G3.1 E-mail DJ.deKoning@BBSRC.AC.UK

Towards Genetical genomics in Livestock

D. J. de Koning, C. P. Cabrera and C. S. Haley
The Roslin Institute, Roslin Biocentre, Roslin, EH25 9PS, United Kingdom

This paper is modified from a paper presented at PSA 2006 and submitted to Poultry Science

ABSTRACT

Microarrays have been widely implemented across the life sciences although there is still debate on the most effective uses of such transcriptomics approaches. In genetical genomics, gene expression measurements are treated as quantitative traits and genome regions affecting expression levels are denoted as expression quantitative trait loci or eQTL. The detected eQTL can either represent a locus that lies close to the gene that is being controlled (*cis*-acting) or one or more loci that are unlinked to the gene that is being controlled (*trans*-acting). One powerful outcome of genetical genomics is the reconstruction of genetic pathways underlying complex trait variation. Because of the modest size of experiments to date, genetical genomics may fall short of its promise to unravel genetic networks. We propose to combine expression studies with fine mapping of functional trait loci. This synergistic approach facilitates the implementation of genetical genomics for species without inbred resources but is equally applicable to model species. Among livestock species, poultry is well placed to embrace this technology with the availability of the chicken genome sequence, microarrays for various platforms as well as experimental populations in which QTL have been mapped. Other species are catching with genome sequences becoming available for cattle and advanced plans for the pig genome. In the build-up towards full-blown eQTL studies, we can study the effects of known candidate genes or marked QTL at the gene expression level in more focussed studies. To demonstrate the potential of genetical genomics, we have identified the *cis* and *trans* effects for a functional body weight QTL on chicken chromosome 4 in breast tissue samples from chickens with contrasting QTL genotypes.

Key words: Experimental Design, Fine Mapping, Gene Expression, Quantitative Trait Locus

INTRODUCTION

Dissecting the genetic control of variation in complex traits and identifying underlying loci controlling such variation has proved to be very challenging. While quantitative trait locus (**QTL**) detection has been successful in identifying chromosomal regions associated with a wide range of complex traits in many different species [e.g. experimental crosses[1], livestock[2-4], humans[5]], these regions are sufficiently large to contain hundreds if not thousands of potential candidate genes. Further fine mapping of these QTL to reduce the size of these regions and hence refine the list of potential candidate genes can be achieved by creating additional recombination events through selective breeding[6] or by exploiting historical recombinations[7]

An approach that has great promise to make a major contribution to the dissection of complex traits is genetical genomics; the combined study of gene-expression and marker genotypes in a segregating population [8,9].

Genetical genomics is aimed at detecting genomic loci that control variation in gene expression, so called expression QTL (eQTL) (to distinguish them from functional QTL that affect traits at the whole-organism level). The detected eQTL can either represent a locus that lies close to the gene that is being controlled (*cis*-acting) or one or more loci that are unlinked to the gene that is being controlled (*trans*-acting) [8]. A major promise

of genetical genomics is that by examining the relationship between transcript location, location of eQTL and pleiotropic effects of eQTL, it might be able to reconstruct genetic pathways that underlie phenotypic variation[8]. Additional information to reconstruct pathways comes from the correlations between and among gene expression measurements and functional traits[10] and the epistatic interactions between eQTL and functional QTL[11]. Therefore, genetical genomics can be exploited as an additional tool to dissect phenotypic variation into its underlying components and elucidate how these components interact. If successful, genetical genomics will enhance and accelerate the characterization of functional QTL, which remains an arduous task, even in model species. At present, several studies have demonstrated the feasibility of eQTL studies and some of these have successfully integrated eQTL and gene expression with data on traditional phenotypes. What all these studies have in common is that, in comparison to 'traditional' QTL studies of functional traits, the sizes of the experiments are modest to small[12]. Consequently, the power of the studies is low and many important QTL will not have been detected and interactions between QTL will have been missed. Thus the results to date have not been very successful at reconstructing genetic pathways or identifying genes underlying functional trait variation. Therefore, more powerful experiments addressing these issues are necessary to realise the full potential of eQTL mapping.

Following a case study of how a QTL experiment has been integrated with microarray analyses in poultry, we outline an experimental strategy to improve the efficiency of future eQTL studies. We subsequently introduce a targeted approach to study the gene expression effects of a marked QTL.

Integrating QTL and Gene Expression studies

In a number of cases, traditional QTL studies have been supplemented with microarray data in an attempt to move from a functional QTL to the underlying gene(s)[13]. Below, we outline a case study where detection of functional QTL was followed up by a gene expression analysis. In this example, microarray experiments were carried out on the founder lines of the study. The underlying idea was that genes that were differentially expressed between the founder lines AND were located in the areas of the QTL that were found in the cross resulting from these lines, would be prime positional candidates for the functional genes underlying the QTL. In genetical genomics terms, this type of analysis explores whether the functional QTL is also a *cis*-acting eQTL. It would be much more difficult for such a study to determine the genetic basis of a QTL that had its functional effect through *trans*-acting regulation of expression of genes located outside the QTL region. This is because there are likely to be many differences in expression between lines for genes across the genome. This study provides no information on where in the genome the control of those expression differences lies and hence which of these genes are associated with the QTL region.

