Muncie, Sarah Alissa. Fine Mapping Quantitative Trait Loci Affecting Health and Reproduction in US Holstein Cattle on Chromosome 18. (Under the direction of Dr. Melissa Ashwell)

Continued genetic improvement for milk production has been associated with decreased fertility in US Holstein cattle. A previous study (Ashwell et al., 2004) identified a putative quantitative trait locus (QTL) affecting daughter pregnancy rate at 54 cM on chromosome 18 in one Holstein grandsire family. The goal of this research is to determine the validity of the putative QTL using additional markers and an extended pedigree. Thirteen microsatellite markers located throughout the chromosome were genotyped in 973 animals that were descendants of the original grandsire in which the QTL was identified. Sons of the grandsire as well as six grandson and six great-grandson families of the original grandsire were selected for this study (range of 16 to 169 sons per family). In analysis of the sons using QTL Express, the same putative QTL affecting daughter pregnancy rate was detected and placed at 45 cM. In a joint analysis of thirteen of the largest families, each containing 10 or more sons, a significant QTL for daughter pregnancy rate was detected at 27 cM. QTL affecting daughter pregnancy rate was detected in two additional sub-families (Family II-5 and Family III-2), indicating a putative QTL affecting daughter pregnancy rate is most likely segregating within this pedigree. Across-family analysis also detected putative QTL on chromosome 18 affecting productive life at 35 cM, somatic cell score at 33 cM and percent difficult births at 72 cM. Analysis of individual families identified eight significant putative QTL and six suggestive putative QTL at the chromosome-wise level affecting somatic cell score, productive life, calving ease, percent difficult births, milk yield, fat yield, protein yield and fat percent. A complex pedigree analysis is underway to make full use of statistical

power to refine the QTL region affecting fertility. Further verification of the QTL effects identified in this study will allow identification of positional candidate genes to be applied in marker assisted breeding programs.
(**Key words**: quantitative trait locus, pregnancy rate, *Bos taurus* chromosome 18)

# Fine Mapping Quantitative Trait Loci Affecting Health and Reproduction in US Holstein Cattle on Chromosome 18

by

Sarah Alissa Muncie

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Department of Animal Science

Raleigh

2005

**Approved by:**

_____

Dr. Melissa Ashwell
Chair of Advisory Committee

_____ _____

Dr. Eugene Eisen Dr. Dahlia Nielsen

## Dedication

To my mother Mary will all my love, whose strength and courage inspire me to persevere

## Personal Biography

Sarah Alissa Muncie was born on May 7, 1980 in Crofton, Maryland, the second child of Herbert Lee Muncie, Jr. and Mary Elizabeth Muncie. Sarah and her brother Eric were raised in the suburban county of Severna Park, MD where they grew up in a loving environment and taught from an early age to never stop striving for the goals they set out to achieve. After graduating from Archbishop Spalding High School in 1998, she went to further her education at a university. In 2002, Sarah graduated from University of Delaware Cum Laude with a Bachelors of Science in animal science with a concentration in pre-veterinary medicine. After graduation, she decided to take a year off to travel around Europe to further decide her career path and to experience other cultures. With careful consideration she decided to attend North Carolina State University to pursue a Master's degree in animal science with a concentration in animal genetics. Following graduation in summer of 2005, she plans on seeking employment in a career that allows her to utilize her scientific knowledge to better the community and help others to learn the wonders of science.

**Acknowledgements**

I would like to thank the many wonderful people who helped guide and support me throughout the completion of this degree. This would never have been possible without them and my appreciation is endless. I would like to express my gratitude to my committee members. Dr. Melissa Ashwell, thank you for your patience, guidance, and tireless use of the hot pink pen. Dr. Eugene Eisen and Dr. Dahlia Nielsen, thank you both for your contributions to this project.

I would also like to express special thanks to Audrey O'Nan and Shelly Nolin for their technical assistance in the laboratory. Thank you for answering all of my questions. A special mention to Maureen Valentine, her contribution to the project was commendable.

There are many graduate students whose friendship will forever be cherished. First, Alaina Sauvé, thank you for everything! I was so lucky to have such a caring and wonderful friend throughout this crazy journey. I am grateful beyond words. Shannon Davidson, your boundless enthusiasm is inspiring. Last but not least, thank you to my fellow genetics girls, Keri Boyette and Jan-Marie Bender. I could not have asked for better study partners.

Finally, a very special thank you goes to my mother. I would never have survived without your "cheerleading", wisdom and love. Thank you for believing in me. Also, to my father, thank you for your support and love in my pursuit of a graduate degree and to my brother Eric, for being a great big brother.

*"The stars are always there, even in the daylight.*
*Sometimes we just can't see them."-Marian Keyes*

## Table of Contents

## List of Tables

## List of Figures

# Literature Review

## United States Dairy Cattle Economics

Milk is a very important farm product and ranks second only to beef within livestock industries (Miller, 2004). In 2000, the United States farm milk sales were approximately 10% over all other farm sales, reaching over $21 billion in revenue (Collins, 2000). Dairy products are vast including fluid milk, cheese, yogurt, butter and ice cream. Other products include dry and condensed milk and whey products, often found in processed foods. All 50 states produce milk products, with California and Wisconsin being the biggest producers. Milk production per cow has reached 19,000 pounds per year in the year 2000, a huge increase over the 9,700 pounds per year average during the 1970's (Miller, 2004). Although milk production has increased per cow during this time period, the number of dairy operations in the United States has decreased from 650,000 to 90,000 during the same time period (Miller, 2004). With more milk per cow, there are fewer cows required to produce the same amount of milk for profit. While the increases in production have been positive, negatively correlated effects have also been observed.

## Dairy Cattle Reproductive Trends

Continued genetic improvement for milk production has been associated with decreased fertility in US Holstein cattle. The lactation cycle is triggered by pregnancy; therefore, milk production is dependent on the cow becoming pregnant (Lucy, 2001). During the last 10 years, milk production has increased by approximately 20% in US dairy cows (Lucy, 2001). Increased milk production is associated with decreased indices for reproductive efficiency, leading to a negative association between reproduction and milk production (Lucy, 2001). In the past, conception rates using artificial insemination methods

in the United States have been reported at approximately 55% (Casida, 1961). Recently, these rates have decreased, ranging from 45% using natural estrus (Darnsfield et al., 1998), to 35% using controlled artificial insemination (Schmitt et al., 1996). The United States is not the only country seeing a decline in dairy cattle reproductive efficiency. The United Kingdom (Royal et al., 2000a), Australia (Macmillan et al., 1996), and Ireland (Roche, 2000) are examples of other countries also experiencing diminishing conception rates.

In the past, genetic selection has been primarily directed towards higher milk production. This has been accomplished by the use of artificial insemination (AI) in most dairy herds. AI allows the elite sires' genetics to be spread globally. The breeding values of elite sires, defined as the value of an individual as a genetic parent, are dependably measured by progeny testing. This is accomplished by evaluating the phenotypic performance of the elite sires' offspring. In dairy cattle progeny testing, new sires resulting from planned matings of superior parents are evaluated based on their daughters' average phenotypic performance (Georges et al., 1995). The animals are then selected for the breeding herd based on these evaluations. Progeny testing can be time consuming, and the costs can be prohibitory. The results of these tests would take an average of three and a half years; thus, the bull would already be approximately 5 years old at the completion of the testing (http://www.aipl.arsusda.gov/publish/presentations/MISC04/Expo_cvt.ppt#1). Also, only about 1 in every 9 bulls successfully passes progeny testing. The testing can cost as much as $25,000 per bull, and this is a substantial loss when bulls do not pass (http://www.aipl.arsusda.gov/publish/presentations/MISC04/Expo_cvt.ppt#1). Even with these unfavorable characteristics of progeny testing, this form of selection has led to the genetic improvement of highly heritable traits (Lucy, 2001).

An increase in milk production is not necessarily associated with decreased fertility and health problems, but higher producing dairy cattle are at the greatest risk for problems (Windig et al., 2005). Fertility is a complex, polygenic trait, influenced by genetics and environment. Decreased fertility could also be due to environmental factors, such as poor management practices, and may be improved through better nutrition and/or estrus detection (Wall et al., 2003). With an improved management method, there may be an increase in costs due to the additional labor requirements; therefore, a management approach of improving fertility may be difficult to maintain (Wall et al., 2003). Upholding high reproductive efficiency in a dairy cattle herd is of utmost importance to today's dairy producers.

The economic importance of fertility becomes evident if one considers that 40% of costs due to health problems are accredited to aspects of reproduction (Shanks et al., 1981). A 500-cow herd producer can incur losses of around $200,000 a year due to reproductive problems (http://www.100daycontract.com). Ignoring decreased fertility will only accentuate the problem. To maintain a successful breeding program, a producer needs to establish an economic equilibrium between costs of high numbers of cows culled from the herd and the economic loss of keeping a less reproductively sound cow (Esslemont, 1993). Culling is defined as the removal of animals from a herd (Monti et al., 1999). Reproduction is the most prevalent reason for culling; however mastitis and low milk production are also incentives for removal from the herd (Bascom and Young, 1998). A decision to cull a cow is considered voluntarily or involuntarily. If the animal is removed from the herd due to low milk production or mastitis, it is considered a voluntarily cull. The animal is removed at the producer's discretion. However, an involuntary cull is the removal of an animal due to

problems with establishing pregnancy or death. The inability of an animal to conceive explains 40% of total culls during the first lactation period (Esslemont and Kossaibati, 1997). When producers make the judgment to cull a cow, a wide variety of factors are included in the analysis; e.g., reproduction, milk production, or disposition (Beaudeau, 1995). A producer must also consider the economic outlook at the time of culling. It might not be wise to cull a cow when the price of culling is less then the cost of replacing the heifers (McCullough and Delorenzo, 1996).

The major question plaguing producers today is whether infertility is inherent to the cow or a result of management practices. Inbreeding occurring in dairy cattle today could also be another factor influencing the decreased reproductive efficiency. In Holstein cattle, inbreeding has been continuously increasing, changing from close to zero in the 1960's to around 5% at the start of the 21$^{st}$ century (Lucy, 2001). In a study of Guernsey cattle, Hermas et al. (1987) found that days open increased by 2 days and conception rate decreased by approximately 4% when inbreeding increased by only 1%.

Regardless of the cause of fertility problems arising in the dairy cattle industry, the overall effect will undoubtedly have negative repercussions on the dairy industry's success (Lucy, 2001). A balance between selection for milk yield and fertility needs to be considered in a breeding scheme to optimize the overall net merit of the animal (Wall et al., 2003). Net merit, as defined by Lush (1948), is the total effect every gene has on all traits a breeder would deem important. Breeders can use net merit to determine a selection objective. It can then be used to measure the advancement towards the predetermined objective (VanRaden, 2004b). With a clear selection objective, a producer can then take the necessary steps to reach this goal.

**Physiology of Reproduction**

A dairy cow is successful when she is healthy, has high milk production and is reproductively efficient. The length of a typical estrus cycle for dairy cattle is 21 days, ranging anywhere from 18-24 days. It takes about 30-32 hours after the onset of estrus for ovulation to occur, and in dairy cattle the ovulation actually occurs after estrus has ended (Dukes et al., 1993). A preovulatory gonadotropin surge initiates ovulation, which occurs with the beginning of sexual maturity, peaks within 6 hours, and is usually complete in approximately 11 hours (Dukes et al., 1993).

An animal must possess two key factors to attain pregnancy--a healthy ovum that can easily be fertilized and a uterus that can sustain the pregnancy after fertilization has occurred (Darwash et al., 1999). A feedback system controls the development of the follicles and ovarian function by means of hormones such as gonadotrophin-releasing hormone, follicle stimulating hormone, luteinizing hormone, estrogens, androgens, and progestins (Pryce et al., 2004).

Ovulation in dairy cattle is affected by elements from the environment. For example, mastitis, a common bacterial infection of the mammary gland, has a large effect on a dairy cow's fertility (Barker et al., 1998; Schrick et al., 2001). The response to the infection causes swelling, redness and increased body temperatures. During the infection, in order to combat the bacteria, there is an increase in somatic cells in the animal's milk (Hansen et al., 2004). This increase in cells is used as a measure of the extent of infection, the somatic cell score. The mastitis infection appears to prevent ovulation due to the increase in body temperatures (Hansen et al., 2004).

**Nutritional Effects on Fertility**

After parturition many dairy cows enter negative energy balance. The animal is unable to consume adequate energy to maintain the production of large quantities of milk (Domecq et al., 1997; Waltner et al., 1993). An animal in negative energy balance uses body fat in order to maintain high milk production (de Vries and Veerkemp, 2000). The intensity and duration of the negative energy balance during the beginning stages of lactation will have varying effects on the cow's ability to establish pregnancy (Butler and Smith, 1989; Nebel et al., 1993). Cows with improper nutrition during this crucial time are inclined to have loss of body reserves and, consequently, major detrimental effects on fertility (Boland and Lonergan, 2003).

The excessive mobilization of fat, which occurs in an animal in a state of negative energy balance, will cause the animal to have a poor body condition. Even if a cow is considered to be genetically superior in terms of reproduction, she can produce poor quality oocytes due to a low body condition, indicating fertility and body condition are closely related (Boland and Lonergan, 2003). Therefore, it is important to have a method to allow producers to determine the body condition of their dairy cattle, on a predetermined scale, to assist in breeding decisions. A widely used method is body condition scoring (BCS), which is used to determine the amount of fat on the body of the animal. A score, from 1 to 5 with 5 being the fattest, is determined for the animal based on visual observation and tangible evaluation (Wildman et al., 1982). The body condition score of the animal will affect reproduction the most during the time period between pregnancies (Garnsworthy, 1988). Both extremes of the body condition scores have negative effects on fertility; cows that are too fat tend be harder to breed back (Wildman et al., 1982), while cows that are too thin have

a tendency to have sporadic ovulations which makes it harder to establish pregnancy (Baishya et al., 1982). A producer should aim to keep an animal's body condition score away from both extremes. Animals with an average amount of fat are in the best condition for reproduction and milk production (Frood and Croxton, 1978).

An animal producing a large amount of milk and not eating properly will have an increased level of circulating growth hormone (GH) in the blood stream. Along with increased circulating GH, there will be lower levels of insulin and liver GH receptors. In association with this, the plasma concentration of insulin-like growth factors (IGFs) will decrease due to a reduction in the production of IGFs and IGF binding proteins in the liver (Webb et al., 1999). Research has identified a connection between the concentration of insulin in the plasma and the initiation of the estrus cycle (Webb et al., 1999).

**Measuring Fertility**

The most common ways to measure the fertility of the dairy cow are interval measures, i.e. days from calving to first service or heat and days open (Pryce et al., 2004). Another common measure is non-return to first service, which is a measure of whether the animal needs a repeat insemination after a preset amount of time (Pryce et al., 2004). These measures are considered 'traditional fertility parameters' (Royal et al., 2000b). Royal et al. (2000b) suggests 'physiological fertility parameters' as a new approach to measuring fertility. This can be accomplished by measuring the level of progesterone in the milk because progesterone is an indicator of luteal activity (Royal et al., 2000b). Using this measure, a producer can evaluate an animal based on her physiological ability to become pregnant and not on her fertility as a result of management practices. However, the

physiology of progesterone levels will still be influenced by the animal's environment and/or body condition at the time of evaluation (Royal et al., 2000b).

In the dairy cattle industry, many functional traits such as fertility and calving ease are difficult to measure because they are often subjective and inconsistent across farms. Data are recorded differently from farm to farm. It may take substantial effort to compile and standardize all information into one central database (Weigel, 2004). This, together with the low heritability of these traits, makes it harder to improve these traits through conventional selection (Kühn et al., 2003).

Fertility is measured differently throughout the world. Interbull (2005) shows that Austria and Germany measure the non-return rate 90 days after the first insemination; however, this trait has a heritability of only 0.02. Israel measures the percent conception index by dividing 100 by number of inseminations per conception, with a heritability of 0.025. The fertility measure used in France is a binary trait, where cattle are either pregnant or not pregnant, instead of as a measure of number of days open or as a percentage of cows pregnant. This trait also has a heritability of 0.02. With fertility being such an important trait to today's dairy producers it is imperative for the United States to develop a method for the inclusion of this trait into their evaluation programs. The United States is working towards this goal by calculating genetic evaluations for daughter pregnancy rate.

**Pregnancy Rate**

Pregnancy rate is defined as the percentage of cows that conceive in a 21-day time period (VanRaden et al., 2004a). The 21-day period represents a normal estrus cycle and allows the cow one opportunity to become pregnant. Pregnancy rate is starting to become the recommended way to evaluate fertility in dairy cattle (VanRaden et al., 2004a) in the United

States. The formula for pregnancy rate is 21 / (days open – voluntary waiting period + 11) (VanRaden et al., 2004a). The waiting period is the time between the start of lactation and the rebreeding of the cow. Length of the waiting period is a subjective decision by the producer and it allows the animal to recover her body condition after parturition. The number 11 is fixed and is used in order to center the conception measure inside each 21-day period. This will allow cows that breed during the first round of insemination to be awarded a higher value over the animals that take longer to conceive. Pregnancy rate is preferred over the more conventional measure of fertility, number of days open (VanRaden et al., 2004a). This may be due to the fact that pregnancy rate evaluations are more up to date and can be predicted. Kuhn and coworkers (2004) recently developed a predictor for determining pregnancy rate. This will allow for earlier evaluation of young sires to be able to incorporate them sooner into breeding programs (Kuhn et al., 2004). Using pregnancy rate as a measure for fertility may be easier for producers to understand and implement. For pregnancy rate, a positive number value will indicate good fertility, however, with days open a large positive value would be considered undesirable (VanRaden et al., 2004a). Marketing the use of this evaluation factor may be easier using a larger positive value as a selection goal. Days open is used in the calculation of daughter pregnancy rate (DPR) and a 1% increase in pregnancy rate is comparable to a decrease of 4 days open (VanRaden et al., 2004a).

