

# Quantitative trait loci mapping in dairy cattle: review and meta-analysis

Mehar S. KHATKAR, Peter C. THOMSON\*, Imke TAMMEN,  
Herman W. RAADSMA

Centre for Advanced Technologies in Animal Genetics and Reproduction (ReproGen), and  
Co-operative Research Centre for Innovative Dairy Products, Faculty of Veterinary Science,  
University of Sydney, PMB 3, Camden NSW 2570, Australia

(Received 4 July 2003; accepted 29 October 2003)

**Abstract** – From an extensive review of public domain information on dairy cattle quantitative trait loci (QTL), we have prepared a draft online QTL map for dairy production traits. Most publications (45 out of 55 reviewed) reported QTL for the major milk production traits (milk, fat and protein yield, and fat and protein concentration (%)) and somatic cell score. Relatively few QTL studies have been reported for more complex traits such as mastitis, fertility and health. The collated QTL map shows some chromosomal regions with a high density of QTL, as well as a substantial number of QTL at single chromosomal locations. To extract the most information from these published records, a meta-analysis was conducted to obtain consensus on QTL location and allelic substitution effect of these QTL. This required modification and development of statistical methodologies. The meta-analysis indicated a number of consensus regions, the most striking being two distinct regions affecting milk yield on chromosome 6 at 49 cM and 87 cM explaining 4.2 and 3.6 percent of the genetic variance of milk yield, respectively. The first of these regions (near marker BM143) affects five separate milk production traits (protein yield, protein percent, fat yield, fat percent, as well as milk yield).

**quantitative trait loci / dairy cattle / review / meta-analysis**

## 1. INTRODUCTION

A major objective of quantitative trait loci (QTL) studies is to find genes/markers that can be implemented in breeding programs *via* marker assisted selection (MAS). There is general agreement from theoretical and simulation studies that application of MAS has the potential to increase the rate of genetic gain especially if traditional selection is less effective [1, 59]. In dairy cattle MAS could be used to pre-select young candidate bulls prior to progeny testing, thus increasing selection differentials, shortening generation

---

\* Corresponding author: PeterT@camden.usyd.edu.au

interval and increasing genetic gain [45]. Once a QTL is identified, it is necessary to identify families in the breeding population which are segregating for that QTL. However, if a QTL has been fine mapped with respect to closely linked markers that are in linkage disequilibrium (LD) with the QTL, the associations between specific marker haplotypes and QTL alleles should hold across populations and need not be re-established for each individual family. Selection for such QTL can be undertaken throughout the population rather than only in the specific families, thereby greatly simplifying the implementation of MAS. Identification of genes underlying QTL can provide not only the most accurate markers for MAS, but also identifies critical biochemical pathways for further investigation and endogenous or exogenous exploitation.

The availability of dense genetic maps of cattle has allowed the whole genome to be evaluated for QTL with major effect. Publication of the results of the first genome-wide scan (in the US Holstein population by Georges *et al.* [24]) was followed by many partial and full genome scans in a number of populations [11,61]. However, apart from the summary provided by Bovenhuis and Schrooten [11], there have been no formal attempts to assemble a consensus map of the QTL derived from different studies.

One major purpose of this article is to review the results of QTL mapping in dairy cattle. The available information in the public domain has an emphasis on milk production and milk composition traits. However, work on other traits is also reviewed. Based on this information, we have developed an on-line QTL map for milk production traits. Furthermore, we have devised and adapted methodologies for undertaking meta-analysis of the published reports to estimate the consensus location of QTL, as well as the effects associated with these QTL.

## **2. REVIEW OF LITERATURE**

### **2.1. Dairy resource populations**

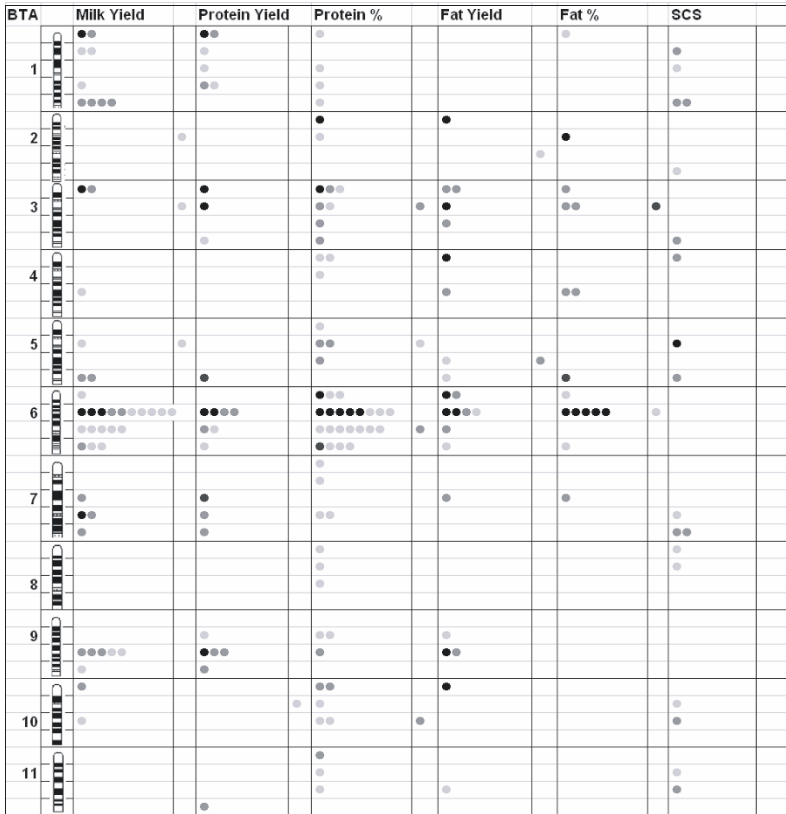
The basic resources critical to mapping of QTL are appropriate pedigreed populations with production records and genomic DNA samples. Weller *et al.* [67] proposed the use of the granddaughter design (GDD) and daughter design (DD) as methods for QTL detection in dairy cattle. For a DD, genotypic information is recorded for sires and their daughters, with phenotypic observations made on the daughters. For a GDD, the grandsires and sires are genotyped, and phenotypic observations are made on the granddaughters. Weller *et al.* [67] demonstrated the increased power of the GDD over the DD as a result of highly accurate estimates of the breeding values of the sires.

Both partial and full genome scans for QTL have been conducted on a number of dairy cattle populations using GDDs. One such population of US Holstein Friesians is the Dairy Bull DNA Repository (DBDR), which has been extensively used for QTL detection [5, 31, 53]. Most of the DBDR sires were used in the 1980s and so this population may not be representative of the present population. A new population termed the Cooperative Dairy DNA Repository (CDDR) is being created for analysis of current generations [4]. In a separate GDD, Georges *et al.* [24] and Zhang *et al.* [71] used 14 half-sib families with a total of 1518 sons from the US Holstein population. QTL detection studies using GDDs have also been published based on the New Zealand and Dutch dairy populations [3, 15, 27, 52, 60, 61], German Holsteins [22, 38, 43], Finnish Ayrshires [63, 65], British black and white cattle [68], Canadian Holsteins [48, 50], Norwegian cattle [49] and French dairy cattle [10]. Lipkin *et al.* [41] and Mosig *et al.* [47] used selective DNA pooling with a DD in Israeli Holstein Friesian cattle. Ron *et al.* [54] used a DD in the Israeli Holstein Friesian population for QTL mapping on BTA6 (*Bos taurus* autosome 6). Grisart *et al.* [27] and Heyen *et al.* [31] also used a DD as a part of their investigations of QTL on BTA14. More flexible designs are now being utilized, thanks to the development of suitable complex pedigree analysis methods [13]. Specific mapping populations for QTL detection in dairy cattle based on intercrossing breeds with extreme differences in lactation performance have also been initiated [39, 69] and will be informative in explaining the genetic differences between breeds as well as providing vital evidence of genes with potentially large effect on dairy production which have become fixed in the specialist dairy breeds.

Fine mapping of QTL for economic traits is at an early stage in livestock [9, 20, 52]. Riquet *et al.* [52] used a fine-mapping approach for QTL affecting milk composition based on the utilization of historical records of recombination and identity-by-descent (IBD) mapping exploiting linkage disequilibrium (LD) in the New Zealand and Dutch Holstein Friesian population [21]. A combined linkage and linkage disequilibrium mapping approach was also implemented in the same population for fine mapping QTL for fat percent [20] and protein percent [9].