Resistance to Marek's disease in chicken. Marek's disease (MD) is an infectious viral disease and a member of the herpes virus family. MD costs the poultry industry about 1 billion USD per annum. To study the genetic control of MD susceptibility, an experimental cross was established between a resistant and a susceptible inbred line of chicken [14]. F_2 offspring from this cross were experimentally challenged and genotyped, providing the data for a QTL analysis that resulted in seven QTL for susceptibility to MD[14,15]. Subsequently, the founder lines of the F2 cross were used for a microarray study to identify genes that were differentially expressed between the two lines following artificial infection. Fifteen of these genes were mapped onto the chicken genome and two of them mapped to a QTL region for Marek's resistance[16]. At the same time, protein interaction studies between a viral protein (SORF2) and a chicken splenic cDNA library revealed an interaction with the chicken growth hormone (GH)[17]. This led to the detection of a polymorphism in the GH gene that was associated with differences in the number of tumours between the

susceptible and the resistant line[17]. The GH gene coincided with a QTL for resistance and also showed up as differentially expressed between the founder lines in the expression study[16]. More recently, the same group describe detection of lymphocyte antigen 6 complex (*LY6E*) as a putative Marek's disease resistance gene, again using the virus-host protein interaction screen[18]. *LY6E* had been demonstrated earlier to be differentially expressed between resistant and susceptible chickens, but its location was not near a MD QTL[16]. Hence one could speculate that one of the MD QTL could act through *trans* acting control of the expression of this locus.

This research has demonstrated nicely how integrating across research disciplines can be very profitable. A limiting factor in the further exploitation of the QTL is the lack of precision and power to detect QTL with only 272 chickens in the F_2. The comparison of gene expression levels on the founder lines showed several potential candidate genes, but the link to the QTL regions is indirect. Scoring the gene expression levels on the F_2 would have provided a more direct link between MD QTL and eQTL and may well have flagged *LY6E* and GH as targets for eQTL. The GH effect coincided with a functional QTL pointing towards a *cis*- effect while the *LY6E* effect would appear to be *trans* regulated, and therefore only traceable to its eQTL in a genetical genomics setting.

Status of eQTL studies. To date, actual eQTL studies have been published for mice[19-21], rats[22], maize[21], yeast[23-25], eucalyptus[26] and human[27,28]. Most of these studies are 'proof of principle' or focus on the regulation of gene expression in itself.

The eQTL studies in yeast started out as a fairly straightforward proof of principle[23], which was followed up by exploring whether *trans*- regulating elements coincided with known transcription factors[24]. More recently, this work was extended to more general questions about the genetic regulation of gene expression in yeast[25] and the relevance of epistasis[29]. Two studies used the same recombinant inbred (**RI**) lines of mice to study eQTL in forebrain[20] and haematopoietic stem cells[19], respectively , could relate their findings to a whole range of phenotypes that have been measured on these mice as part of other studies. These phenotypes, as well as the expression phenotypes, have been made available online at www.genenetwork.org, providing a very valuable resource for the research community. However, the phenotypes, including the expression phenotypes, are only provided for about 33 RI lines available so far, resulting in relatively low power to detect functional QTL and eQTL. With low power to detect QTL, only the largest of QTL effects are detected and most moderate and small QTL will be missed. As a result, integration of eQTL results and functional trait QTL will only identify the largest effects.

A general conclusion from the published eQTL studies, is that the most convincing evidence for eQTL is for the *cis*-acting effects [12] while the reconstruction of genetic networks would require the identification of a larger proportion of *trans*-acting eQTL, including those with moderate effects. In short, current eQTL studies miss many important loci and fail to reconstruct genetic pathways underlying functional variation. At the same time, efforts to find the gene(s) underlying functional QTL via fine mapping and/or gene expression studies would be more effective if they were better integrated.

TOWARDS TARGETED AND INTEGRATED MAPPING

With the continuous improvements in data extraction and normalization, further increase in precision of gene expression measurements can be anticipated. Such a reduction in technical variation in gene expression measurements will increase the power to detect eQTL. Nonetheless, to improve the power and repeatability of eQTL studies it is necessary to increase their size towards those used in QTL studies of other traits.

In addition, combining larger studies with a more focussed approach further improves the power of future eQTL studies. In Figure 1 we outline 'Targeted and Integrated Mapping' of marked phenotypic QTL for functional traits (functional QTL). The central idea is to focus the studies on a relevant functional trait for which QTL have been identified previously.

