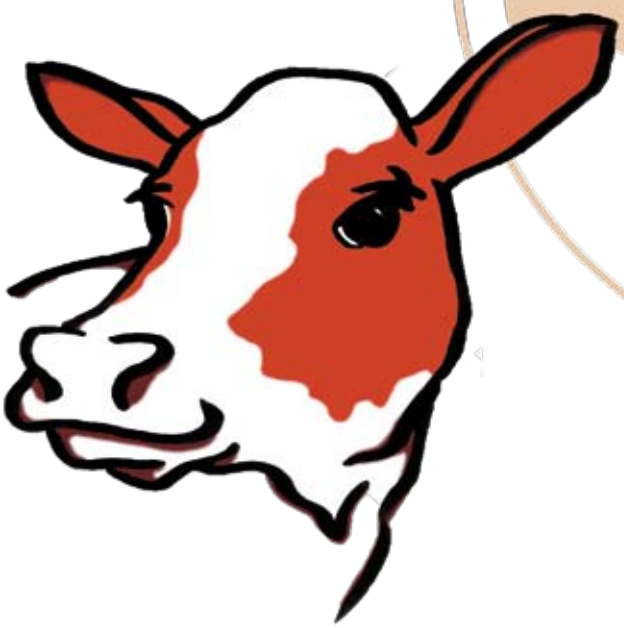
Agrifood Research Reports 109

MTT
Agrifood Research Finland

Quantitative trait loci detection and benefits from marker-assisted selection in dairy cattle

Doctoral Dissertation

Nina Schulman



Animal Production

Agrifood Research Reports 109
60 pages, 5 appendices

Quantitative trait loci detection and benefits from marker-assisted selection in dairy cattle

Doctoral Dissertation

Nina Schulman

Academic Dissertation
To be presented, with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, for public criticism in Auditorium 1041,Biocenter II, Viikinkaari 5, on December 7th 2007, at 12 o'clock.

MTT Agrifood Research Finland

Supervisor:	Dr. Johanna Vilkki
	MTT Agrifood Research Finland
Reviewers:	Professor Theodorus Meuwissen
	Norwegian University of Life Sciences
	Dr. Pekka Uimari
	Jurilab Ltd, Kuopio, Finland
Opponent:	Prorector Lena Andersson-Eklund
	Swedish University of Agricultural Sciences
Custos:	Professor Matti Ojala
	University of Helsinki, Finland

ISBN 978-952-487-131-0 (Printed version)
ISBN 978-952-487-132-7 (Electronic version)
ISSN 1458-5073 (Printed version)
ISSN 1458-5081 (Electronic version)
http://www.mtt.fi/met/pdf/met109.pdf

Distribution and sale
MTT Agrifood Research Finland, Information Management
FI-31600 Jokioinen, Finland, phone + 358 3 4188 2327, e-mail julkaisut@mtt.fi
Printing year
2007
Cover picture
Petri Honkamaa
Printing house
Tampereen Yliopistopaino - Juvenes Print Oy

Quantitative trait loci detection and benefits from marker-assisted selection in dairy cattle

Nina Schulman

MTT Agrifood Research Finland, Department of Biotechnology and Food Research, FIN-31600 Jokioinen, Finland, nina.schulman@mtt.fi

Abstract

Conventional breeding schemes for dairy cattle are based on phenotypic information obtained from individuals and/or their relatives and progeny testing of the young bull candidates. The genetic model used in the evaluation process of the animals does not assume the underlying genes of the quantitative traits to be known. Knowing the chromosomal areas or actual genes affecting the traits would add more information to be used in the selection decisions which would potentially lead to higher genetic response.

The first objective of this study was to map quantitative trait loci (QTL) affecting economically important traits: milk production traits, health traits and fertility traits in the Finnish Ayrshire population. The second objective was to investigate the effects of using QTL information in marker-assisted selection (MAS) on the genetic response and the linkage disequilibrium between the different parts of the genome.

Whole genome scans were carried out on a grand-daughter design with 12 half-sib families and a total of 493 sons. Twelve different traits were studied: milk yield, protein yield, protein content, fat yield, fat content, somatic cell score (SCS), mastitis treatments, other veterinary treatments, days open, fertility treatments, non-return rate, and calf mortality. A total of 150 markers were used in all other studies except for fertility traits where 171 markers were used. The average spacing of the markers was 20 cM with 2 to 14 markers per chromosome. Associations between markers and traits were analyzed with multiple marker regression. Chromosomes were analyzed separately and by using QTL on other chromosomes as cofactors. Significance was determined by permutation and genome-wise *P*-values obtained by Bonferroni correction. The benefits from MAS were investigated by simulation: a conventional progeny testing scheme was compared to a scheme where QTL information was used within families to select among full-sibs in the male path. Two QTL on different chromosomes were modelled. The effects of different starting frequencies of the favourable alleles and different size of the QTL effects were evaluated.

In the whole genome scans of milk, health and fertility traits of Finnish Ayrshire a large number of QTL, 48 in total, were detected at 5% or higher chromosome-

wise significance when the chromosomes were analyzed separately. Milk production QTL were found on 8 chromosomes. There are some interesting yield QTL, for example the QTL affecting fat yield on BTA14 which probably is the DGAT1 gene, and the QTL affecting fat yield on BTA12 and protein yield on BTA5, 12, 25. Quantitative trait loci for SCS were found on BTA3, 11, 14, 18, 27, and 29, for mastitis treatments on BTA18 and for other veterinary treatments on BTA2, 14, 16, 22, and 23. Quantitative trait loci for days open were found on BTA1, 2, 5, 12, 20, 25, and 29, for fertility treatments on BTA1, 5, 10, 14, 15, 19, and 25, for calf mortality on BTA4, 6, 11, 15, 18, and 23 and for non-return rate on BTA10 and 14. The use of cofactors revealed a total of 31 possible QTL for milk production traits and 17 for health traits many of which are likely to be false positives however.

In the simulation study the total genetic response was faster with MAS than with conventional selection and the advantage of MAS persisted over the studied generations. The rate of response and the difference between the selection schemes reflected clearly the changes in allele frequencies of the favourable QTL. The disequilibrium between the polygenes and QTL was always negative and it was larger with larger QTL size. With lower initial allele frequencies the disequilibrium was slightly higher with MAS but with higher initial frequencies it was lower. When selection was continued for four generations, the MAS scheme resulted first in more negative disequilibrium but the disequilibrium decreased slightly faster with MAS than with conventional selection. The disequilibrium between the two QTL was larger with QTL of large effect and it was somewhat larger with MAS for scenarios with starting frequencies below 0.5 for QTL of moderate size and below 0.3 for large QTL. When selection was continued for four generations, the MAS scheme resulted first in more negative values than the conventional scheme but later in less negative values until close to fixation of the favourable allele when the disequilibrium was close to zero in both schemes.

In conclusion, several QTL affecting economically important traits of dairy cattle were detected. Further studies are needed to verify these QTL, check their presence in the present breeding population, look for pleiotropy and fine map the most interesting QTL regions. The results of the simulation studies show that using MAS together with embryo transfer to pre-select young bulls within families is a useful approach to increase the genetic merit of the AI-bulls compared to conventional selection.

Key words: dairy cattle, QTL, genome scan, milk, mastitis, fertility, MAS

Geenikartoitus ja markkeriavusteinen valinta nautakarjalla

Nina Schulman

MTT (Maa- ja elintarviketalouden tutkimuskeskus), Biotekniikka ja elintarviketutkimus, Eläingenomiikka, 31600 Jokioinen, nina.schulman@mtt.fi

Tiivistelmä

Perinteiset lypsykarjan jalostusohjelmat perustuvat eläimeltä ja /tai sen sukulaisilta saataviin fenotyyppitietoihin ja nuorten sonnien jälkeläisarvosteluun. Arvostelussa käytetty geneettinen malli olettaa kvantitatiivisiin ominaisuuksiin vaikuttavien geenien olevan tuntemattomia. Ominaisuuksiin vaikuttavien geenialueiden tai geenien tunteminen lisäisi informaatiota valintapäätösten tueksi, mikä saattaisi lisätä geneettistä edistymistä.

Tämän tutkimuksen tarkoituksena oli ensinnäkin kartoittaa geenialueita (QTL), jotka vaikuttavat taloudellisesti tärkeisiin ominaisuuksiin: maidontuotanto-ominaisuuksiin, terveysominaisuuksiin ja hedelmällisyysominaisuuksiin suomalaisessa Ayrshire populaatiossa. Toiseksi tarkoituksena oli tutkia QTL-informaatiota käyttävän markkeriavusteisen valinnan (MAS) vaikutusta geneettiseen edistymiseen ja eri genominosien väliseen kytkentäepätasapainoon.

Koko genomin kartoitus tehtiin pojantytärmallilla, jossa oli 12 puolisisarperhettä, joissa oli yhteensä 493 poikaa. Tutkittuja ominaisuuksia oli 12: maitotuotos, valkuaistuotos, valkuaisprosentti, rasvatuotos, rasvaprosentti, somaattinen soluluku, utaretulehdushoidot, muiden sairauksien hoidot, tyhjäkausi, hedelmällisyyshoidot, uusimattomuusprosentti ja vasikkakuolleisuus. Markkereita tyypitettiin yhteeensä 150 paitsi hedelmällisyyden kartoituksessa, missä markkereita tyypitettiin 171. Markkereiden välinen keskimääräinen etäisyys oli 20 cM ja niitä oli kahdesta neljääntoista kromosomia kohden. Markkereiden ja tutkittavien ominaisuuksien välinen yhteys analysoitiin usean markkerin regressiomenetelmällä. Kromosomit analysoitiin erikseen sekä käyttämällä muista kromosomeista löydettyjä QTL:iä kofaktoreina. Tilastollinen merkitsevyys määritettiin permutaatiolla ja genomikohtaiset P-arvot Bonferronikorjauksella. Markkeriavusteisen valinnan hyötyä tutkittiin simulaation avulla, missä perinteistä jälkeläisarvostelumallia verrattiin malliin, jossa QTL-informaatiota käytettiin täysveljien välisessä valinnassa perheiden sisällä. Simulaation geneettisessä mallissa ominaisuuteen vaikutti polygeenien lisäksi kaksi QTL:ää, jotka sijaitsivat eri kromosomeissa. Eri kokoa olevien QTL-vaikutusten ja edullisien alleelien alkufrekvenssien vaikutusta tutkittiin.

Suomen Ayrshiren koko genomin kartoituksessa löydettiin useita, yhteensä 48, maito-, terveys- ja hedelmällisyysominaisuuksiin vaikuttavia geenialueita 5%:n kromosomikohtaisella merkitsevyystasolla, kun kromosomit analysoitiin erikseen. Maidontuotantoon vaikuttavia QTL:iä löydettiin yhteensä neljätoista. Muutamia mielenkiintoisia valkuais- ja rasvatuotokseen vaikuttavia geenialueita löytyi. Näitä ovat esimerkiksi rasvatuotokseen vaikuttava QTL kro-

mosomissa 14, joka mahdollisesti on sama kuin Holstein rodusta aiemmin löydetty DGAT1 geeni, rasvatuotokseen vaikuttava QTL kromosomissa 12 sekä valkuaistuotokseen vaikuttavat QTL:t kromosomeissa 5, 12, 25. Somaattiseen solulukuun vaikuttavia QTL:iä löydettiin kromosomeista 3, 11, 14, 18, 27 ja 29. Utaretulehdukseen vaikuttavia QTL:iä löydettiin kromosomista 18 ja muiden sairauksien hoitoihin vaikuttavia QTL:iä kromosomeista 2, 14, 16, 22 ja 23. Tyhjäkauteen vaikuttavia geenialueita löytyi kromosomeista 1, 2, 5, 12, 20, 25 ja 29, hedelmällisyyshoitoihin vaikuttavia geenialueita kromosomeista 1, 5, 10, 14, 15, 19 ja 25, vasikkakuolleisuuteen vaikuttavia alueita kromosomeista 4, 6, 11, 15, 18 ja 23 ja uusimattomuusprosenttiin vaikuttavia alueita kromosomeista10 and 14. Kofaktorianalyysissä maidontuotannon geenialueita löytyi yhteensä 31 ja terveyteen vaikuttavia geenialueita 17, joista useat todennäköisesti kuitenkin ovat vääriä positiivisia tuloksia.

Markkeriavusteisen valinnan simulaatiotutkimuksessa havaittiin, että geneettinen kokonaisedistyminen (polygeeninen edistyminen + QTL-edistyminen) oli nopeampaa MAS:lla kuin perinteisellä valinnalla ja MAS:n hyöty kesti tutkittujensukupolvien ajan. Muutokset hyödyllisten alleelien frekvensseissä vaikuttivat selvästi geneettisen edistymisen nopeuteen ja eroihin valintamenetelmien välillä. Polygeenien ja QTL:ien välinen kytkentäepätasapaino oli aina negatiivinen ja suurempi, kun QTL-vaikutus oli suurempi. Kun hyödyllisien alleelien alkufrekvenssi oli pieni, kytkentäepätasapaino oli hiukan suurempi MAS:lla mutta suuremmilla alkufrekvensseillä pienempi. Kun valintaa jatkettiin neljä sukupolvea, kytkentäepätasapaino oli aluksi MAS:lla negatiivisempaa, mutta väheni sitten nopeammin kuin tavanomaisella valinnalla. QTL:ien välinen kytkentäepätasapaino oli suurempi, kun QTL-vaikutus oli suurempi. Se oli hiukan suurempi MAS:lla, kun alkufrekvenssit olivat alle 0.5 keskikokoisilla QTL:llä ja alle 0.3 suurilla QTL:llä. Kun valintaa jatkettiin neljä sukupolvea, MAS aiheutti ensin negatiivisempaa kytkentäepätasapainoa, mutta myöhemmin vähemmän negatiivisia arvoja kuin perinteinen valinta. Kun hyödylliset alleelit olivat lähes fiksoituneet, kytkentäepätasapaino oli molemmilla valintamenetelmillä lähellä nollaa.

Yhteenvetona: tutkimuksessa löydettiin useita lypsykarjan taloudellisesti merkittäviin ominaisuuksiin vaikuttavia geenialueita. Lisää tutkimuksia tarvitaan näiden QTL:ien varmistamiseksi, segregoitumisen kartoittamiseksi nykyisessä nautapopulaatiossa, pleiotrooppisten vaikutusten määrittämiseksi ja mielenkiintoisten alueiden hienokartoittamiseksi. Simulaatiotutkimuksen tulokset osoittavat, että MAS yhdistettynä alkionsiirtoon, jolloin nuoria sonneja voidaan esivalita perheiden sisällä ja näistä parhaat jälkeläisarvostella, on hyvä, geneettistä edistymistä lisäävä vaihtoehto tavanomaiselle jalostusvalinnalle.

Avainsanat: nauta, lypsykarja, QTL, geenikartoitus, utaretulehdus, hedelmällisyys, markkeriavusteinen valinta

List of original articles

The thesis is a summary and discussion of the following articles, which are referred to by their Roman numerals:

I Viitala, S. M., Schulman, N.F., de Koning, D.J., Elo, K., Kinos, R., Virta, A., Virta, J., Mäki-Tanila, A. and Vilkki, J.H. 2003. Quantitative trait loci affecting milk production traits in Finnish Ayrshire dairy cattle. Journal of Dairy Science 86:1828-1836.

II Schulman, N. F., Viitala, S.M., de Koning, D.J., Virta, J., Mäki-Tanila, A. and Vilkki J. H. 2004. Quantitative trait loci for health traits in Finnish Ayrshire cattle. Journal of Dairy Science 87:443-449.