## 2.2. QTL mapping results

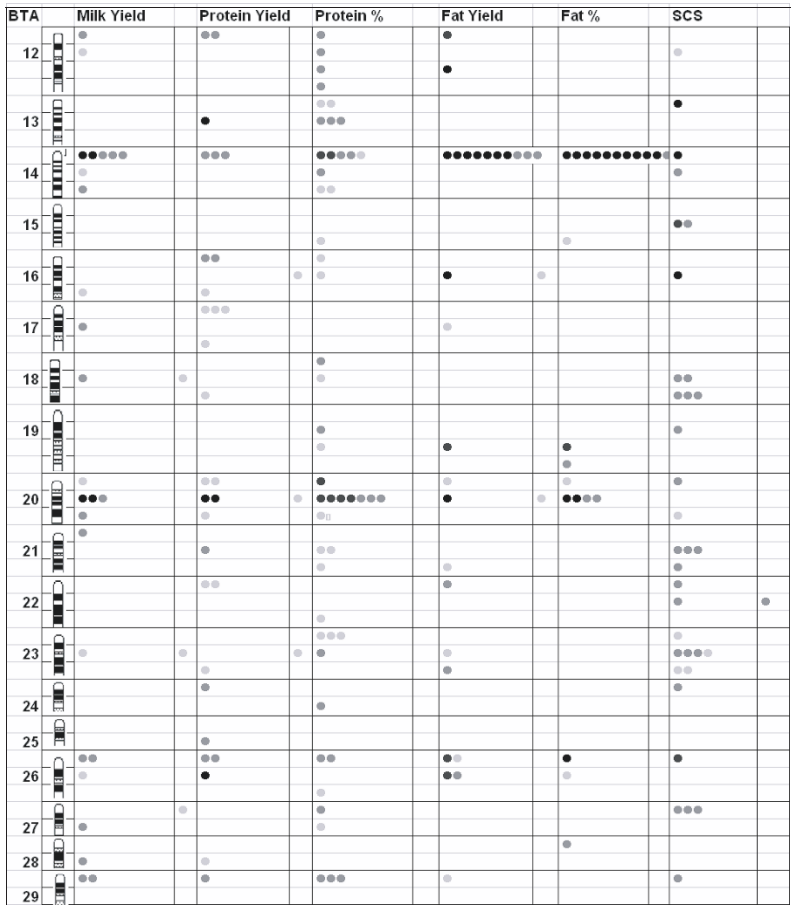
In total, 55 published papers on QTL detection in dairy cattle were reviewed for this study, including milk production, somatic cell score. This included published papers up to May 2003, and the reported QTL must have been



**Figure 1. a** – QTL map for milk production traits in dairy cattle: BTA1-BTA11. Each chromosome has been divided into approximate 30 cM regions, and the location of a QTL reported by a study has been placed in one of these bins, as indicated by a dot. The right hand column for each trait indicates that the location of the QTL was not reported in the study, other than being associated with that chromosome. The level of shading of the dot indicates the statistical significance for support of the QTL: ●  $P < 0.001$  or reported as highly significant; ●  $0.001 < P < 0.01$  or reported as significant; and ●  $0.01 < P < 0.05$  or reported as marginally significant.

statistically significant in some sense. In some cases the results from the same resource population were reported on more than one occasion where different marker density or different statistical approaches were used.

A QTL map summarizing the results from 45 of the above 55 papers for milk yield, milk composition traits and somatic cell score is presented in Figure 1. The map shows the distribution of reported QTL over the entire cattle genome at 30 centimorgan (cM) intervals. The QTL have been categorized into three groups according to significance thresholds, as determined by the



**Figure 1. b** – QTL map for milk production traits in dairy cattle: BTA12-BTA29. •  $P < 0.001$  or reported as highly significant; •  $0.001 < P < 0.01$  or reported as significant; and •  $0.01 < P < 0.05$  or reported as marginally significant.

reported  $P$ -values, whether they be point-wise, chromosome-wide, genome-wide, or unspecified. An online version of this QTL map is available at [http://www.vetsci.usyd.edu.au/reprogen/QTL\\_Map](http://www.vetsci.usyd.edu.au/reprogen/QTL_Map). Clicking on a dot representing a QTL displays a popup table of detailed information about that QTL, namely resource population and design, analytical method, marker map used, map position with confidence interval, closest marker, test statistics, effect size and reference. Note that some of entries in this online map are incomplete, due to a lack of reported information in the cited reference.

### 2.2.1. Milk yield

QTL affecting milk yield (MY) have been identified on 20 of the 29 bovine chromosomes (see Fig. 1). A notable number of studies detected the presence of QTL related to MY on BTA6; however, the position of markers/QTL differed in the various studies. There are also strong indications of the presence of QTL for MY on chromosome 1, 3, 9 and 20 with evidence of QTL at lower reported frequency on other chromosomes (5, 7, 10, 12, 14, 17, 18, 21, 23, 27 and 29).

A consistent finding across studies reporting QTL for MY on BTA6 suggests a primary QTL segregating near the center of BTA6 close to marker BM143 [38, 49, 54, 63] and a second QTL more distant from the centromere [54, 60, 68].

Arranz *et al.* [3] reported a QTL on BTA20 in one family having an allelic substitution effect of 308 kg on MY DYD (daughter yield deviation) and no significant effect on protein yield or fat yield. Probably the same QTL was detected by Georges *et al.* [24] with an allelic substitution effect of 342 kg. Nadesalingam *et al.* [48] indicated the presence of two QTL on BTA1 affecting MY.

A number of attempts have been made to detect an association between the casein gene complex located on BTA6 and milk production [12, 32, 40, 62]. Bovenhuis and Weller [12] used protein genes as markers to detect the linked QTL in the Dutch dairy cattle. Based on a GDD, Lien *et al.* [40] found a significant association of a paternal haplotype having the rare casein ( $\alpha_{s1}$ -CN(C)) allele with an increase in protein yield in a Norwegian cattle family. Velmala *et al.* [62] observed at least one QTL for milk yield and fat yield in the Finnish Ayrshire breed, linked to a casein haplotype segregating in one family.

### 2.2.2. Protein percent and yield

There is strong evidence of QTL on chromosomes 3, 6 and 20 for protein percentage (PP) and on chromosomes 1, 3, 6, 9, 14 and 20 for protein yield (PY). There are also some indications for QTL on other chromosomes (Fig. 1). A QTL for PP in the center of BTA6 has been reported to have an allelic substitution effect of up to 0.07% [54], 0.15% [60], 0.14% [63] and 0.09% [71]. Ashwell *et al.* [6] and Ron *et al.* [54] fine mapped their QTL for PP around the center of BTA6. Another significant QTL on BTA6 around the casein complex affecting PP, MY and FY (fat yield) has been reported by Velmala *et al.* [63].

QTL primarily affecting PP have been detected on BTA20 [3, 10, 24, 71]. Kim *et al.* [34] fine mapped a QTL for PP to the growth hormone receptor

(GHR) gene. On the same chromosome, a QTL for PY was detected at 46–70 cM in Norwegian cattle [49] and at 48 cM in DBDR families [53]. Ashwell *et al.* [5], Boichard *et al.* [10], Heyen *et al.* [31] and Zhang *et al.* [71] reported a QTL for PP towards the centromeric end of BTA3. Heyen *et al.* [31] and Rodriguez-Zas *et al.* [53] reported a QTL for PY on BTA3 near marker IL-STS96 (29.7 cM). Mosig *et al.* [47] employed selective DNA pooling in a DD and found 19–28 QTL affecting PP across the genome in the Israeli Holstein Friesian population.

### **2.2.3. Fat percent and yield**

A genome-wide significant QTL for fat percent (FP) and FY with large effect was detected near the centromeric end of BTA14 using a GDD and confirmed with a DD in an independent population [31]. The same QTL has also been reported in many other studies [5, 10, 13, 15, 43, 52, 55, 71]. This QTL is discussed in more details in Section 4.2. Another genome-wide significant QTL for FP was mapped around 41 cM on BTA3 with an allelic substitution effect of 0.07% [31]. Plante *et al.* [50] and Ron *et al.* [55] also detected a significant QTL for FP on BTA3. QTL for FY have also been identified on this chromosome [49, 53]. Many additional QTL with significant effects on FP and FY have been reported for chromosomes 5, 6, 9, 20 and 26.

### **2.2.4. QTL affecting more than one milk production trait**

Several chromosomes, particularly BTA3, 6, 9, 14, 20 and 23, have been reported to harbor QTL with pleiotropic effects on multiple milk production traits, and this should be expected based on our knowledge of genetic correlations among traits. Coppieters *et al.* [15] and Looft *et al.* [43] detected one QTL in the centromeric region of BTA14 that increases MY and PY while concomitantly reducing FY. This is consistent with the report by Grisart *et al.* [27] where the putative functional SNP in this region of BTA14, with a favorable effect on FY had an unfavorable effect on MY and PY, therefore decreasing the usefulness of such a direct marker for MAS. Wiener *et al.* [68] observed that a QTL on BTA6 had simultaneous effects on MY, FY and PY. Georges *et al.* [24] reported a QTL on BTA6 caused an increase in MY without a concomitant change in FY and PY. However, Zhang *et al.* [71] detected two distinct QTL on BTA6, one affecting MY (40 cM) and another affecting FP and PP (12 cM). Freyer *et al.* [22] fitted a pleiotropic model on BTA6 using a multivariate QTL mapping method, which supported the presence of a significant pleiotropic QTL at 68 cM for FY and PY.

Having evaluated the evidence for QTL of various milk production traits, a range of other relevant traits will now be considered.