Targeted and Integrated Mapping has three components (Figure 1): 1) from a large resource population, individuals that are non-recombinant for markers flanking the QTL region(s), are selected for the eQTL experiment. 2) Individuals that are recombinant for the QTL region(s) are utilised for further fine mapping of the QTL. 3) Additional expression studies are carried out for some of the recombinant individuals to confirm or evaluate positional candidate genes underlying the QTL.

Targeted and Integrated Mapping is applicable to any species for which large segregating populations are either naturally occurring or can be created experimentally, as in many livestock (including poultry), crop and experimental organisms. The approach is particularly appropriate where inbred resources, such as RI lines, are not available or cannot be realistically produced like for most poultry species. In the following sections, we outline this approach in the context of an F_2 study.

The underlying assumption of Targeted and Integrated Mapping is that QTL with major effects on the phenotype for a functional trait will often have major effects on expression of one or more genes. In some cases, the eQTL underlying a functional QTL may act in *cis* to control the expression that causes the phenotypic effect, as recently demonstrated for the *IGF*2 locus in pigs[30]. Alternatively, the phenotypic effect of a QTL and effects on expression of one or more genes may be the downstream consequence of genetic variation acting within a pathway or complex (network) of pathways. In this case we might expect to map one or more *trans*-acting eQTL to the region of the functional QTL. Compared to an unspecific genome scan for eQTL, Targeted and Integrated Mapping will have increased power to detect eQTL underlying functional QTL and to identify genetic networks and gene interactions for target QTL.

Step 1: The eQTL Study

Let us assume that we know from prior information (e.g. a QTL mapping study) that a selected genomic region affects a complex trait, usually because a functional QTL has been mapped there. Markers spaced through the putative functional QTL region (target region) are genotyped prior to phenotyping for fine-mapping or tissue collection for expression studies. This allows contrasting genotypes (e.g. alternative homozygotes in a F_2 population) for one or more functional QTL to be selected for the expression study, whilst individuals that are recombinant in the QTL regions are diverted into the fine-mapping study. Selecting individuals that are homozygous for the target regions increases the power to detect eQTL for these regions and decreases genetic complexity.

This approach improves the power to detect eQTL in three ways: 1) When selecting n homozygous individuals from an F_2, the power to detect the additive effect of an eQTL for the target regions equals that of an F_2 of size $2n$. 2) Because the contrast to estimate the putative eQTL effect is only between classes of homozygous individuals, the genetic test is simpler and uses less degrees of freedom. 3) A targeted study of one or several predetermined QTL regions involves substantially less multiple testing than does a complete genome scan, so the significance threshold for the identified regions could be less stringent than that for the remainder of the genome, increasing the power to detect eQTL even more.

The rest of the genome can also be studied for eQTL albeit with lower power than for the target regions. (For regions unlinked to selected regions the power to detect eQTL should be equivalent to that of an unselected sample of the same size, so selection is not disadvantageous for eQTL mapping in these regions). Furthermore, interactions can be studied between target regions as well as between the target regions and the remainder of the genome[31].

Step2: The Fine Mapping Study

Fine mapping strategies include those in which recombination in the QTL region is increased by targeted breeding (e.g. advanced intercross lines[6]) and those that exploit historical recombination events. Alternatively, a large pedigreed population should provide sufficient recombination to fine-map a QTL without the need for additional generations[32,33]. By typing all

individuals of the population for markers flanking the QTL, all recombinant individuals are identified. These recombinant individuals are available for further focused study to fine map the functional QTL. To decrease the genotyping load of the fine mapping, a subset from these recombinant individuals could be chosen for further study based on their phenotypic values for the functional trait in question[33]. Rather than typing the selected individuals for all available markers in the QTL region, a further decrease of the genotyping load could be obtained by applying genotyping strategies like the half-section algorithm or the golden section algorithm[33].

Step 3: Combining eQTL and Fine Mapping

Any eQTL that are identified in the target regions are potential candidates underlying the functional QTL effect. Given the increased power for the target regions, it is possible that eQTL that are detected in the target region have no direct relation with the functional QTL. The fine mapping study will reduce the confidence interval of the functional QTL, facilitating a more limited selection of positional candidate genes underlying the functional QTL effect.

For *cis*-acting eQTL, the position of the gene with the associated *cis* effect will be accurately known for species with good physical mapping or sequence data. Thus, it can be evaluated whether the gene with an associated *cis* effect still maps to the refined confidence interval of the functional QTL.

For *trans* effects that map to the candidate region, a simple comparison of location of eQTL and fine-mapped functional QTL is unlikely to be conclusive. In this case, additional expression studies using selected individuals from the fine mapping study may be required to resolve which eQTL are most likely correlated with the functional QTL and which are more likely to be linked effects. If the number of positional candidate genes is limited, such a study could evaluate a much smaller number of genes using methods like RT-PCR. The selection of genes that merit additional expression studies can be further limited by selecting those genes that give strong correlations with the phenotypic trait or have a known biological function related to the trait of interest.