DPR is calculated by Animal Improvement Programs Laboratory (AIPL) of the USDA's Agriculture Research Service. The AIPL introduced the use of DPR to be utilized in evaluation programs in February 2003 and was incorporated in the net merit assessment in August of 2003. When a producer makes the decision of which animals to breed, the net merit of each animal is considered; i.e., the factors that cumulatively affect the producer's

decision to breed an animal. Within net merit, DPR is about 7% of the total merit value of the animal in the commercial industry (VanRaden and Seykora, 2003). Including DPR in selection decisions should help improve fertility. Productive life (PL) was introduced into evaluation programs in 1994 and since then has helped to slow the decline in fertility. Productive life is defined as the time period the animal is active in the herd for milking, before she is removed (http://www.aipl.arsusda.gov/reference/scspl/pl.htm). In Holstein cattle, PL receives 11% of the total merit calculation. In the computation of PL, fertility of the animal is included, because an animal cannot be a member of the milking herd if she is unable to achieve and maintain a pregnancy. Due to this factor, it may be more beneficial to select for fertility traits independent of productive life (VanRaden et al., 2004a).

Quantitative measures for fertility are often measured as predicted transmitting ability (PTA) or daughter deviations (DD). Predicted transmitting ability estimates the genetic worth of an animal that will be transmitted to their offspring (Holstein Association USA, Inc.: Glossary of Sire Summaries Terms). DDs are sires' daughters' average phenotypic value in comparison to the entire population. In other words, the difference in trait values of one sire's daughters when compared to all of the daughters in the population is the daughter deviation (VanRaden and Wiggans, 1991).

For selection to continue improving the genetic worth of a dairy cow, fertility needs to be added to the selection decisions. An animal's selection index is based on an overall score, considering how worthy the animal is as a genetic parent and also the worth of the animal's relatives (Pryce et al., 2004). In the end, a dairy producer's success is highly dependent on the overall fertility of the animals. In the future, reproductive health and fertility must be included in breeding programs.

Artificial selection has been used since the 1950's to achieve genetic improvement in cattle. Selection in most breeding programs has used observable phenotypes to obtain superior animals for the breeding herd. When using phenotypes as a basis for selection, environmental and genetic effects are included in the analysis of merit (Dekkers and Hospital, 2002). Phenotypic selection practices are used without knowledge of the genes underlying the observed trait. This limitation can cause problems in that the phenotype may not be a correct indicator of the genetic worth of the animal, and the trait might not be observable in both sexes or at the time of selection (Dekkers and Hospital, 2002). In contrast, the use of molecular genetics may ease some of the problems seen in traditional quantitative selection. One of the greatest benefits of molecular genetics is the ability to evaluate genetic material from both sexes at any stage of life.

**Heritability**

Not only are some traits hard to assess based on the reasons discussed above, some have low heritability, making it difficult to incorporate them into a breeding program. In general, the heritability of a trait is the measure of how accurate the genotype is in predicting a given phenotype. This is termed heritability in the "broad sense". A more commonly used heritability is heritability in the "narrow sense". This value expresses how the genes an offspring receives from its parents will affect the expression of the phenotypes. Since this measure of fertility includes the genes transmitted within family, it also measures how closely relatives will resemble each other. Narrow sense heritability can be very useful in breeding programs to help predict how closely an offspring will perform compared to his or her parent (Falconer and Mackay, 1996).

The heritability of days open in US Holstein cattle is estimated at around 0.04 (± 0.002) (VanRaden et al., 2004a). Heritability of fertility has been estimated using different fertility measures, and the estimates range from 0.01 to 0.06 (VanRaden et al., 2004a). Regardless of how fertility is measured in dairy cattle, the estimates of their heritabilities are, on average, very low; i.e., the phenotype for fertility of an animal is not a strong indicator of the genetic worth of the animal. This outcome may be partially due to the fertility of a dairy cow being greatly influenced by outside factors, such as the environment. Due to the low heritability of this trait, traditional selection methods may not be the best approach for improving this type of trait.

**Quantitative Trait Loci (QTL)**

Economically important traits are often influenced by a large number of genes with relatively small effects (Lande and Thompson, 1990). Much of the current focus of molecular genetics is the search for quantitative trait loci (QTL). Geldermann (1975) came up with the term to refer to the locus underlying a quantitative character. Quantitative traits are affected by the environment along with a large number of polygenes; i.e., gene with small effects that in coordination with additional genes affects the same phenotypic trait, which will result in the continuous distribution of phenotypic expression (Georges et al., 1995). In 1923, Sax was the first to show an association between markers and quantitative trait loci in *Phaseolus vulgaris,* a common bean (Sax, 1923). In order to find QTL, it is necessary to have a phenotypically diverse population, which is segregating at polymorphic, well-dispersed molecular marker loci.

**Experimental Designs**

Two pedigree structures in dairy cattle used in QTL detection studies are the full-sib design and the half-sib design. The full-sib design is less commonly used in dairy cattle experimental populations. A full-sib family is the mating of two individuals and subsequent examination of their offspring in order to determine the inheritance of alleles. In contrast, a single individual is mated to multiple partners in the half-sib design (Soller, 1998). The half-sib family design is most commonly observed in commercial dairy populations due to the wide use of AI in breeding programs. Half-sib designs are beneficial because more animals can be obtained from the multiple matings, which will therefore lead to higher statistical power to detect QTL (Soller, 1998).

The two common half-sib experimental designs used in dairy cattle to detect QTL are the daughter design and the granddaughter design. Both designs have advantages and disadvantages that should be considered before selecting an approach. In the daughter design, daughters of a sire, heterozygous for markers, are assessed for their marker genotype and quantitative trait values. Figure 1 is a visual representation of this approach modified from Sonstegard et al. (2001). In contrast, the granddaughter design uses marker genotypes from sons of heterozygous sires and quantitative trait values from the daughters of the sons. In the granddaughter design the terminology is as follows: the "grandsire" is a sire heterozygous for a particular marker, his sons are the "sons", and the "granddaughters" are the daughters of the sons (Weller et al., 1990). Figure 2 is a diagram representing this concept, modified from Sonstegard et al. (2001). DNA is extracted from the grandsire and his sons for genotyping, and the granddaughters are phenotypically evaluated.

The less frequently used design is the daughter design. In the daughter design, there are fewer animals from which to collect phenotypic data. This design may be beneficial for phenotypes that are expensive to collect or hard to score. On the other hand, since all the daughters need to be genotyped at the selected DNA markers, the genotyping costs will inevitably be higher. In the granddaughter design, there are fewer animals to genotype because the daughters are not genotyped. It is easier to collect DNA samples from the sons to perform genotyping through the use of semen samples as the source of DNA, than from widely dispersed daughters. This can be attributed to AI centers that are concentrated locations of semen. One last benefit of the granddaughter design is there will be higher statistical power due to additional generations in the pedigree with more phenotypic information to include in the analysis (Weller et al., 1990). With an experimental design in place, there are two main approaches to detect QTL; the candidate gene approach and the genome scan approach.

**Candidate Gene Approach**

Researchers using the candidate gene approach utilize knowledge about the underlying mechanism of a phenotypic trait to speculate which genes are affecting the trait. The main objective of this type of study is to determine if there is an association between the different alleles at the candidate gene and the phenotypic trait of interest in the population (Kwon and Goate, 2000). In other words, the goal of the candidate gene approach is to determine if individuals that exhibit the trait of interest are associated with a specific candidate gene allele more often than individuals not displaying the trait. Once a candidate gene is located, it can be applied to breeding schemes relatively easily within populations segregating the favorable allele because the animals can be selected based on the presence or

absence of the superior allele (Soller, 1998). The difficulty in using the candidate gene approach is there must be an understanding of the biological mechanism controlling the trait before being able to select the potential candidate gene (Kwon and Goate, 2000). The physiology underlying a phenotypic trait is often complicated and specific mechanisms are sometimes unknown.

**The Genome Scan Approach**

In contrast, the genome scan approach does not require any prior knowledge of the underlying mechanism controlling the trait and is advantageous when the actual physiology of the trait is not yet elucidated. In a genome scan, markers evenly spaced throughout the genome are selected and genotyped in a segregating population to detect an association between inheritance of the markers and expression of a phenotypic trait. The goal of the genome scan is to discover a region of the chromosome that has an affect on a phenotypic trait (Dekkers and Hospital, 2002). This approach is powerful because even though the physiology is unknown, there is still the potential to detect a locus with a large effect on the trait of interest (Haley and Archibald, 1998). The power to detect QTL using a genome scan is influenced by the sample size of the population used. If there are not enough animals used in the analysis it might be difficult to verify exactly how many genes are affecting the trait and the magnitudes of their effects (Beavis, 1998). The main drawback of a genome scan is that, after the analysis, the specific gene that controls the quantitative trait is still unknown; only a region of the chromosome affecting the trait can be found and it may take a great deal of time to identify the causative gene (Streelman and Kocher, 2000). Another disadvantage of the genome scan is the large number of markers required in order to cover the entire

genome (Haley and Archibald, 1998). Genotyping large quantities of animals with many markers can be fairly time consuming and expensive.

To further resolve the individual loci underlying a QTL, fine-scale mapping of the region containing the QTL must be performed. In fine-scale mapping, the marker density around the QTL region is increased and additional animals are included in the investigation in order to narrow the chromosomal region to a more specific location. The addition of animals to the analysis will add more meioses to help further refine the location of the QTL (Streelman and Kocher, 2000). Fine-mapping the QTL to a 10 cM region will help decrease the search for the actual gene affecting the trait to approximately 500-1,000 genes (Streelman and Kocher, 2000). Once markers closely flanking the QTL have been identified, the information can be incorporated into breeding programs.

**Use of Molecular Markers**

There are two broad types of commonly used genetic markers, Type I and Type II. Type I markers represent functional genes, usually within coding regions and are generally well conserved across species (Haley and Archibald, 1998). Type II markers are not located within coding regions, are highly variable and may have many alleles segregating within a population. Due to Type II markers being highly variable, their inheritance can easily be traced in pedigrees (Haley and Archibald, 1998). Commonly used Type II markers in genome mapping studies are microsatellite markers. Microsatellites are very popular because they are abundant and evenly dispersed in the genome. Microsatellites are tracts of short repeats varying from 1-7 nucleotides in length. Microsatellite markers mutate at a high rate due to DNA polymerase slippage or imperfect DNA repair events that occur during replication resulting in multiple different alleles being present in a population (Beuzen et al.,

2000). The microsatellite repeats can be amplified with flanking primers using the polymerase chain reaction (PCR). The variable lengths can then easily be examined through gel electrophoresis (Soller, 1998).

DNA genetic markers can be used in genetic selection if the markers are linked to a QTL of interest. This approach combines the use of polymorphisms that are abundant in the genome with their linkage to the QTL by determining the association between the marker genotypes and the phenotypic value of that particular trait (Dekkers and Hospital, 2002). The chosen marker loci may not directly affect the trait, but can be used for selection due to the linkage disequilibrium between the marker and the QTL (Lande and Thompson, 1990). Linkage disequilibrium occurs when there is no random association between two loci, and two genes, or a QTL and a DNA marker, are inherited together more frequently than expected by chance. QTL can be detected using genetic markers along with the phenotypes collected from a pedigree that segregates the trait of interest. This information can then be incorporated into breeding schemes.

**Marker Assisted Selection**

Marker assisted selection (MAS) refers to the utilization of marker data in the selection process (Meuwissen and Arendonk, 1992), by using DNA markers to identify superior genotypes in animal populations. This method of selection is advantageous in many ways. First, the environment does not affect the selection decision. Second, the generation interval, defined as the amount of time it takes to replace one generation with the next, can be decreased because selection can be done earlier in life. Last, selection intensity can be increased because more individuals are included in the analysis (Soller, 1998). There is no need to wait for an animal to reach a certain age to be considered in the selection decisions.

Selection intensity is the measure of how strictly breeders select based on a certain criterion. When there are more animals from which to select, those chosen for the breeding herd will be a smaller proportion than those not selected, and consequently, selection intensity will be increased (Bourdon, 2000).

The economic benefits and pitfalls of using molecular genetics in breeding programs must be considered before use. The costs incurred in the use of MAS include DNA collection, genotyping of many animals and statistical analysis (Dekkers and Hospital, 2002). In some situations, the economic benefit of using MAS clearly outweighs the costs. For example, using genotypic data instead of the tedious collection of complicated phenotypic information of difficult to measure traits (i.e. calving ease, pregnancy rate) is clearly favorable (Dekkers and Hospital, 2002). Some traits are only measurable in females and having genetic information on the males would be beneficial during selection processes.

Dairy cattle QTL experiments greatly benefit from the existence of The Dairy Bull DNA Repository (DBDR; Da et al., 1994) as a resource for genetic material. The DBDR is a collection of bull semen available to scientists for discovering loci affecting dairy cattle quantitative traits. The establishment of this collection began in the early 1990's and has assisted scientists in advancing dairy cattle genomics. The collection is located at the University of Illinois at Urbana-Champaign. Various AI organizations throughout North America supplied semen samples to the collection. Semen samples were requested one time from proven sires. There are 35 grandsires in the collection with many families having more than 40 sons (Da et al., 1994).

Another source of semen samples for QTL detection in dairy cattle is the Cooperative Dairy DNA Repository (CDDR; Ashwell and Van Tassell, 1999), which began in the late

1990's. The CDDR is an extension of the DBDR. However, unlike the DBDR, which only accepted bulls that were proven, this repository contains semen from proven and young bulls from AI organizations around the world. The collection contains numerous generations of dairy cattle and currently over 13,000 sires are represented in this collection (Van Tassell, personal communication).

Concentrated collections of semen samples from commercial dairy cattle populations for the use in quantitative genetic analysis help save researchers both time and money. Without them, researchers would need to develop populations independently. Unlike many species, such as mice, which can multiply in a relatively short period of time, the dairy cattle generation interval is long. A population that contains enough animals to obtain statistical significance could take an unrealistic amount of time. Also, it is cost inhibitory for a researcher to obtain enough land and resources to create and maintain this type of population. The use of a commercial population already in place is the ideal solution. One final benefit is if the QTL is identified in a commercial population it can be more easily incorporated into breeding schemes.

**Genotyping**

Genome mapping experiments require polymorphic DNA markers that are dispersed throughout the genome. There are currently over 3,960 unique markers located on the bovine linkage maps (Ihara et al., 2004). These genetic markers can be used to track the inheritance of linked segments of the genome in pedigrees. An association between the inheritance of a marker allele and a phenotype is indicative of linkage between the marker and trait of interest (Haley, 1999).

The ability to perform high throughput genotyping is essential to a successful genome mapping experiment. In the past, restriction fragment length polymorphisms (RFLPs) were used to identify QTL. RFLPs take advantage of nucleotide changes that occur in all genomes, where the nucleotide change results in either the loss or gain of a recognition site for a restriction endonuclease. When the recognition site is located in only one of the alleles, two bands of different size will appear during gel electrophoresis; either a long fragment indicating the loss of a recognition site or a short fragment indicating the gain of a recognition site (Beuzen et al., 2000). A shortcoming of RFLPs is that not all mutations in DNA sequence will cause a loss or gain of a restriction site, and consequently will not be observed through digestion and gel electrophoresis. This technique for genotyping was widely used in the past; however, the approach is not ideal for high throughput genotyping of many animals in a pedigree (Beuzen et al., 2000).

In an entire genome mapping study, a large number of animals need to be genotyped, thus requiring a high throughput method (Li et al., 2001). The development of automated genetic analyzers has facilitated the process of faster genotyping. This method employs the use of fluorescently labeled PCR primers and the sizing of the PCR products to produce individual genotypes. Three different fluorescent dyes are used to label one of the two primers and a fourth dye is used as an internal size standard. Electropherograms that are generated by the allele-calling software present homozygous and heterozygous genotypes in a clearly distinguishable manner. Using the software programs, a user can size alleles with relative accuracy and keep track of all data in one location (Ziegle et al., 1992). Allele sizing is performed electronically (Idury and Cardon, 1997) and the decrease in human intervention will help to prevent errors that may occur in visual analysis of genotypes. With the

development of automated genotyping and its associated software, the rate of genotype processing has increased (Idury and Cardon, 1997). A problem with automated genotyping is the artifact or stutter peaks caused by polymerase slippage during amplification, creating a slightly shorter or longer product. However, the software programs are able to clearly distinguish stutter peaks from major peaks of the complete product (Ziegle et al., 1992).

**GenoProb**

Even with the genotyping software programs currently available, there is the possibility of human error in genotyping experiments. Errors can occur during genotyping for many reasons, including the misloading of samples into wells prior to analysis, mistaken allele sizing due to indistinct additional peaks, and DNA sample contamination, which can occur during DNA isolation (Li et al., 2001). In some cases genotypes can be deduced based on classical Mendelian inheritance. However, many times there are animals in the pedigree that do not have recorded genotypes. Also, in most granddaughter design QTL studies there usually is no genotypic information available on the females. In these instances, information cannot be acquired to infer the inheritance of alleles, causing uninformative genotypes (Thallman et al., 2001). GenoProb is a database program used to calculate probabilities for unknown genotypes within complex pedigrees (Thallman, 2002). This program can identify errors in genotyping and predict unknown genotype probabilities based on the pedigree information.

An algorithm called “iterative allelic peeling” is used to determine the predicted genotype probabilities. It is based on the probabilities of alleles being transmitted from parent to offspring (Thallman et al., 2001). This algorithm is set with a predetermined error term and determines probabilities of individual genotypes conditional on the individuals

nearby in the pedigree. Each time an inheritance is determined, some of the animals can be "peeled" away from the pedigree, or momentarily removed from the current analysis, and another set of probabilities can be determined. This process is repeated with a new set of individuals after each "peel" (Thallman et al., 2001). The allelic peeling is a logical way of looking at the inheritance within the pedigree to speed up the process of determining the individual probabilities. This program helps to identify markers with genotyping problems that might go undetected. Another benefit is the discovery of null alleles. Null alleles are alleles that produce no product during amplification, perhaps caused by mutations within the template, preventing a primer from annealing to its complementary sequence. This will cause the allele calling software to designate the animals as homozygous when they are actually heterozygous. The misrepresentation of the genotypes may skew the experimental results.