### ***2.2.5. Somatic cell score and mastitis***

There are quite a few studies on QTL for somatic cell score (SCS). The US Holstein cattle population exhibited a QTL for SCS on BTA18 [7, 53]. Schrooten *et al.* [56] detected QTL for SCS on BTA18 in Dutch Holsteins. Schulman *et al.* [57] identified a QTL for both SCS and mastitis on the distal end of BTA18 in Finnish cattle. The same QTL was also detected in German cattle [8, 37]. Ashwell *et al.* [7] detected significant marker allele differences for SCS on BTA23 for markers 513, BM1818, BM1443 and BM4505. The QTL for SCS on BTA23 near marker RM033 has been reported in German cattle [51]. Heyen *et al.* [31] also detected an association of SCS with marker MGTG7 on BTA23. This marker is located near the bovine major histocompatibility complex (MHC). Klungland *et al.* [35] and Reinsch *et al.* [51] detected QTL on BTA8, which may be of interest because this region contains four interferon loci. In addition, the presence of QTL for SCS on chromosomes 1, 5, 7, 10, 11, 14, 15, 20, 21, 22, and 27 has been reported in more than one study [5, 10, 31, 37, 51, 53, 56, 71]. Additional QTL have been identified on other chromosomes, but only in single studies. However, the major interest in SCS is as an indicator to susceptibility to mastitis. Klungland *et al.* [35] reported a genome-wide significant QTL affecting clinical mastitis near BM143 on BTA6 and additional QTL for clinical mastitis on BTA3, 4, 14 and 27 in Norwegian cattle. The mastitis QTL on BTA6 is in the region of the QTL for milk production, and this may partially account for the unfavorable genetic correlation between high milk production and increased susceptibility of mastitis. Schulman *et al.* [57] reported QTL for mastitis on BTA14 and BTA18 in Finnish Ayrshire cattle. The distal end of BTA18 showed linkage both for SCS and mastitis. However, in general there seems to be no clear correspondence between the QTL for SCS and mastitis.

### ***2.2.6. Conformation and type traits***

The reports on conformation and type traits are available mainly from DBDR families [5], Dutch Holstein Friesian [56] and French dairy cattle [10] GDD studies. Ashwell *et al.* [5] reported QTL for dairy form on BTA5 and BTA27. Dairy form is a conformation trait based upon body condition, and



has a moderate relationship with milk production [58]. Ashwell *et al.* [5] reported an association between BB709 on BTA16 and udder depth. Biochard *et al.* [10] detected nine putative QTL for udder depth, but no highly significant QTL was found. Schrooten *et al.* [56] detected QTL for dairy character, a composite trait, at the centromeric end of BTA6. A QTL influencing fore-udder attachment was located at the centromeric end of BTA13 and another QTL influencing fore-udder attachment and front-teat placement was found on BTA19 [56]. Schrooten *et al.* [56] reported QTL affecting stature, size, chest width, body capacity and birth weight on BTA5. The same QTL for stature on BTA5, significant at genome-wise level, was also detected in French dairy cattle [10]. Another QTL for stature and size was detected on BTA6 [56]. Elo *et al.* [18] found evidence for a QTL affecting live weight on BTA23 in Finnish Ayrshire cattle. Because the traits are defined differently in each study, the results cannot be directly compared. More studies with consistent trait definitions will be required to refine the location of conformation QTL.

### **2.2.7. Reproduction**

A QTL affecting gestation length was reported in one study on BTA4 [56]. A QTL affecting dystocia and stillbirth is closely linked to the BoLA complex on BTA23 in German Holstein Friesians [28]. Kuhn *et al.* [37] detected QTL for dystocia on BTA8, BTA10 and BTA18, and for stillbirth on BTA6. QTL for *post partum* fertility (success/failure of each insemination of the daughters) were detected on chromosomes 1, 7, 10, 20 and 21 in French dairy cattle [10]. Putative QTL for non-return rate of 90 days were detected on BTA10 and BTA18 in German Holstein Cattle [37].

### **2.2.8. Other traits**

Elo *et al.* [18] reported a genome-wide significant QTL mapped for veterinary treatment (health index which includes all treatments other than for fertility and mastitis) and a QTL affecting ketosis in Finnish Ayrshire cattle on BTA23. There was also some support for QTL for calf mortality and milking speed on the same chromosome by these authors. More recently Schulman *et al.* [57] identified QTL on chromosomes 1, 2, 5, 8, 15, 22 and 23 for veterinary treatment in Ayrshire cattle.

## **2.3. Assessing the QTL mapping results**

The summary map of published QTL (Fig. 1) indicates that there are a large number of reports of QTL for milk production traits. Inspection of these reports

indicates some very interesting similarities among some studies, but also some marked differences in the location and magnitude of the effects of individual QTL. Not surprisingly, there are differences between families, even in the same study, in the level of significance, effect size and location of a particular QTL. There are also differences among studies in the criteria defining the significance thresholds, design methodologies, etc., which make the results of different studies difficult to compare. Consequently, there is a need to determine consensus location(s) of the QTL, as well as consensus estimates of the effects of these QTL. This has been achieved by means of a meta-analysis, as shown in the next section.

### 3. META-ANALYSIS METHODOLOGY

Efforts to combine findings from separate studies have a long history. In 1976, G. Glass proposed a method to integrate and summarize the findings from a body of research. He called the method meta-analysis [25]. Since that time, meta-analysis has become a widely accepted research tool in a variety of disciplines, especially in the medical, social and behavioral sciences [30]. Meta-analysis involves the application of standard statistical principles (hypothesis testing, inference) to situations where only summary information is available (*e.g.* published reports), and not the source unit record data. Well-conducted meta-analysis allows for a more objective appraisal of the evidence, which may lead to resolution of uncertainty and disagreement. Meta-analysis makes the literature review process more transparent, compared with traditional narrative reviews where it is often not clear how the conclusions follow from the data examined [17]. The application of meta-analysis to QTL detection is recent [26, 29]. The combining of the results across studies can provide a more precise and consensus estimate of the location of a QTL and its effect as compared with any single study. However, there are many challenges in combining the results of QTL mapping across studies, namely differences in marker density, linkage map, sample size, study design, as well as statistical methods used.

#### 3.1. Meta-analysis methodology of QTL location

We followed the method described by Goffinet and Gerber [26]. In summary, with a total of  $n$  published reports of a QTL on a particular chromosome, the statistical question is to decide on whether these reports represent a single QTL, two QTL, etc. up to  $n$  separate QTL (one for each publication).

Assessment of the number of QTL can be made on the basis of a likelihood ratio test (LRT), Akaike Information Content (AIC), or adjusted AIC (AIC\*), as in the method outlined by Goffinet and Gerber [26]. This involves selecting from amongst the best-fitting models with 1, 2, ...,  $n$  distinct QTL. As a result each published QTL can then be allocated to its respective consensus QTL. Note that usually, only the latest paper in a publication series on the same study population was included, to avoid duplication of the same QTL report. For a publication to be included in this meta-analysis, it ideally provides the interval map (test statistic profile). As well as providing the estimate of QTL location ( $\hat{d}_i$ ), the interval map also enables estimation of the standard error for QTL location,  $\sigma_i = \text{se}(\hat{d}_i)$ , after conversion of the test statistic to a (approximate) log-likelihood ( $\ln L$ ) scale. We suggest that the standard error can be estimated from the curvature (Fisher information) of the log-likelihood profile at the estimated map position,

$$\sigma_i = \left[ -\partial^2 \ln L / \partial d^2 \Big|_{d=\hat{d}_i} \right]^{-1/2}.$$

In particular, the curvature was estimated by fitting a local quadratic near the maximum of  $\ln L$ , and determining the coefficient of the quadratic term. Note that these standard errors were used to construct a weighted estimate of QTL location, the weights being inversely proportional to the squared standard errors ( $w_i = \sigma_i^{-2}$ ).

For studies that did not include an interval map, we computed average standard errors,

$$\bar{\sigma} = \sqrt{(1/m) \sum_{i=1}^m \sigma_i^2},$$

based on the  $m$  studies where interval maps were available.

Some of the studies used marker distances computed from the observed marker data while others used the USDA MARC cattle map (<http://www.marc.usda.gov/>). For the meta-analysis, QTL positions were rescaled to the USDA map by using the location of the nearest flanking markers on both maps, with a similar linear re-scaling of the standard errors ( $\sigma_i$ ). However, reported (unadjusted) QTL positions were used for the meta-analyses of FY and FP on BTA14, due to the recent work conducted on the centromeric end not covered by the USDA map. Additional details on the implementation of the meta-analysis of QTL location are presented in the Appendix.

