Evaluation of a New Fine-mapping Method Exploiting Linkage Disequilibrium: a Case Study Analysing a QTL with Major Effect on Milk Composition on Bovine Chromosome 14

JongJoo Kim* and Michel Georges
Department of Genetics, Faculty of Veterinary Medicine, University of Liege (B43)
20 Bd de Colonster, 4000-Liege, Belgium

ABSTRACT : A novel fine-mapping method exploiting linkage disequilibrium (LD) was applied to better refine the quantitative trait loci (QTL) positions for milk production traits on bovine chromosome 14 in the pedigree comprising 22 paternal half-sib families of a Black-and-White Holstein-Friesian grand-daughter design in the Netherlands for a total of 1,034 sons. The chromosome map was constructed with the 31 genetic markers spanning 90 Kosambi cM with the average inter-marker distance of 3.5 cM. The linkage analyses, in which the effects of sire QTL alleles were assumed random and the random factor of the QTL allelic effects was incorporated into the Animal Model, found the QTL for milk, fat, and protein yield and fat and protein % with the Lod scores of 10.9, 2.3, 6.0, 25.4 and 3.2, respectively. The joint analyses including LD information by use of multi-marker haplotypes highly increased the evidence of the QTL (Lod scores were 25.1, 20.9, 11.0, 85.7 and 17.4 for the corresponding traits, respectively). The joint analyses including *DGAT* markers in the defined haplotypes again increased the QTL evidence and the most likely QTL positions for the five traits coincided with the position of the *DGAT* gene, supporting the hypothesis of the direct causal involvement of the *DGAT* gene. This study strongly indicates that the exploitation of LD information will allow additional gains of power and precision in finding and localising QTL of interest in livestock species, on the condition of high marker density around the QTL region. ***(Asian-Aust. J. Anim. Sci. 2002. Vol 15, No. 9 : 1250-1256)***

Key Words : Linkage Mapping, Linkage Disequilibrium Mapping, QTL, Milk Production Traits, Dairy Cattle

INTRODUCTION

With the development of genome-wide maps based first on microsatellite markers and more recently Single Nucleotide Polymorphisms (SNPs), it has become possible to initiate the molecular dissection of complex, heritable traits. Quantitative Trait Loci (QTL) accounting for part of the genetic variation are now routinely mapped for a multitude of phenotypes in man, model organisms, and agronomically important plant and animal species (Kim and Park, 2001). However, with the experimental designs that have typically been used so far, detection power and mapping precision/accuracy are rather limited. Confidence intervals for the QTL locations are typically in the 20 to 30 cM range, which is insufficient to envisage positional cloning or effective use of the QTL information in breeding programs. There is therefore an urgent need for the development of more efficient mapping methods.

One option towards that goal is to exploit population-wide linkage disequilibrium (LD) rather than within family LD only. The use of population-wide LD has the potential to increase the detection power by extracting information from chromosomes that are not providing information in conventional family-based linkage analyses. Improvements in mapping precision/accuracy through the use of LD will result from the increased density of informative recombinational events, albeit "historical" recombinants.

It has recently been shown that LD extends over much longer distances in dairy cattle population when compared to human as a result of their unique structure (Farnir et al., 2000). It remains to be established whether this will apply to other livestock species as well. If conclusive, this opens unique opportunities to fine-map genes based on LD without the need for the development of very high density maps in these species.

The most common way to exploit LD for mapping purposes is by single marker association studies or transmission disequilibrium tests (TDT). This approach has not proven to be very effective for fine-mapping because the LD signal does not appear to decrease monotonously with increasing distance from the causal polymorphism. Multipoint mapping methods may help to circumvent some of these limitations. It has recently been noted that in the human LD decays in a stew-wise fashion probably reflecting recombinational hot-spots that delineate blocks of high LD exhibiting limited genetic variation (Daly et al., 2001; Jeffreys et al., 2001). The objective now pursued by human geneticists is therefore to generate a haplotype map of the human genome, and to identify a subset of SNPs accounting for most of the common genetic variation within

* Corresponding Author : Jong-Joo Kim. Tel: +32(4)-366-3821, Fax: +32(4) 366-4122, E-mail: JJKim@ulg.ac.be
Received January 25, 2002; Accepted April 24, 2002

blocks, and to use these in association studies. The efficacy of this approach depends to a large extent on the validity of the Common Disease/Common Variant hypothesis (Reich and Lander, 2001).

