Distribution and location of genetic effects for dairy traits

J. B. Cole,*[1] P. M. VanRaden,* J. R. O'Connell,† C. P. Van Tassell,*‡ T. S. Sonstegard,‡ R. D. Schnabel,§
J. F. Taylor,§ and G. R. Wiggans*
*Animal Improvement Programs Laboratory, ARS, USDA, Beltsville, MD 20705-2350
†University of Maryland School of Medicine, Baltimore 21201
‡Bovine Functional Genomics Laboratory, ARS, USDA, Beltsville, MD 20705-2350
§Division of Animal Sciences, University of Missouri, Columbia 65201

ABSTRACT

Genetic effects for many dairy traits and for total economic merit are evenly distributed across all chromosomes. A high-density scan using 38,416 single nucleotide polymorphism markers for 5,285 bulls confirmed 2 previously known major genes on *Bos taurus* autosomes (BTA) 6 and 14 but revealed few other large effects. Markers on BTA18 had the largest effects on calving ease, several conformation traits, longevity, and total merit. Prediction accuracy was highest using a heavy-tailed prior assuming that each marker had an effect on each trait, rather than assuming a normal distribution of effects as in a linear model, or that only some loci have nonzero effects. A prior model combining heavy tails with finite alleles produced results that were intermediate compared with the individual models. Differences between models were small (1 to 2%) for traits with no major genes and larger for heavy tails with traits having known quantitative trait loci (QTL; 6 to 8%). Analysis of bull recessive codes suggested that marker effects from genomic selection may be used to identify regions of chromosomes to search in detail for candidate genes, but individual single nucleotide polymorphisms were not tracking causative mutations with the exception of diacylglycerol O-acyltransferase 1. Additive genetic merits were constructed for each chromosome, and the distribution of BTA14-specific estimated breeding value (EBV) showed that selection primarily for milk yield has not changed the distribution of EBV for fat percentage even in the presence of a known QTL. Such chromosomal EBV also may be useful for identifying complementary mates in breeding programs. The QTL affecting dystocia, conformation, and economic merit on BTA18 appear to be related to calf size or birth weight and may be the result of longer gestation lengths. Results validate quantitative genetic assumptions that most traits are due to the contributions of a large number of genes of small additive effect, rather than support the finite locus model.

Key words: calving trait, genomic selection, single nucleotide polymorphism, quantitative trait loci

INTRODUCTION

The high-speed genotyping of large numbers of SNP has recently become affordable for dairy cattle, which has allowed development of genomic selection programs (Harris et al., 2008; VanRaden et al., 2009). In addition to increasing rates of genetic improvement and reducing costs of progeny testing (Meuwissen et al., 2001; Schaeffer, 2006), genomic evaluations produce estimates of the contributions of individual markers to additive genetic merit. Although previous studies using microsatellites involved at most a few hundred markers (Schrooten et al., 2000; Kühn et al., 2003; Ashwell et al., 2004; Schnabel et al., 2005; Holmberg and Andersson-Eklund, 2006), SNP-based studies may involve tens of thousands of markers. These data must be evaluated statistically, rather than by individual inspection of markers.

Genetic effects must exist somewhere on the chromosomes for any trait with a nonzero heritability. Marker locations and estimates of effects can be used to assess alternative hypotheses and prior assumptions about the distribution of genetic effects, and to identify chromosomal segments of interest for functional genomic study. Markers with large effects on traits of economic importance may be used to identify regions of the genome that merit further study to identify causative mutations.

Several studies have reported on QTL affecting conformation, calving traits, fertility, and longevity on *Bos taurus* autosome (**BTA**) 18. A QTL affecting longevity, maternal dystocia, and maternal stillbirth in the German Holstein population was described by Kühn et al. (2003), Holmberg and Andersson-Eklund (2006) identified a QTL affecting maternal calving ease in a resource population of Swedish Reds and Swedish Holsteins, and Thomasen et al. (2008) located a QTL with

pleiotropic effects on dystocia, stillbirth, and calf size in a population of Danish Holsteins. A QTL for conformation in US Holsteins was reported by Schnabel et al. (2005), and a similar marker was identified in Canadian Holstein bulls (Kolbehdari et al., 2008). These results suggest that there may be a QTL on BTA18 affecting conformation and calving traits, although no previous study has demonstrated that a single QTL is responsible for these observations or proposed a biological mode of action for the marker effect.

The objectives of this paper were to 1) describe the effect of different priors on the reliability of genomic predictions; 2) demonstrate the value of markers with large effects to identify regions of the genome that merit further study; 3) report the location and size of effects of a marker on BTA18 that is associated with dystocia, conformation, net merit, and productive life; and 4) propose a model of inheritance for the dystocia locus that is consistent with the marker data.

MATERIALS AND METHODS

Data

Genomic data, phenotypic data, and edits were as reported in VanRaden et al. (2009). The Holstein bulls genotyped were categorized as either predictor bulls (3,576 bulls born from 1952 through 1998) or predicted bulls (1,759 bulls born from 1999 through 2002). Information from the predictor bulls was used to compute predictions, which were tested using the predicted bulls. Predictor bulls were required to have a reliability of at least 75% for net merit in August 2003, and predicted bulls were required to have information from ≥10 daughters in their evaluations by April 2008.

The main source of extracted DNA was from semen held in the Cooperative Dairy DNA Repository (Ashwell and Van Tassell, 1999). Semen from significant ancestor bulls was purchased independently or was provided by the National Center for Genetic Resources Preservation, ARS, USDA (Fort Collins, CO).

Genotypes for 38,416 SNP scored in 5,360 Holstein bulls were examined. The selected SNP were from the Illumina Bovine SNP50 BeadChip (Illumina Inc., San Diego, CA; Matukumalli et al., accepted) and had minor allele frequencies greater than 5% in Holsteins. Extraction of DNA and genotyping were performed at 6 locations: Bovine Functional Genomics Laboratory (Beltsville, MD); GeneSeek (Lincoln, NE); Genetics & IVF Institute (Fairfax, VA); Illumina (San Diego, CA); Department of Agricultural, Food and Nutritional Science, University of Alberta (Edmonton, Alberta, Canada); and Division of Animal Sciences, University of Missouri (Columbia, MO). Scoring of marker

genotypes was performed using Illumina's Beadstudio software (v3.2.23).

Allele frequencies in the base (founder) population were estimated using the algorithm of Gengler et al. (2007) that solves for gene content of nongenotyped ancestors and descendants using pedigrees. The pedigree file with all known ancestors of the 5,335 bulls included 41,414 cows and bulls. The genotype file included 205 million known and 2.0 million (1%) unknown genotypes. Allele counts from relatives were used to impute 974,961 (49%) of the 2 million missing genotypes.

The data included multitrait across-country evaluations from the Interbull Centre (Uppsala, Sweden) for bulls that had been progeny-tested in Canada and were expressed on the US scale. Dependent variables for analysis were daughter deviations for 27 traits related to dairy production: milk, fat, and protein yields; fat and protein percentages; productive life; SCS; daughter pregnancy rate; sire (**SCE**) and daughter (**DCE**) calving ease; final score; stature; strength; body depth; dairy form; foot angle; rear legs (side and rear views); rump angle and width; fore udder; rear udder height; udder depth and cleft; front teat placement; teat length; and net merit. Daughter deviations were weighted by reliability from daughters, which was computed from total daughter equivalents minus daughter equivalents from parent average (**PA**).

Genomic Predictions

Predictions were computed using linear and nonlinear genomic models (VanRaden, 2007, 2008). For linear predictions, the traditional additive genetic relationship matrix is replaced by a genomic relationship matrix and is equivalent to assigning equal genetic variance to all markers. Three different nonlinear models were considered: an infinitesimal alleles model with a heavy-tailed prior in which smaller effects are regressed further toward 0 and markers with larger effects are regressed less to account for a nonnormal prior distribution of marker effects (model A), a finite locus model with a normal distribution of marker effects (model B), and a finite locus model with heavy tails (model AB). Infinitesimal allele models assume that all loci have nonzero effects, and finite loci models assume that only a fixed number of alleles have effects. Models A and AB are analogous to the BayesA and BayesB methods of Meuwissen et al. (2001), respectively.

The linear model assumed a normal prior distribution, and model B assumed a normal prior for a proportion of the effects and no effect for the others. The prior used for models A and AB was a heavy-tailed distribution generated by dividing a normal variable by

$h^{|s-2|}$, where h determines departure from normality and s is the size of the estimated marker effect in standard deviations (VanRaden, 2008). Marker effects are normally distributed with no additional weight in the tails when h is 1, and variance in the tails grows with increasing values of h. In addition, the proportion of markers allowed to have effects substantially different from 0 was needed for models B and AB. Optimal values of these parameters for each model were determined by evaluating 3 values of h (1.08, 1.12, and 1.25) and 4 values of marker proportion (6.25, 12.5, 25, and 50%) and comparing the squared correlations of daughter deviations with genomic predictions. A value of 1.25 was used in simulation (VanRaden, 2008), but the optimal value of h using real data was 1.12 for model A and 1.08 for models B and AB. The best value for the proportion of markers was 25% for model B but was 50% for model AB. Model A with a parameter of 1.12 is used for the routine genomic evaluations.