### 3.2. Meta-analysis methodology of QTL effects

After estimating the consensus QTL position using the above approach, we conducted a meta-analysis for the effect size for each consensus QTL. For a paper to be included in this part of the meta-analysis, it needed to include effect information and its standard error (or the ability to derive this). The basis for the QTL effects model is similar to that outlined in Hayes and Goddard [29]. However, in the current application, the focus is on a meta-analysis of effects at the one locus – or at least chromosomal region – rather than across all loci. Suppose that for a consensus QTL, the QTL allelic substitution effects ( $a$ ) differ from sire to sire, and we will assume that  $a \sim N(0, \sigma_A^2)$ . The purpose behind this meta-analysis is to estimate the variance of these effects,  $\sigma_A^2$ . Next assume that for each sire in the available studies, the estimate of the QTL allelic substitution effect,  $a_i$ , is  $\hat{a}_i$  with corresponding standard error  $\varsigma_i = \text{se}(\hat{a}_i)$  and variance  $\varsigma_i^2$ ,  $i = 1, 2, \dots, n$ , where  $n$  is the number of sires. To model the imprecision of  $\hat{a}_i$  estimating  $a_i$ , we assume that  $\hat{a}_i|a_i \sim N(a_i, \varsigma_i^2)$ , and consequently, the unconditional distribution of estimated effects will be  $\hat{a}_i \sim N(0, \varsigma_i^2 + \sigma_A^2)$ . As also considered by Hayes and Goddard [29], there are two other features that need to be modeled in the meta-analysis. Since it is to a certain extent arbitrary which sire allele is labeled as having a positive effect, we will ignore the sign and condition on  $a_i > 0$  and  $\hat{a}_i > 0$ . Secondly, only “significant” QTL tend to be published (resulting in potential publication bias), so we assume that  $\hat{a}_i > c$  where  $c$  is the “threshold” QTL effect that just reaches “publication level”. With these constraints, the probability density function,  $h(\cdot)$ , for the observed QTL effects will be

$$h(\hat{a}_i|\hat{a}_i > c) = n_i(\hat{a}_i)/[1 - N_i(c)], \quad \hat{a}_i > c,$$

say, where

$$n_i(y) = \frac{1}{\sqrt{2\pi(\varsigma_i^2 + \sigma_A^2)}} \exp\left(-\frac{y^2}{2(\varsigma_i^2 + \sigma_A^2)}\right)$$

is the normal probability density function, and  $N_i(y) = \int_{-\infty}^y n_i(t) dt$  is the corresponding cumulative normal distribution function. So there are two parameters to be estimated,  $\sigma_A^2$  and  $c$ , and this is achieved by a maximum likelihood procedure (see Appendix for details).

For those papers where  $\zeta_I$  was not reported, the average value ( $\bar{\zeta}$ ) was computed in a similar way to that of  $\bar{\sigma}$ . However, because the different studies were conducted under different conditions, there was a large variation in the phenotypic standard deviation across studies, for a particular trait. Consequently,

both the effect estimates and their standard errors were re-scaled by dividing by their reported phenotypic standard deviations (where reported), or by appropriate consensus standard deviations used for international evaluations [33] where this was not reported. Consequently, the consensus estimate of  $\sigma_A^2$  will be the *proportion* of the phenotypic variance explained by the consensus QTL.

#### 4. META-ANALYSIS RESULTS

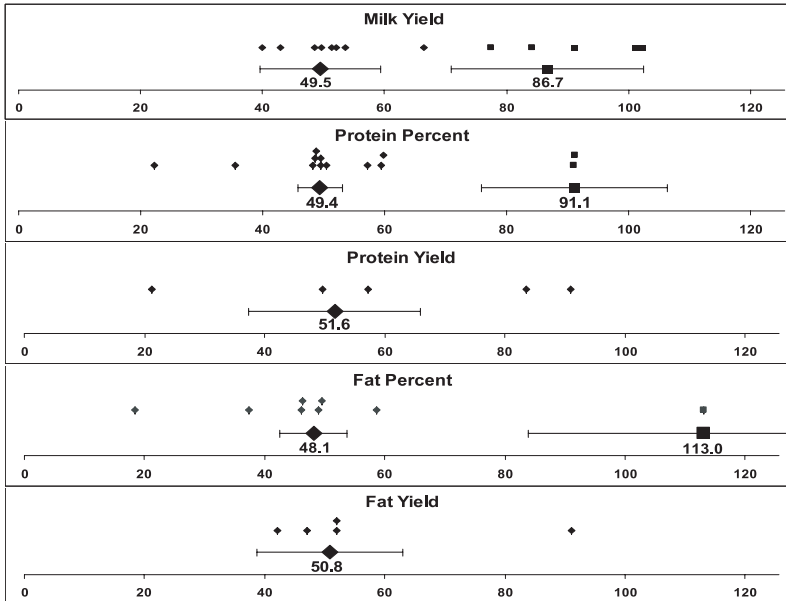
From the 55 published reports on QTL, only 28 contained sufficient information to be included in the meta-analysis (Tab. I). We focused on the meta-analysis for results reported for BTA6 since this chromosome contained the highest number of reported QTL. Meta-analysis was also applied to all other chromosomes where there were sufficient usable studies. Note that the same study may appear more than once for the same chromosome–trait combination, when the different positions were reported for different families.

##### 4.1. Milk production QTL on BTA6

Meta-analysis of studies reporting 13 QTL for MY on BTA6 suggested that the presence of two QTL, one at  $49 \pm 5.0$  cM and another QTL located at  $87 \pm 7.9$  cM, described the data best. The analysis of effect sizes of these two QTL indicated that they explained on average  $4.18 \pm 3.12$  and  $3.63 \pm 5.57$  percent of total phenotypic variance, respectively. For PP, two QTL ( $49 \pm 1.8$  cM and  $91 \pm 7.6$  cM) were identified, again based on 13 QTL reports. The first QTL explained  $1.53 \pm 1.30$  percent of the phenotypic variance. There was evidence of only one QTL for PY ( $52 \pm 7.2$  cM) based on five QTL reports. The eight reports of QTL for FP resulted in two QTL ( $48 \pm 2.8$  and  $113 \pm 14.6$  cM). For FY, the evidence from five reported QTL suggests a single consensus QTL ( $51 \pm 6.0$  cM).

Many studies reported that there was one QTL affecting all five milk production traits, located in the middle (around 50 cM) of BTA6 near marker BM143 (Fig. 2). There was also evidence based on this meta-analysis of second QTL around 87 cM affecting both MY and PP, but the evidence was lacking for PY and FY.

As well as there being support from this meta-analysis for the one QTL on BTA6 affecting multiple traits, there is also direct evidence from individual studies of QTL affecting multiple traits, as outlined previously. The meta-analysis has also found evidence of multiple QTL on a chromosome affecting the one trait, and this is supported by some individual studies. Freyer *et al.* [23],

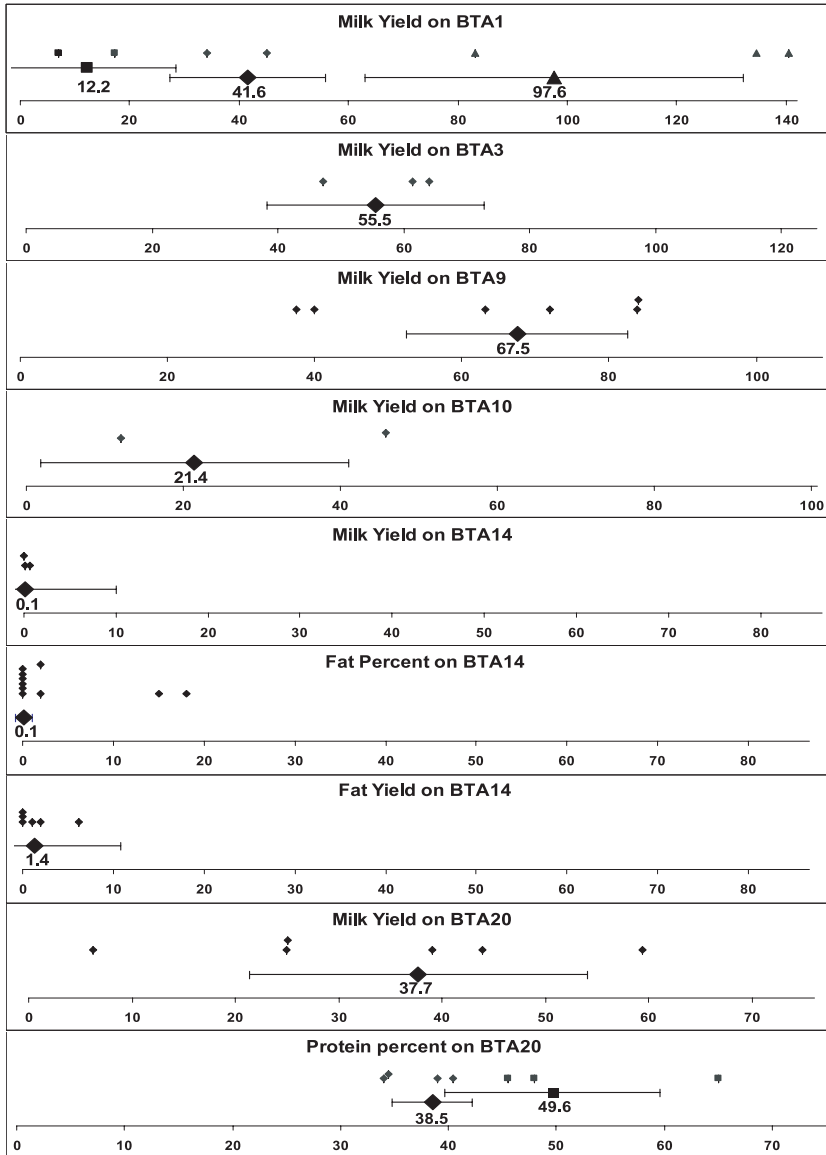


**Figure 2.** Meta-analysis of milk production traits on BTA6. Units of the horizontal axis are in cM. Each small symbol represents a separate study included in the meta-analysis, with the larger symbol indicating the consensus location of the QTL with corresponding 95% confidence interval. Symbol types indicate the grouping of individual studies to the consensus QTL.

