The Use of Genetic Markers to Measure Genomic Response to Selection in Livestock

Luis Gomez-Raya,[*,1] **Hanne Gro Olsen,**[†] **Frode Lingaas,**[‡] **Helge Klungland,**[†] **Dag Inge Våge,**[†] **Ingrid Olsaker,**[‡] **Seblewengel Bekele Talle,**[†] **Monica Aasland**[†] **and Sigbjørn Lien**[†]

Centre UdL-IRTA, Area de Produccio Animal, 25198 Lleida, Spain, †Department of Animal Science, Agricultural University of Norway, N-1432 Ås, Norway and ‡Department of Morphology, Genetics and Aquatic Biology, Norwegian School of Veterinary Science, 0033 Oslo, Norway

ABSTRACT

A method to measure genomic response to natural and artificial selection by means of genetic markers in livestock is proposed. Genomic response through several levels of selection was measured using sequential testing for distorted segregation of alleles among selected and nonselected sons, single-sperm typing, and a test with records for growth performance. Statistical power at a significance level of 0.05 was >0.5 for a marker linked to a QTL with recombination fractions 0, 0.10, and 0.20 for detecting genomic responses for gene effects of 0.6, 0.7, and 1.0 phenotypic standard deviations, respectively. Genomic response to artificial selection in six commercial bull sire families comprising 285 half-sib sons selected for growth performance was measured using 282 genetic markers evenly distributed over the cattle genome. A genome-wide test using selected sons was significant ($P < 0.001$), indicating that selection induces changes in the genetic makeup of commercial cattle populations. Markers located in chromosomes 6, 10, and 16 identified regions in those chromosomes that are changing due to artificial selection as revealed by the association of records of performance with alleles at specific markers. Either natural selection or genetic drift may cause the observed genomic response for markers in chromosomes 1, 7, and 17.

G ENETIC improvement of farm animals has traditionally been done by artificial selection. In short, artificial selection means that animals with variations better fitted to the production conditions are chosen to breed, and, consequently, to pass on favorable characteristics (genes) to their offspring. Theoretically, this process leads to changes in allele frequencies at loci affecting the traits under selection (CROW and KIMURA 1970). However, experimental results showing changes in allele frequency at loci located genome-wide in populations undergoing artificial selection have not yet been reported. The avenue taken in recent years has been to use molecular markers to map quantitative trait loci (QTL) in commercial farm animals (*e.g.*, GEORGES *et al.* 1995) with the final goal of improving the efficiency of artificial selection in the so-called marker-assisted selection.

Genomic response to selection is the change in allele frequency of loci in specific locations of the genome as a result of selection for a quantitative trait. Random anonymous markers such as microsatellites offer the opportunity to investigate genomic response within large progeny groups of bull sires because of the widespread use of artificial insemination in cattle.

The objective of this article is to propose a novel approach to measure genomic response to selection using genetic markers on the basis of sequential testing of the genomic changes taking place through various levels of selection within the offspring or gametes of widely used sires in cattle. It includes (1) testing using selected young bulls for production traits (artificial selection), (2) testing using culled young bulls (zygotic and artificial selection), and (3) testing for selection in sperm cells (gametic selection).

MATERIALS AND METHODS

Population structure and selection: The material for this study is Norwegian cattle, which is a commercial breed specialized in both meat and milk production. Elite sires produce large groups of young bulls. Every year, 400 young bulls (sons of elite sires) are tested at station for growth (measured in grams/day and recorded between 3 and 11 months of age) and conformation. Genetic evaluations for growth utilize the bull's own performance and ~30 half-sibs. Of the 400 bulls entering in the station, 200 are selected for high growth. Of these, ~10 young bulls are eliminated for low fertility and another 70 bulls are eliminated for bad conformation. This process resulted in ~130 young bulls (~30% selected) with 250 daughters each that were progeny tested for milk. Table 1 shows the six bull sires used and the number of selected and culled sons.

DNA isolation and molecular markers: Semen for genotyping of sires and their selected half-sib sons (after selection for growth and conformation has taken place) was available from

[1] *Corresponding author:* Centre UdL-IRTA, Area de Produccio Animal, Av. Alcalde Rovira Roure, 177, 25198 Lleida, Spain. E-mail: luis.gomez@irta.es

TABLE 1

Identification number and name of the six sire families and number of selected and culled sons per sire family

	Size	
Sire	Selected	Culled
2005 Smidesang	71	23
2052 Mauland	32	45
2402 Thorset	54	52
2463 Jørgentvedt	39	42
2946 Bekkevold	42	37
3131 Okkelberg	47	63

the Norwegian breeding organization. Semen from culled sons was not available and DNA was extracted from serum samples routinely utilized in paternity testing of young bulls. Serum samples were lysated at 96° for 10 min and used directly in the PCR. A total of 282 genetic markers covering 2807 cM of the cattle genome (261 microsatellites, one protein polymorphism, and 20 coding genes) were typed in six elite bull sires and 285 sons were selected for high growth and good conformation in the commercial production of Norwegian cattle. A previously constructed genetic map (VÅGE *et al.* 2000; http://www.nlh.no/ihf/Genkartstorfe/) was used for testing for transmission disequilibrium among selected sons.

