J. Dairy Sci. 91:767–783
doi:10.3168/jds.2007-0507

Quantitative Trait Loci Affecting Milk Yield and Protein Percentage in a Three-Country Brown Swiss Population

A. Bagnato,*[1] F. Schiavini,* A. Rossoni,*† C. Maltecca,*‡ M. Dolezal,§ I. Medugorac,# J. Sölkner,§ V. Russo,‖ L. Fontanesi,‖ A. Friedmann,¶ M. Soller,¶ and E. Lipkin¶

*Department of Veterinary Sciences and Technologies for Food Safety, Università degli Studi di Milano, 20133 Milano, Italy
†Italian Brown Breeders Association, Loc. Ferlina 204, Bussolengo 37012, Italy
‡Dairy Science Department, University of Wisconsin, 1675 Observatory Dr., Madison
§Department of Sustainable Agricultural Systems, BOKU—University of Natural Resources and Applied Life Sciences, Vienna, A-1180 Vienna, Austria
#Institute for Animal Breeding, Ludwig-Maximilians University, 80539 Munich, Germany
‖Dipartimento di Protezione e Valorizzazione Agroalimentare, Università di Bologna, 40010 Bologna, Italy
¶Department of Genetics, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel

ABSTRACT

Quantitative trait loci (QTL) mapping projects have been implemented mainly in the Holstein dairy cattle breed for several traits. The aim of this study is to map QTL for milk yield (MY) and milk protein percent (PP) in the Brown Swiss cattle populations of Austria, Germany, and Italy, considered in this study as a single population. A selective DNA pooling approach using milk samples was applied to map QTL in 10 paternal half-sib daughter families with offspring spanning from 1,000 to 3,600 individuals per family. Three families were sampled in Germany, 3 in Italy, 1 in Austria and 3 jointly in Austria and Italy. The pools comprised the 200 highest and 200 lowest performing daughters, ranked by dam-corrected estimated breeding value for each sire-trait combination. For each tail, 2 independent pools, each of 100 randomly chosen daughters, were constructed. Sire marker allele frequencies were obtained by densitometry and shadow correction analyses of 172 genome-wide allocated autosomal markers. Particular emphasis was placed on *Bos taurus* chromosomes 3, 6, 14, and 20. Marker association for MY and PP with a 10% false discovery rate resulted in nominal *P*-values of 0.071 and 0.073 for MY and PP, respectively. Sire marker association tested at a 20% false discovery rate (within significant markers) yielded nominal *P*-values of 0.031 and 0.036 for MY and PP, respectively. There were a total of 36 significant markers for MY, 33 for PP, and 24 for both traits; 75 markers were not significant for any of the traits. Of the 43 QTL regions found in the present study, 10 affected PP only, 8 affected MY only, and 25 affected MY and PP. Remarkably, all 8 QTL regions that affected only MY in the Brown Swiss, also affected MY in research reported in 3 Web-based QTL maps used for comparison with the findings of this study (http://www.vetsci.usyd.edu.au/reprogen/QTL_Map/; http://www.animalgenome.org/QTLdb/cattle.html; http://bovineqtl.tamu.edu/). Similarly, all 10 QTL regions in the Brown Swiss that affected PP only, affected only PP in the databases. Thus, many QTL appear to be common to Brown Swiss and other breeds in the databases (mainly Holstein), and an appreciable fraction of QTL appears to affect MY or PP primarily or exclusively, with little or no effect on the other trait. Although QTL information available today in the Brown Swiss population can be utilized only in a within family marker-assisted selection approach, knowledge of QTL segregating in the whole population should boost gene identification and ultimately the implementation and efficiency of an individual genomic program.

Key words: quantitative trait loci mapping, Brown Swiss, productive trait, selective DNA pooling

INTRODUCTION

Recent developments in molecular genetics have paved the way for the integration of genomic information in livestock breeding (Dekkers, 2004). The minimum requirement for such a task is a map locating the QTL affecting economically important traits to specific chromosomal regions.

Numerous studies (e.g., de Koning et al., 2001; Mosig et al., 2001; Thomsen et al., 2001; Boichard et al., 2003) have found a large number of QTL affecting productive and functional traits in dairy cattle. Khatkhar et al. (2004), in a meta-analysis, considered 45 studies in dairy cattle populations and provided an overview of QTL reported in literature. To the best of our knowl-

Received July 6, 2007.
Accepted October 5, 2007.
[1]Corresponding author: alessandro.bagnato@unimi.it

edge, with the exception of 2 autochthonous French breeds (Boichard et al., 2003), and the Ayrshire breed (Velmala et al., 1999; de Koning et al., 2001), all other QTL mapping studies in dairy cattle have been implemented in the Holstein population. This is probably due to the need for large families, when using either a daughter (**DD**) or a granddaughter designs (**GDD**; Weller et al., 1990). Because QTL mapping studies are traditionally realized on a national basis, this consideration limits QTL mapping studies in the minor breeds due to the difficulty of finding large enough paternal half-sib families. For cosmopolitan minor breeds, however, such as Brown Swiss, Jersey, and Guernsey, the number and size of families on a multinational basis are sufficient to support such studies and in this way provide a basis for genomic approaches to selection in these breeds as well. Carrying out a multinational QTL mapping study based on the GDD should be quite feasible, using semen as source of DNA, and Interbull EBV (http://www-interbull.slu.se/) to bring all sires to a specific country common denominator. On the other hand, a multinational DD for daughters spread over a number of countries would be more difficult with respect to the need to adjust national daughter EBV to a common denominator. To convert different national EBV to a common basis, the best option is to use linear regression conversion formulas method used for several years to provide sire international proofs (Wilmink et al., 1986). This method was abandoned for sires when MACE (multiple across country evaluation) (Schaeffer, 1994) became available. Several studies tried to construct a common denominator genetic evaluation for dams as well, but these attempts have been largely unsuccessful (Canavesi et al., 2001). At present, the only way to obtain consistent EBV for an international female population would be to perform a borderless genetic evaluation beginning with phenotypes (Maltecca et al., 2004). This is the procedure followed by Austria and Germany for the joint genetic evaluation of the Brown Swiss breed across the 2 countries (Emmerling et al., 2002). In the recent past the breeding goals in the 3 countries were similar and sires used in the 3 populations were shared among countries so to permit here the consideration of individuals from Austria, Germany, and Italy as a single European Brown Swiss population.

A selective DNA pooling approach using milk as a source of DNA (Lipkin et al., 1993, 1998; Mosig et al., 2001) was applied in the present study. The use of selective DNA pooling in a daughter design (Darvasi and Soller, 1994) allows an efficient initial screening of marker-QTL association. In addition, because in selective DNA pooling QTL association is based on allele frequency estimation in the pools made of extreme individuals, these pools can readily be constructed without the need of international EBV for the daughters of each sire family, even if sampled in 2 or more countries. In the present study, the use of milk somatic cells as a source of DNA (Lipkin et al., 1993, 1998) allowed collection of a large number of samples through the routine milk recording testing performed by the Brown Swiss breed and milk recording associations in all 3 countries.

Table 1. Total number of milk samples collected by sire and country

Sire	Austria	Germany	Italy	Total
A	943		2,158	3,101
B	841		1,413	2,254
C			1,782	1,782
D	609		646	1,255
E			2,568	2,568
F			947	947
H		1,984		1,984
I		2,192		2,192
J		1,431		1,431
K	1,252			1,252
			Total	18,766

The objective of this work was to map QTL for milk yield (**MY**) and milk protein percent (**PP**) in the European Brown Swiss breed, utilizing a selective DNA pooling approach using milk as source of DNA and sampling individuals from Austrian, German, and Italian populations. Although Brown Swiss population size in each country is relatively small compared with the respective Holstein population, sampling jointly from all 3 countries provided a sufficient number of sire families and daughters for a comprehensive mapping study. To the best of our knowledge this represents the first QTL mapping study reported for this breed, and we believe it will serve as a paradigmatic example for other minor cosmopolitan breeds.

MATERIALS AND METHODS

Sampling of Families

International sampling of 10 large half-sib families (milk from at least 1,000 daughters per family) was performed across the Austrian, German, and Italian Brown Swiss populations. Semen samples of sires were either commercially available or obtained from the Italian Brown Swiss Breeders Association semen bank. Two of the bulls (C and D) were sire and son, and 4 of the remaining bulls (A, F, H, J) had a common grandsire. One sire (K) was sampled in Austria only, 3 families (H, I, J) in Germany only, 3 families (C, E, F) in Italy only, and 3 families (A, B, D) across Italy and Austria. About 40% of targeted daughters could not be sampled because they were dry or had been culled or sold. Table 1 shows the number of daughters sampled by sire and country. On average 1,877 daughters were sampled per sire (range: 947 to 3,101).

In addition to each official milk sample taken during the routine milk recording test, a second sample was collected in 10-mL tubes prepared with an aliquot of milk preservative and labeled with the official breed identification barcode number of the animal. The collected milk samples, along with the official ones, were sent to the central milk testing laboratory and immediately frozen at –20°C.

