Polymorphism of the Bovine *CSN1S1* Promoter: Linkage Mapping, Intragenic Haplotypes, and Effects on Milk Production Traits

E.-M. Prinzenberg,* C. Weimann,* H. Brandt,*

J. Bennewitz,† E. Kalm,† M. Schwerin,‡ and G. Erhardt*

*Institute for Animal Breeding and Genetics,
Justus-Liebig-University, 35390 Giessen, Germany

†Institute for Animal Breeding and Husbandry,
Christian Albrechts University, 24098 Kiel, Germany

‡Research Institute for the Biology of Farm Animals,
18196 Dummerstorf, Germany

ABSTRACT

The bovine *CSN1S1* 5' flanking region (*CSN1S1*-5') was screened for polymorphisms in different cattle breeds. Single-strand conformation polymorphisms (SSCP) and sequence analyses revealed four alleles (1–4), two of them being new allelic forms (3 and 4). Sequences were deposited in GenBank with accession numbers AF549499–502. In alleles 1 and 4, potential transcription factor binding sites are altered by the mutations. Using SSCP analysis, all four alleles were identified in German Holsteins. Six intragenic haplotypes comprising *CSN1S1*-5' (alleles 1, 2, 3, 4) and exon 17 (*CSN1S1**B and C) genotypes were found. Linkage mapping using half-sib families from the German QTL project positioned *CSN1S1* between the markers *FBN14* and *CSN3*, with 5.6 cM distance between *CSN1S1* and *CSN3*. Variance analysis, using family and *CSN1S1* promoter genotypes as fixed effects, of breeding values and deregressed proofs for milk production traits (milk, fat, and protein yield and also fat and protein percentage) revealed significant effects on protein percentage when all families and genotypes were considered. Contrast calculations assigned a highly significant effect to genotype 24, which was associated with highest LS-means for protein percentage breeding values. As *CSN1S1* is one of the main caseins in milk, this could be an effect of mutations in regulatory elements in the promoter region. An effect on milk yield breeding values was indicated for genotype 12, but is probably caused by a linked locus.

(Key words: casein promoter, linkage mapping, candidate gene, variance analysis)

Abbreviation key: *CSN1S1* = casein alpha S1, SSCP = single strand conformation polymorphism.

INTRODUCTION

Effects of milk protein polymorphisms on milk production traits have been investigated during the past decades and, in some cases, results are still conflicting (e.g., Aleandri et al., 1990; Bovenhuis et al., 1992; Lin et al., 1986; Ng-Kwai-Hang et al., 1984, 1986, 1990). Depending on breed or region, the same genes and alleles were associated with different traits and effects. One reason for these contradictions could be a set of different intragenic haplotypes comprising distinct variants in the regulatory elements combined with the same protein variant expressed in milk, as postulated by Ehrmann et al. (1997). Strongest effects on first-lactation milk and protein yields seem to be associated with the *CSN1S1* locus (Lin et al., 1986). Moreover, whole genome and chromosome-wise scans using microsatellites and single nucleotide polymorphisms within the casein genes or casein polymorphisms as markers indicate different QTL for milk yield and components on BTA 6 (e.g., Bovenhuis and Weller, 1994; Freyer et al., 1995, 2002; Spelman et al., 1996; Zhang et al., 1998; Ashwell and van Tassel, 1999; Velmala et al., 1999; Mosig et al., 2001). Most studies found QTL for milk protein content and agree on the most likely QTL position for protein percentage near the marker *BM143* (Mosig et al., 2001; Nadesalingam et al., 2001). Ron et al. (2001) narrowed the QTL region to 4 cM around that marker. Other groups (Freyer et al., 1995, 2000; Ikonen et al., 1999; Aswell and van Tassel, 1999; Mosig et al., 2001) still report QTL effects closer to the casein region. Together, both findings match 2-QTL models applied by Spelman et al. (1996), Zhang et al. (1998), and Velmala et al. (1999), which include both, the *BM143* and the casein region as potential QTL locations. Additional evidence was found for a QTL affecting yield traits near

Received November 25, 2002.

Accepted March 25, 2003.

Corresponding author: E. M. Prinzenberg; e-mail: Eva-maria.prinzenberg@agrar.uni-giessen.de.

or in the casein region (Wiener et al., 2000; Nadesalingam et al., 2001; Thomsen et al., 2001).

The caseins are described as a cluster of closely linked genes (Grosclaude et al. 1965), and a first linkage analysis based on milk protein genotyping revealed recombination fractions of 0.03 for CSN1S1/CSN2, 0.04 for CSN1S1/CSN3, and 0.06 for CSN2/CSN3 (Hines et al., 1981). Later the complex was physically mapped to BTA6 q31-33 (Threadgill and Womack, 1990) and genetically mapped with a marker in CSN3 to 82.6 cM (sex averaged map MARC97, Kappes et al., 1997). Following characterization of cosmid clones, the total length of the bovine casein complex is now estimated to be 200 to 250 kb, with CSN1S1 at the 5' and CSN3 at the 3'-end of the whole region (Rijnkels et al., 1997).

Due to the close linkage of all casein genes, the possibility of a domain regulation was postulated (Threadgill and Womack, 1990). Recently, a regulation conferred at the level of chromatin structure was discussed for the casein locus as well (Rijnkels et al., 2003). Earlier, Rijnkels et al. (1995) failed to identify functional regulatory elements in CSN1S2 and CSN3 gene clones using transgenic mice expression experiments, and therefore suggested, major regulatory sites might be located elsewhere in the casein locus. CSN1S1 is located at the 5' end of the whole casein region, and it is oriented in the sense direction so this gene and its 5' end could be the nearest to a hypothetical distal domain regulation element.

Previous studies were based on the assumption that virtually no recombination between the casein genes occurs. For QTL studies, casein haplotypes and intragenic microsatellites have been used to increase information content in the casein locus region (Lien and Rogne, 1993; Velmala et al., 1995).

CSN1S1 is nearly invariant at protein level in Holstein Friesians with frequencies for CSN1S1*B varying from 0.872 in Finnish (Lien et al., 1999) to 0.985 in Canadian and German Holsteins (Ng-Kwai-Hang et al., 1990; Erhardt, 1993a). Additional protein variants besides CSN1S1*B and C are the rare variants CSN1S1*A, D and F, which were found in US, Canadian, and German Holstein and German Black Pied cattle (Hines et al., 1977; Ng Kwai-Hang et al., 1984, 1986; Erhardt, 1993a, 1993b). Due to very low frequencies of CSN1S1*A, D, F and low frequency of CSN1S1*C, this protein polymorphism provides low information content in Holsteins and its use as genetic marker in the present highly selected Holstein Friesian population is limited. Freyer et al. (1999) found a heterozygosity of only 0.07 for CSN1S1 in Black and White cattle from Germany including pure breed Holstein Friesians, German Black and White and crossbreed cattle of German Black and White, 25% Jersey, and 50%

Holstein Friesian. Therefore, only the more polymorphic genes CSN2 and CSN3 were applied in their linkage analyses.