III Schulman, N. F., Sahana, G., Lund, M. S., Viitala, S. M. and Vilkki, J. H. 2007. Quantitative trait loci for fertility traits in Finnish Ayrshire cattle. Genetics Selection and Evolution. Accepted.

IV Schulman, N. F., de Vries, M.J. and Dentine, M.R. 1999. Linkage disequilibrium in two-stage marker-assisted selection. Journal of Animal Breeding and Genetics 116:99-110.

V Schulman, N. F. and Dentine, M. R. 2005. Linkage disequilibrium and selection response in two-stage marker-assisted selection of dairy cattle over several generations. Journal of Animal Breeding and Genetics 122:110-116.

Contribution of the author

I The author carried out the data editing, preliminary statistical analyses, the main part of the statistical analyses and participated in interpreting the results.

II The author carried out the data editing, did the statistical analyses, participated in interpreting the results and was the main author of the paper.

III The author carried out the data editing, did the whole genome scan, did part of the single trait variance component analyses, participated in interpreting the results and was the main author of the paper.

IV, V The author wrote the simulation program, run the program, participated in interpreting the results and was the main author of the papers.

Symbols and abbreviations

AI	artificial insemination
BAC	bacterial artificial chromosome
BLUP	best linear unbiased prediction
BTA	Bos taurus chromosome
cM	centi Morgan
DGAT1	diacyglycerol acyltransferase
DNA	deoxyribonucleic acid
EBV	estimated breeding value
IBD	identity by descent
LD	linkage disequilibrium
LE	linkage equilibrium
LRT	likelihood ratio test
MA-BLUP	marker-assisted best linear unbiased prediction
MAS	marker-assisted selection
MVN	multivariate normally distributed
PCR	polymerase chain reaction
PIC	polymorphic information content
QTL	quantitative trait loci
RFLP	restriction fragment length polymorphism
RH	radiation hybrid
SCC	somatic cell count (cells/ml)
SCS	somatic cell score (SCC in logarithmic scale)
SD	standard deviation
SNP	single nucleotide polymorphism
YAC	yeast artificial chromosome

Contents

1 Introduction

1.1 Mapping of Quantitative trait loci (QTL) in dairy cattle

1.1.1 Quantitative trait loci (QTL)

Most of the economically important traits in dairy cattle are quantitative. This means that they show continuous variation and are affected by many genes and the environment (e.g., Falconer and Mackay, 1996). Traditional animal breeding has assumed an infinitesimal genetic model where a large number of genes with small effect cause the trait variation. In this model the number of the genes, their effects, locations and allele frequencies are not known. In reality the number of genes is smaller than infinite, few of the genes may have large and moderate effects and most have usually small effects (Hayes and Goddard, 2001). The genes may interact with each other such that the genotype at one locus affects the outcome on the phenotype of the genotype at another locus (Falconer and Mackay, 1996). Hayes and Goddard (2001) predicted that 17% of the genes of largest effect explained 90 % of the genetic variation in dairy cattle. This however depends on the complexity of the trait and a more even distribution of the genes is also possible.

Quantitative trait loci (QTL) (Geldermann, 1975) are polymorphic loci which contain alleles that differentially affect the expression of a quantitative trait (brc.mcw.edu/Crossmap/term.html). They are not necessarily genes but located close to genes that affect the trait. QTL affecting a specific trait are usually found on many different chromosomes (http://encyclopedia.thefreedictionary.com/QTL).

Linkage disequilibrium between the segregating alleles at a QTL and at a marker locus leads to associations between the marker and quantitative phenotype. In QTL mapping, this association is detected with statistical methods such as least squares or maximum likelihood approaches (Soller, 1991). For successful QTL mapping the following are needed: polymorphic markers, linkage maps, suitable populations and designs, and statistical methods.

1.1.2 Markers and linkage maps

A marker is a polymorphic locus that can be typed in the living organism, but does not itself necessarily have any effect on the trait of interest. The first QTL mapping study was carried out by Sax (1923). In this experiment the markers were three loci affecting colour of beans. Later during the 1960's and 1970's allozymes, which are detected by electrophoresis or by their products such as

blood group antigens, were used as markers (e.g., Neimann-Sørensen and Robertson, 1961; Lynch and Walsh, 1998) and they were still common in the early 1990's (e.g., Cowan et al., 1992; Andersson-Eklund and Rendel, 1993). In the 1980's DNA markers became available and QTL mapping became more feasible. The first widely used DNA markers were restriction fragment length polymorphisms (RFLP) (Botstein et al., 1980). These were replaced by microsatellite markers which are more polymorphic, uniformly distributed on the genome and easy to use and therefore more suitable for mapping purposes (Weber and May, 1989). Presently the use of single nucleotide polymorphism (SNP) markers, which are commonly used in QTL mapping of the mouse and human, is increasing especially in fine mapping experiments because of their high density in the genome (e.g., Hu et a., 2004; Liu et al., 2006; Viitala et al., 2006; Edderkaoui et al., 2007).

The markers are ordered on the chromosomes to construct linkage maps (Ott, 1991). The basic principle in mapping is that family members are genotyped and recombination frequencies between the markers passed to offspring are observed. The linkage map shows the genetic distance between the markers on the chromosomes. The distances are based on recombination frequencies between the markers and are transformed to additive map distances using map functions, of which Haldane and Kosambi map functions are most common (Ott, 1991). In addition to genetic mapping also physical mapping using bacterial artificial chromosomes (BAC) or yeast artificial chromosomes (YAC) or radiation hybrid (RH) mapping can be used to order the markers on the chromosomes. The first bovine linkage maps constructed in the early 1990's had only a few hundred markers or less (Barendse et al., 1994; Bishop et al., 1994). The current USDA linkage map has several thousands of markers and the length of the map is approximately 3200 cM (Sonstegard and van Tassell, 2004). Although the early linkage maps were already useful for QTL mapping studies because an average spacing of 20 cM between highly informative markers is quite adequate for whole genome scans (Haley and Andersson, 1997), denser maps are necessary for fine mapping of detected QTL. The ultimate map would be the whole bovine genome sequence with information utilizing the variation at the nucleotide level. The bovine genome has been sequenced with a 7.1 fold coverage, but there are still some gaps, errors and uncertainties in the sequence information that has to be corrected before it can be maximally utilized (http://www.hgsc.bcm.tmc.edu/projects/bovine).

1.1.3 Experimental designs for QTL mapping

In plants and laboratory animals the most widely used experimental design for QTL mapping involves crosses of inbred lines (F2 or backcross) (e.g., Winkelman and Hodgetts, 1992; Collins et al., 1993; Zhang et al., 2004). Here linkage disequilibrium is created in the genome by crossing and is used to find associ-

ations between markers and QTL that differ between the lines (e.g., Lynch and Walsh, 1998). In this situation the F1 animals are all heterozygous for the marker and the QTL and will have the same linkage phase. The mean phenotypic differences of animals with different marker genotypes are then used to map the QTL in the F2 population (Haley and Andersson, 1997). For cattle, inbred lines are not available. Crosses of outbred populations can be used but usually the aim is to detect QTL that explain genetic variation within a population. For this reason QTL mapping of dairy cattle has to be carried out in outbred populations. This is more complicated than with inbred lines and has lower power because both markers and QTL can have several alleles segregating and only part of the animals are heterozygous for markers and QTL. Further, the linkage phase for the marker and the QTL may differ between animals because the markers and the QTL are expected to be in linkage equilibrium in outbred populations (Haley and Andersson, 1997). For this reason QTL mapping in outbred populations is done within families where linkage disequilibrium exists and the association between marker and trait can be detected by looking at the mean differences of groups of progeny receiving alternative marker alleles from a parent (Geldermann, 1975). Because genotyping and in some cases also the collection of phenotypic data is expensive, experimental designs which minimize the cost and which are suitable for dairy cattle, have been developed. The most commonly used are the daughter design (Neimann-Sørensen and Robertson, 1961) and the granddaughter design proposed by Weller et al. (1990). In the dairy cattle breeding system where AI is of great importance, large half-sib families are common. The daughter and granddaughter designs use this half-sib family structure. In the daughter design the sires and daughters are genotyped and the phenotypes are measured in the daughters whereas in the granddaughter design the grandsires and sons are genotyped and the phenotypic data is collected from the granddaughters. The daughter design is more useful in situations where phenotypic data collection is difficult and/or expensive. In other situations the granddaughter design is preferred because it involves less genotyping for the same power. For example with type I error of 0.01, QTL effect of 0.2, measured as half the difference between the mean trait values for the two alternative homozygotes at the QTL devided by the within-QTL genotype standard deviation for the quantitative trait, and h^2 of 0.2, the power with 20 sires with 400 daughters involving 8000 genotypings is 0.93 and the power for a granddaughter design with 20 grandsires with 100 sons and 50 granddaughters per son involving 2000 genotypings is 0.95. The granddaughter design has been later extended by using the relationships between sires and grandsons (Coppieters et al., 1999; Bolard and Boichard, 2002) and full sib information (van der Beek et al., 1995). Also the use of large complex pedigrees has been proposed (Almasy and Blangero, 1998; Bink and van Arendonk, 1999).

1.1.4 QTL mapping methods for dairy cattle

In QTL mapping a QTL affecting some particular trait is assigned to a chromosome location. This is done with statistical methods that use the information of the conditional probability of a QTL genotype given the observed marker genotype (e.g., Lynch and Walsh, 1998).

The simplest mapping approach is to look at associations between markers and trait considering one marker at a time using maximum likelihood (e.g., Weller, 1986, Mackinnon and Weller, 1995) or linear regression methods (e.g., Cowan et al., 1992; Andersson-Eklund and Rendel, 1993). With the single marker approach it is not possible (least squares method) or it is difficult (maximum likelihood method) to separate QTL position and QTL effect (Geldermann, 1975; Haley and Andersson, 1997). Further, the detection of a QTL can be biased towards a more informative marker (Haley et al., 1994) and power of detection may be low because only some of the animals are informative for that particular marker. Therefore, interval mapping methods (Lander and Botstein, 1989) which use information of flanking markers (Georges et al., 1995; Knott et al., 1996) are mostly used for dairy cattle QTL mapping. The method is based on the probability of an offspring receiving one or the other of the sire's two alleles conditional on the marker genotype at particular positions along the chromosome. Least squares and maximum likelihood approaches can both be applied (Xu and Atchley, 1995; Knott et al., 1996). The least squares method is computationally easier and may be more robust (Knott et al., 1996). Other mapping methods proposed for cattle data are the non-parametric rank-based method (Coppieters et al., 1998), variance component methods (e.g., Xu and Atchley, 1995; Grignola et al., 1996; George et al., 2000) and Bayesian methods (e.g., Hoeschele and van Raden, 1993; Uimari et al., 1996). The non-parametric method is suitable for traits that are not normally distributed (Coppieters et al., 1998). Variance component methods assume the QTL is a random effect and marker information is used to compute the IBD (identical by descent) status at the QTL at a certain chromosome position (Lynch and Walsh, 1998). The IBD information is used to compute the (co)variance matrix of the additive effects of the QTL conditional on the markers. It is assumed that the greater the IBD proportion is between the animals, the more similar are the phenotypes (Xu and Atchley, 1995). A mixed linear model is used to estimate parameters. An advantage of the variance component method is that complex pedigrees can be used instead of simpler designs like half-sib families (George et al., 2000).

The first mapping methods handled only one QTL and trait at a time but later extensions have been made to account for multiple QTL on the same chromosome (Jansen and Stam, 1994; Spelman et al., 1996) and simultaneous mapping of multiple traits using inbred crosses (e.g., Jiang and Zeng, 1995; Knott and Haley, 2000) or outbred pedigrees (Sørensen et al., 2003). Multi trait approach-

es have been shown to increase the power to detect QTL and the precision of the location estimate (Knott and Haley, 2000; Sørensen et al., 2003). Analysing two traits simultaneously is especially advantageous when one of the traits has low heritability (Sørensen et al., 2003). Cofactors have been modelled to account for part of the background variation caused by other QTL in inbred line crosses (Jansen, 1993; Jansen and Stam, 1994; Zeng, 1994) where markers are fitted as cofactors, and in half-sib designs where detected QTL are fitted as cofactors (de Koning et al., 2001). The cofactor approach described by de Koning et al. (2001) involves analysing chromosomes individually, identifying possible QTL and choosing them as cofactors, jointly re-estimating the QTL effects, adjusting phenotypes for cofactors and re-analyzing the chromosomes. Use of cofactors as described by de Koning et al., (2001) has been shown to increase the amount of false positives substantially in scenarios with low heritability of the trait and small half-sib family sizes (Sahana et al., 2006). The reason is likely to be that fitting cofactors in non-segregating families will account for part of the non-genetic part of the residual variance (Sahana et al., 2006).

An important question in QTL mapping studies is how to determine the appropriate significance thresholds for QTL detection. Because whole genome scans involve a large number of hypothesis tests, over 3000 with interval mapping in cattle, some of the QTL detected with a point-wise 0.05 significance level will be false positives (Lander and Kruglyak, 1995). Therefore more stringent experiment-wise significance levels should be used to declare significant linkage (Lander and Kruglyak, 1995). Lander and Kruglyak (1995) suggested that two levels of significance should be used: suggestive linkage where one false positive is expected in a whole genome scan and significant linkage where a 5 % genome-wise significance level is applied.

The often used approach to derive significance thresholds for QTL mapping methods is the permutation test by Churchill and Doerge (1994) and Doerge and Churchill (1996). With this approach no assumptions are necessary about the distribution of the test statistic under the null hypothesis or the distribution of the phenotypic trait. In the permutation test the phenotypes of the original data are shuffled and the genotypes kept constant. With a half-sib family structure the trait data is shuffled within sons of each family. This way several data sets are generated and the new data are analyzed to produce test statistics. This gives an empirical distribution of the test statistic under the null hypothesis of no QTL. Chromosome-wise and experiment-wise thresholds can be computed with this method. Alternatively a Bonferroni correction can be used to obtain the experiment-wise threshold (Lynch and Walsh, 1998). Here the chromosome-wise significance level obtained is corrected for multiple comparisons. The correction can be done for number of chromosomes and number of independent traits analyzed. The different lengths of the chromosomes can also be taken into account (de Koning et al., 1999).

Another way to control the genome-wise type I error rate is the method by Piepho (2001). This method computes approximate threshold levels for QTL mapping experiments. This is a general method that allows for any population structure. It assumes normality of the errors under the null hypothesis but is quite robust to this assumption. It is useful in situations when no close form thresholds are available and it is much quicker to compute than permutation thresholds. It needs the LRT values of the mapping experiment as input values.

Even if it is important to know if the QTL are real instead of false positives before they are used in MAS (Spelman and van Arendonk, 1997), or before they are chosen as targets for fine mapping, it is not necessarily clear that the avoidance of false positives is very important in whole genome scans. Very strict thresholds would mean that many real QTL would never be reported. For example Curtis et al., (1996) suggested that especially in low power experiments it would be appropriate to report also less significant QTL. All QTL results should anyway be replicated in independent studies before they can be declared as certain. Even with a true QTL some of the replication studies will possibly not detect the effect, depending on the power of the experiment. A meta-analysis of all studies, including those where the QTL were detected and those were they were not detected, would be the best proof for a QTL to be real (Lander and Kruglyak, 1995).