Distorted segregation of alleles among sperm cells for each of the six sires was tested using single-sperm typing (LI *et al.* 1988), which utilizes one single-sperm cell and repeated cycles of PCR, allowing amplification of the marker alleles. Single-sperm typing allows testing for a potentially unlimited number of meioses per donor. Approximately 50 sperm cells were isolated for each sire and DNA extraction followed by PCR amplification was carried out for the markers showing significant results in the previous tests. The percentage of success in amplifying single-sperm cells was between 89 and 96%.

Testing of genomic response to selection: Sequential testing for distorted segregation of alleles from widely used commercial bull sires is carried out in three stages: (1) among progeny selected for high growth using a full coverage of the genome, (2) among culled progeny on the basis of low growth, and (3) among sperm cells of sires. Typing in steps 2 and 3 was carried out only for markers yielding significant results in step 1 to reduce the amount of genotyping and reduce the problem of multiple testing at the last stage.

Testing for distorted segregation of alleles among either selected or culled young bulls can be attained by the following method. Consider an elite sire heterozygous for a marker with alleles A and a. DNA for testing dams' genotypes was not available. Alleles inherited from dams to sons could be either of the sire alleles (A or a) or any other allele segregating in the population (denoted by a^*). If this marker is nearby a QTL affecting a trait under selection, then the frequency among selected sons inheriting either A or a alleles from their sires would be different from the expected 50:50. This approach is similar to a transmission disequilibrium test developed for mapping loci affecting diseases in which the frequency of transmission of alleles from heterozygous parents to their affected offspring deviates from the expected 50% (SPIELMAN *et al.* 1993). We developed a maximum-likelihood approach to test for transmission disequilibrium of alleles to half-sib sons. Following GOMEZ-RAYA (2001), the likelihood equation to be maximized is

$$L(v, f_A) = K(v f_A)^{n_{AA}}(v(1 - f_A - f_{a^*}) + (1 - v)f_A)^{n_{Aa}}$$
$$\times (v f_{a^*})^{n_{Aa^*}}((1 - v)(1 - f_A - f_{a^*}))^{n_{aa}}((1 - v)f_{a^*})^{n_{aa^*}},$$

where K is a constant, v is the transmission disequilibrium parameter, f_A and f_a are frequencies of alleles A and a in the dam population, f_{a^*} is the frequency of any other allele different from A and a in the dam population, and n_i is the count for genotype i among sons ($i = AA, Aa, aa, Aa^*, aa^*$). Solutions were obtained using the grid search method after maximizing the above equation for v and f_A. The maximum-likelihood solution for f_{a^*} was $(n_{Aa^*} + n_{aa^*})/(n_{Aa^*} + n_{aa^*} + n_{aa} + n_{Aa} + n_{AA})$. Therefore, this method uses the expected frequency of each genotype to estimate the transmission parameter and allele frequencies in the dam population. The null hypothesis (marker not linked to QTL for growth) is $v = 0.5$. Rejection of the null hypothesis (linked QTL) is tested using the likelihood-ratio statistics test (LRT $= -2 \ln(L(0.5, \hat{f}_A)/L(\hat{v}, \hat{f}_A))$), where the numerator is the maximum likelihood under the null hypothesis (fixing $v = 0.5$ and estimating $\hat{f}_A$) and the denominator is the maximum likelihood estimating both $\hat{v}$ and $\hat{f}_A$. The distribution of LRT is a χ^2 with 1 d.f. when the sample size is large. A test was carried out for each marker and sire family. Following GOMEZ-RAYA (2001), an explicit solution for the transmission parameter can be obtained when estimates of allele frequencies in the dam population are available,

$$v = \frac{-(A - 2B) - \sqrt{(A - 2B)^2 - 4B(B + C - A)}}{2(B + C - A)},$$

where

$$A = [(n_{AA} + n_{Aa^*})f_a - (n_{aa} + n_{aa^*})f_A + (1 - 2f_A - f_{a^*})n_{Aa}]$$
$$B = [(n_{Aa} + n_{Aa^*})f_A]$$
$$C = -[(n_{aa} + n_{aa^*})(1 - f_A - f_{a^*})].$$

The genome-wide transmission disequilibrium testing was performed estimating both the transmission parameter and the dam allele frequencies. A joint test for all families for a particular marker was possible using the properties of the χ^2 distribution. That is, the sum of independent variables having central χ^2 distributions follows a central χ^2 distribution with degrees of freedom equal to the sum of the degrees of freedom of the former central χ^2 distributions. A test for selection at the genome level was possible under the assumption of independence of χ^2 across the genome. This is strictly speaking not true because linked markers will have a tendency to cosegregate their alleles. However, the impact of this assumption will be small since each marker will be "unlinked" with most of the other markers in the genome.

A statistical test for distorted segregation of alleles among sperm cells was carried out for those markers and bull sires showing significant results in the testing among selected young bulls. A test for distorted segregation using single-sperm typing was performed by counting the sperm cell carriers of alleles A and a (O_A and O_a, respectively) and by computing $\chi^2 = (O_A - E_A)^2/E_A + (O_a - E_a)^2/E_a$, where $E_A = E_a = \frac{1}{2}(O_A + O_a)$. This test is χ^2 distributed with 1 d.f.

