

Across-Family Marker-Assisted Selection Using Selective Genotyping Strategies in Dairy Cattle Breeding Schemes

S. Ansari-Mahyari,*†‡¹ A. C. Sørensen,* M. S. Lund,* H. Thomsen,* and P. Berg*

*Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, University of Aarhus, Denmark

†Agricultural Biotechnology Research Institute of Iran, Karaj, Iran

‡Department of Large Animal Sciences, Faculty of Life Sciences, University of Copenhagen, Denmark

ABSTRACT

This study investigated the potential loss expected from marker-assisted selection (MAS) when only a proportion of animals are genotyped using several selective genotyping strategies. A population resembling a commercial dairy cattle population over 25 yr was simulated, and the most informative individuals for genotyping were identified among the potential breeding candidates (young bulls and bull-dams). Two strategies were used to identify the most informative animals. The first genotyping strategy was based on selecting individuals for genotyping with predicted total genetic effect [sum of the predicted quantitative trait locus (QTL) and polygenic effects] close to the truncation point for selection. The second strategy used an index that extended the previous strategy to include the variance due to segregation of the QTL in the parents. The 2 strategies for selective genotyping were applied at the 2 different genotyping levels and compared with random selection of candidates for genotyping and complete genotyping of the potential candidates. All selective genotyping strategies at the same proportion of genotyping showed similar cumulative genetic level. The frequency of the favorable QTL allele increased faster with more animals genotyped. Extra response in total genetic effect (polygenic and QTL) was not significantly different between genotyping all candidates (100%), 20%, and 50% genotyping (except for yr 13), but all MAS strategies resulted in significantly higher response than BLUP until yr 18. With 50% (20%) genotyping of candidates for selection within a population, 95% (89%) of maximum cumulative QTL response was achieved in yr 13. All MAS schemes resulted in a 19% decrease in the rate of inbreeding compared with the BLUP scheme. Therefore, it is possible to use selective genotyping in practical dairy cattle breeding and decrease the genotyping

costs with a minimal loss of response compared with complete genotyping of the potential candidates.

Key words: marker-assisted selection, dairy cattle, quantitative trait locus, selective genotyping

INTRODUCTION

Genetic improvements in quantitative traits during the last decades were achieved by selecting genetically superior parents using phenotypic and pedigree information through BLUP. In dairy cattle breeding programs, proven sires are usually selected after progeny testing using evaluation with an animal model (e.g., Powell and Norman, 2006). In addition, selection of superior cows as bull-dams to improve the genetic gain is an important part in breeding programs. In a dairy breeding plan, young bulls can potentially be selected for progeny testing based on information from QTL (Spelman and Garrick, 1998; Schrooten et al., 2005; Schulman and Dentine, 2005). The reason to consider detected QTL [a detected region of the chromosome that contains specific gene(s) affecting a quantitative trait] is to obtain a more accurate evaluation and increase the frequencies of favorable alleles in the population. Therefore, the genetic gain can be increased compared with BLUP by incorporating molecular information on detected QTL in the breeding schemes by marker-assisted selection (MAS). Several studies have shown the benefits of using MAS to increase the genetic merit due to an increased accuracy of genetic evaluation (Lande and Thompson, 1990; Goddard and Hayes, 2002; Villanueva et al., 2005). The MAS is expected to increase genetic gain compared with traditional breeding programs or decrease the cost of progeny testing by prescreening the potential young bulls, especially in situations where selection based on BLUP evaluations has limitations. Examples are traits that are sex-limited, expressed late in life, costly to record, or have a low heritability (Lande and Thompson, 1990; Meuwissen and Goddard, 1996). Several traits with these characteristics are included in dairy cattle breeding schemes such as fertility traits and disease resistance.

Received August 14, 2007.

Accepted December 4, 2007.

¹Corresponding author: SaeidAnsari.Mahyari@agrsci.dk

In previous studies, the selection schemes using marker information for dairy cattle were largely based on information from within families (Kashi et al., 1990; Mackinnon and Georges, 1998; Spelman and Garrick, 1998). Recently, the use of linkage disequilibrium (LD) information to locate QTL has increased (Pérez-Enciso, 2003; Meuwissen and Goddard, 2004; Uleberg and Meuwissen, 2007). The next step after fine-mapping of QTL is to use them in prediction of breeding values. There are several examples in dairy cattle with LD markers used for preselection of candidates (e.g., Dekkers, 2004).

Using DNA information in a population with LD can enhance the accuracy of identifying superior candidates. However, the cost of complete genotype information in MAS is a major limitation in a breeding scheme. Genotyping the whole population is also difficult for practical reasons in commercial dairy cattle populations. Several investigations in QTL experiments attempted to decrease genotyping costs by identifying the most informative individuals based on phenotypic information (Lander and Botstein, 1989; Darvasi and Soller, 1992), segregation analysis (Kinghorn, 1999; Macrossan and Kinghorn, 2003), or combining the phenotypic and genotypic information (Ansari-Mahyari and Berg, 2008).

The objective of this study was to compare genetic gain for several strategies of selective genotyping relative to complete genotype information under the assumption that QTL parameters are known. The 2 genotyping levels (20 and 50%) were considered and compared with genotyping all or no candidates, the latter resulting in selection being based on BLUP. This case was used as a reference. The response to selection at the QTL and polygenes for the different strategies were compared in a simulated dairy cattle breeding scheme for a low-heritability trait only expressed in females (e.g., mastitis). This trait was chosen to maximize the power to detect differences between strategies because it is a trait that is expected to favor MAS over BLUP.

MATERIALS AND METHODS

A population resembling the active breeding part of a commercial dairy cattle population was simulated. Parents were selected based on BLUP or marker-assisted (MA)-BLUP predictions. Genotypes of both potential young bulls and bull-dams were considered in MA-BLUP evaluations. The 25 yr of selection were simulated and the following results were considered: cumulative genetic responses, annual genetic gain, and accuracy of evaluation. In addition, the frequency of the positive QTL allele and the rate of inbreeding were measured.

Genetic Model

The base-population was simulated to generate LD between the markers and QTL according to Meuwissen and Goddard (2000), which assumed that variation in a QTL is due to a mutation that occurred 100 generations ago and increased to a certain frequency due to random drift, given a specific effective population size (200 in this study). Genotypes were simulated for 4 marker loci placed in an identified 4-cM region of a chromosome and positioned at 0, 1.5, 3.5, and 4 cM. Markers with 5 alleles of equal frequencies were assumed. The detected QTL was situated between the second and the third marker at position 2.5 cM. In the first generation of simulating LD, unique alleles were assigned to the QTL for the founders. At the end of simulating LD (generation 100), a founder allele with a frequency of 0.10 (± 0.01) was defined as the favorable QTL allele and the rest of the remaining founder alleles were defined to be the unfavorable QTL allele. Hence, the QTL was biallelic and the frequency of the favorable QTL allele was between 0.09 and 0.11 in the base population. Marker and QTL genotypes were assigned according to Mendelian inheritance. The simulation allowed for recombination within the region, according to the genetic distance between the loci. Haplotypes for base animals were sampled from the haplotypes present in generation 100. In this study, 50 base populations with the above characteristics were generated. Then, the same base-populations were used in all genotyping strategies described below.