Different polymorphisms in the CSN1S1 5' flanking region were described in cows from various breeds (Schild and Geldermann, 1996). The 5' flanking region variants occurred predominantly in combination with distinct exon variants of the same gene (Ehrmann et al., 1997). Comparisons of two CSN1S1 promoter alleles discriminated by a MaeIII RFLP (Koczan et al., 1993) and the two alleles CSN1S1*B and C of the protein coding region revealed up to four haplotypes depending on the breed investigated (Jann et al., 2002). This indicates that the same protein variant may be associated with different promoter variants. Nevertheless, in highly selected milk breeds such as Angler and Ayrshire, only one or two haplotypes were found (Jann et al., 2002).

The aims of our present study were to identify allelic variation in the proximal CSN1S1 promoter region of German Holstein cattle, to use these as markers for linkage mapping of the CSN1S1 gene, to determine combinations with protein coding gene polymorphism (intragenic haplotypes) and to evaluate effects of promoter polymorphisms on milk production traits.

MATERIAL AND METHODS

DNA Samples and Data Collection

Three different sets of DNA samples were used in this study. First, 83 DNA samples of cows from eight breeds (6 German Black and White, 4 German Red, 7 German Yellow cattle, 18 Holstein, 9 Simmental, 13 Jersey, 20 Pinzgauer, 6 Simbrah) including CSN1S1 alleles A, B, C, D, and F were used to assess the general single-strand conformation polymorphism (SSCP; Orta et al., 1989) variability of the selected 5' flanking region and for optimization of the SSCP typing protocol. Intragenic haplotypes (comprising promoter and exon 17 polymorphism) were determined for these samples as well.

Second, individual DNA samples extracted from white blood cells of German Holstein cows were used for promoter allele frequency determinations (n = 490) and for examination of promoter and exon 17 haplotypes (n = 384). The cows (first calving years 1990 to 1997, n = 32, 18, 92, 117, 165, 29, 36, 7) originated from five farms in Hessen and Thuringen (29, 34, 65, 110, and 252 per farm) and were offspring of 139 bulls (1 to 18 daughters per bull).

For linkage mapping, DNA prepared from semen straws of bulls included in the granddaughter design from the German QTL project (details described in Thomsen et al., 2000) was used. Analysis of the effects

of promoter genotypes was done in eight Holstein half-sib families with a total of 678 bulls (average family size 84 sons with a range of 11 to 316 sons). All Holstein bulls' estimated breeding values (EBV) were taken from the national routine sire evaluation at Vereinigte Informationssysteme Tierhaltung (VIT). Reliability of the EBVs ranged from 90–99% (mean 94.55, SD 2.61). The Simmental family was only used for linkage mapping, as breeding values were not available for all traits.

Promoter Polymorphism Analyses

A 655 bp fragment from the proximal *CSN1S1* promoter (*CSN1S1*-5') was PCR amplified using primers selected from the genomic *CSN1S1* sequence (Koczan et al., 1991; GenBank Accession No. X59856) in 15 μ l reactions containing 20–100 ng genomic DNA, 10 pmol of each primer, 0.5 U Taq DNA-polymerase (PqLab Biotechnologie, Erlangen), 50 μ M dNTPs in standard reaction buffer (10 mM Tris-HCl (pH 8.8), 50 mM KCl, 1.5 mM MgCl₂) over 30 cycles: 1 min—93°C (1 $\times$), 40 sec.—91°C, 40 sec. 57°C, 40 sec.—70°C (30 $\times$) with a final 3 min elongation at 70°C. After completion of the PCR reaction, fragments were subjected to SSCP analysis. 25 μ l of denaturing loading dye (95% formamide) were added and the 96-well plate was heated to 93°C for 2 min and chilled on ice immediately. 4 μ l of the denatured samples were loaded on a 16 $\times$ 16 cm 12% non denaturing polyacrylamide gel containing 1% glycerol and were separated overnight at 420V—10°C for 20 hours and silver stained following essentially the protocol of Basam et al. (1991) but using cold 0.04 M EDTA as final stop solution. Alleles were designated numerically starting with the slowest migrating fragment. Primer sequences are available from the authors on request.

Sequence Analyses

For sequencing of the *CSN1S1*-5' variants, PCR products of cows that were homozygous for one of the four alleles were cloned into pCR 2.1 TOPO (Invitrogen B. V., Groningen, The Netherlands) following the manufacturer's protocol. Before sequencing, PCR amplificates of the plasmid DNA were checked by SSCP analysis again and only clones with migration patterns corresponding exactly to the one revealed from genomic DNA amplification were used for sequencing. Three clones per animal were sequenced in both, 5' $\rightarrow$ 3' and 3' $\rightarrow$ 5' directions. Sequences were processed and compared using Chromas Version 1.45 (www.technelysium.com.au/chromas.html), GeneDoc (Nicholas & Nicholas, 1997; www.psc.edu/biomed/genedoc) and DNASIS™ for Windows® (Hitachi Software Engineering America Ltd., San Bruno, USA). Search for putative transcrip-

tion factor binding sites was performed with DNASIS™ and TRANSFAC (Wingender et al., 2000).

Genotyping of Exon 17 Polymorphism

A 265 bp fragment including the variable position discriminating *CSN1S1**B and *CSN1S1**C was PCR amplified using primers 17669f (5'CTC TCT AGC TTT TCA GAC AA 3') and 17933r (5' AAG CAT TTA TGT GTC ATG CT 3') under the same conditions as for the promoter, except 1% DMSO in the reaction mixture and a lower annealing temperature of 54°C. Samples were run under SSCP conditions in a 9% non denaturing polyacrylamide gel (37:1) containing 0.5% glycerol for 5 h at 500 V and 5°C. The *CSN1S1**C specific pattern, confirmed by typing cows with known genotype, showed slower migration than the *CSN1S1**B specific fragments. Rare variants *CSN1S1**A, D and F were not separated from *CSN1S1**B using this test.

