Proc. Assoc. Advmt. Anim. Breed. Genet. Vol 13

BREEDING IN THE 21ST CENTURY: APPLICATION OF MOLECULAR TECHNOLOGY

J. C. M. Dekkers

Department of Animal Science, Iowa State University, Ames, IA, 50011, USA

SUMMARY

To date, most genetic progress in livestock populations has been realized by selection on breeding values estimated from phenotype. Molecular genetic technology is providing knowledge on individual genes affecting traits of economic importance. Over the next decades, knowledge and information on the genetic control of economic traits will increase exponentially. This puts the onus back onto quantitative genetics and animal breeding to devise selection and breeding strategies to optimally incorporate molecular technology in breeding programs and to obtain greatest benefit from the use of information generated by molecular genetics. The objective of this paper is to outline some of the challenges that lie ahead with regard to the integration of molecular with quantitative genetics and animal breeding. These include issues related to detection and estimation of gene effects, incorporating gene and marker information in genetic evaluation, optimization of the use of molecular information in purebred and crossbred selection strategies, and the integration of molecular and reproductive technologies. The main conclusion is that great opportunities exist to enhance rates of genetic improvement in livestock, provided the molecular information is appropriately incorporated in selection programs; misuse of molecular information can lead to suboptimal genetic gains in the shorter term and reduced genetic gain in the longer term. Appropriate integration of molecular genetics in breeding programs will require careful design and optimization of breeding programs, ultimately in conjunction with the design of animal management, feeding and treatment programs.

Keywords: Genetic selection, molecular genetics, breeding programs

INTRODUCTION

To date, most genetic progress for quantitative traits in livestock has been made by selection on phenotype or on estimates of breeding values (EBV) derived from phenotype, without knowledge of the number of genes that affect the trait or the effects of each gene. In this quantitative genetic approach to genetic improvement, the genetic architecture of traits of interest has essentially been treated as a 'black box'. Despite this, the substantial rates of genetic improvement that have been and continue to be achieved in the main livestock species, is clear evidence of the power of quantitative genetics approaches to selection.

The success of quantitative genetic approaches does, however, not mean that genetic progress could not be enhanced if we could gain insight into the black box of quantitative traits. By being able to study the genetic make-up of individuals at the DNA level, molecular genetics has given us the tools to make those opportunities a reality. The objectives of this paper are to review, albeit by no means exhaustive, the issues, opportunities and challenges for incorporation of molecular genetic technologies in genetic improvement programs for livestock.

The eventual application of molecular genetics in livestock breeding programs depends on developments in the following four key areas, which jointly culminate in the successful implementation of strategies for gene-assisted selection (GAS), ie selection on a known gene, or marker-assisted selection (MAS), ie selection on markers linked to a quantitative trait locus (QTL):

i. *Molecular genetics*: identification and mapping of genes and genetic polymorphisms
ii. *QTL detection*: detection and estimation of associations of identified genes and genetic markers with economic traits
iii. *Genetic evaluation*: integration of phenotypic and genotypic data in statistical methods to estimate breeding values of individual animals in a breeding population
iv. *Marker-assisted selection*: development of breeding strategies and programs for the use of molecular genetic information in selection and mating programs.
v. Aspects of each of these areas of research will be reviewed in what follows, with main emphasis on strategies for gene- and marker-assisted selection.

QTL MAPPING APPROACHES

To date, the use of molecular genetics in animal breeding has focused mainly on the detection of genes affecting traits of economic importance (QTL). Two approaches can be distinguished: the candidate gene approach and the genome scan approach (Haley 1999). In the candidate gene approach (Rothschild and Soller 1997), knowledge from species that are rich in genome information (e.g., human, mouse) and/or knowledge of the physiological basis of traits is used to identify genes that are hypothesized to play a role in physiological mechanisms underlying traits of economic importance. Using this information, the candidate genes are identified in the species of interest and polymorphisms in the coding, but usually non-coding, regions of the gene are detected. Associations of these polymorphisms with the trait of interest are then identified using statistical analysis of phenotypic records of an unstructured sample of individuals from the population of interest, relying on population-wide linkage disequilibrium due to close linkage or, ideally, identity of the polymorphism with the causative locus. Using this approach, several genes with major effect have been identified, a prime example being the estrogen receptor gene affecting litter size in pigs (Rothschild 1996).

The genome scan approach to QTL detection uses anonymous genetic markers spread over the genome to identify genes affecting quantitative traits. Unless marker density is high, these studies cannot rely on population-wide linkage disequilibrium between markers and QTL. Instead, they must rely on the linkage disequilibrium that exists within families in outbred populations or that is created in crosses between breeds or lines. Using statistical methods based on interval mapping (Lander and Botstein 1989; Haley and Knott 1992), QTL can then be identified and their position and effect estimated by associating marker data to phenotypic records. The precision of, in particular, estimates of QTL position that can be obtained from these approaches is, however, limited, and large population sizes are needed.

Although the candidate gene and the genome scan approach are often viewed as alternate approaches for identifying genes of interest, it is clear that they can be complementary, with a genome scan identifying regions of the genome that harbor potential QTL, followed by further investigation of genes known to be located in that region using the candidate gene approach.