True breeding values were simulated assuming a mixed inheritance model with an additive QTL effect. A sex-limited trait, like mastitis resistance, was considered with a heritability of 0.04 and a total phenotypic variance of 1.0 in the base population. The polygenic effect of each animal (a_i) in the base population was sampled from $N(0, \sigma_a^2)$, where σ_a^2 is the variance of the additive polygenic component, which was 0.03. The polygenic effect of later generations was obtained from $a_i = \frac{1}{2} a_{\text{sire}} + \frac{1}{2} a_{\text{dam}} + m_i$, where m_i is the Mendelian sampling effect, which was sampled from $N(0, \frac{1}{2}(1-f)\sigma_a^2)$, where f is the average inbreeding coefficient of the sire and dam of i .

It was assumed that the QTL had been identified in earlier QTL mapping experiments, and that the position and variance of the QTL was known with certainty. The variance due to the QTL was $\sigma_{qtl}^2 = 2p(1-p)\alpha^2$ (Falconer and Mackay, 1996), where p is the frequency of the favorable QTL allele in the base population and α is the gene substitution effect. The gene substitution effect was constant and equal to 0.236, so that σ_{qtl}^2 was approximately 0.01 in the base-population, as the initial frequency of the favorable QTL allele was ap-

Table 1. Population breeding scheme in BLUP and MA-BLUP evaluation during each year of simulation

Selection stage ¹	Criteria	Preselection (for genotyping)	Selection	Age (year) of reproduction
Young bulls ²	BLUP	—	200	1
	MA-BLUP	400	200	1
Proven bulls	BLUP	—	10	5–8
	MA-BLUP	—	10	5–8
Bull dams	BLUP	—	1,000	1–5
	MA-BLUP	2,000 ³	1,000	1–5
Cow dams	BLUP	—	3,000	1–5
	MA-BLUP	—	3,000	1–5

¹The candidates of young bulls, proven bulls and bull dams were selected across herds but selection in the candidates of cow dams were within herds.

²The bulls were selected at yr 1 and then selected bulls were progeny tested to identify the proven sires.

³This was the maximum number for genotyping in bull dams.

proximately 0.10. Given that the QTL allele frequency changed, the total genetic variance changed over time.

Population Structure

The numbers of selected candidates in the scenarios are shown in Table 1. After simulating the base-population with LD, 40 herds of equal size (100 cows per herd) were generated. The population size was kept constant.

Selection of Males. In the BLUP scenario, 200 young bulls out of approximately 500 male offspring of 1,000 bull dams were selected at yr 1, based on the average EBV of the parents. In the MA-BLUP scenarios, 400 bull calves were first preselected as candidates for genotyping, according to the average of the parents EBV. Depending on the selective genotyping strategy (described below), all or a fraction of the candidates were genotyped. Then after realizing the genotypic information and redoing the MA-BLUP evaluation, 200 young bulls were selected for progeny testing based on EBV. In both BLUP and MA-BLUP scenarios the selected young bulls were progeny tested with 100 daughters each, and the best 10 progeny-tested bulls across age groups were selected as proven sires.

Selection of Females. In the BLUP approach, 4,000 cows were selected based on EBV including own performance and offspring information and the best 1,000 cows were chosen as bull dams. In the MA-BLUP approach, potential bull dams (maximum 2,000 cows) were first selected for genotyping based on EBV. Again, depending on the selective genotyping strategies, all or a fraction of them were genotyped and MA-BLUP evaluation was performed including the new genotype information. Ultimately, the best 1,000 cows were selected as bull dams. In both BLUP and MA-BLUP, an-

other 3,000 cows were selected, based on the last evaluation to keep the population size constant.

Genotyping Strategies

Several selective genotyping strategies were used under the MA-BLUP approach. In the first year of genotyping in this experiment (yr 4), it was assumed that all founder sires, which were selected for progeny testing, had already been genotyped. Additionally, from the fourth year onward, the sires of animals in the evaluation were assumed to have marker information. The main objective of the selective genotyping strategies was to maximize genetic gain given a reduced number of genotyped animals. Each strategy was replicated 50 times, from 50 different base populations. The criteria in the current study were

1. Complete Genotyping (ALL). Each year, all 400 young bulls that were candidates to be progeny tested and 2,000 cows that were candidates to be bull dams were genotyped. These individuals were identified based on EBV from an MA-BLUP evaluation.

2. Random Genotyping (RAN). A random proportion (20 and 50%) of the 400 young bulls and 2,000 cows were genotyped.

3. Average EBV (AVE). A fraction of the 400 young bulls and 2,000 cows were selected for genotyping based on the estimated QTL and polygenic effects. Animals with EBV closest to the truncation point were chosen for genotyping. This strategy aimed at genotyping the animals most likely to change rank relative to the truncation point.

After an MA-BLUP evaluation, the overall haplotype effect was computed for each candidate as $\hat{Q}_{off} =$

$\frac{1}{2}(\hat{Q}_{S1} + \hat{Q}_{S2}) + \frac{1}{2}(\hat{Q}_{D1} + \hat{Q}_{D2})$, where \hat{Q}_{off} is the total predicted QTL effect in the candidate, and \hat{Q}_{S1} and \hat{Q}_{S2} (\hat{Q}_{D1} and \hat{Q}_{D2}) are the 2 estimated QTL effects in the sire (dam) of the candidate. If one or both of the parents were not included in the evaluation, the average haplotype effects of the grandparents were used or in case they were also not in the evaluation, the expected effects of the parental haplotypes were computed as $\hat{Q}_{parent} = 2 \times [(P_Q \times \alpha_Q) + (P_q \times \alpha_q)]$, where \hat{Q}_{parent} is the sum of the parental haplotype effects, P_Q (P_q) is allele frequency of the favorable (unfavorable) QTL-allele in the birth year of the nongenotyped parent, and α_Q (α_q) is the gene substitution effect of the favorable (unfavorable) QTL allele. In this case, \hat{Q}_{parent} was used for $\hat{Q}_{S1} + \hat{Q}_{S2}$ ($\hat{Q}_{D1} + \hat{Q}_{D2}$) if the sire (dam) was not included in the evaluation. The total genetic value in the offspring was computed as: $\hat{A}_{off} = \hat{Q}_{off} + \frac{1}{2}(\hat{a}_D + \hat{a}_S)$. Given the truncation point (t) among preselected animals, an index for each candidate was calculated to compute the absolute deviation as $Index = |t - \hat{A}_{off}|$.

