"We are searching! What are we finding?"

R.J. Kerr[1] Y. Chen[2], J.M. Henshall[1], B.G. Luxford[3] & C. Moran[2]

[1]Animal Genetics and Breeding Unit, joint Institute of NSW Agriculture and The University of New England, Armidale, NSW, Australia; [2]Department of Animal Science, University of Sydney, NSW, Australia; [3]Bunge Meat Industries Ltd., Corowa, NSW, Australia

Introduction

A locus is a position on a chromosome occupied by genes that determine the state of a particular phenotypic character. When that phenotypic character is something that is measurable, such as mm of fat, or kg of body weight, rather than something that is qualitative, such as eye or coat colour, the locus will be referred to as a quantitative trait locus (QTL). The various occupants of the locus are referred to as alleles. It is intuitive to think that there will be more than one locus involved in the expression of a quantitative character. In fact there will probably be hundreds of loci that are involved. However, it is also reasonable to expect that the occupants or alleles at some loci have more affect on the quantitative trait, than alleles at other loci. Let us call these QTL "leading QTL".

It is now possible to find the positions of these "leading QTL" by QTL mapping. The geneticist, taking each chromosome in turn, peppers the chromosome with a battery of anonymous DNA markers, and pinpoints the leading QTL between two markers. Traditionally, crosses between divergent breeds have been used as the resource population in QTL mapping. This is because in the F_1 progeny the alleles at a QTL are more likely to have contrasting effects. The first allele derives from one breed, the second from the partner breed. The effects of the two alleles are likely to be more pronounced than if the two alleles derived from the one breed. The downside to this is that the allele, that has the more favourable effect on the trait is usually from the breed in which the frequency is already high. For example, Swedish researchers have used a European Wild Boar x Large White cross (Andersson et al. 1994). It is generally found that the "Large White allele" is the favourable allele at any detected QTL, and is most likely fixed in commercial populations.

Here in Australia a PRDC funded project was established to undertake QTL mapping in commercial lines of pigs. This presentation reports on the results of this project.

Materials and methods

Phenotypic data for the study were recorded at Bunge Meat Industries from June 1995 to November 1995. A half-sib design was used in which two Large White boars and two Landrace boars were each mated to a random selection of dams to produce on average 100 progeny. Why are so many progeny required? Think of the 100 progeny as a sample of 100 of the sire's gametes. Each gamete contains a set of paternal chromosomes. The chromosomes in the gamete are produced by meiosis. Each chromosome produced during meiosis will have at least one cross over event. These cross over events are nature's way of ensuring the parental chromosome transmitted to an offspring contains a mixture of grand-paternal and grand-maternal genes. Having 100 gametes per sire in a QTL mapping experiment means there is likely to be a good sample of cross over events at different positions in each of the chromosomes. Correlating changes in phenotypic score with differences in cross over positions on a particular chromosome is in essence the basis of QTL mapping.

The testing procedure started with the recording of animal weight at 21 days which was used to derive average daily gain to 21 days (ADG21). At 18 weeks animals entered the boar test station where they were single penned and fed *ad libitum*. Weight of the animal was recorded at the beginning of the testing period and shortly before slaughter at 22 weeks. The information recorded in the boar test station was used to obtain the following growth and feed efficiency traits: average daily gain from 3 to 18 weeks (ADG1); average daily gain from 18 to 22 weeks (ADG2); lifetime average daily gain (ADG3); daily feed intake (DFDINT); feed conversion ratio (FCR). Carcase characteristics were measured on the live animal as well as in the abattoir and boning room and included: backfat depth at P2 measured with real time ultrasound (FDP2); backfat depth between 3rd/4th last ribs measured with real time ultrasound (FD3/4); muscle depth between 3rd/4th last ribs measured with real time ultrasound (MD3/4); backfat depth at P2 measured with Hennesy Chong machine (HCP2); weight of left back leg (LW); weight of slash boned ham (HAM). Meat quality characteristics measured on the slaughter day and the day after slaughter included the following traits: colour of *m. longissimus dorsi* (L-value) (CLD); colour of *m. superior spiralis* (L-value) (CSP); pH measured 45 minutes and 24 hours after slaughter (PH45 and PH24); drip loss percentage (DLP); intramuscular fat content (IMF). In total 18 traits were measured. See Hermesch (1996) for a detailed description of traits.