The design and statistical analysis of QTL mapping studies has received much attention and debate over the past decade. Several designs for use in outbred populations have been developed and applied, based on family structures within outbred populations (e.g. the two or three generation designs suggested by Weller *et al* (1990) and van der Beek *et al* (1995)) or crosses between divergent outbred lines (e.g., Anderson *et al.* 1994). Population sizes required for sufficient power to detect QTL remains a problem. Required population sizes are greatest for genome scan mapping within outbred populations, smaller but still sizeable for genome scan mapping in crosses between outbred lines, and smallest for candidate gene analyses that rely on strong or complete population-wide linkage disequilibrium. Because of the population sizes required for QTL mapping and associated costs and logistics when such populations must be developed specifically for the purpose of mapping QTL, one of the challenges for QTL mapping in livestock will be for the use of designs that can capitalize on existing population structures in livestock breeding populations. The use of the granddaughter design in dairy cattle (Weller *et al.* 1990), which capitalizes on the large paternal half-sib families created with artificial insemination, is a good example. Selective genotyping with or without DNA pooling (Darvasi and Soller 1994) can be used to further advantage in such designs, minimizing the number of genotyping assays.

Haley (1999) and Coppieters *et al* (1999) suggested the use of population-wide disequilibrium that exists in breeding populations for detection and fine mapping of QTL (see Baret and Hill 1997 for a review of methods). With reference to data from the Holstein population, Coppieters *et al* (1999) indicated that population-wide disequilibrium may in fact be widespread in livestock populations.

Because candidate gene approaches require no specific population structure, in fact an across-population sample is preferred, such approaches can rely on existing animal breeding populations. The same holds for other strategies that rely on population-wide disequilibrium. The use of outbred crosses for QTL mapping with anonymous markers, however, typically requires development of specific resource populations, in particular when they involve breeds that are highly divergent for the trait of interest. Although the use of highly divergent lines improves the power of identifying QTL in such experiments, further research is then needed to ascertain whether identified QTL have an effect within populations that are of commercial interest. Therefore, for within-breed selection, the use of within-breed QTL mapping strategies using existing family and population structures is preferred. Drawbacks to the use of existing breeding populations, especially when this involves animals from the field, are limitations on availability of data, in particular on traits that are difficult to record, which are of specific interest for the use of MAS.

USE OF GENE OR MARKER INFORMATION IN GENETIC EVALUATION

Although candidate gene and QTL mapping experiments can result in identification of genes of interest that can be used in genetic improvement programs, their implementation will require estimation of QTL effects in commercial breeding populations. In particular with the use of anonymous markers, marker and QTL effects must be estimated on a within-family basis and re-estimated on a routinely. This will require routine systems for DNA collection and marker genotyping. Even when the actual gene has been identified, there will be a need to re-estimate gene effects on a regular basis to improve the accuracy and to guard against unfavorable associations with other traits and against epistatic effects with the background genome or environment.

Ideally, estimation of QTL and gene effects would be incorporated in routine animal model genetic evaluations, providing best linear unbiased predictors (BLUP) of EBV for identified QTL and for the collective effect of polygenes. Kennedy *et al* (1992) suggested inclusion of genotype as a fixed effect in an animal model to estimate the effects of known genes, capitalizing on population-wide disequilibrium. Fernando and Grossman (1989) developed BLUP-QTL models for QTL linked to markers, relying on within-family linkage disequilibrium between markers and QTL, by including QTL as random effects along with residual polygenic effects. With increasing marker densities and evidence of extensive linkage disequilibrium across the population (Coppieters *et al.* 1999), it seems clear that methods for genetic evaluation are needed that can simultaneously capture information from within-family disequilibrium and population-wide disequilibrium between markers and QTL.

Although statistical testing is an important issue in QTL mapping and detection, the question is whether statistical significance should also be an issue for the use of molecular data in genetic evaluation and selection. For comparison, animal breeders are very accustomed to selection on BLUP EBV regardless of whether the top bulls have EBV that differ by some level of statistical significance from the next group of bulls. In fact, we never worry about statistical significance when selecting on EBV. The reason for this is that uncertainty about estimates is already incorporated in EBV, which is accomplished by fitting animal genetic effects as random in the BLUP procedures: phenotypic information on bulls with few daughters and, therefore, greater uncertainty on their EBV, is regressed back to the mean to a greater degree than phenotypic information on bulls with many progeny. As a result, by treating animal genetic effects as random, BLUP EBV avoid the problem of fixed effect estimates, for which the effects with the highest estimates are on average overestimated. Instead, BLUP EBV are by definition unbiased: even for the top bulls, their EBV is expected to be equal to their true breeding value.

Based on this analogy, it is important that genetic evaluation procedures that use molecular information properly incorporate uncertainty about QTL parameters and estimates. As a first step, this can be handled by fitting QTL effects as random effects, as was done by Fernando and Grossman (1989), resulting in BLUP estimates of QTL effects that are regressed back to the mean based on the amount of information available to estimate the effect. The same does not occur when the effects of known genes are fitted as fixed effects, as suggested by Kennedy *et al* (1992). Therefore, the use of resulting estimates for selection purposes may require that the effects of known genes, or effects due to population-wide disequilibrium between a marker and a QTL, are also fitted as random effects.

BLUP-QTL approaches, as developed by Fernando and Grossman (1989) and others, assume variance due to the QTL is known without error. Spelman and Van Arendonk (1997) showed that use of improper QTL parameters (variance and location) can reduce efficiency of marker assisted selection. Therefore, as a second step in dealing with uncertainty, methods will be needed that can incorporate uncertainty about QTL parameters. Bayesian methods can be used for this purpose but this will substantially increase the complexity of routine genetic evaluations.

USE OF GENE OR MARKER INFORMATION IN WITHIN-BREED SELECTION

The BLUP-QTL genetic evaluation procedures described above result in BLUP of breeding values for identified QTL and residual polygenic effects. The BLUP of the total breeding value can then obtained by simply summing the EBV for identified QTL and the EBV for polygenic effects, to give the following total QTL-BLUP EBV: EBVTotal = EBVQTL + EBVPolygenes. In theory and under additivity, selection on such EBV will maximize response from the current to the next generation, ie single-generation response.

