceptibility for a hypothetical disease. The box enclosed by the dotted line represents an optional step to target the eQTL study at known QTL for disease resistance.

Secondly, eQTL analyses are performed for thousands of, potentially highly correlated, gene expression levels. Storey and Tibshirani (2003) propose an adaptation of the false discovery rate (FDR) targeted at genome-wide experiments to provide a better balance between statistical stringency and power to detect true effects. Nevertheless, significance thresholds are likely to be more stringent than the level of  $P < 0.001$  used in Table 1, resulting in a further reduction in statistical power. Although efforts have been made to study the properties of eQTL analysis via simulation studies (Perez-Enciso et al., 2003; Perez-Enciso, 2004), the scarce results to date complicate the choice of realistic values for eQTL effects and the effect of expression differences on a physiological trait (Perez-Enciso et al., 2003). Using real expression data and simulated SNP genotypes, Perez-Enciso et al. (2003) showed that, even though simulated genotypes were randomly assigned to the individuals, the resulting eQTL had a very high level of clustering. The occurrence of eQTL hotspots may therefore simply reflect the highly correlated nature of many gene-expression measurements. Whether this high correlation reflects the effect of main hubs of gene regulation in the genome or some underlying technical or environmental correlation is still unclear. It is expected that because of their modest size, most experiments to date only detect the largest of real eQTL effects and may well suffer from a high rate of false positives. For a true evaluation of the utility of eQTL mapping, more powerful experiments are needed than those published to date. A larger number of individuals is necessary to increase the power to detect the simple additive effects, but also to accommodate more advanced genetic models that include interactions between eQTL (Carlborg and Haley, 2004).

### 5. Potential for eQTL mapping in disease resistance

The applications of eQTL mapping to date have mostly focussed on the genetic basis of gene expression differences between lines without linking this information with other phenotypic information. The exception is Schadt et al. (2003) who included obesity related phenotypes and obtained gene expression patterns and eQTL that were specific for one of

two obesity types. Fig. 2 outlines a possible strategy to apply eQTL mapping to identify genomic loci that affect gene expression following disease challenge. Compared to the strategy outlined by Jansen and Nap (2001), the only difference is the inclusion of a disease challenge to monitor differential effects of gene expression that may be related to disease resistance. It must be noted that this approach will also detect loci underlying gene-expression differences in the founder lines that are not related to the disease. While this provides important additional information with regard to genetic control of gene regulation it is important to realise that many eQTL will not be relevant to the targeted disease.

With the absence of recombinant inbred lines for livestock species, other options to increase the power of an eQTL design should be explored. One possibility is to include knowledge about identified QTL into the design of the study. By selecting only individuals that are homozygous for the previously detected QTL (either ++ or --) for disease challenge, expression studies and genome-wide molecular typing, the power to detect eQTL in the selected regions then becomes equal to that of an RI line (i.e. the same power as an  $F_2$  of twice the actual size). Table 2 compares the power of eQTL mapping for  $F_2$  crosses of different size and the power for the regions where only homozygous animals are used. The number of individuals is considerably larger than those reported for existing studies and this is no coincidence (Table 2). In order to unleash the true potential of combining gene-expression and QTL methodology, experiments must be of sufficient size to facilitate meaningful analyses based on standard and more advanced genetical models. From experience in 'traditional' QTL analysis, we feel that an  $F_2$  of 200 is the minimum population size to detect major QTL effects, while an experiment of 400 animals gives good statistical power to detect also more subtle effects (Table 2). Table 2 illustrates very clearly the increased statistical power to detect eQTL for the selected regions. A further advantage of the reduced complexity in the selected regions is the possibility to estimate epistatic interactions between the target regions and between a target region and the remainder of the genome (Carlborg et al., 2003).