1.2 QTL mapping studies in dairy cattle

Several QTL mapping experiments have been carried out. In earlier studies the mapping was done on single or few chromosomes using only few markers (e.g., Geldermann et al., 1985; Andersson-Eklund and Rendel, 1993; Ron et al., 1994; Vilkki et al., 1997) More recently whole genome scans, where associations between markers and traits are searched for on all bovine autosomes, have been carried out (e.g., Heyen et al., 1999; Schrooten et al., 2000; Kühn et al., 2003). Most of the QTL have been detected for milk production traits such as milk yield, protein yield, fat yield, protein content, and fat content (e.g., Bovenhuis and Weller, 1994; Ron et al., 1994; Georges et al., 1995; Zhang et al., 1998; Rodriguez-Zas et al., 2002). This is because milk production traits are routinely recorded in national milk recording schemes in many countries and the records are easily available for research purposes. Several studies have also detected QTL for health traits such as somatic cell score (SCS) and clinical mastitis (Klungland et al., 2001; Holmberg and Andersson-Eklund, 2004), fertility traits such as ovulation rate, days open, non-return rate, and fertility treatments (Blattman et al., 1996; Kühn et al., 2003; Holmberg and Andersson-Eklund, 2006) and udder, body and leg conformation traits (Schrooten et al., 2000; Boichard et al., 2003; Hiendleder at al., 2003; Buitenhuis et al., 2007). The extensive data collection in the Nordic countries has allowed for mapping of traits such as veterinary treatments for clinical mastitis and fertility. For all

traits QTL have been detected on several chromosomes. Many QTL that have been detected in one study have been confirmed in another one but some have been detected only once. In cases where a QTL has been detected for the same trait on the same chromosome in different studies the positions of the highest test statistic may vary considerably (reviewed by Khatkar et al., 2004; http://www.animalgenome.org/QTLdb/cattle.html).

The final aim of QTL mapping is to localize the gene and the mutation that causes the QTL effect. This is a difficult task which includes fine mapping of the chromosome region of interest, comparative genomics over species in order to localize possible candidate genes and expression studies to get functional evidence for the proposed polymorphism (Meuwissen and Goddard, 2000; Mehrabian et al., 2005). The mapping is carried out with a dense set of markers and combined linkage and linkage disequilibrium based statistical methods (Meuwissen et al., 2002) are used to find the association between markers and trait value. Expression studies using microarrays can be applied to further bring down the number of candidate genes in the QTL location (Wayne and McIntyre, 2002). Finally, sequencing of a narrow chromosome segment, harbouring a candidate gene, may be carried out to reveal the possible mutation (Grisart et al., 2002). Finding a polymorphism that is in agreement with the QTL genotypes is not a sufficient proof for the mutation being the real cause of the QTL effect because it can still just be in strong linkage disequilibrium with the causative gene (de Koning, 2006). The only bovine QTL where at least one underlying functional mutation has been detected with certainty is the DGAT1 gene encoding acylCoA:diacyglycerol acyltransferase which is located on chromosome14 and has a major effect on fat content (Grisart et al., 2002; Grisart et al., 2004).

1.3 QTL in selection schemes of dairy cattle

A major objective of QTL mapping, in addition to getting to know the genetic architecture of the quantitative traits, is to find QTL that can be used in breeding schemes (Soller and Beckmann, 1983). A strategy that uses information of markers linked to individual QTL for selection decisions is called marker-assisted selection (MAS). Ideally the marker would be the QTL itself, but in most cases marker brackets would be used.

Breeding schemes traditionally applied to dairy cattle are based on phenotypic information from individuals and/or their relatives and progeny testing of the bull candidates. With MAS the genetic response may be increased compared to conventional breeding by increasing the accuracy of selection, decreasing the generation interval and increasing the intensity of selection (Smith and Simpson, 1986; Kashi et al., 1990; Soller, 1994). The accuracy can be increased by getting information from the QTL in addition to the conventional information. This is especially useful for low heritability traits (Meuwissen and God-

dard, 1996; Ruane and Colleau, 1995) and traits that are difficult or expensive to measure. The generation interval can be decreased for sex-limited traits, such as most of the economically important traits of dairy cattle, if bulls can be selected based on markers instead of the progeny test results (Ruane and Colleau, 1996). In nucleus schemes calves and heifers could be selected more accurately before they have lactation records. Marker information can already be used to select between new born calves or embryos (Peippo et al., 2007). The intensity of selection can be increased by using marker information to select for example among full sibs in nucleus schemes where embryo transfer is applied (Gomez-Raya and Klemetsdal, 1999).

Before QTL can be used in breeding schemes they have to be verified to be real and to exist in the present breeding population (Spelman and Bovenhuis, 1998). This is important because use of non-existing QTL may decrease genetic progress instead of increasing it (Spelman and van Arendonk, 1997). Other factors that should be estimated before implementing MAS are effects of the QTL and allele frequencies of the QTL (Smith and Simpson, 1986). Also sires heterozygous for the QTL should be identified and their linkage phases should be determined. Additionally, it is necessary to find out the effects of the QTL on other important traits than the ones they were first aimed for.

Three different types of markers can be used: direct markers, linkage disequilibrium (LD) -markers and linkage equilibrium (LE) -markers (Dekkers, 2004). Direct markers are the real mutations in the gene that causes the phenotype differences. LD-markers are markers that are very close to the actual mutation and are in population wide linkage disequilibrium with the mutation. LE-markers are located further away from the gene and the linkage disequilibrium exists only within families. For use in MAS, direct markers and LD-markers are more suitable because they can be used in the whole population and the effects persist over generations. The LE-markers can be used only within families because the effect and linkage phase between markers and QTL differ between families. Further the LE-markers have to be re-evaluated at every or almost every generation. Most QTL available at the moment can be traced only using LE markers, but the amount of fine-mapped QTL is increasing.

Two different approaches for MAS have been proposed: within-family selection where QTL information is used to select within families and conventional EBVs are used to select between families (e.g., Kashi et al., 1990) and MAS using BLUP where marker information is included in the mixed model (Fernando and Grossman, 1989). The MAS scheme most suitable for dairy cattle is where QTL information is used to select among young bulls entering progeny test. For this scheme the top down and bottom up strategies have been proposed (Kashi et al., 1990; Mackinnon and Georges, 1998). In the top down scheme the QTL information of the grandsires (bulls sire's sires and/or bull-dam's sires) is used

to pre-select among grandsons, and in the bottom up scheme the QTL of a progeny tested sire is used to pre-select among his sons.

Several simulation studies investigating MAS schemes have been conducted (e.g., Kashi et al., 1990; Meuwissen and Goddard, 1996; Ruane and Colleau, 1996; Mackinnon and Georges, 1998; Spelman and Garrick, 1997; Schrooten et al., 2005). The assumptions about QTL number, QTL effect, QTL allele frequencies and breeding schemes vary a lot between different investigations. Consequently, the additional genetic response achieved by MAS relative to conventional breeding schemes also varies greatly ranging from a few percent to over 20%. The largest responses have been seen for low heritability traits (e.g., Ruane and Colleau, 1995), traits that are recorded late in life (Meuwissen and Goddard, 1996) and MAS combined with embryo transfer schemes (e.g., Schrooten et al., 2005).

In some simulation studies where MAS has been applied for several generations using an index which combines QTL information and phenotypic information, it has been seen that the genetic response using MAS is first higher than with traditional selection and lower in later generations (Gibson, 1994; Hospital et al., 1997). This is because the response of the polygenes is lower with MAS. As reasons to this it has been proposed that negative linkage disequilibrium is building up between the QTL and polygenes (Gibson et al, 1990) and fixing the favourable alleles of the larger QTL would lead to a hitch-hiking effect where unfavourable alleles of QTL of smaller effect would be fixed also (Hospital and Chevalet, 1996). Also, it has been proposed that the variance of the QTL, which is determined by the allele frequency changes, is affecting the selection pressure on the polygenes (Dekkers and van Arendonk, 1998). The loss of long-term response can be avoided by putting different weights on QTL and polygenic information over time (Dekkers and van Arendonk, 1998). Also doing a full genome scan at each generation and selecting for different QTL over time has been shown to maintain the genetic response of MAS in the long-term (Stella et al., 2002).

Marker-assisted selection is already used in practice in some countries. For German Holstein cattle MAS has been started in 1995. Young bulls and bull dam candidates are genotyped for 13 markers on three chromosomes where QTL for milk production traits have been detected (Szyda et al, 2005). In France, a MAS program has been started in 2000 for Holstein, Normande and Montbéliarde cattle breeds (Druet et al., 2005). Animals are typed for 45 markers corresponding to 14 QTL affecting milk production traits, SCC, fertility and udder conformation traits. The QTL information is used to select among young bulls and bull dam candidates. In both MAS schemes marker information is combined with phenotypic information in an index (Fernando and Grossman, 1989). The Danish MAS program was started in the summer 2004. LE-markers associated with mastitis on four chromosome areas have been typed so far in a total of 165 bull

calves and 73 bull mothers. Inherited QTL allele and QTL effect have been reported for these animals to support selection decisions. Currently 58 LE-markers and 1 LD-marker on seven different chromosomes associated with udder health, direct and maternal calving ease, fertility and milk production traits are selected to be typed in the near future. The marker information will be combined with the phenotypic information to give MA-BLUP breeding values (Jørn Rind Thomasen, personal communication, 2007).

Currently much of the research on using marker information in dairy cattle selection schemes is focusing on the development of methods for implementing genome-wide selection (Meuwissen et al., 2001). This involves genotyping of densely located SNP markers throughout the genome and estimating the haplotype effects of SNP pairs. A genomic estimated breeding value with high accuracy can then be obtained for each genotyped animal (Schaeffer, 2006).

2 Objectives of the study

The objectives of this study were (1) to map QTL for economically important traits, i.e. ,milk production traits (I), health traits (II), and fertility traits (III) in Finnish Ayrshire dairy cattle, and (2) to investigate the effects of using QTL in a two-stage selection scheme of dairy cattle, and especially to look at the effect of marker-assisted selection on the genetic response and the linkage disequilibrium between the different parts of the genome (IV,V).

3 Materials and methods

3.1 Detection of QTL for economically important traits

3.1.1 Families and traits

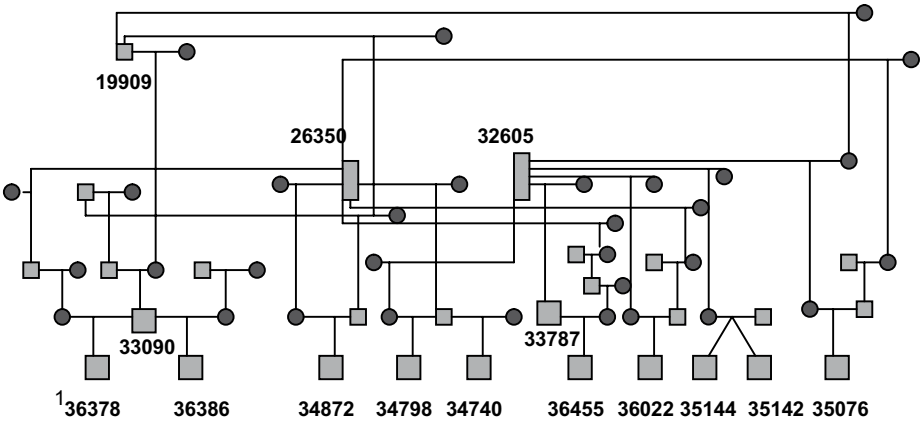
Twelve half-sib families of Finnish Ayrshire (Table 1) were analyzed using a grand daughter design (Weller et al., 1990). These families were chosen because they had the highest number of sons among the families available. The two oldest families (33090 and 33787) have only selected sons available . Many of the bulls are related to each other (Figure 1). For example bulls 35142 and 35144 are full-sibs (dizygotic twins) and bulls 36378 and 36386 are sons of 33090. The total number of sons was 493. The number of sons per bull ranged from 21 to 82 with an average of 41. There was a large variation in the number of daughters per bull, ranging from only a few to several thousands. The average number of daughters per bull was around 500.

Semen samples were provided by the five Finnish Artificial Insemination (AI) stations. Estimated breeding values (EBV) were obtained from the Finnish Animal Breeding Association. For the milk production traits, results from the 1998 evaluation were used. For health and fertility traits EBVs were from the fall 2000 evaluation, except in non-return rate for which they were from the spring 1996 evaluation. This was because there was not enough data for the six oldest families in the year 2000 evaluation for non-return rate.

Table 1. Grandsires in the granddaughter design of Finnish Ayrshire.

id	name	sire	sire of dam	year of birth	sons[1]
33090	Koivuniemen Yllätys	30480	19909	1973	30
33787	Isopuolin Alleri	32605	30551	1974	28
34740	Kytölän Iivari	31331	30635	1977	59
34798	Sorpo Ingvar	31331	32605	1977	41
34872	Kiiskilän Junnu	30992	26350	1978	50
35076	Granudd Joakim	32345	32605	1978	21
35142	Rantalan Jokeri	31838	32605	1978	82
35144	Rantalan Junkkari	31838	32605	1978	29
36022	Peltohaan Laiho	33066	32605	1980	29
36378	Tuomelan Minos	33090	31500	1981	44
36386	Luukkaanmäen Miklaus	33090	32633	1981	40
36455	Kuusiston Mainio	33787	32875	1981	40

[1] Number of sons of the grandsire



[1] Bulls with large squares are the grandsires in the data.

Figure 1. Pedigree of the grandsires from the Finnish granddaughter design data in studies I, II, and III.

Twelve different traits were studied. There were five milk production traits, three health traits, and four fertility traits. The traits were: milk yield, protein yield, protein content, fat yield, fat content, somatic cell score (SCS), mastitis treatments, other veterinary treatments, days open, fertility treatments, non-return rate, and calf mortality. The milk production EBVs were based on 305 day milk yield. The SCS was based on daughters' somatic cell count (SCC) transformed to a logarithmic scale. Mastitis was based on treatments for mastitis done by a veterinarian within 7 days before and 150 days after calving or culling due to udder health disorders within the same time period. Other veterinary treatments included all other treatments except mastitis or fertility treatments within 150 days of calving. Milk fever, ketosis and retained placenta were the most frequent disorders included in that category. Days open was the number of days from calving to next pregnancy. Fertility treatments was based on treatments for fertility disorders done by a veterinarian within 7 days before and 150 days after calving or on culling due to fertility disorders within the same time period. Non-return rate, a paternal trait, was based on inseminations with a bull's semen to a random set of cows. It is measured as the non-return rate within 60 days from insemination. The first 500 inseminations of a bull are included. Calf mortality as a trait of the bull is based on the mortality at birth of the offspring of the daughters.

Mastitis, other veterinary treatments, fertility treatments, and calf mortality were recorded as binary traits 0 or 1. The breeding values used in the mapping studies were estimated by the Finnish Animal Breeding Association as part of the routine genetic evaluation. A repeatability animal model was used to estimate breeding values for the milk production traits, SCS, and days open. A re-

peatability sire model was used for mastitis, other veterinary treatments, and fertility treatments. For calf mortality a repeatability sire-grandsire model was used. Records from the first 3 lactations were used. All bulls had daughters in all three lactations. For non-return rate, a selection index was used for genetic evaluation. Pre-adjustment was done for month and AI-cooperative.