Power, Precision and Population size

The successful implementation of the proposed strategy depends on the power to detect eQTL and the resolution of the fine mapping experiment. Selecting a required number of individuals that are homozygous for a functional QTL region determines the minimum size of the resource population. This in turn then determines the expected precision that can be achieved for fine mapping. Figure 2A shows the predicted statistical power to detect eQTL of with different relative effects (on gene expression) and different numbers of F_2 selected for eQTL mapping. As stated above, for the regions where all selected individuals are homozygous the increase in power is equivalent to doubling the number of individuals. For instance, with 200 F_2 the predicted power to detect an effect of 0.3 phenotypic SD is 0.34 for most of the genome, while for the target region it is 0.84 (Fig. 2A).

When selecting for homozygosity based on markers flanking the confidence interval of the functional QTL, the minimum size of the resource population should take into account the number of QTL that are targeted, the size of the interval between the flanking markers, and random fluctuations in Mendelian proportions. To obtain the number of homozygous individuals for eQTL analysis shown in Figure 2A, we have calculated the required size of the resource population when 1, 2, or 3 functional QTL are targeted with an initial confidence interval of 20 cM (Figure 2B). These population sizes give a 95% probability of yielding the stated number of individuals homozygous for one or the other gamete through each of the selected regions [34]. When focussing on a single QTL the required population size is 700 when aiming at 200 F_2 for the expression study and 1,650 when aiming at 500 F_2 for expression studies. When targeting 2 (3) functional QTL, a population of 2,200 (6,850) is required to provide 200 homozygous F_2 and 5,250 (16,400) to provide 500 homozygous F_2 (Figure 2B). Selecting individuals that are

homozygous for multiple functional QTL improves the ability to map interactions at the expression level between these QTL but the required population size becomes prohibitive for most species when three or more QTL are considered. Assuming infinite map density, the expected confidence interval of a QTL study can be predicted based on the size of the experiment and the QTL effect [6]. The predicted confidence interval for the functional QTL following the fine mapping exercise with 5,000 - 25,000 individuals in the resource population is shown in Figure 2C. For functional QTL of larger effect, sub-cM confidence intervals can be obtained when using a population exceeding 5,000 individuals (Fig 2C). Such an experiment could for instance accommodate eQTL mapping with 400 F_2 that are selected to be homozygous for two QTL flanked by a 20 cM marker bracket (Fig 2B). Based on the numbers presented in Figure 2, targeting (multiple) functional QTL with more modest effects will prove very challenging.

Even though the size of the resource population may seem prohibitive, it is important to realise that not all individuals are fully genotyped. From a resource population of size *N*, all individuals will be typed for 2 markers flanking the *m* targeted functional QTL. The *n* selected F_2 will be used for genome-wide marker analysis in the eQTL study. Linkage mapping does not require high density markers and for most genomes anywhere between 200 and 400 markers should be sufficient for a medium density linkage analysis. The amount of genotyping required for the selective recombinant genotyping depends on the selected fraction, the genotyping strategy and the size of the targeted interval [33]. Figure 2D summarizes the genotyping requirements for resource populations of 5,000 - 20,000 individuals, targeting 1, 2, or 3 functional QTL with an initial confidence interval of 20 cM using the combined golden section / half section algorithm [33] and selecting the top and bottom 25% for the trait of interest.

TOWARDS eQTL IN LIVESTOCK

Genetical Genomics requires genotypes, gene expression measure and a pedigreed population. However, to fully interpret the results, we need to know the location of the genes as well as their function.

Among livestock species, chickens are very well placed to be used in full blown genetical genomics studies. There is a large number of chicken QTL regions in the public domain [4] and the species has the benefit of a full genome sequence [35] and a SNP database [36]. In terms of the gene expression tools, there are a number of tissue specific as well as general two-colour arrays (both spotted cDNA and long oligonucleotide array; http://www.ark-genomics.org/resources/chickens.php) as well as an Affymetrix™ chicken genome array (http://www.affymetrix.com/products/arrays/specific/chicken.affx). Large resource populations of chicken can be bred in a timely fashion or obtained from commercial lines. Populations for fine mapping, like advanced intercross lines (**AIL)** [6], are available in several labs (e.g. Wageningen University, Netherlands; Iowa State, USA and Roslin Institute, UK). Microarrays are also available for other livestock and while the draft sequence of catlle has been released and the pland for sequenceing the pig genome are advancing, the required annotation is still some way of. An area for further development in the immediate future is the ongoing annotation of the genome and other bioinformatics tools like pathway databases that incorporate livestock specific information rather than pathways that are derived from model organisms or humans. However, the most limiting factor in the uptake of genetical genomics in livestock species is the budget required to run microarray studies on large numbers of animals. The recently proposed design of distant pairing[37] for genetical genomics looks promising in that it offers the possibility to array 2*n* individuals using *n* microarrays. In contrast, to reference designs or one-colour arrays, this design is based on the contrast in gene expression between individuals that have been selected a prior on their divergent genotypes. However, this method has been implemented only for RI lines and its efficiency for outbred populations has not been quantified.