Several studies have investigated the extra response to selection that can be achieved with incorporation of marker or gene information in genetic evaluations. Extra responses from MAS depend on the amount of genetic variance explained by the marked QTL, the ability of markers to trace segregation of QTL (linkage and marker informativeness), and the efficiency of selection without marker information. In general, benefits from MAS are greatest in situations where regular selection is limited or inefficient. This includes selection for traits with low heritability and traits with restrictions on phenotypic recording, such as sex limited traits, traits recorded after selection, carcass traits, and traits that are expensive or difficult to evaluate (Meuwissen and Goddard 1996).

Most studies on MAS show greatest extra responses in early generations, followed by a decline in later generations. Gibson (1994) examined the long-term consequences of MAS using computer simulation of selection on a known major QTL (ie gene-assisted selection or GAS). He found that, although GAS resulted in greater cumulative response to selection in the short term, phenotypic selection achieved greater response in the longer term, as is illustrated in Figure 1. This figure also shows that, although GAS fixes the QTL more rapidly, this is at a cost of response in polygenes. With GAS, the polygenic response that is lost in the initial generations, as the QTL is selected toward fixation, is never entirely recovered in later generations.

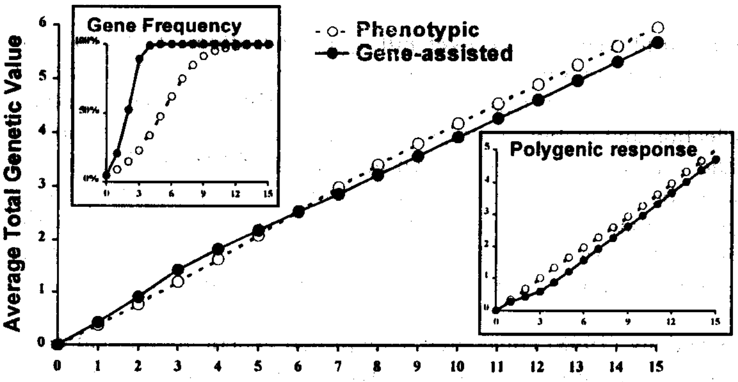

Figure 1. Comparison of phenotypic to genotypic selection on a known additive QTL with a=0.5 phenotypic standard deviations for a trait with h^2=0.25 and 20% of males and females selected.

The results of Gibson (1994) have been confirmed in several other studies that simulated selection on a known major QTL (e.g., Pong-Wong and Woolliams 1998), including studies in which polygenes were simulated based on a finite number of loci instead of based on the infinitesimal genetic model of quantitative genetics (Kuhn *et al.* 1997; Fournet *et al.* 1997). All these studies, however, assumed the effect of the gene was known, compared selection on the QTL to selection on the animal's own phenotype, and assumed phenotype was observed on all animals. Other studies have removed one or more of these limitations by considering selection on linked genetic markers or marker brackets, selection on BLUP of breeding values from an animal model, or selection for a trait that is not measurable in both sexes or on live animals (e.g., Van der Beek and Van Arendonk 1994; Ruane and Colleau 1995, 1996; Meuwissen and Goddard 1996; Larzul *et al.* 1997). These studies did not consistently find that longer-term responses were less for MAS than traditional selection. All studies did, however, find that the advantage of MAS over traditional selection declined in later generations.

The literature described above raises important questions on how identified QTL should be used in selection. Although the main implication from this work is that selection on the QTL-BLUP selection criteria described above may not maximize response in the longer-term, it also raises the question whether this selection criterion makes optimal use of QTL information in the short-term.

The reason that the QTL-BLUP criterion does not maximize response over multiple generations, although it is designed to maximize response from the current to the next generation, is that selection changes not only the population mean but also population parameters, in this case frequency of the identified QTL. Selection in the current generation, thereby, affects responses that can be achieved in subsequent generations. Therefore, optimizing selection response over multiple generations requires strategies that not only focus on maximizing the mean in the next generation but also on controlling the parameters that affect responses in subsequent generations. Such strategies to optimize GAS or MAS over multiple generations are described below.

Optimizing selection on identified QTL over multiple generations. Methods to optimize selection on an identified QTL over multiple generations were developed by Manfredi *et al* (1998) and Dekkers and van Arendonk (1998). Manfredi *et al* (1998) used a standard nonlinear programming routine to derive selection and mating strategies that maximized discounted response over a planning horizon. Dekkers and van Arendonk (1998) used optimal control theory to derive optimal strategies for selection an identified QTL. Optimal control theory is a mathematical method that is used extensively in engineering and economics to optimize multiple-stage decision problems (Bryson and Ho 1975; Lewis 1986). Because of the Markov nature of genetic inheritance, optimal control theory is ideally suited to optimize strategies for selection over multiple generations.

Dekkers and van Arendonk (1998) considered a simplified genetic model and breeding structure to simplify computations; a trait affected by an additive major QTL with two alleles of known effect, along with polygenes, which were modeled based on the infinitesimal model with constant genetic variance; large population size, non-overlapping generations, and equal selection in males in females. Optimal selection strategies were developed by finding weights for an index that combines the (assumed known) breeding value for the identified QTL with an estimate of the polygenic EBV: $I_{Optimal} = b^*(BV_{QTL}) + EBV_{Polygenes}$

Polygenic EBV were estimated based on own phenotype as h^2(phenotype – QTL value), where h^2 is heritability. Under optimal selection, different weights on the QTL were allowed for, depending on QTL genotype and generation. Note that with what will be referred to as standard GAS, weights are equal to one. Phenotypic selection implicitly applies a weight equal to heritability. For optimal selection, the objective was to find the index weights that maximized cumulative response after T generations.