3.1.2 Marker maps and genotypes

All 29 bovine autosomes were included in the analyses. The marker spacing was on average 20 cM. A total of 150 markers were used in all other studies except for fertility traits (study III) for which 171 markers were used. More markers were added for the third study in order to fill gaps in the maps. Three candidate genes were included: growth hormone receptor (Moisio et al, 1998), prolactin receptor (Viitala et al., 2006), and caseine gene haplotypes (Velmala et al., 1995). All other markers were microsatellites. Number of markers varied from 2 to 14 per chromosome and the average number of informative markers ranged from 1.33 (BTA27) to 8.17 (BTA6) per chromosome. Marker maps were constructed with ANIMAP or CRIMAP programs (Green et al., 1990; Georges et al., 1995) (Table 2). Polymorphic information content of markers was calculated. In studies I an II the length of the total genome was 2764 cM. In study III 11 maps, BTA1, 3, 5, 9, 11, 12, 18, 20, 23, 25, and 29, were recalculated (Figure 2). The length of the total genome with the new maps included was 2618 cM. Methods of DNA extraction, PCR reaction protocols and electrophoresis were described by Vilkki et al. (1997) and Viitala et al. (2003). All available sons of the chosen grandsires and all the grandsires with semen samples were genotyped for all markers.

Table 2. Markers in linkage groups BTA 1-29 used in studies I and II, their locations in cM (Haldane), polymorphic information content (PIC) values and number of alleles.

chromosome and marker	cM	PIC	alleles	chromosome and marker	cM	PIC	allelels
BTA1				BTA2			
TGLA49	0	0.69	8	ILSTS026	0	0.67	5
ILSTS104	24	0.39	4	INRA40	1	0.79	9
TGLA57	67	0.58	7	TGLA61	17	0.68	6
BM6506	86	0.49	7	URB42	35	0.64	7
BM864	105	0.72	11	BM4440	79	0.75	8
CSSM32	119	0.46	5	TGLA226	101	0.70	5
CSSM19	154	0.59	5	BM2119	147	0.75	7
MAF46	157	0.56	3	ORFCB11	167	0.81	9
BM3205	157	0.36	2				

continues

chromosome and marker	cM	PIC	alleles	chromosome and marker	cM	PIC	allelels
BTA3				BTA4			
INRA006	0	0.46	4	RM188	0	0.69	9
UWCA7	1	0.42	3	HUJ673	22	0.33	2
FCGR	15	0.42	3	TGLA116	29	0.36	3
BL41	30	0.39	4	BM6458	46	0.63	7
INRA23	31	0.67	8	BM1500	74	0.42	3
HUJ246	63	0.64	6	BMS648	77	0.61	6
HUJ1177	97	0.67	5	BR6303	94	0.68	6
INRA197	130	0.81	8				
BTA5				BTA6			
BM6026	0	0.69	6	ILSTS93	0	0.69	8
BP1	13	0.54	6	INRA133	16	0.56	5
CSSM34	44	0.59	3	ILSTS090	21	0.22	2
ETH10	71	0.79	6	URB016	39	0.23	2
BM1819	83	0.47	4	BM1329	41	0.66	5
ETH152	127	0.74	6	BM143	68	0.83	10
BM2830	131	0.66	9	ILSTS097	84	0.26	2
				BM4528	86	0.39	4
				RM028	89	0.50	3
				BM415	93	0.56	7
				CSN	104	0.68	5
				AFR227	107	0.44	4
				BP7	112	0.47	4
				BM2320	151	0.72	12
BTA7				BTA8			
BM7160	0	0.64	4	IDVGA11	0	0.73	11
BM6105	30	0.78	7	INRAMTT180			
BM6117	57	0.49	3		42	0.77	9
INRA192	87	0.58	4	HEL9	81	1.00	8
ILSTS006	132	0.73	8	BMS2847	130	0.72	8
BL1043	165	0.81	12				
BTA9				BTA10			
CSSM56	0	0.59	5	CSSM038	0	0.79	9
TGLA73	22	0.68	6	ILSTS53	44	0.68	7
UWCA9	55	0.73	7	ILSTS70	100	0.53	4
CSSM25	60	0.43	3	CSSM46	144	0.65	8
ETH225	108	0.77	7				
INRA136	124	0.62	4				

continues

chromosome and marker	cM	PIC	alleles	chromosome and marker	cM	PIC	allelels
BTA11				BTA12			
HEL13	0	0.65	5	BMS2057	0	0.80	11
TGLA438	18	0.61	6	BM6404	27	0.82	9
BMS1048	50	0.63	5	BMS1316	73	0.73	9
HELMTT41	69	0.60	6				
BM7169	86	0.76	8				
INRA177	103	0.68	9				
BM716	131	0.58	5				
BTA13				BTA14			
TGLA23	0	0.68	8	ILSTS039	0	0.63	8
BMS1352	27	0.75	5	BMS1747	16	0.76	6
RM327	57	0.72	7	RM011	50	0.84	9
BMS995	117	0.69	4	BMS740	66	0.58	7
				BM4513	79	0.69	7
BTA15				BTA16			
RM4	0	0.67	7	BM1311	0	0.67	6
HBB	36	0.73	9	IDVGA49	34	0.64	8
HEL1	62	0.60	5	BM1706	62	0.67	7
NCAM	74	0.54	8				
BR3510	100	0.72	6				
MGTG13B	114	0.27	2				
BTA17				BTA18			
BM1233	0	0.67	5	BMS1355	0	0.76	8
ETH185	23	0.73	7	BMS2213	26	0.65	7
BMS941	62	0.79	9	BMS2639	76	0.72	9
				TGLA227	110	0.80	10
BTA19				BTA20			
ETH3	0	0.63	6	BM5004	0	0.64	6
IOBT34	22	0.61	7	BM4107	32	0.30	4
MAP2C	25	0.59	4	ILSTS072	40	0.64	6
CSSM65	34	0.62	4	GHR	49	0.47	3
BP20	57	0.39	3	BM713	61	0.54	4
URB44	67	0.73	7	TGLA304	87	0.53	4
HEL10	90	0.58	6	BM3517	109	0.79	7
BTA21				BTA22			
RM151	0	0.77	8	CSSM026	0	0.78	9
INRA103	40	0.74	6	BM1520	45	0.76	8
TGLA122	68	0.65	8	OARFCB304			
					70	0.23	3

continues

chromosome and marker	cM	PIC	alleles	chromosome and marker	cM	PIC	allelels
BTA23				BTA24			
CSSM005	0	0.70	7	BMS2270	0	0.83	10
RM033	11	0.66	5	BMS466	40	0.80	12
BM1258	21	0.56	5				
BOLA-DRB1							
	31	0.29	3				
RM185	38	0.83	8				
CSSM024	53	0.76	13				
BTA25				BTA26			
BMC4216	0	0.48	3	HEL11	0	0.56	7
BMS130	13	0.62	5	BMS2567	15	0.64	5
BMS1353	59	0.63	8				
AF5	70	0.78	10				
BTA27				BTA28			
BM641	0	0.75	8	BMS510	0	0.79	8
INRA134	29	0.52	3	BMS1714	18	0.55	4
BTA29							
JAB5	0	0.60	5				
ILSTS057	4	0.49	6				
BMC8012	24	0.59	6				
BMC1206	77	0.56	3				

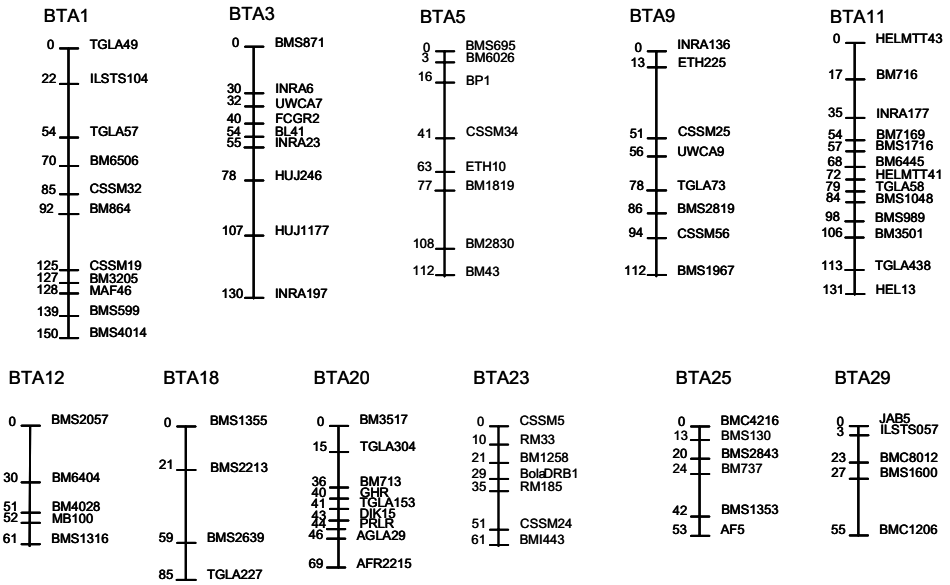


Figure 2. Recalculated linkage maps for study III.

3.1.3 Statistical methods

A multiple marker regression approach (Knott et al., 1996) was used to detect associations between markers and traits. The model used was: $y_{ijk} = a_i + b_i x_{ijk} + e_{ijk}$, where y_{ijk} is the breeding value of bull j, which belongs to family i, and has marker genotype k, a_i is the polygenic effect for half-sib family i, b_i is the allele substitution effect for a QTL within family i, x_{ijk} is the probability for bull j of inheriting the arbitrarily defined first QTL allele from sire i, given the pair of flanking markers k, and e_{ijk} is the residual. The breeding values were weighted by their reliabilities. The most likely linkage phases of the chromosomes of the sire were determined and conditional probability of inheriting the first allele was calculated at 1 cM fixed intervals for each chromosome of the sons. The trait value was regressed on this probability. The regression analysis was nested within families and the test statistic was computed as an F-ratio. The most likely position of a QTL on the chromosome is at the highest value of the F statistic. The regression coefficient is an estimate of the QTL allele substitution effect.

Power of the test was increased by fitting the transmission probabilities of QTL found on other chromosomes as cofactors (de Koning et al., 2001). Quantitative trait loci exceeding the 5 or 10 % nominal significance threshold were chosen as cofactors. A lower threshold value was used for health traits. The presence of two QTL on the same chromosome was tested by a 2-QTL model (Spelman et al., 1996; Velmala et al., 1999). The analyses were performed on chromosomes with more than three informative markers that had some evidence for more than one QTL. That is, the highest values of the test statistic were clearly at different positions on the chromosome in different families.

Significance thresholds and *P*-values were computed by the permutation test (Churchill and Doerge, 1994). Here the genotypes were kept and the phenotypes were shuffled. Then the QTL mapping was done for the shuffled data set. This was repeated 10 000 times and the results were sorted to obtain chromosome-wise significance thresholds. Significance was determined before choosing a new set of cofactors during the analysis process and at the end of the analysis for the final test statistic values. Genome-wise *P*-values were obtained by Bonferroni correction using the formula: $P_{genome} = 1 - (1 - P_{chromosome})^{29}$, where 29 is the number of autosomes. With the 2-QTL model permutations were done to test first 2-QTL vs. no QTL. If this test was significant at 5 %, the *P*-value for 2-QTL vs. 1-QTL was obtained from a standard F table with the number of grandsires as numerator df and the number of offspring minus 3 times (two QTL effects and a polygenic effect) the number of grandsires as denominator df.

In study III part of the single trait and multi-trait QTL mapping was carried out using the variance component method (Sørensen et al., 2003). The following

linear mixed model was used: $\mathbf{y} = \boldsymbol{\mu} + \mathbf{Zu} + \sum_{i=1}^{n_q} \mathbf{Wq}_i + \mathbf{e}$,

where **y** is a vector of breeding values of t traits for each genotyped son, **μ** is a vector of overall trait means, **Z** and **W** are incidence matrices, **u** is a vector of random additive polygenic effects, $\mathbf{q}_i$ is the effect of the i^{th} QTL, n_q is the number of QTL, and **e** is a vector of random residual effects. The random variables **u**, $\mathbf{q}_i$ and **e** are assumed to be multivariate normally distributed (MVN) and mutually uncorrelated. Specifically, **u** is MVN (**0**, $\mathbf{G} \otimes \mathbf{A}$), $\mathbf{q}_i$ is MVN (**0**, $\mathbf{K}_i \otimes \mathbf{IBD}_i$) and **e** is MVN (**0**, $\mathbf{E} \otimes \mathbf{I}$). Matrices **G**, **K** and **E** include variances and covariances among the traits due to polygenic effects, additive QTL effects and residual effects, respectively. The symbol $\otimes$ represents the Kronecker product. Matrix **A** is the additive relationship matrix that describes the covariance structure among the polygenic effects, $\mathbf{IBD}_i$ is the identity by descent (IBD) matrix that describes the covariance structure among the effects for the i^{th} QTL, and **I** is the identity matrix. The tests for significance were based on the asymptotic distribution of the likelihood ratio test (LRT) statistic and the thresholds were obtained by the method by Piepho (2001). Bayesian Information Criterion (Kass and Raftery, 1995) was used to select between the pleiotropic or linked QTL models.

3.2 Two-stage marker-assisted selection of dairy cattle

3.2.1 Simulation schemes

Difference in genetic response between the traditional progeny test scheme and the MAS scheme was studied. Also the disequilibrium between the QTL and the polygenes and between the two QTL were investigated. The main interest was to look at effects of different initial frequencies of the favourable alleles and different QTL effects on the allele frequencies, genetic response and disequilibrium after only one round of selection (IV) or four generations (V). The disequilibrium between QTL was computed using the formula by Hill and Robertson (1968) and the disequilibrium between polygenes and QTL as the correlation between polygenic breeding value and combined QTL breeding value. The simulations were replicated 250 times and the results were summarized for the top sires selected after progeny test.

3.2.2 Genetic model and parameters

Selection was carried out on a sex-limited trait that was modelled with a polygenic component, two independent QTL and a non-genetic component. In the base generation the polygenic component was sampled from a normal distribution. The polygenic heritability of the trait was 0.3. The two QTL were bi-allel-

ic and additive and their effects were either 0.55 or 1.0 σ_{pg} (polygenic standard deviation), considered as moderate and large effects. For most scenarios both QTL had the same effect. The initial allele frequencies of the favourable QTL alleles in the base population varied in different scenarios (0.01, 0.02, 0.05, 0.10, 0.30, 0.50, and 0.90 in study IV and 0.01, 0.02, 0.05, 0.10, 0.30 in study V). The non-genetic component was sampled from a normal distribution. The non-genetic variance was independent of the changes in QTL variance. The sum of the polygenic variance and the non-genetic variance was modelled to be 100. The total genetic variance was allowed to vary because of selection. The Bulmer effect (Bulmer, 1971) decreased the total variance and the changes in allele frequencies of the QTL alleles affected the total variance also. The QTL were assumed to be flanked by close markers and the QTL genotypes of the animals to be known without error. The base population was assumed to be in linkage equilibrium but after the first selection of parents linkage disequilibrium was created.

3.2.3 Selection scheme

A progeny testing scheme utilizing embryo transfer was modelled. A conventional progeny testing scheme was compared to a scheme where QTL information was used within families to select among full-sibs in the male path. In study IV one round of selection was carried out and four rounds instudy V. In order to get linkage disequilibrium in the first parent generation strong selection was first applied where 1000 dams and 100 sires were selected based on an index from a large base population of 200 000 animals which was modelled to be in linkage equilibrium. In study V additionally 2000 dams were selected from the same base population to be used in the dam path. In the male path the best 100 bulls and 1000 cows were mated to produce six full-sib offspring per dam. Sex of the offspring was assigned at random. When MAS was applied the best son within a family was selected according to QTL information, if more than one son was available. In the conventional scheme one son was chosen at random. The chosen 1000 young bulls were progeny tested and the best 100 were selected on estimated breeding values, with a reliability of 0.80, to be sires of the next generation. In the dam path the 3000 (1000 + 2000) selected cows were mated to the same 100 top bulls, as in the male path, to produce a total of 9000 female offspring out of which 2000 cows were selected based on an index that included family information and one own record with a reliability of 0.45. These 2000 cows were mated to the same group of 100 bulls to produce 6000 female offspring out of which 1000 best cows were selected as bull dams for the next bull generation and 3000 best (including the bull dams) as cow dams (Figure 1 in V). The dam path was modelled to be one generation ahead of the male path because in the real breeding schemes bulls have to wait for progeny test records before they are used as bull sires.