Prior to QTL mapping, data were edited to remove outliers and analysed to determine significant environmental effects, using PROC GLM (SAS 1991). The significant fixed effects were then included in a mixed animal model using a pedigree with all known ancestral information. Estimates of the additive genetic and residual variance were taken from a recent variance component estimation experiment performed on another Bunge resource population. Tests for the influence of other random effects such as litter effects and maternal genetic effects were made and these effects were not significant. The overall mean was added to the appropriate individual genetic and residual effects to obtain the adjusted data value, which was used in the QTL mapping.

Statistical issues of QTL mapping

Statistics has two key disciplines. There is the science of gathering and analysing data. There is also the science of drawing statistical conclusions from the data using knowledge of probability. The statistics of QTL mapping are no different. It is not appropriate here to explain in detail the data analysis part of QTL mapping, as it is quite complicated. It is more important for the non-practitioner of QTL mapping to understand how the practitioner draws his conclusions.

Hypothesis testing is all about answering the question, "could these observations really have occurred by chance?" There are four formal steps in hypothesis testing.

1. *Formulate all hypotheses*. The null hypothesis is usually that the observations are the result purely of chance. In QTL mapping this translates to saying any QTL detected is the result of chance alone. The alternate hypothesis is that there is a real effect, that is, there is a real QTL and the observations are partly the result of this QTL, and partly the result of chance variation

2. *Getting the test statistic.* Identify a test statistic that will assess the evidence against the null hypothesis. In QTL mapping the test statistic is usually a likelihood ratio test (LRT). It is essentially a ratio of two probabilities. The probability of the observations occurring given a model that includes a QTL plus random error is compared to the probability of the observations occurring given a model that includes only random error.

3. *Getting the p-value*. Making a probability statement which answers the question: if the null hypothesis is true, then what is the probability of observing a test statistic at least as extreme as the one we observed. The smaller the *p-value*, the stronger the evidence against the null hypothesis.

4. *Comparing the p-value to a fixed significance level,* α. The α acts as a cut off point below which we agree that an effect is statistically significant. That is, if *p-value* $\leq \alpha$ we rule out the null hypothesis and agree that a QTL does exist

In QTL mapping the first two steps are straightforward. Figure 1 shows the results of a scan of a single chromosome. The LRT that is calculated at each analysis point is plotted graphically to create a profile. There are 4 plots, one for each of three traits analysed, and one for a joint mapping of all three traits simultaneously. The highest point in a profile indicates the likely position of a putative QTL. If the *p-value* of the LRT at that highest point is below a fixed significance level, then we can claim the QTL does exist.