A proportion (20 and 50%) of candidates with the smallest index values was genotyped.

4. Combined Average-EBV and Variance-QTL (COMBINE). Whereas the previous strategy used the sum of the estimated QTL and polygenic effects, this strategy also used the segregation variance of the detected QTL. The rationale for this strategy followed from the assumption that the Mendelian sampling variance of the potential candidates depends on the parental genotypes, therefore affecting the probability that genotyping may cause a change of the total genotypic effect relative to the truncation point. The EBV (\hat{A}_{off}) were computed as in the previous strategy. Mendelian sampling variance in the preselected candidates was calculated as follows. Each sire (dam) has 2 estimated haplotype effects as \hat{Q}_{S1} and \hat{Q}_{S2} (\hat{Q}_{D1} and \hat{Q}_{D2}). There are 4 possible combinations of the QTL effects in offspring with equal probability as below:

$$\hat{Q}_{off1} = \hat{Q}_{S1} + \hat{Q}_{D1} \quad \hat{Q}_{off2} = \hat{Q}_{S1} + \hat{Q}_{D2};$$

$$\hat{Q}_{off3} = \hat{Q}_{S2} + \hat{Q}_{D1} \quad \hat{Q}_{off4} = \hat{Q}_{S2} + \hat{Q}_{D2}.$$

Mendelian sampling variance was calculated for all preselected candidates by

$$MenVar_{\hat{Q}_{off}} = \frac{1}{4} \left[\sum_{i=1}^4 (\hat{Q}_i - \hat{Q}_{off})^2 \right],$$

where \hat{Q}_{off} is the mean predicted QTL effect in the offspring $\left(\frac{\sum_{i=1}^4 \hat{Q}_{off,i}}{4} \right)$. Again, if one of the parents or both

were not genotyped, the mean of the haplotype effects were computed as previously described. However, in the situation neither of the parents were evaluated, the average haplotype effects of the grandparents were used. If grandparents were not in the evaluation, the effects in each parent were computed as $\hat{Q}_1 = \frac{1}{2}\hat{Q}_{parent} + \frac{1}{2}\delta$, for the first haplotype effect and $\hat{Q}_2 = \frac{1}{2}\hat{Q}_{parent} - \frac{1}{2}\delta$, for the second haplotype. The δ was computed as the mean of absolute differences of the estimated haplotype effects ($\hat{Q}_{1,i}$ and $\hat{Q}_{2,i}$) in all animals ($i = 1 \dots N$), which were included in the MA-BLUP evaluation by

$$\delta = \frac{\sum_{i=1}^N |\hat{Q}_{1,i} - \hat{Q}_{2,i}|}{N}.$$

The δ was used to compute a more accurate Mendelian segregation variance for ungenotyped parents without affecting \hat{Q}_{off} . Finally, an index was calculated to identify candidates for genotyping as $Index = |P - 0.5|$, where $P = \Phi(\hat{A}_{off}, MenVar_{\hat{Q}_{off}}, t)$. $\Phi(\hat{A}_{off}, MenVar_{\hat{Q}_{off}}, t)$ is the cumulative normal distribution function with mean \hat{A}_{off} and variance $MenVar_{\hat{Q}_{off}}$, which is evaluated at the truncation point. Therefore, P is the probability of an animal being below the truncation point and thus not being selected. Either 20 or 50% of the individuals with the lowest value of the index were genotyped. These were the animals likely to change order relative to the truncation point when genotyped and reevaluated.

Breeding Value Evaluation

The MA-BLUP evaluation was used for selection based on phenotypic, pedigree, and marker information. For progeny tested bulls, daughter yield deviations (DYD) were sampled directly. Twice the DYD was used as the phenotype of the sires in genetic evaluations. During the first 3 yr of all genotyping strategies, animals were selected using BLUP evaluations without genotypic information of the parents.

The MA-BLUP evaluation was based on the following mixed model:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{a} + \mathbf{W}\mathbf{q} + \mathbf{e},$$

where \mathbf{y} = vector of records for cows or $2 \times$ DYD for progeny-tested bulls, \mathbf{X} = incidence matrix relating phenotypes to fixed effects, β = vector of fixed herd-year-season effects, \mathbf{Z} = incidence matrix relating phenotypes to animals, \mathbf{a} = vector of additive polygenic effects for each animal, \mathbf{W} = incidence matrix relating phenotypes to random QTL effects, \mathbf{q} = vector of random additive effects due to the QTL (haplotype effect), and \mathbf{e} = vector of residuals. The QTL was fitted as a random effect because QTL genotypes (alleles) in the animals are unknown. The random vectors are assumed to be normally distributed and mutually independent. Specifically, \mathbf{a} is $N(0, \mathbf{A}\sigma_a^2)$, \mathbf{q} is $N(0, \mathbf{H}_p\sigma_q^2)$, and \mathbf{e} is $N(0, \mathbf{R}\sigma_e^2)$, where \mathbf{A} is the additive relationship matrix, \mathbf{H}_p is the identity by descent (IBD) matrix that contains the IBD probabilities for the QTL position (2.5 cM in this study), and \mathbf{R} is a diagonal matrix with elements of 1 for cows and $\frac{1}{\text{weight}}$ for the bulls with DYD. The weight was calculated as the variance of twice the DYD. The IBD matrix was derived with assumption about historical population structure and clustering approach for LDLA analysis (combined linkage disequilibrium and linkage analysis) across sire family groups using the method described by Meuwissen and Goddard (2000, 2001). The variance component used for prediction of QTL effects was recalculated every year based on the frequency of the QTL alleles.

For the BLUP evaluation, the QTL effect was excluded from the model and the evaluation was based on phenotypic and pedigree information. The total EBV for BLUP and MA-BLUP evaluations were \hat{a} and $\hat{a} + 2\hat{q}$, respectively. Because computation of the gametic relationship matrix was time consuming, data used in the genetic evaluation were constrained to the final 15 yr.

The results of genetic responses, QTL frequency, and rate of inbreeding from the strategies in each year were tested using standard ANOVA with the genotyping strategy method and replicate as 2 fixed independent variables. Replicate was included in the model, as the same 50 base populations were used.

RESULTS

Figure 1 shows that the cumulative genetic response due to the QTL in MAS schemes with the 2 levels of selective genotyping (20 and 50%) was greater than for BLUP. In yr 7, the gain was slightly higher than the ALL strategy. However, the cumulative genetic responses of the all genotyping levels in this study were not significantly different in yr 5 to 8 for the 3 MA-BLUP

scenarios. The responses were significantly higher in the ALL strategy compared with 50% genotyping in yr 12 and 13 (Figure 1). At 20% genotyping, the cumulative response was not significantly different from the ALL strategy in yr 18 and onward. In the BLUP scheme, the annual genetic gain due to the QTL was higher from yr 14 and onward (results not shown), since the variance of the QTL was higher in the BLUP scheme compared with MAS schemes due to slower fixation of the QTL. Until yr 23, the difference in cumulative genetic response between MAS (with different genotyping levels) and BLUP selection was significant ($P < 0.05$). In yr 24 and 25, no significant differences were observed.