Spelman *et al.* [60], Velmala *et al.* [63] and Zhang *et al.* [71] fitted a two-QTL model for milk production traits on BTA6. Velmala *et al.* [63] suggested the presence of two QTL on BTA6, one close to BM143 affecting PP and MY, and another located around the casein complex, affecting PP, MY and FY. Zhang *et al.* [71] indicated that in those families where there was evidence in favor of a two-QTL model, the two loci were in repulsion phase. Cohen *et al.* [14] reported an association between a SNP, mapped in the middle of BTA6, and both protein yield and Israeli breeding index, in the Israeli Holstein sire population.

#### 4.2. Milk production QTL on other chromosomes

For BTA1 the meta-analysis of seven QTL indicated the presence of three QTL for MY at  $12 \pm 8.1$ ,  $42 \pm 7.1$  and  $98 \pm 17.3$  cM (Fig. 3, Tab. I). There was support for one QTL affecting MY at  $56 \pm 8.6$  cM on BTA3 based on three QTL reports and one on BTA9 at  $68 \pm 7.5$  cM based on six QTL reports, the latter explaining about 1.7 percent of the total variance. Single QTL for MY were



**Figure 3.** Meta-analysis of milk production QTL on other chromosomes. Units of the horizontal axis are in cM. Each small symbol represents a separate study included in the meta-analysis, with the larger symbol indicating the consensus location of the QTL with corresponding 95% confidence interval. Symbol types indicate the grouping of individual studies to the consensus QTL.

**Table I.** Summary of meta-analysis results.

BTA	Trait	Position (cM)		% Var. explained		LRT Statistic			No. of Studies	References <sup>a</sup>
		Estimate	SE	Estimate	SE	2 vs. 1	3 vs. 2	3 vs. 1		
6	MY	49.5	4.96	4.18	3.12	15.96	2.24	8	23, 54, 60, 63, 54, 38, 49, 71	
6	MY	86.7	7.89	3.63	5.57			5	23, 48, 68, 54, 60	
6	PP	49.4	1.83	1.53	1.30	28.38	2.18	11	38, 23, 71, 60, 54, 54, 63, 6, 10, 48, 49	
6	PP	91.1	7.60	-				2	6, 54	
6	PY	51.6	7.18			3.00	2.12	5	53, 54, 38, 48, 68	
6	FP	48.1	2.84	-		19.05	1.87	7	38, 71, 23, 54, 6, 54, 49	
6	FP	113.0	14.59	-				1	54	
6	FY	50.8	6.04	-		1.95	0.24	5	71, 54, 38, 54, 68	
1	MY	12.2	8.08	-		14.47	7.46	2	48, 24	
1	MY	41.6	7.12					2	48, 48	
1	MY	97.6	17.3	-				3	48, 46, 53	
3	MY	55.5	8.63			0.78	0.01	3	31, 71, 50	
9	MY	67.5	7.49	1.71	4.72	1.52	0.74	6	50, 65, 24, 68, 53, 68	
10	MY	21.5	9.85			2.36		2	50, 24	
14	MY	0.1	4.97			0.00	0.00	3	10, 43, 15	
14	FP	0.1	0.47			3.14	0.89	11	8, 10, 66, 66, 66, 70, 27, 31, 5, 55, 13	
14	FY	1.4	4.71			0.15	0.02	6	8, 10, 66, 31, 27, 5	
20	MY	37.7	8.18			1.63	0.26	6	46, 3, 50, 9, 24, 46	
20	PP	38.5	1.86			4.32	1.77	4	13, 10, 9, 24	
20	PP	49.6	5.01					3	5, 3, 24	

<sup>a</sup> Note that the reference lists for each QTL consensus region are shown in estimated QTL position order.



also identified on BTA10 (at  $22 \pm 9.9$  cM), at the centromeric end of BTA14, and on BTA20 ( $38 \pm 8.2$  cM) from the meta-analysis of two, three and six QTL reports, respectively.

The meta-analysis of 11 reports indicated the presence of a single QTL for FP on the centromeric end ( $0.1 \pm 0.47$  cM) of BTA14. The QTL for FY was estimated at  $1.4 \pm 4.71$  cM based on six reports. Riquet *et al.* [52] identified a 5 cM common haplotype containing QTL affecting FP at BTA14q11-16 using IBD (identity-by-descent) mapping. Looft *et al.* [43] found linkage disequilibrium between this QTL and the expressed sequence tag KIEL\_E8 mapped to the proximal region of BTA14. Grisart *et al.* [27] identified a strong candidate gene, acylCoA: diacylglycerol acyltransferase (DGAT1), which mapped in this region. They identified a SNP (nonconservative K232A substitution) in the DGAT1 gene which accounted for 51 percent of the DYD variance of FP, and 8 percent of DYD variance of PP. Winter *et al.* [70] also found this mutation associated with milk fat, and a haplotype analysis in different breeds indicated the lysine (K) variant to be ancestral.

The meta-analysis of seven reports for PP on BTA20 suggests the presence of two QTL ( $38 \pm 1.9$  and  $50 \pm 5.0$  cM). Kim *et al.* [34] indicated the growth hormone receptor gene to be a strong candidate gene for the first PP QTL on BTA20. Further, Blott *et al.* [9] found a phenylalanine-to-tyrosine substitution in the transmembrane domain of this gene associated with a major effect on milk yield and composition.

Where there were sufficient studies, the meta-analysis of the location of the QTL could identify the consensus QTL on some of the chromosomes, and by pooling the results could actually refine the estimated positions of these QTL. In general, the meta-analysis of the effect sizes indicated that some of the QTL regions identified explained a relatively large proportion of the phenotypic variance, but this was often accompanied by large standard errors. These may consequently represent inflated estimates of effect size, but these would be expected to be moderated as increased numbers of studies become available. For other trait and chromosome combinations, insufficient information was available to allow for a meta-analysis.

## 5. DISCUSSION

A number of partial and complete dairy cattle genome scans have been undertaken in the last few years. QTL for milk production traits have been identified on all the autosomes of the bovine genome. Comparing the results across studies indicates strong evidence for QTL affecting milk production

traits segregating on chromosomes 1, 3, 6, 9, 14, and 20, while on other chromosomes (2, 11, 15, 22, 24, 25, and 28) QTL have only occasionally been reported (Fig. 1). However, in several studies only a limited selection of candidate chromosomes has been analyzed, *e.g.* several studies targeted BTA6 only for milk production QTL, so there is a clear bias. In any case, the current review and meta-analysis clearly indicates consensus for several QTL affecting dairy production traits. For example, two distinct consensus regions on BTA6 affecting MY have been identified, one near the marker BM143 and another located around the casein complex, and these account for 4.18 and 3.63 percent of phenotypic variance, respectively. In addition, the QTL near BM143 also appears to affect PY, PP, FY, and FP.

Typically, genome scan results indicate that chromosomal regions harboring QTL for dairy-related traits could at best be identified to within a confidence interval of 20 cM or more, which is not sufficient for MAS. However, QTL can be localized more precisely using LD and IBD mapping [52]. Additional analyses of current and new whole-genome scans will be very useful to ascertain the existence of suggestive QTL. However, meta-analysis of all available published results, as shown in this study, can also be used to improve QTL localization, and this has been achieved at no cost, relative to that of undertaking additional laboratory and field-based studies. In particular, the meta-analysis method will always produce a narrower confidence interval for a “consensus” QTL than the smallest of the individual studies included. Of course, further improvements could be made by pooling the unit-record data across studies. Some developments along these lines have already been made: Bennewitz *et al.* [8] reported the results from the joint analysis of two different granddaughter designs.

A thorough meta-analysis of dairy cattle QTL results is hampered by the publishing of only partial information on QTL mapping, and also by the current focus on only the most significant results. Authors should be encouraged to present a broader range of results. At a minimum, this would include (1) the population studied; (2) analytical method for QTL mapping; (3) marker map used; (4) interval map/profile test-statistic (or estimated map position and confidence interval); and (5) estimated effect size of QTL(s) and standard error(s). Clearly, a web-based repository (with appropriate quality control aspects) would be a major step forward, allowing for easy incorporation of additional information with rapid deliberation of QTL regions of interest. Although difficult to orchestrate, a further advance would be a standardization of design and methodologies to allow more transparent comparison and collation of results, as is done for multi-centre clinical trials in human studies [19].

There are striking differences in the magnitude of the effects of individual QTL. There are differences between families even in the same experiment in the level of significance and effect for a particular QTL. This may be partly due to segregation of various alleles of the same QTL or due to segregation of various QTL. The meta-analysis of the effect sizes indicated that some chromosomal regions explain a relatively significant proportion of variation in milk production traits. Based on results from QTL mapping experiments in dairy cattle, Hayes and Goddard [29] fitted gamma distributions of QTL effects. These were moderately leptokurtic, consistent with a few genes of large effect and many of small effect. They predicted 50 to 100 genes affecting a quantitative trait in dairy cattle, 17 percent of which could explain 90 percent of the genetic variance. However, as indicated from the experience of laboratory animals [16, 44], there is a need to investigate interaction among loci (epistasis), genotype-environment interactions, imprinting effects, and linked QTL in dairy cattle to fully understand the genetic architecture of quantitative traits.

