Integration of molecular genetic technology with quantitative genetic technology for maximizing the speed of genetic improvement

Jack C. M. DEKKERS
(Department of Animal Science and Center for Integrated Animal Genomics
239D Kildee Hall Iowa State University Ames IA 50011 USA)

Abstract: Advances in molecular genetics have opened opportunities to enhance strategies for genetic improvement of livestock by directly selecting on genes or chromosomal regions that harbor genes that affect traits of interest so-called quantitative trait loci (QTL). Detection and use of QTL relies on presence of linkage disequilibrium between genetic markers that can be genotyped and QTL. The nature of linkage disequilibrium that exists in populations or that can be created in experimental populations and its use in QTL detection and marker-assisted selection are reviewed. Limitations and opportunities for alternate strategies for marker-assisted selection are outlined. The main conclusion is that opportunities for use of marker-assisted selection to enhance rates of genetic improvement exist in particular using population-wide disequilibrium, but they require careful analysis and implementation in existing breeding programs.

Key words: molecular genetics; quantitative genetic; QTL; MAS
CLC number: S813 **Document code**: A **Article ID**: 1001-411X (2005) S0-0104-021

To date most genetic progress for quantitative traits in livestock has been made by selection on phenotype or on estimates of breeding values (BBV) 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 genetic 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. Molecular data is of interest for use in genetic selection because genotype information has heritability equal to 1 (assuming no genotyping errors), it can be obtained in both sexes and on all animals it can be obtained early in life and it may require the recording of less phenotypic information. Objectives of this paper are to review strategies for the

Biography: Jack C. M. DEKKERS Dr, Professor Department of Animal Science Iowa State University E-mail: jlekkers@iastate.edu

detection and use of genes or markers in genetic improvement.

1 Principles of quantitative trait loci (QTL) detection

Application of molecular genetics for genetic improvement relies on the ability to genotype individuals for specific genetic loci. For these purposes, three types of observable genetic loci can be distinguished, as described by Dekkers[1]:

1) Direct markers: loci for which the functional polymorphism can be genotyped.

2) LD-markers: loci in population-wide linkage disequilibrium with the functional mutation.

3) LE-markers: loci in population-wide linkage equilibrium with the functional mutation.

The use of these loci to detect and use genes that affect quantitative traits (quantitative trait loci or QTL) relies on finding associations of genotype at marker loci with phenotype. Such associations can be detected by contrasting the mean phenotype of individuals that have alternate marker genotypes. A difference in mean phenotype indicates that the marker is linked to a QTL or, in the ideal case, is the QTL (direct marker). However, this does not mean that every marker that is linked to a QTL is expected to show a mean difference in phenotype; besides linkage, the second condition that is needed to create a difference in mean phenotype between alternate marker genotypes is presence of linkage disequilibrium (LD) between the marker and the QTL. The concept of LD is important for both QTL detection and the use of these QTL in selection (marker assisted selection or MAS), and will be explained next.

1.1 Linkage disequilibrium

Consider a marker locus with alleles M and m and a QTL with alleles Q and q that is on the same chromosome as the marker, i.e. the marker and the QTL are linked. An individual that is heterozygous for both loci would have genotype $MmQq$. Alleles at the two loci are arranged in haplotypes on the two chromosomes of a homologous pair that each individual carries. An individual with genotype $MmQq$ could have the following two haplotypes: MQ/mq, where it separates the two homologous chromosomes. Alternative, it could carry the following two haplotypes: Mq/mQ. These alternative arrangements of linked alleles on homologous chromosomes is referred to as the marker-QTL linkage phase. The arrangement of alleles in haplotypes is important because progeny inherit one of the two haplotypes that a parent carries, barring recombination.

Presence of linkage equilibrium or disequilibrium relates to the relative frequencies of alternative haplotypes in a population. In a population that is in linkage equilibrium (LE), alleles at two loci are randomly assorted into haplotypes. In other words, chromomosomes or haplotypes that carry marker allele M are no more likely to carry QTL allele Q than chromosomes that carry marker allele m. In technical terms, the frequency of the MQ haplotypes is equal to the product of the population allele frequency of M and the frequency of Q. Thus, if a marker and QTL are in linkage equilibrium, there is no value in knowing an individual's marker genotype because it provides no information on QTL genotype. If the marker and QTL are in linkage disequilibrium, however, there will be a difference in the probability of carrying Q between chromosomes that carry M and m marker alleles and, therefore, we would also expect a difference in mean phenotype between marker genotypes.

Linkage disequilibrium between markers

and QTL forms the basis for both QTL detection and marker-assisted selection. Thus, an understanding of the factors that affect the presence and extent of LD is important. The main factors that create LD in a population are mutation, selection, drift (inbreeding), and migration or crossing. The main factor that breaks down LD is recombination, which can rearrange haplotypes that exist within a parent in every generation. Fig. 1 shows the effect of recombination an the decay of LD over generations. The rate of decay depends on the rate of recombination between the loci, i.e. on their genetic distance on the chromosome; for tightly linked loci, any LD that has been created will persist over many generations but for loosely linked loci ($r=0.1$), LD will decline rapidly over generations.