An inherent problem to the targeted approach is that in order to obtain animals that are homozygous for multiple QTL, the size of the initial experimental

Table 2

Comparison of statistical power to detect QTL at  $P < 0.001$  for different  $F_2$  designs where animals are selected to be homozygous for identified QTL

|                                          | Statistical power for different QTL effects<br>(QTL effects in phenotypic SD) |      |      |      |      | Number of $F_2$ required<br>(20 cM confidence interval) <sup>a</sup> |      |
|------------------------------------------|-------------------------------------------------------------------------------|------|------|------|------|----------------------------------------------------------------------|------|
|                                          | $N$                                                                           | 0.25 | 0.40 | 0.5  | 0.6  | 1 QTL                                                                | 2QTL |
| Genome scan                              | 100                                                                           | 0.06 | 0.31 | 0.6  | 0.85 |                                                                      |      |
|                                          | 200                                                                           | 0.21 | 0.77 | 0.96 | 0.99 |                                                                      |      |
|                                          | 300                                                                           | 0.41 | 0.95 | 0.99 | 0.99 |                                                                      |      |
|                                          | 400                                                                           | 0.60 | 0.99 | 0.99 | 0.99 |                                                                      |      |
| Selected regions<br>(homozygous for QTL) | 100                                                                           | 0.20 | 0.78 | 0.97 | 0.99 | 350                                                                  | 1150 |
|                                          | 200                                                                           | 0.60 | 0.99 | 0.99 | 0.99 | 700                                                                  | 2200 |
|                                          | 300                                                                           | 0.86 | 0.99 | 0.99 | 0.99 | 1000                                                                 | 3200 |
|                                          | 400                                                                           | 0.96 | 0.99 | 0.99 | 0.99 | 1350                                                                 | 4200 |

All models compared for a point-wise threshold of  $P < 0.001$ , which corresponds to a LOD score of 3.0 for an  $F_2$  design.

<sup>a</sup> The number of initial  $F_2$  required (rounded to nearest 50) to obtain at least  $N$  animals (with  $> 95\%$  confidence) that are homozygous for the two markers flanking the QTL interval of 20 cM.

population needs to be increased dramatically with the number of targeted QTL. For instance, when a QTL has a confidence region of 20 cM that is flanked by fully informative markers (A and B),  $\sim 80\%$  of gametes in  $F_1$  is expected to be A1B1 or A2B2 while the other 20% is recombinant (A1B2 or A2B1). As a result, 16% of the offspring is expected to be homozygous for both markers flanking the QTL and are therefore assumed to be homozygous for the QTL (ignoring double recombination events). To obtain the number of  $F_2$  that are required to get  $n$  individuals that are homozygous for a QTL, the 95% lower cut-off point of a normal distribution was applied (with mean =  $NP$  and standard deviation =  $NP(1 - P)$ , where  $N$  is total number of  $F_2$  and  $P$  the expected proportion of homozygous individuals) (Table 2). These required numbers are based on the total number of homozygous individuals and do not take into account the number of individuals within each genotype class. Including the latter would require even larger sizes. Even though these required numbers may seem quite daunting, the entire  $F_2$  is only typed for markers flanking the QTL region(s), which can be done at a very early stage.

The combination of traditional QTL mapping with eQTL mapping will provide crucial information about the nature of the disease QTL. It will elucidate how a disease QTL affects gene expression following infection: Does it mainly affect expression of genes around the QTL (*cis*-acting) or does it represent a hot spot of

gene regulation for many unlinked genes (*trans*-acting). Identification of genes that are affected by the eQTL provide a first step to unravelling the genetic pathway underlying the response to the disease challenge. The eQTL analysis may identify further regions related to the disease response that remained unnoticed or non-significant in the standard QTL analysis. Obviously, many eQTL will also relate to general line differences that are not related to disease susceptibility and while these eQTL still have tremendous value to elucidate gene regulation networks in general, it will be difficult to disentangle them from those that are related to the disease challenge. One possible solution is a prior analysis of gene expression before and after disease challenge to distinguish between genes that are related to the disease response (differentially regulated before and after challenge) and those that are not.