A *t*-test contrasting growth performance among all sons (selected and culled) inheriting alternative alleles from their sires was carried out by

$$t = \frac{\bar{x} - \bar{y}}{s_{\text{pooled}} \sqrt{1/N_A + 1/N_a}},$$

where

$$s_{\text{pooled}} = \frac{\sqrt{\sum_{i=1}^{N_A} x_i^2 - (1/N_A)\left(\sum_{i=1}^{N_A} x_i\right)^2 + \sum_{j=1}^{N_a} y_j^2 - (1/N_a)\left(\sum_{j=1}^{N_a} y_j\right)^2}}{N_A + N_a - 2}$$

and x_i and y_j are the growth performances of son i inheriting allele A and son j inheriting allele a, respectively. $\bar{x}$ and $\bar{y}$ are the average growths of sons inheriting A and a alleles. The total numbers of sons inheriting alleles A and a were N_A and n_a, respectively.

To account for multiple testing, a permutation test (CHURCHILL and DOERGE 1994) was performed by shuffling 500,000 times the observations on growth performance and by randomly assigning those observations to sons. The shuffling of the observations was carried out within the offspring of each sire but keeping the same genotype at all tested markers for each son. The steps for computing experimentwise P values were, (1) finding the maximum t-test among all tested markers for each permuted datum, (2) ordering of the maximum t-test among all tested markers, (3) computing the number of times (ntimes) out of 500,000 permutations that the t-test in the real data was higher than the value of the permuted data, and (4) computing the experimentwise P value, $(1 - \text{ntimes})/500,000$.

Changes in the transmission disequilibrium parameter by random sampling: The sampling of gametes from a heterozygous sire can lead to changes in the transmission disequilibrium parameter by chance. The approximated accumulated probability of random changes in the transmission parameter can be computed by taking advantage of the normal approximation to the binomial distribution,

$$\text{prob} = 2\int_z^\infty f(x)\,dx,$$

where $f(x)$ is the standard normal density and $z = (x - 0.5n)/\sqrt{0.25n} = (2x - n)/\sqrt{n}$. The integral is multiplied by 2 to account for the probability of distorted segregation of either allele.

Statistical power for a transmission disequilibrium test: For simplicity, it is assumed that a transmission disequilibrium test (TDT) is performed in which inheritance of the allele from the sire can always be traced. Let the sire be heterozygous A/a at a genetic marker and the total number of offspring be n. Assume also that a QTL with alleles Q and q affecting growth is linked to the marker with a recombination fraction c. The linkage phase in the sire is AQ/aq. The distribution of phenotypic values is assumed normal with means μ_Q and μ_q for sons inheriting alleles Q and q from their sire. The expected means among sons inheriting alleles A and a are $(1 - c)\mu_Q + c\mu_q$ and $(1 - c)\mu_q + c\mu_Q$, respectively. Therefore, the difference between sons inheriting alleles A and a is $(1 - 2c)(\mu_Q - \mu_q)$. The sons are evaluated for growth performance using relatives' information. It is assumed that heritability of growth (h^2) was 0.40 and accuracy of estimated breeding values (r) was 0.80. Gene effects are usually given in phenotypic standard deviations rather than in the scale of the estimated breeding values using relatives' information. The variance of estimated breeding values is $r^2\sigma_A^2$. Thus, if g_P is the gene effect in phenotypic standard deviations then the gene effect in standard deviations of estimated breeding values is $g_{\hat{A}} = g_P r/h$. If the marker is linked to a QTL affecting growth then the proportion of sons inheriting either A or a will depart from 50% among sons selected for high growth. A mixture of two normal densities corresponding to the sons inheriting alternative alleles from their sire was used for the power calculation. The means of the two distributions used the expected value of gene effects in standard deviations of estimated breeding values. The variance was assumed to be the same in the two groups. Truncation selection makes different the relative contributions of the two groups of offspring to the selected group of sons. The proportion of sons inheriting A and a alleles can be computed by fixing the total proportion selected (p_s) and by using numerical integration on

$$p_s = \frac{1}{2}p_A + \frac{1}{2}p_a,$$

with $p_A = \int_t^\infty f(x|G_i = A)$, $p_a = \int_t^\infty f(x|G_i = a)$, and $f(x|G_i)$ being the normal density given the allele inherited from the sire ($G_i = A$ or a). In this equation, ½'s are used to weigh the contributions to the total selected proportion of the two normal densities and t is the selection truncation point. Under the null hypothesis, the two normal densities would have the same mean and variance. Under the alternative hypothesis, the two normal densities are assumed to have different means but the same variance. In the latter situation, values p_A and p_a can be found for a unique truncation point, t, analytically. The Simpson rule was used for increasing values of the abscissa until a value of t was found to satisfy the above equation for a given p_s. The values used for computing power corresponded to the actual proportion selected in Norwegian cattle ($p_s = 0.325$). The expected frequencies of A and a alleles among selected sons were