Although a marker and a linked QTL may be in linkage equilibrium across the population, LD will always exist within a family, even between loosely linked loci. Consider a double heterozygous sire with haplotypes MQ/mq (Fig. 2). The genotype of this sire is identical to that of an F_1 cross between inbred lines. This sire will produce four types of gametes: non-recombinants MQ and mq and recombinants Mq and mQ. Because the non-rcoombinants will have higher frequency, depending on recombination rate, this sire will produce gametes that will be in LD, which will extend over larger distance (Fig. 1), because it has undergone only one generation of recombination. This specific type of LD, however, only exists within this family; progeny from another sire, e.g. an Mq/mQ sire, will also show LD, but the LD is in the opposite direction because of the different marker-QTL linkage phase in the sire (Fig. 2). On the other hand, MQ/mQ and Mq/mq sire families will not be in LD because the QTL does not segregate in these families. When pooled across families these four types of LD will cancel each other out, resulting in linkage equilibrium across the population. Nevertheless, the within-family LD can be used to detect QTL and for MAS, provided the differences in linkage phase are taken into account, as will be demonstrated later.

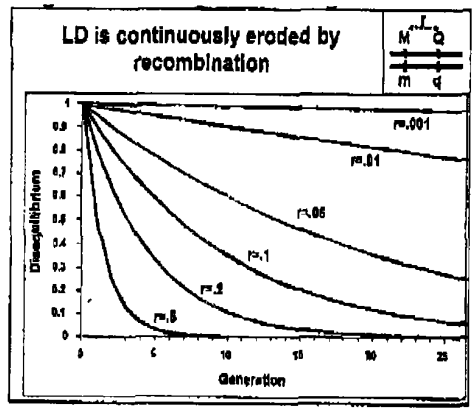

Fig. 1 Break-up of LD over generations

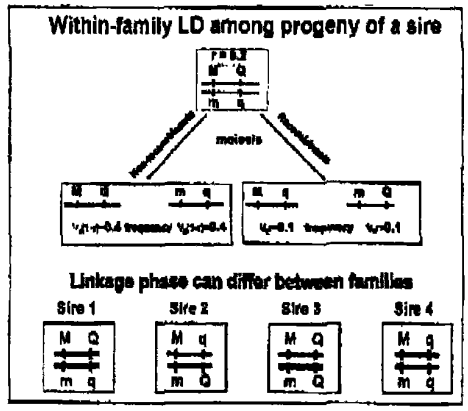

Fig. 2 Within family marker-QTL LD

1.2 Use of LD to detect QTL

Methods to detect QTL using genetic markers rely on identifying markers that are associated/correlated with phenotype. This will only occur for markers that are in LD with a QTL and therefore depends on the type and extent of LD that exist in the population under analysis. Classification of a marker as an LD- versus LE marker depends on the population structure and the extent of LD that is present in that population. For breed crosses, LD is extensive and markers that are 10 – 20 cM from the QTL will still be in extensive population-wide LD. However, within closed outbred populations, the extent of LD will

be limited and LD markers must by necessity be close to the functional mutation (within 1 to 5 cM depending on population history). In closed breeding populations, identification of LD markers, therefore, is more difficult and requires candidate gene approaches or fine-mapping approaches[2] with high-density marker maps. Functional mutations are most difficult to detect and few examples are available[2]. The principles behind QTL detection in each of these scenarios will be described further in the following.

1.2.1 Population-wide LD in crossbred populations Crossing two breeds that differ in gene and, therefore, haplotype frequencies creates extensive LD in the crossbred population that extends over larger distances because the LD present in the F_1 generation has undergone only one generation of recombination in the F_2 (Fig. 1). This enables detection of QTL that differ between the two breeds based on a genome scan with a limited number of markers spread over the genome (every 15 to 20 cM) and has formed the basis for the extensive use of F_2 or backcrosses between breeds or lines for QTL detection (e.g. Malek et al.[3,4]). This extensive LD enables detection of QTL that are some distance from the markers but also limits the accuracy with which the position of the QTL can be determined.

More extensive population-wide LD is also expected to exist in synthetic lines, i.e. lines that were created from a cross in recent history. Depending on the number of generations since the cross, the extent of LD will have eroded some over generations and will, therefore, span shorter distances than in F_2 populations (Fig. 1). This will require a more dense marker map to scan the genome but will enable more precise positioning of the QTL.

1.2.2 Within-family LD in outbred populations Because linkage phases between the marker and QTL can differ from family to family, use of within-family LD for QTL detection requires marker effects to be fitted on a within family basis, rather than across the population. Similar to F_2 or backcrosses, however, the extent of within-family LD is extensive and, thus, genome-wide coverage is provided by a limited number of markers but significant markers may be some distance from the QTL. Thus, LE markers can be readily detected on a genome-wide basis using breed crosses or analysis of large half-sib families, requiring only sparse marker maps (20 cM spacing). Many examples of successful applications of this methodology for detection of QTL regions are available in the literature.