As expected, the frequency of the favorable QTL allele with both BLUP and MA-BLUP selection schemes increased over time but faster with MAS (Figure 2a). All selective genotyping strategies (AVE, COMBINE, and RAN) behaved similarly at each genotyping level and the genetic response was not significantly different between the strategies at any time (Figure 2b). However, the genetic gain at the detected QTL was obtained faster at higher genotyping levels as expected. The maximum difference in QTL allele frequency between genotyping all potential candidates and 50% genotyping of the candidates was 7.6% at yr 11. The scenario ALL and other selective genotyping strategies at 50% genotyping showed frequencies close to fixation (>0.90) in yr 14 and 15, respectively (Figure 2a).

The cumulative values of the total genetic response and the polygenic response relative to the ALL strategy are shown in Figure 3. The total genetic responses (Figure 3a) using marker information with genotyping levels of 20 and 50% were higher in yr 6 to 8 than for the ALL strategy, which was due to extra polygenic response compared with the ALL strategy (Figure 3b). Selection with traditional BLUP resulted in lower total genetic response compared with other MAS schemes until yr 20 (Figure 3a). Total genetic response with 20% genotyping was not significantly different from 50% genotyping in yr 18 and onward. All selective genotyping strategies showed similar total genetic responses at the same genotyping level. The BLUP procedure showed greater annual genetic response than MA-BLUP strategies from yr 14 due to the QTL still not being fixed (results not shown). The maximum superiority in cumulated total genetic response of MA-BLUP strategies over BLUP was achieved in yr 13. Additionally, for the genotyping levels (20 and 50%) the annual genetic gain in most years was larger than for the BLUP scenario until yr 20. However, there were only small differences (nonsignificant) in annual genetic gain between the genotyping scenarios at each genotyping level (results not shown).

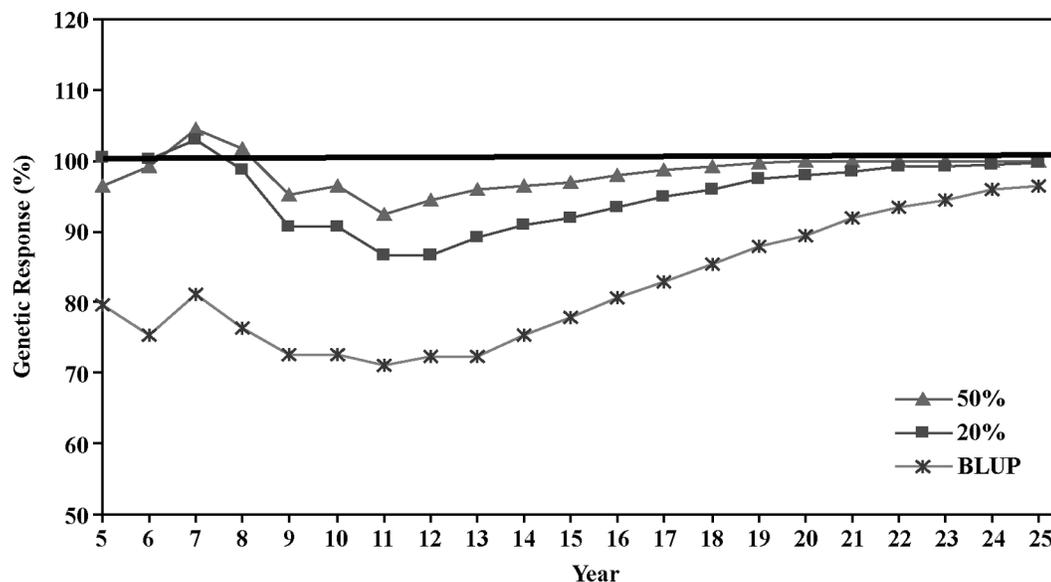


Figure 1. Percentage of cumulative genetic response at the detected QTL in comparison to the complete genotyping (ALL) strategy (the straight line), using BLUP and marker-assisted (MA)-BLUP with different genotyping levels (50 and 20% of the potential candidates using the COMBINE strategy).

The cumulative polygenic response using BLUP evaluations was significantly higher than for MA-BLUP schemes until yr 18 (Figure 3b). The annual polygenic response (results not shown) in BLUP was lower than MA-BLUP from yr 15 where the QTL was close to fixation with MA-BLUP. The QTL was close to fixation with BLUP in yr 24 (Figures 2a and 3b). All selective genotyping strategies led to very similar polygenic responses at a given level of genotyping.

Both genotyping levels (20 and 50%) in the candidates (young bulls and bull dams) were studied to compare the genetic gain achieved relative to the maximum gain obtainable from genotyping all candidates. As explained in Materials and Methods, in each year of selection regardless of the genotyping criteria, the number of genotyped candidates was based on the genotyping levels (20, 50, or 100%) and only nongenotyped animals were candidates for genotyping. Therefore, in the ALL strategy the numbers of genotyped heifers and cows were in most cases fewer than planned as some were already genotyped in a previous year. Figure 4 shows the percentage of cumulative numbers of genotyped animals (males, females, and both sexes) in genotyping levels of 20 and 50% of the potential candidates relative to the ALL strategy. In the first year of MAS (yr 4), the number of genotyped males was greater because of genotyping of progeny-tested bulls from previous years. On the other hand, the proportion of genotyping in females showed that more than 20 and 50% of the candidates were genotyped in this study in comparison to

the ALL strategy until yr 8. This was because more ancestors needed to be genotyped with the identified genotyping levels. Consequently, on average in this study, the actual genotyping proportions relative to the ALL strategy were 32 to 35% for the 20% genotyping level and 72 to 75% for the 50% level.

The inbreeding coefficients until yr 7 were less than 0.60% for all selection schemes and not significantly different until yr 10 (Figure 5). The rate of inbreeding in each year (Δf) was 0.72% in BLUP, whereas the MA-BLUP strategies (with 100% genotyping) resulted in a rate of inbreeding of 0.58%. Genotyping 20% of the potential candidates showed a slightly greater rate of inbreeding than 50 and 100% genotyping during yr 11 to 19 (Figure 5). After fixation of the QTL (Figure 2), Δf with MA-BLUP increased to a value similar to the rate of inbreeding with BLUP.