In the field of livestock genetics, Farnir et al. (2002) recently described a multipoint maximum likelihood (ML) approach for the fine-mapping of QTL in outbred half-sib pedigrees based on a LD model that was previously proposed by Terwilliger (1995) for the analysis of discrete traits. More recently, Meuwissen and Goddard (2000, 2001) proposed a very elegant and general approach based on the coalescent model and using mixed model methodology. We have recently, extended this method by including a hierarchical haplotype clustering step (Kim et al., 2002).

We herein report the results obtained with the latter method when analysing a QTL with major effect on milk yield and composition that was previously mapped to the centromeric end of bovine chromosome 14 (BTA 14) (Coppieters et al., 1998; Heyen et al., 1999; Looft et al., 2001). By applying alternative methods based on LD, we positioned this QTL within a 3cM interval (Riquet et al., 1999; Farnir et al., 2002) and subsequently identified the supposedly causal lysine to alanine substitution (K232A) in the bovine acylCoA:diacylglycerol acyltransferase gene (*DGAT*) shown to be located in the same interval (Grisart et al., 2002).

MATERIALS AND METHODS

Pedigree material and phenotypes

The pedigree material used in this study is a previously described Holstein-Friesian "grand-daughter design" (GDD; Weller et al., 1990) sampled in the Netherlands and composed of 22 paternal half-sib families for a total of 1,034 sons. The phenotypes of the sires were "daughter yield deviations" (DYD) which were obtained directly from CR-Delta (Arnhem-The Netherlands). DYDs correspond to unregressed weighted averages of the daughter's lactation performances adjusted for systematic environmental effects and breeding values of the daughter's dams and expressed as deviations from the population mean (Van Raden and Wiggans, 1991).

Marker genotypes

The sires and their sons composing the GDD were genotyped for 27 previously described chromosome 14 microsatellites and four SNPs within the bovine *DGAT* gene. The corresponding linkage map was constructed as previously described (Coppieters et al., 1998). It covers 90 cM (Kosambi) with an average distance of 3.5 cM between adjacent markers.

The marker linkage phase of the sires and sons were determined as described (Farnir et al., 2002). As a consequence, the marker data consisted of 2×22 sire chromosomes (SC), 1,034 paternally inherited chromosomes of the sons (PC), and 1,034 maternally inherited chromosomes of the sons (MC). From the PC, we can easily compute the probability that son i inherited the "left" (λ_p) or "right" ($\rho_p=1-\lambda_p$) SC from its sire at map position p as described (Coppieters et al., 1998).

QTL fine-mapping exploiting both linkage and LD

The utilised mapping method has been described in detail in Kim et al. (2002), and can be summarised as follows. To test for the presence of a QTL at map position *p* of the studied chromosome:

1. We compute identity-by-descent (IBD) probabilities (ϕ_p) for all pair wise combinations of SC and MC using the method described by Meuwissen and Goddard (2001). This method approximates the probability that two chromosomes are IBD at a given map position conditional on the identity-by-state (IBS) status of flanking markers, on the basis of coalescent theory (Hudson, 1985). A maximum of eight flanking markers was considered to compute ϕ_p
2. Using $(1-\phi_p)$ as a distance measure, we apply a hierarchical clustering algorithm such as UPGMA (Mount, 2001) to generate a dendrogram representing the genetic relationship-at position *p*-between all SC and MC haplotypes encountered in the population.
3. We use the logical framework provided by this dendrogram to group the SC and MC in functionally distinct clusters. In this work, the clusters were defined such that all haplotypes within a cluster had a distance measure $(1-\phi_p)<T$ (Kim et al., 2002).
4. We model the sons' phenotypes (DYDs) using the following linear model:

$$y=Xb+Z_hh+Z_uu+e$$

y is the vector of phenotype records of all sons, b is a vector of fixed effects which in this study reduces to the overall mean. X is an incidence matrix relating fixed effects to individual sons, which in this study reduces to a vector of ones. h is the vector of random QTL effects corresponding to the defined haplotype clusters. Z_h is an incidence matrix relating haplotype clusters to individual sons. In Z_h, a maximum of three elements per line can have non-zero value: "1" in the column corresponding to the cluster to which the MC haplotype belongs, "λ_p" and "ρ_p" in the columns corresponding respectively to the haplotype clusters of the "right" and "left" SC. If either of the SC and/or MC belong to the same cluster, the corresponding coefficients are added.