Linkage Mapping

Genotyping data for all half-sib families were available for 10 microsatellite markers on BTA6, which were earlier used to construct a project specific chromosome map (Thomsen et al., 2000). Genotyping results for the *CSN1S1*-5' polymorphism were integrated in the data set and the map was reconstructed by multipoint analyses using the BUILD option of the CRI-MAP programme package, version 2.4 (Green et al., 1990). The FLIPS 5 option of the same programme was used to determine the marker order with the highest log10 likelihood. Additionally, recombination rates and lod scores between two markers were calculated using the option TWOPOINT. The genetic distances between the markers were calculated using the Kosambi mapping function (Ott, 1991).

Statistical Analyses

Allele frequencies for *CSN1S1*-5' polymorphism in Holstein cows were determined by direct counting for each first calving year and the whole sampling set jointly.

For the bulls, variance analysis was performed using the GLM procedure of the SAS program package (SAS, 1989). Estimated breeding values for first-lactation milk, fat and protein yield, fat, and protein content were analysed using the *CSN1S1*-5' genotype and the family as fixed effects. The same analyses were performed using deregressed EBV. Deregression was done as described by Thomsen et al. (2001). This includes a correction for the additive relationship matrix, number of effective daughters, and residual sire variances. All

genotypes with more than 25 records in the dataset (12, 22, 23, and 24) and all families were included ($n = 635$) in this analysis. Pairwise tests for genotype differences were calculated with the contrast option of the GLM procedure (SAS, 1989).

Statistical model:

$$Y_{ijk} = \mu + CSg_i + F_j + e_{ijk},$$

where

Y_{ijk} = EBV of individual bull (or deregressed EBV),

μ = overall mean,

CSg_i = fixed effect of the *CSN1S1-5'* genotype (12, 22, 23, 24),

F_j = fixed effect of the family (1–8),

e_{ijk} = random residual effect.

RESULTS

Promoter Polymorphism

SSCP analysis of *CSN1S1-5'* in 83 cows of different breeds revealed four different allelic patterns that were named 1, 2, 3, and 4 in the order of increasing mobility. These four alleles were all found both in the randomly selected 490 Holstein cows and the nine paternal half-sib families. Allele frequencies were 0.033 (type 1), 0.738 (type 2), 0.191 (type 3), and 0.038 (type 4) for the cows. Genotype frequencies declined in the order $22 > 23 > 24 > 12 > 33 > 34$. No cows homozygous for 11, 44, or a combination of these rare alleles were found. Allele frequency variation in cows from first calving years 1990 to 1997 did not reveal any clear decline or increase for any of the alleles (results not shown). Linear trends showed constant frequency for alleles 1 and 4 over these 8 yr, a slight decrease of allele 2 and increase of allele 3 was indicated for the years 1992 to 1994, which had the biggest sample sizes.

Genotyping of all sires and grandsires of the granddaughter design revealed 12 of them being homozygous for genotype 22. The remaining 10 heterozygous bulls belong to eight families (seven Holstein, one Simmental) and six of them carried rare alleles 1 or 4 or both 1 and 4. In the eighth Holstein family, also used to calculate the effect of the *CSN1S1-5'* genotype, the father was homozygous for genotype 22, thus all alleles other than allele 2 were inherited from the maternal side. Genotypes 11, 33, and 44 were found in two to three bulls only. Combinations of the rare alleles 1, 3, and 4 were found in 11 to 20 bulls and three or four families only. Total numbers of bulls and distribution

of all genotypes within the nine families are given in Table 1. For the statistical analysis of the effects, rare genotypes 11, 13, 14, 33, 34, and 44 were excluded because of the low number of records and unequal distribution over the families.

Intragenic Haplotypes

Considering *CSN1S1-5'* and exon 17 genotypes in the Holstein cows, genotype combinations 12/BB, 12/BC, 13/BC, 22/BB, 23/BB, 23/BC, 24/BB, 33/BB, 34/BB were found. *CSN1S1-5'* genotype 12 or 13 was found in 26 *CSN1S1*B/CSN1S1*C* heterozygous cows, but as well in two *CSN1S1*B/CSN1S1*B* homozygous cows. Only one *CSN1S1*B/CSN1S1*C* cow did not share a *CSN1S1-5'* allele 1 but was typed 23. These genotypes were confirmed by repeated PCR and SSCP analysis. *CSN1S1-5'* allele 4 was not found in any cow having the *CSN1S1*C* allele. In Simmental cows, from the polymorphism screening sample set, genotype 11 was present with *CSN1S1*B/CSN1S1*C* and *CSN1S1*C/CSN1S1*C*. Thus, at least six intragenic haplotypes (1-B, 1-C, 2-B, 2-C, 3-B, 4-B) are present and exon 17 variant B was found in combination with all four different variants of the 5' flanking region.

Sequence Analyses

The sequences corresponding to the four SSCP detectable alleles were deposited in GenBank with accession numbers AF549499, AF549500, AF549501, and AF549502.

Compared to three other sequences of the 5' flanking region available in GenBank, we found best agreement of the sequence for allele 2 in our study and the sequence X03590. Both differ from AF435922 and X59856 in the length of the oligo-T stretch around position -230, which is (T)₉ in X59856, (T)₈ in our sequence type 2 and X03590 and only (T)₇ in AF435922. In X03590 one extra G (at -75) was present compared to alleles 1 to 4 and the two remaining sequences (AF435922, X59856), so this could be a sequencing error in that particular sequence. Allele 3 in our study is a GTTT deletion variant of allele 2, and the sequence of allele 4 bears one additional base substitution compared with allele 3. The sequence corresponding to allele 1 shows 5 base substitutions (3 A/G, 1 G/A, 1 A/C) and one (T)₂-deletion compared with allele 2 and represents the most divergent sequence in our study. The polymorphic *MaeIII* site in position -175 (Koczan et al., 1993) is disrupted here.

Searches for putative transcription factor binding sites affected by variable sequence positions revealed loss of an Octamer (OCT1) binding-site, a YinYang1 (YY1), and an activator protein (AP1) binding motif in

Table 1. Sizes of the paternal half-sib families, *CSN1S1-5'* genotype of the father, and distribution of *CSN1S1-5'* genotypes in the families.

Family ¹	n	Sire genotype	<i>CSN1S1-5'</i> Genotype									
			11	12	13	14	22	23	24	33	34	44
1	19	23	...	...	...	...	9	10	...	...	...	...
2	110	12	2	48	5	5	37	9	4	...	...	...
3	109	24	...	4	...	3	40	10	38	...	12	2
4	27	13	...	12	3	2	...	10	...	...	...	...
5	12	23	...	...	1	...	5	6	...	...	...	...
6	27	23	...	...	...	...	9	16	1	1	...	...
7	317	22	...	10	...	...	250	49	8	...	...	...
8	57	24	...	1	...	1	22	4	23	...	5	1
9	56	23	...	4	2	...	26	17	3	1	3	...
Total	734		2	79	11	11	398	131	77	2	20	3

¹Breed of families 1–8 = Holstein, family 9 = Simmental.

the sequence of allele 1. In allele 4, the substitution creates a potential ABF1 binding site close to the 18-bp 'milk box' and the proximal MGF/STAT5 binding motif (Schmid-Ney et al., 1991; Watson et al., 1991).