**Advances in QTL Statistical Analysis**

Statistical methods for detecting QTL/marker associations have been progressing throughout the years. Traditional methods involved simple linear regression of phenotype on one marker's genotype, such as analysis of variance. With this method the phenotypic variance between populations containing alternative forms of an allele are compared. If the means are significantly different from zero, there is evidence of a QTL affecting the phenotypic trait (Lander and Botstein, 1989). This method can be used to effectively map QTL; however, there are some problems that need to be addressed. These include underestimating the effect the QTL has on the phenotypic trait and increased number of progeny to estimate the size of the effect, when the QTL does not lie on the marker. Lastly,

the specific location of the QTL is not determined; therefore it is difficult to decipher the extent of the marker/QTL linkage (Lander and Botstein, 1989).

A newer approach to help alleviate some of the problems with single marker analysis is using interval mapping and maximum likelihood methods. With interval mapping, the recombination rate between a set of markers can be used to estimate the phenotypic effect of the QTL. This method can be applied to all locations in the genome where genetic markers are present. Benefits of this approach are that fewer progeny are needed to estimate the effects when compared with the single marker analysis method, the intervals give a better prediction of the location of the QTL and the strength of the QTL is clearly characterized (Lander and Botstein, 1989). QTL studies also use variance components methods with residual maximum likelihood to allow for analysis across families (Zhang et al., 1998). This method determines if adding the additive and dominance effect estimates to a statistical model will minimize residuals. Incorporating linkage disequilibrium data into analysis is another approach used in mapping studies (Meuwissen and Goddard, 2004). This method uses the determination of whether a QTL and a marker are inherited together more likely than expected by chance. This is not ideal for larger populations when there are multiple loci affecting the expression of a trait because the linkage disequilibrium may be incorrectly estimated due to the multiple QTL affects (Lynch and Walsh, 1998).

A different approach has been suggested to help reduce the complexity in the calculations of the maximum likelihood method. This method employs the use of flanking markers, but instead uses a least squares method. Proposed by Haley et al. (1994) this method considers the entire set of informative markers within a linkage group simultaneously. When all markers are taken into consideration at once the statistical power

is increased. In order to make this statistical computation user friendly a computer software program was developed in 2002.

**QTL Express**

QTL Express (http://qtl.cap.ed.ac.uk; Roslin Institute, Edinburgh) is an internet-based software program, which can use outbred populations to map QTL by means of regression interval mapping methods. Interval regression mapping takes into consideration two flanking markers. This method offers an increase in power over single marker analysis and more precise estimates of the QTL effect and position (Lynch and Walsh, 1998). The data set is inputed using three files; a genotype file, a phenotype file, and a map file. A two-step format is utilized during each analysis. First, at specific regions on the chromosome the identity-by-descent (IBD) probabilities are calculated using marker data. The program is determining the probability that individuals received identical copies of an allele from a common ancestor within linkage groups. Next, a linear regression statistical model is determined using the IBD probabilities calculated in the first step (Seaton et al., 2002).

This program is capable of analyzing both half-sib, outbred populations and inbred or outbred $F_2$ populations. Phenotypic values are regressed on genotypic coefficients of a specific QTL location (Haley et al., 1994). Chromosome-wise and genome-wise significant thresholds are determined using permutation tests. Fisher (1935) was the first to suggest the idea of using permutation tests. In QTL mapping experiments permutation tests are used to determine an appropriate significance threshold values (Churchill and Doerge, 1994). In these experiments multiple tests are performed at each marker interval. Consequently, it is likely there will be an association by chance. Permutation is used to overcome this obstacle. This creates a conservative significance level to allow for the multiple tests. These

permutations involve "repeated reshuffling of the quantitative trait values and the generation of a random sample of the test statistic for an appropriate null distribution" (Churchill and Deorge, 1994). The phenotypic data are scrambled multiple times in order to generate a test statistic for a distribution where there is no significant association between a QTL and a phenotypic effect. In order to obtain a 0.05 significant level, at least 1,000 repeated shufflings are necessary (Churchill and Doerge, 1994).

A 95% confidence interval for the QTL location is determined using the statistical bootstrapping method (Seaton et al., 2002). The widths of the confidence intervals are influenced by the size of the sample population, the relative strength of QTL phenotypic effect, and the number of markers used in the study (Lynch and Walsh, 1998). To generate a bootstrap sample, values are drawn from the population numerous times and replaced back into the data set (Visscher et al., 1996). This sample will contain some original values multiple times and some values will not be present. Multiple samples are taken from the data and an estimate of the map position is calculated, in turn generating a distribution of the estimates. Accurate confidence intervals can be determined when using at least 200 bootstrap samples (Lynch and Walsh, 1998).

**Putative QTL Affecting Dairy Cattle Fertility**

There have been numerous genome scans to identify QTL affecting economically important traits in dairy cattle, with at least one QTL reported on every chromosome. Most studies have concentrated on QTL affecting milk production and milk composition. More recent studies have focused on identification of QTL affecting functional traits including fertility. There is a large range of chromosomes (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 14, 15, 16, 17, 18, 19, 20, 21, 27, 28) exhibiting fertility QTL effects throughout the dairy cattle genome.

Chromosome 6 (Schrooten et al., 2000; Kühn et al., 2003; Ashwell et al., 2004) and chromosome 7 (Kühn et al., 2003; Ashwell et al., 2004; Boichard et al., 2003) were indicated in multiple studies; however, the relative position differs in each study. This could be due to the difference in the measurement of the fertility trait. Some examples of these measures include, non-return daughter at 56 days (Schrooten et al., 2000), stillbirth (Kühn et al., 2003), pregnancy rate (Ashwell et al., 2004), dystocia (Kühn et al., 2003), female fertility post partum (Boichard et al., 2003) and gestation length (Schrooten et al., 2000). Recently, Arias and Kirkpatrick (2004) conducted a genome scan to identify QTL affecting fertility by evaluating ovulation rate. Significant QTL affecting ovulation rate were reported in this study on chromosomes 7 (30 cM) and 9 (65 cM).

Another factor that might cause variation in identification of a QTL region affecting fertility is the breed of cattle used for the study. Research has been conducted on an assortment of breeds such as French dairy breeds (Boichard et al., 2003), German Holsteins (Kühn et al., 2003), US Holsteins (Ashwell et al., 2004), Holstein Friesian (Schrooten et al., 2004) and Finnish Ayrshire (Schulman et al., 2004). These breeds may exhibit traits differently, and the minor physiological differences could cause the QTL effects to differ slightly.

Even with the variation in fertility measures and breeds used within studies, it appears that chromosome 18 is reported in multiple studies to exhibit QTL affecting fertility (Ashwell et al., 2004; Boichard et al., 2003; Kühn et al., 2003; and Schrooten et al., 2004).

## Chromosome 18 Putative QTL

There have been several studies that have reported QTL affecting economically important traits on chromosome 18. Table 1 is a compilation of significant results from

various QTL detection studies. Kühn et al. (2003) performed a genome scan for functional traits affecting udder health, functional herd life and fertility in a study of 16 German Holstein half-sib families. All of the phenotypic traits analyzed in this study showed significant effects on chromosome 18, with the interval between BM2078 and TGLA227 having the greatest number of significant effects for both fertility (dystocia, nonreturn rate of 90 days, stillbirth) and udder health (somatic cell count, an indicator of mastitis).

Ashwell et al. (2004) reported a joint analysis of genome scans of 10 Holstein families. A QTL affecting fat yield was found on chromosome 18. A relatively significant QTL affecting pregnancy rate was discovered in two different Holstein families, with the most significant effect (F-statistic=14.1) located in the marker interval BM7109-ILSTS002.

In 2000, Schrooten and coworkers studied 20 Holstein Friesian families to perform a genome scan for functional and conformation traits. The only effect found on chromosome 18 was for somatic cell count in the region between BM7109 and ILSTS002. Schrooten et al. (2004) reanalyzed the data to evaluate two traits simultaneously to account for multiple QTL effects. In this new analysis, QTL affecting multiple traits including milk fat and protein percentages and the interval calving to first insemination trait were identified in the same interval between BM7109 and ILSTS002. This agrees with the finding from Kühn et al. (2003) and Ashwell et al. (2004), who also found a significant QTL effect for stillbirth and pregnancy rate respectively, in this region, suggesting this region may harbor QTL affecting fertility traits.

Genome scans have been performed using several different breeds of dairy cattle. Boichard and coworkers (2003) studied French dairy breeds, Schulman et al. (2004) studied Finnish Ayrshire cattle and Bennewitz et al. (2004) studied German Holstein cattle. In each

of these studies, significant QTL effects were identified on chromosome 18. Boichard et al. (2003) identified QTL affecting a conformation trait (teat distance), Schulman et al. (2004) identified a QTL affecting somatic cell score and Bennewitz et al. (2004) found significance for a milk fat yield. These studies suggest a QTL at the telomeric end of the chromosome.

There appear to be significant effects spanning chromosome 18, however, the marker intervals BM7109-ILSTS002 (Ashwell et al., 2004; Schrooten et al. 2004; Kühn et al., 2003; Boichard et al., 2003) and BM2078-TGLA227 (Kühn et al., 2003; Boichard et al., 2003) seem to be common across studies. This observation may indicate a higher likelihood that one or more QTL may be found in one of these intervals so that fine-mapping of these regions would be advantageous.

## Conclusions

The dairy industry undoubtedly supplies important products to both the consumer and the United States' economy. During the last 50 years, traditional selection methods have been applied to increase milk production per cow, and this was successfully accomplished with milk production nearly doubling. However, this intense selection for milk production has had negative consequences on the fertility of the animals. The fertility of the animal is indisputably linked to milk production, given that the lactation cycle is initiated through parturition. A dairy cow that is unable to achieve and maintain pregnancy will cost a producer a lot of time and money. If the decrease in fertility is not halted, the dairy industry will suffer significant losses.

Advances in molecular genetics, such as the development of the bovine linkage map and progress in high throughput genotyping methods, have assisted in discovery of QTL affecting economically important traits. The incorporation of the daughter pregnancy rate

calculation into national dairy cattle evaluations in 2003 has opened the door for the location of QTL affecting the fertility of the dairy cow. Analysis methods have been developed for examination of genotype and phenotype data. Use of dairy bull DNA collections, such as the DBDR and the CDDR, have helped aid in QTL detection studies by supplying important genetic material from commercial dairy populations in an efficient manner.

The dairy cattle genome has been extensively scanned for QTL affecting economically important traits. Many studies have identified putative QTL, either directly or indirectly, affecting the reproductive efficiency of the dairy cow. More precisely, in the preliminary study by Ashwell et al. (2004) there appears to be a highly significant effect on pregnancy rate on chromosome 18 in a popular US Holstein family. These QTL need to be further evaluated in order to validate their location and effect. The exact benefits of finding a QTL for fertility are yet to be realized, although some improvement decreasing the time and money used in dealing with infertility may consequently have a positive effect on the dairy industry as a whole.

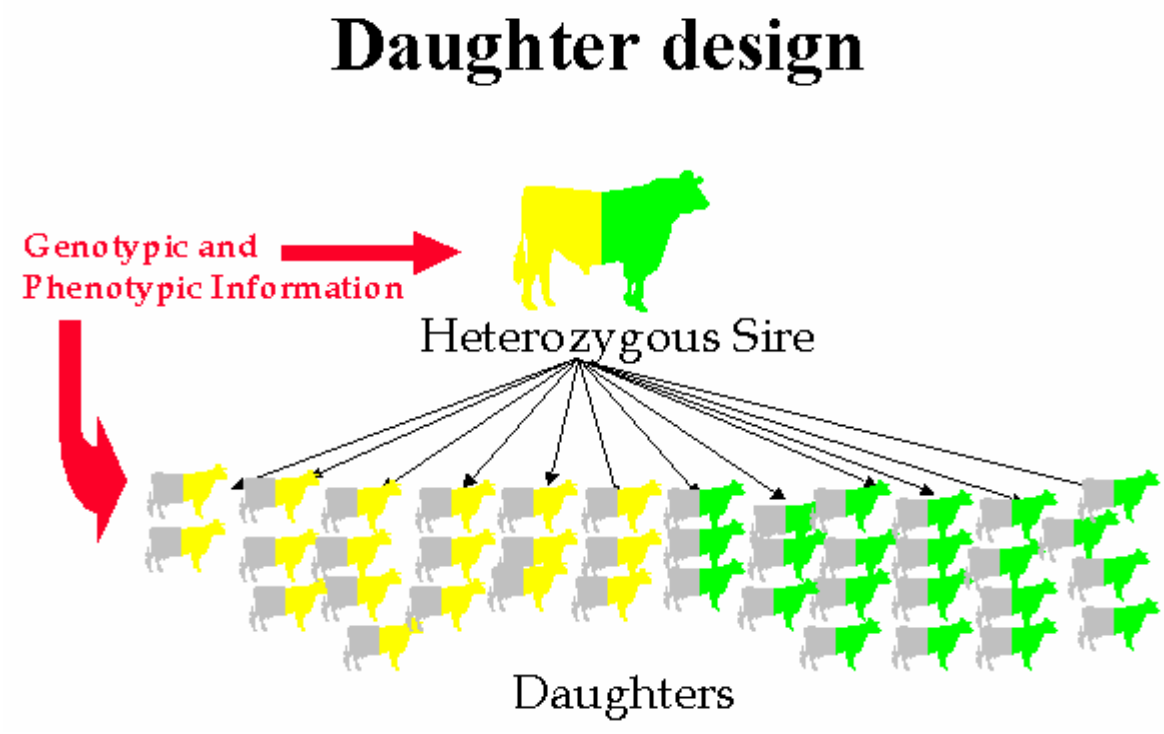


**Figure 1**: Representation of the Daughter design. The sire (top) is shaded in two colors to represent marker heterozygosity. The daughters (bottom) are shaded with a gray allele (from the dam) and either a yellow or green allele depending on inheritance from the sire. In this design, phenotypic and genotypic information is used from the sire and his daughters.

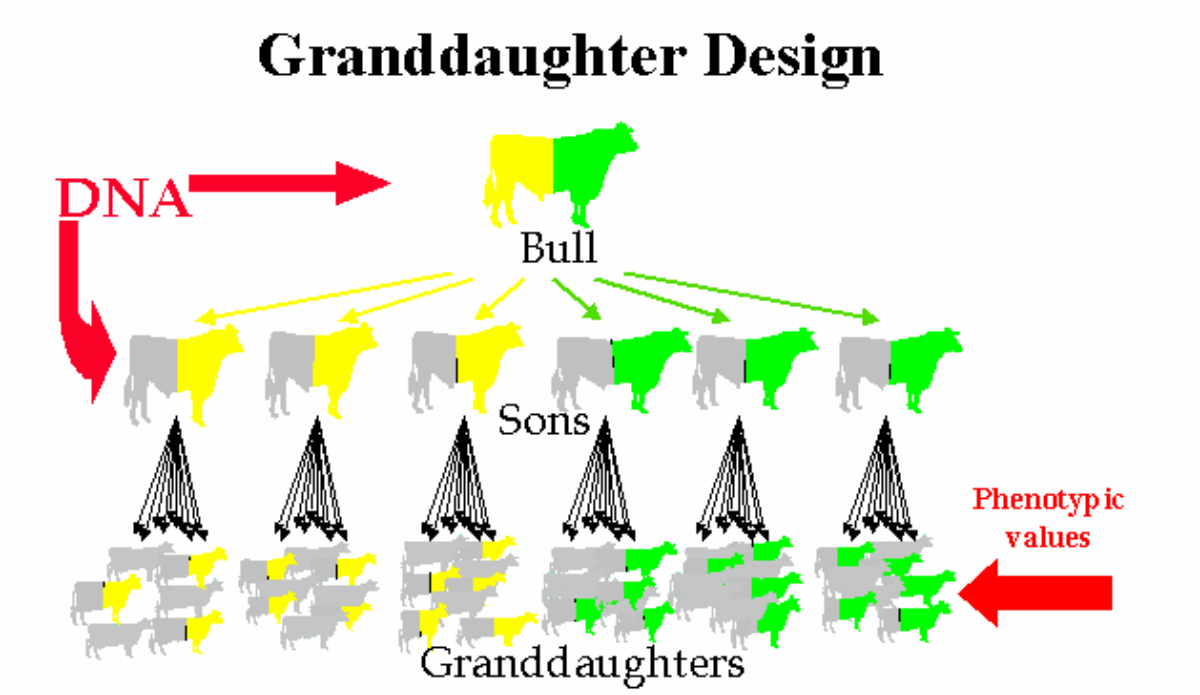


**Figure 2:** Representation of the Granddaughter Design. The Grandsire (top) is shaded in two colors to represent marker heterozygosity. The sons (middle) are shaded in gray (representing the allele from the dam) and in one of the two colors to represent the allele from the grandsire. The granddaughters (bottom) are shaded either completely gray, receiving neither of the grandsire original alleles, or gray and one of the two alternative forms of the alleles coming from their fathers, the sons.