1.2.3 Population-wide LD in outbred populations The amount and extent of LD that exists in the populations that are used for genetic improvement is the net result of all the forces that create and break-down LD and is, therefore, the result of the breeding and selection history of each population, along with random sampling. On this basis, populations that have been closed for many generations are expected to be in linkage equilibrium, except for closely linked loci. Thus, in those populations, only markers that are tightly linked to QTL may show an association with phenotype (Fig. 1), and even then there is no guarantee because of the chance effects of random sampling.

There are two strategies to find markers that are in population-wide LD with QTL:

1) evaluating markers that are in or close to genes that are thought to be associated with the trait of interest (candidate genes).

2) a genome scan using a high-density marker map, with a marker every 0.5 to 2.0 cM.

The success of both these approaches obviously depends on the extent of LD in the population. Studies in human populations rarely gener

ally found that LD extends over less than 1 cM. Thus, many markers are needed to get sufficient marker coverage in human populations to enable detection of QTL based on population-wide LD. Opportunities to utilize population-wide LD to detect QTL in livestock populations may be considerably greater because of the effects of selection and inbreeding. Indeed, Farnir et al[5] identified substantial LD in the Dutch Holstein population, which extended over 5 cM. Similar results have been observed in other livestock species (e.g. in poultry, Heifetz et al[6]). The presence of extensive LD in livestock populations is advantageous for QTL detection, but disadvantageous for identifying the causative mutations of these QTL; with extensive LD, markers that are some distance from the causative mutation can show an association with phenotype.

The candidate gene approach utilizes knowledge from species that are rich in genome information (e.g., human, mouse), effects of mutations in other species, previously identified QTL regions, and/or knowledge of the physiological basis of traits to identify genes that are thought to play a role in the physiology of the trait. Following mapping and identification of polymorphisms within the gene in the pig, the association of genotype at the candidate gene with phenotype can be estimated in a closed swine breeding population (Rothschild et al[7]).

Recent advances in genome technology has enabled sequencing of entire genomes, including of several livestock species; the genomes of the chicken and cattle have been sequenced and public sequencing of the genome of the pig is underway. In addition, sequencing has been used to identify large numbers of positions in the genome that include single nucleotide polymorphisms (SNPs), ie. DNA base positions that show variation. For example, in the chicken, over 2.8 million SNPs have been identified by comparing the sequence of the Red Jungle Fowl to that of three domesticated breeds (International Chicken Polymorphism Map Consortium, 2004). This, combined with reducing costs of genotyping, now enables detection of QTL using LD-mapping with high-density marker maps.

2 Marker-assisted selection(MAS)

Principles and limitations of MAS in both livestock and plant breeding were described by Dekkers et al[8]. The main issues related to livestock will be described in the following.

2.1 Traits with MAS application

Molecular markers have been used to identify loci or chromosomal regions that affect single gene traits and quantitative traits. Single gene traits include genetic defects or disorders and appearance. For the purposes of QTL detection and application, quantitative traits can be categorized into: a) routinely recorded traits, which be further subdivided into traits that are recorded on both sexes, sex-limited traits, and traits that are recorded late in life; b) difficult to record traits (feed intake, product quality); and c) unrecorded traits (disease resistance). The ability to detect QTL decreases in the order a), b), c) because of the availability of phenotypic data. For a similar reason, genome scans, which require more phenotypic data, are often used to detect QTL for traits in category a), whereas candidate gene approaches are more often used to identify QTL for traits that are not routinely recorded (b and c). Potential extra genetic gains from marker-assisted selection are the greatest for traits in category c) and the lowest for traits in category a), in particular for traits that are routinely recorded on both sexes prior to selection, in inverse proportion to the ability to make genetic progress using conventional methods[9].

2.2 General strategies for selection using molecular genetic information

Once markers that are linked to QTL have been identified, their effects can be estimated based on the association between phenotype and genotype. The resulting estimates can be used to assign a 'molecular score' to each selection candidate, which can be used to predict the genetic value of the individual and used for selection. The constitution and method of quantification of the molecular score depends on type of LD that is used and on how the marker will be used in selection.

The three types of molecular loci described previously differ not only in methods of detection but also in methods of application in selection programs, whereas direct and, to a lesser degree, LD-markers, allow selection on genotype across the population, use of LE-markers must allow for different linkage phases between markers and QTL from family to family. Thus, the ease and ability to utilize markers in selection is opposite to their case of detection and increases from functional mutations to LD markers and LE markers. In what follows, selection on each of these three types of markers will be referred to as gene assisted selection (GAS), LD markers assisted selection (LD-MAS), and LE marker assisted selection (LE-MAS).

In addition to a molecular score, individuals can also obtain a regular estimate of the breeding value for the collective effect of all the other genes (polygenes) on the trait. A summary of approaches and strategies for the use of molecular genetic information in genetic improvement is given in Tab. 1. Details of the various uses are provided in the remainder of this paper.