$$p_A^* = 0.5p_A/p_s \quad \text{and} \quad p_a^* = 0.5\,p_a/p_s.$$

The normal approximation to the binomial distribution was used to compute power. It is expected that 50% of the sons inherited either allele A or a from the sire under the null hypothesis. Therefore, the value of the abscissa under the null hypothesis is $z_0 = (0.5 + n_A - (0.5n))/\sqrt{0.25n}$, where n_A is the number of sons inheriting allele A. Numerical integration of the standard normal distribution for increasing values of n_A until the significance level (α) was performed, such as $\alpha = \frac{1}{2}\int_{z_0}^\infty f(x)\,dx$, where $f(x)$ is the standard normal density. For the alternative hypothesis, z_A is computed as if the samples of sons would come from a population with the expected frequencies of alleles A and a among selected sons, p_A^* and p_a^*, respectively. Integration of the normal distribution to compute the probability of type II error was performed by $\beta = \int_{-\infty}^{z_A} f(x)\,dx$, where $z_A = (0.5 + n_A - (p_A^*n))/\sqrt{p_A^* p_a^* n}$, with n_A having the same value as obtained for α under the null hypothesis. Finally, power was computed by $\text{power}_{\text{TDT}} = 1 - \beta$.

Power of the mapping strategy for TDT followed by a daughter design: The mapping strategy consisted, first, of carrying out a TDT using selected offspring. The significant markers and families were then tested using records of performance. The selection at the TDT stage has two main advantages in terms of power: (1) The frequency of heterozygotes at the QTL is increased in the selected families in the t-test and (2) the multiple-testing problem might be reduced since the number of tests in the last stage is much reduced. Power of the experiment attributable to increased heterozygosity at the QTL among sire families can be computed following the approach of WELLER et al. (1990) but with an increased heterozygosity.

The frequency of heterozygotes at the QTL among selected families based on TDT results is

$$f_H^* = \frac{f_H \text{power}_{\text{TDT}}}{f_H \text{power}_{\text{TDT}} + \alpha(1 - f_H)},$$

where f_H is the frequency of heterozygous animals at the QTL before the TDT stage, which under Hardy-Weinberg is $2f_Qf_q$. The numerator is the probability of detecting a real QTL ($\text{power}_{\text{TDT}}$) multiplied by the probability of being heterozygous. The denominator also includes another term for false positives. Power was computed following the daughter design of WELLER et al. (1990) but using the appropriate frequency of heterozygotes (f_H^*) and the expected number of families to be tested using the t-test,

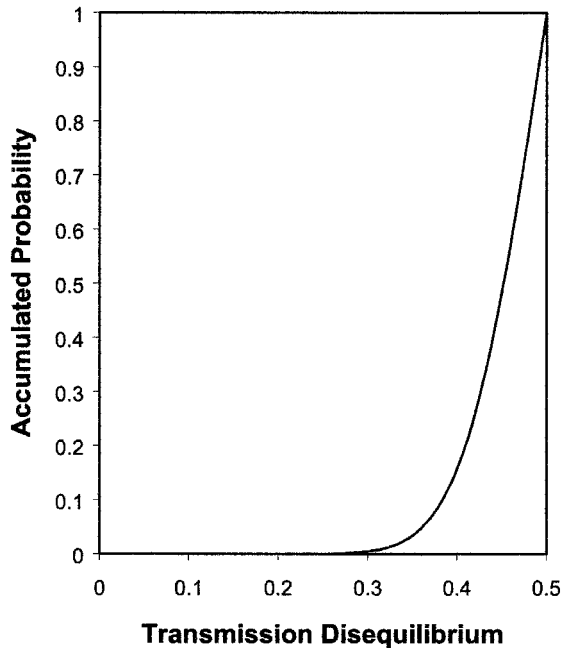$$N_f^* = N_f(f_h \text{power}_{\text{TDT}} + \alpha(1 - f_H)),$$

FIGURE 1.—Accumulated probability of obtaining a given transmission disequilibrium parameter under random sampling of 50 sons.

FIGURE 2.—Power for detecting selection by means of genetic markers using TDT. Recombination fractions (c) were 0, 0.10, and 0.20. The population for typing corresponds to a selected sample of 32.5%. The level of significance is 0.05.

where N_f is the number of families tested at the TDT stage. The expected power of the TDT followed by the daughter design (DD) was calculated by multiplying the power computed following WELLER *et al.* (1990) for each possible family size (according to a binomial with probability f_H^* and size N_i^*) by its probability and summing over all family sizes. The resulting power, referred to as TDT-DD, was compared to the power for testing with a simple daughter design, ignoring the testing at the TDT stage, with six half-sib families and 100 sons (referred to as DD).

RESULTS

The accumulated probability of random changes in the transmission disequilibrium parameter when sampling 50 sons from a heterozygous sire is shown in Figure 1. It is unlikely that the transmission parameter goes beyond the interval 0.3–0.70 when testing distorted segregation for that family size.

Statistical power for transmission disequilibrium testing of selection at a significance level of 0.05 is shown in Figure 2. Power to detect distorted segregation of alleles among offspring is >0.5 for QTL having an effect of >0.6 phenotypic standard deviations with full linkage of the marker to the QTL. Power decreases with increasing distance between the marker and the QTL. For loose linkage power is low. Therefore, a genome-wide scan should be carried out using markers separated from each other by ~10 cM. These results also indicate that the Norwegian cattle population is suitable for measuring genomic response of QTL with moderate to large effect.