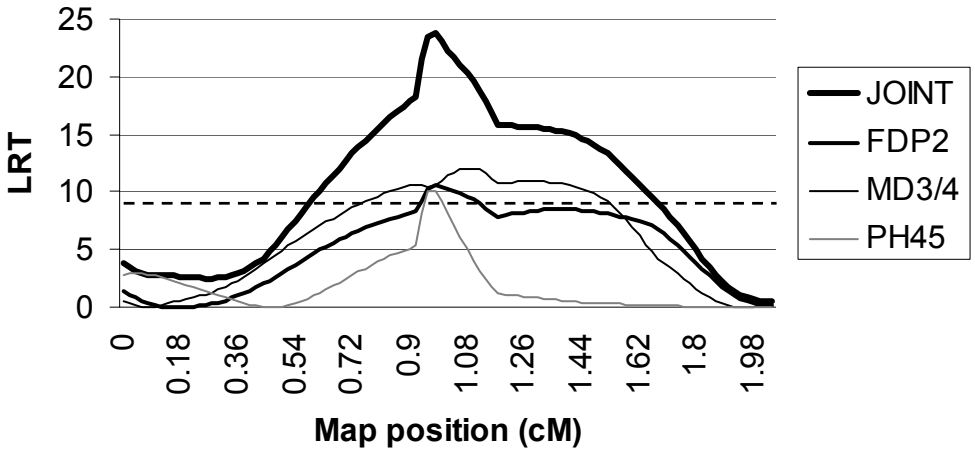

Figure 1. Likelihood ratio test (LRT) statistic against chromosome position, calculated from single trait analyses of FDP2, MD3/4 and PH45, and from a multiple trait analysis (JOINT).

Steps 3 and 4 are the ones that cause a lot of controversy and confusion. It is all to do with how high should you set the "bar", depicted as the dashed horizontal line in Figure 1. For most cases in statistical hypothesis testing a known distribution can be used to specify the probability of observing a test statistic. For example, the binomial distribution can be used to specify the probability of getting 4 tails in 80 coin flips. The F-distribution can be used to specify the probability of observing a particular ratio of mean square errors. Normally a chi-squared distribution is used to specify the probability of observing a particular LRT. However, in QTL mapping there are various reasons given by pure statisticians why this is not strictly correct. There are two courses of action: (i) ignore the pure statisticians and proceed to use the chi-squared distribution; or (ii) determine empirically the distribution of your test statistic, assuming the null hypothesis is true.

Course of action (ii) is actually quite a simple procedure. It is done using a technique known as permutation testing. Consider a table with 2 columns and 100 rows representing 100 progeny. Column 1 is the progeny identification and column 2 is the progeny observation. Suppose we were to jumble the first column only and then say the observation in column 2 really did belong to progeny now named in column 1. This action implies that we don't believe a QTL exists. However a LRT is calculated all the same. If we were to repeat this perhaps 1000 times, we would have 1000 test statistics computed, assuming the null hypothesis is true. We could then compute the relative frequency of the observed test statistic, that is, the number of times the actual LRT occurred divided by the number of permutations, and this would be our *p-value* based on an empiric distribution.

However the controversy doesn't end there. The permutation test allowed us to get a *p-value* computed at *one* analysis point. In scientific work, a fixed α level of .05 or .01 is often used. A fixed α level of .05 can be interpreted as saying 1 time out of 20, results with a significance level of *p-value* $\leq$ .05, are actually false. QTL mapping involves testing at multiple locations across each chromosome considered. There are multiple traits and multiple pedigrees to consider as well. Suppose in total 500 tests were completed and there were twenty occurrences when the null hypothesis was

rejected. However, it is expected that 25 significant results would have been found by chance alone. So how do we then view our significant results?

The traditional approach to dealing with multiple tests has been to consider an 'experiment-wise' *p-value* as opposed to the 'nominal' or 'point-wise' *p-value.* How to actually determine the experiment-wise α level is the subject of many debates. A well-known American geneticist has remarked that the setting of significance thresholds is akin to 'charting a course between Scylla and Charybdis.' A lax approach will bring on a flood of false positives and erode the credibility of genetic research, while an overcautious approach will cause hints of true QTL to be missed. The results of the US36 mapping project illustrate this dilemma.

Results

Table 1 displays various statistics such as numbers of chromosomes analyzed and numbers of markers etc. The last row of the table states the approximate, total number of tests that were performed in the project. This number was calculated by determining the approximate number of intervals analyzed in each family and totaled over the number of families used in the project and the number of traits analyzed.

