Strategies to utilize marker-quantitative trait loci associations

Strategies to Utilize Marker-Quantitative Trait Loci Associations

C. S. HALEY*,[1] and P. M. VISSCHER†

*Roslin Institute (Edinburgh), Roslin, Midlothian, United Kingdom EH25 9PS and
†Institute of Ecology and Resource Management, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh, United Kingdom EH9 3JG

ABSTRACT

Marker-assisted selection holds promise because genetic markers provide completely heritable traits that can be measured at any age in either sex and that are potentially correlated with traits of economic value. Theoretical and simulation studies show that the advantage of using marker-assisted selection can be substantial, particularly when marker information is used, because normal selection is less effective, for example, for sex-limited or carcass traits. Assessment of the available information and its most effective use is difficult, but approaches such as crossvalidation may help in this respect. Marker systems are now becoming available that allow the high density of markers required for close associations between marker loci and trait loci. Emerging technologies could allow large numbers of polymorphic sites to be identified, practically guaranteeing that markers will be available that are in complete association with any trait locus. Identifying which polymorphism out of many that is associated with any trait will remain problematic, but multiple-locus disequilibrium measures may allow performance to be associated with unique marker haplotypes. This type of approach, combined with cheap and high density markers, could allow a move from selection based on a combination of "infinitesimal" effects plus individual loci to effective total genomic selection. In such a unified model, each region of the genome would be given its appropriate weight in a breeding program. However, the collection of good quality trait information will remain central to the use of these technologies for the foreseeable future.

(**Key words**: markers, breeding, quantitative trait loci, selection)

Abbreviation key: **AFLP** = amplified fragment length polymorphism, **MAI** = marker-assisted introgression, **MAS** = marker-assisted selection, **PCR** = polymerase chain reaction, **QTL** = quantitative trait locus.

INTRODUCTION

The science of animal breeding has provided effective tools for modifying the performance of animals through selection and crossbreeding. These tools have been efficiently applied on occasion and have resulted in dramatic changes in animal production efficiency and the quality of livestock products (17). In addition, DNA technologies are now delivering new insights into the genomes of livestock as well as other species. The development of genetic markers and maps has allowed some of the loci that affect qualitative and quantitative traits (i.e., major genes and quantitative trait loci, or quantitative trait loci, respectively) to be mapped, and, in some cases, genes controlling variation have been located and causative mutations have been identified. This work is in its infancy, and many more loci will be mapped and the corresponding genes and DNA sequences will be identified and studied over the coming years. The tools provided by the DNA revolution will thus undoubtedly reveal much about the biology of traits of economic importance. However, much uncertainty exists about how this information will affect livestock breeding. In this paper, we do not discuss the vast knowledge of biology or the opportunities arising from the knowledge that the new technology will bring but instead concentrate on the direct impact of genomic information on livestock breeding. We do not attempt to review all studies of marker-assisted selection (**MAS**); rather, we speculate where the technology might be going and give a broad picture of the opportunities that technological advancements may create for animal breeding.

Much of what we say is framed by two papers written by Charlie Smith. The first, published in 1967 (33), was remarkably prescient in predicting the major findings of much recent work on markers and single loci in animal breeding. This study looked at the value of including known loci in livestock improvement. The major conclusions were that 1) "when normal selection is effective . . . information on known

Received July 24, 1997.
Accepted February 23, 1998.
[1]Corresponding author.

loci can add only a little"; 2) when "normal selection is not very effective, as for characters of low heritability, if indirect selection on relatives must be used (as for sex-limited or carcass traits) then known loci may add significantly to the rate of improvement"; and 3) estimation of the effect of loci was important and poor estimates can lead to a loss of progress. These conclusions are still valid and point to the need to use genomic information in novel ways while evaluating carefully the information actually provided by markers.

The second major influence on our speculations about the future was a 1993 paper by Smith and Smith (34). The thesis of Smith and Smith (34) was that the value of MAS in livestock breeding was currently limited by a number of factors. First, with limited numbers of markers, individual markers are usually some distance away from quantitative trait loci (**QTL**) and, hence, linkage disequilibrium only exists at the level of the family and not at the level of the population. Thus, to use these associations, an estimation of effects had to be at the within-family level and would be family specific. Second, as individual markers are not completely informative and recombination occurs between QTL and markers, information collected within a family can quickly degrade if not constantly reestimated. Finally, Smith and Smith (34) argued that the complexity of this situation made the estimation and use of information more intractable, because complex statistical models and significant computer resources are required.

Smith and Smith (34) argued that with a high density of markers the situation is greatly improved. In this case, any gene or QTL of interest is likely to be located close to one or more markers, and linkage disequilibrium created at the time of a founder mutation or by hybridization, drift, or selection is only eroded very slowly by recombination (Figure 1). Hence, genes of interest are likely to be in extreme or complete disequilibrium with closely linked markers. Thus, marker trait associations can be estimated and utilized at the level of the population rather than within families, which not only eases the statistical and computational problems of estimation, but the larger sample size also allows the single population-wide association to be more accurately estimated than the many within-family associations that are estimated with low density maps. Furthermore, with a marker or haplotype of several markers that are in complete association with a gene of interest, markers can be fully informative, and recombination does not erode associations over the generations. As Smith and Smith (34) have pointed out, this situation approaches direct selection on the genes of interest.

Smith and Smith (34) suggested that the way to proceed is to map QTL to low resolution using standard mapping methods and then to increase the resolution of the map in these regions in order to locate more closely linked markers. In fact, future technological developments should make this approach unnecessary and make possible high resolution maps of the whole genome, even, perhaps, to the level of the DNA sequence. In addition to easing the application of selection on loci with appreciable individual effects, we argue further that the level of genomic information available will have an impact on infinitesimal models. Relationship information derived from marker information will replace the standard relationship matrix; thus, the average relationship coefficients that this represents will be replaced by actual relationships. Ultimately, we can envisage that current models combining a few selected QTL with selection on polygenic or infinitesimal effects will be replaced with a unified model in which different regions of the genome are given weights appropriate to the variance they explain. A unified model will give greater control of additive effects, provide the ability to follow combinations of alleles, and give a greater ability to estimate and predict nonadditive gene effects, such as heterosis and inbreeding. We first briefly review current and coming technology and then discuss potential applications.

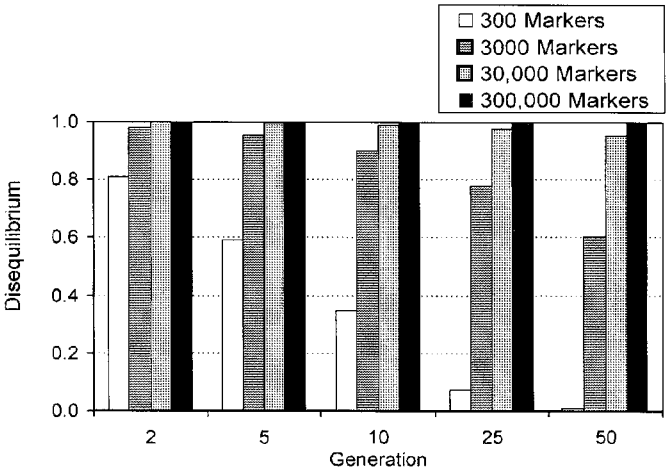

Figure 1. The proportion of retained disequilibrium after generations of random mating. The figures are based on the maximum average distance between marker and quantitative trait locus; the stated number of markers in a typical livestock genome was 30 M and 20 chromosomes. Based on results of Falconer and Mackay (5) and Smith and Smith (34).