By optimizing the weight put on the QTL, Dekkers and van Arendonk (1998) showed that use of information on a QTL can result in greater responses to selection in the longer term than phenotypic selection. While methods developed by Dekkers and van Arendonk (1998) were for a QTL with additive effects and for equal selection among males and females, methods have subsequently been extended to non-additive QTL and different selection for males and females, the latter by allowing different index weights for males versus females. Extra responses from optimal selection over phenotypic and standard GAS, for a QTL with varying degrees of dominance and for varying polygenic heritabilities, are illustrated in Table 1.

Table 1. Extra response (%) from standard (Stand) and optimal (Opt) selection on a known QTL over phenotypic selection for a QTL with different degrees of dominance for a trait with different polygenic heritabilities (h^2). The QTL has an additive effect (a) equal to 1 genetic standard deviation and a starting frequency of 10%. Fractions selected among males and females are 25 and 50%

		Number of generations considered (and optimized over for optimal selection)							
	Degree (d) of	1		3		5		10	
h^2	dominance	Stand	Opt	Stand	Opt	Stand	Opt	Stand	Opt
.05	-0.5a	12.3	18.2	97.9	181.2	128.7	146.3	63.2	74.3
	0	90.4	90.4	126.9	127.4	70.5	90.0	17.9	27.7
	0.5a	108.0	108.0	91.9	93.8	40.6	57.7	3.9	12.7
	a	113.7	113.7	60.3	70.5	18.3	38.5	-3.7	5.2
	1.5a	111.9	111.9	35.0	63.5	4.0	32.7	-8.3	6.8
.25	-0.5a	1.7	2.0	4.3	44.6	22.3	62.0	20.7	27.3
	0	15.4	15.4	20.1	21.6	7.7	18.2	-6.0	1.5
	0.5a	18.9	19.0	8.6	12.1	-3.3	9.6	-7.0	0.8
	a	19.9	20.0	-0.7	7.1	-7.6	8.0	-6.5	2.0
	1.5a	19.2	19.3	-5.7	6.1	-7.8	11.2	-4.6	5.2
.50	-0.5a	1.1	1.2	1.7	13.9	4.1	39.1	14.4	18.5
	0	3.6	3.6	3.7	4.3	0.4	5.1	-4.0	1.3
	0.5a	3.9	4.0	0.0	2.5	-4.5	4.2	-4.3	1.7
	a	3.7	4.0	-3.6	2.4	-4.9	7.8	-3.1	4.0
	1.5a	3.5	3.9	-3.8	5.7	-2.6	13.2	0.3	8.5

Extra response from selection on the QTL through standard GAS was greatest for traits with low heritability and for short planning horizons (Table 1). Standard GAS resulted in similar or less response than phenotypic selection for longer planning horizons and higher heritabilities. Degree of dominance at the QTL affected the extra response that can be expected from standard GAS. It must be pointed out, however, that these results vary depending on the specific parameters used (ie size of the QTL effect, starting QTL frequency, and fractions selected).

The benefit that can be expected from optimizing selection on the QTL over phenotypic selection is also shown in Table 1. Comparing extra responses from optimal and standard GAS over phenotypic selection, response from optimal selection was always at least as high as response from standard GAS and often substantially greater. Benefits from optimal selection over standard GAS increased with degree of dominance at the QTL. For QTL with complete or over-dominance, optimizing selection resulted in substantial extra responses over standard GAS of up to 25% for longer as well as shorter (three generations) planning horizons. For a recessive QTL, extra responses of over 80% were found for a planning horizon of only three generations for a trait with low heritability. Benefits from optimal over standard GAS were minimal or absent over one generation. Again, these results depended greatly on parameters such as size of the QTL, starting frequency, and fractions selected, so the results shown here serve as illustration for the potential benefits that can be expected from optimizing selection.

Figure 2 shows the optimal weights on the breeding value for the QTL for the strategy that maximized cumulative response after 10 generations. For optimal selection, weights changed from generation to generation, as frequency of the QTL changed and the population moved closer to the end of the planning horizon. Except for the last three generations, optimal weights were close to the implicit weight for phenotypic selection (h^2). In any generation, weights were higher for selection of males than females, which is related to the greater selection intensity among males. Weights on the breeding value for the QTL were also greater when the individual was bb versus BB. In essence, this implies that the optimal selection strategy increased the frequency of the favorable QTL allele to a greater degree by selection *against* bb genotypes than by selection *in favor* of BB.

Figure 3 shows trends in gene frequencies for standard and optimal selection on a QTL with over-dominance. For standard GAS selection, frequency of the B allele stabilized around 85% after 10 generations. For optimal selection, gene frequency increased linearly up to the last generation, at which point the optimal strategy moved the gene to fixation in the selected dams and to less than 20% in the selected sires. The optimal strategy thereby created a high frequency of heterozygotes (Bb) in the last generation, consistent with the objective of the strategy, which was to maximize the average genetic level in the *final* generation. Although this would not be a viable strategy to follow in practice, this example illustrates that this procedure can in principle be used to optimize selection in dam and sire lines in a crossbreeding program, where the objective is to maximize performance of crossbred animals. More discussion on optimizing QTL selection in crossbreeding programs follows.