Genomic predictions and PA calculated from August 2003 data of older animals were compared for ability to predict April 2008 evaluations for younger bulls for each of the 27 traits analyzed. Predictive ability was determined by calculating squared correlations between daughter deviations and PA or genomic predictions for the linear and nonlinear models. The experimental design provided an independent, realistic test by separating early daughter information of ancestors used to compute predictions from later daughter information of descendants used to assess prediction accuracy.

The location of causative mutations for simple recessive conditions also can be mapped using the genomic prediction machinery. Recessive codes were used to define phenotypes for red coat color (**RED**; Joerg et al., 1996), complex vertebral malformation (**CVM**; Agerholm et al., 2001), and bovine leukocyte adhesion deficiency (**BLAD**; Shuster et al., 1992). Numbers of genotyped carriers were 141 for CVM and 41 for BLAD, which were binary traits because no homozygous recessive animals were genotyped. All 3 genotypes were present for RED including 41 carriers and 50 homozygous recessives. Heritabilities for all 3 traits were set to 0.99 because reported test results were assumed to contain little error.

Because 2003 PA had not been stored for type traits or for calving ease, 2003 pedigree indices (**PI**) constructed as 0.5 (sire PTA) + 0.25 (maternal grandsire PTA) + 0.25 (birth year mean PTA) were substituted for PA for those traits. Reliability of PI is lower than that of PA, especially for highly heritable traits, because records for the dam are excluded. The 2008 PA was not substituted for the 2003 PA because then the son's information would have added to his dam's reliability.

Final genomic predictions combined 3 terms by selection index: 1) direct genomic prediction, 2) PA computed from the subset of genotyped ancestors using traditional relationships, and 3) published PA or PI. For each animal, a matrix (3×3) is set up with reliabilities for the 3 terms on the diagonals and functions of those 3 reliabilities on the off-diagonals (VanRaden et al., 2009). Reliabilities of direct genomic predictions are calculated by inversion of the mixed model equations for the linear model, and the theoretical reliability from the linear model is also used for the nonlinear models. Regressions and correlations were used to test predictions. A bull's published PTA is a weighted mean of his daughter deviation and his PA, and the use of deregressed evaluations or daughter deviations as dependent variables helps to avoid part-whole correlations with PA. Because daughter deviations as defined by VanRaden and Wiggans (1991) were not available for all traits, daughter deviations were computed as deregressed evaluations. The regression coefficient was calculated from daughter equivalents from progeny, which was obtained by subtracting daughter equivalents from parents from the bull's total daughter equivalents.

Expected and Realized Genetic Effects

Genetic effects were computed for each chromosome as sums of individual allele effects (**SA**), additive genetic variances calculated using base population allele frequencies (**VB**), additive genetic variances calculated using current population allele frequencies (**VC**), and EBV on a within-bull basis (**VWB**). Expected contributions were estimated by multiplying sums or variances by the proportional length of individual chromosomes, with the proportions summing to 1. Ratios of individual allele effects were calculated as:

$$SA_i = \frac{\sum_{c=1}^{n_c} |\alpha_c|}{\sum_{g=1}^{n_g} |\alpha_g| \times \frac{length_i}{length_g}},$$

where SA_i = the ratio of the observed to the expected sums of allele substitution effects for the ith chromosome; n_c = the number of markers on chromosome i; $|\alpha_c|$ = the absolute value of the effect of the cth allele on the ith chromosome; n_g = the number of markers in the genome; α_g is the effect of the gth allele in the genome; and $length_i$ and $length_g$ = the lengths of the ith chromosome and the genome, respectively, in base pairs. The allele effects are equivalent to average allele substitution effects (Falconer and MacKay, 1996).

The additive genetic variance for a trait may be calculated as a function of allele frequencies and squared allele substitution effects as $V_a = \sum_{i=1}^{n} 2p_i\left(1 - p_i\right)\alpha_i^2$, where V_a = the additive genetic variance; n = the number of alleles affecting the trait; p_i = the frequency of the ith alelle affecting the trait; and α_i = the substitution effect of the ith allele (Falconer and MacKay, 1996). Ratios of additive genetic variances were calculated using allele frequencies from either the base (VB) or current (VC) populations:

$$VB_i = \frac{\sum_{c=1}^{n_c} 2p_c\left(1 - p_c\right)\alpha_c^2}{\sum_{g=1}^{n_g} 2p_g\left(1 - p_g\right)\alpha_g^2 \times \frac{length_i}{length_g}},$$

where VB_i = the ratio of the observed to the expected sums of additive genetic variances for the ith chromosome; n_c = the number of alleles on chromosome i; p_c = the frequency of the cth allele on the ith chromosome; α_c = the substitution effect of the cth allele on the ith chromosome; n_g = the number of alleles in the genome; p_g = the frequency of the gth allele in the genome; α_g = the substitution effect of the gth allele in the genome; and $length_i$ and $length_g$ = the lengths of the ith chromosome and the genome, respectively, in base pairs. The values of VC are calculated in the same manner, although the gene frequencies p_c and p_g are replaced with estimates from the current population rather than the base population (VanRaden et al., 2009).

The ratios of additive genetic variances were also calculated on VWB to account for possible linkage disequilibrium. This was done by processing the chromosomal breeding values for each animal receiving an evaluation:

$$VWB_i = \frac{\frac{1}{n_a}\sum_{a=1}^{n_a}\sum_{c=1}^{n_c} 2p_c\left(1 - p_c\right)\alpha_c^2}{\frac{1}{n_a}\sum_{a=1}^{n_a}\sum_{g=1}^{n_g} 2p_g\left(1 - p_g\right)\alpha_g^2 \times \frac{length_i}{length_g}},$$

where VWB_i = the ratio of the observed to the expected sums of allele substitution effects for the ith chromosome; n_a = the number of animals with genomic EBV (**GEBV**); n_c = the number of alleles on chromosome i; p_c = the base population frequency of the cth allele on the ith chromosome; α_c = the effect of the cth allele on the ith chromosome; n_g = the number of alleles in the genome; α_g = the effect of the gth allele in the genome; and $length_i$ and $length_g$ are the lengths of the ith chromosome and the genome, respectively, in base pairs. All of these models are assumed to be purely additive; that is, dominance and epistatic variances are 0.

RESULTS AND DISCUSSION

Genomic Predictions

All markers are assigned equal weights in the linear model. In the heavy-tails model small marker effects are regressed more toward zero than in the linear model, whereas markers with large effects are regressed less. Regressing large marker effects less than small marker effects corresponds to a model under which there is a mixture of many genes with small effects and a few genes with large effects. This model seems appropriate for traits that are known to be affected by a major gene, such as fat percentage, but for which there is a lot of variation not accounted for by the known QTL. Model B is a finite allele model under which only some markers are assumed to have nonzero effects, and the distribution of those effects is normal. Model AB combines heavy tails with finite alleles, which is similar to the BayesB method of Meuwissen et al. (2001), under which most loci have a genetic variance of zero, and the nonzero genetic variances are sampled from an inverse χ^2 distribution. However, models B and AB differ from the BayesB model in that small marker effects are regressed toward, rather than are fixed at, zero.

Model A had little advantage in squared correlation over the linear model except for fat and protein percentages, with increases of 8 and 6%, respectively (Table 1). Gains obtained in simulation averaged 3% but were mostly smaller with real data, indicating that most traits are influenced by more loci than the 100 QTL used in the simulation (VanRaden, 2008). Model B provided similar or poorer fits than did model A for all traits, including fat and protein percentage, for which there are known genes of large effect (Grisart et al., 2004; Cohen-Zinder et al., 2005). Model AB produced fits that were similar to, but slightly poorer than, those for models A and B. The assumption that relatively few loci have nonzero genetic variance is not consistent with these findings, and we did not observe the higher accuracies predicted by Meuwissen et al. (2001) with the BayesB method. Models that assumed that all markers have some effect rather than that most have no effect provided better squared correlations. Slight decreases in squared correlations were noted for most traits in model A when the amount of variance in the tails beyond 2 standard deviations was increased (data not shown).

Table 1. Squared correlations (×100) of daughter deviations with parent average (PA) or with genomic predictions using different numbers of QTL and prior distributions of QTL sizes

		Model[1]			
Trait	PA	Linear	A	B	AB
Net merit	10.8	28.2	28.4	27.8	27.6
Milk	28.2	47.2	48.5	47.7	47.3
Fat	15.0	41.8	44.2	43.7	43.6
Protein	27.3	47.5	47.0	46.7	46.6
Fat percentage	25.1	55.3	63.3	63.1	63.9
Protein percentage	27.6	51.4	57.7	56.7	56.6
Productive life	16.6	25.6	27.4	26.5	26.4
SCS	22.6	37.3	38.3	37.7	37.6
Days open	19.6	29.5	29.0	28.9	29.2

[1]Linear = linear model; A = heavy tails model with a prior of 1.12; B = finite alleles model with 10,000 markers; AB = finite alleles model with 20,000 markers and a heavy tails prior of 1.08.

Largest Effects

The largest marker effects were for fat percentage on BTA14 flanking the diacylglycerol O-acyltransferase 1 (**DGAT1**) gene (Grisart et al., 2004), with lesser effects on both milk and fat yield. Large marker effects for protein percentage were also present on BTA6 flanking the ATP-binding cassette, subfamily G, member 2 gene (Cohen-Zinder et al., 2005). This demonstrates that the genomic predictions work by tracking the inheritance of known causal mutations of large effect. Markers on BTA18 centered on SNP ss86324977 had the largest effects for several traits: SCE, DCE, rump width, stature, strength, and body depth (Table 2; Figure 1). The SNP ss105250374, which is approximately 513 kb downstream of ss86324977, had the largest effect on productive life. Another marker on BTA18 (rs41628662) was associated with the largest effect detected for net merit, in the region previously identified by Ashwell et al. (2004) as having a large effect on daughter pregnancy rate. This marker had a greater effect on economic merit than did *DGAT1*.