Table 1 Various statistics of project

Number of markers genotyped	100
Number of chromosomes tested	18
Number of chromosomes with 3 or more markers	13
Average number of markers per chromosome	7
Average spacing of markers in chomosomes with more than 3 markers	20cM
Number of traits analysed	18
Number of pedigrees	4
Approximate total number of tests performed	3218

The QTL detected in the project are displayed in Table 2. The 2nd column lists what trait the QTL affects, the 3rd and 4th columns state in which sire family and on what chromosome the QTL were detected, respectively. The 5th and 6th columns show the value of the LRT statistic and its nominal *p-value*. The nominal *p-value* is the probability of observing a LRT at least as extreme as the one observed, *at that particular analysis point*. The next column is a value for the experiment-wise *p-value*, which is the probability of observing a LRT at least as extreme as the one observed, *in the whole experiment* (3218 tests). If we were to base our criterion for deeming what is a true QTL on the nominal *p-value* and set α at .05 we would claim that every QTL in the table are true QTL. If we were to base our criterion on the experiment-wise *p-value,* we would claim that there is only one QTL that can be deemed as representing a true effect and worth investigating. The question remains - where to chart a prudent course between Scylla and Charybdis?

Table 2. Detected QTL, ranked in order of their LRT value

QTL[a]	Trait	Family	Chr[b]	LRT	Np-value[c]	Ep-value[d]	EFDR[e]
1	HCP2	LWA	B	19.733	0.00001	0.028	0.029
2	FD3/4	LWA	B	16.284	0.00005	0.161	0.088
3	FDP2	LWB	E	13.851	0.00020	0.471	0.212
4	HCP2	LRA	I	13.578	0.00023	0.521	0.184
5	DFDINT	LRB	D	12.760	0.00035	0.680	0.228
6	PH24	LWA	E	12.290	0.00046	0.769	0.244
7	HCP2	LWA	G	12.283	0.00046	0.770	0.210
8	MD3/4	LWB	E	12.097	0.00051	0.803	0.203
9	HAM	LRA	B	11.873	0.00057	0.840	0.204
10	HCP2	LRA	F	11.861	0.00057	0.842	0.184
11	ADG21	LWA	B	11.336	0.00076	0.913	0.222
12	LW	LRB	F	11.320	0.00077	0.915	0.206
13	ADG3	LRB	K	11.251	0.00080	0.923	0.197
14	PH45	LWB	I	11.189	0.00082	0.929	0.189
15	ADG3	LRB	H	10.973	0.00092	0.949	0.198
16	ADG3	LRB	G	10.909	0.00096	0.954	0.193
17	ADG21	LWB	K	10.795	0.00102	0.962	0.193
18	ADG3	LRB	G	10.690	0.00108	0.969	0.193
19	LW	LWA	B	10.609	0.00113	0.973	0.191
20	ADG3	LRB	F	10.448	0.00123	0.981	0.198
21	FDP2	LWA	B	10.411	0.00125	0.982	0.192
22	HAM	LWA	B	10.396	0.00126	0.983	0.185
23	ADG3	LWA	B	10.364	0.00129	0.984	0.180
24	FDP2	LRA	B	10.264	0.00136	0.987	0.182
25	PH45	LWB	E	10.121	0.00147	0.991	0.189
26	ADG1	LRB	H	10.082	0.00150	0.992	0.185
27	LW	LRB	G	9.828	0.00172	0.996	0.205
28	ADG3	LRB	G	9.812	0.00173	0.996	0.199
29	ADG21	LRA	A	9.706	0.00184	0.997	0.204
30	DLP	LWB	C	9.573	0.00197	0.998	0.212
31	ADG3	LRB	G	9.331	0.00225	0.999	0.234
32	CLD	LRB	E	9.289	0.00231	0.999	0.232
33	ADG2	LRB	D	9.274	0.00232	0.999	0.227
34	LW	LRB	K	9.028	0.00266	1.000	0.252
35	DLP	LWB	E	9.002	0.0027	1.000	0.248
36	HCP2	LRA	K	8.728	0.00313	1.000	0.280
37	MD3/4	LRA	A	8.718	0.00315	1.000	0.274
38	BLW	LWA	A	8.566	0.00343	1.000	0.290
39	MD3/4	LRA	K	8.309	0.00395	1.000	0.326
40	ADG2	LRA	G	8.204	0.00418	1.000	0.336

[a] QTL ranked in order of value for LRT
[b] Chr = Chromosome
[c] p-value N = nominal p-value
[d] p-value E = experiment-wise p-value
[e] EFDR = expected false discovery rate