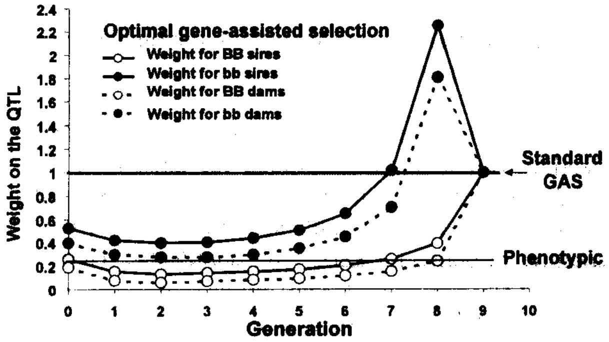

Figure 2. Optimal weights on breeding values for a known additive QTL with a=0.25 phenotypic standard deviations and a starting frequency of 10%, for a trait with h^2=0.25 and 5% of males and 25% of females selected, when the objective is to maximize response after 10 generations.

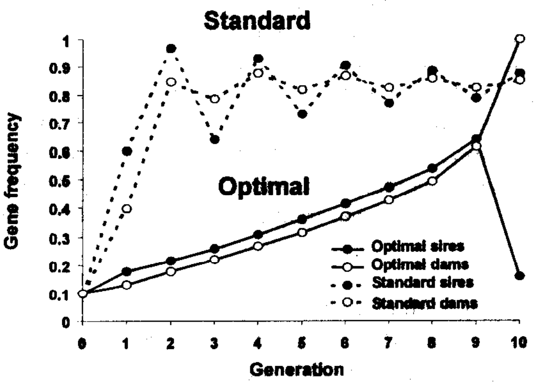

Figure 3. Effect of standard and optimal gene-assisted selection on gene frequency for a known over-dominant QTL with a=0.5 phenotypic standard deviations, d=1.5a, and a starting frequency of 10%, for a trait with h^2=0.25 and 5% of males and 25% of females selected.

In general, the conclusion to be drawn from the results given above is that, when information from a QTL is used properly, by balancing losses in polygenic response against increases in QTL frequency, GAS leads to greater long-term response than phenotypic selection. Thus, the results obtained by Gibson (1994) were caused by the fact that the standard GAS selection criterion used (ie with a weighting of one on the QTL) did not properly weight QTL and polygenic information in order to maximize long-term response. More importantly, however, these results show that the standard GAS criterion is also be suboptimal for shorter planning horizons (e.g. three generations). And, although Table 1 showed no or small benefits from optimizing selection over one generation, there are situations where the standard GAS criterion is substantially suboptimal even over a single generation, as will be demonstrated in the next section.

Optimizing single generation response for QTL with dominance. In table 1, standard GAS was equivalent to optimal selection over a single generation for additive QTL but not for QTL with dominance, although the differences were small in the examples chosen. In other words, selection based on a QTL-BLUP criterion computed by simply summing EBV for the QTL and polygenes did not maximize single-generation response for QTL with dominance. Here, breeding values for the QTL were derived following Falconer and Mackay (1996) as $+2q\alpha$, $(q-p)\alpha$, and $-2p\alpha$for QTL genotypes BB, Bb, and bb, respectively, where α is the average allele substitution effect and is computed based on allele frequencies in the current generation (p_0 and q_0) as $\alpha=a+(q_0-p_0)d$.

The fact that the QTL-BLUP EBV does not maximize single-generation response goes against the objective of EBV, which is to predict the expected phenotype of progeny. Consequently, if derived properly, selection on EBV is expected to maximize response from the current to the next generation. The reason for the discrepancy is that the derivation of BV for the QTL, as described above, assumes parents are mated to an unselected group of individuals from the current population (Larzul *et al.* 1997; Dekkers 1999). Under an additive genetic model, the contribution of a parent to its progeny is independent of the individuals the animal is mated to and, therefore, breeding values are equivalent under random and non-random mating. With dominance, however, breeding values derived following Falconer and Mackay (1996) are equivalent to the expected performance of progeny only under random mating to unselected parents or when changes in gene frequency are small. With substantial changes in gene frequency, as will be the case with direct selection on the QTL, breeding values for the QTL that, when used in combination with EBV for polygenic effects, maximize single generation response to selection, depend on gene frequencies among mates, which themselves will be selected.

It can be shown (Dekkers, 1999) that, with random mating of *selected* parents, optimal breeding values for a QTL are equivalent to breeding values derived based on standard quantitative genetic theory ($+2q\alpha$, $(q-p)\alpha$, and $-2p\alpha$) but with the average effect of allele substitution, α, derived based on gene frequency among selected mates, rather than frequency in the unselected parental generation. For sires, the optimal allele substitution effect depends on gene frequencies among dams ($\alpha_s=a+(q_d-p_d)d$) and for dams, the optimal allele substitution effect depends on gene frequencies among sires ($\alpha_d=a+(q_s-p_s)d$). Because gene frequencies among selected mates depend on the selection practiced, which in turn depends on allele substitution effects, optimal allele substitution effects must be derived in an iterative manner (see Dekkers, 1999).

Response from selection on optimized versus standard BV for the QTL was evaluated for a range of parameters (Dekkers, 1999). Benefits of optimal selection were greatest for intermediate gene frequency and increased with magnitude of additive and dominance effects to up to 9%. Extra response was negligible for gene frequencies less than 0.05 or greater than 0.85. Therefore, depending on degree of dominance, and QTL effect and frequency, there will be a need to optimize selection on QTL even if the objective is to maximize response over a single generation.