The identification of the actual gene and the causative mutation comprising a QTL has been a challenge for several reasons. Causative mutations for quantitative traits are hard to find and difficult to prove [2]. A major hurdle is poor precision of the localization of QTL. Even after successfully fine mapping to localize a QTL to a relatively small segment (<5 cM), the progression from QTL to gene remains a daunting task. The most prudent experimental strategy in such situations is the identification of positional candidate genes by comparative mapping. Comparative map information has been used successfully to identify some genes underlying QTL with large effect [9, 27]. Even using this approach, one may be left with more candidate genes than can feasibly be tested. In addition to QTL mapping, a number of other strategies for gene discovery need to be integrated to understand the genetic architecture of complex traits, namely, genomics (analysis of the transcriptome), proteomics (protein-based gene expression analysis) and metabolomics (analysis of metabolic pathways) [42].

Markers used in all genome scans have been mostly microsatellites although SNPs have also been included in some studies [34, 52]. SNPs will probably replace microsatellites, especially for LD mapping, based on their promising features such as abundance in the genome, low mutation rate and amenability to automation [64]. Efforts to identify SNPs in cattle are underway [36]. Because SNPs are usually biallelic markers, a much larger set of markers will be required to produce the informative haplotype data required for genome analysis.

This review also shows that most QTL mapping effort has so far targeted production traits, particularly milk yield and composition. Much work remains to be done to improve our knowledge of genes regulating other traits, including behavioral, diseases resistance, reproduction, type, feed efficiency, persistency of lactation etc. The scope for MAS on these traits remains extremely high. Defining the most appropriate phenotype will be a crucial factor while hunting genes for these traits. Use of model animals may be more practical for understanding the genetics of some of the more novel traits.

## 6. CONCLUSIONS

To date, many complete and partial genome-wide scans for dairy cattle QTL have been published. Before major new initiatives in this area are undertaken, it is important to “take stock” in the most optimal way, of what has already been found. At virtually no cost, a meta-analysis of the published reports can be undertaken, provided the individual reports contain the necessary information on QTL position and effect with appropriate error estimates. Outputs from such analyses highlight the specific areas of the genome where future resources should be directed to refine characterization of the QTL. In this report meta-analysis has been shown to be effective at identifying a consensus position for a number of QTL for various dairy traits on BTA6, and also across a number of other chromosomes for milk yield.

## ACKNOWLEDGEMENTS

We would like to thank Frank Nicholas for suggestions and improvements to this manuscript. The work reported here was supported with funds from the Co-operative Research Centre for Innovative Dairy Products. M.S.K. is the recipient of an F.H. Loxton Scholarship.

## REFERENCES

- [1] Abdel-Azim G., Freeman A.E., Superiority of QTL-assisted selection in dairy cattle breeding schemes, *J. Dairy Sci.* 85 (2002) 1869–1880.
- [2] Andersson L., Genetic dissection of phenotypic diversity in farm animals, *Nat. Rev. Genet.* 2 (2001) 130–138.
- [3] Arranz J.J., Coppieters W., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Mezer C., Riquet J., Simon P., Vanmanshoven P., Wagenaar D., Georges M., A QTL affecting milk yield and composition maps to bovine chromosome 20: a confirmation, *Anim. Genet.* 29 (1998) 107–115.

- [4] Ashwell M.S., Van Tassell C.P., Sonstegard T.S., The cooperative dairy DNA repository: a new resource for quantitative trait loci detection and verification, Proceedings of the 8th Plant and Animal Genome Conference, San Diego, CA, 2000.
- [5] Ashwell M.S., Van Tassell C.P., Sonstegard T.S., A genome scan to identify quantitative trait loci affecting economically important traits in a US Holstein population, *J. Dairy Sci.* 84 (2001) 2535–2542.
- [6] Ashwell M.S., Schnabel R.D., Sonstegard T.S., Van Tassell C.P., Fine-mapping of QTL affecting protein percent and fat percent on BTA6 in a popular US Holstein family, in: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, 19–23 August 2002, ISBN 2-7380-1052-0, Paper 09-29.
- [7] Ashwell M.S., Rexroad C.E. Jr., Miller R.H., VanRaden P.M., Da Y., Detection of loci affecting milk production and health traits in an elite US Holstein population using microsatellite markers, *Anim. Genet.* 28 (1997) 216–222.
- [8] Bennewitz J., Reinsch N., Grohs C., Leveziel H., Malafosse A., Thomsen H., Xu N., Looft C., Kuhn C., Brockmann G.A., Schwerin M., Weimann C., Hiendleder S., Erhardt G., Medjugorac I., Russ I., Forster M., Brenig B., Reinhardt F., Reents R., Averdunk G., Blumel J., Boichard D., Kalm E., Combined analysis of data from two granddaughter designs: A simple strategy for QTL confirmation and increasing experimental power in dairy cattle, *Genet. Sel. Evol.* 35 (2003) 319–338.
- [9] Blott S., Kim J., Moisisio S., Schmidt-Küntzel A., Cornet A., Berzi P., Cambisano N., Ford C., Grisart B., Johnson D., Karim L., Simon P., Snell R., Spelman R., Wong J., Vilkki J., Georges M., Farnir F., Coppieters W., Molecular dissection of a quantitative trait locus: A phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition, *Genetics* 163 (2003) 253–266.
- [10] Boichard D., Grohs C., Bourgeois F., Cerqueira F., Faugeras R., Neau A., Rupp R., Amigues Y., Boscher M.Y., Leveziel H., Detection of genes influencing economic traits in three French dairy cattle breeds, *Genet. Sel. Evol.* 35 (2003) 77–101.
- [11] Bovenhuis H., Schrooten C., Quantitative trait loci for milk production traits in dairy cattle, in: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, 19–23 August 2002, ISBN 2-7380-1052-0, Paper 09-07.
- [12] Bovenhuis H., Weller J.I., Mapping and analysis of dairy cattle quantitative trait loci by maximum likelihood methodology using milk protein genes as genetic markers, *Genetics* 137 (1994) 267–280.
- [13] Chamberlain A., McParlan H., Balasingham T., Carrick M., Bowman P., Robinson N., Goddard M., Mapping QTL affecting milk composition traits in dairy cattle using a complex pedigree, Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, 19–23 August 2002, ISBN 2-7380-1052-0, Paper 09-08.

- [14] Cohen M., Serrousi E., Ron M., Reichenstein M., Plis-Finarov A., Shani M., Weller J.I., Population-wide linkage disequilibrium between a SNP and a QTL affecting milk protein production on BTA6 in dairy cattle, in: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, 19–23 August 2002, ISBN 2-7380-1052-0, Paper 09-12.
- [15] Coppieters W., Riquet J., Arranz J.J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Nezer C., Simon P., Vanmanshoven P., Wagenaar D., Georges M., A QTL with major effect on milk yield and composition maps to bovine chromosome 14, *Mamm. Genome* 9 (1998) 540–544.
- [16] Corva P.M., Medrano J.F., Quantitative trait loci (QTLs) mapping for growth in the mouse: A Review, *Genet. Sel. Evol.* 33 (2001) 105–132.
- [17] Egger M., Smith G.D., Meta-analysis bias in location and selection of studies, *Brit. Med. J.* 316 (1998) 61–66.
- [18] Elo K.T., Vilkki J., de Koning D.J., Velmala R.J., Maki-Tanila A.V., A quantitative trait locus for live weight maps to bovine chromosome 23, *Mamm. Genome* 10 (1999) 831–835.
- [19] Elwood J.M., Critical appraisal of epidemiological studies and clinical trials, 2nd edn., Oxford University Press, Oxford, 1998.
- [20] Farnir F., Grisart B., Coppieters W., Riquet J., Berzi P., Cambisano N., Karim L., Mni M., Moisisio S., Simon P., Wagenaar D., Vilkki J., Georges M., Simultaneous mining of linkage and linkage disequilibrium to fine map quantitative trait loci in outbred half-sib pedigrees. Revisiting the location of a quantitative trait locus with major effect on milk production on bovine chromosome 14, *Genetics* 161 (2002) 275–287.
- [21] Farnir F., Coppieters W., Arranz J.J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Mni M., Nezer C., Simon P., Vanmanshoven P., Wagenaar D., Georges M., Extensive genome-wide linkage disequilibrium in cattle, *Genome Res.* 10 (2000) 220–227.
- [22] Freyer G., Sorensen P., Kuhn C., Weikard R., Hoeschele I., Search for pleiotropic QTL on chromosome BTA6 affecting yield traits of milk production, *J. Dairy Sci.* 86 (2003) 999–1008.
- [23] Freyer G., Kuhn C., Weikard R., Zhang Q., Mayer M., Hoeschele I., Multiple QTL on chromosome six in dairy cattle affecting yield and content traits, *J. Anim. Breed. Genet.* 119 (2002) 69–82.
- [24] Georges M., Nielsen D., Mackinnon M., Mishra A., Okimoto R., Pasquino A.T., Sargeant L.S., Sorensen A., Steele M.R., Zhao X., Womack J.E., Hoeschele I., Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing, *Genetics* 139 (1995) 907–920.
- [25] Glass G.V., Primary, secondary and meta-analysis of research, *Educ. Res.* 5 (1976) 3–8.
- [26] Goffinet B., Gerber S., Quantitative trait loci: a meta-analysis, *Genetics* 155 (2000) 463–473.
- [27] Grisart B., Coppieters W., Farnir F., Karim L., Ford C., Berzi P., Cambisano N., Mni M., Reid S., Simon P., Spelman R., Georges M., Snell R., Positional candidate cloning of a QTL in dairy cattle: Identification of a missense mutation