It is also of interest to know the gain in power for QTL detection by using the strategy TDT-DD *vs.* a simple DD.
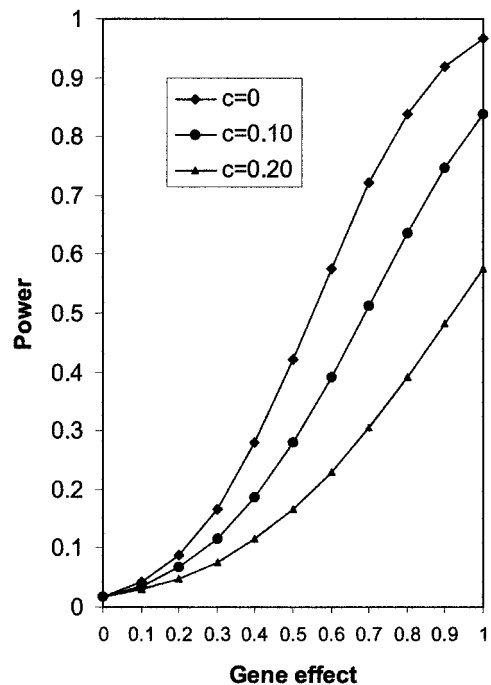
Power for the detection of QTL segregating at moderate or low frequencies increases by the selection of markers and families upon TDT results (Figure 3). On the contrary, QTL segregating at an intermediate frequency do not benefit for the strategy including pretesting with TDT (Figure 3).

Testing of transmission disequilibrium of alleles among selected sons was performed with 282 genetic markers evenly distributed over the genome. The genome-wide test was highly significant ($P < 0.001$), indicating that natural and/or artificial selection for growth and conformation induces detectable changes in the genetic makeup of commercial populations (Table 2). Chromosomes 2, 6, 16, and 21 were significant at $P < 0.01$ whereas chromosomes 1 and 18 were significant at $P < 0.05$. Values of the likelihood-ratio statistic test for markers showing significant results in the genome-wide search are depicted in Table 3. The expected number of markers showing significant results at $P < 0.01$ is ~3 because 282 markers were used, which suggests that there is an excess of significant markers that cannot be attributed to chance. However, the method cannot distinguish between transmission disequilibrium of alleles among selected offspring due to selection (*i.e.*, a marker linked to a QTL influencing growth) or chance because of the amount of multiple testing performed.

Figure 4 shows the likelihood-ratio statistic test for markers in chromosome 2 and sire family 2005 when using estimates of allele frequencies in the dam popula-
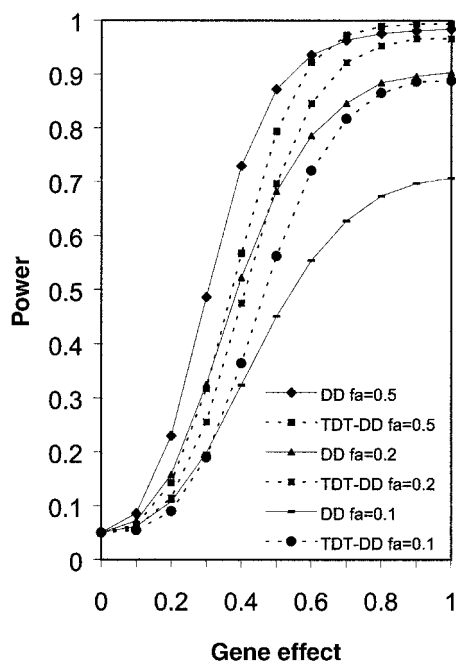
FIGURE 3.—Power of TDT followed by a daughter design *vs.* a simple daughter design using the population structure of Norwegian cattle.

TABLE 2

Likelihood-ratio statistic test (LRT) for the segregation of alleles among selected offspring for all markers within each of the 29 autosomal chromosomes

Chromosome	d.f.	LRT
1	68	97.72*
2	47	73.4**
3	45	50.5
4	51	44.28
5	81	74.5
6	42	67.42**
7	46	60.95
8	31	34.91
9	44	58.67
10	30	54.16
11	30	38.82
12	31	38.58
13	43	52.01
14	20	27.89
15	41	58.54
16	35	58.63**
17	56	72.15
18	34	48.83*
19	38	27.88
20	26	30.88
21	29	61.06**
22	31	26.06
23	44	60.16
24	34	20.88
25	36	44.04
26	35	34.03
27	35	55.20
28	30	29.53
29	15	18.63
Total	1128	1420.31***

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

tion. The chromosomal fragments near INRA40 are transmitted in disequilibria among selected young bulls ($P < 0.01$). A graphic display of the transmission parameters for this chromosome and sire family is given in Figure 5. A transmission map is constructed using all transmission parameters for all the markers after they are oriented in the two homologous chromosomes. Transmission maps illustrate how different regions of the chromosome are transmitted from sire to selected offspring.

The next step was to test for distorted segregation of alleles among nonselected offspring. Table 3 also shows the likelihood-ratio test statistic for nonselected sons. Markers INRA40 in family 2005 ($P < 0.01$), BM1237 in family 2052 ($P < 0.10$), BM1237 in family 3131 ($P < 0.05$), and BM6430 in family 2463 ($P < 0.10$) were significant. The lack of more significant results could be attributed to the weaker selection intensity among nonselected sons, which corresponded to a selected percentage of 66%.