Linkage Mapping

With the BUILD option of CRI-MAP, two alternative marker orders resulted and *CSN1S1-5'* was either mapped between markers *ILSTS097* and *FBN14* or *FBN14* and *CSN3*. FLIPS 5 analysis then assigned *CSN1S1-5'* to the interval *FBN14* – *CSN3*, with the highest log10 likelihood. Recombination frequency for *CSN1S1/CSN3* was 0.06 and the calculated genetic distance was 5.6 cM. All markers and their positions are shown in Table 2 in comparison to the marker positions of the ADR2000 map without *CSN1S1* (Thomsen et al.,

2000), MARC97, and IBRP linkage maps (assessed via www.theark.db.org). Our new map is referred to as ADR+. The two-point analysis revealed recombination frequencies of 0.04 for *CSN1S1/FBN14* (lod score 29.56) and 0.07 for *CSN1S1/CSN3* (lod score 9.36), 0.10 for *CSN1S1/BP7* (lod score 17.86) and 0.13 for *FBN14/CSN3* (lod score 4.68).

Statistical Analysis

Variance analysis for all Holstein bulls from the granddaughter design with one of the four predominant genotypes resulted in highly significant effects of the family ($P < 0.0001$) on all traits investigated and significant ($P = 0.025$) effects of *CSN1S1-5'* genotype on EBV for protein percentage (BV_PP) (Table 3) as well as on the deregressed EBV protein percentage

Table 2. Markers used for linkage mapping on BTA6, number of informative meioses, and map position calculated with CRI-MAP (map named ADR+) compared to published linkage maps ADR2000, MARC97, and IBRP97.

Marker	Informative meioses	Map position (cM)			
		ADR+	ADR2000 ¹	MARC97	IBRP97
<i>ILSTS93</i>	477	0.0	0	0.0	16.0
<i>ILSTS90</i>	224	13.8	13	11.8	0.0
<i>URB16</i>	437	n.a. ²	39	n.a.	40.0
<i>BM1329</i>	302	44.9	42	35.5	45.0
<i>DIK82</i>	767	61.8	58	n.a.	67.0
<i>ILSTS097</i>	228	83.6	80	67.2	89.0
<i>FBN14</i>	256	89.7	88	n.a.	n.a.
<i>CSN1S1-5'</i>	299	93.9	n.a.	82.6 ³	103.0 ³
<i>CSN3</i>	260	99.5	97	82.6	103.0
<i>BP7</i>	409	109.7	107	91.2	n.a.
<i>BMC4203</i>	510	130.0	127.7	112.9	n.a.
Total length		130.0	127.7	125.6	160.0

¹The ADR2000 map is published in a graphical mode only, thus the marker positions are given approximately and without decimals.

²n.a. = position not assigned/marker not included in that map.

³*CSN1S1* is shown in the IBRP97 and MARC97 maps at the same position as *CSN3*, but was not genetically mapped independently.

Table 3. Effect of *CSN1S1-5'* genotypes on EBV and deregressed proofs for milk production traits. Significant *P* values are in bold letters, the highest LS means for these traits are underlined, and the magnitude of the substitution effect is given in units standard deviation (SD) below the *P* value in the contrast columns.

Variable	<i>P</i> for Effect of <i>CSN1S1</i>	LS Means in <i>CSN1S1-5'</i> genotype group ($\pm$ SE)						<i>P</i> for Contrast		
		SD	12	22	23	24	12 vs. 22	23 vs. 22	24 vs. 22	
BV_MY	0.165	375.870	555.520	379.681 $\pm$ 75.326	240.009 $\pm$ 43.649	310.927 $\pm$ 52.025	176.406 $\pm$ 75.460	0.082	0.234	0.408
BV_FY	0.287	6.710	19.060	5.726 $\pm$ 2.594	3.352 $\pm$ 1.503	7.187 $\pm$ 1.792	3.592 $\pm$ 2.599	0.391	0.062	0.928
BV_PY	0.159	8.500	14.970	9.059 $\pm$ 2.012	4.469 $\pm$ 1.166	6.338 $\pm$ 1.390	5.741 $\pm$ 2.016	0.033*	0.240	0.536
							0.307			
							diff. 12-22 units SD			
BV_FP	0.382	-0.105	0.259	-0.118 $\pm$ 0.036	-0.076 $\pm$ 0.021	-0.065 $\pm$ 0.025	-0.036 $\pm$ 0.036	0.271	0.681	0.263
BV_PP	0.025*	-0.049	0.111	-0.043 $\pm$ 0.015	-0.042 $\pm$ 0.009	-0.049 $\pm$ 0.010	0.001 $\pm$ 0.015	0.956	0.565	0.004**
										0.393
										diff. 24-22 units SD
DRG_MY	0.311	161.679	138.554	184.589 $\pm$ 19.696	162.309 $\pm$ 11.415	149.998 $\pm$ 13.603	139.118 $\pm$ 19.731	0.288	0.429	0.249
DRG_FY	0.405	6.148	6.348	6.858 $\pm$ 0.905	6.315 $\pm$ 0.524	5.629 $\pm$ 0.625	5.180 $\pm$ 0.906	0.573	0.338	0.219
DRG_PY	0.156	5.160	4.698	5.969 $\pm$ 0.671	5.058 $\pm$ 0.389	4.267 $\pm$ 0.464	4.601 $\pm$ 0.673	0.203	0.137	0.504
DRG_FP	0.459	-0.6*10 ⁻³	1.3*10 ⁻³	-0.63*10 ⁻³ $\pm$ 0.19*10 ⁻³	-0.42*10 ⁻³ $\pm$ 0.11*10 ⁻³	-0.36*10 ⁻³ $\pm$ 0.13*10 ⁻³	-0.24*10 ⁻³ $\pm$ 0.19*10 ⁻³	0.286	0.680	0.343
DRG_PP	0.028*	-0.3*10 ⁻³	0.6*10 ⁻³	-0.24*10 ⁻³ $\pm$ 0.08*10 ⁻³	-0.24*10 ⁻³ $\pm$ 0.05*10 ⁻³	-0.28*10 ⁻³ $\pm$ 0.05*10 ⁻³	0.01*10 ⁻³ $\pm$ 0.08*10 ⁻³	0.983	0.462	0.006**
										0.386
										diff. 24-22 units SD

P* < 0.05; *P* < 0.01; ****P* < 0.001.