Information on SCC for each official sample was made available through the milk recording agencies. In some instances a reported SCC for a particular sample was missing or the date recorded in the breed association database did not match the sampling date. In these cases, SCC was obtained directly from the collected sample. As a quality assurance, SCC values for a random cross section of 100 samples were verified by comparing the SCC information from the milk recording agencies with those obtained directly from the collected samples. The maximum observed difference was 6%, which is below the 7% tolerance accepted among instruments in official testing. Additionally, a pilot study was conducted to compare SCC obtained from fresh and frozen samples: for fresh samples SCC was recorded after sampling; for frozen-unfrozen samples SCC was measured weekly for up to 6 wk. In the frozen-unfrozen collected samples, SCC had a maximum variation of 5% with respect to the cell count reported for the original fresh sample (data not shown).

Pool Constitution

For MY and PP EBV were available for each daughter according to the national genetic evaluation procedure in Italy (Interbull, 2000) and in Austria and Germany (Emmerling et al., 2002). Daughter EBV were corrected (**cEBV**) by subtracting half of the dam's EBV to remove the dam effect on daughter EBV and optimize the ranking of high and low individuals on their sire's genetic contribution (Dolezal, 2007). The highest and lowest 200 individuals of the cEBV distribution (high and low "tails", respectively) for each sire-family-by-trait combination were identified for inclusion in the pools. Each high or low tail group was divided into 2 independent subpools each consisting of 100 randomly selected daughters (henceforth, replicate subpools), by assigning animals with even rank position to subpool-1, and those with odd rank position to subpool-2. In addition each of the replicate subpools was duplicated from the same samples in a separate independent pooling procedure, performed by different operators on a different day (henceforth, duplicate subpools). Thus, a total of 8 subpools (2 tails [high, low] × 2 replicate subpools per tail × 2 duplicate subpools per replicate subpool) were constructed for each sire-by-trait combination. With 10 sires and 2 traits, there were a total of 20 sire-by-trait combinations, giving a total of 160 subpools. Pool construction for sires sampled across Italy and Austria was proportional to the number of daughters collected in each country. This was done to control for possible differences in genetic variability and average genetic level among the different populations.

Sample Preparation

Based on the SCC, a volume of milk to attain the desired numbers of cells (20,000 per individual) was included in each subpool. The subpool was then divided into 10 pool-fractions to decrease possible risk of contamination or loss of an entire subpool due to mishandling. Each pool-fraction contained 2,000 cells from each of the 100 daughters included to guarantee approximately equal genomic contribution from each individual to the PCR. Following Lipkin et al. (1993, 1998), cell lysate was used in the PCR reaction instead of purified DNA. An aliquot of milk was diluted with saline (NaCl 0.9%) and centrifuged at 685 × *g* for 5 min to pellet somatic cells. Cells were resuspended in saline and recentrifuged (2 to 15 times) until a clear pellet of milk somatic cells was obtained. Somatic cells were resuspended in Tris and EDTA to a volume of 50 μL (a final concentration of 4,000 cells per μL) and incubated in thermal cycler (MJ Research PTC-200, GMI Inc., Ramsey, MN) at 100°C for 5 min, 50°C for 5 min, and again at 100°C for 5 min to lyse the cells. A commercial kit (http://www.spin.it/talent/genomix.html) was utilized to obtain DNA from semen.

Markers and Genotyping

A panel of 187 dinucleotide microsatellite markers (www.marc.usda.gov/genome/genome.html) covering the entire autosomal genome was chosen. Primers were labeled with fluorescent dye (FAM, HEX, TET) according to the standard (TAMRA350, Applied Biosystems Foster City, CA) used for fragment analysis on the ABI377 sequencer (Applied Biosystems).

The PCR reaction was optimized for each fluorescent primer starting from a standard protocol consisting of 1 μL of PCR buffer with $MgCl_2$ (15 m*M*), 0.2 μL of dNTP (10 m*M* each), 0.3 μL of each primer (10 μ*M* each), 2 μL of 5X *Taq*Master PCR Enhancer (Eppendorf AG, Hamburg, Germany) and 0.5 U of *Taq* DNA polymerase (Eppendorf AG). One microliter of semen DNA 1:10 diluted or 1 μL of cell lysate (to provide a DNA average contribution of 4,000 cells per subpool) was used in each PCR reaction. Water was added to a final volume of 10 μL. Amplification was performed in MJ Research PTC-200 thermal cycler (GMI Inc.), programming 1 step of

10 min at 95°C, 35 cycles comprised of 95°C for 1 min, 54°C to 64°C for 1 min (according to specific primer annealing temperature) and 72°C for 1 min; a final step of 45 min at 72°C completed the amplification process to eliminate the half-step bands.

One-half microliter of each PCR product was added to a mix of TAMRA 350 (Applied Biosystem), blue-dextran, and formamide and separated by electrophoresis on an ABI377 automatic sequencer of Applied Biosystems using a 5% acrylamide 0.2-mm gel. Densitometric values of sires and pool genotypes (peak height) of each detectable fragment were obtained using Genescan and Genotyper softwares (Applied Biosystems). Among the initial 187 markers, 3 were homozygous in all sires, 10 did not amplify properly, and 2 systematically produced by-products that affected densitometric values. This left a total of 172 markers that were used in the genome scan, yielding an average spacing of about 20 cM, except for BTA 3, 6, 14, and 20 where a denser map (about 2 to 5 cM spacing) was used: for chromosome 6, because of the strongly confirmed presence of QTL on that chromosome (reviewed by Khatkar et al., 2004; Polineni et al., 2006; Hu et al., 2007); for chromosomes 3, 14, and 20, because they were found significant across all populations studied in the BovMAS-EU-funded project (of which this is a part) and were targeted for more detailed study. For the same reason, to confirm significance in these 3 chromosomes, individual genotyping was carried out for a set of 42 markers each with a sample of 96 individuals per tail, and results provided a correlation larger then 0.95 among pools and individual frequencies (results not shown). Further confirmation of the effectiveness of the selective DNA pooling approach of this study comes from the consistency of trait-specific results obtained in the present study and those reported in 3 Web-based cattle QTL databases (see below).

The 10 sires were genotyped for each marker, and only markers that were heterozygous in the sires were subsequently genotyped in the pools of the sire. In the initial scan, one of the duplicates of each of the 4 subpools per sire-trait combination was genotyped. Following the initial analysis in suggestive QTL regions, the subpools were typed again with their duplicates to confirm the putative association, by an independent set of reactions.

Estimating Frequency of Sire Alleles in the Pools

To obtain frequencies of sire marker alleles in the pools, densitometric values were corrected for shadow bands as in Lipkin et al. (1998). The relative intensity of a shadow band with respect to the main band is a linear function of the number of repeats of the dinucleotide in the microsatellites. The number of dinucleotide repeats was estimated from the genotypes of the sires of the 10 studied families, plus 1 additional Brown sire and 10 Holstein sires, according to the methodology described in Lipkin et al. (1998). Only densitometric values of homozygous individuals or of heterozygous individuals for which alleles were separated by at least 4 repeats (8 bp) were used in repeat number estimation. Holstein sires were considered only when fewer than 2 Brown sires were available for repeat number estimation. In these cases additional Brown Swiss females were later genotyped to confirm results and to test for breed difference in allele repeat number calculation. For alleles for which suitable individuals were not available, interpolation was used.

Pool densitometry and estimation of frequency of long (l) and short (s) sire alleles were performed separately for each of the 2 subpools in each tail (see Appendix for details). Consequently, denoting the 4 subpools: H1 and H2, L1 and L2, respectively (**H** = high tail, **L** = low tail), a total of 8 allele frequency estimates were obtained for each sire-marker-trait combination, denoted H1l, H2l, H1s, H2s, L1l, L2l, L1s, L2s. It should be emphasized that there is no intrinsic relationship between the 2 replicate-subpools in the H tail and the 2 replicate-subpools in the L tail, and the assignment of indices 1, 2 within tails was arbitrary.

Frequency estimates for long and short alleles in replicate-subpools in the same tails were compared and their variance calculated. If the calculated variance within sire-trait-marker-replicate-subpool-allele was greater than 0.012 (about 4 times the observed empirical variance between replicate-subpools in the same tail; see appendix for definition of empirical variance), the test was tagged as an outlier, and PCR amplification and gel run were performed again on the same samples and on the duplicates of the samples. If variance among the duplicates was reduced below the above threshold, the new values were used. If variance remained high, the frequency estimates from the individual subpools were compared with those from other subpools (replicates and duplicates) of the same sire-marker-trait-tail combination, in an attempt to identify the outlier subpool. When this was successful, the test was recalculated excluding the data from the outlier subpool. When it was not possible to clearly identify a specific outlier subpool, the entire test was discarded. The outlier subpools were examined with respect to sire, trait, and tail to identify subpools that were problematic in multiple instances. In this case, the entire subpool was eliminated from all subsequent calculations, even when individual data points of that subpool were not aberrant.