MARKER TECHNOLOGY

Marker technology and genetic maps for livestock have developed rapidly in recent years. The first DNA-based markers to be used to any great extent

were restriction fragment length polymorphisms. These types of markers were relatively slow to be generated and had little information usually containing only two alleles in any population. Typing of such markers was slow and tedious, and relatively large amounts of sample DNA were required. The development of the polymerase chain reaction (**PCR**) and microsatellite markers moved the technology forward substantially. With PCR, only very small amounts of sample DNA are needed. Microsatellite markers may have several alleles in a livestock population (good for information content, but increasing the difficulties of interpretation). The use of semi-automated DNA sequencing systems allows increasing throughput of markers and electronic data collection. Microsatellite markers predominate in the current maps of major livestock species (e.g., pigs, cattle, and chickens), and over 1000 microsatellite markers have been mapped in each species. In addition, several hundred markers of other types are also mapped. Thus, around 1500 markers are available in the major species, giving an average marker spacing of around 2 cM (equivalent to about 2% recombination). The distribution of markers is, however, not uniform, so some areas of the genome are more densely marked, and some are less densely marked. The available markers have not yet been fully integrated into single maps for each species, but there is an increasing ability to align maps through subsets of common markers combined with tools to facilitate alignment in databases, such as PiGBASE, which is accessible via the World Wide Web (http://www.ri.bbsrc.ac.uk/pigmap/pig_genome_mapping.html).

The currently available tools have already been put to use. The first genome scans of livestock have been completed and published (1, 9), and many more are underway. These studies have shown that genes with relatively large effects on quantitative traits are segregating both within and between livestock populations. In addition, a number of other loci affecting traits of economic value have been identified and some are already being utilized in breeding programs.

The current marker technology can be effectively applied to map genes of interest. It has been estimated that around 50,000 to 100,000 microsatellite markers may be present in the genome of mammals such as the pig (45), so more markers could be isolated, and map density could be increased. However, identification of each microsatellite is tedious, and, even with multiple marker typing on semi-automated systems, the cost and speed of genotyping are limiting factors. Another gel-based technology, amplified fragment length polymorphisms (**AFLP**) (39), may increase data throughput in some circumstances. This approach combines cutting the DNA with specific restriction enzymes and amplifying it with specific PCR primers prior to generating a fingerprint of around 100 loci simultaneously; a proportion of the loci are variable and, thus, are suitable as markers in any population. Fingerprints made up of different sets of loci can be derived relatively easily by using different primers or different primers and restriction enzymes. Large numbers of different primer combinations and, hence, even larger numbers of marker loci can be produced. Thus, although the method has some limitations, it is possible to generate very large numbers of markers segregating in a particular cross very rapidly. For example, AFLP of pigs have been used to generate rapidly a high density map of the genome and to identify markers linked to the dominant white locus (30). In chickens, a set of seven AFLP primer combinations was used on bulked samples of DNA from either broiler or layer types, and the products were displayed on a single gel. These seven AFLP primer combinations generated around 50 separate markers that apparently were fixed for alternative alleles in the two types of lines (M. Clinton, 1997, personal communication).

GeneChip™ (Affymetrix, Inc., Santa Clara, CA) and related technology may allow an even higher density of information to be extracted from the genome. Photolithographic synthesis can produce a chip with ordered arrays of oligonucleotide sequences that are capable of rapidly interrogating and resequencing individual DNA samples. This technology can be used to resequence long stretches of DNA or to genotype large numbers of simple sequence polymorphisms simultaneously. Already chips have been produced that are capable of resequencing substantial portions of the human BRCA1 gene (15) and of the genome of HIV (24). As the resolution of the technology increases, it may be possible to put the entire sequence of the human genome on a few such chips (2). Harnessing the power of this technology may first come about through chips that can genotype diallelic single nucleotide polymorphisms. Such polymorphisms in which there are two relatively common alleles may occur every 1000 bp in humans. Variability, in Holstein cattle at least, has been estimated to be one-third of that existing in humans (8), but, even in this case, around 3 million of such polymorphisms are available across the genome. Their diallelic nature makes such polymorphisms intrinsically less informative than microsatellites, but increasing the number of diallelic polymorphisms used would easily compensate for this. A chip is planned that will be capable of simultaneously genotyping 1000 single nucleotide polymorphisms in the human genome and thus, provide a tool for rapid genome scanning.

The ultimate information is provided by the DNA sequence itself. A conceivable progression of the technology is to the provision of the complete sequence of individuals, which, of course, not only produces neutral polymorphic markers, but also provides information on the variants with a phenotypic effect. If this technology develops, the challenge, just as great in humans as it is in livestock, is to discriminate between those variants with an effect and those without an effect. Once such technologies have taken firm root in human genomics and their developmental costs have been absorbed, these technologies will be transferred to other species, including livestock, providing more information at decreasing cost. The question remains, however, whether animal breeders who are given access to such vast amounts of information, at costs equivalent or less than those associated with a performance test, will be able to use it, and if so, how.

MAS

A number of theoretical and simulation studies have been performed to date that focus on the manipulation of genes or QTL of large effect using markers. In addition, studies have been performed that have used markers to confer information about the background genotype (i.e., that part of the genome containing genes affecting traits of interest but not known to contain any loci of large enough effect to be worth manipulating individually). When genes of appreciable effect have been marked, studies can broadly be divided into those in which disequilibrium between markers and genes is assumed to exist at the population level and those in which such disequilibrium only exists at the within-family level. We start with the latter case, which provides some lessons for the value of MAS, and move to look at studies that assume population-wide disequilibrium, which, as we have argued, may be more readily accessible in future.

Within-family Selection

A common assumption is that only a low marker density is possible and, thus, that within an outbred population, markers and QTL are likely in linkage equilibrium. Thus, even for a marker that is closely linked to a QTL across the population, no overall association exists between the marker and the trait affected by the QTL. If the focus, however, is not on the population level but on the family level, then the linkage disequilibrium reappears. Therefore, MAS can be applied by aiming at within-family associations.

Because the breeding value of an animal is the family average plus its individual deviation from the family average, usually sufficient information is available to estimate the family average. The within-family deviation is estimated based on own performance only, is estimable only after waiting a long period (progeny testing), or may not be estimable at all (e.g., in nucleus breeding for sex-limited or carcass-limited traits for which no animal record is available). It follows that the efficiency of MAS in an outbred population may not be greatly hampered by the fact that between-family differences are not picked up by the marker, as these are already relatively well estimated. The efficiency of MAS depends only on the ability to explain within-family variance. The variance explained by the marker depends on two factors: 1) the effect of the QTL linked to the marker and 2) the available information to estimate the effects of the marker alleles segregating within a family. The ability to estimate the effect of a marker is a major determinant of the efficiency of MAS. We can, therefore, safely predict that within families in which the number of offspring is limited the amount of information is a critical parameter.