With regard to response to disease challenge, it is important to note it may not be necessary to sacrifice and dissect the animal for RNA collection. It is possible to collect blood samples after infection and measure gene expression levels on lymphocytes (Liu et al., 2001; Min et al., 2003). Alternatively, immune cells like macrophages can be collected from healthy animals and challenged in vitro (Stephen Bishop, Roslin Institute, personal communication). These studies of gene expression at the cellular level have the advantage that the animals are still alive and can be used to monitor progress of the disease or used to collect other

phenotypes later in life. Another alternative is the use of cell lines to measure expression levels as illustrated for humans by Schadt et al. (2003). The choice between studying immune cells and/or specific tissues is partly dictated by the research question: Gene expression levels on the immune cells will reflect the response to infection for genes related to the actual immune response while measurements on specific tissues will also reflect the effects of differences in severity of disease between animals.

## 6. Concluding remarks

While the studies to date have been useful indicators of the potential of genetical genomics, their size is too small to assess the full merit of eQTL mapping. High costs, especially for the expression studies, will be associated with sizing-up these experiments to obtain the same statistical power as conventional QTL experiments we feel this is necessary to allow meaningful analyses. The high cost also poses further emphasis on careful experimental design and efficient use of animal resources. Many studies will require dissection of animals, while only one or two tissues can be targeted for expression studies. We strongly recommend that additional tissues are stored and made available to collaborators. In this way, other investigators can do additional analyses (gene expression or other) and utilise existing genotype information and eQTL knowledge for the system under study. Both the multi-disciplinary nature of eQTL projects as well as the high cost make this field of research particularly attractive for scientific consortia that subsequently make (part of) their resources available to the research community.

## Acknowledgements

DJK and CSH are supported by the BBSRC while ÖC acknowledges support from the Knut and Alice Wallenberg foundation.

## References

Andersson, L., 2001. Genetic dissection of phenotypic diversity in farm animals. *Nat. Rev. Genet.* 2, 130–138.

- Andersson, L., Georges, M., 2004. Domestic-animal genomics: deciphering the genetics of complex traits. *Nat. Rev. Genet.* 5, 202–212.
- Brem, R.B., Yvert, G., Clinton, R., Kruglyak, L., 2002. Genetic dissection of transcriptional regulation in budding yeast. *Science* 296, 752–755.
- Butte, A., 2002. The use and analysis of microarray data. *Nat. Rev. Drug Discov.* 1, 951–960.
- Carlborg, O., Kerje, S., Schutz, K., Jacobsson, L., Jensen, P., Andersson, L., 2003. A global search reveals epistatic interaction between QTL for early growth in the chicken. *Genome Res.* 13, 413–421.
- Carlborg, O., Haley, C.S., 2004. Epistasis: too often neglected in complex trait studies? *Nat. Rev. Genet.* 5, 618–625.
- Chesler, E.J., Lu, L., Wang, J., Williams, R.W., Manly, K.F., 2004. WebQTL: rapid exploratory analysis of gene expression and genetic networks for brain and behavior. *Nat. Neurosci.* 7, 485–486.
- Churchill, G.A., 2002. Fundamentals of experimental design for cDNA microarrays. *Nat. Genet.* 32 (Suppl.), 490–495.
- Coussens, P.M., Jeffers, A., Colvin, C., 2004. Rapid and transient activation of gene expression in peripheral blood mononuclear cells from Johne's disease positive cows exposed to *Mycobacterium paratuberculosis* in vitro. *Microb. Pathog.* 36, 93–108.
- Dyck, P.A., Hoda, F., Osmer, E.S., Green, R.M., 2003. Microarray analysis of hepatic gene expression in gallstone-susceptible and gallstone-resistant mice. *Mamm. Genome* 14, 601–610.
- El Sayed, N.M., Hegde, P., Quackenbush, J., Melville, S.E., Donelson, J.E., 2000. The African trypanosome genome. *Int. J. Parasitol.* 30, 329–345.
- Guda, K., Cui, H., Garg, S., Dong, M., Nambiar, P.R., Achenie, L.E., Rosenberg, D.W., 2003. Multistage gene expression profiling in a differentially susceptible mouse colon cancer model. *Cancer Lett.* 191, 17–25.
- Hanotte, O., Ronin, Y., Agaba, M., Nilsson, P., Gelhaus, A., Horstmann, R., Sugimoto, Y., Kemp, S., Gibson, J., Korol, A., Soller, M., Teale, A., 2003. Mapping of quantitative trait loci controlling trypanotolerance in a cross of tolerant West African N'Dama and susceptible East African Boran cattle. *Proc. Natl. Acad. Sci. U.S.A.* 100, 7443–7448.
- Hernandez-Sanchez, J., Waddington, D., Wiener, P., Haley, C.S., Williams, J.L., 2002. Genome-wide search for markers associated with bovine spongiform encephalopathy. *Mamm. Genome* 13, 164–168.
- Hill, E.W., O'Gorman, G.M., Agaba, M., Gibson, J.P., Hanotte, O., Kemp, S.J., Naessens, J., Coussens, P.M., MacHugh, D.E., 2005. Understanding bovine trypanosomiasis and trypanotolerance: the promise of functional genomics 105, 247–258.
- Huang, X., Hazlett, L.D., 2003. Analysis of *Pseudomonas aeruginosa* corneal infection using an oligonucleotide microarray. *Invest Ophthalmol. Vis. Sci.* 44, 3409–3416.
- Jansen, R.C., 2003. Studying complex biological systems using multifactorial perturbation. *Nat. Rev. Genet.* 4, 145–151.
- Jansen, R.C., Nap, J.P., 2001. Genetical genomics: the added value from segregation. *Trends Genet.* 17, 388–391.
- Kemp, S.J., Iraqi, F., Darvasi, A., Soller, M., Teale, A.J., 1997. Localization of genes controlling resistance to trypanosomiasis in mice. *Nat. Genet.* 16, 194–196.