Several other SNP associated with large effects were identified, including putative QTL for fore udder attachment and udder depth (94.8 Mbp), fat yield (98.7 Mbp), and stature (112.3 Mbp) on BTA5; foot angle (62.6 Mbp) on BTA16; protein percentage (34.0 Mbp) on BTA20; and stature (51.5 Mbp) on BTA29. The SNP at 94.8 Mbp associated with fore udder attachment and udder depth is bracketed by microsatellites associated with udder depth at 90.8 and 109 Mbp reported by Schrooten et al. (2000) and Ashwell et al. (2004), respectively. The stature SNP at 112.3 Mbp on BTA5 is located slightly downstream of a microsatellite associated with body size at 123 Mbp (Schrooten et al., 2000). Kolbehdari et al. (2008) reported a SNP (rs41634810) at 65.9 Mbp on BTA16 associated with foot angle, which is within 40 kbp of the SNP (ss86295756) at 62.2 Mbp reported in this study. Khatkar et al. (2004) reported the presence of a QTL at 38.5 cM on BTA20 based on a meta-analysis of data from several earlier studies. No reports of QTL associated with fat yield on BTA5 in the region of 98.7 Mbp or stature anywhere on BTA29 were found in the literature.

Table 2. Single nucleotide polymorphisms on *Bos taurus* autosome 18 with the largest effects (in additive genetic SD) on net merit (NM), longevity (PL), sire (SCE) and daughter calving ease (DCE), sire (SSB) and daughter (DSB) stillbirth, stature (Stat), strength (Str), body depth (BDep), and rump width (RWid)

SNP[1]	Position (bp)	Trait	Effect (SD)	Hetero[2]	Functional class	Gene name
rs41628662	41,453,097	NM	0.07	0.19	Unknown	Unknown
ss86324977	57,125,868	BDep	0.32	0.13	Intron	Sialic acid binding Ig-like lectin 5
		DCE	0.44			
		DSB	0.02			
		RWid	0.32			
		SCE	0.53			
		SSB	0.12			
		Stat	0.25			
		Str	0.32			
ss105250374	57,639,664	PL	0.14	0.18	Intron	Zinc finger protein 577

[1]National Center for Biotechnology Information accession number (when known).
[2]Heterozygosity.
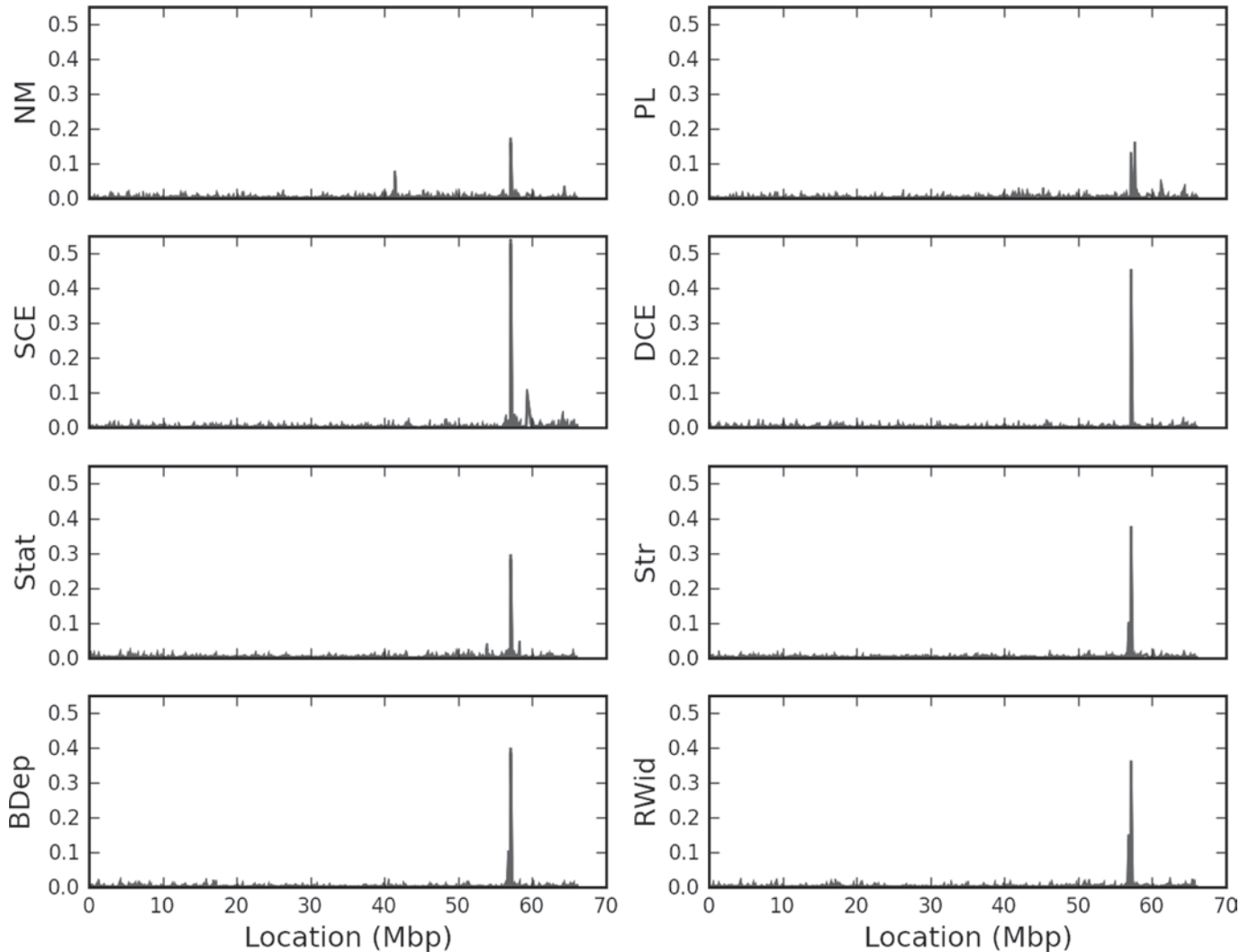
Figure 1. Size (in additive genetic SD) and location of marker effects on *Bos taurus* autosome 18 affecting net merit (NM), longevity (PL), sire (SCE) and daughter calving ease (DCE), stature (Stat), strength (Str), body depth (BDep), and rump width (RWid).

Information about individual SNP with large associated effects was obtained from the National Center for Biotechnology Information's Single Nucleotide Polymorphism Database, and all results are based on version 4.0 of the *Bos taurus* assembly (Table 2). The SNP ss86324977 is located at 57,125,868 bp on BTA18 and has the largest effects of all tested SNP on SCE and DCE, strength, stature, body depth, and rump width. A BLAST analysis of the 5′ flanking sequence indicates that the SNP is located in an intron of the sialic acid binding Ig-like lectin (**Siglec**)-5 gene. A related molecule, Siglec-6, is expressed in the human placenta and may be involved in the regulation of signals related to the initiation of parturition (Brinkman-Van der Linden et al., 2007). It has been found that Siglec-5 and 6 both bind to leptin, with the latter exhibiting tighter binding

(Patel et al., 1999). Human Siglec-9 also is expressed at high levels in many tissues, including the placenta and fetal liver (Foussias et al., 2000). Siglec-6 sequesters leptin (Brinkman-Van der Linden et al., 2007), which is secreted by adipocytes and circulates in the bloodstream in a protein-bound form (Houseknecht et al., 1998), and leptin-deficient mice show a delayed onset of parturition (Mounzih et al., 1998). High levels of bound leptin also signal that bodily energy reserves are high, resulting in decreased feed intake and increased rates of weight loss. Animals homozygous for the rare SNP may express high levels of Siglec-5, resulting in a leptin deficiency leading to delayed parturition. The association of ss86324977 with SCE, DCE, strength, stature, body depth, and rump width may be due to increased calf size associated with longer gestation lengths.

Hansen et al. (2004) estimated genetic correlations among direct and maternal gestation length, calving difficulty, and calf size, which showed that calves with longer gestation lengths are larger at birth and have higher rates of dystocia. Similarly, cows that carry their calves longer more frequently bear larger calves with a greater risk of calving difficulty. Genes within the Siglec family show a high rate of evolution within species; therefore, human and bovine physiology may differ. Significant additional work is needed to characterize the bovine Siglecs and their effects on metabolism.

The SNP ss105250374 is located within an intron of the zinc finger protein 577 gene at 57,539,664 bp on BTA18 and has the largest association with productive life of all tested SNP. Zinc finger protein 577 is a member of the superfamily of Krüppel-associated box proteins, which are well-known transcriptional repressors (Bellefroid et al., 1991; Margolin et al., 1994). Ganss and Jheon (2004) also reported in a review of literature that several zinc finger transcription factors play important roles in mammalian bone and skeletal development. Additional research is needed to better understand how Krüppel-associated box proteins are involved in metabolic processes that affect productive life.