IMPLICATIONS FOR MAS IN PURE BRED POPULATIONS

The above shows that for selection on known QTL, selection on a simple sum of the breeding value of the QTL and an estimate of the breeding value for polygenes does not maximize response to selection and that substantial extra responses can be achieved from optimizing selection, not only in the longer term but also in the short term (even over a single generation), especially for QTL with dominance. These results were, however, obtained under a simplified genetic model and breeding structure. Specifically, selection was on a single identified QTL with known effects, polygenic effects followed the infinitesimal model with variance assumed unaffected by selection, a population of infinite size (no inbreeding) with non-overlapping generations was considered, phenotype was observed on all selection candidates and estimates of polygenic breeding values were based on phenotype assumed adjusted without error for the known QTL and other fixed effects. Additional research is, therefore, required (and ongoing) to further elucidate opportunities for optimization of selection on identified QTL in practical situations, where many of these assumptions do not hold.

The fact that the results given above show that selection on the standard GAS criterion is suboptimal, does not imply that BLUP-QTL procedures should not be used for genetic evaluation. Indeed, methods such as BLUP-QTL are needed to provide the best estimates of breeding values for the QTL and polygenic effects. What is at issue here is how the QTL and polygenic EBV that result from such models are combined into an overall criterion for selection. The results described above indicate that simply summing QTL and polygenic EBV, which would be conform BLUP procedures, does not provide in a criterion that maximizes response to selection. Instead, QTL and polygenic EBV must be weighted appropriately, in accordance with the selection objective. It must be noted that optimization of QTL selection does require knowledge of the degree of dominance at the QTL, as well as the frequency of QTL alleles. This will require further development of BLUP-QTL procedures, which currently mainly focus on estimation of additive QTL breeding values.

In the above, it was assumed that the QTL was identified and its effects known without error. In general, it can be anticipated that uncertainty about estimates of QTL effects, including as a result of linkage to a marker, will reduce opportunities to optimize selection and make the standard QTL-BLUP criterion less suboptimal. However, as information on effects and mode of action of identified genes increases, including epistatic interactions, gametic imprinting and polar overdominance (Cockett *et al.* 1996), opportunities to obtain greater responses to selection from optimizing selection will increase.

Another important limitation of the research described above is that the objective considered was to maximize cumulative response after a fixed number of generations. This will rarely be a realistic goal in practical breeding programs but helped to illustrate opportunities for optimizing selection in the present scenario. A more practical objective would be to maximize the discounted sum of responses in each generation of a planning horizon. Depending on the discount rate, such an objective can be short-term (high discount rate) or longer-term (low discount rate), thereby affecting the opportunities for optimizing selection. Manfredi *et al.* (1998) considered maximization of such an objective for selection and mating on a known QTL and found limited benefits from optimal over standard GAS for the dairy goat selection program they considered.

Although MAS or GAS typically implies selection on loci that have a direct effect on the quantitative trait of interest, it must be noted that the same concepts also apply to selection on undesirable (or desirable) single gene traits or genetic defects that have no direct effect on the quantitative trait (Dekkers and van Arendonk, 1998), e.g. the stress gene in swine or the BLAD gene in dairy cattle, and even coat colour and the polled gene. After all, the objective in most breeding programs is to maximize the quantitative trait 'profitability' and, by contributing to profitability, genetic defects or other undesirable single gene traits are QTL for profit as a quantitative trait. In many cases, there is great incentive to eliminate undesirable genes from breeding stock. It must be realized, however that intense selection against undesirable genes is equivalent to fixing a QTL for a quantitative trait rapidly, with consequent losses in rates of improvement in other genes that affect profitability. Therefore, there is scope for optimizing selection on single gene traits and considering such genes as QTL for profitability appears useful.

In practice, selection will not be on a single identified QTL over the entire planning horizon. Instead, new QTL will be identified on a regular basis or multiple QTL may be selected for simultaneously. Meuwissen and Goddard (1996) argued that with new QTL identified on a regular basis, selection on any given QTL will only be for a relatively short-term and that, therefore, the long-term consequences of selection on a QTL, as described by Gibson (1994), would not apply. In essence, one long-term planning horizon is converted to a series of short-term planning horizons. Although it is clear that regularly switching selection to new QTL will reduce the long-term negative consequences of QTL selection, it is unclear whether they will entirely disappear and whether in this situation selection on QTL-BLUP breeding values is indeed optimal. After all, any selection on a QTL will change genetic parameters (frequency and variance at the QTL) and, even if any given QTL is only selected on for a relatively short term, after which it becomes part of the polygenic variation, this selection does affect future variances and opportunities for selection. Such changes in variance and future opportunities for selection can considered when optimizing selection. In addition, as shown above, QTL-BLUP selection may not be optimal even in the shorter-term, especially for QTL with dominance.

A related question is how to maximize response from simultaneous selection on multiple QTL. Although strategies to optimize selection for such a situation has not been solved, some of the issues involved can be illustrated by consider the extreme case when all QTL are known without error and the objective is to fix *all* QTL for the favorable allele as rapidly as possible. Under a simple genetic model, it can be shown that in this case, the emphasis placed on each individual QTL depends not on the effect of the QTL but also on the allele frequencies at the QTL; greater emphasis must be placed on QTL for which the favorable allele has low frequency because it will require greater selection pressure to move that allele to fixation, regardless of the effect of the QTL. Although this is an extreme situation, it illustrates that the need for optimization remains with simultaneous selection on multiple QTL. The need for optimization will further increase as information on epistatic interactions between QTL accumulates.

INTEGRATING MOLECULAR AND REPRODUCTIVE TECHNOLOGIES

The previous indicates that one of the main challenges for selection on a QTL is that it reduces selection on polygenes and, unless selection on the QTL is properly balanced against lost response in polygenes, QTL selection can be detrimental in the long-term and suboptimal in the short term. In the cases considered, however, QTL selection was incorporated into traditional selection stages, where QTL selection competes with polygenic selection. Several authors have suggested that the most effective manner to capitalize on QTL information is by incorporating it at stages, where selection was not possible previously.