4 Results

4.1 Detection of QTL for economically important traits

4.1.1 QTL for milk production traits

In the original across family analysis where chromosomes were analyzed separately 14 QTL associated with milk production traits were suggested at 5% chromosome-wise significance (I). QTL affecting milk yield were found on *Bos taurus* chromosome (BTA) 1, 5, 6, 12, and 20, fat yield on BTA12 and BTA14 and protein yield on BTA5, 12, and 25. QTL associated with fat percentage was detected on BTA14 and protein percentageon BTA6, 14, and 23. Out of these, the QTL for fat percentage on BTA14 and fat yield on BTA12 were genome-wise significant at 5%. When putative QTL, chromosome-wise significant at 5 %, on other chromosomes were used as cofactors in the analysis in order to increase statistical power, a total of 31 QTL were detected at 5% genome-wise significance. Out of these, 22 QTL were genome-wise significant at 1%.

When investigating the genome-wise significant QTL from the single chromosome analysis in detail, it was found that the position of the QTL for fat percentage was close to ILSTS039 at the centromeric end of BTA14. The within families analysis suggests that there are three families with the QTL segregating. The QTL substitution effect was 0.65, 0.36 and 0.29 %-units in these families. The largest of these effects is almost 2.5 standard deviations of EBV. The fat yield QTL on BTA12 was located close to marker BM6404. The QTL was seen in two families and the allele substitution effects were 15.2 kg and 11.7 kg which are close to one standard deviation of EBV.

The results from the cofactor analyses show that several of the chromosomes have QTL for more than one milk production trait (Table 3). On BTA12, QTL were detected for milk yield, fat yield, protein yield, and protein percentage. The positions for these QTL were between markers BMS2057 and BM6404, except the QTL for fat yield which was located in proximity of marker BM6404. On BTA3 and BTA6, QTL were detected for milk yield and both milk percentage traits. On BTA14, QTL were found for fat yield and both percentage traits, and on BTA25, QTL were detected for milk yield, protein yield and protein percentage. In addition, two QTL affecting different milk production traits were detected on five chromosomes.

Table 3. Detected QTL for milk production traits on Bos taurus chromosomes (BTA).

BTA	milk yield	fat yield	protein yield	fat %	protein %
1	o[1]				
2	x				
3	x			x	x
5	o		o		
6	o			x	o
12	o	o	o		x
14		o		o	o
19				x	
20	o				x
21	x				
23	x				o
25	x		o		x
26				x	
27	x	x			
29	x	x			

[1]o indicates a QTL detected in the single chromosome analysis and cofactor analysis
x indicates a QTL detected only in the cofactor analysis

4.1.2 QTL for health traits

In the first stage when the chromosomes were analyzed without cofactors, a total of 12 QTL were located at 5% chromosome-wise significance. Quantitative trait loci for somatic cell score (SCS) were found on BTA3, 11, 14, 18, 27, and 29, for mastitis treatments on BTA18 and for other veterinary treatments on BTA2, 14, 16, 22, and 23 (II). Only the QTL for other veterinary treatments on BTA23 was close to being genome-wise significance. When cofactors were added to increase power to detect QTL, a total of 17 genome-wise significant QTL were detected: 8 for SCS, 2 for mastitis, and 7 for other veterinary treatments.

The number of segregating families per chromosome and trait was quite low, i.e., 2 to 4 for SCS, 1 to 3 for mastitis, and 2 to 5 for other veterinary treatments in the analyses where cofactors were used. For SCS the most likely positions for the QTL in different families were similar, but for the other traits there were considerable differences. The effects of the QTL were 0.5 to 1.7 SD of EBV for SCS, 0.7 to 1.4 SD of EBV for mastitis, and 0.6 to 1.4 SD of EBV for other veterinary treatments.

4.1.3 QTL for fertility traits

In the single trait regression analysis 22 fertility QTL were detected at 5% chromosome-wise significance (III). Quantitative trait loci for days open were found on BTA1, 2, 5, 12, 20, 25, and 29, for fertility treatments on BTA1, 5, 10, 14,

15, 19, and 25, for calf mortality on BTA4, 6, 11, 15, 18, and 23, and for non-return rate on BTA10 and 14. None of these QTL were genome-wise significant. The significant chromosome and trait combinations were re-analyzed with the single trait variance component method. Only three of the QTL were detected with this method: the QTL for days open on BTA1 and 12, and the QTL for fertility treatments on BTA1. Though not significant, many of the locations of the highest LRT-values were at the same positions as the positions of the highest F-values in the regression analysis.

Multi-trait analysis were carried out if QTL for two different fertility traits were detected on the same chromosome in the regression analysis or if a fertility trait QTL was found on the same chromosome with a milk production trait QTL. Chromosomes with milk production QTL were chosen from the results of study I and they were reanalyzed using the variance component method. Only milk QTL that were chromosome-wise significant at 5% with the variance component method were considered in the multi-trait analysis.

For chromosomes where two fertility traits were analyzed together, indication of one pleiotropic QTL affecting both traits was seen on BTA1. A pleiotropic QTL affecting days open and fertility treatments was detected at 1% chromosome-wise significance with highest F-value at 148 cM. For chromosomes where a fertility trait and a milk trait were analyzed together, linked QTL affecting days open and milk yield were identified on BTA1 at 1% chromosome-wise significance with peak positions at 144 cM for days open and at 104 cM for milk yield. For the other analyzed chromosomes it was not possible to conclude if there were two linked QTL or one pleiotropic QTL affecting the traits.

4.2 Two-stage marker-assisted selection of dairy cattle

4.2.1 Gains in selection response due to MAS

The total genetic response consisted of two parts: the response due to the polygenes and the response due to the two QTL. It was based on the true breeding values of the best bulls selected after progeny test.

In general, the polygenic response was quite similar with MAS and the traditional selection scheme. The response was always lower when the QTL had large effects compared to the moderate size QTL and when the biallelic QTL had intermediate allele frequencies. With low starting frequencies of the favourable QTL alleles the polygenic response was lower with MAS and with higher starting frequencies the opposite was true. The same effect of QTL allele frequencies on the polygenic response was seen when selection was continued for four generations. With a QTL of moderate size at very low starting frequen-

cies (0.01) the polygenic response was still slightly lower with MAS after four generations of selection. The higher the starting frequencies were, the faster the MAS scheme became superior to conventional selection.

The combined response of the two QTL was higher with MAS than with traditional selection in all scenarios. With both selection schemes the response of the QTL was higher with the larger QTL effects and with medium starting frequencies of the alleles. The largest advantage of MAS compared to conventional selection with moderate QTL size was at starting frequencies of 0.10 when the difference was 0.22 SD_{pg} (polygenic standard deviation) in favour of MAS. With the larger QTL size the largest difference was 0.35 SD_{pg} in favour of MAS at starting frequencies of 0.05. When selection was continued for four generations, the response in the QTL was higher with MAS until the allele frequencies of the favourable alleles became very high. The difference between the schemes was greatest with QTL of large size and low (0.01) starting frequencies of the favourable alleles.

The total genetic response was always higher with MAS than with conventional selection. The difference between the schemes was greater with the larger size QTL. The largest difference in total genetic response was seen with starting frequencies of the favourable alleles at 0.10 when it was 0.37 SD_{pg} for a QTL with large effect and 0.22 SD_{pg} with a QTL of moderate effect. When selection was continued for four generations, MAS was always superior to conventional selection. The difference between the selection schemes increased over generations except for scenarios with large starting frequencies (≥ 0.30) where the difference finally decreased when the allele frequencies approached fixation. For example with the scenario of large QTL effects and starting frequencies of the favourable alleles at 0.05, the difference between the selection schemes was 0.48 SD_{pg} at generation 4.

4.2.2 Influence of MAS on linkage disequilibrium

The disequilibrium was calculated both between polygenes and QTL and between the two QTL. The disequilibrium between the polygenes and QTL was always negative and it was larger with larger QTL size. With lower initial allele frequencies the disequilibrium was slightly higher with MAS but with higher initial frequencies lower. When selection was continued for four generations, starting at low frequencies of the favourable allele with a QTL of medium size, the disequilibrium became increasingly negative until allele frequencies reached intermediate values when it started to become less negative. The MAS scheme resulted first in more negative disequilibrium but the disequilibrium decreased slightly faster with MAS than with conventional selection. For example with medium size QTL and starting frequencies of 0.02, the disequilibrium was -0.45 with MAS and -0.41 with conventional selection at genera-

tion 1 compared to -0.52 and -0.54 at generation 4. With QTL of large size the disequilibrium was always smaller with MAS except for the scenario with very low (0.01) initial allele frequencies.

The disequilibrium between the QTL was negative for both selection schemes with all scenarios except when the initial frequencies of the favourable alleles were very high (0.90) when there was no disequilibrium in the progeny tested bulls. The disequilibrium was larger with QTL of large effect and it was somewhat larger with MAS for scenarios with starting frequencies below 0.50 for QTL of moderate size and below 0.3 for large QTL. The disequilibrium had the largest values -0.08 with starting frequencies at 0.10 for moderate size QTL and -0.15 for large size QTL in the MAS scheme. When selection was continued for several generations, the disequilibrium was always negative. The MAS scheme resulted first in more negative values than the conventional scheme but later in less negative values until close to fixation of the favourable allele when the disequilibrium was close to zero in both schemes.

4.2.3 Changes in allele frequencies due to MAS

Selection increased the allele frequencies of the favourable alleles of both QTL. The increase was larger with the QTL of large effect. The frequencies were clearly increased in the young bulls compared to the base generation with both MAS and traditional selection where only parental selection was applied, but the increase with MAS was greater. After the final round of selection when the progeny test results were used, the allele frequencies still increased with both schemes and the values were higher with MAS. As an example when the initial allele frequencies were 0.10 in the base population, the frequencies in the bulls selected after progeny test were 0.51 when MAS was applied compared to 0.41 without MAS. When selection was continued for four generations it was seen that the increase was faster with larger QTL where the frequencies were over 0.70 in generation 4 for all starting frequencies when MAS was used. With all scenarios the increase was first faster with MAS and later when both schemes approached very high frequencies the increase with the traditional selection was somewhat faster until both schemes came close to fixation of the favourable alleles.

5 Discussion

5.1 Detection of QTL for economically important traits

5.1.1 QTL for milk production traits

QTL for milk production were suggested on 14 chromosomes in the single chromosome analysis at 5% or higher chromosome-wise significance (I). In the literature, QTL affecting milk production traits have been reported on all these chromosomes (e.g., Mosig et al., 2001; Khatkar et al., 2004). The positions of the QTL on the same chromosomes are varying between experiments, but the confidence intervals are likely to be large and therefore it is not possible to decide if these are separate or the same QTL. The two genome-wise QTL that were found were a QTL for fat percentage on BTA14 and a QTL for fat yield on BTA12. The QTL for fat percentage on BTA14 was located at the centromeric end of the chromosome and it is likely that this is the DGAT1 gene with a major effect on fat percentage and other milk composition traits detected by Winter et al. (2002) and Grisart et al. (2002). A QTL for fat yield has earlier been detected on BTA12 in one Holstein family closer to the distal end of the chromosome (Rodriguez-Zas et al., 2002).

When putative QTL on other chromosomes were used as cofactors in the analysis in order to increase statistical power, a total of 31 QTL were detected at 5% genome-wise significance. In most of the chromosomes the difference in the positions corresponding to the highest F-value was within 10 cM between the analyses with and without cofactors. For protein yield on BTA5, the difference was much higher (54 cM between positions 77 and 131). The position observed with the cofactor analysis is supported by the results from the variance component method in study III where a QTL for protein yield first was observed at 68 cM in the same sire families. The cofactor method increased the number of detected QTL substantially. However, these results should be taken with caution because it has been shown that the number of false positives can increase when cofactors are applied especially if the family size is small and the heritability of a trait is low (Sahana et al., 2006). In the present QTL mapping study the low heritability should not be of great concern because the reliability due to the progeny test is likely to be high. The family sizes are quite small however.

From the cofactor results it can be seen that several of the QTL seem to have an effect on milk yield and one or both of the content traits without affecting the solid yields (Table3). One probable explanation is that the QTL increases the amount of water and thus decreases the proportion of fat and/or protein. BTA12 and 25 are the only chromosomes with QTL for at least 2 yield traits and

1 content trait and the QTL on BTA14 seems to affect fat and protein contents without having an effect on milk yield.

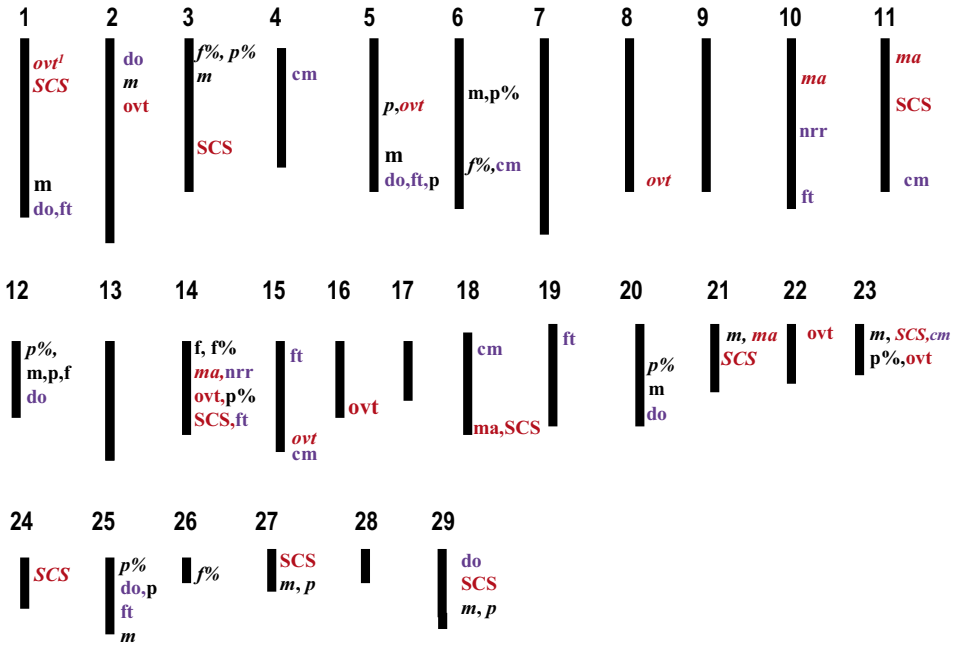
5.1.2 QTL for health traits

The analyzed health traits were SCS, clinical mastitis treatments and other veterinary treatments. SCS and mastitis treatments are both measuring udder health, but they are not the same trait. The genetic correlation between SCS and mastitis has been reported to be moderate to high, but not close to unity (e.g., Pösö and Mäntysaari, 1996; Rogers et al., 1998; Rupp and Boichard, 1999). Therefore, it is not expected that all QTL that are found for one of the traits should be detected for the other, too. However, when QTL for both traits are found at the same location, it gives more confidence in the QTL. In this study only 4 chromosomes (BTA11, 14, 18, and 21) had some evidence for QTL for both traits. The QTL affecting clinical mastitis and SCS were genome-wise significant in the cofactor analysis and located exactly at the same position on BTA18. These results are supporting the idea that SCS and clinical mastitis are measuring different aspects of udder health. Persisting high SCC levels are shown to mainly be a sign of subclinical mastitis which is most often caused by contagious bacteria such as Streptococcus aureus and Streptococcus agalactiae (de Haas et al., 2002). Incidences of clinical mastitis are most often caused by environmental bacteria such as Escherichia coli and in these infections the SCC levels increase rapidly but are soon dropping to normal level when the infection is cured. Therefore high SCC levels are not detected for these cows in many cases. There is also variation in the base level of SCC in healthy cows and there is some evidence that cows with very low SCC are more susceptible to clinical mastitis than cows with higher SCC levels (Schukken et al., 1998). However, there is also evidence for the opposite conclusion that cows with low SCC are always less likely to get clinical mastitis (Rupp and Boichard, 2000).