Location of Causative Mutations

The availability of low-cost, high-density genotyping arrays may allow faster identification of causal mutations associated with harmful recessives by identifying SNP in strong linkage disequilibrium with the recessive locus. The SNP with the strongest association may not necessarily be located in the underlying gene but should identify a region of a specific chromosome that can more closely be investigated. This approach was tested using recessive codes for 3 well-known phenotypes in Holsteins, BLAD, CVM, and RED, for which the causative mutations are known. Statistical models for binomial responses (Tempelman, 1998) were not used because our goal was to determine the suitability of the existing genomic prediction system for this task.

The locations of the SNP with the strongest associations with each trait are presented in Table 3, as are the names and locations of the genes in which the causal mutations are known to lie. The National Center for Biotechnology Information's Single Nucleotide Polymorphism Database was used to determine if the most strongly associated SNP were located in known genes. None of these SNP were located in the protein in which the causal mutations lie, and the SNP associated with CVM and RED were not located in known genes. The BLAD-associated SNP is located in the pituitary tumor-transforming 1 interacting protein gene, and the RED-associated SNP lies in an intron of vacuolar sorting protein 35. There are 2, 42, and 17 SNP closer to the causal mutations within the genes involved in BLAD, CVM, and RED, respectively, than the SNP most strongly associated with the respective phenotype. The SNP with the strongest associations with RED and BLAD lie within about 100 kbp of the causal mutation, which is expected given the resolution of the assay. However, the SNP strongly associated with CVM is about 1 Mb away from the causal mutation. Given that very little linkage disequilibrium (**LD**) exists in cattle at the 1 Mb range (McKay et al., 2007; Sargolzaei et al., 2008), and that this association was detected despite the 42 putatively closer SNP, it is likely that the SNP are not correctly ordered in this region of the genome.

Most of the markers closer to the causal mutation than the SNP with the largest effect were discovered by the Bovine Functional Genomics Laboratory during the SNP discovery project for the Illumina BovineSNP50 chip (Van Tassell et al., 2008), which were selected to have relatively high minor allele frequencies, and the SNP receiving large solutions may be those with allele frequencies more like the recessives, rather than those closest to the causal mutations. Such a scenario is consistent with the findings of Sargolzaei et al. (2008), who demonstrated how high levels of LD are associated with similar allele frequencies. A few families provide most of the phenotypic information because of the low frequency of carriers in the population, so the SNP solutions also may represent the presence of LD blocks in the carrier families.

The distribution of marker effects (Figure 2) shows that most SNP have small effects (contribute random noise), whereas markers in regions in which causative mutations lie have much larger effects. Although the markers with the largest associated effects from a genomic analysis may not track perfectly the causative mutations, they are potentially useful tools for identifying chromosomal regions associated with recessive Mendelian phenotypes. Animals exhibiting a previously unknown recessive condition, such as brachyspina syndrome (Agerholm et al., 2006), could be genotyped and the data processed using the existing genomic evaluation system. The resulting marker effects would identify specific chromosomal regions in which the causative mutation may be located.

Chromosomal Breeding Values

Chromosomal EBV (**CEBV**) were calculated as the sums of individual marker effects on each chromosome for each bull, and the CEBV were summed to obtain the GEBV. The distributions of CEBV for fat percent-

Table 3. Location of known causative mutations responsible for simple recessive conditions in Holstein bulls and QTL of known large effect identified by SNP associations

Gene[1]	BTA[2]	SNP Name	SNP Location	Gene Name	Gene Location	NCBI[3] SNP Assignment
ABCG2	6	ss61523660	37,454,410	ATP-binding cassette, subfamily G, member 2	37,351,168	None[4]
BLAD	1	ss46527085	146,829,619	Integrin, β 2	146,755,368	Pituitary tumor-transforming 1 interacting protein
CVM	3	ss86341096	45,261,894	UDP-N-acetylglucosamine transporter	46,235,536	None[4]
DGAT1	14	ss86329615	443,936	Diacylglycerol O-acyltransferase 1	444,097	Diacylglycerol O-acyltransferase 1
RED	18	rs41579939	14,058,411	Melanocortin 1 receptor	13,776,888	Vacuolar protein sorting 35 homolog

[1]ABCG2 = ATP-binding cassette, subfamily G, member 2; BLAD = bovine leukocyte adhesion deficiency; CVM = complex vertebral malformation; DGAT1 = diacylglycerol O-acyltransferase 1; RED = red coat color.

[2]BTA = Bos taurus autosome.

[3]NCBI = National Center for Biotechnology Information.

[4]The SNP is not located within a known or predicted gene.

age and SCE on BTA14 are shown in Figures 3 and 4 and were selected because there is a known QTL for fat percentage and no known QTL for SCE on this chromosome. In the absence of a segregating QTL of large effect, the mean CEBV have an expected value of 0 and should be normally distributed. The distributions of CEBV for both traits deviate slightly from normality ($P < 0.05$) and are positively skewed. The apparent drop in frequency of observations near the mean for fat percentage is due to binning, and the mean CEBV is slightly positive (0.01; $P < 0.01$). The mean CEBV for SCE on BTA14 is 0 ($P > 0.05$), as expected. Selection for higher milk and components yields in the presence of a known QTL for fat percentage has not produced significantly skewed CEBV for fat percentage.

When haplotypes become routinely available, which will allow the tracing of parental origin of chromosomal segments, it may be possible to develop mate selection strategies that optimize expected progeny EBV by selecting complementary chromosomes. For example, cows with positive CEBV for chromosomes 13, 14, 16, 17, 19, or 20 might be selected for breeding to the bull O-Bee Manfred Justice-ET (7HO6417; Figure 5). Optimal solutions to that problem will require thorough knowledge of relationships among effects on individual chromosomes, as well as their relationships to aggregate genotypes. For example, Sonstegard et al. (2008) reported that 29 SNP had antagonistic associations with milk yield and fertility in Holsteins, suggesting that it may not be possible to produce genotypes in which all chromosomes have favorable values for a selection objective that includes both traits.

Expected and Realized Genetic Effects

Ratios of observed versus expected allele effects for yield were higher than any of the 3 variance ratios (VB, VC, or VWB) for autosomes with no evidence of QTL (Table 4) and were much lower for chromosomes with evidence of QTL effects (e.g., BTA6 and BTA14). Given the lack of predictive power of SA for known QTL, it was not calculated for the merit, calving, and conformation traits (Tables 5 and 6). Results were similar for the VB and VC ratios; the results presented below focus primarily on the VC and VWB variance ratios.

All 3 variance-derived methods identified the QTL affecting fat percentage on BTA14; protein percentage on BTA6 and BTA14; fertility, longevity, economic merit, dystocia, and conformation on BTA18; and fertility on BTA24. Variance ratios were similar for the QTL on BTA6, BTA14, and BTA24 when calculated using base rather than current allele frequencies, suggesting that QTL allele frequencies have changed little in response to selection based primarily on yield. Differences be-
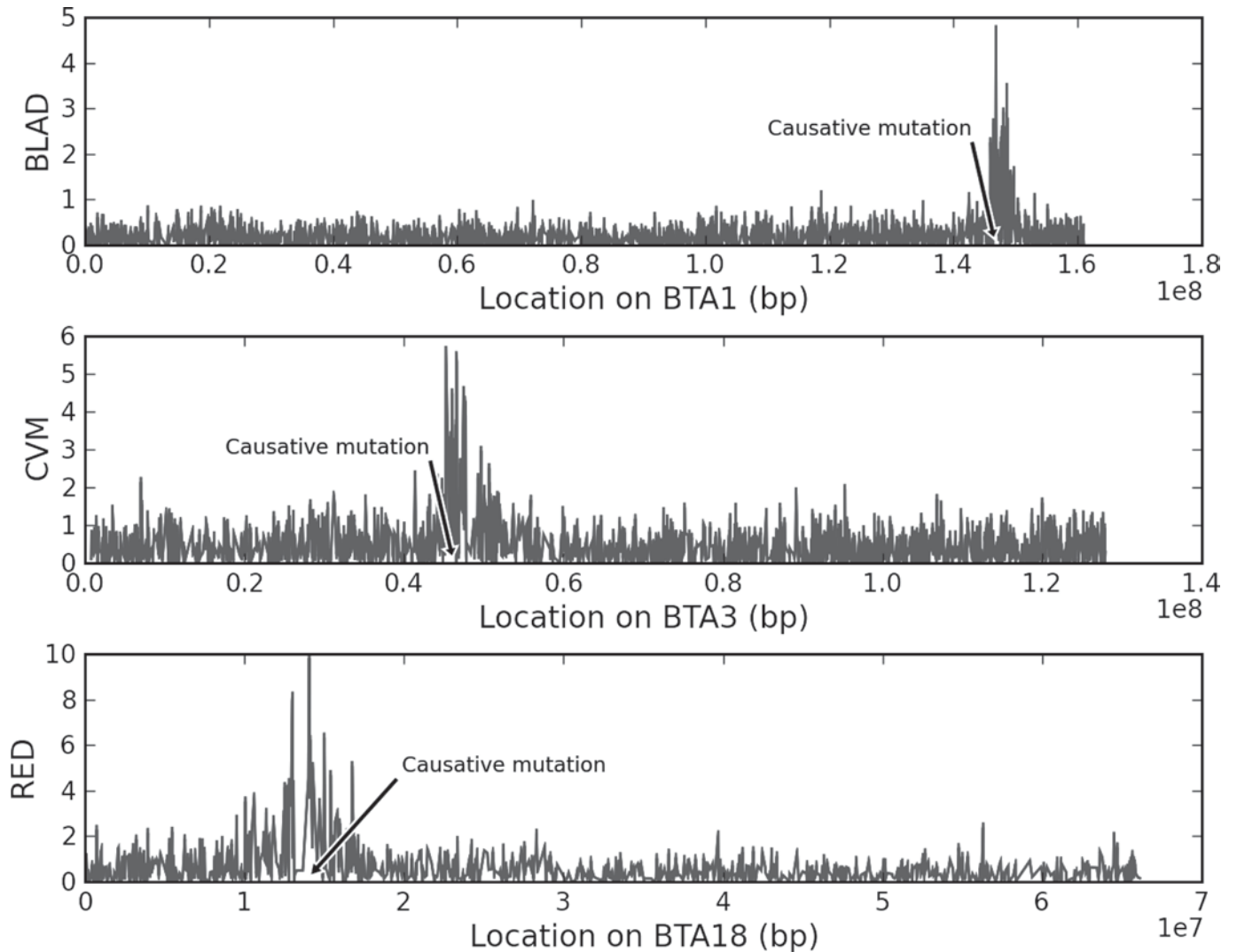
Figure 2. Size (in SD of marker effects) and location of marker effects on *Bos taurus* (BTA) autosome 1 affecting bovine leukocyte adhesion deficiency (BLAD), BTA3 affecting complex vertebral malformation (CVM), and BTA18 affecting red coat color (RED).