The last step was to test for gametic selection at significant markers in step 1 using single-sperm typing. The results are also given in Table 3. Only marker BM1237 (family 2052) was significant ($P < 0.05$).

A summary of the transmission parameters in the three testing points (selected sons, nonselected sons, and sperm cells) is given in Figure 6. The transmission parameter is the frequency of transmission of one of the sire's alleles to his sons. The number in parentheses in the histogram corresponds to the allele that is highly represented among selected sons. The relation between artificial selection and the observed distorted segrega-

tion among selected and nonselected sons can be established after performing a *t*-test contrasting growth performance of sons inheriting alternative alleles from their sire (Table 4). Alleles at a high frequency among selected sons should also increase growth. Markers BM4528 in families 2402 and 2946, BM1237 in family 2052, and BM6340 in family 2463 yielded significant results ($P < 0.01$), showing increased performance for sons carrying the allele highly represented among selected sons. INRA40 in family 2005 also yielded higher performance for bulls inheriting the allele highly represented among selected sons and was close to being significant. In all cases except marker BM4528 in sire family 2402, the pattern of segregation of alleles was reversed among nonselected sons. INRA128 was not significant for the *t*-test contrasting growth performance but showed distorted segregation of alleles among selected sons. Marker BM8125 was nonsignificant. These results suggest that sampling might play an important role in the genomic changes in populations undergoing artificial selection.

Markers BM1237 in family 3131 and BM8125 in family

TABLE 3

Likelihood-ratio statistic test (LRT) using selected and nonselected half-sib sons and single-sperm typing

Chromosome	Marker	Family	Selected		Nonselected		Sperm typing	
			n	LRT	n	LRT	n	χ^2
1	INRA128	2005	70	7.72**	23	0.83	44	0.82
2	INRA40	2005	71	10.76**	21	7.80**	43	0.21
2	ETH121	2402	53	11.43**	51	0.75	50	0.08
6	BM4528	2402	54	9.86**	52	0.06	51	1.59
6	BM4528	2946	40	9.85**	46	2.32	44	0.36
7	INRA53	2463	38	11.16**	42	0.92	43	1.14
10	BM1237	2052	32	5.56*	45	3.70***	44	4.45*
10	BM1237	3131	44	7.56**	63	3.85*	44	0
16	BM6430	2463	37	9.14**	41	3.60***	41	1.19
17	BM8125	2005	70	16.76**	23	0.84	43	1.14

***$P < 0.10$; **$P < 0.01$; *$P < 0.05$. n, the size of the sample used in each analysis.

2005 were not associated with growth performance. As shown in Figure 6, the allele highly represented among selected sons was also the most abundant among nonselected sons. These results suggest that zygotic selection may have operated prior to the test for growth.

Another round of typing was carried out by (1) single-sperm typing using marker BM1237 (family 2052) with another sample of sperm cells and (2) typing for markers BM1237 (family 3131) and BM8125 (family 2005) using DNA from 48 half-sib daughters. The purpose of the first was to confirm gametic selection and of the second was to verify if natural selection has operated in that
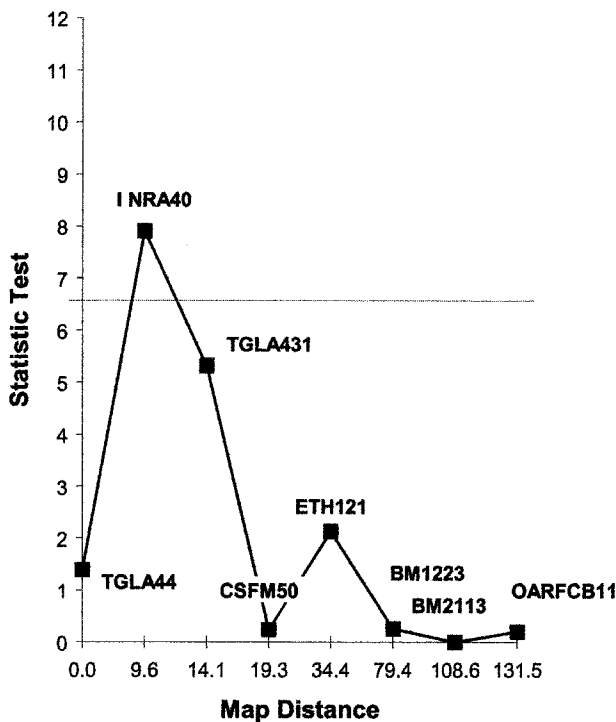