When the chromosomes were analyzed separately, a total of 12 QTL were suggested at 5% chromosome-wise significance: 6 for SCS, 1 for mastitis and 5 for other veterinary treatments (II). When cofactors were added to increase power, 17 genome-wise significant QTL were detected. Even though the cofactor approach is likely to produce many false positive results (Sahana et al., 2006), almost all of the QTL findings for SCS are located on chromosomes where QTL for SCS have been detected in other studies (e.g., Ashwell et al., 1997; Reinsch et al., 1998; Heyen et al., 1999; Klungland et al., 2001). Only on BTA24 no QTL for SCS has earlier been found. This increases the confidence in the cofactor results. The QTL for mastitis detected on BTA11 and 14 in the cofactor analysis are also supported by the fact that QTL for SCS were detected on the same chromosomes in the single chromosome analysis. Also Holmberg and Andersson-Eklund (2004) detected a mastitis QTL on BTA11, although at a different location.

Other veterinary treatments includes treatments for many different diseases. It is not possible to draw conclusions about which diseases the detected QTL are affecting. Because milk fever, ketosis, and retained placenta have been the most common registered disorders (Rautala, 2001) it is likely that many of the QTL may be associated withthese traits. The results for other veterinary treatments should thus be regarded as preliminary information that could be used in further studies focussing on specific disease traits. Holmberg et al. (2004) also analyzed other veterinary treatments and only on BTA15 a QTL was seen in both studies. The location of this QTL is close to the marker RM4.

Because the ultimate goal of QTL experiments is to find QTL that could be used in breeding schemes to improve the genetic merit of the animals, it is important to know if the detected health QTL are overlapping with QTL affecting milk traits. When comparing the results of this study with the milk QTL detected in (I), some chromosomes were seen to carry QTL for health traits and milk production traits (Figure 3). In most cases the positions were not overlapping,



[1]ovt = other veterinary treatments, SCS = somatic cell score, m = milk yield, do = days open,
ft = fertility treatments, f% = fat percentage, p% = protein percentage, cm = calf mortality,
p = protein yield, ma = clinical mastitis, nrr = non-return rate, f = fat yield
QTL in italics are detected only in the cofactor analysis.
Milk production traits are marked in black, health traits in red and fertility traits in purple.

Figure 3. Quantitative trait loci (QTL) detected in the whole genome scans of Finnish Ayrshire cattle. Numbers refer to the different chromosomes.

but we have to bear in mind that the precision of the most likely QTL positions are not likely to be very good. Multi trait analysis of these chromosomes would give more information about the location of these QTL.

5.1.3 QTL for fertility traits

In recent years the fertility of high yielding dairy cows has been declining especially among Holstein cattle (e.g., Royal et al., 2000; Washburn et al., 2002). One reason for this is the unfavourable genetic correlation between milk yield and fertility traits (e.g., Kragelund et al., 1979; Pösö and Mäntysaari, 1996). Some genes that affect milk production also affect reproduction or are linked to reproduction genes. It has become necessary to pay more attention to fertility in the breeding schemes. Because fertility traits have low heritability (e.g., Jansen et al., 1987; Weigel and Rekaya, 2000) and are therefore difficult to improve using traditional breeding, use of fertility QTL could bring new tools to improve the breeding strategies.

Out of the 22 fertility QTL exceeding the 5% chromosome-wise significance threshold in study III none were genome-wise significant and only 6 were chromosome-wise significant at 1%. When days open and calf mortality were analyzed using cofactors, two of the QTL for days open and four of the QTL for calf mortality reached genome-wise significance (Table 4). Typical for the fertility traits was that the peak positions of the detected QTL varied a lot between families. This may be due to the complexity of the traits or to the fact that there are more than one QTL affecting the same trait on the chromosomes. However, the results from the 2-QTL regression analyses indicated there being more than one QTL only on BTA1, 5, and 14 for days open. It is more difficult to detect QTL on chromosomes were the most likely QTL positions vary greatly between families.

When the data was re-analyzed with the variance component method, only three of the QTL previously found with the regression analysis were detected. The reason for this may be that the variance component method may not detect QTL with low allele frequency whereas the regression method does (de Koning et al, 2003). Also the peak positions of the test statistics seen with the variance component method were in most cases at the same position as with the regression analysis and when analysing only the segregating families with the variance component method, the same QTL were detected. This increases the confidence in the results from the regression analysis.

Multi-trait analyses were carried out in order to separate between one pleiotropic QTL and two linked QTL if there were two fertility trait QTL or one fertility trait QTL and one milk QTL on the same chromosome. Because of the low number of segregating families, the different peak positions of the test statistics

Table 4. Location and significance level (P-value) of QTL for days open (DO) and calf mortality (CM) when the mapping was done using cofactors.

BTA[1]	Trait	cM	P_{chr}[2]	P_{gen}[3]
1	DO	150	<0.01	ns
2	DO	2	<0.01	ns
4	CM	11	0.01	ns
5	DO	113	0.03	ns
11	CM	132	<0.01	0.01
12	DO	47	<0.01	0.01
15	CM	115	<0.01	<0.01
16	CM	16	<0.01	ns
18	CM	4	<0.01	0.03
23	CM	4	<0.01	<0.01
25	DO	54	<0.01	ns
29	DO	6	<0.01	0.04

[1]Bos taurus chromosome,

[2]P-value at chromosome-wise significance level,

[3]P-value at genome-wise significance level

in different families and the sparse maps, it was possible to separate the different models, pleiotropic or linked, in only two cases.

5.1.4 General discussion on the whole genome scans

Even though the power of the Finnish grand daughter design has not been quantified, it is probably not very high mostly because of relatively small family sizes (Weller et al., 1990; van der Beek et al., 1995). The average number of sons per bull was 41 ranging from 21 to 82. It would be better to have family sizes of 100 sons per bull, but unfortunately such large families were not available in the Finnish Ayrshire population. The power could also have been increased with genotyping more markers especially on the smaller chromosomes where only two or three markers were now typed. With few markers on a chromosome, some families are likely to have only one or even no informative markers on a chromosome even when highly polymorphic microsatellite markers are used. In this kind of low power studies the estimated QTL effects are likely to be overestimated because only effects of large effect would be detected and these large effect are more likely to be overestimated than correctly estimated or underestimated (Georges et al., 1995).

In the whole genome scans of milk, health and fertility traits of Finnish Ayrshire a large number of QTL were detected. With the low significance level (chromosome-wise 5%) of many of the found QTL, some of them are likely to be false positive results, but there is still a great number of QTL that could potentially be

used in the breeding scheme if MAS would be applied. Many of the milk QTL seem to affect the water content in the milk. There are some interesting QTL affecting milk production traits, for example the QTL affecting fat yield on BTA14 which probably is the DGAT1 gene (Grisart et al., 2004), and the QTL affecting fat yield on BTA12 and protein yield on BTA5, 12, 25 and possibly 27 and 29. Among the health QTL, the most interesting are those on BTA11, 14 and 18. On these chromosomes the QTL seem to affect both SCS and mastitis which gives confidence in the results. For the fertility traits, the most interesting QTL are those on BTA1, 5 and 25. On these chromosomes there are QTL for both days open and fertility treatments.

Multi-trait mapping was carried out only among the fertility traits, and the fertility and milk traits. Unfortunately the mapping method was able to separate between the pleiotropic and linked QTL models in only two cases mainly because of the data structure, i.e., few segregating families and few markers. However, before QTL can be used in MAS, it is necessary to know if they have unfavourable effects on some other traits and if a possible correlation is due to one pleiotropic QTL or two linked QTL. In the latter case the QTL could still be used in MAS. Therefore, it would be of interest to carry out multi-trait mapping also on the milk and health traits.

In order to use some of the QTL for breeding purposes it should be checked if they are segregating in the present Ayrshire population and especially among the present AI bulls. In the bull families of the Finnish grand daughter design most of the QTL were segregating only in few families and this is likely to be the case also in the present breeding population which may make the use of these QTL less attractive. Because the positions of the detected QTL are not accurate and the confidence intervals are likely to span almost the whole chromosome, fine mapping of the potentially interesting QTL is necessary before they can be used in MAS. In the near future large scale genotyping using SNP arrays, including tens of thousands or maybe even hundreds of thousands of SNP markers, will make a large amount of new genotypic data available. This data can be used to fine map the QTL areas detected in the whole genome scan and to find previously undetected QTL using LD or LDLA mapping methods (e.g., Meuwissen and Goddard, 2000; Meuwissen et al., 2002). With the use of high density markers and LD mapping methods it would not be necessary for the collected data to follow a family structure. Because the number of large half-sib families in Finnish Ayrshire is limited, this approach would make data collection easier.

5.2 Two-stage marker-assisted selection of dairy cattle

In the simulation studies (IV, V) the effects of using QTL information in selection of young bulls prior to progeny test were investigated. Especially the dif-

ferences in genetic response and disequilibrium between the different genetic components (polygenes and two QTL) were evaluated.

A two stage selection approach which uses MAS among young bulls and progeny records for final selection was chosen because this is the selection strategy best suited to dairy cattle, easiest to carry out and most often used in practice (e.g., Mackinnon and Georges, 1998; Druet et al., 2005). The scheme utilizes embryo transfer in order to increase the number of full-sib offspring and the selection is done between full brothers. In this way the Mendelian sampling part which is not utilized in parental selection can be used to increase selection intensity. The selection of polygenes is not affected by using the marker information at this stage. This kind of embryo transfer scheme can be applied in nucleus herds, but it can also be used when possible bull dams are located at different farms.

In these simulation studies MAS was started in a population in LD. This LD was created with one round of intense selection. In most other studies MAS has been started in populations in LE between polygenes and QTL (e.g., Gibson et al., 1990; Ruane and Colleau, 1996; Spelman and Garrick, 1997). There should be LD in the population if the QTL affects a trait that is under selection. However, it may have been more correct to do several generations of moderate selection instead of one intense selection step in order to create the LD. One intense selection step may have created too much LD. In reality LD is mostly created in populations of limited size by drift but also by selection.

Inbreeding was ignored for computational reasons. The rate of inbreeding would however probably be quite small, except for the QTL, because the animals in the base population were assumed to be unrelated and later mating among relatives was avoided. One of the assumptions made was that there was no recombination between the marker and the QTL and no error in the estimated QTL effect. This implies an optimal situation for MAS. By using markers further away from the functional mutation the advantage of MAS over traditional selection would decrease. If the estimated QTL effect would be larger than the real effect, the long-term response of MAS would be negatively affected (Spelman and van Arendonk, 1997). Also an incorrect estimate of the position of the QTL as well as selection for a nonexistent QTL would reduce the benefits from MAS (Spelman and van Arendonk, 1997). The assumption that the marker is the QTL was done in order to generate maximum disequilibrium. With the dense markers available at present and more QTL fine-mapped, this is a realistic assumption however. Four generations of selection was carried out. For the lower starting frequencies of the favourable allele this was quite short but with the higher starting frequencies the favourable alleles approached fixation even after four generations. Looking at the results of the scenarios with different starting frequencies together would give an idea of the outcome of selection over many gen-

erations. The different frequencies of the favourable alleles in the base generation used varied from very low (0.01) to high (0.90). The wide range of starting values in this study was used in order to get a good overview of the mechanisms of allele frequency changes on the variables. In practice QTL of very low frequencies would be hard to detect in the population.

There were only small differences in the disequilibrium between the QTL and the polygenes between MAS and the conventional selection scheme. This is not surprising since the MAS step of selection is not creating any disequilibrium. A small increase in LD at the second selection step using the MAS scheme can be due to the fact that more animals with the genotypes having the favourable alleles will be available for selection and the selected animals happen to have LD between the QTL and polygenes.

The total genetic response was faster with MAS than with conventional selection and the advantage of MAS persisted over several generations. The rate of response and the difference between the selection schemes reflected clearly the changes in allele frequencies of the favourable QTL. When MAS was used among the young bulls the QTL that were in low frequencies were detected more often than with conventional selection which lead to faster increase in QTL response. At this time the selection pressure was lower for the polygenes in the final selection with MAS because there were more animals with good QTL genotypes available for selection. Later when the QTL frequencies were higher with MAS, the selection operated more on the polygenes than in the conventional selection scheme where more of the selection pressure still was operating on the QTL. Clear advantage of MAS has been found also in other studies that applied MAS with a within-family approach (e.g., Spelman and Garrick, 1997; Gomez-Raya and Klemetsdal, 1999).

Several studies have reported that the response with MAS is faster in the first few generations, but later MAS results in lower gains compared to conventional selection (e.g., Gibson, 1994; Dekkers and van Arendonk, 1998). One explanation for the long-term loss has been the negative disequilibrium between polygenes and QTL due to selection (Gibson, 1994). In the present study negative disequilibrium was introduced by selection but the difference between the methods was small. Also other studies have shown that the negative disequilibrium is not causing the long-term loss seen with MAS in some cases (e.g., Ruane and Colleau, 1995; Dekkers and van Arendonk, 1998). Rather the so called 'Gibson effect' is caused by putting too little selection pressure on the polygenes in the beginning of the selection, whilst later the selection pressure is on the polygenes as the QTL become fixed (Dekkers and van Arendonk, 1998). The unequal selection pressure on the polygenes is caused by the changes in allele frequencies of the QTL which leads to chance in QTL variance over time and changes in selection pressure on the QTL. This means that a selection scheme

that fixes the QTL with least relaxation of the polygenic selection differential is optimal. In the present study (V) the Gibson effect was not seen and MAS was always ahead of conventional selection until the favourable alleles approached fixation and both schemes gave similar response. This is due to the within-family selection approach where the selection on QTL is done in a separate selection step and does not heavily affect the selection pressure of the polygenes. Selection was continued for only four generations, but using different starting frequencies would give an idea of the outcome of selection for more generations. The loss in long-term response has been seen when selection has been done using an index including different weights for QTL and polygenes (Gibson, 1994).

The results of these simulation studies show that using MAS together with embryo transfer to pre-select young bulls within families is a useful approach to increase the genetic merit of the AI-bulls compared to conventional selection. Even though a more appealing way to incorporate marker information in the breeding scheme might be to use them in an index together with the phenotypic information and using appropriate weighting to avoid long-term loss (Dekkers and van Arendonk, 1998), the two-stage approach is computationally less demanding and easier to implement. The results of these studies also give insight in how the selection applied with the different selection schemes change the allele frequencies of the favourable QTL alleles and how these changes are affecting the genetic response and disequilibrium between the polygenes and QTL.

Currently most of the research on using genetic markers in breeding schemes is focusing on genome-wide selection (Meuwissen et al., 2001). Genome-wide selection differs from MAS because it gets information from the whole genome by using dense SNP markers, whereas MAS uses only information of some of the QTL having an effect on the trait under selection.

