Quantitative Genetics, Molecular Markers, and Plant Improvement

Kendall R. Lamkey[a,b] and Michael Lee[b]

[a]Field Crops Research Unit, USDA-Agricultural Research Service
[b]Department of Agronomy, Iowa State University, Ames, IA 50011

Abstract

Quantitative genetics in conjunction with statistics has provided much of the scientific framework for modern plant breeding. Although there has been no specific review of the contributions of quantitative genetics and statistics to plant breeding, the contributions have been undoubtedly profound and lasting. Quantitative genetic theory in many ways is robust to and naive of modern genetic principles. Little is known about the biology or the genetic architecture of quantitative traits. In this paper, five major areas of quantitative genetics -- number of loci controlling quantitative traits, nature of quantitative trait loci, gene action and effects, epistasis, and genotype x environment interaction -- relevant to plant improvement and to molecular marker applications to such improvement are reviewed. Beyond generalities, the conclusion is that quantitative genetics has provided little specific information on the biology or the architecture of quantitative traits. Molecular markers may complement plant breeding in three broad areas. Molecular markers provide reliable estimates of genetic diversity, may improve screening efficiency for many traits through their linkage with alleles with small (quantitative traits) and with large (qualitative traits) effects, and will provide the first understanding of biology and architecture of quantitative traits at the DNA level. Generalities about the usefulness of molecular markers in plant improvement are difficult to make.

Introduction

Quantitative genetics in conjunction with statistics and Mendelian genetics has provided the scientific framework for much of modern plant breeding. The primary contribution of quantitative genetics to plant improvement has been the prediction of response to artificial selection (Barton, 1990), which has allowed the comparison of alternative breeding methods and the development of new methods. Although some major revolutions in plant improvement, such as Shull's (1908, 1909) proposal for the inbred-hybrid concept in maize (*Zea mays* L.), came about before the advent of quantitative genetic theory and much of modern statistical theory, there has been consistent and concerted effort to describe plant-breeding methodology and to predict selection responses by using quantitative genetic and statistical models with only a limited understanding of the biology and the architecture of quantitative traits. This should come as no surprise because understanding of the nature of genes and their regulation has come about only recently and is still incomplete.

Plant breeders have been criticized for being slow to adopt modern molecular techniques into their breeding programs and for being satisfied with the status quo (Helentjaris, 1992). It also has been argued that plant breeding is more art than science and that there is no usable scientific framework for plant breeding (Helentjaris, 1992). These undeserved accusations reflect a lack of

Citation: Lamkey, K. R. and M. Lee. 1993. Quantitative genetics, molecular markers, and plant improvement. p. 104-115. *In*: B. C. Imrie and J. B. Hacker (ed.) Focused plant improvement: Towards responsible and sustainable agriculture. Proc. 10th Australian Plant Breeding Conf., Gold Coast, 18-23 April 1993. Organising committee, Australian Convention and Travel Service: Canberra

understanding of plant breeding, a failure to recognize the enormous contributions of quantitative genetics and statistics to plant improvement (Hallauer and Miranda, 1988; Falconer, 1981), and a failure to recognize the heavy use that plant breeders make of supporting sciences such as agronomy, soil science, plant pathology, entomology, physiology, computer science, etc. Plant breeders also have been plagued with unkept promises of revolutions from allied sciences (Simmonds, 1991) and have justifiably grown weary of newer and bigger promises. Many scientists fail to recognize that plant breeders are heavy users of technology and will adopt rapidly any new technology that can be proved to augment and to improve the efficiency and the cost effectiveness of their breeding programs.

Although plant breeders have been slow to adopt molecular marker techniques, we believe that breeders have not rejected this technology. There have been several reports of the potential applications of molecular markers to plant improvement (Burr et al., 1983; Helentjaris et al., 1985; Beckman and Soller, 1986), but few reports have analyzed breeding programs in detail to identify applications having the greatest potential. Plant breeding in its simplest form consists of (a) generation of genetic variation through recombination and (b) selection to identify superior recombinants for advancement in the breeding program. All breeding methods share these two steps, but the protocols used within each step are complicated and differ with breeding method, breeding objective, and species (Hallauer, 1990; Fehr, 1987).

Understanding the application of molecular markers to plant breeding requires a detailed understanding of plant-breeding methodology (resource allocation [economics], progeny testing, reproductive biology [cycle time]), quantitative genetics, and statistics. Knowledge of breeding methodology is required to understand how current cultivars and hybrids are developed. Knowledge of quantitative genetics and statistics is required to understand the theoretical basis of plant-breeding methodology; the design, implementation, and interpretation of plant breeding experiments; and the statistical power required to detect differences among cultivars and hybrids. For the program to be successful, economics of cultivar development must be considered simultaneously with design of the breeding program. These areas represent a vast knowledge base that has expanded over the past 90 years. Many molecular marker experiments conducted today could not have been done without this knowledge base.