A prime example is preselection of young dairy bulls for entry into progeny testing (e.g., Kashi *et al.* 1990); with use of multiple ovulation and embryo transfer (MOET) technologies, often more bull calves are produced from superior bull dams than can be entered into progeny testing. Because the EBV of full brothers is identical when based on pedigree information alone, QTL information allows informative decisions to be made on which full brother should be entered into progeny testing. The effect of this additional QTL selection stage on polygenic response depends on the availability of excess bull calves that is not capitalized on in current selection strategies; if currently the choice of which bull calf from a flush to enter into progeny testing is at random, the QTL selection stage will have minor consequences on polygenic response and there is no need for optimization, apart from consideration of costs. However, if the excess of bull calves to allow QTL selection must be created by increasing the number of bull dams selected, then QTL selection will reduce polygenic response by reducing selection intensity among bull dams. In this situation, there will be a need to balance polygenic against QTL response through optimization.

Opportunities for QTL selection in this manner can be increased through the use of reproductive technologies. The use of MOET in preselection of dairy bulls, as discussed above, is a good example. Other technologies, such as oocyte recovery and in vitro fertilization can be used to further enhance reproductive rates.

In the example of pre-selection of bull calves for progeny testing, QTL selection capitalized on extra selection space (Soller and Medjugorac 1999) that was created within a generation through increased reproductive rates. Georges and Massey (1991) proposed designs they named velogenetics, in which short generations with selection based on QTL information alone are introduced, with generations with regular phenotypic alternated by several generations of QTL selection. Without the need to produce phenotypic data on selection candidates for QTL selection, the generations with QTL selection can be very rapid through oocyte recovery from the foetus, in-vitro embryo production and embryo transfer (Georges and Massey 1991). Variations upon velogenetic designs were proposed by Haley and Visscher (1998). Such designs do require optimization of QTL versus polygenic responses, along with costs. An additional challenge in these designs is generation of QTL information which, unless QTL are known, must be re-estimated on a regular and within family basis using phenotypic information generated in the phenotypic selection generations. Nevertheless, these examples show great opportunities exist for the development of enhanced breeding programs based on integrating molecular and reproductive technologies.

USE OF GENE OR MARKER INFORMATION IN CROSSBREEDING PROGRAMS

Although breeding programs primarily rely on selection within purebred populations, in many cases the objective is to improve crossbred performance. This raises important additional questions on how to incorporate molecular genetic information in selection programs within pure breeds that contribute to crossbreeding programs, in particular for QTL that exhibit non-additive effects. Alternative scenarios for selection on a single identified QTL within sire and dam breeds for a two-way cross are presented in Table 2. From these simplified scenarios, it is clear that emphasis on the QTL may not be the same for sire and dam breeds and that there will be a need to simultaneously optimize selection on identified QTL within both breeds in order to maximize both the rate of improvement within the pure breeds and the level of performance in the crossbreds. While table 2 considers only a single QTL, it is clear that design of selection criteria and strategies will be further complicated when multiple QTL are available, each with their own mode of action and epistatic interactions.

Table 2. Possible directions for selection within sire and dam breeds that contribute to a two-way cross on an identified QTL for a trait expressed in crossbreds, depending on mode of action at the QTL. In all cases, selection emphasis on the QTL must be optimized against emphasis on polygenes

Mode of gene action at QTL	Direction for QTL selection on favorable allele in	
	Sire breed[1]	Dam breed[1]
Additive	Increase	Increase
Partial dominance	Increase	Increase but at slower rate
Negative dominance	Increase	Increase
Over dominance	Increase	Decrease

[1] Choice of selection strategy within sire versus dam breeds may depend on the accuracy of estimates of polygenic breeding values within each breed, population size, inbreeding, etc.

THE FUTURE

The contribution of molecular genetics to enhance the knowledge on genetics of economic traits will not cease with identification of genes based on QTL or candidate gene searches. Although substantial additional genetic gains can be achieved with selection on linked genetic markers, the ultimate aim will be to directly identify the main genes involved in a trait and to elucidate the function of each of these genes. Technologies and approaches to advance molecular genetic knowledge to this level are being developed and applied to the human genome and model organisms (e.g. Schmitt, 1999). With the wealth of information that is generated in this research, bio-informatics plays an increasingly important role to organize, analyze, and interpret this information (Sobral, 1999). Although research on gene identification, function, expression and regulation in livestock can greatly benefit from genome research in humans and model organisms, much additional work will be needed elucidate the specific role of each gene involved in traits of economic importance in livestock and their interactions with other genes and the environment. Ultimately, however, this research will provide additional knowledge that can be used to enhance selection programs. With the additional information on genetic control of traits will come additional knowledge on the environmental control (e.g. feeding) of traits of economic importance in livestock. The interaction between environment, management and genetics will become more important. Opportunities to select and manage for niche markets will increase.

As information on the genetic and environmental control of animal production increases, sophisticated methods can be developed to control the performance of individual animals in the production process through environmental and management factors, as well as to control and enhance the performance of populations of animals through designed genetic improvement. Approaches akin to the control of engineering systems can then be used to control the expression of individual genes and combinations of genes to enhance performance of animals through tailored management, feeding and treatment systems and controls. In addition, engineering systems approaches can be used to improve a population of animals in the desired direction through the controlled selection on individual genes or combinations of genes that will optimize the performance of progeny and subsequent generations, given tailored management, feeding and treatment systems and controls. It is clear that this will require close interaction between genetic improvement and environmental control systems, requiring the need to optimize large engineering-type systems in an integrated manner. In principle, these systems could even be extended to forage production and genetic improvement systems, leading to the integration of genetic improvement and production of animals with genetic improvement and production of animal feeds. The sky will be the limit.