Haplotype cluster effects with corresponding variance, σ^2_H, individual polygenic effects with corresponding variance, σ^2_A, and individual error terms with corresponding variance, σ^2_E, were estimated using AIREML program (developed according to the protocol of Johnson and Thompson (1995)), by maximising the log restricted likelihood function L:

$$L = -.5\log|\mathbf{V}| - .5\log|\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}| - .5(\mathbf{y}-\mathbf{X}\hat{\mathbf{b}})^T\mathbf{V}^{-1}(\mathbf{y}-\mathbf{X}\hat{\mathbf{b}})$$

In this, **V** equals:

$$\mathbf{V} = \sigma^2_H \mathbf{Z}_h \mathbf{H} \mathbf{Z}_h^T + \sigma^2_A \mathbf{Z}_u \mathbf{A} \mathbf{Z}_u^T +$$

Because we assume that the covariance between the QTL effects of the different haplotype clusters is zero, H reduces to an identity matrix. This differentiates our approach from that of Meuwissen and Goddard (2000), in which H is the matrix of IBD probabilities between haplotypes. A is the additive genetic relationship matrix.

5. Steps 3 and 4 are repeated for all possible values of T (from 1 to 0 by 0.01 decrement unit), in order to identify a restricted maximum likelihood (REML) solution for map position p. By analogy with Farnir et al. (2002) we will denote the hypothesis corresponding to this REML solution as H_2.

QTL mapping exploiting linkage only

Note that the previous model can be extended with minor modifications to map QTL by exploiting linkage information only. This is simply achieved by ignoring all MCs and considering that all SCs belong to distinct haplotype clusters, irrespective of their marker genotype. REML solutions for the different parameters can be found as described in the previous section. Again by analogy with Farnir et al. (2002), we refer to the corresponding hypothesis as H_1.

Hypothesis testing and significance thresholds

The likelihood of the data under the H_2 and H_1 hypotheses are compared with that under the null hypothesis, H_0, of no QTL at map position p. The latter is computed as described above but using the reduced model:

$$y=Xb+Z_u u+e$$

Evidence in favour of a QTL at map position, p, can then be expressed as a lod score:

$$z_p = 0.43\times(ML_{H1/2} - ML_{H0})$$

As customary when performing interval mapping, the hypothetical position of the QTL is slided throughout the chromosome map, and lod scores are computed at each map position as described to generate chromosome-wide lod score profiles.

Kim et al. (2002) have shown by simulation that when performing a whole genome scan (29 Morgan) with a marker density of one marker every 5cM, $2\times ln(10)\times z_p$ has (under the null hypothesis) a chi-squared distribution with 2 degrees of freedom corrected (Bonferroni correction) for 58 and 174 independent traits when testing respectively H_1 and H_2. Experiment-wide significance levels were computed from these distributions in this study.

RESULTS

Including LD information increases the evidence in favour of a QTL with major effect on milk yield and composition at the centromeric end of bovine chromosome 14

Figure 1 and Table 1 report the location scores that were obtained when searching QTL influencing milk yield and composition on bovine chromosome 14, using linkage information only or linkage plus linkage disequilibrium information. As expected from the results of our previous analyses (Coppieters et al., 1998), linkage information alone yielded highly significant, experiment wide evidence for the presence of a QTL at the proximal end of BTA14 for all traits except fat yield. The most likely position for the QTL corresponds to the first marker interval (between *BULGE9* [0.0 cM] and *BULGE11* [1.0 cM]) for all of the traits. Exploiting linkage disequilibrium (the joint linkage and LD analyses excluding the four *DGAT* SNPs in the map) information considerably increased the experiment-wide significance levels for all traits (Table 1). This indicates that this novel mapping approach has the potential to increase the power to detect QTL without the need to increase the sample size. As in the previous analysis, the most likely position of the QTL corresponded to the first marker interval for fat yield and fat % yielding the highest lod scores, while being in the second or even fifth marker interval for the three remaining traits (milk and protein yield and protein %) (Table 1). It is not obvious therefore, that inclusion of linkage disequilibrium information has dramatically increased the mapping precision in this instance.