Despite this knowledge base regarding plant and animal improvement, important questions remain intractable or inadequately answered. Plant-breeding methodology is still a long-term process requiring 5 to 10 years to place a new cultivar or hybrid on the market. Improvements still are needed, and molecular markers offer one avenue of achieving them. Our objectives are (1) to review some intractable or inadequately answered questions remaining after four decades of quantitative genetic studies, and (2) to assess how molecular markers can have an effect on these questions and ultimately on plant improvement.

Quantitative genetics and plant improvement

Current status

Quantitative genetics had its origin in the controversy, occurring first during the early part of

this century, regarding whether the continuous variation observed for metric traits could be reconciled with the discrete processes of Mendelian genes and of inheritance laws (Kempthorne, 1977). The foundations of quantitative genetics often are attributed to Fisher (1918), Haldane (summarized 1932), and Wright (1921) (Falconer, 1981). Since 1920, there has been a great expansion of research in quantitative genetics, and numerous books (Pollak et al., 1977; Weir et al., 1988; Hallauer and Miranda, 1988; Falconer, 1981) and conferences have summarized the current status of the science.

The success of modern plant breeding in improving crops of economic importance is unquestionable and well documented (Fehr, 1984). Several factors, both genetic and nongenetic, have been responsible for this improvement, and quantitative genetics undoubtedly has been one of these factors. To our knowledge, no review of the specific contributions of quantitative genetics and the closely allied field of statistics to plant improvement has been written. We believe, however, that quantitative genetics and statistics have been important in the development of systematic progeny-testing schemes and of breeding methodology, and, perhaps most important, has provided analytic tools for comparing responses to selection for various breeding methods and progeny-testing schemes. Because modern plant breeding and quantitative genetics were developed more or less simultaneously, it is difficult to separate the two or to imagine the current status of plant breeding without the aid of either quantitative genetics or statistics. What is important to recognize is the profound and lasting contributions of quantitative genetics to plant improvement.

Despite the merits and the contributions of quantitative genetics, many aspects of the theory are naive in light of modern genetic principles (Lewontin, 1977; Kempthorne, 1977, 1988). Lewontin (1977) described quantitative genetics as an attempt to produce knowledge by a systemization of ignorance. He noted that all quantitative geneticists know is that phenotypes are manifestations of genotypes expressed in environments. And genotypes result from the actions of genes organized into chromosomes that behave regularly during gametogenesis. Beyond these generalities, the biology and the architecture of quantitative traits, e.g., the number of loci controlling a trait; the number of alleles segregating per locus; the allelic frequencies; the effects of allelic substitutions; the linkage relations among loci; the epistatic interactions between loci; and the expression and regulation of genes are poorly understood. These factors are the basic building blocks of quantitative genetics, and all quantitative genetic models must make assumptions about these factors. Molecular biology has the potential to answer many relevant questions. The challenge faced by quantitative geneticists is to incorporate emerging information from molecular biology into models or to demonstrate that this information is irrelevant because of the robustness of the theory (Lewontin, 1977).

Many questions concerning the biology and the architecture of quantitative traits remain intractable or inadequately answered. Molecular biology has begun to provide information pertinent to some of these questions, but much remains to be discovered. The next sections consider five major areas relevant to plant improvement, and to which basic molecular marker research may provide deeper insight. These areas are loci number controlling quantitative traits, quantitative trait loci (QTL) nature, gene action and effects, epistasis, and genotype x environment (G x E) interaction. Each section will review briefly our assessment of what is known currently.

How many loci?

The question of the number of loci controlling quantitative traits was asked as early as 1921 by Castle (Wright, 1968). The answer to this question has important implications for both plant breeding and QTL-mapping studies. The fact remains, however, that there is no good evidence regarding whether the number of loci controlling individual quantitative traits is small (5 to 20) or large (100 or more) (Barton and Turelli, 1989). Identification of only a few loci explaining a large proportion of the variance may greatly enhance selection efficiency (Lande and Thompson, 1990; Lande, 1992), whereas, if most of the variation is accounted for by large numbers of loci with small effects, only chromosomal regions with large effects can be mapped, and there will be little hope of understanding the biology of quantitative traits by means of current molecular techniques (Barton and Turelli, 1989).