Marker-QTL Linkage Tests

The test for a linkage between a marker and each of the 2 studied traits was carried out at 2 levels: i) the

sire-marker level (within sire analysis), which tests for association of marker and trait at the level of the individual sire; and ii) the marker level (across sires analysis), which tests for association of marker and trait across all sires heterozygous at a marker.

Individual Sire-Marker-Trait Combinations. For each sire-marker-trait combination for which the sire was heterozygous at the marker, the comparisonwise error rate (**CWER**), P_{ijk}, for the ith sire-jth marker-kth trait combination was obtained as twice the area of the normal curve from $|Z_{ijk}|$ to $+\infty$, with

$$Z_{ijk} = D_{ijk}/SE(D'),$$

where D_{ijk} represents the difference in sire-allele frequencies between the high and low daughter pools of the ith sire with respect to the jth marker and the kth trait, and SE(D′) is the empirical standard error of D_{ijk}. The D_{ijk} and Se(D′) were estimated as described in Appendix 1.

Individual Marker-Trait Combinations. The CWER, P_{jk}, for the jth marker and kth trait combination was obtained as the area of the χ^2 distribution from χ^2_{jk} to $+\infty$, with

$$\chi^2_{jk} = \sum_{i=1}^{s_{jk}} Z^2_{ijk} \text{ with df} = s_{jk},$$

where s_{jk} is the number of heterozygous sires tested for the jth marker, kth trait combination (Weller et al., 1990, Lipkin et al., 1998; Mosig et al., 2001).

Experimentwise Linkage Tests Proportion of False Positives. To overcome the multiple-test situation but retain power, the proportion of false positive (**PFP**) criterion (Mosig et al., 2001; Fernando et al., 2004) was used to set threshold CWER levels for declaration of significance at the sire-marker-trait and marker-trait levels. To calculate the PFP, it is necessary to first estimate the number of tests for which the null hypothesis is true (denoted n_2) because, by definition, a false positive can come only from among this set of tests. Following Mosig et al. (2001), n_2 was estimated by iteration from the frequency distribution of the obtained P-values. This was done separately for marker level tests (n_{2M}) or for sire-marker level tests (n_{2S}). Using these values for n_2, PFP values were calculated separately for marker level tests (**PFP_M**) and for sire-marker level tests (**PFPs**) as follows:

$$PFP_{Mi} = (P_{Mi} \times n_{2M}) / R_{Mi},$$

where P_{Mi} is the CWER P-value of the ith marker test, R_{Mi} is the rank number of the ith marker test, when the marker tests are ranked by their P-values from lowest to highest; and similarly,

$$PFP_{Si} = (P_{Si} \times n_{2S}) / R_{Si},$$

substituting the corresponding expressions for sire-marker tests in place of marker tests.

Significance Thresholds. For each trait, marker-tests having CWER P-values corresponding to $PFP_M \leq 0.10$ were considered to be significant. With respect to sire-marker tests, PFP values were calculated for all sire-marker-trait tests, whether the marker was significant. However, sire-marker tests were considered significant only if they had CWER P-values corresponding to PFP ≤ 0.20 and were also within significant markers. The less stringent threshold for the sire-marker level reflects an attempt to rebalance type I and type II error, taking into consideration the weaker power curve of the individual sire × marker tests compared with the marker tests. The less stringent threshold also takes into account the prior Bayesian probability of about 0.40 that a sire will be heterozygous at the QTL, given a significant marker.

Estimating Power of the Test

Because the PFP analysis for each trait provides an estimate of the number of tests involving that trait for which the null hypothesis is true (n_{2M} or n_{2S}), the total number of tests for which the null hypothesis is falsified [i.e., the number of tests for the trait representing true marker-QTL association (n_{1M}) or sire-marker-QTL (n_{1S}) association], can be estimated as

$$n_{1M} = N_M - n_{2M}$$

and

$$n_{1S} = N_S - n_{2S},$$

where N_M and N_S are the total number of marker or sire-marker tests, respectively, for each trait. On this basis, adapting the notation of Mosig et al. (2001) to the notation of the present paper, the power of the test for a particular trait can be calculated as

$$P_M = o_M (1 - PFP_M)/n_{1M}$$

for the marker level of analysis, and as

$$P_S = o_S (1 - PFP_S) / n_{1S}$$

for the sire-marker level of analysis, where o_M and o_S are the observed number of effects declared significant

at the marker and sire-marker levels, respectively; and $(1 - PFP_M)$ and $(1 - PFP_S)$ are the proportion of true significant effects among all declared significant effects for marker and sire-marker tests, respectively; so that $o_M(1 - PFP_M)$ and $o_S(1 - PFP_S)$ are the observed number of declared significant effects at the marker and sire-marker levels, respectively, corrected for the proportion of true positives.

Allele Substitution Effects

Allele substitution effects were calculated as described in Lipkin et al. (1998) and Mosig et al. (2001), based on shadow-corrected allele frequency estimates. The population standard deviation of EBV needed for this calculation was estimated from the difference in the mean EBV of the individuals in the high and low pools, as described in Appendix 2. Allele substitution effects were calculated only for sire-marker combinations defined as significant following the above PFP criterion.

Identifying QTL Regions

An attempt was made to define chromosomal regions harboring QTL (**QTLR**) and chromosomal regions that did not harbor QTL according to the results of the genome scan to highlight them over the entire genome. All markers and sire combinations were considered in QTLR definition, thus adding novel information with respect to significant markers only. This information allows exclusion of chromosomal regions from further QTL search (e.g., to refine QTL location) in the same breed for MY and PP. Given that interval mapping methods that would permit analysis of multiple QTL on the same chromosome with selective DNA pooling are still under development (Dekkers et al., 1999; Dolezal, 2007), to identify multiple QTL on the same chromosome we used the qualitative approach described by Mosig et al. (2001). This approach is based on the analysis of all sires by marker combinations (significant and not significant) on the chromosome. It identifies as multiple well-separated QTL on a chromosome an interval where one or more sires showed, with respect to a marker, significance for more proximal markers, with clear lack of significance for more distal markers, and the opposite was true for other sires (i.e., nonsignificant at proximal but significant at distal). On the other hand it also identified as multiple separated QTL on a chromosome an interval where one or more sires showed significance at proximal and distal markers but not at the intervening markers, and the opposite was true for the other sires (as above). We identified QTL for MY and PP separately to define trait specific-QTL regions and then combined the 2 analyses defining as a unique chromosomal region across the 2 traits those where both traits showed significance. Some of the defined regions include more than 1 flanking marker indicating a chromosomal span, whereas others are based on a single marker.

Definition of QTLR and non-QTLR on the same chromosome was based exclusively on the results of the present study and did not include a comparative approach with literature as this would require a meta-analysis as developed by Khatkar et al. (2004), which is beyond the scope of this paper.

The QTLR here defined were compared with 3 Web-based QTL maps (http://www.vetsci.usyd.edu.au/reprogen/QTL_Map/; http://www.animalgenome.org/QTLdb/cattle.html; http://bovineqtl.tamu.edu/) reporting results of published literature. The first Web QTL map is based on the meta-analysis of Khatkar et al. (2004). The second (Hu et al., 2007) and the third (Polineni et al., 2006) collect and update all publicly available QTL reported for a large variety of traits.

Comparison with Web-based QTL maps (i.e., literature) is not straightforward because the approaches used to map QTL are different, and information on QTL description is quite variable (e.g., some QTL are reported as a point location based on a single marker, some are point locations presented with a confidence interval, others do not provide a point location but only suggest a possible QTL-containing region, and so on). In any event we indicated the corresponding QTLR as previously reported if in the databases a QTL location is reported (i) within plus or minus 5 cM from our QTLR defined as an interval, or (ii) within plus or minus 10 cM of our QTLR defined by a single marker.

RESULTS

Descriptive Statistics

Based on the genotypes of the 10 Brown Swiss sires (plus 1 additional sire) across all markers, average marker heterozygosity and polymorphic information content were 0.61 and 0.56, respectively. On average, the difference between the H and L tails among all sires was of 552.5 kg for MY and 0.15% for PP (Table 2). The overall proportion selected to each tail was 0.12, which corresponds to an average selection intensity of 1.67. Thus, we estimate the within-sire standard deviation of cEBV values as 165.6 kg for MY and 0.046% for PP, which yield estimates for the standard deviation of EBV(MY) and EBV(PP) equal 190.3 kg and 0.053%, respectively.