Including *DGAT* SNPs supports the hypothesis that this gene causes the QTL effect

Figure 1 and Table 1 also report the results obtained when including four previously described *DGAT* SNPs in the analysis. The evidence supporting a QTL on proximal BTA14 dramatically increased again for all analysed traits.

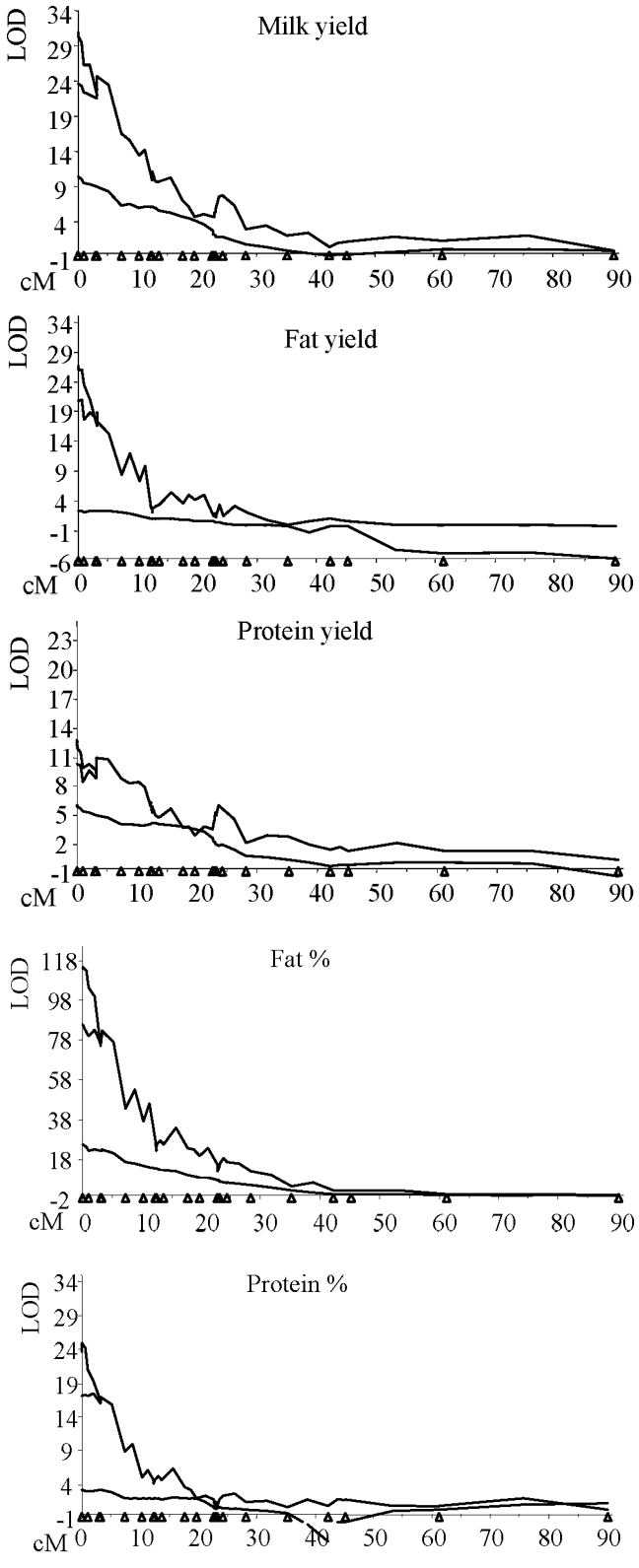


Figure 1. Lod profiles of the QTL influencing milk yield and compositions on bovine chromosome14. The upper and middle lines of the Lod profiles of each trait were generated by the joint linkage and linkage disequilibrium (LD) analyses after including and excluding the four DGAT SNP markers of the gene, acylCoA:diacylglycerol acyltransferase, respectively and the lower line by the linkage analyses only. Open triangles below X-axis indicate marker positions.

compared to the joint linkage and LD analyses without the four *DGAT* SNPs. It is also noteworthy that the most likely position of the QTL now coincided with the position of the *DGAT* gene for all of the traits except protein %. These observations support the hypothesis of the direct causal involvement of the *DGAT* gene.

Table 2 reports for each trait the estimates of the three variance components as obtained at the corresponding most likely map position. These estimates are in reasonable agreement with previous reports (Farnir et al., 2002; Grisart et al., 2002), bearing in mind that both sets of estimates were obtained using very different methodology and do not measure exactly the same parameters.