Tab. 1 Strategies for the use of molecular data in genetic improvement programs

type	selection Program	marker requirements	information required to compute molecular score	composition of molecular score	selection criterion
between breed[1)]	introgression · foreground	< 10 cM from QTL	line origin of markers alleles at target loci	presence /absence target alleles	mol score
	· background	~ 20 cM intervals genome wide	line origin of marker alleles	% recipient breed alleles[2)]	mol score (recipient trait phen.)
	· intercross	< 5 ~ 10 cM from QTL	line origin of marker alleles at target loci	# taget alleles	mol score
	synthetic line development	< 10 cM from QTL	line origin of QTL / marker alleles Estimates of QTL /marker effects	sum. of QTL /marker estimates	mol score phen. EBV
within breed selection Ⅱ	LD-MAS /GAS direct markers cand. genes LD-markers	< 1 ~ 2 cM from QTL	genotypes at QTL /markers estimates of QTL /marker effects	sum. of QTL /marker esrimates	mol score phen. EBV
	LE-MAS	< 10 cM from QTL	parental origin marker alleles within-family estimates of QTL /marker effects	sum. of QTL /marker estimates	mol score phen. EBV

1) All LD-MAS, capitalizing on extentive LD in cross;

2) Greater emphasis on markers linked to target loci to reduce linkage drag.

In general, at the time of selection, both molecular and phenotypic information is available for use in selection. The following three selection strategies can then be distinguished:

1) Select on the molecular score alone.

2) Tandem selection, with selection on molecular score, followed by selection on phenotype based EBV.

3) Selection on an index of the molecular score and the regular EBV.

Selection on molecular score alone ignores information that is available on all the other genes (polygenes) that affect the trait and is expected to result in the lowest response to selection, unless all genes that affect the trait are included in the molecular score. This strategy does, however, not require additional phenotypes, other than those that are needed to estimate marker effects, and can be attractive when phenotype is difficult or expensive to record (e.g. disease traits, meat quality, etc.).

If both phenotypic and molecular information is available on selection candidates, index selection is expected to result in greater response to selection than tandem selection. The reason is similar to why two-trait selection using independent culling levels is expected to give lower multiple-trait response than index selection; two-stage selection does not select individuals for which a low molecular score may be compensated by a high phenotype-based EBV. The choice between these strategies (and other alternatives) also depends on other factors, such as market and cost considerations.

2.3 Marker-assisted improvement using between-breed variation

Molecular information can enhance the process of integrating superior qualities of different breeds in a number of ways. All of these rely on crosses which, as described previously, results in extensive LD, which can be capitalized on using MAS. If a large proportion of the breed difference in the trait of interest is due to a small number of genes, introgression strategies can be used. If a larger number of genes is involved, marker-assisted selection within a synthetic line is the preferred method of improvement. These strategies will be further described below.

2.3.1 Marker-assisted introgression Within the context of meat quality, the aim of an introgression program is to introduce one or more meat quality genes (target genes) from a breed that is superior for meat quality but inferior for performance (the donor breed) into a high performance line that lacks the target genes (the recipient breed). This is done through an initial F_1 cross followed by multiple backcrosses to the recipient breed and one or more generations of intercrossing. The aim of the backcross generations is to generate individuals that carry one copy of the donor QTL allele but that are similar to the recipient breed for the rest of the genome. This is accomplished by successive backcrosses to the recipient breed to 'dilute' the donor genome, while maintaining the donor allele at the QTL by selecting only carriers as parents of the next generation. The aim of the intercrossing phase is to fix the donor allele at the QTL. The end result is a population that is similar to the recipient breed, except for retrying two copies of the donor allele at the QTL.

The effectiveness of introgression schemes is limited by the ability to identify backcross or intercross individuals that carry the target gene(s) and by the ability to identify backcross individuals that have a high proportion of the recipient genome, in particular in the region(s) around the target gene(s). The latter affects the number of backcross generations required to recover the recipient genome. Molecular genetics can enhance the effectiveness of both phases of an introgression program. Effectiveness of the

backcrossing phase can be increased in two ways: i) by identifying carriers of the target gene(s) (foreground selection), and ii) by enhancing recovery of the donor genetic background (background selection). Effectiveness of the intercrossing phase can also be enhanced through foreground selection on the target gene(s). Note that the use of markers in either fore- or background selection does not require estimation of the marker or QTL effects. Instead, their use relies on breed differences and the association of marker alleles with these breed differences as a result of extensive LD.

For foreground and intercross selection, selection is on a molecular score that is based on presence or absence of the target allele (only individuals that carry the allele are selected) (Tab. 1). If the target gene cannot be directly genotyped, carrier individuals can be identified based on markers that flank the QTL at <10 cM, because of the extensive LD in crosses. The markers must have breed-specific alleles, such that line origin can be identified. The effectiveness of foreground selection depends on the number of target genes and on the confidence interval for the position of those genes. The latter determines the size of the genomic region that must be introgressed. Both factors have a large impact on the number of individuals that is required to find individuals that are carriers for all target genes during the backcrossing phase and homozygous during the intercrossing phase. For the introgression of multiple target genes, gene pyramiding strategies can be used during the backcrossing phase to reduce the number of individuals required[10,11].