DISCUSSION

In this study, several selective genotyping strategies were compared to investigate the potential loss of genetic gain when only a fraction of the candidates are genotyped. These strategies were designed to maximize genetic gain by identifying and genotyping the animals that contribute most information in a population with population-wide linkage disequilibrium (**LD**) between markers and a QTL. The cumulative responses due to the QTL increased at higher levels of genotyping under all selective genotyping strategies. Allele frequency of

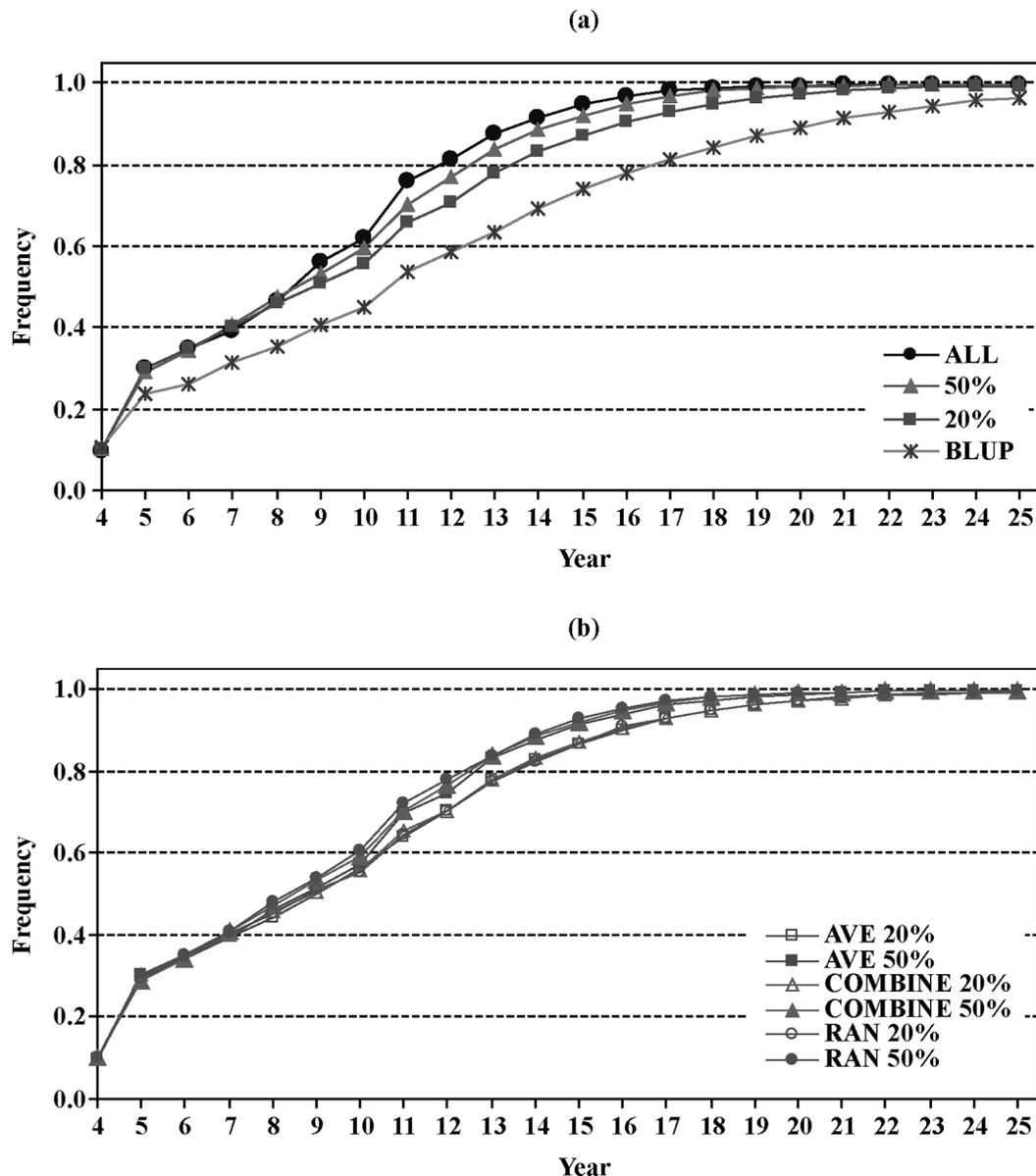


Figure 2. Frequency of the positive QTL allele: a) BLUP and marker-assisted (MA)-BLUP with 3 levels of genotyping [100%, 50 and 20% of the potential candidates each year of selection using the combined average-EBV and variance-QTL (COMBINE) strategy], b) genotyping of the potential candidates under different selective genotyping criteria [average EBV (AVE), COMBINE, and random genotyping (RAN)] using different genotyping levels (50 and 20%).

the favorable QTL surpassed 0.90 in yr 16 when 20% of the candidates were genotyped and in yr 14 for MA-BLUP selection with complete genotype information (Figure 2a).

Three selective genotyping criteria were used to choose animals for genotyping from the potential candidates each year. The AVE and COMBINE strategies chose animals with the highest uncertainty on whether to be selected or not. In the RAN strategy, the animals were sampled randomly from the preselected group. All

proposed genotyping strategies showed similar cumulative genetic level each year, at similar proportion of genotyping. The computational time demand was not different between the genotyping strategies using the same genotyping level and, therefore, any of them can be used in a MAS scheme based on selective genotyping.

Increased initial response from MAS is mostly due to increased frequency of the favorable allele and strong emphasis on the detected QTL (Spelman and Garrick, 1997). However, in practical breeding schemes, more