tween variance ratios for the QTL on BTA18 ranged from 2.0 to 6.1, and the VC ratios were always larger than VWB (Table 6). The frequency of the unfavorable allele for SNP ss86324977 decreased from 13% in the base population to 6% in the most recent animals, perhaps as a result of selection against dystocia or for net merit.

Variance ratios calculated on a within-sire basis may be preferred to VB or VC as a measure of the magnitude of genetic effects because they account for LD among adjacent loci. The VB and VC methods assume that adjacent loci are statistically independent, implying that covariances among all loci are 0, which is not the case for markers that are in LD. In the VWB approach, CEBV and GEBV were calculated for each bull and then averaged over all genotyped bulls to cal-

culate ratios for each chromosome. In some cases, the VWB ratio for a known QTL is larger than for VB or VC, and in some cases, it is smaller. Values of VWB substantially different from VB or VC may identify regions in which there are multiple loci contributing to variation in a trait and the transmission of LD blocks is being tracked, rather than the individual causative mutations.

Many association studies have reported upward biases in estimates of marker effects when the marker location and effect size are jointly estimated. Göring et al. (2001) first described this problem in detail, showing that the bias is a function of several parameters, including the true effect size and the sample size. They also noted that in many common situations, estimates of marker effects are essentially independent of the true marker
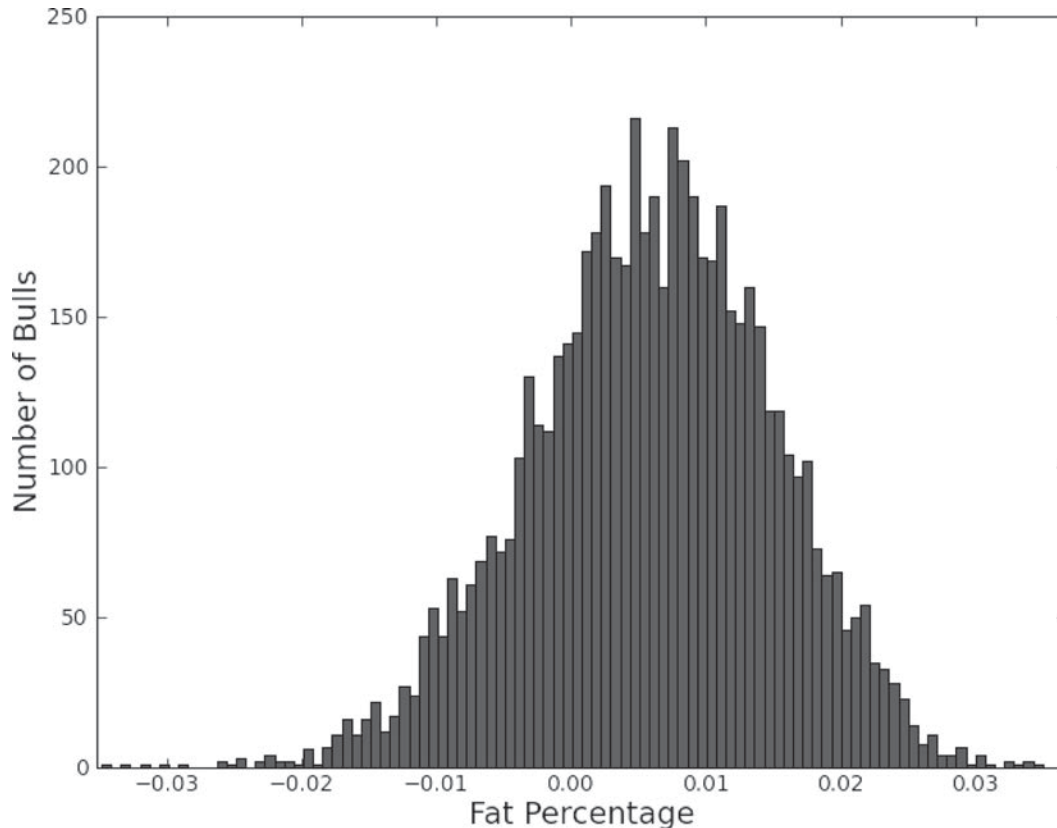
Figure 3. Distribution of chromosomal EBV of fat percentage on *Bos taurus* autosome 14; the diacylglycerol O-acyltransferase QTL is located on this chromosome.

effects. Biased marker effects would have an effect on all 4 calculated ratios (SA, VB, VC, and VWB), with the variances being most affected, and this may partially explain the differences among VB, VC, and VWB discussed above. Several approaches for minimizing the problem have been proposed in the literature, but no particularly satisfying solutions have been developed. Simultaneous estimation of all marker effects, such as performed in this study, may reduce but not eliminate upward biases. Allison et al. (2002) also suggested that the use of empirical Bayes approaches may help reduce bias. Given the large sample size in this study and the high correlation between marker effects from the linear and nonlinear models, it is tempting to assume that the marker effects are unbiased, but this may not be the case. The magnitude of marker effects should be interpreted cautiously until confirmed by follow-up studies.

The expected proportions of genetic variance for each trait accounted for by SNP on a chromosome were calculated based on chromosome lengths assuming that all markers had equal effects. Based on its length, BTA18 was expected to account for 2.2% of the additive genetic variance for each trait, but it actually accounted for 2.9% (economic merit) to 7.6% (SCE) of the addi-

tive genetic variation. Markers on BTA1 accounted for considerably less variance than expected based on its length, although it is unclear why this would be the case.

The active X chromosome in female humans and mice, and the sole X chromosome in males, is upregulated by a doubling of global gene expression (Nguyen and Disteche, 2006). As a result, X-linked loci are expected to explain as much variance as autosomal loci. However, the X chromosome accounts for much less variance than expected when using VC but is closer to expectation when VWB is used, suggesting that there may be large LD blocks on that chromosome. Vicoso and Charlesworth (2006) reported that DNA sequences of X-linked genes under positive selection change faster than those under negative selection, which may lead to higher rates of fixation and lower-than-expected variance for genes on the X chromosome.

Marker association effects for most other traits were evenly distributed across all chromosomes, with only a few regions having larger effects. This may explain why the infinitesimal model and standard quantitative genetic theories have worked well. The distribution of marker effects indicates that favorable alleles will not
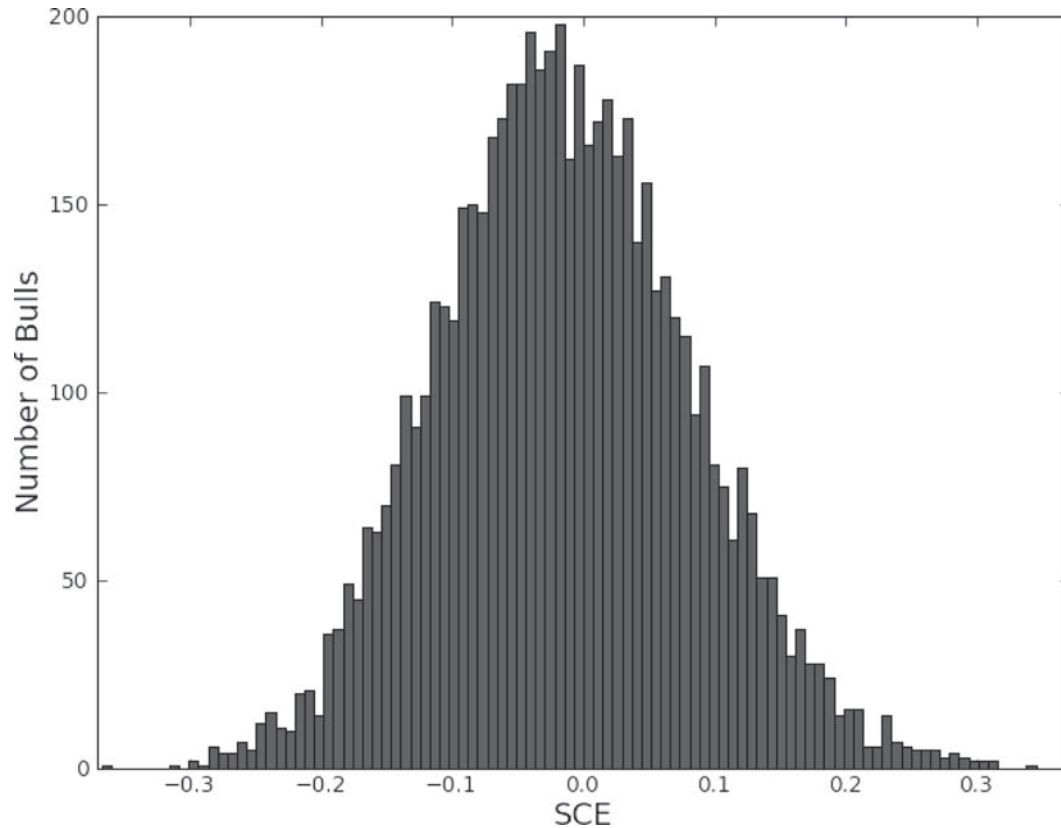
Figure 4. Distribution of chromosomal EBV of sire calving ease (SCE) on *Bos taurus* autosome (BTA) 14; there are no QTL on BTA14 affecting this trait.

become quickly fixed within a population, and genetic variation will remain even after intense selection. Thus, dairy cattle breeders may expect genetic progress to continue for many generations.