For background selection, markers are used that are spread over the genome at <20 cM intervals, such that most genes that affect the trait will be within 10 cM from a marker. Again, markers must have breed-specific alleles to allow the tracing of alleles back to their breed origin. Marker genotypes are then used to estimate the proportion of the recipient genome present in an individual, which is used as the molecular score (Tab. 1). Individuals with the highest proportion are selected. Combining foreground and background selection, selection will be for the donor breed segment around the target locus but for recipient breed segments in the rest of the genome. Foreground selection will result in selection for not only the target locus but also for donor breed loci that are linked to this locus, some of which could have an unfavorable effect on performance. To reduce this so-called linkage drag around the target locus, greater emphasis can be given in the molecular score used for background selection to markers that are in the neighborhood of the target locus (apart from the flanking markers which are used in foreground selection).

Most studies have considered marker-assisted introgression (MAI) of single QTL[10] but often several QTL must be introgressed simultaneously. Several studies (e.g. Koudande et al[11]) have shown that large population sizes are needed to obtain sufficient individuals that are heterozygous for all QTL in the backcrossing phase. This would make MAI not feasible in livestock breeding programs. In many cases, however, immediate fixation of introgressed QTL alleles may not be required. Instead, the objective of the backcrossing phase can be to enrich the recipient breed with the favorable donor QTL alleles at high enough frequency such that they can be selected in following backcrossing. In this case, individuals can be selected during the backcross phase based on a molecular score computed as the expected number of donor alleles at the n introgressed QTL, as determined from marker genotypes: $MS = \sum_{i=1}^{n} P(Q_i)$, where $P(Q_i)$ is the probability that the individual car-

ries the donor allele for QTL *I*. Probabilities P(Q_i) *can be set equal to* 1, 0.5, and 0 if the individual carries 2, 1, and 0 donor alleles at the two markers that flank the QTL, ignoring double recombinants. Selection can be on a similar score during the intercrossing phase. The effectiveness of these strategies was evaluated by Chaiwong et al.[12] These studies showed that although it may not be possible to maintain a frequency of 50% during backcrossing in populations of limited size, MAI can introduce multiple QTL alleles at frequencies that will enable their selection following backcrossing.

2.3.2 ***Marker assisted synthetic line development*** Lande et al.[13] proposed a strategy for marker-assisted selection within a hybrid population created by crossing two inbred lines. The strategy capitalizes on population-wide LD that initially exists in crosses between lines or breeds. Thus, marker-QTL associations identified in the F_2 generation can be selected on for several generations, until the QTL or markers are fixed or the disequilibrium disappears. Zhang et al.[14] evaluated the use of markers in such a situation with selection on BLUP EBV. They compared the following three selection strategies:

MAS: selection on a molecular score derived from marker effects.

BLUP: selection on BLUP EBV derived from phenotype.

COMB: combined selection on an index of the EBV based on markers and phenotype.

Data for a cross between inbred lines were simulated on the basis on 100 QTL and 100 markers in a genome of 2 000 cM. Marker effects were estimated in the F_2 generation using a two-step procedure. In the first step, a separate F_2 population from the same cross was used to identify markers with the largest effects. Then, to obtain unbiased estimates, the effects of those markers were re-estimated in the F_2 population under selection. The latter were used to obtain marker-based EBV, which were used as the molecular score throughout the selection process. Zhang et al.[14] found that index selection (COMB) resulted in the greatest response, followed by selection on BLUP EBV and selection on markers alone. Rates of response declined over generations for all strategies because data were simulated using a finite number of loci which were moved to fixation by selection. Rates of response declined faster for the MAS strategy because recombination eroded the disequilibrium between the markers and QTL. Nevertheless, substantial rates of response were obtained using selection on markers alone. The latter has potential for selection for meat quality traits because it does not require continuous phenotypic evaluation of meat quality traits, in contrast to the BLUP and COMB strategies. Zhang et al.[14] considered the ideal situation of a cross with inbred lines. Although the lines were not divergent for the trait of interest, they were homozygous at alternate alleles for all loci. Breeds used in a cross to enhance meat quality will typically have different means, which will increase the extent of linkage disequilibrium in the cross. However, both breeds will likely segregate for most QTL, which will reduce the disequilibrium. Nevertheless, even in crosses between commercial breeds of swine, substantial numbers of QTL have been found for which the breeds have sufficient differences in frequency to allow their detection[3,4,15]. In addition, favorable effects have been found to originate from the breed with the lower mean for a number of QTL[4].

2.4 Genetic improvement using within-breed variation

Most selection programs for swine focus on genetic improvement within a breed or line and the subsequent use of that line within a cross

breeding strategy. Within-breed selection requires information that captures differences between individuals within a breed, rather than the between-breed differences that were discussed previously. The purpose of this section is to describe opportunities for genetic improvement of meat quality based on within-breed selection programs, starting with conventional selection. This will be divided into selection using LE- and LD-markers.