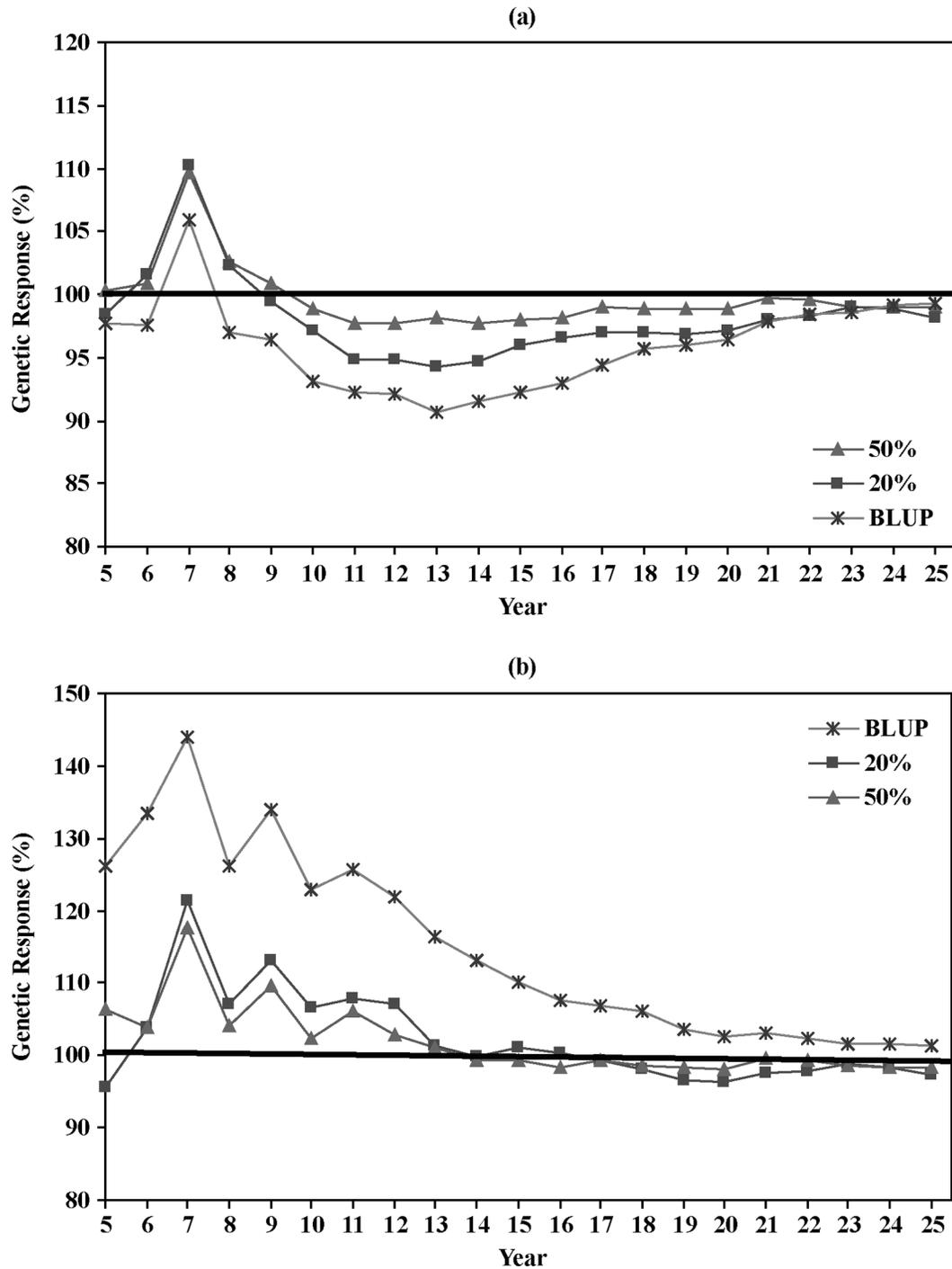


Figure 3. Percentage of cumulative genetic response in comparison to the complete genotyping (ALL) strategy (the straight line), using BLUP and marker-assisted (MA)-BLUP with different genotyping levels [50 and 20% of the candidates using the combined average-EBV and variance-QTL (COMBINE) strategy]: a) total genetic response (polygenic + QTL), b) polygenic response.

than one trait or QTL is used and therefore a slower fixation of the favorable QTL allele is expected. The strategies for selective genotyping presented in this study are easily extendable to multiple QTL and multiple traits.

In this study, selective genotyping criteria were applied for both males and females across families. However, other studies of multistage MAS in dairy cattle considered only sires for genotyping (e.g., Schrooten et al., 2005). One could consider using the genotyping

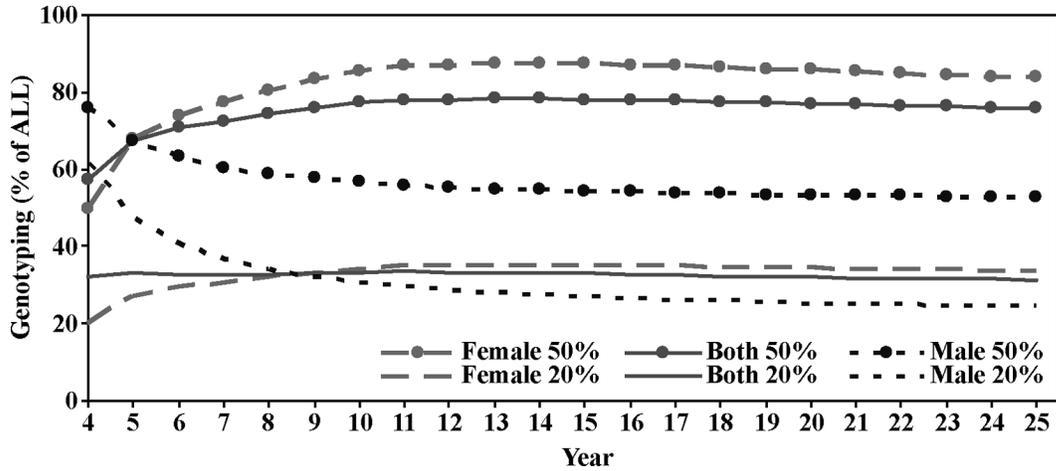


Figure 4. Percentage genotyped animals relative to the complete genotyping (ALL) strategy at 20 and 50% genotyping levels based on the combined average-EBV and variance-QTL (COMBINE) strategy for males, females, and both sexes.

criteria developed in this study in the sires only. It is expected that selective genotyping of males is less superior to BLUP because associations between phenotypes and haplotypes come exclusively from older, progeny-tested sires and not from animals being both phenotyped and genotyped.

Cumulative polygenic response using BLUP evaluations were higher than for MA-BLUP schemes. However, extra polygenic response in BLUP decreased considerably compared with MA-BLUP from yr 13, when the QTL was close to fixation with MA-BLUP selection.

Other studies showed less polygenic response to selection based on pedigree and marker information compared with pedigree information only (e.g., Spelman and Garrick, 1997). In contrast, Sonesson (2007) used within-family MAS and showed that polygenic genetic gain was not significantly different when using marker information.

Increased gain in MAS was due to more information in genetic evaluations compared with BLUP (e.g., Villanueva et al., 2005), which results in increased accuracy of predicted breeding values. The accuracy of pre-