Calving Complex

Bulls that sire large calves transmit genes to their daughters, which, in turn, support high fetal growth rates. McDermott et al. (1992) reported that calf birth weight is the single most important factor in predicting dystocia, and Johanson and Berger (2003) showed that birth weight is also a valuable predictor of the probability of stillbirth. Comparison of genome-wide and BTA18-specific correlations among SCE, DCE, longevity, net merit, stature, strength, body depth, and rump width suggests that the putative QTL on BTA18 is associated with calf growth rate.

Correlations among BTA18-specific EBV (Table 7, upper diagonal) reveal favorable correlations among longevity and economic merit ($r = 0.88$), undesirable correlations among conformation and calving ease traits ($r = 0.78$ to 0.95), and unfavorable correlations of calving ease with net merit, longevity, and conformation ($r = -0.44$ to -0.72). Comparison of BTA18 specific to genome-wide correlations (Table 7, lower diagonal) strongly suggests that the QTL on BTA18 affects calf size and that selection for extreme conformation (larger body size) has resulted in larger calves and increased rates of dystocia, which are largely attributable to feto-pelvic incompatibility (Meijering, 1984).

The genome-wide correlation among SCE and DCE is positive ($r = 0.58$), suggesting that selection for improved SCE will also improve DCE. The QTL on BTA18 has the same directional effect on SCE as DCE and a larger correlation ($r = 0.81$), indicating a much larger direct than maternal effect. Large heifer calves grow into large cows, but phenotypic correlations among external body measures and the size of the internal pelvic opening are essentially 0 (Meijering, 1984), and correlations between external body dimensions and dystocia are near 0 (Naazie et al., 1989; Sieber et al., 1989). Large body size is not accompanied by large internal pelvic openings that are capable of easily birthing large calves. The genetic correlation between BTA18-specific and genome-wide sire EBV for SCE is moderate ($r = 0.42$), as is that between BTA18-specific and genome-wide sire EBV for DCE ($r = 0.30$; Table 8), suggesting that

Table 4. Ratio of observed versus expected effects for milk (MILK), fat (FAT), and protein (PROT) yields and component percentages by chromosome calculated using sums of allele effects (SA), additive genetic variances calculated using current population allele frequencies (VC), and calculated from individual marker effects on a within-bull basis (VWB)

| BTA[1]	MILK			FAT			PROT			FAT%			PROT%		
	SA	VC	VWB	SA	VC	VWB	SA	VC	VWB	SA	VC	VWB	SA	VC	VWB
1	0.9	0.7	0.4	0.9	0.6	0.5	0.9	0.8	0.5	0.9	0.3	0.2	0.9	0.5	0.3
6	1.0	0.9	0.8	0.9	0.5	0.4	1.0	0.9	0.8	1.1	0.6	0.8	1.6	4.5	2.7
14	1.3	5.9	5.4	1.4	9.1	8.6	1.1	1.4	1.8	1.9	18.6	27.0	1.5	7.3	7.6
18	1.2	1.2	1.4	1.2	1.1	1.6	1.3	1.5	1.9	1.3	0.7	1.8	1.1	0.7	0.9
23	1.3	1.2	2.3	1.2	0.9	1.9	1.2	1.3	2.0	1.1	0.5	1.2	1.2	0.8	1.7
24	1.1	0.9	1.2	1.1	0.8	0.9	1.1	1.1	1.2	1.0	0.4	0.7	1.0	0.6	0.9
28	1.2	1.2	2.4	1.2	1.0	2.4	1.2	1.6	2.7	1.1	0.5	1.5	1.2	0.8	1.6
X	0.5	0.5	0.6	0.5	0.5	1.0	0.5	0.6	0.7	0.5	0.3	0.8	0.5	0.4	0.7

[1]BTA = *Bos taurus* autosome.

that there are several other loci across the genome that affect dystocia and that many of them have opposite effects on direct and maternal traits.

Genome-wide correlations among SCE and conformation are moderate and positive, ranging from 0.23 to 0.30, indicating that calving difficulty will increase as calves increase in size. The BTA18-specific correlations are large and positive, ranging from 0.69 to 0.79. Such large differences in genetic correlations suggest that many factors, including conformation, have an effect

on dystocia on a genome-wide basis but that the QTL on BTA18 is related specifically to calf size.

The relationships among DCE and conformation follow a similar pattern to SCE. Genome-wide correlations among DCE and conformation are moderate and positive, ranging from 0.23 to 0.30, indicating that over time larger cows will themselves have more calving difficulty. The BTA18-specific correlations are large and positive, ranging from 0.69 to 0.79, and reflect the effect of dam genes on calf growth rate in utero. Dam genes
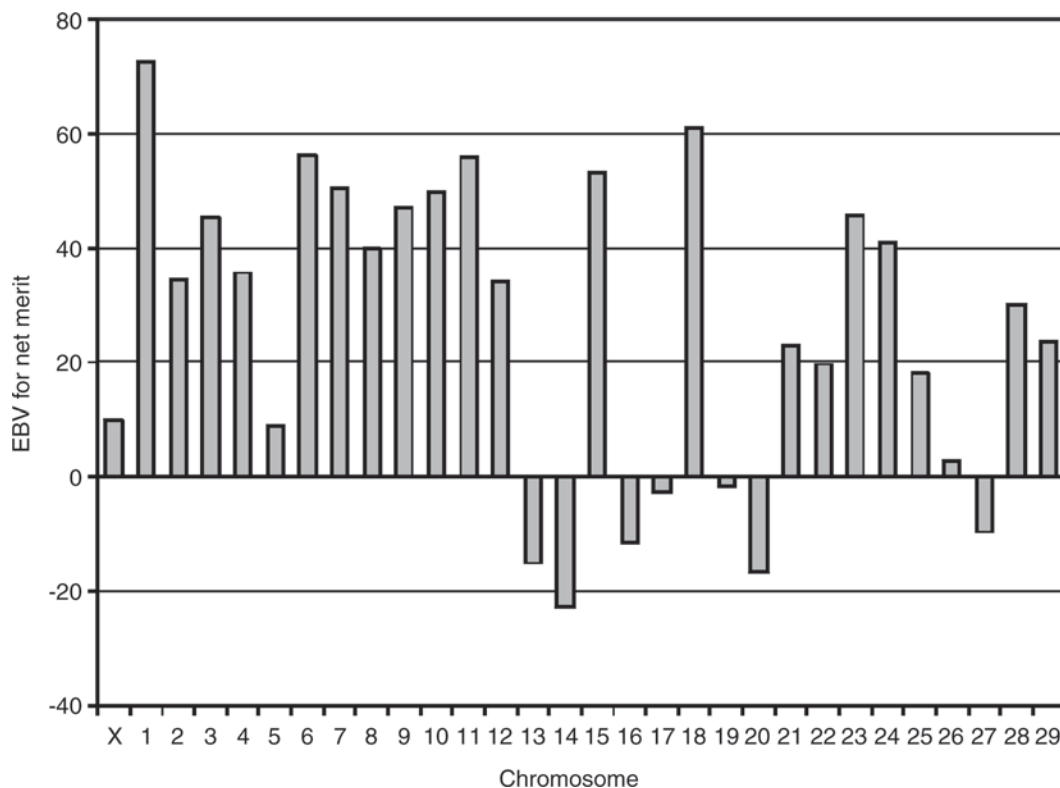
Figure 5. Estimated breeding values of individual chromosomes for net merit (NM) of the Holstein bull O-Bee Manfred Justice-ET (7HO6417), whose official May 2008 proof for NM is +$781.

Table 5. Ratio of observed versus expected effects of SCS, daughter pregnancy rate (DPR), productive life (PL), and net merit (NM) by chromosome for additive genetic variances calculated using current population allele frequencies (VC) and calculated from individual marker effects on a within-bull basis (VWB)

BTA[1]	SCS		DPR		PL		NM	
	VC	VWB	VC	VWB	VC	VWB	VC	VWB
1	0.6	0.4	0.7	0.6	0.7	0.4	0.8	0.4
6	1.7	1.4	0.9	1.1	1.1	1.6	0.8	0.8
14	1.1	1.4	0.7	1.0	1.0	1.4	1.8	2.5
18	1.2	1.3	1.7	3.4	7.0	5.4	3.2	3.7
23	1.0	2.3	0.9	2.0	1.2	3.1	1.5	3.6
24	3.9	1.3	4.9	2.5	0.9	0.9	1.2	1.3
28	0.9	1.4	1.1	1.5	0.9	1.3	1.1	2.2
X	0.5	0.7	1.6	1.2	0.7	0.9	0.7	0.7

[1]BTA = *Bos taurus* autosome.

can affect calf birth weight by their expression in the growing fetus as well as in the uterine environment.