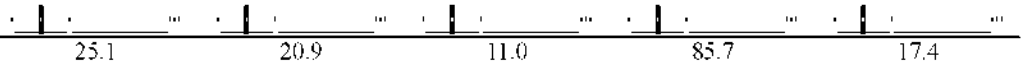
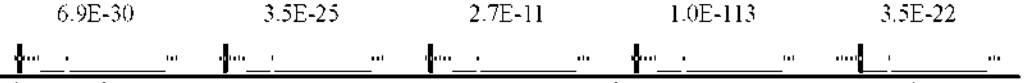
Figure 2A shows the haplotype dendrogram obtained when including the *DGAT* SNPs in the defined haplotypes at the mid-point (0.005 cM) of the first marker interval, as well as the clustering (T=0.99) that yielded two haplotype groups and the highest lod score (115 Lods) for milk fat percentage. It can be seen that this optimal clustering partitioned the tree into the two haplotype groups corresponding to the perfect segregation of the two alleles of *DGAT* SNP defining the *K232A* mutation, previously suggested to be the causal mutation (Grisart et al., 2002).

Figure 2B shows the dendrogram obtained at the location (0.01 cM) of the first marker interval when excluding the *DGAT* SNPs from the definition of haplotype unit, with the clustering (T=0.20) that yielded 77 haplotype groups and the highest lod score (85.7 Lods) for milk fat percentage. Obviously, this clustering produced redundant haplotype groups assuming that *K232A* is the causal mutation. However, it can be seen that the majority of clusters (haplotype groups) are generally homogeneous with respect to *K232A* allele, as expected if this were the causal mutation.

DISCUSSION

In this paper, we evaluated a novel fine-mapping method that exploited linkage disequilibrium using the data for the QTL with major effect on milk yield and composition that was previously mapped on proximal bovine chromosome 14 (Coppieters et al., 1998). The proposed method has been described in detail in Kim et al. (2002). It is derived from an approach that was recently developed by Meuwissen and Goddard (2000, 2001), to which it adds a hierarchical haplotype clustering step. In our hands, this haplotype clustering has effectively solved numerical computation problems that were encountered when implementing the original method. We also believe that the hierarchical clustering-in addition to being a very intuitive way of representing the kinship of individual chromosomes-opens a number of new avenues for research. It offers, for instance, a way to select functionally distinct

Table 1. The most likely (ML) positions, Lod scores, experimental-wide significance p-values for the detected QTL influencing milk yield and compositions on bovine chromosome 14 by the linkage (L) mapping and the joint linkage and linkage disequilibrium (L+LD) mapping analyses

Mapping[a]	Trait	Milk yield (kg)	Fat yield (kg)	Protein yield (kg)	Fat %	Protein %
L	Lod score[b]	10.9	2.3	6.0	25.4	3.2
	p-value	8E-10	1.8E-1	5.4E-5	2.4E-24	3.3E-2
	ML position[c]					
L+LD (-DGAT)	Lod score[b]	25.1	20.9	11.0	85.7	17.4
	p-value	1.7E-22	2.9E-19	2.7E-9	3.5E-84	8.7E-16
	ML position[c]					
L+LD (+DGAT)	Lod score[b]	31.4	26.7	12.8	115.0	24.9
	p-value	6.9E-30	3.5E-25	2.7E-11	1.0E-113	3.5E-22
	ML position[c]					

[a] The L−LD analyses were performed after including (+DGAT) or excluding (-DGAT) the four DGAT genetic markers of the gene, acylCoA:diacylglycerol acyltransferase.

[b] The highest Lod scores were obtained at the ML positions.

[c] Thin arrows in order indicate marker positions [cM] of *BULGe9* [0.0]-(*DGAT1* [0.01]-*DGAT2* [0.02]-*DGAT3* [0.03]-*DGAT4* [0.04] for −DGAT L−LD analyses)-*BULGe11* [1.0]-*BULGe13* [3.0]-*ILSTS39* [3.1]-*BULGe50* [3.2], respectively. The arrows with black background indicate the most likely QTL positions.