Table 2. The proportion selected from each tail, mean EBV values on the Italian scale in each of the tails (low and high), and difference between the means of the 2 tails (Diff.) for the 2 traits considered

Sire family	Percent selected	Milk yield			Protein percent		
		Low	High	Diff.	Low	High	Diff.
A	0.06	87.4	609.7	522.3	–0.14	0.08	0.22
B	0.09	54.2	725.8	671.6	–0.16	0.02	0.18
C	0.11	–119.6	549.4	669.0	0.00	0.29	0.29
D	0.16	65.0	617.9	552.8	–0.06	0.08	0.14
E	0.08	86.1	712.5	626.4	0.05	0.33	0.28
F	0.21	–125.1	263.2	388.3	–0.02	0.13	0.15
H	0.10	159.5	706.5	547.1	0.02	0.09	0.07
I	0.09	–67.9	502.1	570.0	0.07	0.14	0.07
J	0.14	–110.2	355.1	465.3	0.06	0.13	0.07
K	0.15	–170.5	341.9	512.3	–0.02	0.06	0.08
Overall	0.12	–14.1	538.4	552.5	–0.02	0.13	0.15

Genome Scan

Empirical SE(D′) and Identification of Outlier Pools. The empirical SE(D′) calculated on the basis of 976 sire-marker tests for MY and 970 sire-marker tests for PP was equal to 0.059. Using this value, pools were identified as outliers if the variance among frequency estimates for subpools in the same tails was ≥0.012. For sire B, one of the high PP subpools and one of the low MY subpools were consistently aberrant and therefore were eliminated from the subsequent analyses. For sire A, more than 50% of the tests were outliers, yet specific aberrant pools were not identified. Nevertheless, sire A was also excluded from subsequent analyses. On average, for the remaining 9 sires, including sire B, less than 15% of the tests were discarded as outliers. After the editing, among the 976 sire-marker tests initially available for MY and 970 for PP, 776 and 762, respectively, were included in the analyses for marker QTL association. It is important to note the effective manner in which the replicated subpools of each tail identified outlier values. These would have been retained in the analysis had this been based on only a single pool per tail. After elimination of outliers, the SE(D′) was reduced to 0.056 and this value was used to calculate the *Z*-test statistic. The empirical value of 0.056 obtained for SE(D′) was considerably larger than the value of 0.043 calculated according to the deterministic procedure of Lipkin et al. (1998), apparently because it accounts for sources of errors such as sample cell count in addition to those included in the deterministic procedure (technical error and binomial sampling).

The Proportion of False Null Hypotheses Among All Null Hypotheses

Across both traits, there were a total of 340 marker-trait tests and 1,538 sire-marker-trait tests (Table 3). Both at the marker level and sire-marker level, there was a marked excess of low (significant) *P*-values, compared with the null-hypothesis expectation of equal numbers (i.e., 10 % of all tests) in each bin. The excess of low *P*-values was less evident at the sire-marker-trait level than at the marker-trait level. This is due to the fact that when a sire is tested at a marker in linkage to a QTL, a significant sire-marker effect is obtained only if the linked QTL is also in the heterozygous state and fewer data are available for each sire, compared with the data available across sires. At both the marker-trait and sire-marker-trait levels, the excess of low *P*-values was similar for PP and MY.

At the sire-marker level, estimates of the proportion of false null-hypotheses, (i.e., n_1 out of all N sire-marker tests) were $P_S = 0.205$ and 0.224 for PP and MY, respectively (Table 3). This represents the estimated proportion of true marker linkage to a heterozygous QTL for the individual sires. At the marker level (Table 3), the

Table 3. Distribution of confidence-wise error rate (CWER) *P*-values for marker-trait tests and for sire-marker-trait tests; estimated number of false (n_1) and true (n_2) null hypotheses at the 2 levels of comparison; critical *P*-value, threshold *P*-values for PFP < 0.10 (marker tests) and PFP < 0.20 (sire-marker tests)

CWER-P Bin	Marker		Sire-marker	
	Milk yield	Protein percent	Milk yield	Protein percent
0.1	65	70	185	188
0.2	21	20	69	79
0.3	23	15	78	65
0.4	10	15	82	65
0.5	8	14	61	62
0.6	5	6	66	69
0.7	11	12	56	63
0.8	11	9	60	60
0.9	8	4	52	54
1	6	7	67	57
Total	168	172	776	762
n_1	86	96	174	156
n_2	82	76	602	606
Critical *P*-value	0.0708	0.0727	0.0314	0.0358

Table 4. Total number of tests at marker and sire-marker level, and number and proportion of significant tests (marker level, PFP ≤ 0.10; sire-marker level PFP ≤ 0.20 within significant markers)

	Total tests			Significant tests					
Item	Milk yield	Protein percent	Milk yield and protein percent	Milk yield only	Protein percent only	Milk yield and protein percent	Milk yield all[1]	Protein percent all[1]	None
Marker level (across sires)	168	172	168	36	33	24	60	58[2]	75
PFP 10%				21.4%	19.6%	14.3%	35.7%	34.1%	44.6%
Sire-marker level (within sire)	776	762	663	76	73	10	86	83	504
PFP 20% within significant markers				11.5%	11.0%	1.5%	11.1%	10.8%	76.0%

[1]All significant tests for the given trait, including instances significant for both traits.

[2]This includes 1 significant marker among the 4 markers tested for protein percent only.

comparable estimates were P_M = 0.558 and 0.512 for PP and MY, respectively. This represents the estimated proportion of true marker-QTL linkage. Because the difference between the estimates for the sire-marker level and the marker level is due solely to the fact that the sire level requires QTL heterozygosity as well as marker heterozygosity, the proportion of QTL heterozygosity can be estimated as $Q_H = P_S/P_M$ = 0.37 and 0.44 for PP and MY, respectively.

Critical CWER P-Values and Power of the Tests

At both test levels (marker and sire-marker) critical PFP values were similar for PP and MY, but critical *P*-values at the marker level were about twice those at the sire-marker level (Table 3).

Marker Level. There were a total of 340 tests, of which 118 were significant at PFP ≤ 0.l0 (Table 4). Of these, 60 markers were significant for MY and 58 were significant for PP. The estimated numbers of marker-QTL linkages (Table 3) were 86 for MY and 96 for PP. The resulting power of the marker tests (Pm), adjusted for presence of false positives, was 0.63 and 0.54 for MY and PP, respectively.

Sire-Marker-Trait Level. A total of 1,538 tests were performed, of which 204 were significant at PFP ≤ 0.20. These can be divided into 2 groups. The first, representing 534 tests (34.7% of the total), includes tests within the 118 marker-trait combinations that were significant at PFP ≤ 0.10. Among these, 169 sire-marker tests showed significance at PFP = 0.20. The second group of tests, comprising 1,004 tests (65.3% of the total), includes tests within the 222 marker-trait combinations that were not significant at PFP ≤ 0.10. Among these, there were only 35 significant sire-marker-trait tests at PFP ≤ 0.20. The total estimated numbers of sire-marker combinations heterozygous at the QTL (n_{1S}, listed as n_1 under sire-marker in Table 3), were then 174 and 156 for MY and PP, respectively. Thus, effective power at the sire-marker level (Ps), corrected for false positives, was 0.39 for MY and 0.43 for PP.

Relationships Between the Two Traits

Of the 168 markers tested for both traits, 75 (44.65%) were not significant for either of the traits (Table 4). Of the 93 markers significant for at least 1 trait, 36 (38.7%) were significant at PFP ≤ 0.1 for MY alone, 33 (35.5%) for PP alone, and 24 (25.8%) for both (Table 4). Significant markers were distributed over 24 chromosomes (all except for BTA5, 15, 24, 25, 29) for MY; and 26 chromosomes (all except for BTA9, 15, 18) for PP.

Of the 663 sire-marker combinations tested for both traits, 504 (76.0%) were not significant for either trait (Table 4). Of the 159 significant sire-marker combinations, 76 (47.8%) were significant for MY, 73 (45.9%) were significant for PP, and only 10 (6.3%) were significant for both traits (Table 4). Part of the disproportion between significant tests for both traits obtained within family and those found across families can be attributed to the lower power of the sire-marker tests, which are only about half those of the marker tests. Nevertheless, here too, the large number of tests significant for a single trait only indicates that at least some of these loci are indeed specific to a single trait. Thus, the indication from both levels of analysis is that a significant proportion of loci affects both MY and PP, but an appreciable proportion of loci affects a single trait only.

Among the 10 sire-marker combinations for which both traits were significant, 9 showed opposite effect direction, as expected from the biometrical relationships between milk yield and protein percent. However, one of the sire-marker combinations showed effect estimates of the same sign.