(DRG_PP, *P* = 0.028). The same analysis including reliability of the breeding values as weighting factor resulted in nearly unchanged *P*-values, especially for the significantly affected traits (i.e., 0.0257 instead of 0.0254 for BV_PP; results not shown in detail). Highest LS-mean values for BV_PP and also DRG_PP were found in the group having *CSN1S1-5'* genotype 24 (Table 3).

Calculation of 24 vs. 22 contrasts revealed highly significant effects of *CSN1S1-5'* on protein content (DRG_PP and BV_PP, *P* = 0.006 and 0.004). Contrast 12 vs. 22 showed significant effects on BV_PY (*P* < 0.05), but not on DRG_PY (*P* = 0.203) and indicated an effect on BV_MY (0.05 < *P* < 0.1). For 23 vs. 22, an effect on breeding value for fat yield (BV_FY) approached significance (0.05 < *P* < 0.1), but was not detected when deregressed EBV were used in the analysis (DRG_FY, *P* = 0.338). Other contrasts were not significant. Probabilities for the contrasts are shown in Table 3. The difference in BV_PP between genotype 24 and 22 (indicating the allele substitution effect 2 $\rightarrow$ 4) was 0.39 standard deviations, which is 0.043 units of the breeding value.

DISCUSSION

All four *CSN1S1-5'* alleles that were found in the polymorphism screening samples from breeds selected for different production purposes were also present in German Holstein cattle. Previous examinations based on protein and coding region polymorphisms found a near fixation of *CSN1S1*B* in Holstein Friesian and other specialized milk breeds (Lin et al., 1986). In contrast to these findings, no loss of variability seems to be present in the 5' flanking *CSN1S1* gene region of German Holsteins. Moreover, no changes in allele frequencies over 8 yr (1990 to 1997) were detected. This would have been expected for a gene or region under selective pressure. This could indicate that milk yield, the main selection criterion, is probably not directly correlated to this polymorphism.

Using the SSCP generated *CSN1S1-5'* genotypes, we were able to perform linkage mapping of the *CSN1S1* 5' flanking region. This mapping revealed a 5.6 cM distance to *CSN3*, which is the most distant gene within the casein cluster. Previous assignments of *CSN1S1* in genetic linkage maps were not separated from *CSN3* or the casein cluster, mainly because of a lack of informative *CSN1S1* polymorphism. Based on the physical distance of about 200 to 250 kb (Ferretti et al., 1990; Threadgill and Womack, 1990; Rijnkels et al., 1997) a genetic distance of more than 5 cM is rather surprising. For human linkage maps, 1 cM is estimated to cover roughly 1 million base pairs (White et al., 1989) and

the same estimate results from the total size of the bovine genome in cM or base pairs (about 2900 cM and 3×10^9 base pairs; Kappes et al., 1997). Based on these assumptions, less than 0.5 cM would have been expected for the genetic distances within the casein cluster, which is exceeded 10-fold in our mapping results. Nevertheless, the distance in our linkage map agrees roughly with the estimated recombination frequency of 0.04 based on protein polymorphism based linkage studies about 20 yr ago (Hines et al., 1981).

A study comparing genetic and physical distances in the fungus *Neurospora crassa* revealed big differences in physical to genetic distance ratio depending on the number and position of markers considered (Bowring and Catcheside, 1999). Within a 650-kb contig, ratios of 30 kb/cM to 300 kb/cM depending on the marker choice and density applied were found, with an average ratio for the whole *Neurospora* genome of 40 kb/cM. DeWan et al. (2002) compared genetic and physical maps in humans and discussed several reasons for inconsistencies, including errors in physical maps, low marker heterozygosity, low number of informative meioses or closely linked markers. These authors conclude that it is not possible to interpolate between physical map distances and recombination frequencies due to possible hot spots of recombination and the unknown effects of sex and age of parent at conception on recombination rates.

Regarding the casein locus, three independent studies (Ferretti et al., 1990; Threadgill and Womack, 1990; Rijnkels et al., 1997) lead to the same gene order and more or less identical physical distances, thus errors in the physical map are unlikely. The ADR2000 linkage map was in accordance with published maps regarding length (intermediate between MARC97 and IBRP97) and marker order (Kappes et al., 1997, www.thearkdb.org). Moreover, the total length of the chromosome and length of the flanking marker interval *FBN14-CSN3* remain nearly unchanged after integration of *CSN1S1-5'* in the map compared to the data of Thomsen et al. (2000). The whole dataset for all markers was in agreement with the Mendelian laws of inheritance, thus the number of undetected typing errors is presumably low. Low marker heterozygosity and close linkage are a factor, that could influence mapping results, but as the number of informative meioses for *CSN1S1-5'* was higher than for several microsatellite markers, mapping results should be reliable. The high recombination rates detected now within the casein cluster, are most probably a consequence of differences in the marker choice. Earlier studies used coding region polymorphism, which is fairly low in Holstein cows. Thus, for these older linkage analyses, low marker heterozygosity could have masked the true recombination rate.

In our study, we found four different intragenic haplotypes with the predominant coding region allele *CSN1S1*B*, thus information suitable for linkage mapping increased substantially.

Sequence analysis of the 5' flanking region revealed two new variants (alleles 3 and 4) and two alleles most likely corresponding to the wild type (allele 2) and previously described sequence polymorphisms (allele 1). Haplotype combinations and the sequence analyses suggest, that type 1 corresponds to the type *c* as described by Koczan et al. (1993), even if a smaller fragment was studied there. The five different point mutations and the (T)₂ deletion present in allele 1 were as well described by Schild and Geldermann (1996). The promoter alleles 3 and 4 detected in Holstein cows now were not described in these earlier studies, which might be a consequence of the low number of animals investigated.

Comparing *CSN1S1-5'* and the polymorphism in exon 17 shows that the 5' flanking variant does not reliably predict the protein type and vice versa. This is in contrast to Koczan et al. (1993) and Ehrmann et al. (1997), who suggested one polymorphic site in the gene could serve as marker for other polymorphic regions in the same gene. A minimum of six intragenic haplotypes (*1-B*, *1-C*, *2-B*, *2-C*, *3-B*, *4-B*; additionally *3-C* could not be excluded from the genotype combination *23/BC*) are present in German Holsteins, whereas in other specialized milk breeds such as Angler and Ayrshire only a low number of different haplotypes was described (Jann et al., 2002). In the latter study, an RFLP-based typing and a smaller fragment were used, thus only a single variable position was resolved and alleles 3 and 4 could not be detected. Therefore, the number of haplotypes in these breeds needs to be determined by a more powerful test before postulating a loss of haplotypes or variability of the promoter region in consequence of milk yield selection.