An important question to ask is why many plant breeders and others consider it given that quantitative traits are controlled by many loci with small effects. The most obvious answer is that classical quantitative genetics, with the assumptions of many loci with small effects, has adequately described short-term selection response, resemblance of relatives, inbreeding depression, and the genetic variance structure of populations (Falconer, 1981; Hallauer and Miranda, 1988). These characteristics of quantitative genetics are robust to number of loci, allelic frequencies, allelic effects, and linkage. In fact, concepts such as short-term selection response originated primarily in statistics rather than in Mendelian genetics (Barton and Turelli, 1989) and are, therefore, robust with respect to assumptions concerning the genetic architecture of a trait.

The assumption of many loci with small effects was made to simplify the models and mathematics of quantitative genetics. As Thompson (1975) has suggested these simplifying assumptions have come to be taken, over time, as established fact with little or no supporting evidence. He presents from studies with *Drosophila* evidence suggesting the actual number of loci controlling quantitative traits is much smaller than commonly assumed in quantitative genetic models. He also showed that models with three loci, with two loci controlling 90% of the variation, could give continuous distributions resembling those for characters commonly thought to be controlled by a large number of loci.

There are two ways of estimating the number of loci: biometrical techniques and chromosomal mapping studies using genetic markers (Barton and Turelli, 1989). Biometrical approaches in plant species have included the Castle-Wright formula (and modifications) (Lande, 1981) and long term selection response theory (Dudley and Lambert, 1992), both of which are limited by requisite assumptions. Most estimates of the effective number of genetic factors made by using the Castle-Wright formula have ranged from 5 to 20, depending upon trait, species, and genetic material (Lande, 1981; Barton and Turelli, 1989). Estimates obtained using the Castle-Wright approach cannot exceed the recombination index of Darlington (1937), which equals the haploid number of chromosomes plus the mean number of recombination events per gamete (Lande, 1981).

Other methods of determining the number of loci controlling a trait involve genetic or morphological markers, chromosomal rearrangements (translocations and inversions), and aneuploids. Although these methods have been used frequently in many species, there are many

limitations, particularly in the degree of resolution obtained. Chromosomal rearrangements and aneuploids are limited to mapping whole chromosome arms or whole chromosomes. For good mapping studies, neutrality of genetic markers with respect to fitness is often a requirement, which these three methods rarely can meet. Determining the number of loci controlling a quantitative trait is an important first step towards understanding the genetic architecture of quantitative traits, but there is no clear evidence that the methods used so far have effected plant improvement.

What is a QTL?

What a QTL is is an important question because one's answer will affect how one designs and interprets QTL-mapping studies. It is interesting that staunch positions are held at all in this regard, because there is very little good direct evidence bearing on the subject. At least four proposals have been made concerning the nature of loci controlling variation in quantitative traits (Mackay et al., 1992). Quantitative trait loci (QTL) (1) are "major" genes having pleiotropic effects on other traits (see Barton, 1990 for a review); (2) are different from major genes in that alleles at QTL are constrained to having only small effects on the character (Mather, 1941); (3) are modifiers of the expression of major loci (Mukai and Cockerham, 1977); and (4) have alleles with a range of effects: alleles with large effects cause recognition of the locus as a gene with major effects, and segregation of alleles with small effects gives rise to quantitative variation (Robertson, 1985).

Robertson's (1985) proposal perhaps first made by Thompson (1975), that quantitative genetic variation is the result of variation (mutations) in wildtype alleles (isoalleles), currently has the most appeal. Genetic maps of loci identified by alleles with major effects usually have resulted because of the appearance of a null or a near-null allele dramatically altering the phenotype of the plant. Robertson's proposal predicts that there should be phenotypic variation in a population due to segregation of isoalleles at these major loci. With conventional plant-breeding methods, this hypothesis would be difficult to test, except through the use of isogenic lines. Results probably would be inconclusive because of linkage drag even if the isoalleles could be identified and backcrossed. The absence of genes with major effects for quantitative traits in species with well-populated genetic maps may be due to the absence of alleles with effects great enough to be detected, to the lack of a screen for some traits, or simply to the fact that alleles with large effects have not been looked for.

Gene action

Gene action type (additive, dominant, overdominant [Falconer, 1981]) and gene effects have been studied extensively in many crop species. The type of gene action controlling a trait is very important in decisions regarding breeding method, cultivar type (inbred, hybrid, population, etc.), and interpretation of data from quantitative genetic experiments. The study of gene action has been approached in two ways (Sprague, 1966). One characterizes the predominant types of genetic variance (additive vs. dominant) in populations, an activity leading to development and analysis of mating designs, including the North Carolina mating designs (see Hallauer and Miranda, 1988 for a review). Because of the difficulties in artificial hybridization, the variance component approach is not used frequently in self-pollinated crops instead generation mean analysis has been the most prominent approach to determining gene action in these species. The results of these studies lead to

the proposal of many breeding methods that capitalize on different types of gene action, including recurrent selection for general combining ability and inbred per se selection (additive effects), recurrent selection for specific combining ability (dominance effects), and reciprocal recurrent selection (both additive and dominance effects).