2.4.1 Selection using LE-MAS based on within family LD Use of within-family LD between a QTL and a linked marker requires marker effects or, at a minimum, marker-QTL linkage phases to be determined separately for each family. This requires marker genotypes and phenotypes on family members. If linkage between the marker and QTL is loose, phenotypic records must be from close relatives of the selection candidate because associations will erode through recombination. With progeny data, marker-QTL effects or linkage phases can be determined based on simple statistical tests that contrast the mean phenotype of progeny that inherited alternate marker alleles from the common parent. Alternatively, marker-assisted animal models have been developed to incorporate marker data in genetic evaluation for complex pedigrees[16]. These models result in BLUP EBV of QTL effects along with polygenic BBV. Because selection is on performance traits along with meat quality, these estimates should be combined with EBV for performance traits into an economic index.

Implementation of LE-MAS requires extensive phenotyping and genotyping, which calls the economic feasibility of such programs into question. In addition, data should be available for several generations prior to initiating MAS to accurately estimate QTL effects. Another obstacle for the use of within-family LD is that it requires knowledge of QTL regions that segregate within the population. Since most QTL mapping studies in pigs are based on the breed cross model, information about within-breed segregation of QTL is limited. Thus, within-breed QTL mapping studies must be conducted prior to implementation of MAS. Although such studies could concentrate on QTL regions previously identified in breed cross studies, substantial population sizes will be required to detect or confirm their segregation within a breed. Such a study was recently conducted by Evans et al[17]. They found that QTL regions identified in a cross between divergent breeds could indeed be confirmed to segregate within commercial lines. Related issues were discussed by Spelman et al[18]. in the context of implementing QTL knowledge in dairy cattle breeding programs.

2.4.2 Selection using LD-MAS based on population-wide LD Although markers that are not within the functional gene are not expected to be in extensive LD with a QTL within a closed population, markers that are tightly linked to a QTL have a substantial probability to be in partial population-wide LD with that QTL because of the effects of drift, selection, mutation, and population admixture[19-21]. This probability is higher in selected populations of small effective size, which is the case for livestock, as demonstrated by Farnir et al[5]. for dairy cattle. The extent of LD can often be enhanced through the use of haplotypes of tightly linked markers. High-density marker maps with, e. g., a marker every 1 or 2 cM, will also include markers that are in tight linkage with the QTL and that have the potential to be in substantial population-wide LD, as was recently demonstrated by Meuwissen et al[21]. through simulation. They showed that for populations with an effective population size of 100 and a 1 or 2 cM spacing between markers across the genome

sufficient disequilibrium was present that genetic values could be predicted with substantial accuracy for several generations on the basis associations of marker haplotypes with phenotype on as few as 500 individuals. New high-throughput technologies now enable to conduct this for reasonable costs. In addition, opportunities may exist to utilize this approach on a limited scale by saturating previously identified QTL regions with markers.

For markers that are in population-wide LD with the QTL, selection can be directly on marker genotype or on marker haplotype if multiple linked markers are used to track the QTL. It is, however, essential to estimate the effects of the markers within the population under selection to capture the degree of LD and linkage phases that are present in the population and to guard against potential interactions of the QTL with the background genome. For the same reason, it will also be prudent to re-estimate the effects on a regular basis. Estimation requires marker genotypes and meat quality phenotypes on a random sample of individuals in the population and should be based on an animal model with marker genotypes or haplotypes included as fixed effects[22,23].

2.4.3 LD-MAS *versus* LE-MAS An important consideration for the use of molecular genetics in breeding programs is whether to work toward the application of LE-MAS, LD-MAS, or GAS. Requirements for detection are least for LE markers and greatest for identification of functional mutations. However, once a functional mutation is identified, requirements for estimation and confirmation of effects in other populations are much lower than for LE-markers because the latter requires phenotypes and genotypes on pedigreed populations versus a random sample. Requirements for integration of genotype data in routine genetic evaluation procedures are also much greater for LE-MAS than for LD-MAS and GAS, both with regard to requirements of individuals that must be phenotyped and genotyped and with regard to methods of analysis. Genetic evaluation requirements are slightly greater for LD-MAS than GAS because LD-MAS requires identification and analysis of marker haplotypes and confirmation of marker-QTL linkage phases.

Whereas the previous refer to requirements for a given QTL, LE-MAS allows for genome-wide analysis and evaluation of QTL with a limited number of markers. This is also possible for LD-MAS with high-density genotyping. Meuwissen et al.[21] demonstrated that MV of high accuracy could be obtained from a Bayesian mixed model analysis of marker haplotypes with high-density genotyping.

Opportunities for increases in genetic gain from a given QTL are lowest from LE-MAS because of the limited information that is available to estimate effects on a within-family basis, while for both LD-MAS and GAS, effects are estimated from data across families. Accuracy of estimates may be slightly lower for LD-MAS than GAS as a result of incomplete marker-QTL disequilibrium and a greater number of effects (marker haplotypes versus QTL genotypes)[24]. Opportunities for intellectual property protection and product differentiation are greatest for GAS but limited for LE-MAS.