- Khatkar, M.S., Thomson, P.C., Tammen, I., Raadsma, H.W., 2004. Quantitative trait loci mapping in dairy cattle: review and meta-analysis. *Genet. Sel. Evol.* 36, 163–190.
- Klungland, H., Sabry, A., Heringstad, B., Olsen, H.G., Gomez-Raya, L., Vage, D.I., Olsaker, I., Odegard, J., Klemetsdal, G., Schulman, N., Vilkki, J., Ruane, J., Aasland, M., Ronningen, K., Lien, S., 2001. Quantitative trait loci affecting clinical mastitis and somatic cell count in dairy cattle. *Mamm. Genome* 12, 837–842.
- Lee, S.H., Dimock, K., Gray, D.A., Beauchemin, N., Holmes, K.V., Belouchi, M., Reelson, J., Vidal, S.M., 2003. Maneuvering for advantage: the genetics of mouse susceptibility to virus infection. *Trends Genet.* 19, 447–457.
- Liu, H.C., Cheng, H.H., Tirunagaru, V., Sofer, L., Burnside, J., 2001. A strategy to identify positional candidate genes conferring Marek's disease resistance by integrating DNA microarrays and genetic mapping. *Anim. Genet.* 32, 351–359.
- Lynch, M., Walsh, J.B., 1998. *Genetics and Analysis of Complex Traits*. Sinauer Associates Inc., Sunderland.
- Min, W., Lillehoj, H.S., Kim, S., Zhu, J.J., Beard, H., Alkharouf, N., Matthews, B.F., 2003. Profiling local gene expression changes associated with *Eimeria maxima* and *Eimeria acervulina* using cDNA microarray. *Appl. Microbiol. Biotechnol.* 62, 392–399.
- Moreno, C.R., Lantier, I., Lantier, F., Andreolotti, O., Vaiman, D., Sarradin, P., Echeynne, F., Cribru, E.P., Cossedu, G., Elsen, J.M., 2002. Transposition to sheep of mouse quantitative trait loci (QTL) influencing susceptibility to prion diseases. In: *Proceedings of the Seventh World Congress on Genetic Application on Livestock Production*. Montpellier, France (CD-ROM, contribution 13–21).
- Moreno, C.R., Lantier, F., Lantier, I., Sarradin, P., Elsen, J.M., 2003. Detection of new quantitative trait loci for susceptibility to transmissible spongiform encephalopathies in mice. *Genetics* 165, 2085–2091.
- Perez-Enciso, M., Toro, M.A., Tenenhaus, M., Gianola, D., 2003. Combining gene expression and molecular marker information for mapping complex trait genes: a simulation study. *Genetics* 164, 1597–1606.
- Perez-Enciso, M., 2004. In silico study of transcriptome genetic variation in outbred populations. *Genetics* 166, 547–554.
- Quackenbush, J., 2002. Microarray data normalization and transformation. *Nat. Genet.* 32 (Suppl.), 496–501.
- Rathod, P.K., Ganesan, K., Hayward, R.E., Bozdech, Z., DeRisi, J.L., 2002. DNA microarrays for malaria. *Trends Parasitol.* 18, 39–45.
- Schadt, E.E., Monks, S.A., Drake, T.A., Lusk, A.J., Che, N., Colinayo, V., Ruff, T.G., Milligan, S.B., Lamb, J.R., Cavet, G., Linsley, P.S., Mao, M., Stoughton, R.B., Friend, S.H., 2003. Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422, 297–302.