Increases in selection response from MAS within outbred populations obtained from the simulation studies described in the literature vary from zero to more than 60%. Ruane and Colleau (32) found no additional response when selecting for a trait that was independent of sex and carcass and when performance was tested before selection. Meuwissen and Goddard (27) found an increase in response of 64% when MAS was for a carcass trait and when selection was performed before the trait was measured. In the latter case, an important assumption was that, at the moment MAS was implemented, marker genotype information was available for five ancestral generations. For a trait expressed in all animals but recorded after selection, the additional response was 37% when five ancestral generations were recorded, but only 6% when only the grandparent generations were marker typed. The results of Ruane and Colleau (32) and Meuwissen and Goddard (27) were for nucleus breeding. Van der Beek and Van Arendonk (36) studied an outbred poultry breeding nucleus with selection for a sex-limited trait by which only a limited number of male full sibs were allowed to be selected. In their study, the additional response to MAS was 6 to 12%, and most of the additional response was due to the replacement of a random selection of male full sibs by a marker-assisted preselection step within each full-sib family.

One potential for MAS is the preselection of young dairy bulls. The preselection of dairy bulls before they

enter a progeny test is very appealing, because this procedure applies selection where previously choices would have been random. For example, if only pedigree information on estimated breeding values is available, the choice between any number of full-sib young bulls is completely random. The use of marker information makes it possible to distinguish among these brothers and to select the most promising ones. A particular attraction is that such preselection is associated with relatively low risk. If associations with markers are actually spurious, then preselection based on markers is equivalent to random selection and no less progress is made than would have been made without use of markers (13). Kashi et al. (23) studied a progeny-testing scheme in which markers were used to select the young sires that enter the progeny test, (i.e., markers were used in an extra selection step to preselect young bulls). This procedure resulted in an extra response of up to 30%. However, by the time the QTL are found, the granddaughters will already have completed a lactation, and any preselection of bulls can only take place among the great-grandprogeny of the elite sire. An alternative that avoids this problem is provided by Mackinnon and Georges (26), who proposed the use of MAS in progeny-testing schemes using preselection of young bulls based upon marker genotypes. In this scheme, QTL are presumed to have been detected a priori (for example, in a granddaughter design), and those sires that are heterozygous for QTL are identified from daughter information. This information is then used to preselect which sons of heterozygous sires undergo progeny testing, and no preselection is practiced in the sons of homozygous sires. Additional progress from this scheme is predicted to be between 8 to 23% (with 1 to 5 QTL selected) over that obtained from progeny testing alone.

The results of these studies emphasize that, even when within-family associations must be used, substantial extra progress can be achieved. They also confirm the point made by Smith (33) that appreciable extra progress is made only when marker information is used at times when normal selection is less effective.

Within-population Selection

As we have noted, Smith and Smith (34) argued that effort should be devoted to finding markers that are so close to the QTL that recombination between them and the QTL can be ignored, and selection can occur across families. One way to achieve this goal with a relatively low density marker map is by crossing two different lines, which may generate widespread disequilibrium, at least for the early generations. Crossing two inbred lines generates complete disequilibrium initially. In livestock, crosses between outbred lines also generate complete disequilibrium when markers and QTL are fixed for alternative alleles in the two founder lines, even if the lines themselves are outbred. For markers and QTL that are segregating within a line, disequilibrium is unlikely to be complete unless high density marker maps are available such that the disequilibrium created by past hybridization and founder events has not decayed. Studies that assume complete disequilibrium generated by a line cross will become more relevant to livestock as increasing marker density provides more opportunity to exploit such disequilibrium both within and between lines.

Marker-assisted selection from a line cross has been studied by Lande and Thompson (25) using theoretical derivations and by Zheng and Smith (47, 48) and Gimelfarb and Lande (11) using simulation results. All studies assumed linkage disequilibrium throughout the population and considered the simplest situation, a cross between inbred lines. Marker-assisted selection from the intercross (an F_2 progeny in these studies) then proceeds in two phases: 1) a number of markers associated with the trait or traits of interest are chosen based on estimates from a marker quantitative trait association study (e.g., by analysis of associations in the F_2 population), and 2) animals are selected based on marker genotype and phenotypes (or the phenotypes of their relatives). For the first phase, an important question is how to select the markers that explain some of the genetic variance in the population. A simple multiple regression approach results in an overestimate of the total variance explained by the markers (46). Therefore, Lande and Thompson (25) proposed a two-stage procedure. In the first stage, a set of promising markers is selected from all available markers. In the second stage, estimates of the regression coefficients for the selected markers are obtained from a new (independent) sample of animals. In this case, the marker effects are unbiased (25). However, as pointed out by Visscher (37), the total amount of variance explained by the selected group of markers may still be biased upward.

Lande and Thompson (25) give examples of the theoretical relative efficiency of MAS (including both marker information and phenotypic observations) compared with traditional index selection and conclude that the benefits may be very large. A number of simulation studies suggest that the theoretical predictions of Lande and Thompson (25) are too optimistic, especially if the time horizon is more than one generation. Using simulation, Gimelfarb and

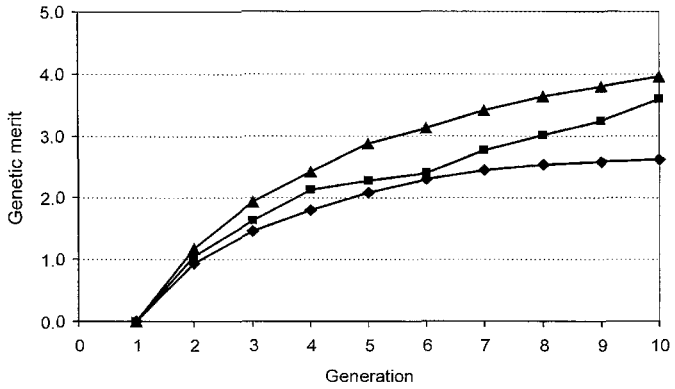

Figure 2. Simulated selection in a synthetic population starting from the F_2 population created by crossing two inbred lines for a trait with an initial heritability of 0.1. Selection was on markers only (♦), on BLUP based on phenotypic information (■), or an index combining these two sources of information (▲). Based on results of Zhang and Smith (47).

Lande (11) contrasted MAS with phenotypic selection alone and found that MAS was more efficient for at least 5 generations of selection for a single trait. If the regression coefficients for selected markers were reestimated for each generation, MAS was more efficient for the first 10 generations of selection. Surprisingly, almost all of the weight in the index combining marker information and phenotypes was put on the marker information (11). Even with reestimation of the marker effects, this emphasis is unexpected because, after a few generations of selection, most genetic variance will be within rather than between marker genotypes.