3 Integration of MAS in breeding programs

It is clear that successful implementation of a MAS program requires a comprehensive integrated approach that is closely aligned with business goals and markets. Components of such an approach include DNA collection, phenotyping, pedigree recording, genotyping, establishment of genotypic and phenotypic data bases, and ana-

lytical tools and quality control systems. Application of MAS also requires careful consideration of economic aspects and business risks. Economic analysis of MAS requires a comprehensive approach that aims to evaluate the economic feasibility and optimal implementation of MAS. An excellent example of such an analysis is in Hayes et al[25], who conducted a comprehensive economic analysis of the implementation of LE-MAS in the nucleus breeding program of an integrated pig production enterprise. QTL detection and MAS on identified QTL regions for a multi-trait breeding goal and associated genotyping costs and extra returns from the production phase of the integrated enterprise were considered in the economic assessment. They concluded that implementation of LE-MAS was feasible for the assumed cost and price parameters. They also found that, in particular if QTL detection was based on small sample sizes, stringent thresholds should be set during the QTL detection phase such that genotyping costs during the implementation phase are reduced and selection of false positives is minimized.

Whereas Hayes et al.[25] evaluated economic returns from MAS from increased profit at the production level, which is proportional to extra genetic gain, most commercial breeding programs derive profit from increased market share of breeding stock or germ plum. In general, implementation of MAS will have a greater impact on market share than on genetic gain. Success of MAS also depends on the consistency of QTL effects across populations and environments. Results from introgression programs in plants have found that effects tend to be consistent for major genes for simple traits but not for QTL for complex traits (e.g. yield)[26]. Inconsistent effects have also been observed for some well-studied genes in livestock. For example, for the ESR gene for littersize in pigs some studies have found no effect and interactions with line and environment have been identified[7]. Reasons for inconsistent results across studies and populations include statistical anomalies such as false positive or negative results (small sample sizes) and overestimation of significant QTL effects, as well as true effects such as inconsistent marker-QTL linkage phases across populations for LD-markers, genotype by environment interactions and epistatic effects. This points to the need to continuously evaluate and monitor gene or QTL effects in the target population and environment, which requires continuous emphasis on phenotypic recording. In addition, strategies must be developed to estimate gene effects at the commercial level for nucleus breeding programs, in particular if they involve crossbreeding. This also opens opportunities to use markers to capitalize on non-additive effects and assignment of specific matings.

Given the uncertainties about the sustainability of marker effects, it appears prudent to use molecular genetic information in a manner that does not prevent progress toward the overall breeding goal that can be achieved through conventional selection. A crucial concept in this regard is to apply MAS in selection space that is not or under-utilized by conventional selection[27]. A prime example is pre-selection on the basis of markers among members of a full-sib family for further testing, prior to availability of individual or progeny records. In such situations conventional selection has no basis for selection because EBV are derived from pedigree information, which is the same for all members of a full-sib family. Family members can, however, differ for the markers they inherited, which then provides a basis for selection, instead of having to make a random choice.

Recent gene and QTL mapping studies have also revealed that QTL may not be expressed in a

Mendelian fashion. In particular, several studies have detected genes and QTL in pigs that are subject to gametic imprinting[28,29]. Future studies will undoubtedly identify other epigenic phenomena that affect the inheritance and expression of QTL. These effects will need to be taken into account when designing selection programs. Although they may on the one hand complicate selection programs, they may also provide opportunities. For example, De Koning[30] suggested that utilization of a combination of imprinted and sex-linked QTL would allow a diverse set of markets to be targeted through strategic crosses between a single set of breeds.

References:

[1] DEKKERS J C M. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons[J]. J Anim Sci, 2004, 82: 313-328.

[2] ANDERSSON L. Genetic dissection of phenotypic diversity in farm animals[J]. Nature Reviews Genetics, 2001, 2: 130-138.

[3] MALEK M, DEKKERS J C M, LEE H K, et al. A molecular genome scan analysis to identify chromosomal regions influencing economic traits in the pig. I. Growth and body composition[J]. Mamm Genome, 2001, 12: 630-636.

[4] MALEK M, DEKKERS J C M, LEE H K, et al. A molecular genome scan analysis to identify chromosomal regions influencing economic traits in the pig. II. Meat and muscle composition[J]. Mamm Genome, 2001, 12: 637-645.

[5] FARNIR F, COPPIEWS W, ARRANT J J, et al. Extensive genome-wide linkage disequilibrium in cattle[J]. Genome Res, 2000, 10: 220-227.

[6] HELFETZ E M, FULTTON J E, SULLIVAN N, et al. Extent and consistency across generations of linkage disequilibrium in commercial layer chicken breeding populations[J]. Genetics, 2005.

[7] ROTHSCHILD M F, PLASTOW G S. Andvances in pig genomics and industry applications[J]. AgBiotechNet, 1999, 1: 1-7.

[8] DEKKERS J C M, HOSPITAL F. The use of molecular genetics in improvement of agricultural populations[J]. Nature Reviews Genetics, 2001, 3: 22-32.

[9] MEUWISSAN T H E, GODDARD M E. The use of marker haplotypes in animal breeding schemes[J]. Genet, 1996, 28: 161-176.

[10] HOSPITAL F, CHARCOSSET A. Marker-assisted introgression of quantitative trait loci[J]. Genetics, 1997, 147: 1 469-1 485.