- in the bovine DGAT1 gene with major effect on milk yield and composition, *Genome Res.* 12 (2002) 222–231.
- [28] Grupe S., Schwerin M., Mapping of quantitative trait loci on chromosome 23 in German Holstein cattle families, *Arch. Tierzucht* 41 (1998) 225–235.
- [29] Hayes B., Goddard M.E., The distribution of the effects of genes affecting quantitative traits in livestock, *Genet. Sel. Evol.* 33 (2001) 209–229.
- [30] Hedges L.V., Olkin I., *Statistical methods for meta-analysis*, Academic Press Inc., 1985.
- [31] Heyen D.W., Weller J.I., Ron M., Band M., Beever J.E., Feldmesser E., Da Y., Wiggans G.R., VanRaden P.M., Lewin H.A., A genome scan for QTL influencing milk production and health traits in dairy cattle, *Physiol. Genom.* 1 (1999) 165–175.
- [32] Ikonen T., Bovenhuis H., Ojala M., Ruottinen O., Georges M., Associations between casein haplotypes and first lactation milk production traits in Finnish Ayrshire cows, *J. Dairy Sci.* 84 (2001) 507–514.
- [33] Interbull, Genetic evaluations, 2002, <http://www-interbull.slu.se/eval/appen022.html> [consulted: 12 Feb. 2003].
- [34] Kim J.J., Farnir F., Coppieters W., Johnson D., Georges M., Evaluation of a new QTL fine-mapping method exploiting linkage disequilibrium on BTA14 and BTA20 in a dairy cattle, in: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France, 19–23 August 2002, ISBN 2-7380-1052-0, Paper 21-23.
- [35] 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., Quantitative trait loci affecting clinical mastitis and somatic cell count in dairy cattle, *Mamm. Genome* 12 (2001) 837–842.
- [36] Konfortov B.A., Licence V.E., Miller J.R., Re-sequencing of DNA from a diverse panel of cattle reveals a high level of polymorphism in both intron and exon, *Mamm. Genome* 10 (1999) 1142–1145.
- [37] Kuhn C., Bennewitz J., Reinsch N., Xu N., Thomsen H., Looft C., Brockmann G.A., Schwerin M., Weimann C., Hiendleder S., Erhardt G., Medjugorac I., Forster M., Brenig B., Reinhardt F., Reents R., Russ I., Averdunk G., Blumel J., Kalm E., Quantitative trait loci mapping of functional traits in the German Holstein cattle population, *J. Dairy Sci.* 86 (2003) 360–368.
- [38] Kuhn C., Freyer G., Weikard R., Goldammer T., Schwerin M., Detection of QTL for milk production traits in cattle by application of a specifically developed marker map of BTA6, *Anim. Genet.* 30 (1999) 333–340.
- [39] Larroque H., Gallard Y., Thauinat L., Boichard D., Colleau J.J., A crossbreeding experiment to detect quantitative trait loci in dairy cattle, in: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France, 19–23 August 2002, ISBN 2-7380-1052-0, Paper 01-42.
- [40] Lien S., Gomez Raya L., Steine T., Fimland E., Rogne S., Associations between casein haplotypes and milk yield traits, *J. Dairy Sci.* 78 (1995) 2047–2056.
- [41] Lipkin E., Mosig M.O., Darvasi A., Ezra E., Shalom A., Friedmann A., Soller M., Quantitative trait locus mapping in dairy cattle by means of selective milk



DNA pooling using dinucleotide microsatellite markers: Analysis of milk protein percentage, *Genetics* 149 (1998) 1557–1567.

- [42] Liu H.C., Cheng H.H., Integrating molecular approaches with QTL to identify positional candidate genes, in: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France, 19–23 August 2002, ISBN 2-7380-1052-0, Paper 21-22.
- [43] Looft C., Reinsch N., Karall-Albrecht C., Paul S., Brink M., Thomsen H., Brockmann G., Kuhn C., Schwerin M., Kalm E., A mammary gland EST showing linkage disequilibrium to a milk production QTL on bovine Chromosome 14, *Mamm. Genome* 12 (2001) 646–650.
- [44] Mackay T.F.C., Quantitative trait loci in *Drosophila*, *Nat. Rev. Genet.* 2 (2001) 11–20.
- [45] Mackinnon M.J., Georges M.A.J., Marker-assisted preselection of young dairy sires prior to progeny-testing, *Livest. Prod. Sci.* 54 (1998) 229–250.
- [46] Maki-Tanila A., De Koning D.J., Elo K.T., Moisisio S., Velmala R., Vilkkii H.J., Mapping multiple quantitative trait loci by regression in half sib designs, in: *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production*, Armidale, Vol. 26, 1998, 6 WCGALP Congress Office, University of New-England, Armidale, pp. 269–272.
- [47] Mosig M.O., Lipkin E., Khutoreskaya G., Tchourzyna E., Soller M., Friedmann A., A whole genome scan for quantitative trait loci affecting milk protein percentage in Israeli-Holstein cattle, by means of selective milk DNA pooling in a daughter design, using an adjusted false discovery rate criterion, *Genetics* 157 (2001) 1683–1698.
- [48] Nadesalingam J., Plante Y., Gibson J.P., Detection of QTL for milk production on Chromosomes 1 and 6 of Holstein cattle, *Mamm. Genome* 12 (2001) 27–31.
- [49] Olsen H.G., Gomez-Raya L., Vage D.I., Olsaker I., Klungland H., Svendsen M., Adnøy T., Sabry A., Klemetsda G., Schulman N., Kramer W., Thaller G., Rønningen K., Lien S., A genome scan for quantitative trait loci affecting milk production in Norwegian dairy cattle, *J. Dairy Sci.* 85 (2002) 3124–3130.
- [50] Plante Y., Gibson J.P., Nadesalingam J., Mehrabani-Yeganeh H., Lefebvre S., Vandervoort G., Jansen G.B., Detection of quantitative trait loci affecting milk production traits on 10 chromosomes in Holstein cattle, *J. Dairy Sci.* 84 (2001) 1516–1524.
- [51] Reinsch N., Thomsen H., Looft C., Kalm E., Grupe S., Kuhn C., Schwerin M., Leyhe-Horn B., Hiendleder S., Erhardt G., Medjugorac I., Russ I., Forster M., Brenig B., Reents R., Averdunk G., First results on somatic cell count loci from the ADR bovine mapping project, in: *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production*, Armidale, Vol. 26, 1998, 6 WCGALP Congress Office, University of New-England, Armidale, pp. 426–428.
- [52] Riquet J., Coppieters W., Cambisano N., Arranz J.J., Berzi P., Davis S.K., Grisart B., Farnir F., Karim L., Mni M., Simon P., Taylor J.F., Vanmanshoven P., Wagenaar D., Womack J.E., Georges M., Fine-mapping of quantitative trait loci by identity by descent in outbred populations: application to milk production in dairy cattle, *Proc. Natl. Acad. Sci. USA* 96 (1999) 9252–9257.