6 Conclusions

The first aim of this study was to map quantitative trait loci (QTL) affecting traits that are important to the milk producers and dairy industry. These are the milk production traits (milk yield, fat yield, protein yield, fat percentage and protein percentage) that directly affect the income of the dairy farmer and are also important for the dairies, and the functional health and fertility traits (somatic cell score, mastitis treatments, other veterinary treatments, days open, fertility treatments, calf mortality and non-return rate) that affect the income through effects on the milk quality and costs of production in general. The functional traits are also related to sustainable and ethical milk production.

Several QTL affecting these economically important traits were detected. Among the milk production QTL, the most interesting are the QTL affecting milk and protein yield on BTA5, milk yield, fat yield, protein yield, and protein percentage on BTA12 and fat yield, fat percentage and protein percentage on BTA14. For the functional traits, a total of 34 QTL were detected, at chromosome-wise 5% significance, when the chromosomes were analyzed separately. Out of these 7 affected udder health, 5 treatments for other diseases, and 22 fertility traits. Almost all the milk production QTL were detected on chromosomes where also QTL for functional traits are located. In only few cases their locations are overlapping, but confidence intervals are likely to be large. Multi-trait analyses were carried out on the chromosomes where QTL for milk production traits and fertility traits are detected at overlapping positions. The results indicate that there seems to be two linked QTL on BTA1, one affecting milk yield and one days open. Because of the sparse maps and small population size, it was not possible in the other cases to distinguish between a pleiotropic QTL and two linked QTL.

It is potentially possible to use these QTL in the breeding program. Before this can be done further studies are needed however. The QTL should be verified in an independent sample and their segregation in the present breeding population should be determined as well as their effects and allele frequencies. Because it is important to know about correlated effects of QTL on other traits more multi-trait analyses should be done and using more and larger families and denser maps in order to achieve greater power. Even though markers further away from the functional mutation can be used in MAS, it would be of great importance to fine map the interesting regions in order to get markers in population wide linkage disequilibrium. The use of LD-markers would lead to much more efficient MAS.

The second aim of this thesis was to investigate the effects of using MAS in a breeding scheme where embryo transfer was implemented and marker information was used to select among full brothers the best candidates for progeny test-

ing. This was carried out as a stochastic simulation study. The MAS scheme was compared to a conventional scheme where one full brother would be chosen at random for the progeny test. Differences between the schemes in genetic response and linkage disequilibrium between the different parts of the genome were investigated. The size of the QTL effects and starting frequencies of the favourable allele were varied.

The results show that in general the total genetic response is higher with MAS both in the first generation of selection and in the long run when within family selection is applied. Both selection methods lead to negative disequilibrium between the QTL and the polygenes and the two QTL. Both the genetic response and disequilibrium are affected by the initial allele frequencies of the favourable alleles in the population and the changes in allele frequencies during later generations. MAS is increasing the allele frequencies of the favourable alleles faster in the early generations compared to traditional selection which explains the differences in the results seen between the methods. Using MAS together with embryo transfer to pre-select young bulls within families is a useful approach to increase the genetic merit of the AI-bulls compared to conventional selection.

7 Acknowledgements

First I want to thank Johanna Vilkki from MTT Agrifood Research Finland for being my supervisor and having patience during the time when my family always came first. It has been very pleasant to work with Johanna.

Asko Mäki-Tanila is thanked for introducing the subject of QTL mapping to me and encouraging me to work on it during my stay in Wisconsin. Asko is also thanked for encouraging me to write this thesis.

Margaret Dentine is warmly thanked for letting me work in her group at the University of Wisconsin-Madison and letting me participate in her lectures and seminars. I enjoyed working with Gretel.

I would like to thank the whole EGE team for a nice working environment and especially Sirja Viitala for good discussions. I would also like to thank Elina Laihonen, Pirkko Kallenautio and Outi Kasari at our office who always provide help when needed.

Special thanks are expressed to Dirk-Jan de Koning who has kindly answered hundreds of questions regarding the QTL mapping program.

Jukka Pösö is thanked for promptly providing all needed information from the Finnish Animal Breeding Association.

Mogens Lund is warmly thanked for the opportunity to visit Research Centre Foulum. Mogens is also especially thanked for the inspiration and energy that helped me finish this thesis.

Goutam Sahana is warmly acknowledged for enjoyable collaboration with the third paper.

Brian Kirkpatrick is thanked for introducing me to the practical QTL mapping in his lab and letting me follow his lectures during my stay in Wisconsin.

Matti Ojala from the University of Helsinki is thanked for being my professor and always helping with questions related to my studies.

Pekka Uimari and Theo Meuwissen are warmly thanked for reviewing this thesis.

The Finnish Academy is thanked for funding the completion of this thesis. Also the Finnish Cultural Foundation and Tiura Foundation are thanked for financial support.

Finally many warm thanks are expressed to my family Ismo, Isabella, Teresia and Jonas who have had patience with an absent and absent minded wife and mother during the last few months. I would also like to thank my parents, brother and sister for being there for me.

November 2007

Nina Schulman

8 References

Almasy, L. and Blangero, J., Multipoint quantitative-trait linkage analysis in general pedigrees, Am. J. Hum. Gen. 62 (1998) 1198-1211.

Andersson-Eklund, L. and Rendel, J., Linkage between amylase-1 locus and a major gene for milk fat content in cattle, Anim. Genet. 24 (1993) 101-103.

Ashwell, M.S., Rexroad Jr., C.E., Miller, R.H., VanRaden, P.M. and 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.

Barendse, W., Armitage, S.M., Kossarek, L.M., Shalom, A., Kirkpatrick, B.W., Ryan, A.M., Clayton, D., Li, L., Neibergs, H.L. and Zhang, N., A genetic linkage map of the bovine genome, Nat. Genet. 6 (1994) 227-235.

Bink, M.C.A. and van Arendonk, J.A.M., Detection of quantitative trait loci in outbred populations with incomplete marker data, Genetics 151 (1999) 409-420.

Bishop, M.D., Kappes, S.M., Keele, J.W., Stone, R.T., Sunden, S.L.F., Hawkins, G.A., Toldo, S.S., Fries,R., Grosz, M.D., Yoo, J. and Beattie, C.W., A genetic linkage map for cattle, Genetics 136 (1994) 619-639.

Blattman, A.N., Kirkpatrick, B.W., Gregory, K.E., A search for quantitative trait loci for ovulation rate in cattle, Anim. Genet. 27 (1996)157-162.

Boichard, D., Grohs, C., Bourgeois, F., Cerqueira, F., Faugeras, R., Neau, A., Rupp, R., Amigues, Y., Boscher, M.Y.and Leveziel, H., Detection of genes influencing economic traits in three French dairy cattle breeds, Genet. Sel. Evol. 35 (2003) 77-101.

Bolard, M. and Boichard, D., Use of maternal information for QTL detection in a (grand)daughter design, Genet. Sel. Evol. 34 (2002) 335-352.

Botstein, D., Wh, R.L., Skolnick, M. and Davis, R.W., Construction of a genetic linkage map in man using restriction fragment length polymorphisms, Am. J. Hum. Gen. 32 (1980) 314-331.

Bovenhuis, H. and 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.

Buitenhuis, A.J., Lund, M.S., Thomasen, J.R., Thomsen, B., Nielsen, H., Benn, C. and Guldbrandtsen, B., Detection of quantitative trait loci affecting lameness and leg conformation traits in Danish Holstein cattle, J. Dairy Sci. 90 (2007) 472-481.

Bulmer, M.G., The effect of selection on genetic variability, Am. Natur. 105 (1971) 201-211.

Churchill, G.A. and Doerge, R.W., Empirical threshold values for quantitative trait mapping, Genetics 138 (1994) 963-971.

Collins, A.C., Martin, I.C.A. and Kirkpatrick, B.W., Growth quantitative trait loci (QTL) on mouse chromosome 10 in a Quackenbush-Swiss x C57BL/6J backcross, Mamm. Genome 4 (1993) 454-458.

Coppieters, W., Kvasz, A., Arranz, J.J., Grisart, B., Riquet, J., Farnir, F. and Georges, M., The great-grand daughter design: a simple strategy to increase the power of a grand-daughter design for QTL mapping, Genet. Res. Camb. 74 (1999) 189-199.

Coppieters, W., Kvasz, A., Farnir, F., Arranz, J.J., Grisart, B., Mackinnon, M. and Georges, M., A rank-based nonparametric method for mapping quantitative trait loci in outbred half-sib pedigrees: Application to milk production in a granddaughter design, Genetics 149 (1998) 1547-1555.

Cowan, C.M., Dentine, M.R. and Coyle, T., Chromosome substitution effects associated with κ-casein and β-lactoglobulin in Holstein cattle, J. Dairy Sci. 75 (1992) 1097-1104.

Curtis, D., Response to genetic dissection of complex traits, Nature Gen. 12 (1996) 356-357.

de Haas, Y., Barkema, H.W. and Veerkamp, R.F., The effect of pathogen-specific clinical mastitis on the lactation curve for somatic cell count, J. Dairy Sci. 85 (2002) 1314-1323.

Dekkers, J.C.M., Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons, J. Anim. Sci. 82(E. Suppl.) (2004) E313-E328.

Dekkers, J.C.M. and van Arendonk, J.A.M., Optimizing selection for quantitative traits with information on an identified locus in outbred populations, Genet. Res. Camb. 71 (1998) 257-275.

de Koning, D.J., Conflicting candidates for cattle QTLs, Trends in genetics 22 (2006) 301-305.

de Koning, D.J., Janss, L.L.G., Rattink, A.P., van Oers, P.A.M., deVries, B.J., Groenen, M.A.M., van der Poel, J.J., de Groot, P.N., Brascamp E.W. and van Arendonk , J.A.M., Detection of quantitative trait loci for backfat thickness and intramuscular fat content in pigs (Sus scrofa), Genetics 152 (1999) 1679-1690.

de Koning, D.J., Pong-Wong, R., Varona, L., Evans, G.J., Giuffra, E., Sanchez, A., Plastow, G., Noguera, J.L., Andersson, L. and Haley, C.S., Full pedigree quantitative trait locus analysis in commercial pigs using variance components, J. Anim. Sci. 81 (2003) 2155-2163.

de Koning, D.J., Schulman, N.F., Elo, K., Moisio, S., Kinos, R., Vilkki, J. and Mäki-Tanila, A., Mapping of multiple quantitative trait loci by simple regression in half-sib designs, J. Anim. Sci. 79 (2001) 616-622.

Doerge, R.W. and Churchill, G.A., Permutation tests for multiple loci affecting a quantitative character, Genetics 142 (1996) 285-294.

Druet, T., Fritz, S., Boichard, D. and Colleau, J.J., Estimation of genetic parameters for quantitative trait loci for dairy traits in the French Holstein population, J. Dairy Sci. 89 (2005) 4070-4076.

Edderkaoui, B., Baylink, D.J., Beamer, W.G., Wergedal, J.E., Porte, R., Chaudhuri, A. and Mohan, S., Identification of mouse duffy antigen receptor for chemokines (Darc) as a BMD QTL gene, Genome Res. 17 (2007) 577-585.

Falconer, D.S. and Mackay, T.F.C., Introduction to quantitative genetics, 4th ed. Longman, Harlow, UK (1996) 436 p.

Fernando, R.L. and Grossman, M., Marker assisted selection using best linear unbiased prediction, Genet. Sel. Evol. 21 (1989) 467-477.

Geldermann, H., Investigations on inheritance of quantitative characters in animals by gene markers I. Methods, Theor, Appl. Genet. 46 (1975) 319-330.

Geldermann, H., Pieper, U. and Roth B., Effects of marked chromosome sections on milk performance in cattle, Theor. Appl. Genet. 70 (1985) 138-146.

George, A.W., Visscher, P.M. and Haley, C.S., Mapping quantitative trait loci in complex pedigrees: A two –step variance component approach, Genetics 156 (2000) 2081-2092.

Georges, M., Nielsen, D., Mackinnon, M., Mishra, A., Okimoto, R., Pasquino, A.T., Sargeant, L.S., Sorensen, A., Steele, M.R. and Zhao X., Mapping quantitative trait loci controlling milk production in dairy cattle exploiting progeny testing, Genetics 139 (1995) 907-920.

Gibson, J.P., Short term gain at the expense of long-term response with selection of identified loci, Proc. 5th World Congress Genet. Appl. Livest. Prod. 21 (1994) 201-204.

Gibson, J.P., Jansen, G.B. and Rozzi, P., The use of κ-casein genotypes in dairy cattle breeding, Proc. 4th World Congress Genet. Appl. Livest. Prod. 14 (1990) 163-166.

Gomez-Raya, L. and Klemetsdal, G., Two-stage selection strategies utilizing marker-quantitative trait locus information and individual performance, J. Anim. Sci. 77 (1999) 2008-2018.

Green, P., Falls, K. and Crooks, S., Documentation of CRIMAP, version 2.4, Washington University School of Medicine, St Louis, MO, 1990.

Grignola, F.E., Hoeschele, I. and Tier, B., Mapping quantitative trait loci in outcross populations via residual maximum likelihood. I. Methodology, Genet. Sel. Evol. 28 (1996) 479-490.

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges, M. and 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.

Grisart, B., Farnir, F., Karim, L., Cambisano, N., Kim, J.J., Kvasz, A., Mni, M., Simon, P., Frère, J.M., Coppieters, W. and Georges, M., Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition, PNAS. 101 (2004) 2398-2403.

Haley, C. and Andersson, L., Linkage mapping of quantitative trait loci in plants and animals, In: Hames, B.D. ed., Genome mapping A practical approach, New York, USA,Oxford university press (1997) 49-71.

Haley, C.S., Knott, S.A. and Elsen, J.M., Mapping quantitative trait loci in crosses between outbred lines using least squares, Genetics 136 (1994) 1195-1207.

Hayes, B. and Goddard, M.E., The distribution of the effects of genes affecting quantitative traits in livestock, Genet. Sel. Evol. 33 (2001) 209-229.

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

Hiendleder, S., Thomsen, H., Reinsch, N., Bennewitz , J., Leyhe-Horn, B., Looft, C., Xu, N., Medjugorac, I., Russ, I., Kühn, C., Brockmann, G.A., Blümel, J., Brenig, B., Reinhardt, F., Reents, R., Averdunk, G., Schwerin, M., Förster, M., Kalm, E. and Erhardt, G., Mapping of QTL for body conformation and behavior in cattle, J. Hered. 94(2003) 496-506.

Hill, W.G. and Robertson, A., Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38 (1968) 226-231.

Hoeschele, I. and vanRaden, P.M., Bayesian analysis of linkage between genetic markers and quantitative trait loci. I. Prior knowledge, Theor. Appl. Genet. 85 (1993) 953-960.

Holmberg, M. and Andersson-Eklund, L., Quantitative trait loci affecting health traits in Swedish dairy cattle, J. Dairy Sci. 87 (2004) 2653-2659.

Holmberg M. and Andersson-Eklund, L., Quantitative trait loci affecting fertility and calving traits in Swedish dairy cattle, J. Dairy Sci. 89 (2006) 3664-3671.

Hospital, F. and Chevalet, C., Interactions of selection, linkage and drift in the dynamics of polygenic characters, Genet. Res. Camb. 67 (1996) 77-87.

Hospital, F., Moreau, L., Lacoudre, F., Charosset, A. and Gallais, A., More on the efficiency of marker-assisted selection, Theor. Appl. Genet. 95 (1997) 1181-1189.