Limited resources

With limited resources and a more focused objective, the principle of targeted eQTL mapping can still be applied. In the context of an F_2 study or similar, increasing the power from a smaller study can be used if the main focus is the identification of *cis* and *trans* acting eQTLs that underlie the QTL peak and a whole genome eQTL scan is not of interest. The increased power from selection of homozygous individuals and the less stringent significance threshold required in a focused study, as opposed to a genome scan, require a more modest sized resource population and correspondingly fewer individuals to be assayed for gene expression. For an experiment with 200 F_2 that are homozygous for the selected region(s), the power to detect eQTL is > 95% for any effect larger than 0.3 phenotypic SD (using a less stringent threshold of $P < 0.01$.) Recombinant individuals can be used to increase the mapping accuracy of the QTL, but the improved resolution will be more modest and correspond to the smaller overall size of the resource population. The smaller sized study may mean that a genome scan is less worthwhile (although if genotyping costs are modest, a genome scan for the largest effect eQTL can be undertaken with little additional input as the expression data are already recorded).

Using Genetical genomics for a marked QTL

To illustrate potential of genetical genomics we describe a pilot study in chickens [38]. The crucial part is the focused study of a particular putative QTL, in this case one affecting body weight segregating in an inter-cross of broilers and layers. Our objective was to identify candidate genes through the effect of the QTL at the gene expression level: what genes are affected, where do they map and in what kind of pathways are they involved?

We identified individuals that were homozygous for markers flanking a QTL region on chromosome 4 (GGA4) from the seventh generation of an advanced inter-cross between a single broiler and a single layer chicken. These were inferred to be either QQ (broiler allele) or qq (layer allele) for the QTL and matings were set up to provide birds with 'known' QTL genotypes. From the resulting offspring, QQ males and qq males were slaughtered at 21 days of age and a sample of the breast muscle was taken for RNA isolation and microarray studies. The microarrays design was a direct comparison of QQ versus qq for eight independent samples with a dye-swap (16 arrays used in total). The microarray was a chicken cDNA array with 12,877 functional features, spotted in duplicate (Ark-Genomics©, 2004). Using five alternative normalization procedures, we defined a consensus set of results consisting of 45 (895) differentially expressed genes when applying a false discovery rate (**FDR**) of 5% (20%) [38]. This implies that out of 45 (895) results we expect less than 3 (180) false positive results. The genes that are differentially expressed seem evenly distributed over the genome and there appears to be no enrichment for affected genes in the QTL area on GGA4. However, there are 12 differentially expressed genes (FDR < 20%) that map to the QTL region and should be considered positional candidate genes for the QTL. Among these, AADAT (FDR < 5%) is involved in lysine degradation, lysine biosynthesis and tryptophan metabolism, making it a promising candidate gene. At present, we are performing pathway analyses to see what pathways are enriched for differentially expressed genes and thus providing further clues on the way in which the QTL affects body mass. Further annotation of the microarray and dedicated pathway databases for chicken will further improve the characterization of this QTL. This demonstrates how a focussed study can aid the dissection of a QTL using limited resources.

CONCLUDING REMARKS

Although the existing eQTL studies demonstrate the utility of genetical genomics, they do not show its full potential because they miss many moderate effects and provide little opportunity to unravel genetic pathways due to a lack of *trans*-acting effects that would provide tangible links between eQTL and genes. Targeted and Integrated Mapping is applicable to any species for which populations with a few thousand or more pedigreed individuals can be accessed and has

distinct advantages over an untargeted genome scan for eQTL. If the targeted eQTL study identifies *cis*-acting eQTL underlying the functional QTL, this provides a direct route to the candidate loci controlling the functional QTL [13,16]. Targeted and Integrated Mapping is specifically aimed at unravelling genetic pathways underlying a functional QTL; by contrast, non-targeted studies of similar size would identify eQTL relating to many pathways, but with too few interconnected QTL to reconstruct a pathway. For example, the studies on BXD mice consider a very wide range of phenotypes and gene expression measures, but limited statistical power reduces the number of meaningful inference that can be drawn [19,20]. With a Targeted and Integrated Mapping approach the fine mapping will reduce the size of the region containing the functional QTL, which in turn can be used to re-evaluate the eQTL that map to the functional QTL, further refining the list of potential candidate genes and the possible gene networks underlying the functional QTL.

While the utility of inbred resources like RI lines for fine mapping and (e)QTL mapping has been demonstrated elsewhere[19,20,22], we want to emphasize that genetical genomics should not be restricted to model species and we make the case that poultry is very well placed among livestock species to pioneer these approaches.