- [53] Rodriguez-Zas S.L., Southey B.R., Heyen D.W., Lewin H.A., Interval and composite interval mapping of somatic cell score, yield, and components of milk in dairy cattle, *J. Dairy Sci.* 85 (2002) 3081–3091.
- [54] Ron M., Kliger D., Feldmesser E., Seroussi E., Ezra E., Weller J.I., Multiple quantitative trait locus analysis of bovine chromosome 6 in the Israeli Holstein population by a daughter design, *Genetics* 159 (2001) 727–735.
- [55] Ron M., Heyen D.W., Weller J.I., Band M., Feldmesser E., Pasternak H., Da Y., Wiggans G.R., VanRaden P.M., Ezra E., Lewin H.A., Detection and analysis of a locus affecting milk concentration in the US and Israeli dairy cattle populations, in: *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production*, Armidale, Vol. 26, 1998, 6 WCGALP Congress Office, University of New-England, Armidale, pp. 422–425.
- [56] Schrooten C., Bovenhuis H., Coppieters W., van Arendonk J.A.M., Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle, *J. Dairy Sci.* 83 (2000) 795–806.
- [57] Schulman N.F., Moisio S.M., De Koning D.J., Elo K., Mäki-Tanila A., Vilkki J., QTL for health traits in Finnish Ayrshire cattle, in: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France, 19–23 August 2002, ISBN 2-7380-1052-0, Paper 09-19.
- [58] Sonstegard T.S., Van Tassell C.P., Ashwell M.S., Dairy cattle genomics: Tools to accelerate genetic improvement?, *J. Anim. Sci.* 79 (2001) E307–E315.
- [59] Spelman R., Garrick D., Utilisation of marker assisted selection in a commercial dairy cow population, *Livest. Prod. Sci.* 47 (1997) 139–147.
- [60] Spelman R.J., Coppieters W., Karim L., van Arendonk J.A.M., Bovenhuis H., Quantitative trait loci analysis for five milk production traits on chromosome six in the Dutch Holstein-Friesian population, *Genetics* 144 (1996) 1799–1808.
- [61] Spelman R.J., Coppieters W., Grisart B., Blott S., Georges M., Review of QTL mapping in the New Zealand and Dutch dairy cattle populations, *Proc. Assoc. Advmt. Anim. Breed. Genet.* 14 (2000) 11–16.
- [62] Velmala R., Vilkki J., Elo K., Maki-Tanila A., Casein haplotypes and their association with milk production traits in the Finnish Ayrshire cattle, *Anim. Genet.* 26 (1995) 419–425.
- [63] Velmala R.J., Vilkki H.J., Elo K.T., De Koning D.J., Maki-Tanila A.V., A search for quantitative trait loci for milk production traits on chromosome 6 in Finnish Ayrshire cattle, *Anim. Genet.* 30 (1999) 136–143.
- [64] Vignal A., Milan D., Sancristobal M., Eggen A., A review on SNP and other types of molecular markers and their use in animal genetics, *Genet. Sel. Evol.* 34 (2002) 275–305.
- [65] Vilkki H.J., De Koning D.J., Elo K.T., Velmala R., Maki-Tanila A., Multiple marker mapping of quantitative trait loci of Finnish dairy cattle by regression, *J. Dairy Sci.* 80 (1997) 198–204.
- [66] Viitala S.M., Schulman N.F., de Koning D.J., Elo K., Kinos R., Virta A., Virta J., Maki-Tanila A., Vilkki J.H., Quantitative trait loci affecting milk production traits in Finnish Ayrshire dairy cattle, *J. Dairy Sci.* 86 (2003) 1828–1836.

- [67] Weller J.I., Kashi Y., Soller M., Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle, *J. Dairy Sci.* 73 (1990) 2525–2537.
- [68] Wiener P., Maclean I., Williams J.L., Woolliams J.A., Testing for the presence of previously identified QTL for milk production traits in new populations, *Anim. Genet.* 31 (2000) 385–395.
- [69] Williams J., Woolliams J., Bovine genome analysis, Annual Report 97-98, Roslin Institute, UK, 1998, pp. 43–46.
- [70] Winter A., Kramer W., Werner F.A.O., Kollers S., Kata S., Durstewitz G., Buitkamp J., Womack J.E., Thaller G., Fries R., Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content, *Proc. Natl. Acad. Sci. USA* 99 (2002) 9300–9305.
- [71] Zhang Q., Boichard D., Hoeschele I., Ernst C., Eggen A., Murkve B., Pfistergenskow M., Witte L.A., Grignola F.E., Uimari P., Thaller G., Bishop M.D., Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree, *Genetics* 149 (1998) 1959–1973.

## APPENDIX

### Algorithmic details for the meta-analysis methodology of QTL location

The following method is effectively that of Goffinet and Gerber [26], but provides some of the algorithmic details for this meta-analysis. Suppose we have  $n$  reports of QTL on a chromosome at estimated positions  $\hat{d}_1, \hat{d}_2, \dots, \hat{d}_n$  and assume that these are estimates of the true QTL positions at  $d_1, d_2, \dots, d_n$ , although they are not necessarily all distinct QTL. Suppose further that we also have standard errors of these estimates, say  $\sigma_1 = \text{se}(\hat{d}_1)$ ,  $\sigma_2 = \text{se}(\hat{d}_2)$ ,  $\dots$ ,  $\sigma_n = \text{se}(\hat{d}_n)$ . As in Goffinet and Gerber [26], we assume that the  $\sigma_i = \text{se}(\hat{d}_i)$  are consistent estimates of the true standard errors, and will take these as assumed parameters. The appropriate number of QTL is determined by iteration, starting from a common QTL ( $n_q = 1$ ). Next, all partitions into two groups are considered ( $n_q = 2$ ), then all partitions into three groups ( $n_q = 3$ ), etc, up to  $n_q = n$  QTL. For a model with  $n_q$  distinct QTL, assume the true map positions are  $\delta_1, \delta_2, \dots, \delta_{n_q}$  with estimates  $\hat{\delta}_1, \hat{\delta}_2, \dots, \hat{\delta}_{n_q}$ . Note that if we order the published map positions from “smallest” to “largest”,  $\hat{d}_{(1)} < \hat{d}_{(2)} < \dots < \hat{d}_{(n)}$ , there are  $\frac{(n-1)!}{(n_q-1)!(n-n_q)!}$  partitions that need to be evaluated for  $n_q$  possible QTL. The procedure is as follows:

For  $n_q = 1, \dots, n$ .

For partition  $j = 1, \dots, \frac{(n-1)!}{(n_q-1)!(n-n_q)!}$ .

Calculate the weighted mean QTL position for each assigned QTL of the partition,  $q = 1, \dots, n_q$ ,  $\hat{\delta}_q = \sum_{i \in S_q} \sigma_i^{-2} \hat{d}_i / \sum_{i \in S_q} \sigma_i^{-2}$ , where  $S_q$  is the set of QTL estimates assigned to belong to QTL  $q$ .

Calculate standard error for each  $\hat{\delta}_q$ ,  $se(\hat{\delta}_q) = \left[ \sum_{i \in S_q} \sigma_i^{-2} \right]^{-1/2}$   $q = 1, \dots, n_q$ .

Calculate log-likelihood for this partitioning,  $\ln L = \sum_{i=1}^n \ln f(\hat{d}_i)$ , where  $f(\hat{d}_i)$  is the probability density function of a normal variable  $N(\hat{\delta}_{q(i)}, \sigma_i^2)$ , and  $q(i)$  is the QTL ( $q = 1, \dots, n_q$ ) that report  $i$  belongs to.

Next  $j$ .

Determine partition  $j$  with maximum log-likelihood, say  $\ln L_{n_q}$ .

Next  $n_q$ .

The optimal model can be selected by inspection of the change in log-likelihoods between successive numbers of QTL ( $n_q = 1, 2, \dots, n$ ), although in practice it is not usually necessary to evaluate beyond a few QTL on a chromosome. As indicated in Goffinet and Gerber [26], the change can be evaluated by various means, a likelihood ratio test (LRT), Akaike Information Content (AIC), or adjusted AIC (AIC\*).

### Parameter estimation for the meta-analysis methodology of QTL effects

For this model, there are two parameter to estimate, namely  $\sigma_A^2$  and  $c$ . The data consist of the (standardized) estimated QTL effects for each of the  $n$  sire reports,  $\hat{a}_i$ ,  $i = 1, \dots, n$ , with associated standard errors,  $s_i = se(\hat{a}_i)$ , taken as being known. Using the expression for the probability density function,  $h(\hat{a}_i | \hat{a}_i > c)$ ,  $\hat{a}_i > c$ , given in the main text, the log-likelihood for these parameters, given the data, is

$$\ln L(\sigma_A^2, c; \hat{\mathbf{a}}) = \sum_{i=1}^n \ln n_i(\hat{a}_i) - \sum_{i=1}^n \ln[1 - N_i(c)].$$

The maximum likelihood estimate of  $c$  is simply  $\hat{c} = \min(\hat{a}_i)$ . The log-likelihood can then be maximized numerically (e.g. using Excel's Solver) to return the estimate of  $\sigma_A^2$ . Alternatively, by setting  $\partial \ln L(\sigma_A^2, \hat{c}; \hat{\mathbf{a}}) / \partial \sigma_A^2$  to zero, we solve

$$\sum_{i=1}^n \frac{\hat{a}_i^2 - \sigma_A^2 - s_i^2}{(\sigma_A^2 + s_i^2)^2} - \hat{c} \sum_{i=1}^n \frac{n_i(\hat{c})}{(\sigma_A^2 + s_i^2)[1 - N_i(\hat{c})]} = 0.$$

An approximate standard error for  $\hat{\sigma}_A^2$  can be obtained from the observed Fisher information,  $-\partial^2 \ln L(\sigma_A^2, c; \hat{\mathbf{a}}) / \partial \sigma_A^2$  evaluated at the maximum

likelihood estimates,  $\sigma_A^2 = \hat{\sigma}_A^2$ ,  $c = \hat{c}$ , giving

$$\text{se}(\hat{\sigma}_A^2) = \left\{ \sum_{i=1}^n \frac{\hat{a}_i^2 - 1/2 (\hat{\sigma}_A^2 + \varsigma_i^2)}{(\hat{\sigma}_A^2 + \varsigma_i^2)^3} - \frac{\hat{c}}{4} \sum_{i=1}^n \frac{\hat{c}n_i(\hat{c})}{(\hat{\sigma}_A^2 + \varsigma_i^2)^2 [1 - N_i(\hat{c})]} \right. \\ \left. \times \left[ \frac{\hat{c}n_i(\hat{c})}{1 - N_i(\hat{c})} - \frac{\hat{c}^2}{\hat{\sigma}_A^2 + \varsigma_i^2} + 3 \right] \right\}^{-1/2} .$$