Hu, F.Z., Donfack, J., Ahmed, A., Dopico, R., Johnson, S., Post, J.C., Ehrlich, G.D. and Preston, R.A., Fine mapping a gene for pediatric gastroesophageal reflux on human chromosome 13q14, Hum. Genet. 114 (2004) 562-572.

Jansen, R.C., Interval mapping of multiple quantitative trait loci. Genetics 135 (1993) 205-211.

Jansen, R.C. and Stam, P., High resolution of quantitative traits into multiple loci via interval mapping, Genetics 136 (1994) 1447-1455.

Jansen, J., van der Werf, J. and de Boer, W., Genetic relationships between fertility traits for dairy cows in different parities, Livest. Prod. Sci. 17 (1987) 337-349.

Jiang, C., Zeng, Z.B., Multiple trait analysis of genetic mapping for quantitative trait loci, Genetics 140 (1995) 1111-1127.

Kashi, Y., Hallerman, E. and Soller, M., Marker-assisted selection of candidate bulls for progeny testing programmes, Anim. Prod. 51 (1990) 63-74.

Kass, R.E. and Raftery, A.E., Bayes factors, J. of Am. Stat. Assoc. 90 (1995) 773-795.

Khatkar, M.S., Thomson, P.C., Tammen, I. and Raadsma, H.W., Quantitative trait loci mapping in dairy cattle: review and meta-analysis, Genet. Sel. Evol. 36 (2004) 163-190.

Klungland, H., Sabry, A., Heringstad, B., Olsen, H.G, Gomez-Raya, L. Våge, D.I., Olsaker, I., Ødegård, J., Klemetsdal, G., Schulman, N.F., Vilkki, J., Ruane, J., Aasland, M., Rønningen, K. and Lien, S., Quantitative trait loci affecting clinical mastitis and somatic cell count in dairy cattle, Mamm. Genome 12 (2001) 837-842.

Knott, S.A. and Haley, C.S., Multi trait least squares for quantitative trait loci detection, Genetics 156 (2000) 899-911.

Knott, S.A., Elsen, J.M. and Haley, C.S., Methods for multiple-marker mapping of quantitative trait loci in half-sib populations, Theor. Appl. Genet. 93 (1996) 71-80.

Kragelund, K., Hillel, J. and Kalay, D., Genetic and phenotypic relationship between reproduction and milk production, J. Dairy Sci. 62 (1979) 468-474.

Kühn, C., Bennewitz, J., Reinsch, N., Xu, N., Thomsen, H., Looft, C., Brockmann, G.A., Schwerin, M., Weimann, C., Hiendleder, S., Erhardt, G., Medjugorac, I., Förster, M., Brenig, B., Reinhardt, F., Reents, R., Russ, I., Averdunk, G, Blümel, J. and Kalm, E., Quantitative trait loci mapping functional traits in the German Holstein cattle population, J. Dairy Sci. 86 (2003) 360-368.

Lander, E.S. and Botstein, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps, Genetics 121 (1989) 185-199.

Lander, E. and Kruglyak, L., Genetic dissection of complex traits: guidlines for interpreting and reporting linkage results, Nature Gen. 11 (1995) 241-247.

Liu, P., Wang, Y., Vikis, H., Maciag, A., Wang, D., Lu, Y., Liu, Y. and You, M., Candidate lung tumor susceptibility genes identified through whole-genome association analyses in inbred mice, Nature Gen. 38 (2006) 888-895.

Lynch, M. and Walsh, J.B., Genetics and analysis of quantitative traits, 1st edn., Sinauer Associates, Sunderland, USA (1998) 980 p..

Mackinnon, M.J. and Georges, M.A.J., Marker-assisted preselection of young dairy sires prior to progeny-testing, Livest. Prod. Sci. 54 (1998) 229-250.

Mackinnon, M.J. and Weller, J.I., Methodology and accuracy of estimation of quantitative trait loci parameters in a half-sib design using maximum likelihood, Genetics 141 (1995) 755-770.

Mehrabian, M., Allayee, H., Stockton, J., Lum, P.Y., Drake, T.A., Castellani, L.W., Suh, M., Armour, C., Edwards, S., Lamb, J., Lusis, A.J. and Schadt, E.E., Integrating genotypic and expression data in a segregating mouse population to identify 5-lipoxygenase as a susceptibility gene for obesity and bone traits, Nature Gen. 37 (2005) 1224-1233.

Meuwissen, T.H.E. and Goddard, M.E., The use of marker haplotypes in animal breeding schemes, Genet. Sel. Evol. 28 (1996) 161-176.

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

Meuwissen, T.H.E., Hayes, B.J. and Goddard, M.E., Prediction of total genetic value using genome-wide dense marker maps, Genetics 157 (2001) 1819-1829.

Meuwissen, T., Karlsen, A., Lien, S., Olsaker, I., Goddard, M.E., Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping, Genetics 75 (2002) 15-24.

Moisio, S., Elo, K., Kantanen, J. and Vilkki, J., Polymorphism within the 3'flanking region of the bovine growth hormone receptor gene, Anim. Genet. 29 (1998) 55-57.

Mosig, M.O., Lipkin, E., Khutoreskaya, G., Tchourzyna, E., Soller, M. and 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 adjusted false discovery rate criterion, Genetics 157 (2001) 1683-1698).

Neimann-Sørensen, A. and Robertson, A., The Association between blood groups and several production characteristics in three Danish cattle breeds. Acta Agric. Scand. 11 (1961) 163-196.

Ott, J., Analysis of human genetic linkage, Rev. ed., Baltimore, London, The John Hopkins university press (1991) 302 p.

Peippo, J., Viitala, S., Virta, J., Räty, M., Tammiranta, N., Lamminen, T., Aro, J., Myllymäki, H. and Vilkki, J., Birth of correctly genotyped calves after multiplex marker detection from Bovine embryo microblade biopsies, Molec. Repr. Devel. (2007) April 17 (Epub ahead of print).

Piepho, H.P., A quick method for computing approximate threshold for quantitative trait loci detection, Genetics 157 (2001) 425-432.

Pösö, J. and Mäntysaari, E.A., Genetic relationships between reproductive disorders, operational days open and milk yield, Livest. Prod. Sci. 46 (1996) 41-48.

Rautala, H., Terveystarkkailun tulokset 2000, Nauta 31 (2001) 22-23.

Reinsch, N., Xu, N. Thomssen, H., Looft, C., Kalm, E., Grupe, S., Kuhn, C., Schwerin, M., Leyhe, B., Hiendleder, S., Erhard,G., Medjugorac, I., Russ, I., Foster, M., Brenig, B., Reents, R. and Averdunk, G., First results on somatic cell count loci from the ADR bovine mapping project, Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia 26 (1998) 426-428.

Rodriguez-Zas, S.L., Southey, B.R., Heyen, D.W. and 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.

Rogers, G.W., Banos, G., Sander Nielsen, U. and J. Philipsson, Genetic correlations among somatic cell score, productive life, and type traits from the United States and udder health measures from Denmark and Sweden, J. Dairy Sci. 81 (1998) 1445-1453.

Ron, M., Band, M., Yanai, A. and Weller, J.I., Mapping quantitative trait loci with DNA microsatellites in a commercial dairy cattle population. Anim. Genet. 25 (1994) 259-264.

Royal, M.D., Darwash, A.O., Flint, A.P.F., Webb, R., Wooliams, J.A. and Lamming, G.E., Declining fertility in dairy cattle: changes in traditional and endocrine parameters of fertility, Anim. Sci. 70 (2000) 487-501.

Ruane, J. and Colleau, J.J., Marker assisted selection for genetic improvement of animal populations when a single QTL is marked, Genet. Res. 66 (1995) 71-83.

Ruane, J. and Colleau, J.J., Marker assisted selection for a sex-limited character in a nucleus breeding population, J. Dairy Sci. 79 (1996)1666-1678.

Rupp, R. and Boichard, D., Genetic parameters for clinical mastitis, somatic cell score, production, udder type traits, and milking ease in first lactation Holsteins, J. Dairy Sci. 82 (1999) 2198-2204.

Rupp, R. and Boichard, D., Relationship of early first lactation somatic cell count with risk of subsequent clinical mastitis, Livest. Prod. 62 (2000) 169-180.

Sahana, G., de Koning, D.J., Guldbrandtsen, B., Sørensen, P., Lund, M.S., The efficiency of mapping of quantitative trait loci using cofactor analysis in half-sib design. Genet. Sel. Evol. 38 (2006) 167-182.

Sax, K., Association of size differences with seed-coat patterns and pigmentation in Phaseolus vulgaris, Genetics 8 (1923) 552-560.

Schaeffer, L.R., Strategy for applying genome-wide selection in dairy cattle, J. Anim. Breed. Genet. 123 (2006) 218-223.

Schrooten, C., Bovenhuis, H., Coppieters, W. and Arendonk, J.A., Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle, J. Dairy Sci. 8 (2000) 795-806.

Schrooten, C., Bovenhuis, H., van Arendonk, J.A.M. and Bijma, P., Genetic progress in multistage dairy cattle breeding schemes using genetic markers, J. Dairy Sci. 88 (2005) 1569-1581.

Schulman, N. F., Viitala, S.M., de Koning, D.J., Virta, J., Mäki-Tanila, A. and Vilkki J.H, Quantitative trait loci for health traits in Finnish Ayrshire cattle, J. Dairy Sci. 87 (2004) 443-449.

Smith, C. and Simpson, P., The use of genetic polymorphisms in livestock improvement, J. Anim. Breed. Genet. 103 (1986) 205-217.

Soller, M., Mapping quantitative trait loci affecting traits of economic importance in animal populations using molecular markers. In: Eds.L.B. Schook, H.A. Lewin and D.G. McLaren, Gene mapping techniques and applications, Univ. of Illinois, Illinois, Marcel Dekker Inc., New York, USA (1991) 21-49.

Soller, M., Marker assisted selection – an overview, Anim. Biot. 5 (1994) 193-207.

Soller, M. and Beckmann, J.S., Genetic polymorphisms in varietal identification and genetic improvement, Theor. Appl. Genet. 67 (1983) 25-33.

Sonstegard, T.S. and van Tassell, C.P. Bovine genomics update: making a cow jump over the moon, Genet. Res. Camb. 84 (2004) 3-9.

Sørensen, P., Lund, M.S., Guldbrandtsen, B., Jensen J. and Sorensen D., A comparison of bivariate and univariate QTL mapping in livestock populations, Genet. Sel. Evol. 35 (2003) 605-622.

Spelman, R. and Bovenhuis, H., Moving from QTL experimental results to the utilization of QTL in breeding programmes, Anim. Genet. 29 (1998) 77-84.

Spelman, R.J., Coppieters, W., Karim, L., van Arendonk, J.A. and 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.

Spelman, R.J. and Garrick, D.J., Genetic and economic responses for within-family marker-assisted selection in dairy cattle breeding schemes, J. Dairy Sci. 81 (1997) 2942-2950.

Spelman, R.J. and van Arendonk, J.A.M., Effect of inaccurate parameter estimates on genetic response to marker-assisted selection in an outbred population, J. Dairy Sci. 80 (1997) 3399-3410.

Stella, A., Lohuis, M.M., Pagnacco, G. and Jansen, G.B., Strategies for continual application of marker-assisted selection in an open nucleus population, J. Dairy Sci. 85 (2002) 2358-2367.

Szyda, J., Liu, Z., Reinhardt, F. and Reents, R., Estimation of quantitative trait loci parameters for milk production traits in German Holstein dairy cattle population, J.Dairy Sci 88 (2005) 356-367.

Uimari, P., Thaller, G. and Hoeschele, I., The use of multiple markers in a Bayesian method for mapping quantitative trait loci, Genetics 143 (1996) 1831-1842.

van der Beek, S., van Arendonk, J.A.M. and Groen, A. F., Power of two- and three-generation QTL mapping experiments in an outbred population containing full-sib or half-sib families, Theor. Appl. Genet. 91 (1995) 1115-1124.

Velmala, R. J., Vilkki, J.H., Elo, K.T., de Koning, D.J. and Mäki-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.

Velmala, R., Vilkki, J., Elo, K. and Mäki-Tanila, A., Casein haplotypes and their association with milk production traits in the Finnish Ayrshire cattle, Anim. Genet. 26 (1995) 419-425.

Viitala, S.M., Schulman, N.F., de Koning, D.J., Elo, K., Kinos, R., Virta, A., Virta, J., Mäki-Tanila, A. and Vilkki, J.H., Quantitative trait loci affecting milk production traits in Finnish Ayrshire dairy cattle, J. Dairy Sci. 86 (2003)1828-1836.

Viitala, S., Szyda, J., Blott, S., Schulman, N., Lidauer,M., Mäki-Tanila, A., Georges, M. and Vilkkki, J., The role of the bovine growth hormone receptor and prolactin receptor genes in milk, fat and protein production in Finnish Ayrshire dairy cattle, Genetics 173 (2006) 2151-2164.

Vilkki, J.H., de Koning, D.J., Elo, K., Velmala, R. and Mäki-Tanila, A., Multiple marker mapping of quantitative trait loci of Finnish dairy cattle by regression, J. Dairy Sci. 80 (1997)198-204.

Washburn, S.P., Silvia, W.J., Brown, C.H., McDaniel, B.T. and McAllister, A.J., Trends in reproductive performance in southeastern Holstein and Jersey DHI herds, J. Dairy Sci. 85 (2002) 244-251.

Wayne, M.L. and McIntyre, L.M., Combining mapping and arrying: An approach to candidate gene identification, PNAS 99 (2002) 14903-14906.

Weber, J.L. and May, P.E., Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction, Am. J. Hum. Genet. 44 (1989) 388-396.

Weigel, K.A. and Rekaya, R., Genetic parameters for reproductive traits of Holstein cattle in California and Minnesota, J. Dairy Sci., 83 (2000) 1072-1080.

Weller, J.I., Maximum likelihood techniques for the mapping and analysis of quantitative trait loci with the aid of genetic markers, Biometrics 42 (1986) 627-640.

Weller, J.I., Kashi, Y. and Soller, M., Power of daughter and granddaughtes designs for determining linkage between marker loci and quantitative trait loci indairy cattle, J. Dairy Sci. 73 (1990) 2525-2537.

Winkelman, D.C. and Hodgetts, R.B., RFLPs for somatotropin genes identify quantitative trait loci for growth in mice, Genetics 131 (1992) 929-937.

Winter, A., Kramer, W., Werner, F.A.O., Kollers, S., Kata, S., Durstewitz, G., Buitkamp, J., Womack, J.E., Thaller, G. and 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. 99 (2002) 9300-9305.

Xu, S. and Atchley, W.R., A random model approach to interval mapping of quantitative trait loci, Genetics 141 (1995) 1189-1197.

Zeng, Z.B., Precision mapping of quantitative trait loci, Genetics 136 (1994) 1457-1468.

Zhang, Q., Boichard, D., Hoeschele, I., Ernst, C., Eggen, A., Murkve, B., Pfister-Genskow, M., Witte, L.A., Gringola, F.E., Uimari, P., Thaller, G. and 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.

Zhang, W.K., Wang, Y.J., Luo, G.Z., Zhang, J.S., He, C.Y., Wu, X.L., Gai, J.Y. and Chen, S.Y., QTL mapping of ten agronomic traits on the soybean (Glycine max L. Merr.) genetic map and their association with EST markers, Theor. Appl. Genet. 108 (2004) 1131-1139.

Agrifood Research Reports 109

109

Quantitative trait loci detection and benefits from marker-assisted selection in dairy cattle

MTT
Agrifood Research Finland