[11] KOUDAND O D, IRAQI F, THOMSON P C, et al. Strategies to optimize marker-assisted introgression of multiple unlinked QTL[J]. Mammalian Genome, 2000, 11: 145-150.

[12] CHAIWONG N, DEKKERS J C M, FEMANDO R L, et al. Introgressing multiple QTL in backcross breeding programs of limited size[A]. Proc 7th W Wld Congr Genet Appl Livest Prod: Electronic Communication 22-08[C]. 2002.

[13] LANDE R, THOMPSON, R. Efficiency of marker-assisted selection in the improvement of quantitative traits[J]. Genetics, 1990, 124: 743-756.

[14] ZHANG W, SMITH C. Computer simulation of markers-assisted selection utilizing linkage disequilibrium[J]. Theor Appl Genet, 1992, 83: 813-820.

[15] GRINDFLEK E, SZYDA J, LIU Z, et al. Detection of quantitative trait loci for meat quality in a commercial slaughter pig cross[J]. Mamm Genome, 2001, 12: 299-304.

[16] FERNANDO R L, GROSSMAN M. Marker-assisted selection using best linear unbiased prediction[J]. Genet Sol, 1989, 21: 467-477.

[17] EVANS G, GIUFFRA E, SANEHEZ A, et al. PIGQTECH-Transferring QTL technology to the pig breeding industry[A]. Proc 7th W Wld Congr Genet Appl Livest Prod: Electronic Communication 3-02[C]. 2002.

[18] SPELMAN R J, BOVENHUIS H. Moving from QTL experimental results to the utilisation of QTL in breeding programmes[J]. Anim Genet, 1998, 29: 77-84.

[19] SVED J A. Linkage disequilibrium and homozy-

gosity of chromosome segments in finite populations[J]. Theor Point Biol, 1971, 2: 125 – 141.

[20] GODDARD M E. Mapping genca for quantitative traits using linkage disequilibrium[J]. Genet Sel, 1991, 23 (Suppl 1): 131 – 134.

[21] MEUWISSEN T H R, HAYES B, GODDARD M E. Prediction of total genetic value using genome-wide dense marker maps[J]. Genetics, 2001, 157: 1 819 – 1 829.

[22] SHORT T H R, SOUTHWOOD M F, McLAREN D O, et al. Effect of the estrogen receptor locus on reproduction and production traits in four commercial pig lines[J]. J Anim Sci, 1997, 75: 3 138 – 3 142.

[23] ISRAEL C, WELLER J I. Estimation of candidate owe effects in dairy cattle populations[J]. J Dairy Sci, 1998, 81: 1 653 – 1 662.

[24] HAYES B, BOWMAN P J, GODDARD M E. Linkage diseqtdlibrium and accuracy of predicting breeding values from marker haplotypes[J]. Proc Assoc Advmt Anim Breed Genet, 2001, 14: 269 – 272.

[25] HAYES B, GODDARE M E. Evaluation of marker assisted selection in pig enterprises[J]. Livest Prod Sci, 2003, 81: 197 – 211.

[26] HOSPITAL F A B I, CAUSSE L M, CHARCOSSET A. Use of markers in plant breeding: lessons from genotype building experiments[A]. Proc 7th Wld Congr Genet Appl Liveat Prod, Electronic Communication 22 – 05 [C]. 2002.

[27] SOLLER M, MEDJUGOME I. A successful marriage: making the transition from quantitative trait locus mapping to marker assisted selection[A]. DEKKERS, C M, LAMONT S J, ROTHSCILD M F. From Jay Lush to genomics: visions for animal breeding and genetics. Dept Animal Science[C]. Ames: Iowa State University, 1999.

[28] JEON J, CARLBORG O, TORNSTEN A, et al. A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the IGF2 locus[J]. Nat Genet, 1999, 21: 157 – 158.

[29] De KONING D, RATTINK A P, van ARENDONK H B, et al. Genome-wide scan for body composition in pigs reveals important role of imprinting[J]. Proc Nail Acad Sci, 2000, 97: 7 947 – 7 950.

[30] De KONING D J. Identification of (non-)Mendelian factors affecting pork production[D]. The Netherlands: Wageningen University, 2001.

(Edited by ZHOU Zhi-hong)

译文

分子遗传技术与数量遗传技术结合使遗传改良速度最大化

Jack C. M. DEKKERS

(Department of Animal Science and Center for Integrated Animal Genomics, 239D Kildee Hall, Iowa State University, Ames, IA, 50011, USA)

摘要: 随着分子遗传学的发展, 为家畜的遗传改良提供了新的手段, 可以直接选择影响性状的基因或包含基因的染色体区域, 即所谓数量性状基因座(QTL)。QTL的检测和应用依赖于可以分型的遗传标记和 QTL之间是否存在连锁不平衡. 本文对畜群中存在的或可以在试验畜群中产生的连锁不平衡的特性、及其在 QTL检测和标记辅助选择中的作用进行了综述, 并且概述了标记辅助选择的不同策略的优缺点. S主要的观点是利用标记辅助选择加快遗传改良的速度是可行的, 特别是可以利用群体范围的连锁不平衡时更是如此, 但是这需要认真地分析并在现有育种方案中落实。