**Table 1:** Results from QTL detection studies on chromosome 18

| Trait[1] | Relative Position (cM) | Marker Interval | Breed |
|---|---|---|---|
| Milk Protein Percent[a] | 10 | ABS13 | French Dairy Breeds |
| Dystocia-maternal effect[b] | 53 | TGLA357-BM7109 | German Holstein |
| Pregnancy Rate[c] | 54 | BM7109-ILSTS002 | US Holstein |
| Protein percentage and interval calving-first insemination[d] | 67 | BM7109-ILSTS002 | Holstein Friesian |
| Fat percentage and interval calving-first insemination [d] | 68 | BM7109-ILSTS002 | Holstein Friesian |
| Somatic Cell Count[e] | 70 | BM7109-ILSTS002 | Holstein Friesian |
| Heel Depth[a] | 74 | ILSTS002 | French Dairy Breeds |
| Stillbirth-paternal effect[b] | 75 | BM7109-ILSTS002 | German Holstein |
| Milk Fat Yield[c] | 84 | BM6507-TGLA227 | US Holstein |
| Udder Balance[a] | 98 | IDVGA55 | French Dairy Breeds |
| Functional Herd Life[b] | 104 | EAC | German Holstein |
| Dystocia- direct effect[b] | 107 | BM2078 | German Holstein |
| Stillbirth-maternal effect[b] | 110 | BM2078-TGLA227 | German Holstein |
| Non-return rate at 90 days-maternal[b] | 111 | BM2078-TGLA227 | German Holstein |
| Somatic Cell Score[g] | 113 | Not reported | Finnish Ayrshire |
| Teat Distance[a] | 115 | TGLA227 | French Dairy Breeds |
| Somatic Cell Count[b] | 117 | TGLA227 | German Holstein |
| Non-return rate at 90 days-paternal[b] | 117 | TGLA227 | German Holstein |
| Milk Fat Yield[f] | 139 | Not reported | Holstein |

[1] Boichard et al. (2003) = a, Kühn et al. (2003) = b, Ashwell et al. (2004) = c, Schrooten et al. (2004) multiple trait analysis = d, Schrooten et al. (2000) = e, Bennewitz et al. (2004) = f, and Schulman et al. (2004)= g.

# Literature Cited

Arias, J. and B. Kirkpatrick. 2004. Mapping of bovine ovulation rate QTL; an analytical approach for three generation pedigrees. Animal Genetics 35: 7-13.

Ashwell, M.S., and C.P. Van Tassell. 1999. The Cooperative Dairy DNA Repository-a new resource for quantitative trait loci detection and verification. J. Dairy Sci. 82 (Suppl. 1): 54.

Ashwell, M.S., D.W. Heyen, T.S. Sonstegard, C.P. Van Tassell, Y. Da, P.M. VanRaden, M. Ron, J.I. Weller, and H.A. Lewin. 2004. Detection of quantitative trait loci affecting milk production, health, and reproductive traits in Holstein cattle. J. Dairy Sci. 87: 468-475.

Baishya, N., S.V. Morant, G.S. Pope, J.D. Leaver. 1982. Rearing of dairy cattle. 8. Relationships of dietary energy intake, changes in live weight, body condition and fertility. Anim. Prod. 34:63.

Barker, A.R., F.N. Schrick, M.J. Lewis, H.H. Dowlen, S.P. Oliver. 1998. Influence of clinical mastitis during early lactation on reproductive performance of Jersey cows. J. Dairy Sci. 81: 1285-1290.

Bascom, S.S., A.J. Young. 1998. A summary of the reasons why farmers cull cows. J. Dairy Sci. 81: 2299-2305.

Beaudeau, F. 1995. Cow's health and farmer's attitude towards the culling decision in dairy herds. Ph.D. Diss., Wageningen Agric. Univ., Wageningen, The Netherlands.

Beavis, W.D. 1998. QTL Analysis: power, precision, and accuracy. In A.H. Paterson (ed.). Molecular Dissection of Complex Traits. CRC Press, Boca Raton, Fl. 145-162.

Bennewitz, J., N. Reinsch, V. Guiard, S. Fritz, H. Thomsen, C. Looft, C. Kühn, M. Schwerin, C. Weimann, G. Erhardt, F. Reinhardt, R. Reents, D. Boichard, and E. Kalm. 2004. Mulitple quantitative trait loci mapping with cofactors and application of alternative variants of the false discovery rate in an enlarged granddaughter design. Genetics 168: 1019-1027.

Beuzen, N.D., M.J. Stear, and K.C. Chang. 2000. Molecular markers and their use in animal breeding. The Veterinary Journal 160: 42-52.

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

Boland, M.P. and P. Lonergan. 2003. Effects of nutrition on fertility in dairy cows. Adv. In Dairy Tech. 15: 19-33.

Bourdon, R.M. 2000. Understanding Animal Breeding. Prentice Hall. Upper Saddle River, NJ. p 199.

Butler, W.R. and R.D. Smith. 1989. Interrelationships between energy balance and postpartum reproductive function in dairy cattle. J. Dairy Sci. 72: 767-783.

Casida, L.E. 1961. Present status of the repeat-breeder cow problem. J. Dairy Sci. 44: 2323-2329.

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

Collins, K. 2000. U.S. Dairy Policy Issues. National Dairy Leadership Conference, U.S. Department of Agriculture. Williamsburg, Va.

Da, Y., M. Ron, A. Yanai, M. Band, R.E. Everts, D.W. Heyen, J.I. Weller, G.R. Wiggans, and H.A. Lewis. 1994. The Dairy Bull Repository: A resource for mapping quantitative trait loci. Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, ON, Canada 21: 229-232.

Darnsfield, M.B., R.L. Nebel, R.E. Pearson, and L.D. Warnick. 1998. Timing of insemination for dairy cows identified in estrus by a radio telemetric estrus detection system. J. Dairy Sci. 81: 1874-1882.

Darwash, A.O., G.E. Lamming, J.A. Woolliams. 1999. The potential for identifying heritable endocrine parameters associated with fertility in post-partum dairy cows. Anim. Sci. 68: 333-347.

Dekkers, J. C.M. and F. Hospital. 2002. The use of molecular genetics in the improvement of agricultural populations. Genetics: Nature Reviews. 3: 22-32.

de Vries, M.J., R.F. VeerKamp. 2000. Energy balance of dairy cattle in relation to milk production variables and fertility. J. Dairy Sci. 83: 62-69.

Domecq, J.J., A.L. Skidmore, J.W. Lloyd, and J.B. Kaneene. 1997. Relationship between body condition scores and conception at first artificial insemination in large dairy herd of high yielding Holstein cows. J. Dairy Sci. 80: 113-120.

Dukes, H.H., M.J. Swenson, and W.O. Reece. 1993. Dukes' Physiology of Domestic Animals. Cornell University Press, Ithaca, NY.

Esslemont, R. J. 1993. Relationship between calving to conception interval and culling rate for failure to conceive. Vet. Rec. 133: 163-164.

Esslemont, R.J. and M.A. Kossaibati. 1997. Culling in 50 dairy herds in England. Vet. Rec. 140: 36-36.

Falconer, D.S. and T.F.C. Mackay. 1996. Introduction to Quantitative Genetics. Pearson Education Limited. Essex, England.

Fisher, R.A. 1935. The Design of Experiments, Ed. 3. Oliver & Boyd Ltd., London.

Frood, M.J., and D. Croxton. 1978. The use of condition scoring in dairy cows and its relationship with milk yield and life weight. Anim. Prod. 27: 285.

Garnsworthy, P.C. 1988. The effect of energy reserves at calving on performance of dairy cows. In: Garnsworth, P.C. (Ed.), Nutrition and Lactation in the Dairy Cow. Butterworths, London, p. 157-170.

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

Georges, M., D. Nielsen, M. Nackinnon, A. Mishra, R. Okimoto, A.T. Pasquino, L. S. Sareant, A. Sorensen, M. R. Steele, X. Zhao, J.E. Womack, and I. Hoeschele. 1995. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics. 139: 907-920.

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

Haley, C. and A. Archibald. Progress in Pig Gene Mapping. In: Progress in Pig Science. J. Wiseman, M.A. Varley, and J.P. Chadwick (eds). Nottingham University Press. 1998. pp. 5-27.

Haley, C. 1999. Advances in quantitative trait locus mapping. In: Dekkers JCM, Lamont SJ, Rothschild MF, eds. From JL Lush to genomics: visions for animal breeding and genetics. Ames, IA, USA: Iowa State University Press, 47-59.

Hansen, P.J., P. Soto, and R.P. Natzke. 2004. Mastitis and fertility in cattle-possible involvement of inflammation or immune activation in embryonic mortality. Am. J. Reprod. Immunol. 51(4): 294-301.

Hermas, S.A., C.W. Young, and J.W. Rust. 1987. Effects of mild inbreeding on productive and reproductive performance of Guernsey Cattle. J. Dairy Sci. 70: 712-715.

Holstein Association USA, Inc.: Glossary of Sire Summaries Terms. http://www.holsteinusa.com/html/ss_terms.html. Accessed on May 25, 2005.

Idury, R.M. and L.R. Cardon. 1997. A simple method for automated allele binning in microsatellite markers. Genome Methods. 7: 1107-1109.

Ihara,N., A. Takasuga, K. Mizoshita, H. Takeda, M. Sugimoto, Y. Mizoguchi, T. Hirano, T. Itoh, T. Watanabe, K.M. Reed, W.M. Snelling, S.M. Kappes, C.W. Beattie, G.L. Bennett, and Y. Sugimoto. 2004. A comprehensive genetic map of the cattle genome based on 3802 microsatellites. Genome Res. 14: 1987-1998.

Interbull. 2005. Description of National Genetic Evaluation Systems for dairy cattle traits as applied in different Interbull member countries. http://www-interbull.slu.se/national_ges_info2/framesida-ges.htm. Accessed June 21, 2005.

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

Kuhn, M.T., P.M. VanRaden, and J.L. Hutchison. 2004. Use of early lactation days open records for genetic evaluation of cow fertility. J. Dairy Sci. 87: 2277-2284.

Kwon, J.M. and A.M. Goate. 2000. The candidate gene approach. Alcohol Research and Health. 24: 164-168.

Lande, R. and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics. 124: 743-756.

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

Li, J.L., H. Deng, D.B. Lai, F. Xu, J. Chen, G. Gao, R.R. Recker, and H.W. Deng. 2001. Toward high-throughput genotyping: Dynamic and automatic software for manipulating large-scale genotype data using fluorescently labeled dinucleotide markers. Genome Res. 11: 1304-1314.

Lucy, M.C. 2001. ADSA Foundation Scholar Award: Reproductive loss in high-producing dairy cattle: where will it end? J. Dairy Sci. 84: 1277-1293.

Lush, J.L. 1948. The Genetics of Populations. Mimeo. Iowa State Univ., Ames.

Lynch, M. and B. Walsh. 1998. Genetics and Analysis of Quantitative Traits. Sinauer Associates, Inc. Sunderland, MA.

Macmillan, K.L., I.J. Lean, and C.T. Westwood. 1996. The effects of lactation on the fertility of dairy cows. Aust. Vet. J. 73: 141-147.

McCullough, D. A. and M.A. Delorenzo. 1996. Effect of price and management level on optimal replacement and insemination decisions. J. Dairy Sci. 79: 242-253.

Meuwissen, T.H.E. and M.E. Goddard. 2004. Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data. Genet. Sel. Evol. 36: 261-279.

Meuwissen, T.H.E., and J.A.M. Van Arendonk. 1992. Potential Improvements in rate of genetic gain from marker-assisted selection in dairy cattle breeding schemes. J. Dairy Sci. 75: 1651-1659.

Miller, J. 2004. Subject: Dairy economics. http://www.ers.usda.gov/Briefing/Dairy/. Accessed June 15, 2005.

Monti, G. B.A. Tenhagen, and W. Heuwieser. 1999. Culling policies in dairy herds: A Review. J. Vet. Med. 46: 1-11.

Nebel, R. L. and M.L. McGilliard. 1993. Interaction of high milk yield and reproductive performance in dairy cows. J. Dairy Sci. 76: 3257-3268.

Pryce, J.E., M.D. Royal, P.C. Garnsworthy, I.L. Mao. 2004. Fertility in the high-producing dairy cow. Liv. Prod. Sci. 86:125-135.

Roche, J.F., D. Mackey, and M.D. Diskin. 2000. Reproductive management of postpartum cows. Anim. Reprod. Sci. 60-61: 703-712.

Royal, M.D., A.O. Darwash, A.P.F. Flint, R. Webb, J.A. Woolliams, G.E. Lamming. 2000a. Declining fertility in dairy cattle: changes in tradition and endocrine parameter of fertility. Anim. Sci. 70: 487-501.

Royal, M.D., G.E. Mann, A.P.F. Flint. 2000b. Strategies for reversing the trend towards subfertility in dairy cattle. Vet. J. 160: 53-60.

Sax K. 1923. The association of size differences with seed-coat pattern and pigmentation in Phaseolus vulgaris. Genetics 8:552–60

Schmitt, E.J., T. Diaz, M. Drost, and W.W. Thatcher. 1996. Use of gonadotropin-releasing hormone agonist or human chorionic gonadotropin for timed insemination in cattle. J. Anim. Sci. 74: 1084-1091.

Schrick, F.N., M.E. Hockett, A.M. Saxton, M.J. Lewis, H.H. Dowlen, S.P. Oliver. 2001. Influence of subclinical mastitis during early lactation on reproductive parameters. J. Dairy Sci. 84: 1407-1412.

Schrooten, C., H. Bovenhuis, W. Coppieters, J.A.M. van Arendonk. 2000. Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle. J. Dairy Sci. 83: 795-806.

Schrooten, C., M.C.A.M. Bink, and H. Bovenhuis. 2004. Whole genome scan to detect chromosomal regions affecting multiple traits in dairy cattle. J. Dairy Sci. 87:3550-3560.

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

Seaton G., C.S. Haley, S.A. Knott, M. Kearsey, P.M. Visscher. 2002. QTL Express: mapping quantitative trait loci in simple and complex pedigrees. Bioinformatics 18: 339-340.

Shanks, R.D., A.E. Freeman, and F.N. Dickinson. 1981. Postpartum distribution of costs and disorders of health. J. Dairy Sci. 64:683.

Soller, M. 1998. Marker-assisted selection in animal biotechnology. pp 387-406 In: Agriculture Biotechnology: Books in soils, plants and the environment. A. Altman (ed), New York Marcel Decker, Inc., NY.

Sonstegard, T.S., C.P. Van Tassell, and M.S. Ashwell. 2001. Dairy Cattle Genomics: Tools to accelerate genetic improvement? J. Anim. Sci. 73(E. Suppl.): E307-E315.

Streelman, J.T. and T.D. Kocher. 2000. From phenotype to genotype. Evolution and Development. 2/3: 166-173.

Thallman, R.M., G.L. Bennett, J.W. Keele, and S.M. Kappes. 2001. Efficient computation of genotype probabilities for loci with many alleles: I. Allelic peeling. J. Anim. Sci. 79: 26-33.

Thallman, R.M. 2002. User's Manual for GenoProb Version 2.0: Computation of Genotype and phase probabilities in complex pedigrees by iterative allelic peeling. USDA-MARC, Clay Center, NE.

VanRaden, P.M. and G.R. Wiggans. 1991. Derivation, calculation and use of national animal model information. J. Dairy Sci. 74: 2737-2746.

VanRaden, P.M. and A.J. Seykora. 2003. Net merit as a measure of lifetime profit: 2003 revision. USDA AIPL Annual Research Report. http://www.aipl.arsusda.gov/reference/nmcalc.htm

VanRaden, P.M., A.H. Sanders, M.E. Tooker, R.H. Miller, H.D. Norman, M.T. Kuhn, and G.R. Wiggans. 2004a. Development of a National Genetic Evaluation of cow fertility. J. Dairy Sci. 87: 2285-2292.

VanRaden, P.M. 2004b. Invited Review: Selection on net merit to improve lifetime profit. J. Dairy Sci. 87: 3125-3131.

Van Tassell, C. Genetic Selection tools in a genomics era. http://www.aipl.arsusda.gov/publish/presentations/MISC04/Expo_cvt.ppt#1. Accessed June 15, 2005

Visscher, P.M., R. Thompson, and C.S. Haley. 1996. Confidence intervals in QTL mapping by bootstrapping. Genetics. 143: 1013-1020.

Wall, E., S. Brotherstone, J.A. Woolliams, G. Banos, and M.P. Coffey. 2003. Genetic evaluation of fertility using direct and correlated traits. J. Dairy Sci. 86: 4093-4102.

Waltner, J.P., J.P. McNamara, and J.K. Hillers. 1993. Relationship of body condition score to production variables in high producing Holstein dairy cattle. J. Dairy Sci. 76: 3410-3419.

Webb, R., M.D. Royal, J.G. Gong, P.C. Garnsworthy. 1999. The influence of nutrition on fertility. Cattle Practice, Vol. 7, Part 3.

Weigel, K.A. 2004. Improving the reproductive efficiency of dairy cattle through genetic selection. J. Dairy Sci. 87: (E.Suppl): E86-E92.

Weller, J.I., Y. Kashi, and M. Soller. 1990. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. J. Dairy Sci. 73: 2525-2537.

Windig, J.J., M.P.L. Calus, and R.F. Veerkamp. 2005. Influence of herd environment on health and fertility and their relationship with milk production. J. Dairy Sci. 88: 335-347.

Wildman, E.E., G.M. Jones, P.E. Wagner, R.L. Boman, H.F. Trout, Jr., and T.N. Lesch. 1982. A dairy cow body condition scoring system and its relationship to selected production characteristics. J. Dairy Sci. 65: 495.

Zhang, Q., D. Boichard, I. Hoeschele, C. Ernst, A. Eggen, B. Murkve, M. Pfister-Genskow, L.A. Witte, F.E. Grignola, P. Uimari, G. Thaller and M.D. Bishop. 1998. Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. Genetics 149: 1959-1973.

Ziegle, J.S., Y. Su, K.P. Corcoran, L. Nie, P.E. Mayrand, L.B. Hoff, L.J. McBride, M.N. Kronick, and S.F. Diehl. 1992. Application of Automated DNA sizing technology for genotyping microsatellite loci. Genomics. 14: 1026-1031.