The QTL on BTA18 had the largest effects genome-wide on both sire and daughter stillbirth. The effect was greatest on sire stillbirth and is consistent with previously reported relationships of dystocia with risk of stillbirth (Meyer et al., 2000). Larger calves are more likely to experience dystocia, and dystocia is genetically associated with stillbirths (Hansen et al., 2004). The effect of the QTL on daughter stillbirth was small, implying that it does not significantly increase the likelihood of a bull's daughters delivering stillborn calves.

Several studies have reported QTL for conformation traits, calving traits, fertility, and longevity on BTA18. A QTL affecting longevity, maternal dystocia, and maternal stillbirth in the German Holstein population was described by Kühn et al. (2003). Holmberg and Andersson-Eklund (2006) identified a QTL affecting maternal calving ease in a resource population of Swedish Reds and Swedish Holsteins. Thomasen et al. (2008) located a QTL with pleiotropic effects on dystocia, stillbirth, and calf size in a population of Danish Holsteins.

Schnabel et al. (2005) performed a genome scan in 2 half-sib US Holstein families to identify QTL for yield and conformation and reported microsatellite associations with foot angle, fore udder attachment, and teat length. Their analysis included similar calving ease and conformation data to the current study but did not report significant QTL effects on BTA18, possibly due to the limited number of parental chromosomes screened in that study. Their results were similar to the earlier findings of Schrooten et al. (2000), who genotyped 949 young bulls in the Netherlands belonging to 22 half-sib families and reported no significant QTL effects for conformation or dystocia on BTA18.

More recently, Kolbehdari et al. (2008) genotyped 462 Canadian Holstein bulls and identified SNP rs41636734 at 53,743,293 bp on BTA18 to be associated with conformation, the mammary system, overall rump score, and direct (sire) calving ease. They reported no significant associations with dairy strength, stature, herd life, or fertility, possibly because of the limited number of genotyped bulls. This SNP is located in an intron of the cluster of differentiation 37 (CD37) gene, which is a member of the transmembrane 4 superfamily, whose members are cell-surface proteins that mediate signal transduction events (van Spriel et al., 2004). The SNP reported in this study (ss86324977) is located in an intron of the gene encoding the Siglec-5 protein at 57,125,868 bp, which is far

Table 6. Ratio of observed versus expected effects of sire calving ease (SCE), daughter calving ease (DCE), stature (STAT), strength (STR), body depth (BDEP), and rump width (RWID) by chromosome for additive genetic variances calculated using current population allele frequencies (VC) and calculated from individual marker effects on a within-bull basis (VWB)

BTA[1]	SCE		DCE		STAT		STR		BDEP		RWID	
	VC	VWB	VC	VWB	VC	VWB	VC	VWB	VC	VWB	VC	VWB
1	0.7	0.5	0.9	0.6	0.5	0.5	0.6	0.6	0.7	0.6	0.5	0.4
6	0.6	0.6	0.8	0.9	0.5	0.5	0.5	0.6	0.5	0.6	0.6	0.7
14	0.9	1.1	1.1	1.5	0.6	0.6	2.8	1.1	0.7	1.1	0.6	0.9
18	9.7	7.3	3.9	3.8	4.4	2.7	8.0	5.2	10.5	6.3	11.0	5.2
23	0.9	1.3	1.6	3.6	0.8	1.8	0.6	1.2	0.8	1.9	0.9	1.5
24	1.0	1.0	1.0	1.0	0.5	1.0	0.6	1.8	0.7	1.6	0.6	0.8
28	1.1	3.0	1.0	1.2	1.5	3.2	1.0	2.4	0.9	1.8	1.0	1.7
X	0.8	0.7	1.2	1.1	0.5	0.9	0.4	1.2	0.5	1.1	0.5	0.9

[1]BTA = *Bos taurus* autosome.

Table 7. Correlations among BTA18[1]-specific (above the diagonal) and genome-wide (below the diagonal) sire estimated breeding values for net merit (NM), longevity (PL), sire (SCE) and daughter calving ease (DCE), stature (STAT), strength (STR), body depth (BDep), and rump width (RWid)

Item	NM	PL	SCE	DCE	STAT	STR	BDep	RWid
NM		0.88	−0.64	−0.67	−0.44	−0.53	−0.60	−0.46
PL	0.71		−0.72	−0.71	−0.48	−0.62	−0.67	−0.55
SCE	−0.32	−0.31		0.81	0.69	0.79	0.79	0.79
DCE	−0.56	−0.44	0.58		0.57	0.73	0.75	0.69
STAT	0.16	0.00	0.25	−0.13		0.82	0.83	0.78
STR	0.02	−0.09	0.28	0.03	0.72		0.95	0.90
BDep	−0.01	−0.18	0.30	0.05	0.78	0.91		0.88
RWid	0.11	−0.02	0.23	−0.07	0.68	0.71	0.72	

[1]BTA18 = *Bos taurus* autosome 18.

enough from the CD37 gene that there should be no significant LD between them if the assembly is correct in this region of the genome.

These results support our findings that there is a QTL on BTA18 that has significant effects on body size, dystocia, longevity, and economic merit. The QTL results in high calf birth weights, which are in turn associated with greater risks of dystocia, reduced dam fertility, and reduced lifetime economic merit of dams (Meijering, 1984; Mee, 2008). Large cows, in turn, bear large calves with increased risks of dystocia. The large, positive genetic correlations among the conformation and dystocia traits imply that the QTL has a large effect on calf size, and the association of SNP ss86324977 with the Siglec-5 gene suggests a possible underlying mechanism. Although GEBV for stillbirth have yet to be calculated, and were unavailable for analysis, the literature is clear that high calf birth weights substantially increase the risk of stillbirth (Johanson and Berger, 2003).

CONCLUSIONS

A nonlinear model with heavy tails allows markers with large effects to have predictions far away from the mean and regresses markers with small effects toward zero, reflecting the biology of large-effect QTL such as *DGAT1*. These models provide a better fit to the data than do models that allow only a finite number of loci to have nonzero effects. No advantage was gained by including interactions between heavy tails and the number of nonzero marker effects. Linear model predictions were intermediate to the heavy-tailed and finite loci models. The superior predictive ability of heavy-tailed models when compared with finite locus models validates quantitative genetic theory by demonstrating that most genes have small effects, although there are genes of large effect associated with some traits.

Marker associations on BTA18 were centered on SNP ss86324977, which had large effects on economic merit, longevity, calving ease, and conformation. Relationships among those traits may be attributable to a gene product or regulatory element associated with calf birth weight, and the SNP associated with the QTL is located in an intron of the Siglec-5 gene, which may regulate the onset of parturition through leptin sequestration. The presence of significant marker effects resulted in greater-than-expected proportions of explained genetic variance being assigned to BTA18. Analysis of known recessive genes suggests that marker solutions from

Table 8. Correlations of BTA18[1]-specific[2] with genome-wide sire EBV for net merit (NM), longevity (PL), sire (SCE) and daughter calving ease (DCE), stature (STAT), strength (STR), body depth (BDep), and rump width (RWid)

Item	NM_{GW}	PL_{GW}	SCE_{GW}	DCE_{GW}	$STAT_{GW}$	STR_{GW}	$BDep_{GW}$	$RWid_{GW}$
NM_{18}	0.34	0.28	−0.32	−0.27	−0.11	−0.13	−0.17	−0.10
PL_{18}	0.28	0.32	−0.35	−0.27	−0.11	−0.15	−0.20	−0.10
SCE_{18}	−0.25	−0.23	0.42	0.30	0.05	0.14	0.16	0.12
DCE_{18}	−0.22	−0.23	0.29	0.30	0.05	0.12	0.16	0.10
$STAT_{18}$	−0.17	−0.19	0.27	0.19	0.16	0.20	0.22	0.18
STR_{18}	−0.16	−0.20	0.29	0.22	0.09	0.21	0.21	0.14
$BDep_{18}$	−0.18	−0.22	0.31	0.23	0.14	0.22	0.27	0.17
$RWid_{18}$	−0.13	−0.16	0.30	0.21	0.07	0.16	0.17	0.15

[1]BTA18 = *Bos taurus* autosome 18.

[2]Subscripts differentiate between BTA18-specific (18) and genome-wide (GW) estimated breeding value.

genomic evaluations may be useful for identifying chromosomal regions that merit further study.

ACKNOWLEDGMENTS

REFERENCES

Agerholm, J. S., C. Bendixen, O. Andersen, and J. Arnbjerg. 2001. Complex vertebral malformation in Holstein calves. J. Vet. Diagn. Invest. 13:283–289.

Agerholm, J. S., F. McEvoy, and J. Arnbjerg. 2006. Brachyspina syndrome in a Holstein calf. J. Vet. Diagn. Invest. 18:418–422.