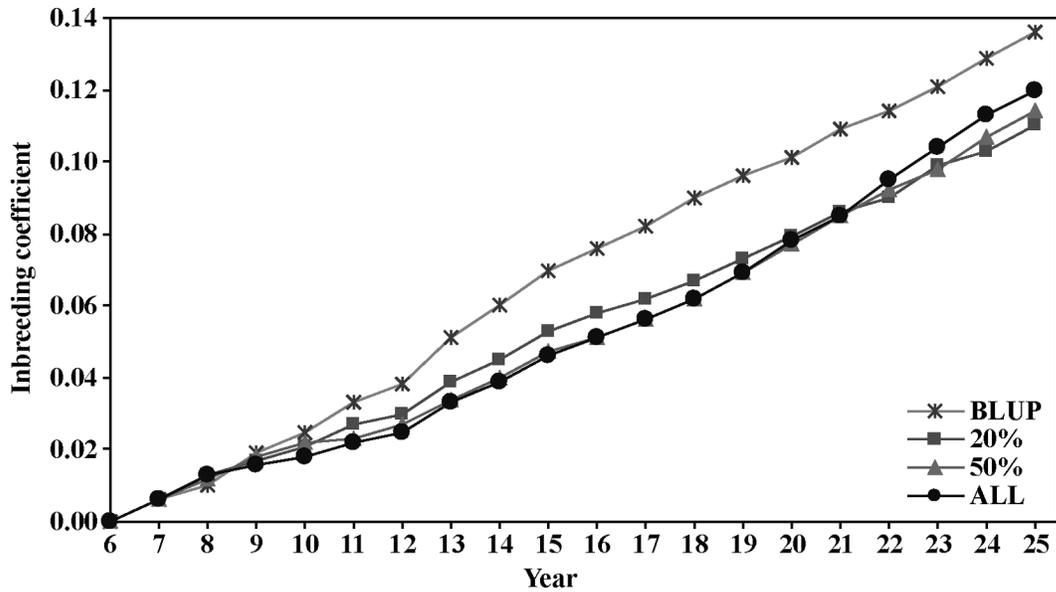


Figure 5. Inbreeding coefficient vs. birth year resulting from BLUP and marker-assisted selection using 3 levels of genotyping [100, 50, and 20% of the potential candidates using the combined average-EBV and variance-QTL (COMBINE) strategy].

dicted QTL effects increased during the first years of selection because the increasing number of genotyped offspring gave more information per haplotype. Accuracies for 50% of the potential candidates being genotyped were very close to the accuracies in the ALL strategy, whereas accuracies were lower with 20% of candidates genotyped. In BLUP and MA-BLUP scenarios, the accuracy decreased over time because of decreasing genetic variance due to the Bulmer effect (Bulmer, 1971), inbreeding, and reduced heterozygosity at the QTL (after yr 7 to 10).

Previous simulation studies investigated the benefit of using markers linked to known genes in dairy cattle breeding schemes (Meuwissen and van Arendonk, 1992; Stella et al., 2002; Schrooten et al., 2005) and showed higher genetic response. The genetic progress increased with 5 to 64% compared with BLUP selection, depending on the trait being selected and the genetic models used (Meuwissen and Goddard, 1996; Hayes and Goddard, 2003). The MAS studies assumed complete genotype information, but genotyping the whole population or even the potential candidates alone is expensive and not a practical approach for dairy cattle breeding (Dekkers and Van der Werf, 2007). For reasons mainly related to the genotyping costs, dairy cattle breeding programs based on MAS should identify and genotype the most informative breeding animals in multistage breeding programs. This simulation study attempted to clarify the usefulness of a proper genotyping criterion in a real situation of a dairy cattle breeding scheme.

Marshall et al. (2002) used several genotyping strategies and assumed an advanced breeding scheme (e.g., closed nucleus system) with complete information on pedigree and phenotypes. Based on the predicted marker genotypes with reasonable certainty (Kinghorn, 1999), Marshall et al. (2002) used an index based on genotype information with a single linked marker and EBV of 2 traits to select a fraction of the population for genotyping. Marshall et al. indicated that the gain at a detected QTL with selective genotyping was close to the gain with complete genotyping. However, in more extensive breeding schemes (e.g., sheep and beef), there is incomplete recording of pedigree relationships. In addition, in this approach it is more complicated to compute the genotype probabilities when several markers or complex pedigrees are considered. The method developed by Kinghorn (1999) to find the most informative individuals for genotyping does not take into account the phenotypic information and only uses the results from segregation analysis.

In dairy cattle breeding, it is necessary to develop more flexible selective genotyping methods for existing outbred populations. The results showed that the selec-

tive genotyping approaches could replace complete genotyping, with a minimal loss in genetic progress. In addition, the considered methods were quite simple and used the results from the previous MAS evaluation. With 20% genotyping, the total cumulative genetic level per year was not significantly different from 50% genotyping, but 20% genotyping was significantly lower than complete genotyping in yr 11 to 16. With a genotyping level of 50% (20%), 95% (89%) of the maximum response at the QTL was obtained in yr 13.

The MAS schemes in dairy cattle in France and Germany, which are based on within family selection, have been described in the literature (Weller, 2007). These schemes use linkage information and need to genotype either all young bulls and bull dams or all sires and their daughters in the German and French breeding plans, respectively. The use of LD to locate genes that affect quantitative traits has increased recently (Meuwissen and Goddard, 2004; Uleberg and Meuwissen, 2007). The classical linkage and LD analysis are complementary, and combined methods could simultaneously use linkage and LD information. Therefore, to increase the accuracy of the selection based on fine-mapped QTL and decrease genotyping costs in breeding programs, it is necessary to develop MAS across families in a population using selective genotyping methods. This study indicated that one method to decrease genotyping costs in a MAS scheme across families was genotyping the candidates close to the truncation point, when QTL effect estimation was based on both LD and linkage analysis information. Therefore, the markers that are close enough to the causative mutation and have consistent associations across families can be used in animal breeding programs (Dekkers, 2004). In addition, selection across families does not require any specific family structure (Dekkers, 2004). However, the benefits of implementing marker-linked QTL, using selective genotyping or in general for MAS, depend on the net result of extra costs, savings, and amount of genetic progress of the chosen breeding program (Schrooten et al., 2005).

Another benefit of MAS schemes over BLUP in this study was a reduction in the rate of inbreeding, as using additional marker information allows selection of high-ranking animals within families due to reduced covariances between EBV of genotyped animals within families. In a practical breeding scheme, restricting Δf can be achieved through selecting parents from more families [e.g., optimum contribution selection for MAS (Villanueva et al., 2002)]. Potentially the attempts to restrict Δf in MAS is expected to increase genetic progress as has been shown for BLUP selection (Villanueva et al., 2002).

The markers in this study were assumed in LD with the QTL. Additionally, the developed genotyping criteria used LD information to find the most informative animals for genotyping across families. Using dense markers in LD with a QTL (as a haplotype), such that the marker QTL associations persist across families in the population, makes the implementation of MAS much simpler, and the linkage phase for each sire family does not have to be established (Hayes and Goddard, 2003). Exceptional marker information content can be obtained using a dense marker map of the QTL region. Therefore, the haplotypes with the strongest link to the quantitative trait can be detected. Consequently, the most valuable haplotypes are determined not only by proximity to the QTL, but also by the relative information content and degree of linkage disequilibrium between the markers and QTL alleles.

In this study, LD between markers and QTL was simulated in the base populations. In the first years of selection based on marker information, the frequency of the favorable QTL allele was low (~0.10). In MAS schemes, the breeding animals are selected based on estimated effects of the QTL and polygenic effects, and therefore during the first years of selection, this could result in lower polygenic gain in MAS procedures in comparison to BLUP selection, where selection depends on EBV not directly accounting for the QTL. Consequently, total genetic gain was higher with MA-BLUP compared with BLUP evaluation.