# Fine Mapping Quantitative Trait Loci Affecting Health and Reproduction in US Holstein Cattle on Chromosome 18

S.A. Muncie and M.S. Ashwell
North Carolina State University Raleigh, NC

## Abstract

Continued genetic improvement for milk production has been associated with decreased fertility in US Holstein cattle. A previous study (Ashwell et al., 2004) identified a putative quantitative trait locus (QTL) affecting daughter pregnancy rate at 54 cM on chromosome 18 in one Holstein grandsire family. The goal of this research is to determine the validity of the putative QTL using additional markers and an extended pedigree. Thirteen microsatellite markers located throughout the chromosome were genotyped in 973 animals that were descendents of the original grandsire in which the QTL was identified. Sons of the grandsire as well as six grandson and six great-grandson families of the original grandsire were selected for this study (range of 16 to 169 sons per family). In analysis of the sons using QTL Express, the same putative QTL affecting daughter pregnancy rate was detected and placed at 45 cM. In a joint analysis of thirteen of the largest families, each containing 10 or more sons, a significant QTL for daughter pregnancy rate was detected at 27 cM. QTL affecting daughter pregnancy rate was detected in two additional sub-families (Family II-5 and Family III-2), indicating a putative QTL on chromosome 18 affecting daughter pregnancy rate is most likely segregating within this pedigree. Across-family analysis also detected putative QTL affecting productive life at 35 cM, somatic cell score at 33 cM and percent difficult births at 72 cM. Analysis of individual families identified eight significant putative QTL and six suggestive putative QTL at the chromosome-wise level affecting somatic cell score,

productive life, calving ease, percent difficult births, milk yield, fat yield, protein yield and fat percent. A complex pedigree analysis is underway to make full use of statistical power to refine the QTL region affecting fertility. Further verification of the QTL effects identified in this study will allow identification of positional candidate genes to be applied in marker assisted breeding programs.

(**Key words**: quantitative trait locus, pregnancy rate, *Bos taurus* chromosome 18)

# Introduction

Milk is a very important farm product and ranks second only to beef within livestock industries (Miller, 2004). In 2000, the United States farm milk sales were approximately 10% over all other farm sales, reaching more than $21 billion in revenue (Collins, 2000). Milk production per cow has nearly doubled since the 1970's (Miller, 2004), so with more milk per cow, there are fewer cows required to produce the same amount of milk for profit. While the increases in production have been positive, negatively correlated effects have also been observed.

Continued genetic improvement for milk production has been associated with decreased fertility in US Holstein cattle. In the past, conception rates using artificial insemination methods in the United States have been reported at approximately 55% (Casida, 1961). These rates are declining and are now reported between 35% and 45% (Schmitt et al., 1996; Darnsfield et al., 1998). The United States is not the only country seeing a decline in dairy cattle reproductive efficiency. The United Kingdom (Royal et al., 2000), Australia (Macmillan et al., 1996), and Ireland (Roche, 2000) are examples of other countries experiencing diminishing conception rates. The profitability of a dairy cattle herd is greatly affected by reproductive performance. The lactation cycle is triggered by pregnancy; therefore, milk production is dependent on the cow becoming pregnant (Lucy, 2001). The economic importance of fertility becomes evident if one considers that 40% of direct health costs can be attributed to reproductive factors (Shanks et al., 1981). Approximately $200,000 can be lost each year due to reproductive problems in a 500-cow herd (http://www.100daycontract.com).

In the past, genetic selection has been primarily directed towards higher milk production using traditional selection methods. This type of selection makes use of breeding values, determined through progeny testing, and phenotypic observations to achieve genetic improvement. Progeny testing can be time consuming and expensive. Also, this method may not be ideal for fertility traits because fertility is a complex, polygenic trait, influenced by genetics and environment and has a low heritability of around 0.04 (VanRaden et al., 2004). The ability to use marker-assisted selection for a lowly heritable trait like pregnancy rate would help improve reproduction while maintaining higher milk yields.

There have been numerous genome scans to identify QTL affecting economically important traits in dairy cattle, with at least one QTL reported on every chromosome. Most studies have concentrated on QTL affecting milk production and milk composition. The incorporation of the daughter pregnancy rate (DPR) calculation into national dairy cattle evaluations in 2003 has opened the door for the detection of QTL affecting the fertility of the dairy cow. Recently, genome scan studies have successfully identified putative fertility QTL on various chromosomes throughout the bovine genome. Specifically, there have been numerous studies that have identified QTL affecting fertility on chromosome 18. For example, Kühn et al. (2003) performed a genome scan for functional traits affecting udder health, functional herd life and fertility in a study of 16 German Holstein half-sib families. The interval between BM2078 and TGLA227 appeared to have the most significant effects for fertility (dystocia, nonreturn rate of 90 days, stillbirth). Schrooten et al. (2004) studied 20 Holstein Friesian families and reported QTL affecting multiple traits including milk fat and protein percentages along with the interval calving to first insemination trait in the interval between BM7109 and ILSTS002. In the same region, a preliminary study by

Ashwell and coworkers (2004) identified a putative QTL for daughter pregnancy rate at 54 cM on chromosome 18 in a popular US Holstein sire.

There appear to be significant effects spanning chromosome 18; however, the marker intervals BM7109-ILSTS002 (Ashwell et al., 2004; Schrooten et al. 2004; Kühn et al., 2003; Boichard et al., 2003) and BM2078-TGLA227 (Kühn et al., 2003; Boichard et al., 2003) seem to be common across studies for fertility traits on chromosome 18. This observation may indicate a higher likelihood that one or more QTL may be found in one of these intervals, and fine-mapping these regions would be advantageous.

The object of this study is to validate and refine the location of the putative QTL affecting DPR found in our preliminary study (Ashwell et al., 2004) through genotyping additional descendents from the popular Holstein family at additional microsatellites chosen based on proximity to the putative QTL.

## Materials and Methods

### Source of Materials

Semen samples from the Dairy Bull DNA Repository (DBDR; Da et al., 1994) and the Cooperative Dairy DNA Repository (CDDR; Ashwell and Van Tassell, 1999) were obtained for all available descendents of a popular Holstein bull in which a QTL affecting DPR has been detected. Both contain semen from proven (DBDR, CDDR) and unproven (CDDR) bulls donated by AI organizations in North America for QTL studies. The pedigree structure with number of descendents in each family used in this study is shown in Figure 3. DNA was extracted from semen using a lysis/phenol-chloroform protocol as previously described (Ashwell et al., 1996).

Phenotypic observations from daughter deviations (DD) and PTA of milk yield and composition (e.g., protein and fat yields and protein and fat percentages), somatic cell score (SCS), productive life (PL), and pregnancy rate were obtained from Animal Improvement Programs Laboratory (AIPL) of the United States Department of Agriculture-Agricultural Research Service (USDA-ARS). Phenotypic information on calving ease and percent difficult births were supplied by the National Association of Animal Breeders (NAAB).

**Microsatellite Marker Amplification**

Thirteen microsatellite markers (MS) were selected from the USDA Meat Animal Research Center (MARC) bovine map (http://www.marc.usda.gov/genome/genome.html). Markers were selected in the region previously identified to contain the QTL for pregnancy rate in a preliminary study and based on possibility of many observable alleles segregating in the population. Each marker was amplified using PCR as previously described (Ashwell et al., 1997). The appropriate annealing temperatures were optimized using a temperature gradient ranging from 50-65°C. A 1% agarose gel was used to visualize bands to assess the optimal annealing temperature. Marker information, including PCR annealing temperatures, primer sequences, and linkage map locations can be reviewed in Appendix 1.

**Genotyping**

PCR products were analyzed using the Applied Biosystems 3100 Genetic Analyzer (ABI3100; Foster City, CA). For each set of primers, a dilution series, ranging from undiluted to a 1:100 dilution, was evaluated on the ABI3100 to ensure optimal peak heights. Over 12,500 genotypes were generated in this study. Appropriate dilutions for each of the markers are presented in Appendix 1. Genotypes were determined for all animals using GeneMapper® version 3.7 (ABI, Foster City, CA; Appendix 2) and the GS-HD400 ROX

lane standard. Questionable alleles were manually analyzed to discern if a genotype could be determined. Animals with questionable alleles were re-amplified and reanalyzed on the ABI3100. Genotypes were not assigned to animals whose alleles were still undistinguishable after two genotyping attempts.

**Statistical Analysis**

All genotyping data were first analyzed using the software program GenoProb (USDA-MARC, Clay Center, NE). GenoProb is a database program used to calculate probabilities for unknown genotypes within complex pedigrees (Thallman et al., 2001). Data files were input into the program including the animals' genotypes, a complete pedigree, and a map file as previously described (Thallman, 2002). After the probabilities for the genotypes were calculated, a cut-off value of genotype probabilities >0.95 was used. Animals whose genotype probabilities did not reach this cut-off were removed from further analysis.

Corrected data were analyzed using a regression interval mapping program (Seaton et al., 2002; QTL Express; http://qtl.cap.ed.ac.uk/). Half-sib analysis was used to investigate the data within and across the families. One QTL was fitted in the model. Chromosome-wise significance thresholds ($P< 0.05$ and $P<0.01$) were calculated using 1000 permutations. One thousand re-samples were selected for bootstrapping to determine the 95% confidence intervals. Analysis was conducted at 1 cM intervals and phenotypic values were weighted using their respective reliabilities.

## Results and Discussion

Previously we identified (Ashwell et al., 2004) a QTL affecting DPR in a popular Holstein family. Here we genotyped additional descendents of the popular bull at more

microsatellites to try to validate and fine-map the DPR QTL. Thirteen markers covering *Bos taurus* (BTA) chromosome 18 were genotyped in 973 animals to generate over 12,500 genotypes for this study (Table 2). On average, 92% of all the animals were genotyped for every marker. Out of the 27 sires genotyped, the average heterozygosity was approximately 47%. Only families that had more then 10 descendents could be included in the analysis, for a total of thirteen families (940 out of the 973 genotyped). In analysis of 17 economically important traits in the 13 families, 221 within-family tests and 17 across-family tests were performed; a total of 238 individual tests. Therefore, it is expected that there will be 12 significant QTL effects at a $P<0.05$ and 3 significant QTL effects at a $P<0.01$ detected by chance due to multiple testing. QTL effects at a chromosome-wise significance of $P<0.05$, expressed here as suggestive QTL, and a $P<0.01$, expressed here as significant QTL, from within- and across-family analysis, are reported in Tables 3 and 4, respectively.

**GenoProb**

Analysis of genotyping data with the GenoProb program resulted in discovery of two markers with genotyping errors, which without this further analysis would have gone undetected. The genotype probabilities for BMS1355 were, on average, very low. After reanalysis of the genotyping data using GeneMapper the alleles were not spaced as would be expected from a dinucleotide repeat microsatellite marker. There appeared to be only a 1-nucleotide difference between the alleles. This could have possibly been caused by a random mutation in the microsatellite template or PCR amplification errors. Subsequently, many alleles were called incorrectly and the correct genotypes were undistinguishable. Attempts to further optimize PCR amplification conditions failed to correct this problem. Consequently, BMS1355 was excluded from all subsequent analysis.

The GenoProb analysis also identified that the marker DIK2779 most likely has null alleles. GeneMapper does not recognize null alleles, thus causing incorrect designation of heterozygous animals as homozygotes, which may lead to skewed experimental results. The affected genotypes were changed to include a zero to represent the null allele (i.e. 0, 221) in order to keep this marker's genotyping data in the analysis.

**Marker Effects on Somatic Cell Score and Productive Life**

QTL affecting SCS were detected in this study. Somatic cell score is an indirect measure of mastitis, a bacterial infection of the udder. A significant QTL affecting PTA SCS was detected in the across-family analysis at 34 cM (Figure 4). Analysis of individual families identified a suggestive QTL effect on SCS in Family I-1 at 10 cM and a significant QTL affecting PTA SCS in Family II-1 at 34 cM (Figure 5), in the same location as the QTL detected in the across-family analysis (INRA121 to BM8151) (Figure 4). Previous studies have identified SCS QTL effects, but the location of the QTL is more telomeric. Both Schulman et al. (2004) and Kühn et al. (2003) identified SCS QTL effects in similar regions, at 113 cM and 117 cM respectively. Schrooten and coworkers (2000) identified a SCS QTL effect at 70 cM. The dissimilarity in QTL location might be due to differences in markers chosen for analysis, method of statistical analysis, or the linkage map used in the studies. The 95% confidence interval from within-family analysis in our study ranged from 0 cM to 78 cM. With the confidence interval taken into account, the findings from Schrooten and coworkers (2000) would then fall into the same range, suggesting at least one QTL is being detected on chromosome 18.

Productive life is defined as the time period the animal is active in the herd for milking, before she is removed from the herd

(http://www.aipl.arsusda.gov/reference/scspl/pl.htm). A significant QTL effect on productive life in the across-family analysis was also identified at 35 cM (Figure 4). DD for productive life QTL effects were found in three families in similar locations on BTA18 (Figure 6). Suggestive QTL effects for productive life were identified in Family I-1 (47 cM). Significant QTL effects for productive life were seen in Family III-3 (33 cM) and Family III-6 (41 cM). The variance seen in SCS QTL effects may be related to the effects seen in the productive life analysis. Reproduction is the most prevalent reason for culling; however, mastitis and low milk production are also incentives for removal from the herd (Bascom and Young, 1998). Consequently, an animal that has mastitis has a greater chance of being removed from the herd, causing her to have a shorter productive life. The across-family PL QTL is located in the same region as the SCS QTL seen in Family II-1. Since SCS is incorporated into the PL calculation we may be detecting the same DNA region causing the observed variation in all four families.

It may be difficult to compare productive life findings between studies due to variations in measurements. For example, Kühn and colleagues (2003) reported a putative QTL for functional herd life at marker BM2078 on BTA18, which is a more distal location than that reported here. In Germany, PL evaluations are adjusted for animals that are voluntarily culled due to low milk production (Kühn et al., 2003). In the US, PL evaluations include voluntary and involuntary culls, and may explain the difference in placement of the two PL QTL reported by both groups.

Traits such as calving ease and percentage of difficult births may also be associated with the QTL affecting productive life. Percentage difficult birth measures how hard for the daughters of a particular sire to give birth compared to the average daughter and calving ease

is the tendency of a sire's progeny to be easier or more difficult for the dams to birth than an average calf. An animal that has a difficult time during parturition will have a higher likelihood of being removed from the herd. Significant QTL affecting calving ease and percentage of difficult births were detected in Family III-6 at approximately 72 cM (Data not shown). In the same region, a suggestive QTL effect was also detected in the across-family analysis (Data not shown). In the study by Kühn and coworkers (2003), QTL affecting dystocia, or difficult birthing, were identified in the marker interval TGLA357 to BM7109 and at marker BM2078 on BTA18. These QTL placements are in agreement with the QTL affecting productive life reported here. Our study and previous studies provide additional evidence that the marker intervals from BM7109-ILSTS002 and BM2078-TGLA227 may harbor QTL for a trait affecting the length of time an animal will be active in a milking herd.

**Marker Effects on Milk Traits**

Five PTA and five DD traits were evaluated to identify QTL affecting milk production and composition on BTA18. No QTL effects for milk traits were detected in the across-family analysis. If some of the sires are homozygous for the possible QTL then it is possible for the cumulative affect to not be detected during across family analysis. When data were analyzed within families, a significant putative QTL affecting milk yield was identified in Family II-2 towards the centromere between BMS1355 and BMS2213 (Figure 7). In the same family, a suggestive QTL affecting protein yield was identified in the same interval. QTL affecting milk protein yield and milk yield on BTA18 were also identified in a genome scan of Norwegian dairy cattle (Olsen et al., 2002). In that study QTL were placed in the middle of the chromosome for milk yield (39 cM; INRA121-BM7109) and towards the telomere for protein yield (79 cM; ILSTS002-EAC). Boichard et al. (2003) also identified a

QTL affecting protein percent close to the centromere using three French dairy breeds (10 cM; ABS13). Results from these studies indicate that at least one QTL affecting milk traits lies on BTA18. Inconsistencies in QTL placement may be explained by different detection methods and breeds used in the three studies. The 95% confidence interval for the QTL affecting milk yield and protein yield in our study is large, ranging from 3 cM to 79 cM, encompassing those reported by Olsen et al. (2002) and Boichard et al. (2003). Therefore, the significant effects observed in all three studies may be the result of the same QTL.

A suggestive QTL effect was identified in Family I-1 for fat yield at 52 cM between marker BMS2914 and ILSTS002. In the preliminary study in Family I-1, Ashwell and coworkers (2004) reported a QTL affecting fat yield more distal to the one reported here in the same family (84 cM; BM6507-TGLA227). The accumulation of additional descendents, a change in marker order and the addition of more markers to the new analysis may explain the discrepancy in QTL placement. A suggestive QTL affecting fat percent was also identified in Family III-4 at 53 cM. To date no other studies have reported QTL affecting fat components in this region of the chromosome and further verification of the QTL is necessary.

**Marker Effects on Pregnancy Rate**

The daughter pregnancy rate trait was introduced into United States genetic evaluation programs in August of 2003 and has allowed for detection of QTL affecting fertility. In a preliminary study by Ashwell and coworkers (2004), a QTL was detected on BTA18 at 54 cM in the same DBDR family used here (I-1). In our study, a QTL affecting DD daughter pregnancy rate was again detected. The across-family analysis identified a significant QTL, but appeared in two significant QTL peaks in the interval BMS2213-

INRA121 (Figure 4). The average QTL location in the across-family analysis was 27 cM, with a 95% confidence interval ranging from 8.0 cM to 74.0 cM, spanning 66 cM. When analyzed separately, the significant QTL effect in Family I-1 was placed closer to the centromere then previously reported (Ashwell et al., 2004), at 45 cM within the interval DIK2128-BM7109 (Figure 8). It appears that two sub-families may be segregating for the DPR QTL; Family II-5 at 53 cM (BMS2914-ILSTS002) and Family III-2 at 44 cM (DIK2128-BM7109).

Other groups have reported QTL affecting fertility traits on BTA18. Kühn and coworkers (2003) reported a QTL affecting nonreturn rate after 90 days, a measure of repeat inseminations, in the interval between BM2078-TGLA227 when their data were analyzed across sixteen families. This location is distal to the DPR QTL reported here, but falls within the 95% confidence interval. Our study provides further evidence of a QTL affecting fertility on BTA18.