Table 2. The magnitudes of the detected QTL at the most likely positions for milk yield and compositions on bovine chromosome 14 by the linkage (L) mapping and the joint linkage and linkage disequilibrium (L+LD) mapping analyses

Mapping		Milk yield (kg)	Fat yield (kg)	Protein yield (kg)	Fat %	Protein %
L	σ^2_H [a]	7,965	2.7	3.8	52.6	2.7
	σ^2_A [b]	49,930	74.3	42.3	246.8	41.3
	σ^2_E [c]	14,350	16.2	8.6	16.3	6.6
	r^2_{QTL} [d]	0.20	0.06	0.13	0.29	0.10
L+LD (-DGAT)	σ^2_H [a]	4,326	9.2	1.7	62.8	1.6
	σ^2_A [b]	46,540	67.3	42.3	144.2	42.4
	σ^2_E [c]	14030	13.3	8.3	33.8	3.7
	r^2_{QTL} [d]	0.13	0.19	0.06	0.41	0.07
L+LD (+DGAT)	σ^2_H [a]	34,700	8.0	12.6	57.4	5.4
	σ^2_A [b]	50,660	60.8	42.3	116.3	42.1
	σ^2_E [c]	10,850	17.5	8.3	42.5	3.4
	r^2_{QTL} [d]	0.25	0.17	0.12	0.42	0.19

[a,b,c] Three variance components were estimated by the REML (restricted maximum likelihood) analyses when fitting a putative QTL: σ^2_H (QTL allelic variance), σ^2_A (polygenic variance), σ^2_E (residual variance).

[d] Proportion of the trait variance explained by the QTL ($2\times\sigma^2_H/(2\times\sigma^2_H+\sigma^2_A-\sigma^2_E)$).

chromosomes for detailed molecular analysis, or to measure the extent of genetic variation across chromosomes in the population of interest thereby potentially revealing signatures of selection effects.

We demonstrate that the proposed method extracts more information from the available data than conventional linkage analysis, thus having the potential to enhance the power to detect QTL. This gain in power is, however, expected to be a function of marker density around QTL location, which can be supported by the results of the fat % QTL. Even if the two QTL detected from linkage plus LD mapping after including and excluding *DGAT* gene from haplotype unit, were localized at (close to) the same position (0.005 cM and 0.01 cM, respectively), different magnitude of the two corresponding Lod scores is substantial (Table 1). Thus to exploit linkage disequilibrium efficiently, the proposed method is very likely to require a higher marker density when compared to linkage analysis, and therefore a higher genotyping load per individual. Cost benefit analyses of experimental designs combining different marker densities and sample size are being conducted (Kim et al., 2002).

The proposed method is likely to result in an increase in mapping accuracy as well. The gain in mapping accuracy is being evaluated more quantitatively using simulated data (Kim et al., 2002). The marked increase in significance that was obtained in this analysis when including SNPs located within *DGAT*, a gene with strong candidacy, is in support of the possible causal involvement of the corresponding gene and even polymorphisms as previously suggested (Grisart et al., 2002).

ACKNOWLEDGEMENTS

This work was funded by a research grants from Vialactia Biosciences (Auckland, New Zealand), Cr-Delta, Livestock Improvement Corporation, the Vlaamse Rundvee Vereniging, the Belgian Ministry of Agriculture and the

Figure 2. Phylogeny trees among fat % QTL alleles specific to unique haplotypes (N=133 for *A* and 359 for *B*) including (*A*) or excluding (*B*) the four *DGAT* SNPs in the first marker interval of bovine chromosome 14 by the joint linkage and LD analyses. The black or white square in the first column that is assigned to each haplotype (terminal node) in the built tree represents alternative *DGAT* SNP alleles defining the *K232A* mutation. The black or white block with the length equal to the number of haplotypes in the second column represents a defined haplotype group (total two haplotype groups for *A*, seventy-seven for *B*).

European Union. We are very grateful to Wouter Coppieters and Fred. Farnir for their major contributions to this work. Our gratitude is extended to Theo Meuwissen, Richard Spelman and Dave Johnson for providing us with the AIREML program and helpful comments.

REFERENCES

Coppieters, W., J. Riquet, J. Arranz, P. Berzi, N. Cambisano, B. Grisart, L. Karim, F. Marcq, L. Moreau, C. Nezer, P. Simon, P. Vanmanshoven, D. Wagenaar and M. Georges, 1998. A QTL with major effect on milk yield and composition maps to

bovine chromosome 14. Mamm. Genome 9:540-544.