The best-known example of the importance of gene action in plant improvement has been the long debate about the types of gene action involved in heterosis of grain yield in maize. Most of the literature about maize, the most extensively studied plant species, suggests that additive effects of genes with partial to complete dominance are more important than dominance effects in determining grain yield. Breeders still contend, however, that dominance effects caused by genes with overdominant gene action are also important (Horner et al., 1989). Despite five decades of research on the types of gene action and gene effects in maize and in other species, there is still debate about the types of gene action predominating for important traits. The reason for this stalemate has to do with the very nature of the quantitative genetic models themselves; that is, all estimates of gene action and gene effects have been averages over the whole genome. Estimates for individual loci have been impossible and there is most likely a distribution of gene action and effects influencing a given quantitative trait. Determining the nature of this distribution could be very important in the design of breeding programs to capitalize on the types of gene action that may exist for a trait.

Epistasis

The possibility that epistasis accounts for a significant proportion of the genetic variance of quantitative traits has been investigated extensively. The same types of experiments used to detect gene action or gene effects, that is, variance components and generation mean analysis, also were used to investigate epistasis. The role of epistasis in breeding methodology has been intriguing because there is clear evidence that genes interact. The types of interactions possible are best illustrated by the complexity of the anthocyanin pathway in maize (Coe et al., 1988). The assumption of no epistasis is one of the most common made in quantitative genetic models, and even two-locus epistatic models with linkage become nearly intractable mathematically (Weir and Cockerham, 1977). Amount and type of epistasis present in crop species can have major consequences on both the reliability of predictions and the design of breeding programs.

From both a breeding methodology and statistical point of view, epistasis is difficult to estimate. Reports of estimates of epistasis in self-pollinated species are rare and most epistasis studies have been conducted with maize. Studies estimating epistasis in maize are too numerous for comprehensive review, but a few interesting conclusions can be drawn. Studies estimating epistasis using generation mean analysis generally have reported significant epistatic effects. Estimates made using the analysis of variance (covariance of relatives) approach generally have reported nonsignificant epistatic effects. Studies with open pollinated varieties generally have shown additive effects to be more important than dominance or epistatic effects, and studies with elite inbred lines generally have reported dominance and epistatic effects to be more important than additive effects.

Statistically, there is more power to detect epistasis using generation mean analysis than by

using the analysis of variance approach and the reference population for generation mean analysis is often much narrower than for the analysis of variance. This difference has lead to a confounding of statistical power with the method of estimation and reference population such that the inability to detect epistasis using the analysis of variance may be due to lack of statistical power or to the absence of epistasis. With both methods, the inability to detect epistasis cannot be taken as evidence for the absence of epistasis because of the canceling of epistatic effects among loci. New approaches and methods for estimating epistasis for quantitative traits are needed.

Genotype x environment interaction

If the number of published papers is any indication of the importance of a topic, then G x E interaction must be one of the most important problems facing plant breeders. There are numerous reviews of G x E interaction (see Crossa, 1990 for a review) that illustrate the complexity of the subject. One noticeable theme throughout the literature, is that G x E interaction has been approached primarily from a statistical point of view rather than from a biological one. It would be quite difficult to assess the role that this research has played in plant improvement, and we will not attempt to do this. Rather, we wish to make the point that our understanding of G x E interaction will improve only with knowledge of how genes are regulated by environments. The understanding of G x E interaction for quantitative traits then will require an understanding of the genetic architecture of these traits.

Not a problem confined to plant breeding and its experiments, G x E interaction also must be dealt with in molecular marker research. Whenever plants are evaluated in the field, there is a G x E interaction; and because all applications of molecular markers will require field evaluation, e.g., a linkage of phenotype to genotype, potentially at many stages, the same challenge exists as for plant breeding: a lack of biological causes for the G x E interaction. G x E interaction will not be a simple hurdle for molecular marker research unless the regulation of genes by environments proves very simple. Plant breeding has dealt with G x E interaction by evaluating commercial cultivars and hybrids in hundreds of environments over many years. For this reason alone, some applications of molecular markers will have limited impact on some aspects of plant breeding. Molecular markers, however, may provide insights into G x E interaction that statistical research so far has been unable to provide.