Significant Markers and Sire-Marker Families

Tables 5 and 6 have results for significant MY (Table 5) and PP (Table 6) single marker tests. When the marker was significant at PFP < 0.10 but none of the sire-marker tests was significant at the PFP < 0.20 level, an allele substitution effect was not calculated. This occurred for 5 of the 60 significant markers for

Table 5. Milk yield—chromosome (BTA = *Bos taurus* autosome), location, marker name, and number of heterozygous/significant sire-families tested (in parentheses), significance level, and mean allele substitution effect (alpha, in units of the trait SD for EBV) calculated within significant families[1]

BTA	cM	Marker	*P(Marker)*	Alpha (SD)	BTA	cM	Marker	*P(Marker)*	Alpha (SD)	BTA	cM	Marker	*P(Marker)*	Alpha (SD)
1	15.4	BMS0574 (6/2)	2.65E-03	0.41	9	24.1	BM1227 (4/1)	2.14E-02	0.40	19	95.0	RM388 (4/1)	2.19E-02	0.31
	73.1	BMS4001 (7)	5.53E-02		10	100.0	BMS0614 (8/3)	4.60E-04	0.38	20	38.4	GHRIA (2/1)	3.99E-02	0.38
	83.8	BM8246 (8/3)	3.49E-03	0.33	11	19.4	BM0716 (4/2)	6.57E-04	0.32		55.1	AGLA29 (6/1)	3.23E-03	0.24
	122.1	CSSM019 (6/2)	6.12E-04	0.38		54.5	BMS1716 (6/1)	7.92E-03	0.43	21	62.7	TGLA122 (7/1)	2.42E-02	0.44
	142.2	BMS4044 (7/2)	2.07E-02	0.28		81.8	IDVGA-3 (6/1)	2.40E-02	0.43		75.3	BMS743 (5/1)	1.77E-03	0.47
2	11.9	TGLA431 (2/1)	2.33E-03	0.43		92.2	BMS989 (6/1)	2.46E-02	0.37	22	2.9	INRA026 (4)	2.42E-02	
	74.5	TEXAN01 (7/2)	5.32E-06	0.55		112.3	BMS0607 (5/2)	5.55E-02	0.32	23	52.3	RM185 (6/1)	1.23E-02	0.37
	96.1	TEXAN04 (5/2)	7.43E-03	0.33	12	50.4	BM0860 (7/1)	8.02E-03	0.42		56.3	BM7233 (4/2)	4.41E-02	0.30
	115.4	BM2113 (6)	6.76E-02			83.6	BM4028 (7/1)	1.47E-02	0.49	26	2.8	BMS651 (4/2)	1.05E-04	0.48
3	0.0	BMS871 (3/1)	5.32E-02	0.30		102.0	BMS1316 (4/1)	9.80E-03	0.43		41.6	BM4505 (6/1)	4.82E-02	0.48
	17.1	INRA006 (4)	6.88E-02		13	23.0	BMS1742 (4/2)	2.26E-02	0.32	27	0.0	BM3507 (6/2)	1.48E-02	0.36
	68.0	HUJ246 (7/3)	1.94E-03	0.34		27.6	BMC1222 (4/1)	1.45E-02	0.28		55.8	RM209 (6/2)	1.84E-02	0.36
	125.8	BMC4214 (3/1)	8.90E-03	0.46		81.0	BL1071 (7/1)	6.79E-03	0.59	28	49.4	BMS1714 (7/1)	1.88E-02	0.47
4	12.5	BMS1788 (5/2)	7.57E-09	0.59	14	5.1	CSSM066 (6/1)	5.01E-02	0.40		50.5	BMS2200 (7/1)	1.66E-02	0.25
	87.3	OBESE (4/1)	1.80E-02	0.37		10.5	BMS1747 (5/1)	4.85E-02	0.35					
6	8.2	INRA133 (4/1)	1.52E-02	0.34		69.8	BMS947 (7/2)	1.48E-03	0.39					
	35.5	BM1329 (4/2)	5.83E-03	0.44		100.0	BL1036 (5/2)	3.11E-04	0.45					
	91.5	CSN3 (6/3)	3.38E-04	0.36	16	54.1	CSSM028 (2/1)	3.08E-03	0.31					
7	25.4	RM006 (7)	5.50E-02		17	5.5	BMS499 (8/5)	2.11E-05	0.36					
	72.9	INRA112 (5/2)	2.17E-02	0.32		38.3	CSSM9 (7/1)	4.07E-03	0.49					
	90.7	BMS1331 (6/1)	7.08E-02	0.38		67.3	CSSM033 (4/1)	3.08E-03	0.41					
	101.1	BM9065 (8/2)	5.63E-02	0.25		92.1	BM1233 (5/1)	2.64E-02	0.36					
8	41.6	BMS0678 (2/1)	2.12E-02	0.28	18	54.7	ILSTS002 (4/1)	4.91E-02	0.37					

[1]Where alpha is missing none of the sires was significant at the sire-marker level. *P*-value, comparison-wise error rate at the marker level.

Table 6. Protein percent—chromosome (BTA = *Bos taurus* autosome), location, marker name, and number of heterozygous/significant sire-families tested (in parentheses), significance level, and allele substitution effect (alpha, in units of the trait SD for EBV values) calculated within significant families[1]

BTA	cM	Marker	*P(Marker)*	Alpha (SD)	BTA	cM	Marker	*P(Marker)*	Alpha (SD)	BTA	cM	Marker	*P(Marker)*	Alpha (SD)
1	122.1	CSSM019 (6/3)	5.35E-04	0.34	8	2.7	BMS1864 (6/1)	3.56E-02	0.24	21	13.5	HEL05 (3/1)	1.76E-02	0.43
	142.2	BMS4044 (6/2)	3.87E-03	0.47		31.4	BM0310 (5/1)	1.87E-02	0.34		29.7	BM0103 (4/2)	5.07E-03	0.39
	148.2	BMS0922 (5/1)	4.12E-02	0.31	10	100.0	BMS0614 (3/1)	2.32E-02	0.37	22	2.9	INRA026 (3/2)	1.57E-04	0.61
2	11.9	TGLA431 (3/1)	1.70E-02	0.31	11	19.4	BM0716 (4/2)	6.26E-04	0.38	23	42.9	BMS468 (3/1)	5.44E-02	0.36
	74.5	TEXAN01 (7/3)	2.68E-04	0.36		92.2	BMS989 (3)	4.36E-02			52.3	RM185 (8/2)	3.06E-03	0.37
3	59.4	INRA003 (3/2)	6.49E-03	0.27		117.0	BMS655 (5/1)	6.46E-02	0.35	24	8.1	BMS2526 (6/1)	4.59E-02	0.32
	68.0	HUJ246 (7/3)	4.27E-03	0.34	12	83.6	BM4028 (6/1)	4.79E-02	0.35		43.9	BMS1743 (6/3)	2.16E-04	0.36
	77.6	BMS1266 (3/1)	8.84E-03	0.35	13	27.6	BMC1222 (6/2)	8.91E-03	0.36	25	14.4	BMS130 (1/1)	2.00E-02	0.28
	116.5	BMS0896 (3)	4.42E-02			51.7	HUJ616 (6/1)	3.88E-02	0.50	26	27.0	BM1314 (6/2)	3.38E-03	0.51
4	87.3	OBESE (3/2)	1.11E-04	0.36		81.0	BL1071 (8/1)	1.56E-02	0.41	27	55.8	RM209 (6/1)	2.96E-02	0.30
5	17.3	BP1 (5/1)	1.56E-02	0.30	14	0.0	ILSTS039 (7/3)	1.14E-02	0.40	28	49.4	BMS1714 (7/2)	1.06E-02	0.44
6	35.5	BM1329 (4/1)	3.56E-04	0.58		5.1	CSSM066 (6/1)	4.50E-03	0.83	29	21.1	BMC8012 (4/2)	2.95E-03	0.41
	63.9	BM4322 (6/1)	1.83E-02	0.31		8.1	DIK2201 (3/1)	1.50E-02	0.59		24.2	ILSTS089 (4/1)	4.97E-02	0.52
	67.4	BMS0470 (4/2)	3.01E-03	0.31		93.7	BMS2055 (5/1)	1.20E-02	0.29					
	72.4	ILSTS097 (8/4)	7.15E-05	0.39	16	30.2	BM4025 (3/1)	3.94E-02	0.27					
	82.0	BM0415 (3/2)	7.64E-04	0.45	17	67.3	CSSM033 (4/2)	4.94E-04	0.36					
	101.4	BM8124 (4/1)	1.97E-02	0.33		92.1	BM1233 (6/1)	2.68E-02	0.21					
7	16.8	BMS713 (7)	6.41E-02		19	70.2	Map2C (1)	3.95E-02						
	25.4	RM006 (7/2)	3.14E-02	0.42	20	8.2	BM1225 (3)	3.37E-02						
	72.9	INRA112 (4/1)	1.10E-02	0.45		19.1	BMS1282 (4/1)	1.47E-02	0.67					
	90.7	BMS1331 (6)	7.27E-02			58.7	BMS1120 (4/1)	8.84E-03	0.34					
	101.1	BM9065 (8/2)	1.85E-02	0.48		63.2	DIK4527 (2/2)	4.54E-03	0.38					
	116.6	ILSTS006 (5/1)	6.01E-02	0.37										

[1]Where alpha is missing none of the sires was significant at the sire-marker level. *P*-value, comparison-wise error rate at the marker level.