Daly, M. J., J. D. Rioux, S. F. Schaffner, T. J. Hudson and E. S. Lander. 2001. High-resolution haplotype structure in the human genome. Nature Genet. 29:229-232.

Farnir, F., W. Coppieters, J. Arranz, P. Berzi, N. Cambisano, B. Grisart, L. Karim, F. Marcq, L. Moreau, M. Mni, C. Nezer, P. Simon, P. Vanmanshoven, D. Wagenaar and M. Georges. 2000. Extensive genome-wide linkage disequilibrium in cattle. Genome Res. 10:220-227.

Farnir, F., B. Grisart, W. Coppieters, J. Riquet, P. Berzi, N. Cambisano, L. Karim, M. Mni, S. Moisio, P. Simon, D. Wagenaar, J. Vilkki and M. Georges. 2002. Simultaneous mining of linkage and linkage disequilibrium to fine-map QTL in outbred half-sib pedigrees: revisiting the location of a QTL with major effect on milk production on bovine chromosome 14. Genetics 161:275-287.

Grisart, B., W. Coppieters, F. Farnir, L. Karim, C. Ford, N. Cambisano, M. Mni, S. Reid, R. Spelman, M. Georges and R. Snell. 2002. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT gene with major effect on milk yield and composition. Genome Res. 12:222-231.

Heyen, D. W., J. L. Weller, M. Ron, M. Band, J. E. Beever, E. Feldmesser, Y. Da, G. R. Wiggans, P. M. VanRaden and H. A. Lewin. 1999. A genome scan for QTL influencing milk production and health traits in dairy cattle. Physiol. Genomics 1:165-175.

Hudson, R. R. 1985. The sampling distribution of linkage disequilibrium under an infinite alleles model without selection. Genetics 109:611-631.

Jeffreys, A. J., L. Kauppi and R. Neumann. 2001. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. Nature Genet. 29:217-222.

Johnson, D. L. and R. Thompson. 1995. Restricted maximum likelihood estimation of variance components for univariate Animal Models using sparse matrix techniques and average information. J. Dairy. Sci. 78:449-456.

Kim, J. J. and Y. I. Park. 2001. Current status of quantitative trait locus mapping in livestock species. Asian-Aust. J. Anim. Sci. 14(4):587-596.

Kim, J. J., F. Farnir, W. Coppieters and M. Georges. 2002. A novel joint linkage and linkage disequilibrium mapping enhanced QTL detection power and mapping precision in outbred half-sib pedigrees: I. methodology and simulation. (In preparation).

Looft, C., N. Reinsch, C. Karall-Albrecht, S. Paul, M. Brink, H. Thomsen, G. Brockmann, C. Kuhn, M. Schwerin and E. Kalm. 2001. A mammary gland EST showing linkage disequilibrium to a milk production QTL on bovine Chromosome 14. Mamm. Genome 12:646-650.

Mount, D. W. 2001. Bioinformatics: Sequence and Genome analysis. Cold Spring Harbor Laboratory Press, New York, New York.

Meuwissen, T. H. E. and M. E. Goddard. 2000. Fine mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci. Genetics 155:421-430.

Meuwissen, T. H. E. and M. E. Goddard. 2001. Prediction of identity by descent probabilities from marker-haplotypes. Genet. Sel. Evol. 33:605-634.

Reich, D. E. and E. S. Lander. 2001. On the allelic spectrum of human disease. Trends In Genetics 17:502-510.

Riquet, J., W. Coppieters, N. Cambisano, J. Arranz, P. Berzi, S. K. Davis, B. Grisart, F. Farnir, L. Karim, M. Mni, P. Simon, J. F. Taylor, P. Vanmanshoven, D. Wagenaar, J. E. Womack and M. Georges. 1999. 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 96:9252-9257.

Terwilliger, J. D. 1995. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. Am. J. Hum. Genet. 56:777-787.

Van Raden, P. M. and G. R. Wiggens. 1991. Derivation calculating and use of National Animal Model Information. J. Dairy Sci. 74:2737-2746.

Weller, J. I., Y. Kashi and M. Soller. 1990. Power of daughter and granddaughter designs for genetic mapping of quantitative traits in dairy cattle using genetic markers. J. Dairy Sci. 73:2525-2537.