Zhang and Smith (47, 48) compared the efficiency of MAS (using only marker information) with selection on BLUP breeding values and selection on an optimum combination of marker and phenotypic information. Marker-assisted selection clearly showed an increased efficiency of 10 to 30% of combined selection over the best alternative, BLUP selection, in the early generations. By generation 5, combined selection could have reached a phenotypic value not achieved by BLUP selection for two to three more generations (Figure 2). Nevertheless, the relative advantage of using markers was less in later generations, in part because recombination cased a decay of the associations between markers and QTL in the low density map used (five markers per Morgan). Zhang and Smith (47) concluded that MAS using linkage disequilibrium has limitations until close linkages of markers and QTL are available. Zhang and Smith (47) showed the loss in efficiency of estimating the marker effects with imprecision. For example, reducing the population size in which marker QTL associations were estimated from 1000 to 100 reduced response to selection by about 50%. Similarly, Gimelfarb and Lande (11) concluded that population size was the main parameter in determining the relative efficiency of MAS. In contrast, Lande and Thompson (25) concluded that estimating the marker effects with error did not significantly reduce the response to selection. However, Lande and Thompson (25) assumed that the appropriate markers (i.e., the ones most closely associated with QTL) were included in the selection index and that only the genetic variance associated with the markers was subject to error.

Whittaker et al. (41, 42) have looked at ways to select the appropriate markers and to estimate their effects. Others (25, 47, 48) have suggested identifying which markers to use in one set of data and then estimating marker effects in a second set. However, this procedure can be practically carried out only in the first generation and even then is wasteful of data. Gimelfarb and Lande (11) have shown that reestimating effects every generation is more effective than estimating them only once at the start of the selection. Whittaker et al. (42) employed a crossvalidation procedure that can be used to select markers and estimate their effects for every generation. In this procedure, the data are split into halves. In the first half, markers are selected, and the marker effects are estimated. These estimates are used to calculate the marker scores for individuals in the second half of the data, and the covariance between the marker score and the genotype is estimated as the phenotypic covariance in this second half of the data. The procedure can be repeated by swapping the roles of the data halves, and an overall estimate can be obtained by combining the results from the halves.

Whittaker et al. (42) showed that an index that was calculated using crossvalidation gave better responses than previously used indices (Figure 3). Indeed, the crossvalidation index gave as good a selection response as an index in which the true covariances between the selected markers and genotype were used. However, even better responses could have been obtained if markers had been both selected and used on the basis of knowledge of the true covariances between markers and genotype. This last approach shows that better methods might make even more progress, although even in this latter approach the recombination between markers and QTL would have reduced progress in later generations.

The studies of selection from hybridized populations have tended to focus on using MAS to complement selection in cases in which the use of phenotypic information is relatively efficient (e.g., when phenotypic information is available on both sexes at the

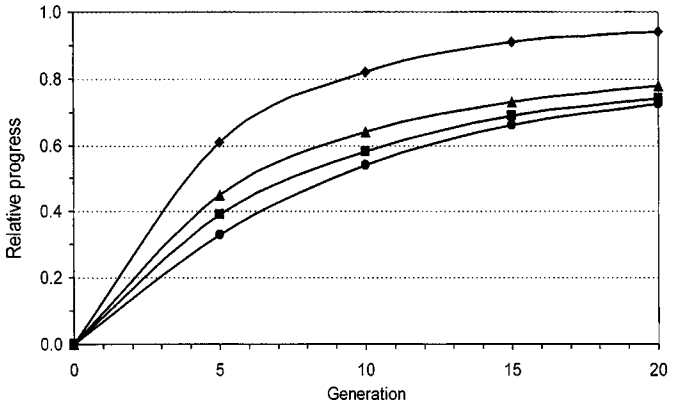

Figure 3. Simulated selection in a synthetic population starting from the F_2 population created by crossing two inbred lines for a trait with an initial heritability of 0.1. Selection was on phenotype (●), on a crossvalidation index (▲), on the type of index used by Lande and Thompson (25) and Gimelfarb and Lande (11) (■), or on the true covariances between markers and genotype (♦). The curve of the crossvalidation index overlies that obtained using the actual covariances between the selected markers and genotype. Progress is shown relative to the absolute maximum attainable if all beneficial alleles were fixed. Data are from Whittaker et al. (42), except for the Lande and Thompson (25) index, which is from Whittaker et al. (41).

time of selection). As noted previously, this situation is the least favorable for use of marker information. Even in these cases, substantial extra progress can be achieved from the use of markers. Although studies of the hybridization of inbred lines are not directly relevant to livestock, such studies do indicate what may be possible once the marker density increases to a level at which linkage disequilibrium between markers and QTL can be relied upon to exist even within a population. At this time, the focus is on detecting which of the many available markers are in disequilibrium with the genes of interest and on estimating the marker contributions to variation. The results from the crossvalidation approach of Whittaker et al. (42) provide a promising approach to the problem of correctly weighting marker information and potentially for correctly weighting information from other indirect indicators, such as type traits.

Marker-assisted Introgression

Introgression is appropriate when the only requirement is to transfer alleles at one or a few loci from one breed to another. In a program using marker-assisted introgression (**MAI**), genetic markers could be used in two ways: 1) to help identify animals carrying the allele that is to be introgressed and 2) to select for (or against) a particular background genotype. The route usually proposed (19, 20, 14) is the backcrossing of a number of generations of animals that are carrying the desired allele (from the donor population, the inferior breed) with a recipient population (the commercial breed). This process is followed by an inter se cross to make the desired allele homozygous (Table 1).

The efficiency of the MAI program depends on the frequency of the introgressed allele in the final population and the genetic progress for traits of economic benefit. In a backcross selection program in which no selection is practiced, the proportion of the genes from the donor line would halve each generation. Many studies have assumed that the allele to be introgressed can be identified without error from the phenotype; thus, the frequency of the allele during the backcross phase remains at 50%. In practice, the allele to be introgressed often cannot be directly identified from the phenotype of the animal. For example, the allele may be the recessive allele of a major gene, its expression may be limited by age or sex, or it may be at a QTL. In this case, the marker or markers linked to the allele of interest may be used in an attempt to maintain its frequency. Using only a single marker, the frequency of the desired allele may drop below 50% in the backcross generations because of recombination between the marker and the introgressed locus (38). Using flanking markers is more efficient because a double recombination event is needed before the desired allele is lost from between the markers. The problem of maintaining an allele using linked markers is made more difficult when the QTL have an imprecisely known position. In Figure 4, an example is given of the effect of not knowing the exact location of a QTL to be introgressed. We assumed that the location of the trait gene was estimated with a standard deviation of 6 cM and that the population size was 800. Clearly, selection on a single marker that is thought to be close to the gene to be introgressed can be risky because the frequency of the number of animals that carry one copy of the desired

TABLE 1. Crosses performed and genotypes produced in a backcross introgression program.

Cross[1]	Population	Genotypes produced	Genotype selected
D × R	F_1	Qq	Qq
F_1 × R	BC_1	Qq qq	Qq
BC_1 × R	BC_2	Qq qq	Qq
BC_{k-1} × R	BC_k	Qq qq	Qq
BC_k × BC_k	G_{k+1}	Qq qq QQ	QQ
G_{k+1} × G_{k+1}	G_{k+2}	QQ	QQ

[1]Donor (D) genotype = QQ, recipient (R) genotype = qq, BC = backcross, and G = generation.