MY, and for 6 of the 58 significant markers for PP. For marker Map2C on BTA19 for PP there was only a single heterozygous sire having CWER *P*-value of 0.0395. This was significant at the marker level (critical *P*-value for PFP ≤ 0.10 = 0.073), but just below significance threshold at the sire-marker level (critical *P*-value for PFP ≤ 0.20 = 0.036).

Allele Substitution Effects

Allele substitution effects in units of the SD of EBV (190 kg for MY and 0.053% for PP) were 0.38 (range 0.24 to 0.59) for MY and 0.39 (range 0.21 to 0.83) for PP. These findings are higher than the values generally found as they represent fractions of an approximation to a genetic standard deviation, rather than a phenotypic standard deviation.

Defining QTLR

In the present study, a total of 55 QTLR affecting MY, PP, or both were identified (Table 7). These QTLR were distributed across 28 of the 29 bovine autosomes (all, except for BTA15). On 10 chromosomes a single QTLR was identified and 9 chromosomes showed evidence for 2 and 3 QTLR. Of the total of 55 QTLR, 16 (29%) were for MY alone, 13 (23.6%) were for PP alone, and 26 (47.2%) were for MY and PP. The proportion of regions affecting both MY and PP is considerably higher than that found for the individual marker level and sire-marker level tests. This is due to the fact that for many of the QTLR multiple markers were tested, increasing the power of the test, so that loci had more of an opportunity to show significance for both of the traits. Nevertheless, an appreciable number of regions affecting only one of the traits remain, suggesting that at least some of these QTLR are indeed limited in their effect to only one of the traits. This has important implications for marker-assisted selection which will be considered in the Discussion section. For convenience, Table 7 also compares the QTLR identified in the present study with those reported in the 3 Web databases. This, too, will be addressed in more detail in the Discussion section.

DISCUSSION

Comparison of the Results of the Present Study and Those Reported Previously

Of the 55 QTLR found in the present study, 43 (78.2%) were also reported in at least one of the 3 Web-based QTL maps (http://www.vetsci.usyd.edu.au/reprogen/QTL_Map/; http://www.animalgenome.org/QTLdb/cattle.html; http://bovineqtl.tamu.edu/). Of these, 13 (30.2%) were reported in all 3 databases, 12 (27.9%) were reported in 2 of the databases, and 18 (41.9%) were reported in a single database only.

Of the 43 QTLR found in the present study that also had reported results in one or more of the 3 Web databases, 10 affected PP only, 8 MY only, and 25 both MY and PP. Remarkably, all 10 QTLR of the present study affecting PP only and all 8 QTLR of the present study affecting MY only were also reported in the Web databases as affecting the same single trait only (see later for further discussion of these findings), whereas of the 25 QTLR affecting both traits in the present study, 11 were reported also in the databases as affecting both MY and PP, 10 were reported as affecting PP only, and 4 as affecting MY only. Nine of the 12 QTLR identified in the present study that were not represented in the databases, and 10 of the 18 QTLR that were represented in a single database only, were assigned to chromosomes BTA16 through BTA29. These are the smaller bovine chromosomes and have been less intensively studied than chromosomes BTA1 through BTA15. Thus, in some part, the lack of representation of the results of the present study in the databases may be due to paucity of reported results in the literature for these chromosomal regions. In this study QTLR affecting both MY and PP, for which effects on only one of the traits (PP or MY) were reported in the databases, were found. Part of this discrepancy can be attributed to sampling variation. However, in 2 instances (BTA13 region 1 and BTA14 region 4) all 3 databases reported effects on PP only, and the present study found effects on both PP and MY. This may indicate different effects of the alleles segregating in the Brown Swiss, compared with those segregating in the populations (close to 100% Holstein) screened for the databases, but could also represent false positives. Because the databases draw partly on the same literature, they are correlated.

Aside from the few differences outlined above, the overall concordance between the results presented in this study and those available in the Web databases is very high and indicates that most of the QTL segregating in the reference populations (mainly Holstein) from which the databases are drawn are also segregating in the Brown Swiss population. The high general correspondence in QTLR distribution between the Brown Swiss and that reported in the databases is somewhat unexpected. The origin and selection history for the Brown Swiss and the Holstein breeds differ. The Holstein breed originated in the lowlands (the Netherlands) of Europe and has been generally selected to increase milk yield, and only in the last 2 decades or so have milk quality and functional traits become a target for selection. The Brown Swiss was originally a dual-purpose Alpine breed, and only in the past 4 de-

Table 7. Investigated *Bos taurus* autosomes (BTA), number of markers tested in parentheses, identification of QTL region (QTLR) within chromosome, location (cM on USDA cattle genome map), trait significant in this study (MY = M, PP = P), and trait significant for milk yield (M) or milk protein percentage (P) in the studies reported in the 3 databases available on the Web

BTA	Markers, n	Region	Location	This study	Web site		
					QTL Map	qtlDB	qtlViewer
1	1	1	15.4	M	M	M	M
	1	2	46.8	—			
	5	3	73.1–148.2	MP	MP	M	MP
2	1	1	11.9	MP	P	P	M
	3	2	38.9–61.7	—			
	3	3	74.5–115.4	MP	P		
3	2	1	0.0–17.1	M	M		
	3	2	27.4–46.0	—			
	4	3	59.4–77.6	MP	M	M	MP
	1	4	87.3	—			
	2	5	116.5–125.8	MP	P		M
4	1	1	12.5	M			
	2	2	28.4–56.6	—			
	1	3	87.3	MP	M		
	1	4	99.7	—			
5	1	1	17.3	P		P	
	4	2	56.3–112.4	—			
6	1	1	8.2–35.5	MP	MP		MP
	2	2	43.9–53.7	—			
	5	3	63.9–82.0	P	P	P	P
	3	4	91.5–101.4	MP	MP	MP	M
	2	5	119–127.3	—			
7	1	1	0	—			
	2	2	16.8–25.4	MP	P		P
	2	3	32.0–58.5	—			
	2	4	72.9–90.7	MP	MP	P	M
	3	5	90.7–116.6	MP	MP	MP	M
8	1	1	2.7	P			
	1	2	21.1	—			
	2	3	31.4–41.6	MP	P		
	3	4	63.0–122.9	—			
9	1	1	24.1	M			
	3	2	50–90.7	—			
10	3	1	35–82.2	—			
	1	2	100.0	MP	M	MP	
11	1	1	19.4	MP	P		
	1	2	54.5	M			
	5	3	81.8–117.0	MP	P	P	
12	1	1	15.1	—			
	2	2	20.8–50.4	M	M	M	M
	2	3	83.6–102.0	MP	P		
13	2	1	23.0–27.6	MP	P	P	P
	1	2	41.7	—			
	1	3	51.7	P	P	P	
	1	4	60.0	—			
	1	5	81.0	MP	M	M	
14	4	1	0.0–10.5	MP	MP	MP	MP
	5	2	17.8–69.0	—			
	1	3	69.8	M	M	M	
	2	4	93.7–100.0	MP	P	P	P
16	1	1	14.2	—			
	1	2	30.2	P			
	1	3	54.1	M			
	2	4	80.0–90.8	—			
17	1	1	5.5	M			
	1	2	38.3	M			
	1	3	54.7	—			
	2	4	67.3–92.1	MP		M	M
18	1	1	24.5	—			
	1	2	54.7	M		M	M
	1	3	84.1	—			

Continued

Table 7 (Continued). Investigated *Bos taurus* autosomes (BTA), number of markers tested in parentheses, identification of QTL region (QTLR) within chromosome, location (cM on USDA cattle genome map), trait significant in this study (MY = M, PP = P), and trait significant for milk yield (M) or milk protein percentage (P) in the studies reported in the 3 databases available on the Web

					Web site		
BTA	Markers, n	Region	Location	This study	QTL Map	qtlDB	qtlViewer
19	2	1	16.0–43.3	—			
	2	2	70.2–95.0	MP			M
20	1	1	0	—			
	2	2	8.2–19.1	P			P
	5	3	26.3–37.8	—			
	4	4	38.4–55.1	M	M	M	M
	2	5	56.9–57.34	—			
	2	6	58.7–63.2	P		P	P
	4	7	71.8–82.94	—			
21	2	1	13.5–29.7	P	P	P	
	2	2	35.9–45.2	—			
	2	3	62.7–75.3	M			
22	1	1	2.9	MP			
	3	2	48.9–82.9	—			
23	2	1	11.8–20.7	—			
	3	2	42.9–56.3	MP	M	P	P
	1	3	71.6	—			
24	1	1	8.1	P			
	1	2	36.9	—			
	1	3	43.9	P	P		
	1	4	61.2	—			
25	1	1	14.4	P		P	
	1	2	46.4	—			
26	1	1	2.8	M			
	1	2	27.0	P	P		
	1	3	41.6	M		M	
	3	4	52.4–74.7	—			
27	1	1	0.0	M	M		
	2	2	13.2–45.2	—			
	1	3	55.8	MP			P
28	2	1	8.0–29.2	—			
	2	2	49.4–50.5	MP			P
29	2	1	21.1–24.2	P	P		
	1	2	40.2	—			

cades has it been selected primarily for milk production, retaining however, a strong emphasis on functional traits, with weighting about 30% in the selection indices of the Italian Brown Swiss population, and as much as 45% in the German and Austrian Brown Swiss populations.