The application of a false discovery rate (FDR) has been recently advocated as an alternative to controlling the experiment-wise error rate (Weller et al. 1999). The FDR is defined as the expected number of true null hypotheses within the class of rejected null hypotheses. In other words, of the null hypotheses so far rejected, what proportion were wrongly rejected. The last column in Table 2 displays this expected proportion. The experimenter has to decide on some level. For example, let us decide that 25% is a tolerable false discovery rate. A dotted line is drawn after the

QTL listed as number 35. The expected FDR exceeds .25 beyond this point. Thus we can conclude that of the 35 QTL deemed as representing true QTL, we can expect that 8 or 9 QTL won't be detected in another experiment, because they in fact do not represent true QTL. The FDR seems to offer a good compromise between the too stringent experiment-wise error rate and the too lax nominal error rate. This is perhaps the right course between Scylla and Charybdis.

It is unrealistic to expect that definitive results can be obtained from this single experiment. Confirmation that the QTL are indeed true effects is best provided by statistical significance in at least two independent studies. Thus, this experiment should be viewed as a means to determine which chromosomal regions should be re-sampled. A new PRDC funded project has already begun that will examine selected chromosomal regions in other pedigrees. The new project will also evaluate the feasibility and economic benefit of marker assisted selection (MAS). Using data on QTL positions and effects, obtained from both projects, as well as relevant estimates of genetic and phenotypic parameters from Australian pigs, the benefits of MAS will be weighed against the costs of genotyping. The number of QTL so far detected from this project and the magnitude of their effects gives hope that MAS will make a useful contribution in practical breeding programs in Australia.

An aspect of Figure 1 not yet discussed is the plot labeled 'JOINT' in the legend. This is the plot of the LRT against chromosome position, calculated from the joint mapping of the three traits in one analysis. As well as being generally more powerful than single trait mappings, a joint mapping analysis affords the possibility of answering biologically interesting questions such as, is the QTL pleiotropic, or are there in fact two QTL in the same region, each affecting different traits. Statistical tests suggested the QTL indicated in Figure 2 does have a pleiotropic effect on the traits pH45, FDP2, MD3/4. Table 3 shows the effect of this QTL on each of the traits. This illustrates that QTL mapping will help us resolve at the genetic level the basis of the genetic correlations we routinely predict.

Table 2. Estimated effects of a putative QTL (±SE) with pleiotropic effects on 3 traits indicated by joint mapping

Boar	Chromosome	Traits	Effect of allele substitution (δ)[1]
LWB	E	PH45	0.25 (±..09)
		FDP2	1.65 mm (±..49)
		MD3/4	-3.56 mm (±.1.25)

[1]Effect in phenotypic standard deviation units

Summary

A QTL is a position on a chromosome. The occupants or alleles at this position have a significant effect on a quantitative trait. This article has discussed the results of a collaborative QTL mapping project involving the Dept. Animal Science at the

University of Sydney, AGBU and Bunge Meat Industries and funded by the PRDC. At a nominal level many QTL appear significant. An important issue for the scientist is the setting of a sufficiently stringent threshold above which QTL can be claimed as being 'true' QTL, and not the result of chance. This QTL mapping project has identified at least 30 QTL that are considered as representing true effects and are currently being re-sampled in further pedigrees.

References

Andersson, L., Haley, C.S., Ellegren, H., Knott, S.A., Johansson, K., Andersson, K., Andersson-Eklund, L., Edfors-Lilja, I, Fredholm,M., Hansson, I., Hakansson, J. and Lundstrom, K. Science 263: 1771.

Hermesch, S. (1996) PhD Thesis, University of New England, Armidale.

SAS (1991) SAS/STAT User's Guide. Release 6.03 edition, SAS Institute Inc.

Weller, J.I., Song, J.Z., Heyen, D.W., Lewin, H.A. and Ron, M. (1998) A new approach to the problem of multiple comparisons in the genetic dissection of complex traits. Genetics 150: 1699.