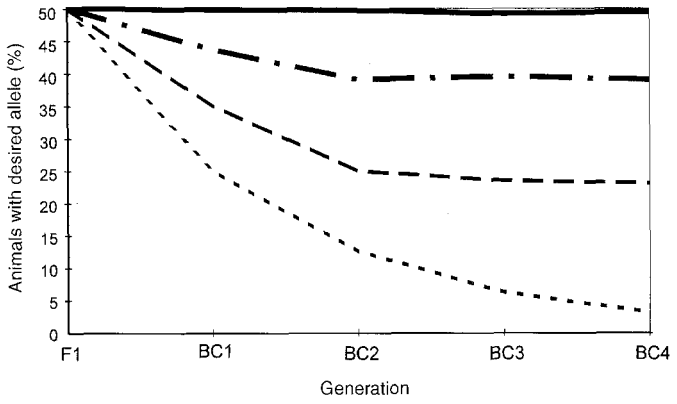

Figure 4. Marker-assisted introgression. The frequency of an introgressed QTL allele in generations from F_1 to backcross (BC) generation BC_4. Lines represent random selection (– – –), selection using one linked marker (— — —), selection using two linked markers (— · —), and selection using a four-marker haplotype (————). Data are from Visscher et al. (38).

allele has dropped to about 20% by generation 5. Recently, Hospital and Charcosset (20) have extended these studies to look at the control of multiple QTL in an introgression program. As the number of QTL being manipulated increases, there are fewer individuals carrying the desired allele at all loci. Hospital and Charcosset (20) concluded that up to four QTL could be controlled in populations of reasonable size (a few hundred individuals).

During the backcrossing phase and intercrossing phases, candidates for selection (i.e., animals with the allele to be introgressed) vary in proportions of the genome that is from the donor and recipient lines. For example, the theoretical range in the proportion of the genome that comes from the recipient line is 81 to 94% in the second backcross generation (Table 2). To speed recovery of the recipient genome, selection can take place on this background genotype. The use of markers for this type of selection has the advantages that no phenotypic data need to be collected and that markers can be scored on individuals of both sexes early in life. Hospital et al. (21), Visscher et al. (38), and Hospital and Charcosset (20) showed that progress that was equivalent to one or two generations of backcrossing can be gained by using this approach.

The donor line or breed may be inferior for other traits of economic importance. For example, although the Meishan pig is superior to commercial European breeds for litter size and related traits, it is inferior with respect to lean growth and fatness traits. One criterion for the efficacy of an MAI program is the performance of the animals carrying the introgressed allele compared with the mean of the recipient line at the start of the program. However, during the backcrossing and intercrossing phases, the recipient (commercial) line undergoes selection, and, thus, a better comparison is one between the population carrying the allele and the commercial population at the same time point. This structure is analogous to the problem studied by Haley (16) and studied extensively by Gama et al. (6), who calculated the genetic lag for economic performance for various backcross and intercross programs when a transgene was introgressed into a nucleus population of pigs. Gama et al. (6) assumed that the transgene genotype was known and that genetic markers had not been used to distinguish between the background genome of the founder transgenic animal and the rest of the population. Gama et al. (6) concluded that a gene would need an economic effect equivalent to one to two generations of selection to make its introgression worthwhile. The major contribution to the overall genetic lag of the final new commercial product (in which animals carry two copies of the desired allele) is not, as commonly thought, from the initial breed difference or the selection lost during backcrossing, but rather is from the last two generations of intercrossing to make the animals homozygous for the desired allele. Hence, to reduce that lag, increasing the population size or the intensity of selection, for example, through new reproductive technologies such as embryo transfer should be considered even for a prolific species such as pigs.

Genomic Selection

The ability to use markers to speed the recovery of the recipient genotype in a MAI program demonstrates that the standard relationship matrix describes an average appropriate for the infinitesimal model. Under the finite locus reality, variation occurs around this average, and marker information can be used to estimate this variation. This process is exactly that followed when MAS is implemented for in-

TABLE 2. Variation in the proportion of genome from the recipient in a backcrossing program for a typical livestock genome (20 chromosomes, each 100 cM long) (18).

Backcross	$\bar{X}$	95th Percentile range
	(%)	
1	75	66.6–83.4
2	87.5	80.7–94.3
3	93.75	88.8–98.8
4	96.88	93.5–100.0

dividual loci—the marker information is used to predict how the inheritance of the locus deviates from the average expectation. At the level of individual loci, the deviation from the average expectation is more extreme (i.e., an allele is inherited or not) than across the genome as a whole (Table 2). However, using the actual relationship matrix might improve the accuracy of breeding value estimation and, hence, improve genetic progress even in models in which large numbers of loci of small effect are acting.

Nejati-Javaremi et al. (28) looked at a situation in which all loci contributing to the phenotype of interest were identified. They simulated a small population with a limited number of half-sibs and a trait controlled by 5, 10, or 100 loci. Use of the actual allelic relationships between individuals rather than the average value increased the correlation between the estimated and the true breeding values by over 60% when 5 loci controlled the trait and by around 9% when 100 loci controlled the trait.

The results of Nejati-Javaremi et al. (28) must be considered as upper limits because of the assumption that genotypes at QTL themselves were known. Jørgensen and Jensen (22), however, simulated the use of markers to predict genomic relationships. Marker density and information content were used at a range from 50-cM spaced markers with two alleles up to 10-cM spaced markers with 16 alleles and with 10 or 1000 QTL. The correlation between the estimated and true breeding values were generally increased from 1 to 3% by the use of marker information. The higher values were found for the more informative marker maps, and the results from the two studies of Nejati-Javaremi et al. (28) and Jørgensen and Jensen (22) may converge at high marker densities. Furthermore, those two studies did not look at the use of the marker information in novel selection schemes. The ability to estimate breeding values under such models prior to the availability of phenotypic information could open the door to breeding schemes in which the marker information is used when normal selection is less effective. As we have demonstrated previously, this ability can lead to substantial additional genetic progress.

Results from genome-based selection depend not only upon the marker information that is assumed but also on the genetic model for the traits of interest. Rather than consider separately the marker-based selection of individual loci and of many loci of small effect, we can anticipate their unification in a common framework. With high density marker information giving essentially complete information on relationships, different genomic regions could be given the weight appropriate to the variation controlled. Thus, selection for a limited number of detectable QTL can be complemented by genomic selection aimed at the residual genetic variation that is spread over the remainder of the genome.

Other Issues

All studies report a decline in additional response from MAS over generations, which is expected because additional response is due to an accelerated increase of the frequencies of favorable QTL alleles. As soon as the favorable allele reaches a frequency above 0.5, an accelerated increase of QTL allele frequency results in an accelerated reduction of variance explained by the QTL and, therefore, in reduced impact of MAS. Meanwhile, because of improved selection for the marked QTL, more selection intensity is focused on the QTL, and, consequently, less is focused on the other genes. Therefore, over the long term, MAS might even result in a reduced cumulative genetic gain because MAS only results in an accelerated fixation of the QTL, which is an improvement in time, but not in the absolute level (as phenotypic selection would also ultimately fix the QTL), and decreases the polygenic response. This phenomenon was most clearly demonstrated by Gibson (10). From this observation, no general conclusion on the effect of MAS on long-term response should be drawn. First, reduced long-term response does not occur when MAS is applied at a stage or within a group of animals where normally no selection or random selection is applied. For example, preselection of young bulls entering a progeny test does not influence the selection pressure or the response for genes not associated with a marker. Second, adapted selection strategies can be introduced. Instead of selecting animals with the maximum estimated breeding value, an index can be created in which the weight of selection on QTL is reduced. In addition, in most breeding situations, the short-term responses to selection determine the success of the breeding program. Even if conventional selection is equivalent by 10 generations, this fact is of little relevance to most commercial breeders.