Although we might have expected somewhat greater differentiation between QTL found in the Brown Swiss and those reported in the Web databases for the Holstein than actually observed, the present results suggest that the same genes may be involved in genetic variation in milk production in the 2 breeds. The appreciable difference in production levels between them would then be more a matter of allele frequencies at the responsible genes than differences in the genes that are segregating. This would be consistent with the fact that what really differs in the selection history of the 2 breeds is primarily the intensity of selection for the 2 traits here investigated and not the direction of selection.

Nevertheless, some of the novel QTLR reported here, and particularly the 2 above mentioned QTLR with different effects in the Web databases and the present study (BTA13 region 1, and BTA14 region 4) may represent true differences in the QTL segregating in the Brown Swiss and Holstein populations.

Specific QTLR for PP and for MY and MAS

The complete concordance of results between the present study and the databases with respect to QTLR with effects limited to MY or PP supports the assumption that these specific QTLR are, indeed, affecting only 1 of the 2 traits, PP or MY. Furthermore, of the 43 QTLR of the databases that had corresponding QTLR in the present study, 11 had effects on both MY and PP. All 11 of these QTLR corresponded to QTLR of the present study that also affected both traits. Thus, this supports the assumption that these QTLR do indeed have effects on both traits. Thus, of the 43 QTLR re-

ported in the present study and also in the databases, 29 (10 for PP, 8 for MY, and 11 for PP and MY) show complete concordance between the present study and the databases. Considering that the present study is completely independent of the databases, these QTLR and their specific effects can be regarded as confirmed QTL, using the terminology of Lander and Kruglyak (1995). With respect to the remaining 14 QTLR for which the present study reported effects on both traits, but the databases reported effects on a single traits only, part may represent incomplete power of the experiments reported in the databases (it should be noted that all 3 databases are based mostly on the same literature, and hence are highly correlated), but some may indeed, as noted above, represent different effects in the Brown Swiss breed.

The loci that apparently affect PP only are of particular interest because they would appear to provide an avenue to increase milk protein yield through MAS without increasing milk yield and consequently without increasing functional stress on the cow. These loci should be prime candidates for high resolution mapping followed by candidate gene analysis. Similarly, the loci that affect MY without affecting PP would appear to provide loci that could act powerfully to increase overall protein yield. In contrast, the many loci that affect both MY and PP would be expected to act in opposite directions on the 2 traits, so having little net effect on protein yield.

Selective DNA Pooling – The Case of the Brown Swiss as a Paradigm for an International QTL Mapping Project in the Cosmopolitan Minor Breeds

Mapping QTL in minor cosmopolitan breeds having a different genetic makeup than the predominant breeds due to different selection criteria and environment, can lead to the detection of QTL that are not segregating in the Holstein population and paves the way to detect genes related to very important economic traits under selection in specific genomic regions.

Generally, the limited size of outbred populations structured in large half-sib families is a limiting factor for mapping QTL for economically important traits in these minor breeds. In the present experiment, an international sampling design was used to collect DNA samples of individuals from 10 sire families across national boundaries. International sampling of the largest families was carried out evaluating the family size jointly for the 3 countries involved in the project, Austria, Germany, and Italy. In Europe some autochthonous breeds in the Alps (e.g., Alpine Gray) represent a similar situation: they are located across countries, and as a population as a whole they have a sufficient numbers and family structure for a QTL mapping study based on daughter design, but not if they are considered as resource population within any single country.

One concern of mapping QTL with a DD is the possibility to have available EBV of each daughter on the same scale base. However, an international EBV service is not presently available for female populations. Even if the cosmopolitan minor breed mainly comes from the same original strain, the different selection goals may have changed the genetic background of the female populations. The selective DNA pooling design used in this project can overcome the necessity to convert phenotypes to the same scale to use them jointly in the association analyses. Pooling individuals from different genetic strains does not need the direct comparison of EBV calculated by different systems and methods. In fact, selective DNA pooling allows mixing of individuals from the trait-tail of each country distribution of the same sire family, taking a proportional number of individuals according to country origin.

This approach is based on the assumption that the different populations are at the same genetic evolutionary moment (i.e., active male reproducers and close female individuals ancestors are of the same generation), or in other words that the average genetic value of the sire family in country A is equivalent to the average genetic value in country B. This assumption deals very effectively with the reality of the international market of semen that has made the same genes available across many national populations for several decades. Calculation of EBV for the daughters on an international scale is only needed for the estimation of allele substitution effect and not to test the QTL association. This is not the case for a DD with individual genotyping, for which association is detected with multiple regression approach and interval mapping.

CONCLUSIONS

This work is the first QTL genome scan in the Brown Swiss population for productive traits. Knowledge of the QTLR is a basis for further studies to better identify location and effect of the QTL with the use of high-resolution mapping using novel approaches such as the fractioned pooling design with selective DNA pooling (Frenkel et al., 2005; Korol et al., 2007), or the utilization of linkage disequilibrium mapping methods (Farnir et al., 2000; Cohen et al., 2002; Meuwissen et al., 2002).

Even if the QTL information presently available in the Brown Swiss population can be utilized only in a within family MAS approach, we believe that the possibility for the breed organizations to use available knowledge of detected QTL should motivate them to implement a structure for a genomic selection of indi-

viduals: new expected findings in gene identification can be expected to increase the efficiency of genomic assisted selection in the near future (Meuwissen et al., 2001, Schaeffer, 2006). In the Brown Swiss population the limited size of each national breeding program may motivate the stakeholders of the selection to join their effort internationally to increase their chances for success in the application of genomic technologies. Today the stakeholders of the genetic improvement in dairy cattle have indeed the possibility to be the prime movers in the gene discovery activity and in their use for selection in livestock. The confirmed QTLR identified in the Brown Swiss and the 3 Web-based databases represent a basic resource for such gene discovery.

ACKNOWLEDGMENTS

We thank all the Italian Provincial and Regional Breeders Associations; the milk recording agencies of Austria and Germany; the milk processing laboratories of the official milk recording of AUT, DEU, and ITA; and the Austrian, German, Italian Brown Breeders organizations for the precious help in sampling the population, conducting the cell count, and for the use of the herd books. This research was supported by the FP5 program of the EU (BovMAS project QLK5-CT-2001-02379).

REFERENCES

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

Canavesi, F., D. Boichard, V. Ducroq, N. Gengler, G. De Jong, and Z. Liu. 2001. Production traits European joint evaluation (PROTEJE). Interbull Bull. 27:32–34.

Cohen, M., E. Serrousi, M. Ron, M. Reichenstein, A. Plis-Finarov, M. Shani, and J. I. Weller. 2002. Population-wide linkage disequilibrium between a SNP and a QTL affecting milk protein production on BTA6 in dairy cattle. No. 09-12 in Proc. 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France.

Darvasi, A., and M. Soller. 1994. Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. Genetics 138:1365–1373.

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

Dekkers, J. C. M., J. Wang, and R. L. Fernando. 1999. Quantitative trait locus mapping using selective DNA pooling data. Page 165 in Proc. Conf. "From Jay L. Lush to Genomics: Vision for Animal Breeding and Genetics," Ames, IA.

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

Dolezal, M. 2007. Improvements in selective DNA pooling for quantitative trait locus mapping in dairy and dual purpose cattle. PhD Thesis. University of Natural Resources and Applied Life Sciences, Vienna, Austria.

Emmerling, R., M. Lidauer, and E. A. Mäntysaari. 2002. Multiple lactation random regression test-day model for Simmental and Brown Swiss in Germany and Austria. Interbull Bull. 29:111–117.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to Quantitative Genetics. 4th ed. Longmans Green, Harlow, Essex, UK.

Farnir, F., W. Coppieters, J. Arranz, P. Berzi, N. Cambisano, B. Grisart, L. Karim, F. Marcq, L. Moreau, M. Mni, C. Nezer, P. Simon, P. Vanmanshoven, D. Wagenaar, and M. Georges. 2000. Extensive genome-wide linkage disequilibrium in cattle. Genome Res. 10:220–226.

Fernando, R. L., D. Nettleton, B. R. Southey, J. C. M. Dekkers, M. Rothschild, and M. Soller. 2004. Controlling the proportion of false positives (PFP) in a multiple test situation. Genetics 166:611–619.

Frenkel, Z., L. Cohen, E. Lipkin, M. Soller, and A. Korol. 2005. A new approach of QTL mapping analysis of selective DNA pooling experiments. Page 857 in Proc. Plant & Animal Genomes XIII Conference, San Diego, CA.