Acknowledgements. The authors acknowledge financial support from the BBSRC.

Reference List

1 Doerge,R.W. (2002) Mapping and analysis of quantitative trait loci in experimental populations. *Nat. Rev. Genet.* 3, 43-52

2 Andersson,L. and Georges,M. (2004) Domestic-animal genomics: deciphering the genetics of complex traits. *Nat. Rev. Genet.* 5, 202-212

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

4 Hocking,P.M. (2005) Review of QTL results in chicken. *World's Poultry Science Journal* 61, 215-226

5 Flint,J. and Mott,R. (2001) Finding the molecular basis of quantitative traits: successes and pitfalls. *Nat. Rev. Genet.* 2, 437-445

6 Darvasi,A. (1998) Experimental strategies for the genetic dissection of complex traits in animal models. *Nat. Genet.* 18, 19-24

7 Cardon,L.R. and Bell,J.I. (2001) Association study designs for complex diseases. *Nat. Rev. Genet.* 2, 91-99

8 Jansen,R.C. and Nap,J.P. (2001) Genetical genomics: the added value from segregation. *Trends Genet.* 17, 388-391

9 Jansen,R.C. (2003) Studying complex biological systems using multifactorial perturbation. *Nat. Rev. Genet.* 4, 145-151

10 Hitzemann,R. *et al.* (2003) A strategy for the integration of QTL, gene expression, and sequence analyses. *Mamm. Genome* 14, 733-747

11 Carlborg,O. *et al.* (2004) Simultaneous mapping of epistatic QTL in chickens reveals clusters of QTL pairs with similar genetic effects on growth. *Genet. Res.* 83, 197-209

12 de Koning,D.J. and Haley,C.S. (2005) Genetical genomics in humans and model organisms. *Trends Genet.* 21, 377-381

13 Wayne,M.L. and McIntyre,L.M. (2002) Combining mapping and arraying: An approach to candidate gene identification. *Proc. Natl. Acad. Sci. U. S. A* 99, 14903-14906

14 Vallejo,R.L. *et al.* (1998) Genetic mapping of quantitative trait loci affecting susceptibility to Marek's disease virus induced tumors in F2 intercross chickens. *Genetics* 148, 349-360

15 Yonash,N. *et al.* (1999) High resolution mapping and identification of new quantitative trait loci (QTL) affecting susceptibility to Marek's disease. *Anim Genet.* 30, 126-135

16 Liu,H.C. *et al.* (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

17 Liu,H.C. *et al.* (2001) Growth hormone interacts with the Marek's disease virus SORF2 protein and is associated with disease resistance in chicken. *Proc. Natl. Acad. Sci. U. S. A* 98, 9203-9208

18 Liu,H.C. *et al.* (2003) Identification of chicken lymphocyte antigen 6 complex, locus E (LY6E, alias SCA2) as a putative Marek's disease resistance gene via a virus-host protein interaction screen. *Cytogenet. Genome Res.* 102, 304-308

19 Bystrykh,L. *et al.* (2005) Uncovering regulatory pathways that affect hematopoietic stem cell function using 'genetical genomics'. *Nat. Genet.* 37, 225-232

20 Chesler,E.J. *et al.* (2005) Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. *Nat. Genet.* 37, 233-242

21 Schadt,E.E. *et al.* (2003) Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422, 297-302

22 Hubner,N. *et al.* (2005) Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nat. Genet.* 37, 243-253

23 Brem,R.B. *et al.* (2002) Genetic dissection of transcriptional regulation in budding yeast. *Science* 296, 752-755

24 Yvert,G. *et al.* (2003) Trans-acting regulatory variation in Saccharomyces cerevisiae and the role of transcription factors. *Nat. Genet.* 35, 57-64

25 Brem,R.B. and Kruglyak,L. (2005) The landscape of genetic complexity across 5,700 gene expression traits in yeast. *Proc. Natl. Acad. Sci. U. S. A* 102, 1572-1577

26 Kirst,M. *et al.* (2005) Genetic architecture of transcript-level variation in differentiating xylem of a eucalyptus hybrid. *Genetics* 169, 2295-2303

27 Monks,S.A. *et al.* (2004) Genetic inheritance of gene expression in human cell lines. *Am. J. Hum. Genet.* 75, 1094-1105

28 Morley,M. *et al.* (2004) Genetic analysis of genome-wide variation in human gene expression. *Nature*

29 Brem,R.B. *et al.* (2005) Genetic interactions between polymorphisms that affect gene expression in yeast. *Nature* 436, 701-703

30 Van Laere,A.S. *et al.* (2003) A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* 425, 832-836

31 Carlborg,O. and Haley,C.S. (2004) Epistasis: too often neglected in complex trait studies? *Nat. Rev. Genet.* 5, 618-625