In a multitrait analysis study by Schrooten et al. (2004), significant chromosomal regions affecting interval between calving and first insemination and milk protein and fat percentage traits were identified on BTA18. The QTL was placed at 68 cM. The results from their study may indicate QTL affecting milk traits and fertility are not independent, but the authors could not distinguish between one QTL with pleiotropic effects and multiple QTL that are closely linked (Schrooten et al., 2004). In our study, Family I-1 that showed the largest QTL effect on DPR also had a QTL affecting milk fat yield in a similar location (52 cM) to that reported by Schrooten and coworkers, providing further evidence that a QTL having pleiotropic effects may be in this location.

In addition to a possible link between genes affecting fertility and milk components, there may also be an indirect effect of disease status on animal fertility. In a study investigating how the disease mastitis affects ovarian function in dairy cattle, Huszenicza et al. (2005) found that cows suffering from mastitis had a delayed ovarian cycle when compared to healthy animals. The bacteria responsible for causing these infections release cytokines and have been suggested to be the causative agent in the delay of ovulation in postpartum dairy cattle (Sheldon and Dobson, 2004). As previously suggested, cows with mastitis will have greater problems with fertility (Huszenicza et al., 2005), and thus, detection of SCS QTL may be related to the DPR QTL. In Family II-1 a QTL affecting SCS is located at 34 cM (Figure 5), in the same region as the QTL affecting DPR and SCS in across-family analysis. Therefore a QTL affecting both traits may exist in this region.

Further analysis is needed to extract the full power achieved by this complex pedigree. In using QTL Express, each half-sib family is considered to be unrelated to the other families; however, this assumption is incorrect given the pedigree structure in Figure 3. In order to fully harness the statistical power of this complex, extended pedigree, different statistical analysis is needed. Markov Chain Monte Carlo methods or variance component methods that are able to handle a complex pedigree may be useful in further analysis (Meuwissen and Goddard, 2004; Zhang et al., 1998). This will allow for more specific positioning of the QTL, estimation of QTL effects and identification of the number of QTL affecting fertility traits on BTA18 in these animals. After further refinement of the QTL region to approximately a 10 cM interval, positional candidate genes can be identified. The completion of the bovine genome sequence and comparative radiation hybrid maps already in place will assist in this process.

## Conclusions

Previously, a QTL affecting DPR was identified in one half-sib Holstein family. In this study that family was expanded to include all available descendents that were genotyped at 13 microsatellite markers. Putative QTL on BTA18 affecting SCS, PL, milk traits and DPR were identified. Putative QTL affecting PL and SCS, which is incorporated into the PL calculation, were identified at approximately 34 cM. QTL affecting milk traits were detected towards centromere and interstitial regions of the chromosome, providing evidence of a QTL affecting milk traits segregating in this pedigree. Genotyping of additional sons of the popular Holstein sire provided additional evidence of a QTL affecting DPR identified in our preliminary study, placing the QTL at 45 cM. Many studies have identified putative QTL either directly or indirectly affecting reproductive health of the dairy cow. This study provides additional evidence that QTL affecting fertility exist on BTA18. Further analysis is underway to refine the location of the QTL to a 10 cM interval so positional candidate genes can be identified and evaluated for causative mutations.

## Acknowledgements

The authors would like to thank the contributing AI organizations for providing semen to the Dairy Bull DNA Repository and the Cooperative Dairy DNA Repository. The authors would also like to thank Audrey O'Nan, Shelly Nolin, and Keri Boyette for their outstanding technical assistance in the laboratory.

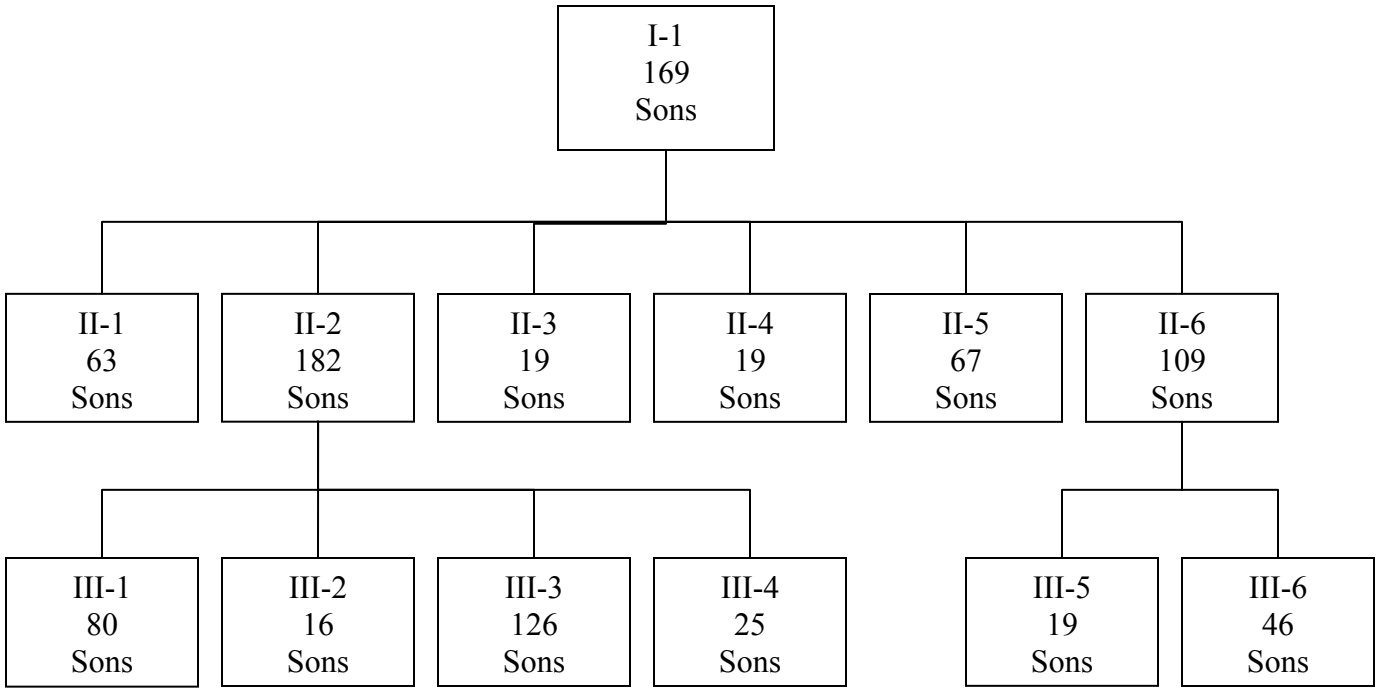


**Figure 3**: Pedigree of original sire and descendents, representing four generations from the original sire. Each box indicates the number of sons per sire genotyped in this study. Some sons were genotyped but did not have available samples from their sons and are not shown in this pedigree.

**Table 2.** Genotyping statistics from 973 animals genotyped for fine-mapping project.

| Marker Name | Relative Position | Alleles | Sires Genotyped (No.)[a] | Heterozygous Sires (No.) | Total Sons Genotyped (No.) | Percent of Total Genotyped |
|---|---|---|---|---|---|---|
| BMS1355[b] | 2.9 | 6 | 18 | 14 | 874 | 90% |
| BMS2213 | 24.5 | 8 | 17 | 13 | 848 | 87% |
| INRA121 | 30.2 | 8 | 17 | 12 | 937 | 96% |
| BM8151 | 40.2 | 6 | 16 | 8 | 919 | 94% |
| DIK2128 | 40.2 | 7 | 16 | 6 | 884 | 90% |
| BM7109 | 47.0 | 6 | 16 | 11 | 955 | 98% |
| BMS2914 | 50.0 | 6 | 18 | 14 | 910 | 93% |
| ILSTS002 | 54.7 | 5 | 19 | 19 | 891 | 91% |
| DIK2779 | 56.3 | 5 | 17 | 14 | 867 | 89% |
| BMS929 | 61.2 | 2 | 15 | 8 | 919 | 94% |
| BB710 | 62.1 | 5 | 17 | 12 | 931 | 95% |
| BM2078 | 76.8 | 8 | 16 | 13 | 891 | 91% |
| TGLA227 | 84.0 | 10 | 18 | 17 | 851 | 87% |

[a]Out of 27 total sires.
[b]BMS1355 was not used in the final analysis due to allele calling problems.

**Table 3.** Chromosome-wise QTL effects on chromosome 18 from within-family analysis.

| Family | Trait[a] | Location (cM) | F Statistic | Estimated allelic difference | SE | Marker Interval |
|---|---|---|---|---|---|---|
| I-1 | DD_SCS | 0 | 8.55 | 0.14 | 0.05 | Centromere-BMS1355 |
| I-1 | PTA_SCS | 10 | 7.83 | 0.08 | 0.03 | BMS1355-BMS2213 |
| **I-1** | **DD_DPR** | **45** | **14.62** | **0.09** | **0.02** | **DIK2128-BM7109** |
| **I-1** | **PTA_DPR** | **47** | **18.94** | **0.61** | **0.14** | **BM7109-BMS2914** |
| I-1 | PTA_PL | 47 | 7.79 | 0.50 | 0.18 | BM7109-BMS2914 |
| I-1 | PTA_fatyld | 52 | 7.19 | 10.62 | 3.96 | BMS2914-ILSTS002 |
| I-1 | DD_fatyld | 52 | 7.01 | 11.63 | 4.39 | BMS2914-ILSTS002 |
| **II-1** | **PTA_SCS** | **34** | **12.75** | **0.16** | **0.04** | **INRA121-BM8151** |
| II-1 | DD_SCS | 34 | 9.45 | 0.18 | 0.06 | INRA121-BM8151 |
| **II-2** | **PTA_milk** | **4** | **11.36** | **412.75** | **122.45** | **BMS1355-BMS2213** |
| **II-2** | **DD_milk** | **4** | **10.14** | **424.77** | **133.42** | **BMS1355-BMS2213** |
| **II-2** | **PTA_proyld** | **4** | **11.60** | **10.47** | **3.08** | **BMS1355-BMS2213** |
| II-2 | DD_proyld | 4 | 9.26 | 10.28 | 3.38 | INRA121-BM8151 |
| II-5 | DD_DPR | 53 | 7.38 | 0.10 | 0.04 | BMS2914-ILSTS002 |
| III-2 | DD_DPR | 44 | 8.98 | 0.20 | 0.07 | DIK2128-BM7109 |
| **III-3** | **DD_PL** | **33** | **16.81** | **1.66** | **0.40** | **INRA121-BM8151** |
| III-4 | PTA_fat% | 53 | 9.72 | 0.10 | 0.03 | BMS2914-ILSTS002 |
| III-4 | DD_fat% | 53 | 8.76 | 0.10 | 0.03 | BMS2914-ILSTS002 |
| **III-6** | **DD_PL** | **41** | **11.23** | **2.42** | **0.72** | **DIK2128-BM7109** |
| **III-6** | **CE** | **71** | **14.48** | **0.11** | **0.03** | **BB710-BM2078** |
| **III-6** | **PDB** | **72** | **15.09** | **1.88** | **0.48** | **BB710-BM2078** |

[a]Predicted transmitting ability (PTA) and daughter deviations (DD) for productive life (PL), fat percentage (fat%), protein percentage (pro%), milk yield (milk), fat yield (fatyld), protein yield (proyld), somatic cell score (SCS), calving ease (CE), percent difficult births (PDB) and daughter pregnancy rate (DPR). **Bold** type indicates significant QTL ($P<0.01$) and normal type indicates suggestive QTL ($P<0.05$).

**Table 4.** Chromosome-wise QTL effects on chromosome 18 from analysis across 13 families.

| Trait[a] | Location (cM) | F Statistic | Marker Interval |
|---|---|---|---|
| **DD_DPR** | **27** | **2.78** | **BMS2213-INRA121** |
| **PTA_SCS** | **33** | **2.6** | **INRA121-BM8151** |
| **DD_PL** | **35** | **2.45** | **INRA121-BM8151** |
| **PTA_DPR** | **47** | **3.07** | **BM7109-BMS2914** |
| PDB | 72 | 2.43 | BB710-BM2078 |

[a]Predicted transmitting ability (PTA) and daughter deviations (DD) for productive life (PL), somatic cell score (SCS), percent difficult births (PDB) and daughter pregnancy rate (DPR). **Bold** type indicates significant QTL ($P<0.01$) and normal type indicates suggestive QTL ($P<0.05$).

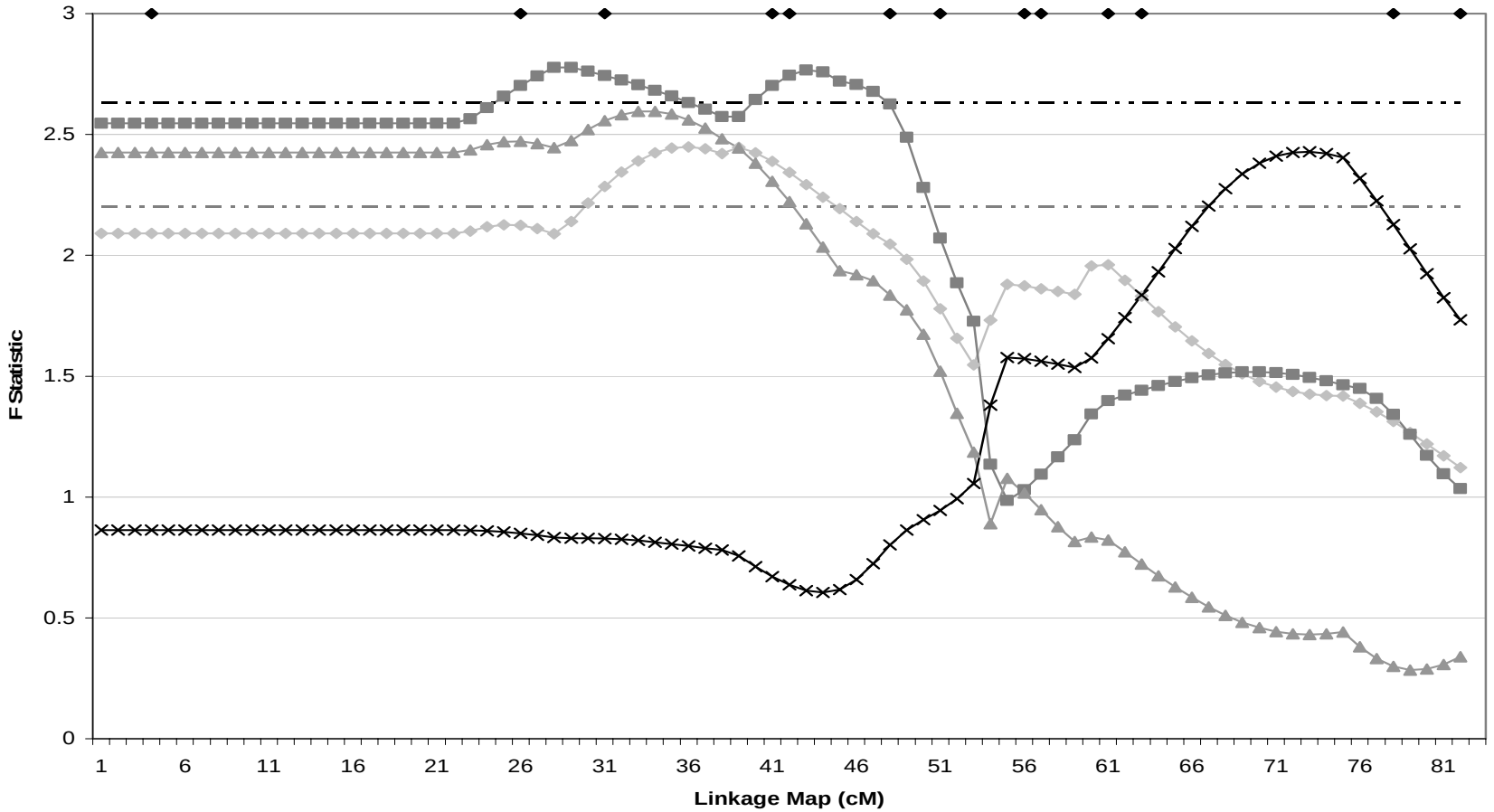


**Figure 4**: QTL effects on BTA18 from across-family analysis. Chromosome-wise thresholds for P< 0.05 and P< 0.01 are represented by grey and black dashed lines, respectively. Thresholds were averaged to view on a combined graph, with values ranging from 2.11-2.31 for P<0.05 and from 2.54-2.78 for P<0.01. The marker locations correspond to black diamonds (♦) at the top of the graph. DD traits are represented as follows: Productive life (♦), daughter pregnancy rate (■), somatic cell score (▲), and percent difficult births (**x**).