There are several issues regarding gene action that may assume particular relevance in the exploitation of QTL. Obviously, the breeder needs to consider the possibility of deleterious pleiotropic effects of a selected QTL. Such effects could in some instances explain why QTL with large effects on traits under selection have not been fixed. The halothane-ryanodine receptor gene provides a good example of such a situation (40). More difficult is the possibility of interactions of a QTL with the environment or with

the background genotype. Interactions of genotype and environment have not proven to be a major problem for livestock kept in intensive production systems. However, major interactions of genotype and environment can occur when genotypes are placed in more challenging extensive environments. Interactions between genotypes possible will be a problem for QTL of a major effect. For example, will a QTL with a major effect on body fat have the same effect in a lean genetic background as it does in a fat genetic background? There is little evidence on these questions for QTL in livestock, and we need to await the accumulating results of ongoing research. However, experience with major genes, such as the halothane gene (40) and the Booroola gene (29), would suggest that interactions with both environment and genotype can occur, but the main effects of the loci occur across a range of environments and genetic backgrounds.

We have thus far ignored the possibilities of manipulating nonadditive variation using markers. It is this area, however, in which the infinitesimal model is by its very nature poorly suited. With the detailed knowledge of genomic relationships provided by high density marker coverage, there are possibilities for estimating and controlling inbreeding and, hence, inbreeding depression and heterosis in the whole genome or in particular regions of the genome. With a model incorporating a finite number of loci, dominance at individual loci can be made consistent with the occurrence of inbreeding depression and heterosis in a way that is not possible under the infinitesimal model (31). Incorporation of genomic information into such a model may allow better estimation of the dominance effects and also better predict the consequences of selecting and mating two individuals. The work of Gavora et al. (7) shows that surprisingly high correlations can be obtained for genetic distances between lines that are estimated using markers and heterosis that results when the lines are crossed. It will be a challenge to see whether such information can be used both at the level of line crossing and at the level of mating individuals within or between lines.

MAS AND NEW TECHNOLOGIES

As shown in the previous sections, the benefits from MAS were highest when used at stages or in groups of animals where no selection was previously possible. The introduction of a new technology such as MAS requires a reoptimization of the breeding program and should provoke breeders to alter their breeding strategies. One of the first alterations suggested was preselection of young bulls entering the progeny test. A second alteration was to select before recording instead of after recording, which shows much promise. Georges and Massey (8) took this idea to the extreme by suggesting "velogenetics". This scheme uses "velogenesis" (3) in which the generation interval is greatly reduced by harvesting oocytes from calves while still in utero. The harvested oocytes are matured and fertilized in vitro prior to being transferred to a recipient female. The process can be repeated by harvesting oocytes from these second generation animals with the generation interval being reduced to around 3 to 6 mo. In velogenetics, markers are added to such a scheme either to select the in utero calves from which to harvest oocytes or to type one or two cells from embryos while they are still in vitro to determine which to transfer to recipient females (Figure 5a). In principle, this procedure could be used for several successive generations, with each generation using markers for selection without ever generating an adult animal or measuring a phenotype. Hence, velogenetics might be used to introgress rapidly a gene from one breed to another using semen from the recipient line each generation.

The combination of MAS and embryo technologies could be further enhanced by technologies currently under development, such as nuclear transfer (4, 44). In a velogenetic scheme, the selection of embryos based on marker information prior to transfer may be difficult because of the limited number of cells that can be harvested without damaging the embryo. An alternative would be to culture the embryo in vitro until a sufficient number of cells are available for effective marker genotyping. After the desired cultures have been identified on the basis of the marker genotype, nuclear transfer would be used for the remaining cells into an enucleated oocyte to regenerate one or more of the desired embryos for transfer into a recipient female (Figure 5b).

A major potential drawback in difficulty, cost, and welfare of these velogenetic schemes is the need to harvest oocytes from calves in utero. If the technology eventually develops to a stage at which cell differentiation can be controlled in vitro, then in vitro meiosis, followed by fertilization, may become possible. In this case, the step requiring transfer to the recipient female, followed by oocyte harvesting, would become redundant. Cell cultures derived from fertilized oocytes could be selected using markers and then induced to undergo meiosis. After fertilization (not necessarily via a true oocyte), the resulting cultures could again be selected on marker information, and the process could be repeated (Figure 5c). After a number of generations, once the desired genotypes

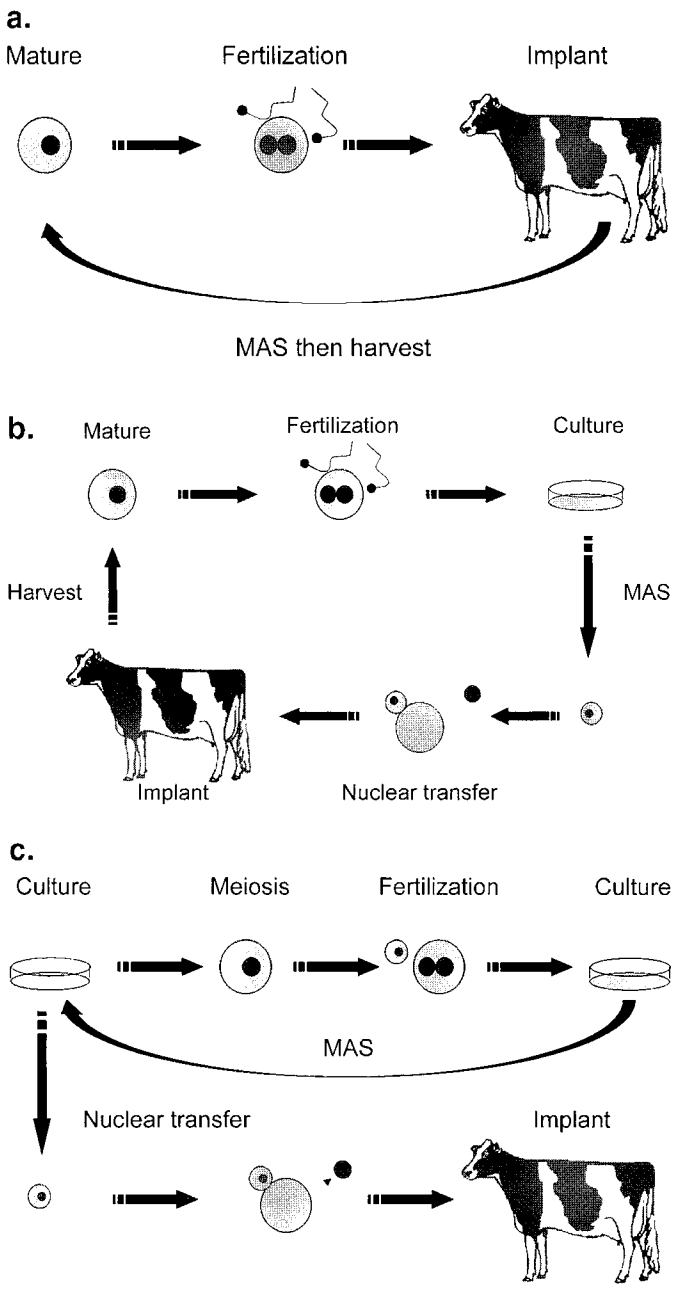

Figure 5. Cell technologies and marker-assisted selection (MAS). a. Velogenetics. Calves are selected in utero using marker information, and oocytes are harvested. The oocytes are matured, fertilized, and implanted, and MAS again is applied to in utero calves to repeat the cycle (8). b. Nuclear velogenetics. Embryos are cultured in vitro and selected using marker information. Nuclear transfer from selected cultures is used to generate new embryos for implantation. Oocytes are harvested from calves in utero and matured, fertilized, and cultured in vitro to repeat the cycle. c. Whizzogenetics. Embryos are cultured in vitro and selected using marker information. Selected cultures are induced to undergo meiosis, and the resulting cells are fertilized and recultured in vitro. Marker information is used to select cultures to repeat the process. Once desired genotypes are achieved, nuclear transfer from selected cultures is used to generate new embryos for implantation (43).