The role of the *CSN1S1* promoter in casein expression implies effects on milk protein production in the udder. Therefore the *CSN1S1* promoter could be a functional candidate for milk protein content, regardless of position of known QTL regions. Variance analysis using a granddaughter design indicated a highly significant effect of *CSN1S1* promoter genotype on BV_PP (and DRG_PP), which supports the candidate locus hypothesis.

Genotype wise LS-means and contrast calculations each identified positive effects on protein percentage for genotype 24, which means, that allele 4 probably is the most positive allele for protein content. The highest LS-means for milk and protein yield (BV_MY and BV_PY) were found in the group with genotype 12. Contrast calculations for genotypes 12 and 22 indicated,

that there might be effects on milk yield and protein yield associated with genotype 12. The highest LS-mean for fat yield was found in genotype group 23, and contrast 23 vs. 22 showed a *P*-value near 5% significance level. Nevertheless, these findings for genotypes 12 and 23 were not supported by the variance analysis using deregressed EBV and therefore are probably influenced by single sires or families.

Quantitative trait loci for milk and fat yield close to the casein locus were postulated earlier (Bovenhuis and Weller, 1994; Velmala et al., 1999). However, because significant effects on yield traits were found only for single genotypes and were not detected in the variance analysis including all four genotypes and the analysis with deregressed EBV, a direct gene effect on yield traits seems less probable than an effect of a linked locus. This interpretation is in accordance with Zhang et al. (1998), who clearly discriminated QTL positions for milk yield and protein percentage on BTA 6 and is also supported by the two QTL hypothesis suggested by Spelman et al. (1996), Lipkin et al. (1998), Velmala et al. (1999) and Ron et al. (2001).

Allele substitution effects reach close to 40% of the overall standard deviation in breeding values for protein percentage when genotypes 22 and 24, representing the wild type and the most positive heterozygous genotype, respectively, are considered. The number of bulls with homozygous genotype 44 was too low to calculate effects for this group. It remains unclear whether the allele substitution effect for homozygous animals is in the same range.

With respect to potential selection benefit, the low frequency of allele 4 could be an advantage because potential gain following selection on favorable alleles is expected to be significant if this allele is present at low frequency in the starting population (Ron et al., 2001). The main selection criterion for production traits over the past decades has been milk yield and other yield traits. In Germany, from 1990 to 2002 only fat and protein yield were considered in the milk production index RZM. Therefore, the Holstein population used in our study can be regarded as more or less unselected for protein content. Recent developments and milk payment systems induced a moderate change in breeding objectives and now importance of protein content is increasing. Since August 2002 fat and protein content EBV are included in the German index RZM, because producing the same amount of fat and protein with more milk and lower contents is expensive due to the high energy consumption for (constant) lactose synthesis, which requires the same feed energy as for 3.4% protein (Rensing et al., 2002). Production of a reduced milk amount with higher fat and protein contents therefore seems economically reasonable.

Other QTL regions and positional candidate genes in the region around *BM143* on BTA6 have been recently identified by comparative mapping and bioinformatics. Some are related to bone formation, others are involved in cellular transport and regulation but have not been associated with lactogenesis (Ron et al., 2001).

Our results support the candidate locus hypothesis for the *CSN1S1* 5' flanking region, even if a final confirmation will require additional detailed studies of the mutations underlying the different SSCP patterns. Functional relevance of the mutations found in sequencing of the four alleles is suggested for allele 1, which lacks putative OCT1, AP1, and YY1 binding sites. YY1 is a multifunctional protein, which may activate or repress transcription. For *CSN2*, YY1 acts in the absence of lactogenic hormones as repressor of transcription and is counteracted by STAT5 (Meier and Groner, 1994; Rosen et al., 1999). The AP1 family of transcription factors is rather heterogeneous (Karin et al., 1997) and function in milk protein gene expression not yet proven. Expression studies in bovine mammary gland cell culture indicate prolactin-induced activation of a collagenase gene promoter containing an AP1 site (Olazabal et al., 2000), so prolactin sensitivity seems to be present. The meaning of the potential ABF1 binding site that is created by the mutation specific for allele 4, is not yet investigated for milk protein genes. This transcription factor and its binding motif seem to be involved in the process of chromatin remodelling and are important for transcription initiation (Bodmer-Glavas et al., 2001; Miyake et al., 2002). This is especially interesting in context with the recent hypothesis of casein expression regulation mechanisms at chromatin level of Rijnkels et al. (2003).

We hypothesize that the mutations identified in our study affect functional transcription factor binding sites and affect protein content directly through regulatory sequences. For a better causal understanding of allelic effects, further research to clarify functional impact of the four *CSN1S1*-5' variants and differences in putative transcription factor binding sites will be necessary.

CONCLUSIONS

The 5' flanking region of bovine *CSN1S1* is a highly polymorphic region and is therefore suitable for linkage analyses, association studies, or chromosome and genome wide QTL analyses. With a Holstein granddaughter design using EBV and de-regressed EBV for milk production traits, *CSN1S1*-5' genotypes proved highly significant for protein content. Sequence analysis of the different *CSN1S1*-5' alleles revealed variation within possible transcription factor binding sites. Together with the effects on protein content, this gives strong

support for the hypothesis of *CSN1S1-5'* as functional candidate locus for protein content in the bovine. *CSN1S1-5'* could be a useful marker to increase protein content in Holstein cattle.

ACKNOWLEDGMENTS

The authors thank Sandra Stein for skillful technical assistance in performing SSCP typings and plasmid cloning experiments. DNA from half-sib families was provided by the German cattle breeders association (Arbeitsgemeinschaft Deutscher Rinderzüchter, ADR) and cofunded by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF). Individual milk recording data of all cows and EBV for the bulls were kindly provided by Vereinigte Informationssysteme Tierhaltung (VIT).