32 Thaller,G. and Hoeschele,I. (2000) Fine-mapping of quantitative trait loci in half-sib families using current recombinations. *Genet. Res.* 76, 87-104

33 Ronin,Y. *et al.* (2003) High-resolution mapping of quantitative trait loci by selective recombinant genotyping. *Genetics* 164, 1657-1666

34 Jansen, R. C. On the selection for specific genes in doubled haploids. Heredity 69, 92-95. 1992.
Ref Type: Generic

35 Hillier,L.W. *et al.* (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432, 695-716

36 Wong,G.K. *et al.* (2004) A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* 432, 717-722

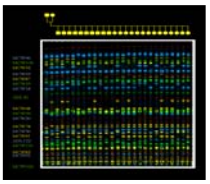
37 Fu,J. and Jansen,R.C. (2006) Optimal design and analysis of genetic studies on gene expression. *Genetics* 172, 1993-1999

38 Cabrera, C. P., Dunn, I. C., Fell, M., Wilson, P. W., Burt, D. W., Waddington, D., Talbot, R. T., Hocking, P. M., Law, A., Haley, C. S., Knott, S. A., and de Koning, D. J. **Application of genetical genomics to a marked QTL in poultry**. Proceeding

of 8th WCGALP in press. 2006. Brazil. Ref Type: Conference Proceeding

39 Lynch M. and Walsh J.B. (1998) *Genetics and analysis of complex traits*, Sinauer associates, Inc.

(a) Large segregating population

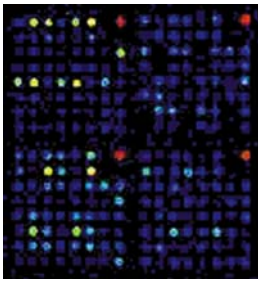


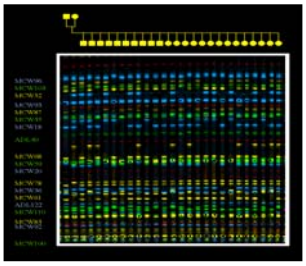
(b) Molecular typing of population for known QTL

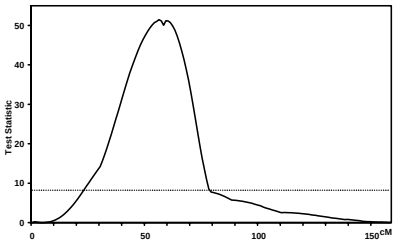
Experimental challenge

(c) Individuals that are homozygous for target region(s)

Tissue sampling, gene expression analysis and genome-wide genotyping



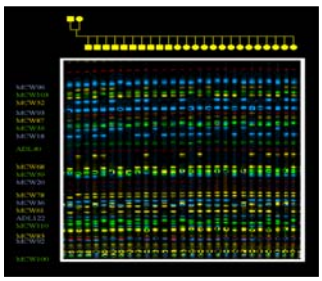


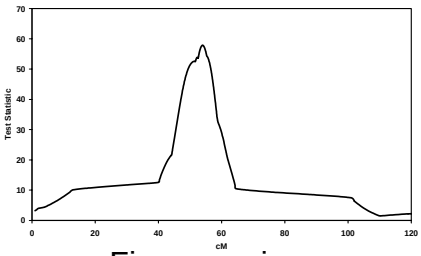


eQTL mapping

(d) Individuals that are recombinant for target region(s)

Additional phenotyping and (selective) recombinant genotyping





Fine mapping

(e) Positional candidate genes in target regions, effects of QTL on gene expression in cis and trans

eQTL fine mapping

Figure 1. Targeted and Integrated Mapping. The design requires a resource population of a few thousand individuals or more (a) that is segregating for phenotypic traits of interest and in which QTL affecting this trait have been discovered or confirmed. (b) The entire population is genotyped for markers flanking the previously identified QTL. Individuals that are homozygous for the region(s) of interest (c) will be used for tissue collection and gene expression analyses, possibly following an experimental challenge. They will also be genotyped for markers spanning the entire genome. This will identify whether a marked QTL region: affects expression of genes in the same area of the QTL (in *cis*), affects the expression of genes elsewhere in the genome (in *trans*) or interacts with each other marked QTL regions or other regions of the genome. Individuals that are recombinant for the QTL region(s) (d) can be further phenotyped for the trait of interest. Combined with a genotyping strategy that is aimed at identifying all recombinants in the QTL region(s) the QTL region can be narrowed down. If experimentally feasible, a sub-set of the individuals that were used for fine mapping could be used for limited gene expression analysis in order to validate the eQTL results via eQTL fine mapping.