are attained, animals are regenerated, possibly via nuclear transfer (43). Such a scheme would allow for rapid introgression when genotypic information alone is used, for example, introgression of the polled gene into Holstein-Friesian cattle with little genetic lag. With high density marker maps and knowledge of close marker-QTL associations, more generalized selection objectives could also be tackled in periods of less than one natural cattle generation.

CONCLUSIONS

The work we have reviewed in this paper shows some of the potential advantages of using marker information in breeding programs. Even with relatively low density maps, the use of marker information when more traditional selection is less effective can lead to substantial additional progress (26, 27). The results from studies of selection from synthetic populations, although not directly relevant to livestock, show that marker information can enhance progress even when phenotypic information is relatively efficient if disequilibrium is present throughout the population. We have argued that, as marker technologies improve, high density marker information gives access to disequilibrium between markers and trait loci for the entire population. In addition, such high density marker information provides good estimates of the true genomic relationships between animals that are known to be related (and possibly those that are not known to be related also). In addition to improving heritabilities and selection accuracy through the elimination of pedigree errors, such information could be used to allow selection based on a model that unites detectable QTL with background polygenic effects. In such a unified model, appropriate weight in breeding programs can be given to those genomic regions associated with detectable effects and to those regions that account for the undetectable residual genetic variance.

With information about very high density markers, the estimation of marker effects remains a large problem. With thousands or even tens or hundreds of thousands of available markers, selecting those markers most associated with QTL effects is difficult. One solution could be to use an approach analogous to that of Terwilliger (35) to identify the region most associated with the effect. This approach looks for disequilibrium with multiple-marker polymorphisms, producing a likelihood plot similar to that obtained from interval mapping that has peaks identifying the strongest associations in the genome. This sort of approach is required because, unlike the situation for

crosses between inbred lines, in an outbred population, the variance is high between the marker loci in their disequilibrium with a trait locus. Thus, a method is needed to account for this variance, and that method cannot simply be to look at associations with each marker locus in turn. Goddard (12) investigated regression onto several linked loci jointly and found that, with close linkage, marker haplotypes around a QTL could explain a high proportion of the QTL variance (up to 60%). However, Goddard looked at only three linked markers and recombination fractions between loci down to 2×10^{-3} (optimistic parameters at the time he was writing). If the high density of polymorphisms that are present in the genome are used, there may be 200 markers within this distance of a QTL with the closest being at an average recombination frequency of 10^{-5} from the QTL. Potentially, nearly all of the variance could be explained by such markers, but further investigation of these scenarios by theory and simulation is required. Following these approaches, we can envisage identifying a number of multiple-marker haplotypes associated with phenotypic performance throughout the genome, followed by the use of a crossvalidation procedure to give them appropriate weight. These haplotypes would then represent the very close linkages envisaged by Smith and Smith (34), which could be used in breeding programs (with refined estimates of their effect as data accumulates) for which normal selection is less effective.

High density marker maps will provide a wealth of new information on the genome. Animal breeders need to be flexible and creative in the use of this information. This information will not replace information that is already being collected because it will be a long time before phenotype can be predicted solely from DNA sequence. Thus, collection of good performance information remains crucial for the foreseeable future. However, high density marker information may to some extent supplement or even replace pedigree recording. We think that MAS has much to offer animal breeding when used with other new technologies and when used carefully to complement selection based on phenotype. Rapidly developing genome technologies mean that the close linkages envisaged by Smith and Smith (34) will soon be the rule rather than the exception. The challenge for animal breeders is to be imaginative in designing ways of using this information. When the technology has developed to the stage that every analysis of milk composition is returned with the complete DNA sequence of the animal concerned at little extra cost, dairy scientists need to be ready to use this information in the most effective manner.

ACKNOWLEDGMENTS

This work was stimulated by, and owes much to, the writings and ideas of Charles Smith. We are also grateful to others, including Philippe Baret, Keith Campbell, John Gibson, Mike Goddard, Sara Knott, Graham Plastow, Ian Wilmut,and John Woolliams for past stimulating discussions. Our work is supported by the Marker Assisted Selection Consortium of the British Pig Industry, Biotechnology and Biological Sciences Research Council, Ministry of Agriculture Fisheries and Food, and the University of Edinburgh.

REFERENCES

1 Andersson, L., C. S. Haley, H. Ellegren, S. A. Knott, M. Johansson, K. Andersson, L. Andersson-Eklund, I. Edfors-Lilja, M. Fredholm, I. Hansson, J. Håkansson, and K. Lundström. 1994. Genetic mapping of quantitative trait loci for growth and fatness in pigs. Science (Washington, DC) 263:1771–1774.

2 Anonymous. 1996. To affinity and beyond. Nat. Genet. 14: 367–370.

3 Betteridge, K. J., C. Smith, R. B. Stubbings, K. P. Xu, and W. A. King. 1989. Potential genetic improvement of cattle by fertilization of fetal oocytes *in vitro*. J. Reprod. Fertil. (Suppl.)34: 87–98.

4 Campbell, K.H.S., J. McWhir, W. A. Ritchie, and I. Wilmut, 1996. Sheep cloned by nuclear transfer from a cultured cell line. Nature (Lond.) 380:64–66.

5 Falconer, D. S., and T.F.C. Mackay. 1996. Introduction to Quantitative Genetics. 4th ed. Longman Sel., Harlow, United Kingdom.

6 Gama, L. T., C. Smith, and J. P. Gibson. 1992. Transgene effects, introgression strategies and testing schemes in pigs. Anim. Prod. 54:427–440.

7 Gavora, J. S., R. W. Fairfull, B. F. Benkel, W. J. Cantwell, and J. R. Chambers. 1996. Prediction of heterosis from DNA fingerprints in chickens. Genetics 144:777–784.

8 Georges, M., and J. M. Massey. 1991. Velogenetics, or the synergistic use of marker assisted selection and germ-line manipulation. Theriogenology 35:151–159.

9 Georges, M., D. Nielsen, M. Mackinnon, A. Mishra, R. Okimoto, A. T. Pasquino, L. S. Sargeant, A. Sorensen, M. R. Steele, X. Zao, J. E. Womack, and I. Hoeschle. 1995. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics 139:907–920.