REFERENCES

- Aleandri, R., L. G. Buttazoni, J. C. Schneider, A. Caroli, and R. Davoli 1990. The effects of milk protein polymorphisms on milk components and cheese-producing ability. *J. Dairy Sci.* 73:241–255.
- Ashwell, M. S., and C. P. van Tassel 1999. Detection of putative loci affecting milk, health, and type traits in a US Holstein population using 70 microsatellite markers in a genome scan. *J. Dairy Sci.* 82:2497–2502.
- Bassam, B. J., G. Caetano-Anollés and P. M. Gresshoff 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal. Biochem.* 196:80–83.
- Bodmer-Glavas, M., K. Edler, and A. Barberis. 2001. RNA polymerase II and III transcription factors can stimulate DNA replication by modifying origin chromatin structures. *Nucleic Acids Res.* 29:4570–4580.
- Bovenhuis, H., J. A. M. Van Arendonk, and S. Korver. 1992. Associations between milk protein polymorphisms and milk production traits. *J. Dairy Sci.* 75:2549–2559.
- Bovenhuis, H., and J. I. Weller 1994. Mapping and analysis of dairy cattle quantitative trait loci by maximum likelihood methodology using milk protein genes as genetic markers. *Genetics* 137:267–280.
- Bowring, F. J., and D. E. A. Catcheside. 1999. Recombinational landscape across a 650-kb contig on the right arm of linkage group V in *Neurospora crassa*. *Curr. Genetics* 36:270–274.
- DeWan, A. T., A. R. Parrado, T. C. Matise, and S. M. Leal. 2002. The map problem: A comparison of genetic and sequence-based physical maps. *Am. J. Hum. Genet.* 70:101–107.
- Ehrmann, S., H. Bartenschlager, and H. Geldermann. 1997. Quantification of single gene effects on milk proteins in selected groups of dairy cows. *J. Anim. Breed. Genet.* 114:121–132.
- Erhardt, G. 1993a. Allele frequencies of milk proteins in German Cattle breeds and demonstration of α_{s2} -casein variants by isoelectric focusing. *Arch. Tierz. Dummerstorf* 36:145–152.
- Erhardt, G. 1993b. A new α_{s1} -Casein allele in bovine milk and its occurrence in different breeds. *Anim. Genet.* 24:65–66.
- Ferretti, L., P. Leone, and V. Sgaramella. 1990. Long range restriction analysis of the bovine casein genes. *Nucleic Acids Res.* 18:6829–6833.
- Freyer, G., Z. Liu, L. Panicke, and G. Erhardt 1995. Untersuchungen der genetischen Kopplung zwischen möglichen QTL für Milchleistungseigenschaften und Milchproteinmarkern mittels markergestützter komplexer Segregationsanalyse. *Arch. Tierz. Dummerstorf* 38:379–393.
- Freyer, G., Z. Liu, G. Erhardt, and L. Panicke 1996. Schätzung von QTL-Effekten für Milchleistungseigenschaften und Untersuchung der genetischen Kopplung mit Caseinmarkern unter Anwendung der Multipointanalyse. *Arch. Tierz. Dummerstorf* 39:369–385.
- Freyer, G., Z. Liu, G. Erhardt, and L. Panicke 1999. Casein polymorphism and relation between milk production traits. *J. Anim. Breed. Genet.* 116:87–97.
- Freyer, G., C. Kühn, Q. Zhang, M. Mayer, and I. Hoeschele. 2002. Multiple QTL on chromosome six in dairy cattle affecting yield and content traits. *J. Anim. Breed. Genet.* 119:69–82.
- Green, P., K. Falls, and S. Crooks. 1990. Documentation of CRI-MAP, Version 2.4. Washington University School of Medicine, St. Louis, MO.
- Grosclaude, F., J. Pujolle, J. Garnier, and B. Ribadeau-Dumas. 1965. Déterminisme génétique des caséines κ du lait de vache; étroite liaison du locus κ -Cn avec les loci α -Cn et β -Cn. *C. R. Acad. Sci. (Paris)* 261:5229–5232.
- Hines, H. C., G. F. W. Haenlein, J. P. Zikakis, and H. C. Dickey. 1977. Blood antigen, serum protein and milk protein gene frequencies and genetic interrelationships in Holstein cattle. *J. Dairy Sci.* 60:1143–1151.
- Hines, H. C., J. P. Zikakis, G. F. W. Haenlein, C. A. Kiddy, and C. L. Trowbridge 1981. Linkage relationships among loci of polymorphisms in blood and milk of cattle. *J. Dairy Sci.* 64:71–76.
- Ikonen, T., M. Ojala, and O. Ruottinen. 1999. Associations between milk protein polymorphism and first lactation milk production traits in Finnish Ayrshire cows. *J. Dairy Sci.* 82:1026–1033.
- Jann, O., E.-M. Prinzenberg, H. Brandt, J. L. Williams, P. Ajmone-Marsan, P. Zaragoza, C. Özbeyaz, and G. Erhardt. 2002. Intragenic haplotypes at the bovine *CSN1S1* locus. *Arch. Tierz. Dummerstorf* 45:11–19.
- Kappes, S. M., J. W. Keele, R. T. Stone, R. A. McGraw, T. S. Sonstegard, T. P. L. Smith, N. L. Lopez-Corrales, C. W. Beattie. 1997. A second-generation linkage map of the bovine genome. *Genome Res.* 7:235–249.
- Karin, M., Z.-G. Lin, and E. Zandi. 1997. AP-1 function and regulation. *Curr. Opin. Cell Biol.* 9:240–246.
- Koczan, D., G. Hobom, and H.-M. Seyfert 1991. Genomic organization of the bovine alpha s1 casein gene. *Nucleic Acids Res.* 19:5591–5596.
- Koczan, D., G. Hobom, and H.-M. Seyfert 1993. Characterization of the bovine α_{s1} -casein gene C-allele based on a MaeIII-polymorphism. *Anim. Genet.* 24:74.
- Lien, S., and S. Rogne 1993. Bovine casein haplotypes: Number, frequencies and applicability as genetic markers. *Anim. Genet.* 24:373–376.
- Lien, S., J. Katanen, I. Olsaker, L.-E. Holm, E. Eythorsdottir, K. Sandberg, B. Dalsgaard, and S. Adalsteinsson. 1999. Comparison of milk protein allele frequencies in Nordic cattle breeds. *Anim. Genet.* 30:85–91.
- Lin, C. Y., A. J. McAllister, K. F. Ng-Kwai-Hang, and J. F. Hayes. 1986. Effects of milk protein loci on first lactation production in dairy cattle. *J. Dairy Sci.* 69:704–712.
- Lipkin, E., M. Mosig, A. Darvasi, E. Ezra, A. Shalom, A. Friedmann, and M. Soller 1998. Quantitative trait locus mapping in dairy cattle by means of selective milk DNA pooling using dinucleotide microsatellite markers: Analysis of milk protein percentage. *Genetics* 149:1557–1567.
- Meier, V. S., and B. Groner 1994. The nuclear factor YY1 participates in repression of the β -casein promoter in mammary epithelial cells and is counteracted by mammary gland factor during lactogenic hormone induction. *Mol. Cell. Biol.* 14:128–137.
- Miyake, T., C. M. Loch, and R. Li. 2002. Identification of a multifunctional domain in Autonomously Replicating Sequence Binding Factor 1 required for transcription activation, DNA replication and gene silencing. *Mol. Cell. Biol.* 22:505–516.
- Mosig, M. O., E. Lipkin, G. Khutoreskaya, E. Tchourzyna, E. Soller, and A. Friedmann 2001. A whole genome scan for quantitative trait loci affecting milk protein percentage in Israeli-Holstein cattle, by means of selective DNA pooling in a daughter design, using an adjusted false discovery rate criterion. *Genetics* 157:1683–1698.