FIGURE 4.—Likelihood-ratio statistic test for markers in chromosome 2 and sire 2005 along genetics distances (in centimorgans using the Kosambi mapping function). The horizontal line represents the threshold for hypothesis testing at $P < 0.01$.
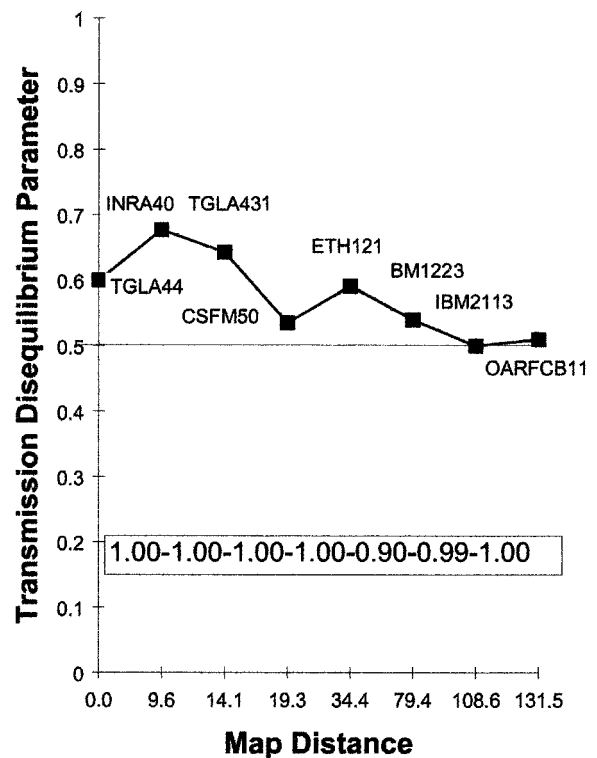

FIGURE 5.—Transmission map for chromosome 2 and sire 2005. Estimates of the transmission parameter for each marker along the genetic distance (in centimorgans using the Kosambi mapping function) were obtained assuming known allele frequencies in the dam population. The conditional probability of the most likely linkage phase given the observations for each pair of consecutive markers TGLA44, INRA40, TGLA431, CSFM50, ETH121, BM1223, BM2113, and OARFCB11 ranged from 0.90 to 1 and is shown in the box in the bottom.
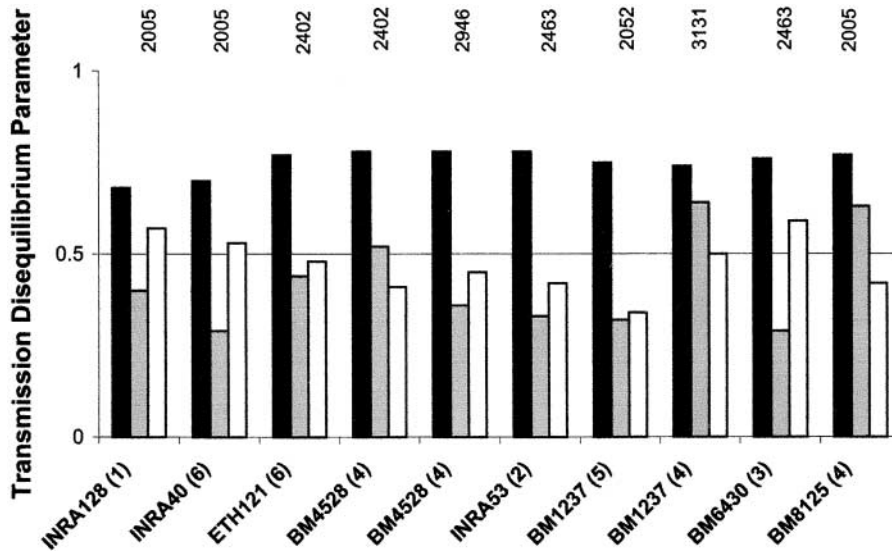
FIGURE 6.—Transmission disequilibrium parameter in selected sons, nonselected sons, and single-sperm typing. The values in parentheses are the corresponding alleles to the estimated transmission disequilibrium parameters (alleles highly represented among selected sons). Sire families are given at the top. (■) Selected, (▨) nonselected, (□) single-sperm typing.

family and marker among the female offspring. The results for either single-sperm typing or half-sib daughters were nonsignificant (χ^2 values ranging between 0.00 and 0.05).

DISCUSSION

Livestock represent a unique opportunity to follow genomic changes through the various levels of selection because of their large progeny groups. Sequential test-

ing for distorted segregation of alleles among sperm cells and selected and culled sons is proposed to establish the main components of selection in commercial livestock at the gametic, zygotic, and human-made levels. Our results indicate that gametic selection (selection of gametes before fertilization), if it exists, is not strong enough to be detected, which could be attributed to the industry testing for fertility traits before young bulls are used to inseminate the cow population.

Zygotic selection is the selection from fertilization

TABLE 4

Values of the *t*-test contrasting growth performance between sons inheriting alternative alleles from their sires

Chromosome	Marker	Family	Allele	Growth	d.f.	*t* value	EW *P* value
1	INRA128	2005	2	37.291	69	0.510	1.00
			1	30.831			
2	INRA40	2005	6	47.330	66	1.3956	0.84
			4	27.558			
2	ETH121	2402	6	0.350	86	0.421	1.00
			1	−7.287			
6	BM4528	2402	4	18.188	80	2.937**	0.04
			1	−33.284			
6	BM4528	2946	4	60.882	66	1.841***	0.51
			1	15.123			
7	INRA53	2463	1	25.479	36	0.0797	1.00
			2	23.350			
10	BM1237	2052	5	35.241	62	3.165**	0.02
			3	−20.106			
10	BM1237	3131	3	−2.130	86	0.5297	1.00
			4	−10.624			
16	BM6430	2463	3	56.738	63	3.178**	0.02
			2	−5.382			
17	BM8125	2005	1	47.877	64	1.4127	0.83
			4	25.939			