CONCLUSIONS

Selective genotyping can be implemented by using any of the criteria in this study as a preselection step when selecting young bulls for progeny testing or bull dams. Marker-assisted selection schemes using LD markers increased genetic response at a detected QTL as well as the total genetic gain and decreased the inbreeding rate compared with traditional BLUP. By genotyping the more informative candidates for selection, a reduction in genotyping costs can be achieved with a minimal loss of response compared with complete genotyping and reduced rates of inbreeding compared with traditional BLUP.

ACKNOWLEDGMENTS

This research was financed by a grant from Agriculture Research and Education Organization of Iranian Ministry of Jihad-e-Agriculture and "DNA based selection to improve disease resistance, fertility, calf survival and production in Danish dairy cattle" from the Danish Directorate for Food, Fisheries and Agri Business j. no. 3401-04-00853.

REFERENCES

- Ansari-Mahyari, S., and P. Berg. 2008. Combined use of phenotypic and genotypic information in sampling animals for genotyping in detection of quantitative trait loci. *J. Anim. Breed. Genet.* (accepted)
- Bulmer, M. G. 1971. The effect of selection on genetic variability. *Am. Nat.* 105:201–211.
- Darvasi, A., and M. Soller. 1992. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. *Theor. Appl. Genet.* 85:353–359.
- 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., and J. H. J. Van der Werf. 2007. Strategies, limitations, and opportunities for marker-assisted selection in livestock. Pages 165–228 in *Marker-Assisted Selection: Current Status and Future Perspectives in Crops, Livestock, Forestry and Fish.* E. P. Guimarães, J. Ruane, B. D. Scherf, A. Sonnino, and J. D. Dargie, ed. FAO, Rome, Italy.
- Falconer, D. S., and T. F. C. Mackay. 1996. *Introduction to Quantitative Genetics.* 4th ed. Longman Group, Essex, UK.
- Goddard, M. E., and B. J. Hayes. 2002. Optimisation of response using molecular data. *Proc. 7th World Congr. Genet. Appl. Livest. Prod., Montpellier, France* 33:3–10. Hayes, B., and M. E. Goddard. 2003. Evaluation of MAS in pig enterprises. *Livest. Prod. Sci.* 81:197–211.
- Kashi, Y., E. Hallerman, and M. Soller. 1990. Marker assisted selection of candidate bulls for progeny testing programs. *Anim. Prod.* 52:21–31.
- Kinghorn, B. P. 1999. Use of segregation analysis to reduce genotyping costs. *J. Anim. Breed. Genet.* 116:175–180.
- Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124:743–756.
- Lander, E., and D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185–199.
- Mackinnon, M. J., and M. A. J. Georges. 1998. Marker-assisted preselection of young dairy sires prior to progeny testing. *Livest. Prod. Sci.* 54:229–250.
- Macrossan, P. E., and B. P. Kinghorn. 2003. Cyclic genotyping strategies. I. A comparison of ranking criteria. *J. Anim. Breed. Genet.* 120:303–311.
- Marshall, K., J. Henshall, and J. H. J. van der Werf. 2002. Response from marker-assisted selection when various proportions of animals are marker typed: A multiple trait simulation study relevant to the sheep meat industry. *Anim. Sci.* 74:223–232.
- Meuwissen, T. H. E., and M. E. Goddard. 1996. The use of marker-haplotypes in animal breeding schemes. *Genet. Sel. Evol.* 28:161–176.
- Meuwissen, T. H. E., and M. E. Goddard. 2000. Fine scale mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci. *Genetics* 155:421–430.
- Meuwissen, T. H. E., and M. E. Goddard. 2001. Prediction of identity by descent probabilities from marker-haplotypes. *Genet. Sel. Evol.* 33:605–634.
- Meuwissen, T. H. E., and M. E. Goddard. 2004. Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multi-trait data. *Genet. Sel. Evol.* 36:261–279.
- Meuwissen, T. H. E., and J. A. M. van Arendonk. 1992. Potential improvements in rate of genetic gain from marker assisted selection in dairy cattle breeding schemes. *J. Dairy Sci.* 75:1651–1659.
- Pérez-Enciso, M. 2003. Fine mapping of complex trait genes combining pedigree and linkage disequilibrium information: A Bayesian unified framework. *Genetics* 163:1497–1510.
- Powell, R. L., and H. D. Norman. 2006. Major advances in genetic evaluation techniques. *J. Dairy Sci.* 89:1337–1348.
- Schrooten, C., H. Bovenhuis, J. A. M. van Arendonk, and P. Bijma. 2005. Genetic progress in multistage dairy cattle breeding schemes using genetic markers. *J. Dairy Sci.* 88:1569–1581.

- Schulman, N. F., and M. R. Dentine. 2005. Linkage disequilibrium and selection response in two-stage marker-assisted selection of dairy cattle over several generations. *J. Anim. Breed. Genet.* 122:110–116.
- Sonesson, A. K. 2007. Within-family marker-assisted selection for aquaculture species. *Genet. Sel. Evol.* 39:301–317.
- Spelman, R. J., and D. J. Garrick. 1997. Utilisation of marker assisted selection in a commercial dairy cow population. *Livest. Prod. Sci.* 47:139–147.
- Spelman, R. J., and D. J. Garrick. 1998. Genetic and economic responses for within-family marker-assisted selection in dairy cattle breeding schemes. *J. Dairy Sci.* 81:2942–2950.
- Stella, A., M. M. Lohuis, G. Pagnacco, and G. B. Jansen. 2002. Strategies for continual application of marker-assisted selection in an open nucleus population. *J. Dairy Sci.* 85:2358–2367.
- Uleberg, E., and T. H. E. Meuwissen. 2007. Fine mapping of multiple QTL using combined linkage and linkage disequilibrium mapping—A comparison of single QTL and multi QTL methods. *Genet. Sel. Evol.* 39:285–299.
- Villanueva, B., R. Pong-Wong, J. Fernández, and M. A. Toro. 2005. Benefits from marker-assisted selection under an additive polygenic genetic model. *J. Anim. Sci.* 83:1747–1752.
- Villanueva, B., R. Pong-Wong, and J. A. Woolliams. 2002. Marker assisted selection with optimised contributions of the candidates to selection. *Genet. Sel. Evol.* 34:679–703.
- Weller, J. L. 2007. Marker-assisted selection in dairy cattle. In *Marker-Assisted Selection: Current Status and Future Perspectives in Crops, Livestock, Forestry and Fish.* E. P. Guimarães, J. Ruane, B. D. Scherf, A. Sonnino, and J. D. Dargie, ed. FAO, Rome, Italy.