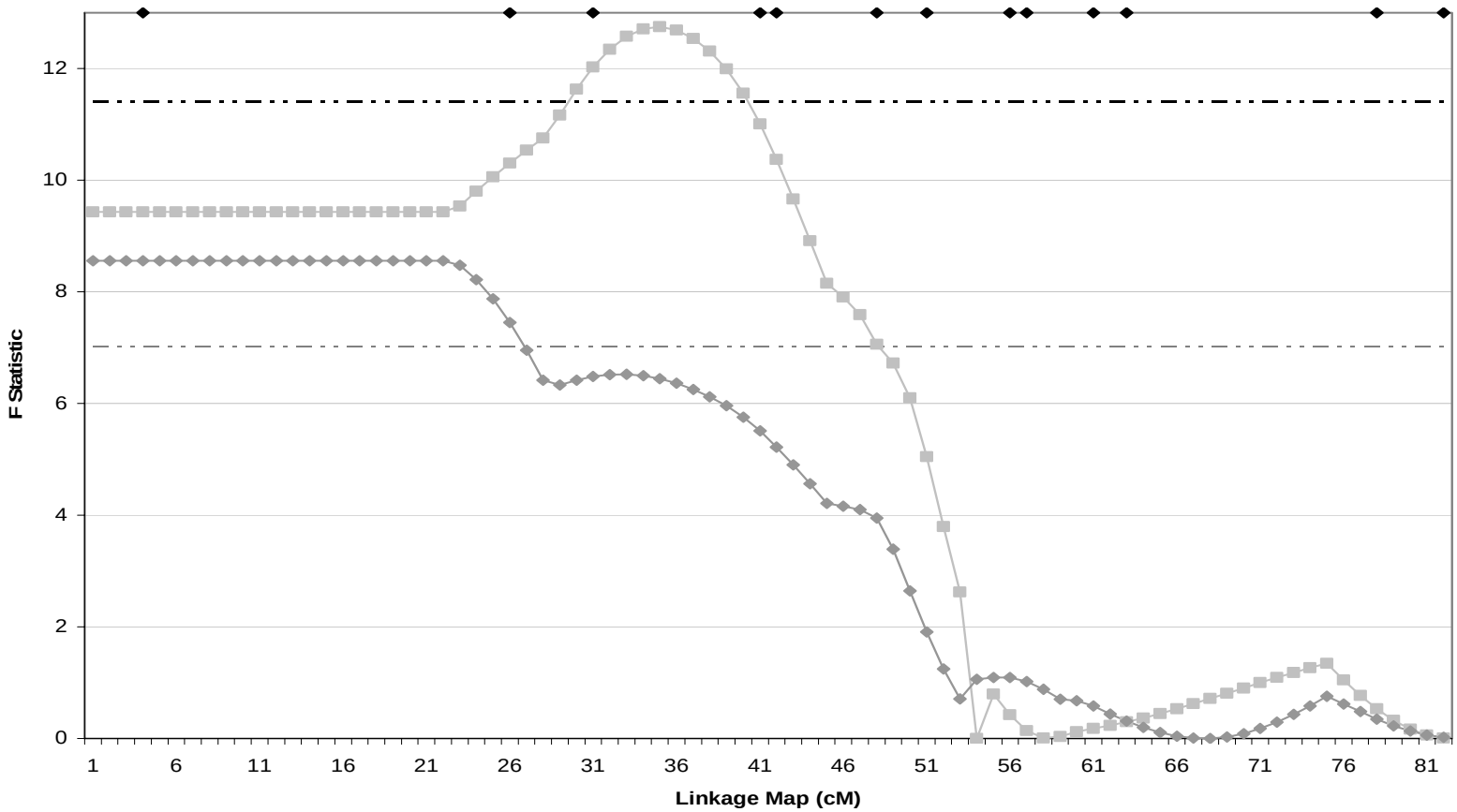


**Figure 5**: QTL effects on BTA18 from within-family analysis of somatic cell score (SCS). Chromosome-wise thresholds for P< 0.05 and P< 0.01 are represented by grey and black dashed lines, respectively. Thresholds were averaged to view on a combined graph, with values ranging from 6.57-7.54 for P<0.05 and from 10.2-12.59 for P<0.01. The marker locations correspond to black diamonds (♦) at the top of the graph. Families are represented as follows: Family I-1 (♦) and Family II-1 (■).

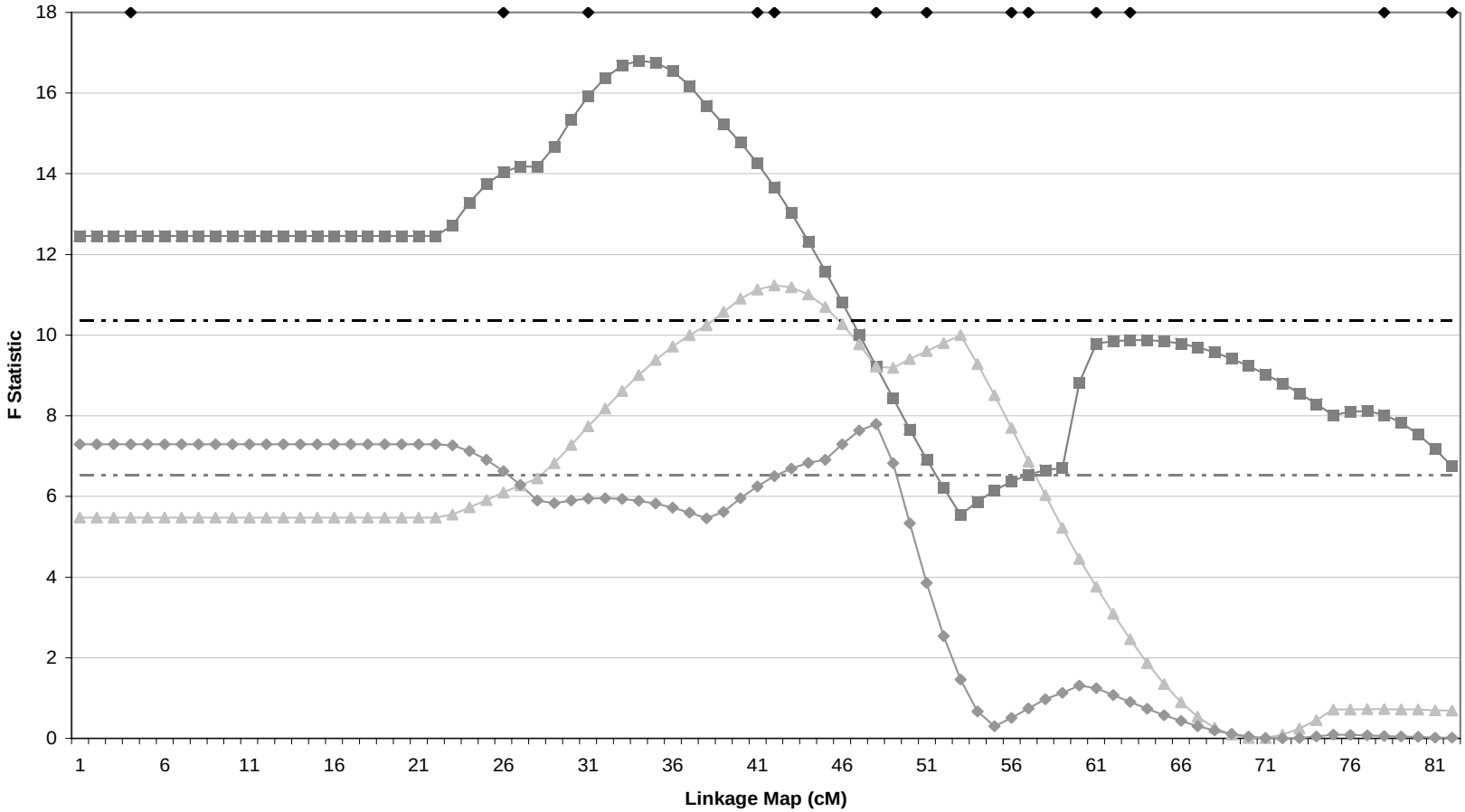


**Figure 6**: QTL effects on BTA18 from within-family analysis of Productive Life (PL). Chromosome-wise thresholds for P< 0.05 and P< 0.01 are represented by grey and black dashed lines, respectively. Thresholds were averaged to view on a combined graph, with values ranging from 5.78-7.22 for P<0.05 and from 9.51-10.70 for P<0.01. The marker locations correspond to black diamonds (♦) at the top of the graph. Families are represented as follows: Family I-1 (♦), Family III-6 (▲) and Family III-3 (■).

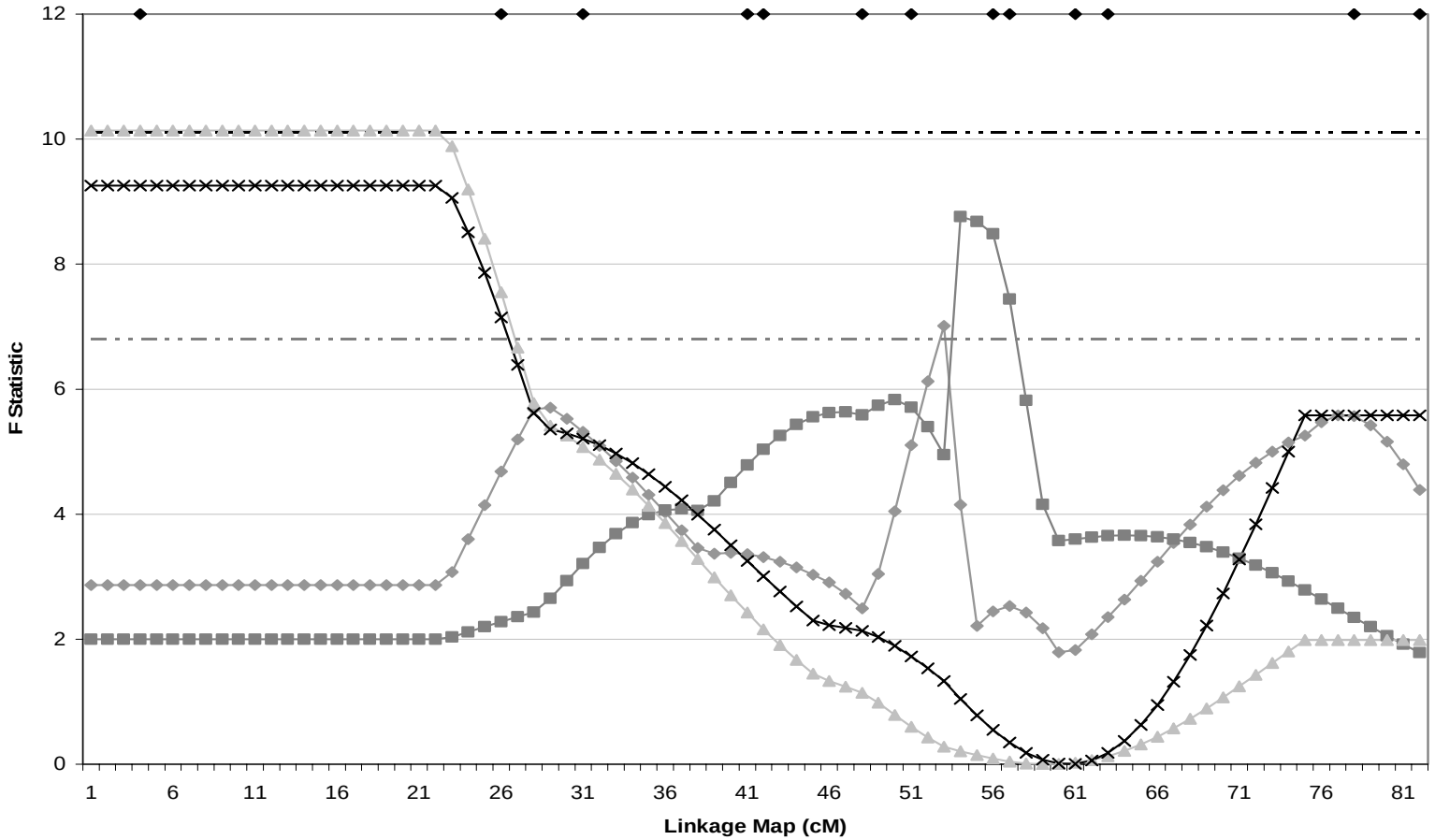


**Figure 7**: QTL effects on BTA18 from within-family analysis of milk production and composition traits. Chromosome-wise thresholds for P< 0.05 and P< 0.01 are represented by grey and black dashed lines, respectively. Thresholds were averaged to view on a combined graph, with values ranging from 6.12-8.17 for P<0.05 and from 9.43-11.5 for P<0.01. The marker locations correspond to black diamonds (♦) at the top of the graph. Families and DD traits are represented as follows: Family I-1 fat yield (♦), Family III-4 fat percentage (■), Family II-2 milk yield (▲), and Family II-2 protein yield (**x**).

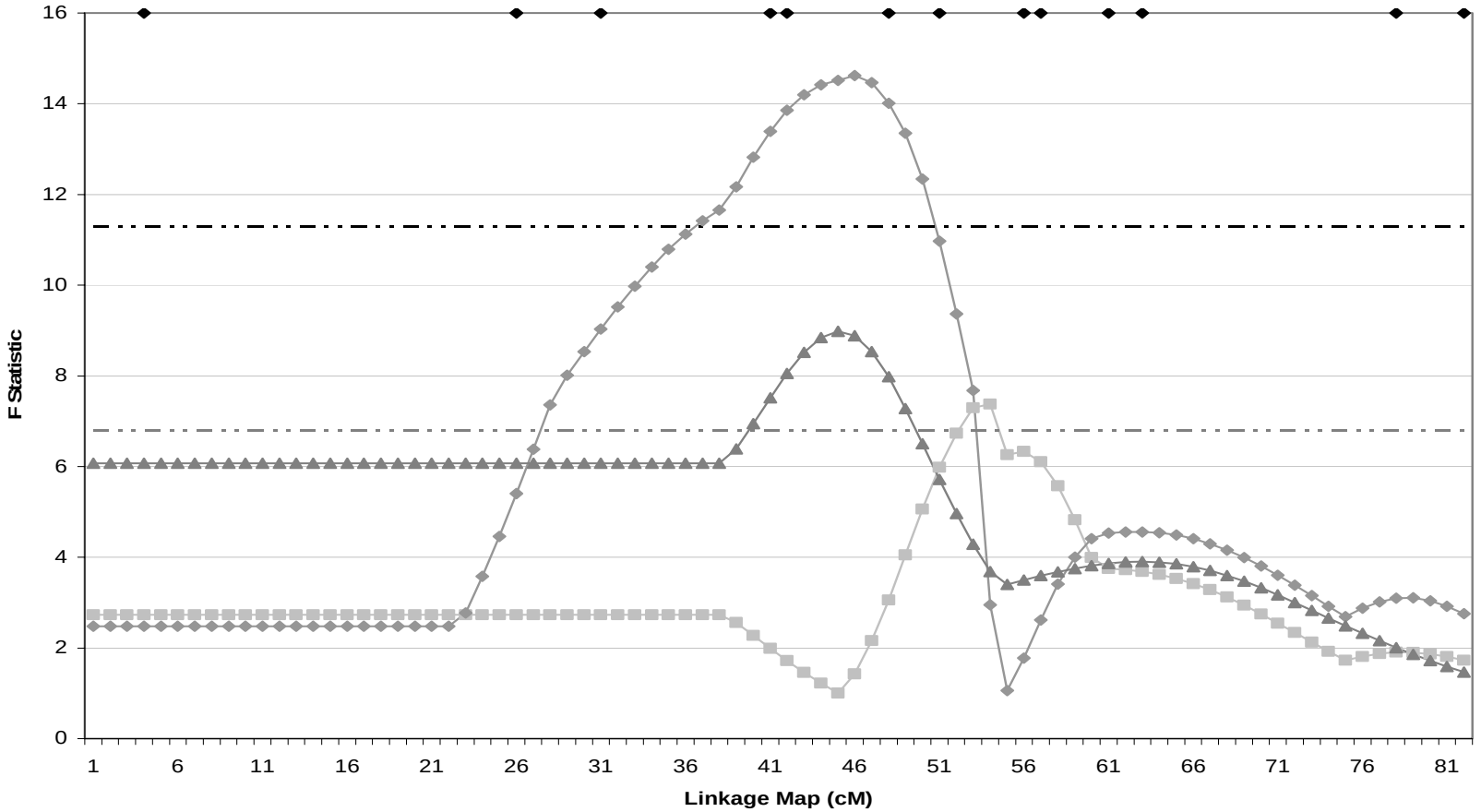


**Figure 8**: QTL effects on BTA18 from within-family analysis of daughter pregnancy rate (DPR). Chromosome-wise thresholds for P<0.05 and P<0.01 are represented by grey and black dashed lines, respectively. Thresholds were averaged to view on a combined graph, with values ranging from 6.44-7.6 for P<0.05 and from 9.43-13.5 for P<0.01. The marker locations correspond to black diamonds (♦) at the top of the graph. Families are represented as follows: Family I-1 (♦), Family II-5 (■), and Family III-2 (▲).

## Literature Cited

Ashwell, M.S., D.W. Heyen, T.S. Sonstegard, C.P. Van Tassell, Y. Da, P.M. VanRaden, M. Ron, J.I. Weller, and H.A. Lewin. 2004. Detection of quantitative trait loci affecting milk production, health, and reproductive traits in Holstein cattle. J. Dairy Sci. 87: 468-475.

Ashwell, M.S. and C.P. Van Tassell. 1999. The Cooperative Dairy DNA Repository-a new resource for quantitative trait loci detection and verification. J. Dairy Sci. 82 (Suppl. 1):54.

Ashwell, M.S., C.E. Rexroad Jr., R.H. Miller, P.M. VanRaden, and Y. Da. 1997. Detection of loci affecting milk production and health traits in an elite US Holstein population using microsatellite markers. Anim. Genet. 28: 216-222.

Ashwell, M.S., C.E. Rexroad, R.H. Miller, and P.M. VanRaden. 1996. Mapping economic trait loci for somatic cell score in Holstein cattle using microsatellite markers and selective genotyping. Anim. Genet. 27: 235-242.

Bascom, S.S. and A.J. Young. 1998. A summary of the reasons why farmers cull cows. J. Dairy Sci. 81: 2299-2305.

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

Casida, L.E. 1961. Present status of the repeat-breeder cow problem. J. Dairy Sci. 44: 2323-2329.

Collins, K. 2000. U.S. Dairy Policy Issues. National Dairy Leadership Conference, U.S. Department of Agriculture. Williamsburg, Va.

Da, Y., M. Ron, A. Yanai, M. Band, R.E. Everts, D.W. Heyen, J.I. Weller, G.R. Wiggans, and H.A. Lewis. 1994. The Dairy Bull Repository: A resource for mapping quantitative trait loci. Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, ON, Canada 21: 229-232.

Darnsfield, M.B., R.L. Nebel, R.E. Pearson, and L.D. Warnick. 1998. Timing of insemination for dairy cows identified in estrus by a radio telemetric estrus detection system. J. Dairy Sci. 81: 1874-1882.