- Nadesalingam, J., Y. Plante, and J. P. Gibson 2001. Detection of QTL for milk production on Chromosomes 1 and 6 of Holstein cattle. *Mamm. Genome* 12:27–31.
- Ng-Kwai-Hang, K. F., J. F. Hayes, J. E. Moxley, and H. G. Monardes 1984. Association of genetic variants of casein and milk serum proteins with milk, fat and protein production by dairy cattle. *J. Dairy Sci.* 67:835–840.
- Ng-Kwai-Hang, K. F., J. F. Hayes, J. E. Moxley, and H. G. Monardes 1986. Relationships between milk protein polymorphisms and major milk constituents in Holstein Friesian cows. *J. Dairy Sci.* 69:22–26.
- Ng-Kwai-Hang, K.-F., H. G. Monardes, and J. F. Hayes 1990. Association between genetic polymorphism of milk proteins and production traits during three lactations. *J. Dairy Sci.* 73:3414–3420.
- Nicholas, K. B., and H. B. Nicholas, Jr. 1997. GeneDoc: A tool for editing and annotating multiple sequence alignments. Distributed by the author. www.psc.edu/biomed/genedoc
- Olazabal, I., J. Munoz, S. Ogueta, E. Obregón, J. García-Ruz 2000. Prolactin (PRL)-PRL-Receptor System increases cell proliferation involving JNK (c-Jun Amino Terminal Kinase) and AP-1 Activation: Inhibition by glucocorticoids. *Mol. Endocrinol.* 14:564–575.
- Orita, M., Y. Suzuki, T. Sekiya, and K. Hayashi 1989. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* 5:874–879.
- Ott, J. 1991. Analysis of Human Genetic Linkage. rev. ed. Johns Hopkins University Press, Baltimore.
- Rensing, S., E. Pasman, F. Reinhardt, and F. Feddersen 2002. New total merit index RZG for Holsteins in Germany with more emphasis on herd life. *Interbull Bull.* 29:147–149.
- Rijnkels, M., P. M. Kooiman, P. J. A. Krimpenfort, H. A. de Boer, and F. R. Pieper 1995. Expression analysis of the individual bovine β -, α s2- and κ -casein genes in transgenic mice. *Biochem. J.* 311:929–937.
- Rijnkels, M., P. M. Kooiman, H. A. de Boer, and F. R. Pieper. 1997. Organization of the bovine casein gene locus. *Mamm. Genome* 8:148–152.
- Rijnkels, M., L. Elnitskij, W. Miller, and J. Rosen. 2003. Multi species comparative sequence analysis of the casein gene cluster region. *Genomics* (In press).
- Ron, M., D. Kliger, E. Feldmesser, E. Seroussi, E. Ezra, and J. I. Weller 2001. Multiple quantitative trait locus analysis of bovine chromosome 6 in the Israeli Holstein population by a daughter design. *Genetics* 159:727–735.
- Rosen, J. M., S. L. Wyszomierski, D. Hadsell, D. 1999. Regulation of milk protein gene expression. *Annu. Rev. Nutr.* 19:407–436.
- SAS/STAT User's Guide, Version 6, Edition 1989. SAS Institute Inc., Cary, NC.
- Schild, T. A., and H. Geldermann. 1996. Variants within the 5'-flanking regions of bovine milk protein encoding genes. III genes encoding the Ca-sensitive caseins α s1, α s2 and β . *Theor. Appl. Genet.* 93:887–893.
- Spelman, R. J., W. Coppieters, L. Karim, J. A. M. van Arendonk, and H. Bovenhuis. 1996. Quantitative trait loci analysis for five milk production traits on chromosome six in the Dutch Holstein Friesian population. *Genetics* 144:1799–1808.
- Thomsen, H., N. Reinsch, N. Xu, C. Looft, S. Grupe, C. Kühn, G. A. Brockmann, M. Schwerin, B. Leyhe-Horn, S. Hiendleder, G. Erhardt, I. Medjugorac, I. Russ, M. Förster, B. Brenig, F. Reinhardt, R. Reents, J. Blümel, G. Averdunk, and E. Kalm. 2000. A male bovine linkage map for the ADR granddaughter design. *J. Anim. Breed. Genet.* 117:289–306.
- Thomsen, H., N. Reinsch, N. Xu, C. Looft, S. Grupe, C. Kühn, G. A. Brockmann, M. Schwerin, B. Leyhe-Horn, S. Hiendleder, G. Erhardt, I. Medjugorac, I. Russ, M. Förster, B. Brenig, F. Reinhardt, R. Reents, J. Blümel, G. Averdunk, and E. Kalm. 2001. Comparison of estimated breeding values, daughter yield deviations and de-regressed proofs within a whole genome scan for QTL. *J. Anim. Breed. Genet.* 118:357–370.
- Threadgill, D. W., and J. E. Womack. 1990. Genomic analysis of the major bovine milk protein genes. *Nucleic Acids Res.* 18:6935–6942.
- Velmalä, R. J., H. J. Vilkky, K. T. Elo, D. J. deKoning, and A. V. Mäki-Tanila. 1999. A search for quantitative trait loci for milk production traits on chromosome 6 in Finnish dairy cattle. *Anim. Genet.* 30:136–143.
- White, R., J. M. Lalouel, M. Leppert, M. Lathrop, Y. Nakamura, and P. O'Connell. 1989. Linkage mapping of human chromosomes. *Genome* 31:1066–1072.
- Wiener, P., I. Maclean, J. L. Williams, and J. A. Wooliams. 2000. Testing for the presence of previously unidentified QTL for milk production traits in new populations. *Anim. Genet.* 31:385–395.
- Wingender, E., X. Chen, R. Hehl, H. Karas, I. Liebich, V. Matys, T. Meinhardt, M. Prüß, I. Reuter, and F. Schacherer. 2000. TRANSFAC: An integrated system for gene expression regulation. *Nucleic Acids Res.* 28:316–319.
- 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.