- Schulman, N.F., Viitala, S.M., de Koning, D.J., Virta, J., Maki-Tanila, A., Vilkki, J.H., 2004. Quantitative trait Loci for health traits in Finnish Ayrshire cattle. *J. Dairy Sci.* 87, 443–449.
- Siwek, M., Buitenhuis, A.J., Cornelissen, S.J., Nieuwland, M.G., Bovenhuis, H., Crooijmans, R.P., Groenen, M.A., Vries-Reilingh, G., Parmentier, H.K., van der Poel, J.J., 2003a. Detection of different quantitative trait loci for antibody responses to keyhole limpet hemocyanin and *Mycobacterium butyricum* in two unrelated populations of laying hens. *Poult. Sci.* 82, 1845–1852.
- Siwek, M., Cornelissen, S.J., Nieuwland, M.G., Buitenhuis, A.J., Bovenhuis, H., Crooijmans, R.P., Groenen, M.A., Vries-Reilingh, G., Parmentier, H.K., van der Poel, J.J., 2003b. Detection of QTL for immune response to sheep red blood cells in laying hens. *Anim. Genet.* 34, 422–428.
- Sonstegard, T.S., Gasbarre, L.C., 2001. Genomic tools to improve parasite resistance. *Vet. Parasitol.* 101, 387–403.
- Stear, M.J., Bishop, S.C., Mallard, B.A., Raadsma, H., 2001. The sustainability, feasibility and desirability of breeding livestock for disease resistance. *Res. Vet. Sci.* 71, 1–7.
- Storey, J.D., Tibshirani, R., 2003. Statistical significance for genome-wide studies. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9440–9445.
- Vallejo, R.L., Bacon, L.D., Liu, H.C., Witter, R.L., Groenen, M.A., Hillel, J., Cheng, H.H., 1998. Genetic mapping of quantitative trait loci affecting susceptibility to Marek's disease virus induced tumors in F<sub>2</sub> intercross chickens. *Genetics* 148, 349–360.
- Wayne, M.L., McIntyre, L.M., 2002. Combining mapping and arraying: an approach to candidate gene identification. *Proc. Natl. Acad. Sci. U.S.A.* 99, 14903–14906.
- Yonash, N., Cheng, H.H., Hillel, J., Heller, D.E., Cahaner, A., 2001. DNA microsatellites linked to quantitative trait loci affecting antibody response and survival rate in meat-type chickens. *Poult. Sci.* 80, 22–28.
- Yvert, G., Brem, R.B., Whittle, J., Akey, J.M., Foss, E., Smith, E.N., Mackelprang, R., Kruglyak, L., 2003. Trans-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcription factors. *Nat. Genet.* 35, 57–64.
- Zhang, C., de Koning, D.J., Hernandez-Sanchez, J., Haley, C.S., Williams, J.L., Wiener, P., 2004. Mapping of multiple quantitative trait loci affecting bovine spongiform encephalopathy (BSE). *Genetics* 167, 1863–1872.
- Zhu, J.J., Lillehoj, H.S., Allen, P.C., Van Tassell, C.P., Sonstegard, T.S., Cheng, H.H., Pollock, D., Sadjadi, M., Min, W., Emara, M.G., 2003. Mapping quantitative trait loci associated with resistance to coccidiosis and growth. *Poult. Sci.* 82, 9–16.