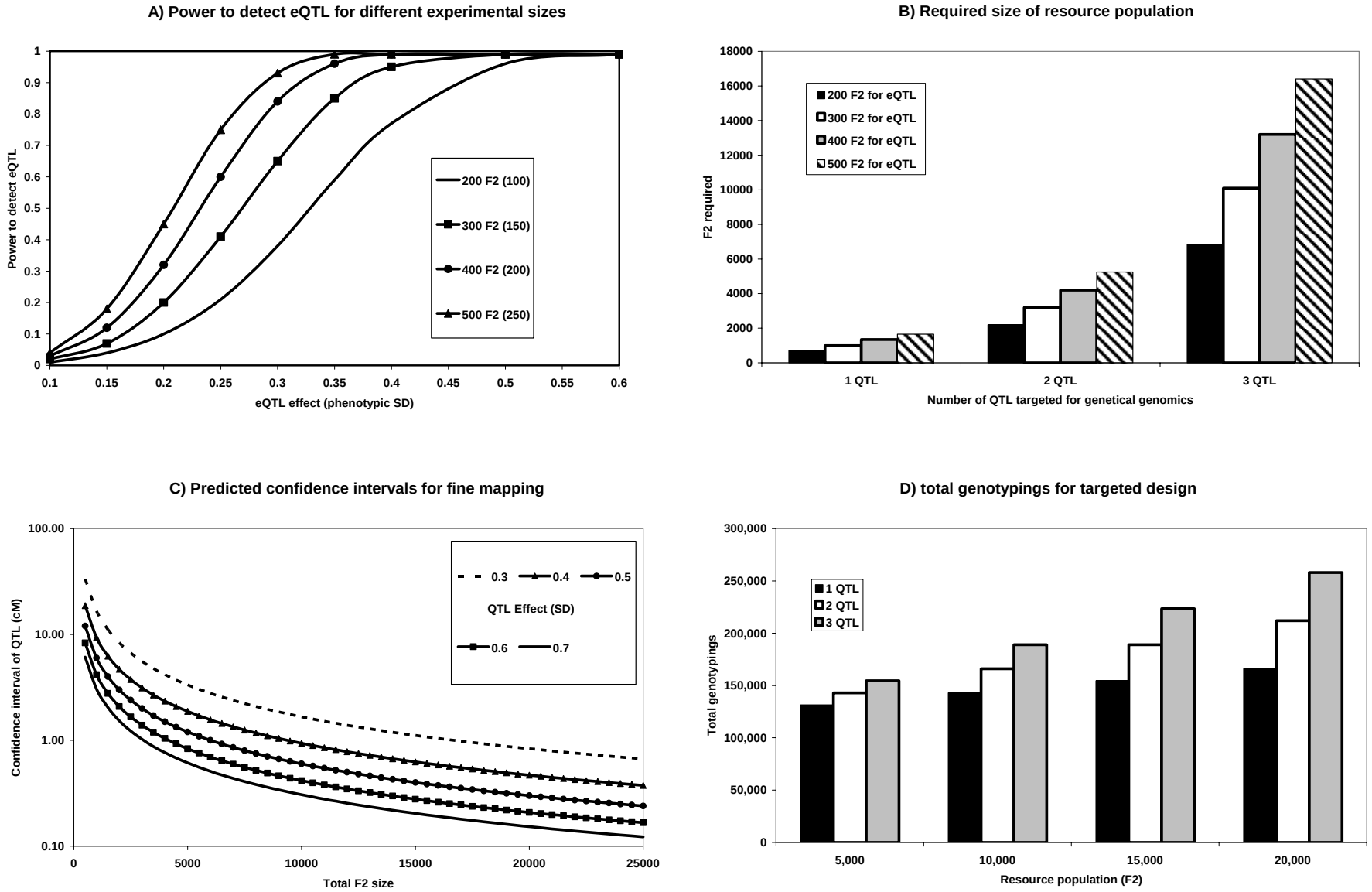


Figure 2 Statistical power and precision for Targeted and Integrated Mapping. A) Power to detect eQTL at $P < 0.001$ (LOD = 3.0) for different eQTL effects and F_2 population sizes[39]. Between brackets is the equivalent number of selected F_2 that are homozygous for the QTL or RI lines for the same statistical power. B) The size of the resource population that is required to obtain a given number of F_2 individuals that are homozygous (with the same line origin) for 1, 2, or 3 functional QTL with a confidence interval of 20 cM taking into account potential deviations from Mendelian ratios. The numbers are based on a 95% probability to have the required number of animals homozygous for the QTL[34]. C) The expected resolution (confidence interval) from fine mapping given the size of the resource population and the original complex trait QTL effect [6]. D) The total number of genotyping experiments for the combined strategy targeting 1, 2, or 3 QTL with an initial confidence interval of 20 cM using a golden section / half section selective genotyping strategy on the 25% top and tails of the resource population[33]. The genotyping for the eQTL, assuming a genome scan, is fixed at 120,000 (i.e. 200 individuals for 600 markers, 300 for 400 markers etc.). The concept is illustrated for an F_2 experimental cross, based on the outline that is presented in Figure 1