10 Gibson, J. P. 1994. Short-term gain at the expense of long-term response with selection of identified loci. Proc. 5th World Congr. Genet. Appl. Livestock Prod., Guelph, ON, Canada 21:201–204.

11 Gimelfarb, A., and R. Lande. 1994. Simulation of marker assisted selection in hybrid populations. Genet. Res. 63:39–47.

12 Goddard, M. E. 1991. Mapping genes for quantitative traits using linkage disequilibrium. Genet. Sel. Evol. 23(Suppl. 1): 131s–134s.

13 Gomez-Raya, L., and J. P. Gibson. 1993. Within-family selection at an otherwise unselected locus in dairy cattle. Genome 36: 433–439.

14 Groen, A. F., and C. Smith. 1995. A stochastic simulation of the efficiency of marker-assisted introgression in livestock. J. Anim. Breed. Genet. 112:161–170.

15 Hacia, J. G., L. C. Brody, M. S. Chee, S.P.A. Fodor, and F. S. Collins. 1996. Detection of heterozygous mutations in BRCA1

using high-density oligonucleotide arrays and 2-color fluorescence analysis. Nat. Genet. 14:441–447.
16 Haley, C. S. 1991. Considerations in the development of future pig breeding programs. Australasian J. Anim. Sci. 4:305–328.
17 Havenstein, G. B., P. R. Ferket, S. E. Scheideler, and D. V. Rives. 1994. Carcass composition and yield of 1991 vs. 1957 broilers when fed typical 1957 and 1991 broiler diets. Poult. Sci. 73:1795–1804.
18 Hill, W. G. 1993. Variation in genetic composition in backcrossing programs. J. Hered. 84:212–213.
19 Hillel, J., T. Schaap, A. Haberfield, A. J. Jeffreys, Y. Plotzky, A. Cahaner, and U. Lavi. 1990. DNA fingerprints applied to gene introgression in breeding programs. Genetics 124:783–789.
20 Hospital, F., and A. Charcosset. 1997. Marker-assisted introgression of quantitative trait loci. Genetics 147:1469–1485.
21 Hospital, F., C. Chevalet, and P. Mulsant. 1992. Using markers in gene introgression breeding programs. Genetics 132: 1199–1210.
22 Jørgensen, J. N., and P, Jensen. 1996. Genome relationship in breeding value estimation. 47th Mtg. Eur. Assoc. Anim. Prod., Lillehammer, Norway. Circulated paper.
23 Kashi, Y., E. Hallerman, and M. Soller. 1990. Marker-assisted selection of candidate bulls for progeny testing programmes. Anim. Prod. 51:63–74.
24 Kozal, M. J., N. Shah, N. P. Shen, R. Yang, R. Fucini, T. C. Merigan, D. D. Richman, D. Morris, E. R. Hubbell, M. Chee, and T. R. Gingeras. 1996. Extensive polymorphisms observed in HIV-1 clade-B protease gene using high-density oligonucleotide arrays. Nat. Med. 2:753–759.
25 Lande, R., and R. Thompson. 1990. Efficiency of marker assisted selection in the improvement of quantitative traits. Genetics 124:743–756.
26 Mackinnon, M. J., and M.A.J. Georges. 1998. Marker-assisted pre-selection of young dairy sires prior to progeny testing. Livest. Prod. Sci. 54:229–250.
27 Meuwissen, T.H.E., and M. E. Goddard. 1997. Estimation of effects of quantitative trait loci in large complex pedigrees. Genetics 146:409–416.
28 Nejati-Javaremi, A., J. P. Gibson, and C. Smith. 1994. Gain in accuracy of evaluation by including total allelic identity. Proc. 5th World Congr. Genet. Appl. Livestock Prod., Guelph, ON, Canada 19:171–174.
29 Piper, L. R., B. M. Bindon, and G. H. Davis. 1985. The single gene inheritance of the high litter size of the Booroola Merino. Pages 115–125 *in* Genetics of Reproduction in Sheep. R. B. Land and D. W. Robinson, ed. Butterworths, London, United Kingdom.
30 Plastow, G. S., M. Kuiper, R. Wales, A. L. Archibald, C. S. Haley, and K. W. Siggens. 1998. AFLP for mapping and QTL detection in commercial pigs. Proc. 6th World Congr. Genet Appl. Livest. Prod., Armidale, Australia 26:209–212.
31 Pong-Wong, R., F. Shaw, and J. A. Woolliams. 1998. Estimation of dominance variation using a finite-locus model. Proc. 6th World Congr. Genet Appl. Livest. Prod., Armidale, Australia 26: 41–44.
32 Ruane, J., and J. J. Colleau. 1995. The value of marker assisted selection when one QTL is marked. Genet. Res. 66:71–83.
33 Smith, C. 1967. Improvement of metric traits through specific genetic loci. Anim. Prod. 9:349–358.
34 Smith, C., and D. B. Smith. 1993. The need for close linkages in marker-assisted selection for economic merit in livestock. Anim. Breed. Abstr. 61:197–204.
35 Terwilliger, J. D. 1995. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. Am. J. Hum. Genet. 56:777–787.
36 Van der Beek, S., and J.A.M. Van Arendonk. 1996. Marker-assisted selection in an outbred poultry breeding nucleus. Anim. Sci. (Pentcaitland) 62:171–180.
37 Visscher, P. M. 1995. Bias in multiple genetic correlation from half-sib designs. Genet. Sel. Evol. 27:335–345.
38 Visscher, P. M., C. S. Haley, and R. Thompson, 1996. Marker assisted introgression in backcross breeding programs. Genetics 144:1923–1932.
39 Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP—a new technique for DNA-fingerprinting. Nucleic Acids Res. 23:4407–4414.
40 Webb, A. J., A. E. Carden, C. Smith, and P. Imlah. 1982. Porcine stress in pig breeding. Proc. 2nd World Congr. Genet Appl. Livest. Prod., Madrid, Spain 6:493–497.
41 Whittaker, J. C., R. N. Curnow, C. S. Haley, and R. Thompson. 1996. Using marker-maps in marker-assisted selection. Genet. Res. 66:255–265.
42 Whittaker, J. C., C. S. Haley, and R. Thompson. 1997. Optimal weighting of information in marker-assisted selection. Genet. Res. 69:137–144.
43 Wilmut, I., C. S. Haley, and J. A. Woolliams. 1992. Impact of biotechnology on animal breeding. Anim. Reprod. Sci. 28: 149–162.
44 Wilmut, I., A. Schnieke, J. McWhir, A. Kind, and K.H.S. Campbell. 1997. Viable offspring derived from fetal and adult mammalian cells. Nature (Lond.) 385:810–813.
45 Winterø, A. K., M. Fredholm, and P. D. Thomsen, 1992. Variable (dG-dT)n-(dC-dA)n sequences in the porcine genome. Genomics 12:281–288.
46 Wishart, J. 1931. The mean and second moment coefficient of the multiple correlation coefficient in samples from a normal population. Biometrika 22:353–361.
47 Zhang, W., and C. Smith, 1992. Computer simulation of marker-assisted selection utilizing linkage disequilibrium. Theor. Appl. Genet. 83:813–820.
48 Zhang, W., and C. Smith, 1993. The use of marker assisted selection with linkage disequilibrium: the effects of several additional factors. Theor. Appl. Genet 86:492–496.