Hu, Z. L., E. R. Fritz, and J. M. Reecy. 2007. AnimalQTLdb: A livestock QTL database tool set for positional QTL information mining and beyond. Nucleic Acids Res. 35:D604–609.

Interbull. 2000. Italy – BSW. National genetic evaluation programmes for dairy production traits practiced in Interbull member countries 1999–2000. Interbull Bull. 24:60.

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

Korol, Z. Frenkel, L. Cohen, E. Lipkin, and M. Soller. 2007. Fractioned DNA pooling: A new cost-effective strategy for fine mapping of quantitative trait loci. Genetics 176:2611–2623.

Lander, R., and L. Kruglyak. 1995. Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. Nat. Genet. 11:241–247.

Lipkin, E., M. O. Mosig, A. Darvasi, E. Ezra, A. Shalom, A. Friedmann, and M. Soller. 1998. Mapping loci controlling milk protein percentage in dairy cattle by means of selective milk DNA pooling using dinucleotide microsatellite markers. Genetics 149:1557–1567.

Lipkin, E., A. Shalom, H. Khatib, M. Soller, and A. Friedmann. 1993. Milk as a source of desoxyribonucleic acid and substrate for the polymerase chain reaction. J. Dairy Sci. 76:2025–2032.

Maltecca, C., A. Bagnato, and K. A. Weigel. 2004. Comparison of international dairy sire evaluations from meta-analysis of national estimated breeding values and direct analysis of individual animal performance records. J. Dairy Sci. 87:2599–2605.

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.

Meuwissen, T. H. E., A. Karlsen, S. Lien, I. Olsaker, and M. E. Goddard. 2002. Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. Genetics 161:373–379.

Mosig, M. O., E. Lipkin, G. Khutoreskaya, E. Tchourzyna, M. 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 milk DNA pooling in a daughter design, using an adjusted false discovery rate criterion. Genetics 157:1683–1698.

Polineni, P., P. Aragonda, S. R. Xavier, R. Furuta, and D. L. Adelson. 2006. The bovine QTL viewer: A web accessible database of bovine quantitative trait loci. BMC Bioinformatics 7:283.

Schaeffer, L. R. 1994. Multiple-country comparison of dairy sires. J. Dairy Sci. 77:2671–2678.

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

Thomsen, H., N. Reinsch, N. Xu, C. Looft, S. Grupe, C. Kuhn, G. A. Brockmann, M. Schwerin, H. B. Leyhe, S. Hiendleder, G. Erhardt, I. Medjugorac, I. Russ, M. Forster, B. Brenig, F. Reinhardt, R. Reents, J. Blumel, G. Averdunk, and E. Kalm. 2001. Comparison of estimated breeding values, daughter yield deviations and deregressed proofs within a whole genome scan for QTL. J. Anim. Breed. Genet. 118:357–370.

Velmala, R. J., H. J. Vilkki, K. T. Elo, D. J. De Koning, and A. V. Maki-Tanila. 1999. A search for quantitative trait loci for milk production traits on chromosome 6 in Finnish Ayrshire cattle. Anim. Genet. 30:136–143.

Weller, J. I., Y. Kashi, and M. Soller. 1990. Daughter and granddaughter designs for mapping of quantitative trait loci in dairy cattle. J. Dairy Sci. 73:2525–2537.
Wilmink, J. B. M., A. Meijering, and B. Engel. 1986. Conversion of breeding values for foreign populations. Livest. Prod. Sci. 14:223–229.

APPENDIX 1. ESTIMATION OF D_{IJK} AND SE(D′)

Estimation of D_{ijk}

Using the notation for the allele frequency as in the text in each of the subpools, the D_{ijk} was estimated as (modified from Lipkin et al., 1998; Mosig et al., 2001):

$$D_{ijk} = (D1_{ijk} + D2_{ijk})/2,$$

where (omitting the subscripts *ijk* for clarity), D1 = [(H1l – L1l) – (H1s – L1s)]/2 is the difference in frequency of sire marker alleles between the high and low tails, averaged across the 2 sire marker alleles, for a single pair of arbitrarily chosen subpools, one from each tail; and

$$D2 = [(H2l – L2l) – (H2s – L2s)]/2$$

is the difference in frequency of sire marker alleles between the high and low tails, averaged across the 2 sire marker alleles for the remaining pair of subpools, one from each tail.

The minus sign of the short alleles (H1s, L1s, H2s, L2s) in the expressions for D1 and D2 derives from the fact that in the event of marker-QTL linkage, selection will have opposite direction of effect on frequency of the alternative sire alleles in the 2 tails. In the initial scan, one of the duplicates of each of the 4 subpools per sire-trait combination was genotyped, and each D_{ijk} was based on frequencies estimated on these 4 subpools per sire-marker-trait combination. In suggestive QTL regions, the subpools were typed again with their duplicates, producing a D_{ijk} based on 8 subpool analyses, for a replicated and more robust comparison-wise marker test. In this case allele frequencies of each tail were obtained by averaging the corrected intensity (Lipkin et al., 1998) of the 2 duplicate subpools of each replicate subpool.

Estimating SE(D′), the Empirical SE(D)

Because pool densitometry and estimation of the frequency of the L and S sire alleles was performed separately for each of the 2 replicate subpools in each tail, it was possible to use the variance among D values estimated from the 2 replicate subpools of the same tail to provide an empirical estimate of SE(D), denoted SE(D′). The estimate of SE(D′), computed across markers and traits, was used in the present study in place of the rather complex deterministic calculation of SE(D) used in Lipkin et al. (1998) and Mosig et al. (2001). The empirical SE(D) was based on a tail D-value denoted DH for the H pools or DL for the L pools, obtained between the 2 independent subpools at the same tail for the same allele.

According to the notation in the text, calculation of DH and DL is as follows:

$$DH = [(H1l – H2l – (H1s – H2s)]/2$$

and

$$DL = [(L1l – L2l) – (L1s – L2s)]/2.$$

The DH and DL values include all of the factors causing sampling variation between high and low subpools of the same allele (e.g., between H1l and L1l; that is, all sources of technical error, binomial sampling, and other undefined factors). But because the DH, DL values are from the same tail, they do not include effects due to QTL linkage. Thus, the DH, DL values represent D values for a single pair of high-low pools under the null hypothesis: H_0 = absence of marker-QTL linkage. Note, that DH and DL have exactly parallel structure to D1 and D2, except that the 2 elements separated by a minus sign come from the same tail; whereas for D1 and D2, the 2 elements separated by a minus sign come from alternative tails.

Whenever DH or DL were larger than 0.2, the pools used in their estimation were considered as outliers and eliminated from the subsequent calculations only for this specific marker (a total of 24 DH and DL values were eliminated in this manner).

Because the actual D-values used to test marker-QTL linkage for any particular sire-marker-trait combination, have the composition D = (D1 + D2)/2, the corresponding D′ expression, substituting DH for D1 and DL for D2, has the composition

$$D′ = (DH + DL)/2.$$

In principle, SE(D′) is an estimate of SE(D) under the null hypothesis. In practice, SE(D′) was estimated from SE(DH,DL), the pooled standard deviation across all DH and DL values, as

$$SE(D′) = SE(DH,DL)/\sqrt{2}.$$

Because it is based on many more degrees of freedom, SE(D′) calculated in this way was found to provide a more stable estimate of SD(D′) than SE(D′) calculated

directly from D′ values. This SE(D′) was used in all subsequent calculations of Z values.

APPENDIX 2. ESTIMATION OF EBV STANDARD DEVIATION

The German/Austrian daughter EBV values were converted to the Italian values, using the Interbull conversion formulas available through the Breeders Association. The population SD of trait EBV values was the difference in mean value of the high and low pools for each of the 2 traits, with expectation equal to $2I_P\sigma_i$, where I_P is the selection intensity corresponding to selection of the extreme P proportion of the population, and σ_i is the population standard deviation for trait i (Falconer and Mackay, 1996). Based on this, it is possible to estimate the within-sire population standard deviation for the studied traits, from the known proportion selected to the high and low tails (P_i), and the observed difference between the mean of the high and low tails (T_i).

The within-sire standard deviation is an underestimate of the population standard deviation, because it includes only three-fourths of the genetic variance. Thus, assuming that EBV variance is completely genetic (i.e., EVB accuracy equal 1), the estimates of the within-sire standard deviation in EBV values were increased by a factor of $1/(0.75)^{0.5} = 1.15$ to give an estimate of the population standard deviation of EBV values.

In the ideal case, when the accuracy of EBV is 1, SD of EBV will be equal to the genetic SD. When accuracy is less than 1, estimates of allele effects are expected to be regressed to a similar degree as the SD of EBV. Given the average accuracy of daughter EBV is 0.49, this undoubtedly gives a somewhat conservative estimate of the standardized allele substitution effects.