Huszenicza, G., S. Jánosi, M. Kulcsár, P. Kóródi, J. Reiczigel, L. Kátai, A.R. Peters and F.D. Rensis. Effects of clinical mastitis on ovarian function in post-partum dairy cows. Reprod. Dom. Anim. 40: 199-204.

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

Lucy, M.C. 2001. ADSA Foundation Scholar Award: Reproductive loss in high-producing dairy cattle: where will it end? J. Dairy Sci. 84: 1277-1293.

Macmillan, K.L., I.J. Lean, and C.T. Westwood. 1996. The effects of lactation on the fertility of dairy cows. Aust. Vet. J. 73: 141-147.

Miller, J. 2004. Subject: Dairy economics. http://www.ers.usda.gov/Briefing/Dairy/. Accessed June 15, 2005.

Meuwissen, T.H.E. and M.E. Goddard. 2004. Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data. Genet. Sel. Evol. 36: 261-279.

Olsen, H.G., L. Gomez-Raya, D.I. Våge, I. Olsaker, H. Klungland, M. Svendsen, T. Ådnøy, A. Sadry, G. Klemetsdal, N. Schulman, W. Krämer, G. Thaller, K. Rønningen, and S. Lien. 2002. A genome scan for quantitiative trait loci affecting milk production in Norwegian dairy cattle. J. Dairy Sci. 85: 3124-3130.

Roche, J.F., D. Mackey, and M.D. Diskin. 2000. Reproductive management of postpartum cows. Anim. Reprod. Sci. 60-61: 703-712.

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

Schmitt, E.J., T. Diaz, M. Drost, and W.W. Thatcher. 1996. Use of gonadotropin-Releasing hormone agonist or human chorionic gonadotropin for timed insemination in cattle. J. Anim. Sci. 74: 1084-1091.

Schrooten, C., H. Bovenhuis, W. Coppieters, and J.A.M. van Arendonk. 2000. Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle. J. Dairy Sci. 83: 795-806.

Schrooten, C., M.C.A.M. Bink, and H. Bovenhuis. 2004. Whole genome scan to detect chromosomal regions affecting multiple traits in dairy cattle. J. Dairy Sci. 87:3550-3560.

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

Seaton G., C.S. Haley, S.A. Knott, M. Kearsey, P.M. Visscher. 2002. QTL Express: mapping quantitative trait loci in simple and complex pedigrees. Bioinformatics 18: 339-340.

Shanks, R.D., A.E. Freeman, and F.N. Dickinson. 1981. Postpartum distribution of costs and disorders of health. J. Dairy Sci. 64:683.

Sheldon, I.M. and H. Dobson. 2004. Postpartum uterine health in cattle. Anim. Reprod. Sci. 82-83: 295-306.

Thallman, R.M. 2002. User's Manual for GenoProb Version 2.0: Computation of Genotype and phase probabilities in complex pedigrees by iterative allelic peeling. USDA-MARC, Clay Center, NE.

Thallman, R.M., G.L. Bennett, J.W. Keele, and S.M. Kappes. 2001. Efficient computation of genotype probabilities for loci with many alleles: I. Allelic peeling. J. Anim. Sci. 79: 26-33.

VanRaden, P.M., A.H. Sanders, M.E. Tooker, R.H. Miller, H.D. Norman, M.T. Kuhn, and G.R. Wiggans. 2004. Development of a National Genetic Evaluation of cow fertility. J. Dairy Sci. 87: 2285-2292.

Zhang, Q., D. Boichard, I. Hoeschele, C. Ernst, A. Eggen, B. Murkve, M. Pfister-Genskow, L.A. Witte, F.E. Grignola, P. Uimari, G. Thaller and M.D. Bishop. 1998. Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. Genetics 149: 1959-1973.

# APPENDICES

**Appendix 1**. Primers sequence, label, map position, annealing temperature for each marker[a].

| Marker Name | Primer Sequence[b] | Fluorescent Label | Relative Position | Dilution[c] | Ethanol Precipitation | Annealing Temperature |
|---|---|---|---|---|---|---|
| BMS1355 | AAACCCCAAAAAGAACCC<br>ATTTGCGACATTGGATGAA | NED | 2.9 | 1:80** | No | 57°C |
| BMS2213 | ATGGGCAGCTTAGGGATTG<br>CTTCAAGAGCCTTCAGTGGG | HEX | 24.5 | 1:80** | No | 63°C |
| INRA121 | GGAAACCCATTGGAGGATTTG<br>CTTCACTATTCCCCACAAAGC | 6-FAM | 30.2 | 1:50* | No | 57°C |
| BM8151 | TACCTCTAGTGCCACCTGGG<br>TACAAATTCGGGAGAGTGCC | NED | 40.2 | undiluted* | Yes | 59°C |
| DIK2128 | CACCATGTTCTTTGGCAAGTT<br>AGGCAATGGGCTAAGTGCT | HEX | 40.2 | 1:20*+ | No | 57°C |
| BM7109 | CAGGTAAAAGAGCGGCTTTG<br>CAGCTTCATGCCCTAGAAGG | 6-FAM | 47.0 | 1:50* | No | 57°C |
| BMS2914 | AATATGGTCTGGAAAAGAACACA<br>GATGCTGTGTCCCTGAATG | 6-FAM | 50.0 | 1:25* | No | 51°C |
| ILSTS002 | TCTATACACATGTGCTGTGC<br>CTTAGGGGTGAAGTGACACG | HEX | 54.7 | 1:10* | No | 59°C |
| DIK2779 | AGCGACTAAATAACAACAGCAA<br>TGAACCATCAGCTTCTCCTG | 6-FAM | 56.3 | 1:20*+ | No | 57°C |
| BMS929 | CTCCCAACTGAATGTGCATG<br>GTGAGAGTGACTGGGGCTG | NED | 61.2 | 1:100** | No | 61°C |

**Appendix 1** (Continued)

| Marker Name | Primer Sequence[b] | Fluorescent Label | Relative Position | Dilution[c] | Ethanol Precipitation | Annealing Temperature |
|---|---|---|---|---|---|---|
| BB710 | GATCTGTGGGTTGAGTGTGTG<br>AGATTTGTGCCTGCGTGTC | NED | 62.1 | 1:50** | No | 59°C |
| BM2078 | CCCAAAAGAAGCCAGGAAG<br>TCAGAGTTTGGGGTCCTCAG | 6-FAM | 76.8 | 1:25* | No | 63°C |
| TGLA227 | CGAATTCCAAATCTGTTAATTTGCT<br>ACAGACAGAAACTCAATGAAAGCA | HEX | 84.0 | Undiluted** | No | 64°C |

[a]All primer sequences are listed from 5' to 3'

[b]Forward primer sequence is listed first and reverse primer sequence is listed second

[c]Dilution factor * = 12 μl PCR reaction,** = 5 μl PCR reaction, + = 1.2 μM primer concentration

# Appendix 2
# Genemapper 3.7 Instructions

*Compiled by K Boyette*
*Revised by S Muncie and A O'Nan*

A few "Genemapper" terms to become familiar with before starting:
**Panel Manager**-Where you create and set up all your marker information.
**Bin Set**-The individual marker used. You will want to name this the same thing you named the analysis method.
**Kit**-The overall name for your project (ex. Chromosme 18 Cattle Genotyping Project)
**Panel**-The markers you used for that particular project. Just so you know, you can use more than one marker when genotyping animals as long as they have a different allele size OR dye color. This is called multiplexing.
**Autobin**-when the computer looks at your data and selects where the alleles should be based on the allele information you set up in the Panel Manager.

## STEP ONE-ADDING SAMPLES TO GENEMAPPER

1. Open Genemapper 3.7 software.
2. Click FILE and ADD SAMPLES TO PROJECT.
3. Highlight the appropriate folder. At the bottom of the window, click ADD TO LIST. Your sample folders should appear on the right half of the window.
4. Click ADD (on bottom right side of screen).

## STEP TWO-SETTING UP ANALYSIS PARAMETERS

1. On the toolbar, click TOOLS and PANEL MANAGER.
2. Click on PANEL MANAGER in the left window so that it is highlighted in blue.
3. Go to FILE and NEW KIT. Name the KIT the overall name of your project. Leave the kit type as Microsatellite and click OK.
   **WRITE DOWN KIT NAME_____.**
4. Highlight the **Kit** you just made by clicking on it. It should be highlighted in blue when you click on it.
5. Next, you are going to create a PANEL. A PANEL is what marker(s) you ran on that specific plate. You add new PANELS by going to FILE and NEW PANEL. Note to multiplexers: You will name your panel the name of the two markers you used-EXA. BMS2078_BMS360
   **WRITE DOWN PANEL NAME_____.**
6. In the left window, click on the PANEL you just made. This will highlight it in blue. Go to FILE and NEW MARKER. Add each individual marker you used for the PANEL you made.
   **WRITE DOWN MARKER NAME:_____.**
7. Enter the MARKER information-You need to enter information only for the Marker Name, Dye Color, and Minimum and Maximum allele size. That is all you need to worry about here. *If you are entering more than one marker, repeat step 6.*
8. Go to BINS (on top toolbar).
9. Highlight in the left window the kit you created in step 3. (Remember, the kit pertains to the overall project name).

10. Go to BINS and NEW BIN SET.
11. Name the BIN SET the same thing you named the PANEL in step 5. I know that this is redundant but just go with it.
    **WRITE DOWN BIN SET NAME:_____.**
12. Click APPLY and OK. This should take you back to the main SAMPLES Screen.

**STEP THREE-SETTING UP ANALYSIS METHOD**

1. Locate the ANALYSIS METHOD column and click on the first row. A pull down screen will appear. Select NEW ANALYSIS METHOD.
2. Under SELECT ANALYSIS TYPE click microsatellite and then OK.
3. A window will pop up under the "General" tab. Name the Method something relevant to the plate you are analyzing. We commonly use the Marker Name. *Note to Multiplexers: You would enter both marker names (EXA. BMS2078_BMS360).* That is all you need to change in this window.
   **WRITE DOWN ANALYSIS METHOD NAME:_____.**
4. Go to the "Allele" tab. Click on BINSET and you will see a pull down menu. Select the Bin Set for you sample. This will have the same name as the Analysis Method. That is all you need to do on this screen.

**STEP FOUR-FILLING IN SAMPLE INFORMATION**

1. In the Main SAMPLES screen, click on the column named "Size Standard." A pull down menu should appear. Select GS400HD. Fill down the column by clicking the column header and holding down CTRL-D.
2. Click the "Analysis Method" column. Choose what ever you named the Analysis Method in the Step Three of the Step Three Section.
3. Click the "Panel" column. Click on the first cell and a pull down menu should appear. Click on the KIT folder and double click the PANEL name. Fill down the column by clicking the column header and holding down CTRL-D.

**STEP FIVE-SAMPLE ANALYSIS AND AUTOBINNING**

1. Click on the Green Arrow on the action toolbar.
2. Name the PROJECT (*you can name this the same thing you named the KIT)*
   **WRITE DOWN PROJECT NAME:_____.**
3. Wait for the data to be analyzed.
4. Now you will need to AUTOBIN. Don't fret, you are almost done!!!!! To autobin, go to TOOLS and PANEL MANAGER. Clcik on KIT> PANEL>BINS>ADD REFERENCE DATA. Select a couple of runs to use a reference data. Click ADD (bottom right side of window).
5. Go back to main SAMPLES screen and Click on the Green Arrow on the action toolbar (this too is redundant).
6. After it has reanalyzed your sampled, you can click on the GENOTYPES tab and see your analyzed data.

# Appendix 3
# Instructions for GenoProb

**These instructions should be used in association with the User's Manual for GenoProb Version 2.0**

1. Unzip the three programs files using a decompressing software program
2. After all three files have been unzipped, open folder one and double-click on the setup.exe and follow instructions to install the GenoProb program.
3. Access files might need to be converted to the correct version
    a. Open one of the two access files, GenoProb analysis or GenoProb inputs (they are stored in the directory named \templates)
    b. Go to Tools, Database utilities, convert database, to current Microsoft Access database version, rename the files and safe them to a new folder
    c. Repeat with the other file
4. In order to use GenoProb the user must set up tables similar to the ones in the original input file
    a. Tables must be made in Microsoft Excel and imported into a new access database for storage.
    b. Set up the files exactly how they are seen in the GenoProb Inputs file with your data. The tables needed include: Species, Map, Projects, Pedigree and Marker Data tables.
        i. The project name will be consistent for all tables and can only be 10 characters long
    c. For the **Pedigree table**, a terminal animal of the pedigree the sire and dam can be listed as zero. For the sort column, you can use consecutive numbers to ensure that the parents precede their offspring. Grandparents listed as 1, parents listed as 2, sons listed as 3 and so on.
    d. For the **Marker Data table**, each animal needs to be listed separately for each marker; the status for the markers can be listed as verified (1). All other fields in this table should be self-explanatory.
    e. For the **Map Table**, for the GD_crit, you can use –1 (no records are written to genotype distribution detail) for the entire column
    f. After your tables are made in excel you want to import them into access
        i. To import excel files into access, Go to File, get external data, Import
        ii. Pick the correct excel file, and follow the prompting (usually you want to click the box that makes the first row the header, and also you **do not** want a primary key)
        iii. Repeat for all tables
5. Before you can append tables the field properties between your data and the GenoProb inputs file need to match.
    a. Under the tables view, you need to then select "Design" View. This will allow you to see each field and the associated information
    b. Select each field name and look at the bottom to the field properties. Both Microsoft Access files will need to be open simultaneously. Go through

each field and make sure that your data field properties match the field data properties for the same field in the GenoProb inputs file.
    i. Exception: If you want to use a long registration number for the animal ID you will need to change the length allowed for that character in the **GenoProb inputs** file and then make sure it matches the name length in your file.

6. Now the tables made with your data need to be appended separately to the GenoProb Inputs access database.
    a. In the Microsoft Access file you created with your data, go to the queries tab. (It will be on the left side of your screen)
    b. Double click to go create queries in design view.
    c. On the "show table" window select the table you want to append first. Then drag the asterisk (*) at the top of the table onto the field section of the query window. Then close the "show table" window.
    d. Go to Query on the tool bar menu and to append query.
    e. Select "another database"
    f. In the space provided, write the file path name exactly as it appears in the properties for the window. (If the path name does not match exactly then it will not work!)
    g. After you have the file name exactly correctly the pull down menu for the table name will now be filled with the list of tables in the GenoProb Inputs access file.
    h. Select the correct table you want to append with your information.
    i. Then click OK.
    j. You must then click the red exclamation point to run your query.
    k. Nothing will happen in the database file you have open. However, you can go check the GenoProb inputs file and your data should now be in the files already in place.
    l. Repeat these steps for all the tables you created with your data.
7. Now you should be ready to actually use the GenoProb program. Double click the icon to get to the opening screen.
    a. You will first need to "set database paths". Click on the icon and it will open to two different entry spots: 1. Analysis results path, 2. Data source path
        i. Make sure you type the exact file name from the new inputs file you just appended all your data to and the analysis file that was converted to the current access version. (If the file name is not **EXACTLY** correct this will not work!)
    b. Next, you want to click on "Add new job". Select your information from all the pull down windows. Local Data will be used. Name your project something for easy reference.
    c. You want to unclick the complete penetrance box and also pick a null prior of 0.0001.
    d. Select the job you just created from the current job pull down menu.
    e. Then click run current job. The entire job should now run. If not, then you did something wrong.

8. After the job has been completely run, the output data can be viewed by opening the Output Summary in the GenoProb Analysis file. This gives you a summary of the predicted genotypes and the probability the genotype is correct.
    a. Gmax0 and Gmax1 are the predicted genotypes for the animal
    b. pGmax is the probability that the genotype is correct. A common cut off value is < 0.98.
    c. oGmax is the likelihood that the Gmax0 came from the dam and the Gmax1 came from the sire, given that they are the true genotypes.
9. For easier viewing of the results from the GenoProb analysis, it may be helpful to concatenate the animal ID with the marker name. This will help to sort the samples easier, since each animal is duplicated for each marker
    a. Open design view for a query
    b. In the field row type [title of the marker name column] &“ ”& [title of animal ID column]
    c. Then drag which columns you want to be displayed in your query in the following columns

**Trouble shooting:**

1. The animal ID name needs to match **EXACTLY** in the marker data table and the pedigree table. If there are extra spaces, the program will not recognize your data and will dump the run.
2. Ensure there are no duplicates in data. Any duplicates that are there will dump the analysis. This is **VERY** important. Double check this.
3. If an animal is in the Marker Data table, it must be in the pedigree. However, an animal can be in the pedigree that is not in the marker data table.
4. You cannot specify an animal as 0/0. This will imply to the program that it is a null allele. Therefore, 0/0 can not specify and unknown genotype. If the genotype is unknown, leave the animal out of the marker data table.
5. You do not need to make an error rates table. Using the error rates table provided in the GenoProb inputs file will be sufficient.

# Appendix 4
# Project Materials

**Where:** 4 degrees (Refrigerator)
**What:**

1. Extracted DNA from Ned Project (top shelf left-hand side):
   Styrofoam boxes labeled Sarah Ned project/Box number and in blue tape
   Please see database file for itemized list of where to find each DNA sample
2. Sarah's Ned project Diluted DNA Plates (top shelf right-hand side):
   Each Sample Diluted to 20 ng/ul using Tris HCl pH 8
   Please see database file for itemized list of the animal location in each plate
3. Sarah's Diluted Primers (top shelf right-hand side):
   In a box labeled Sarah Chromosome 18 diluted primers
   Each primer is diluted to 20 mM and labeled with the primer name
   Primers include:
   BM7109, INRA121, BMS929, BM8151, TGLA227, BB710, ILSTS002, BMS1355, BMS2213, DIK2128, DIK2779, BMS2914, BM2078

**Where:** -20 degrees (Freezer on the right side closest to the entrance)
**What:**

Ned frozen semen (The second to bottom shelf):
   Labeled Sarah Ned Project/box number and in Yellow tape
   Please see database file for itemized list for where to find each Semen Sample