Allison, D. B., J. R. Fernandez, M. Heo, S. Zhu, C. Etzel, T. M. Beasley, and C. I. Amos. 2002. Bias in estimates of quantitative-trait-locus effect in genome scans: Demonstration of the phenomenon and a method-of-moments procedure for reducing bias. Am. J. Hum. Genet. 70:575–585.

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. (Abstr.)

Bellefroid, E. J., D. A. Poncelet, P. J. Lecocq, O. Revelant, and J. A. Martial. 1991. The evolutionarily conserved Krüppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. Proc. Natl. Acad. Sci. USA 88:3608–3612.

Brinkman-Van der Linden, E. C. M., N. Hurtado-Ziola, T. Hayakawa, L. Wiggleton, K. Benirschke, A. Varki, and N. Varki. 2007. Human-specific expression of Siglec-6 in the placenta. Glycobiology 17:922–931.

Cohen-Zinder, M., E. Seroussi, D. M. Larkin, J. J. Loor, A. Everts-van der Wind, J. H. Lee, J. K. Drackley, M. R. Band, A. G. Hernandez, M. Shani, H. A. Lewin, J. I. Weller, and M. Ron. 2005. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. Genome Res. 15:936–944.

Falconer, D. S., and T. F. C. MacKay. 1996. Introduction to Quantitative Genetics, 4th ed. John Wiley and Sons, New York, NY.

Foussias, G., G. M. Yousef, and E. P. Diamandis. 2000. Identification and molecular characterization of a novel member of the Siglec family (Siglec9). Genomics 67:171–178.

Ganss, B., and A. Jheon. 2004. Zinc finger transcription factors in skeletal development. Crit. Rev. Oral Biol. Med. 15:282–297.

Gengler, N., P. Mayeres, and M. Szydlowski. 2007. A simple method to approximate gene content in large pedigree populations: Application to the myostatin gene in dual-purpose Belgian Blue cattle. Animal 1:21–28.

Göring, H. H. H., J. D. Terwilliger, and J. Blangero. 2001. Large upward bias in estimation of locus-specific effects from genomewide scans. Am. J. Hum. Genet. 69:1357–1369.

Grisart, B., F. Farnir, L. Karim, N. Cambisano, J. J. Kim, A. Kvasz, M. Mni, P. Simon, J. M. Frere, W. Coppieters, and M. Georges. 2004. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. Proc. Natl. Acad. Sci. USA 101:2398–2403.

Hansen, M., M. S. Lund, J. Pedersen, and L. G. Christensen. 2004. Gestation length in Danish Holsteins has weak genetic associations with stillbirth, calving difficulty, and calf size. Livest. Prod. Sci. 91:23–33.

Harris, B. L., D. L. Johnson, and R. J. Spelman. 2008. Genomic selection in New Zealand and the implications for national genetic evaluation. Pages 325–330 in Proc. 36th ICAR Biennial Session, Niagara Falls, NY. Interbull Bull., Uppsala, Sweden.

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

Houseknecht, K. L., C. A. Baile, R. L. Matteri, and M. E. Spurlock. 1998. The biology of leptin: A review. J. Anim. Sci. 76:1405–1420.

Joerg, H., H. R. Fries, E. Meijerink, and G. F. Stranzinger. 1996. Red coat color in Holstein cattle is associated with a deletion in the MSHR gene. Mamm. Genome 7:317–318.

Johanson, J. M., and P. J. Berger. 2003. Birth weight as a predictor of calving ease and perinatal mortality in Holstein cattle. J. Dairy Sci. 86:3745–3755.

Khatkar, M. S., P. C. Thompson, I. Tammen, and H. W. Raadsma. 2004. Quantitative trait loci mapping in dairy cattle: Review and meta-analysus. Genet. Sel. Evol. 36:163–190.

Kolbehdari, D., Z. Wang, J. R. Grant, B. Murdoch, A. Prasad, Z. Xiu, E. Marques, P. Stothard, and S. S. Moore. 2008. A whole-genome scan to map quantitative trait loci for conformation and functional traits in Canadian Holstein bulls. J. Dairy Sci. 91:2844–2856.

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

Margolin, J. F., J. R. Friedman, W. K.-H. Meyer, H. Vissing, H.-J. Thiesen, and F. J. Rauscher. 1994. Krüppel-associated boxes are potent transcriptional repression domains. Proc. Natl. Acad. Sci. USA 91:4509–4513.

Matukumalli, L. K., C. T. Lawley, R. D. Schnabel, J. F. Taylor, M. Allan, M. P. Heaton, J. R. O'Connell, T. S. Sonstegard, T. P. L. Smith, S. S. Moore, and C. P. Van Tassell. Development and characterization of a high density SNP genotyping assay for cattle. Genome Res. (accepted)

McDermott, J. J., O. B. Allen, S. W. Martin, and D. M. Alves. 1992. Patterns of stillbirth and dystocia in Ontario cow-calf herds. Can. J. Vet. Res. 56:47–55.

McKay, S. D., R. D. Schnabel, B. M. Murdoch, L. K. Matukumalli, J. Aerts, W. Coppieters, D. Crews, E. D. Neto, C. A. Gill, C. Gao, H. Mannen, P. Stothard, Z. Wang, C. P. Van Tassell, J. L. Williams, J. F. Taylor, and S. S. Moore. 2007. Whole genome linkage disequilibrium maps in cattle. BMC Genet. 8:74–85.

Mee, J. F. 2008. Prevalence and risk factors for dystocia in dairy cattle: A review. Vet. 176:93–101.

Meijering, A. 1984. Dystocia and stillbirth in cattle – A review of causes, relationships, and implications. Livest. Prod. Sci. 11:143–177.

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

Meyer, C. L., P. J. Berger, and K. J. Koehler. 2000. Interactions among factors affecting stillbirths in Holstein cattle in the United States. J. Dairy Sci. 83:2657–2663.

Mounzih, K., J. Qiu, A. Ewart-Toland, and F. F. Chehab. 1998. Leptin is not necessary for gestation and parturition but regulates

maternal nutrition via a leptin resistance state. Endocrinology 139:5259–5262.

Naazie, A., M. M. Makarechian, and R. T. Berg. 1989. Factors influencing calving difficulty in beef heifers. J. Anim. Sci. 67:3243–3249.

Nguyen, D. K., and C. M. Disteche. 2006. Dosage compensation of the active X chromosome in mammals. Nat. Genet. 38:47–53.

Patel, N., E. C. M. Brinkman-Van der Linden, S. W. Altmann, K. Gish, S. Balasubramanian, J. C. Timans, D. Peterson, M. P. Bell, J. F. Bazan, A. Varki, and R. A. Kastelein. 1999. OB-BP1/Siglec-6. A leptin- and sialic acid-binding protein of the immunoglobulin superfamily. J. Biol. Chem. 274:22729–22738.

Sargolzaei, M., F. S. Schenkel, G. B. Jansen, and L. R. Schaeffer. 2008. Extent of linkage disequilibrium in Holstein cattle in North America. J. Dairy Sci. 91:2106–2117.

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

Schnabel, R. D., T. S. Sonstegard, J. F. Taylor, and M. S. Ashwell. 2005. Whole-genome scan to detect QTL for milk production, fertility and functional traits in two US Holstein families. Anim. Genet. 36:408–416.

Schrooten, C. S., 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.

Shuster, D. E., M. E. Kehrli Jr., M. R. Ackermann, and R. O. Gilbert. 1992. Identification and prevalence of a genetic defect that causes leukocyte adhesion deficiency in Holstein cattle. Proc. Natl. Acad. Sci. USA 89:9225–9229.

Sieber, M., A. E. Freeman, and D. H. Kelley. 1989. Effects of body measurements and weight on calf size and calving difficulty of Holsteins. J. Dairy Sci. 72:2402–2410.

Sonstegard, T. S., L. Ma, J. B. Cole, G. R. Wiggans, C. P. Van Tassell, G. Liu, B. Mariani, B. Crooker, P. M. VanRaden, M. Silva, and Y. Da. 2008. Genomic signatures of artificial selection in U.S. Holstein cows. ISAG Proc. Poster 2098. http://www.isag.org.uk/ISAG/all/2008_ISAG_Amsterdam_P2000.pdf Accessed March 11, 2009.

Tempelman, R. J. 1998. Generalized linear mixed models in dairy cattle breeding. J. Dairy Sci. 81:1428–1444.

Thomasen, J. R., B. Guldbrandtsen, P. Sørensen, B. Thomsen, and M. S. Lund. 2008. Quantitative trait loci affecting calving traits in Danish Holstein cattle. J. Dairy Sci. 91:2098–2105.

van Spriel, A. B., K. L. Puls, M. Sofi, D. Pouniotis, H. Hochrein, Z. Orinska, K. P. Knobeloch, M. Plebanski, and M. D. Wright. 2004. A regulatory role for CD37 in T cell proliferation. J. Immunol. 172:2953–2961.

Van Tassell, C. P., T. P. L. Smith, L. K. Matukumalli, J. F. Taylor, R. D. Schnabel, C. T. Lawley, C. D. Haudenschild, S. S. Moore, W. C. Warren, and T. S. Sonstegard. 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. Nat. Methods 5:247–252.

VanRaden, P. M. 2007. Genomic measures of relationship and inbreeding. Interbull Bull. 37:33–36.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423.

VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. J. Dairy Sci. 92:16–24.

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

Vicoso, B., and B. Charlesworth. 2006. Evolution on the X chromosome: Unusual patterns and processes. Nat. Rev. Genet. 7:645–653.