REFERENCES

Anderson, L., Haley, C.S., Ellegren, H., Knott, S.A., Johansson, M., Anderson, K., Anderson-Eklund, L., Edfors-Lilja, I, Fredholm, M., Hansson, I., Hadansson, J. and Lundstrom, K. (1994) *Science* **263**:1771

Baret, V. and Hill W.G. (1997) *Anim. Breed. Abstr.* **65**:309

Bryson, A.E. and Ho, Y-C. (1975). "Applied Optimal Control: Optimization, Estimation and Control" Wiley, New York

Cockett, N.E., Jackson, S.P., Shay, T.L., Farnir, F, Berghmans, S., Snowder, G.D., Nielsen, D. and Georges, M. (1996) *Science* **273**:236

Coppieters, W., Blott, S., Farnir, F., Grisart, B, Riquet, J. and Georges, M. (1999) In "From Jay Lush to Genomics: Visions for Animal Breeding and Genetics" p.70, editors J.C.M. Dekkers, S. Lamont and M.F. Rothschild, Iowa State University, Ames

Darvasi, A. and Soller, M. (1994) *Genetics* **138**:1365

Dekkers, J.C.M. (1999) *Genet. Sel. Evol.* (Submitted)

Dekkers, J.C.M. and van Arendonk, J.A.M. (1998) *Genetical Research* **71**:257

Falconer, D.S. and Mackay, T.F.C (1996) "Introduction to Quantitative Genetics" Longman, Harlow

Fernando, R.L. and Grossman, M. (1989) *Genet. Sel. Evol.* **21**:467

Fournet, F., Elsen, J.M., Barbieri, M.E. and Manfredi, E. (1997) *Genet. Sel. Evol.* **29**:35

Georges, M. and Massey, J.M. (1991) *Theriogenology* **35**:151

Gibson, J.P. (1994). *Proc. 5th Wld. Congr. Genet. Appl. Livest. Prod.* **21**:201

Haley, C. (1999) In "From Jay Lush to Genomics: Visions for Animal Breeding and Genetics" p.47, editors J.C.M. Dekkers, S. Lamont and M.F. Rothschild, Iowa State University, Ames

Haley, C.S. and Knott, S.A. (1992) *Heredity* **69**:315

Haley, C.S. and Visscher, P.M. (1998) *J. Dairy. Sci.* **81**(2):85

Kashi, Y., Hallerman, E. and Soller, M. (1990) *Anim. Prod.* **51**:63

Kennedy, B.W., Quinton, M. and Van Arendonk, J.A.M. (1992) *J. Anim.Sci.* **70**:2000

Kuhn, M.T., Fernando, R.L. and Freeman, A.E. (1997) *J.Dairy Sci.* **80**(1):228

Lander, E.S. and Botstein, D. (1989) *Genetics* **121**:185

Larzul, C., Manfredi, E. and Elsen, J.M. (1997) *Genet. Sel. Evol.* **29**:161

Lewis, F.L. (1986) "Optimal Control" Wiley, New York

Manfredi, E., Barbieri, M., Fournet, F. and Elsen, J.M. (1998) *Genet. Sel. Evol.* **30**:127

Meuwissen, T.H.E. and Goddard, M. E. (1996) *Genet. Sel. Evol.* **28**:161

Pong-Wong, R., and Woolliams J.A. (1998) *Genet. Sel. Evol.* **30**:313

Rothschild, M.F. and Soller, M. (1997) *Probe* **8**:13

Rothschild, M.F., Jacobson, C., Vaske, D.A., Tuggle, C., Wang, L., Short, T., Eckardt, G., Sasaki, S., Vincent, A., McLaren, D.G., Southwood, O., van der Steen, H., Mileham, A. and Plastow, G. (1996) *Proc. Natl. Acad. Sci.* (USA) **93**:201

Ruane, J. and Colleau, J.J. (1995) *Genetical Research* **66**:71

Ruane, J. and Colleau, J.J. (1996) *J. Dairy Sci.* **79**:1666

Schmitt, K. (1999) In "From Jay Lush to Genomics: Visions for Animal Breeding and Genetics" p.97, editors J.C.M. Dekkers, S. Lamont and M.F. Rothschild, Iowa State University, Ames

Sobral, B.W.S. (1999) In "From Jay Lush to Genomics: Visions for Animal Breeding and Genetics" p.115, editors J.C.M. Dekkers, S. Lamont and M.F. Rothschild, Iowa State University, Ames

Soller, M. and Medjugorac, I. (1999) In "From Jay Lush to Genomics: Visions for Animal Breeding and Genetics" p.85, editors J.C.M. Dekkers, S. Lamont and M.F. Rothschild, Iowa State University, Ames

Spelman, R.J. and Van Arendonk, J.A.M. (1997) *J. Dairy Sci.* **80**:3399

Van Arendonk, J.A.M., Bink, M.C.A.M., Bijma, P., Bovenhuis, H., De Koning, D.J. and Brascamp, E.W. (1999) In "From Jay Lush to Genomics: Visions for Animal Breeding and Genetics" p.60, editors J.C.M. Dekkers, S. Lamont and M.F. Rothschild, Iowa State University, Ames

Van Der Beek, S. and Van Arendonk, J.A.M (1994) *Proc. 5th Wld. Congr. Genet. Appl. Livest. Prod.* **21**:237

Van Der Beek, S., Van Arendonk, J.A.M. and Groen, A.F. (1995) *Theor. Appl. Genet.* **91**:1115

Weller, J.I., Kashi, Y. and Soller, M. (1990) *J. Dairy Sci.* **73**:2525