***$P < 0.10$; **$P < 0.01$. The *t*-test was carried out only for markers yielding significant results for the likelihood-ratio statistic test among selected sons. Experimentwise *P* values (EW *P* value) were computed using a permutation test.

until the animals are ready for artificial insemination. This could occur if the same allele is more frequent among selected sons and among nonselected sons. This, in fact, occurred for markers BM1237 and BM8125 in our material. Consequently, those markers could be under zygotic selection. For example, it could happen when the marker is linked to a gene reducing viability of embryos or provoking high mortality among young animals. Another explanation is sampling, in which the allele is highly represented in selected and nonselected sons just by chance. Records on how zygotic selection occurred are needed to clearly establish how selection took place.

The main conclusion of this study is that the genetic makeup of commercial cattle is being systematically modified. Not surprisingly, the main cause for these changes is human intervention. The use of records of performance allowed us to establish the main forces influencing genomic response to artificial selection. Both direct and stochastic changes in the allele frequencies are important components of the response. Regions of chromosomes 6, 10, and 16 responded directly to artificial selection for growth in commercial cattle. This knowledge brings two applications: (1) to identify QTL responsible of the observed response and (2) to monitor other known loci with an important physiological role (*e.g.*, immune response), which are located nearby those chromosomal areas responding to artificial selection. The first is comparable to today's efforts for mapping QTL in commercial cattle (*e.g.*, Georges *et al.* 1995). The second opens new possibilities for controlling genomic manipulation of farm animals.

Testing of linkage disequilibrium was performed within families. This was a necessary assumption because we used mostly microsatellites (random anonymous markers) and, therefore, linkage phase of marker and QTL alleles could be different in different sires. Consequently, population-wide linkage disequilibrium was not tested. It is clear from the numbers given in Table 1 that the contribution to the next generation of bull sires was not equal. In particular, family 2005 was highly selected (~25% of sons were selected) whereas selection pressure on family 2052 was low (~11%). Genomic response across families could be tested only if polymorphisms in coding genes or markers strongly linked to them were available. This would require a highly dense genetic map.

The research performed in this article was carried out using Norwegian cattle, which is a relatively small population. The American dairy cattle population is much larger with several elite bull sires having between 100 and 300 sons' progeny tested for dairy traits (E. B. Burnside, personal communication). The power for detection would be much higher than for European populations and would allow identification of the areas of the genome under selection for milk traits. This information could be compared with available information from QTL mapping in the dairy cattle population of America (Georges *et al.* 1995). It is a paradox that QTL of relatively large effect are segregating in the commercial dairy population since fixation at those loci after a few cycles of selection is expected (Crow and Kimura 1970; Gomez-Raya and Klemetsdal 1999). The approach proposed in this article could be used to elucidate if the chromosomal areas where QTL have been detected are also under selection pressure in the American cattle population. For example, mapped QTL in the American population, which are not modifying their allele frequencies, might provide evidence of other forces maintaining genetic variability.

During the past several years great progress has been achieved in building better comparative maps between livestock and humans/mouse (Band *et al.* 2000). Identifying breakpoints between conserved syntenic groups in different species is extremely important for the extrapolation of positional information from the highly developed human map to lower density maps in cattle. In the near future, it will be possible to identify positional candidate genes for the loci responding to artificial selection by taking advantage of this wealth of information and thus to establish the genetic nature of the genomic changes operating under selection in farm animals.

LITERATURE CITED

Band, M. R., J. H. Larson, M. Rebeiz, C. A. Green, D. W. Heyen *et al.*, 2000 An ordered comparative map of the cattle and human genomes. Genome Res. **10:** 1359–1368.

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

Crow, J. F., and M. Kimura, 1970 *An Introduction to Population Genetics Theory.* Burgess Publishing, Minneapolis.

Georges, M., D. Nielsen, M. Mackinnon, A. Mishra, R. Okimoto *et al.*, 1995 Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics **139:** 907–920.

Gomez-Raya, L., 2001 Biased estimation of the recombination fraction using half-sib families and informative offspring. Genetics **157:** 1357–1367.

Gomez-Raya, L., and G. Klemetsdal, 1999 Two-stage selection strategies utilizing marker-QTL information and individual performance. J. Anim. Sci. **77:** 2008–2018.

Li, H., U. B. Gyllensten, X. Cui, R. K. Saiki, H. A. Erlich *et al.*, 1988 Amplification and analysis of DNA sequences in single human sperm and diploid cells. Nature **335:** 414–417.

Spielman, R. S., R. E. McGinnis and W. J. G. Ewens, 1993 Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am. J. Hum. Genet. **52:** 506–516.

Våge, D. I., I. Olsaker, H. Klungland, L. Gomez-Raya and S. Lien, 2000 A male genetic map designed for QTL mapping in Norwegian cattle. Acta Agric. Scand. Sect. A Anim. Sci. **50:** 56–63.

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

Communicating editor: J. B. Walsh