Applications of molecular markers

Genetic diversity

Assessment of genetic diversity is important in plant breeding if there is to be improvement by selection. For assessment of genetic diversity, molecular markers have been generally superior to morphological, pedigree, heterosis, and biochemical data (isozymes and chromatography) (Melchinger et al., 1991; Melchinger, 1993). Genetic diversity commonly is measured by genetic distance (GD) or genetic similarity (GS = 1 - GD), both of which imply that there are either differences or similarities at the genetic level (Weir, 1990, p. 162).

Published applications of molecular marker-based GD in plant breeding have been limited

primarily to maize, but results should apply to other species. Melchinger (1993) reviewed the application of molecular marker-based GD for assigning maize inbred lines to heterotic groups, determining the relation between inbred lines and hybrids, and predicting hybrid performance. Data showed that GS calculated from molecular marker data faithfully separated inbred lines into their heterotic groups. There also seems to be promise of assigning inbreds of unknown pedigree to heterotic groups although a large number of markers (> 100) and well-characterized reference populations may be needed to obtain an accurate assessment of GS. Strong correlations (0.61 to 0.95) between Malecot's coancestry coefficient *(f)* and GS for related ($f > 0$) genotypes indicated that pedigree data provide reliable estimates of GS. Genetic similarity estimates based on molecular markers are expected to be superior to estimates of f because of unreliable or incomplete pedigree data and because of the assumptions required to calculate f. Molecular marker-based GD has some potential for predicting hybrid performance of related lines, but in typical hybrid-breeding programs, in which hybrids are produced from unrelated lines from different heterotic groups, molecular marker-based GD has been of no value in predicting hybrid performance. These results suggest the use of molecular marker-based GD for predicting hybrid performance in crops in which either hybrids are being explored, such as wheat (*Triticum aestivum* [Vill., Host] Mackey), and soybean (*Glycine max* (L.) Merr.), or hybrid breeding is being practiced, but distinct heterotic groups have not been developed.

Molecular marker-based GD also has potential for assessing changes in genetic diversity over time (Duvick, 1984), protection of intellectual property rights (Smith et al., 1990), registration of germplasm in countries having ratified the rules of the UPOV convention, and evaluation of new sources of germplasm for their potential to increase genetic diversity (Smith and Smith, 1992). Direct applications to plant breeding, however, have been limited so far to prediction of hybrid performance. But accurate prediction of hybrid performance does not seem likely unless gene action is primarily dominant or overdominant, complementary heterotic groups are established, trait heritability is high, at least 30 to 50% of the markers are linked to QTL, and no more than 20 to 30% of the markers are dispersed randomly (Bernardo, 1992). Molecular markers, however, may be useful for early generation testing in hybrid-breeding programs. If individual markers or marker intervals associated with combining ability can be identified when a plant or progeny is crossed onto a given tester, then these markers could be used as a first screen to identify the top 50% of the progenies for field evaluation. Although this procedure would not decrease the time to cultivar development, it would decrease the amount of material tested or permit the evaluation of a wider range of germplasm for the same amount of field resources.

An equally important application of molecular marker-based GD may be in the selection of parents to cross in a breeding program. This application deserves serious attention because breeders currently rely primarily on pedigree and performance data for choosing parents in breeding programs. Using molecular markers to select parents has the potential to allow simultaneous maintenance of genetic diversity and performance. Dudley et al. (1992) presented one application of molecular markers for choosing parents, and additional research is needed in this area. Using molecular markers to choose parents likely will require establishment of a relation between GD and genetic variation, and many of the same conditions necessary for predicting hybrid performance may be required for choosing parents. Using molecular markers as a diagnostic tool to survey new or exotic germplasm for novel genetic diversity also may be possible. It is unlikely, however, that

this use will be possible with random genomic or cDNA clones because molecular marker-based genetic diversity will not guarantee genetic diversity for the traits of interest. Screening with probes of expressed genes with known function offers the greatest potential in this area.

Mapping QTL

The vast majority of molecular marker research in quantitative traits has been devoted to mapping QTL. Mapping QTL is really a misnomer, because what is actually being done is the mapping of chromosomal regions containing one or more putative QTL. With current mapping technology, the existence of a single QTL between two flanking markers cannot be resolved clearly. Most studies reported to date have detected, localized, and estimated genetic effects in the same experiment because of resource limitations. Genetic effects of mapped QTL regions are overestimated by this procedure because of sampling errors (Lande and Thompson, 1990; Lande, 1992). Furthermore, few researchers have followed up with the necessary experiments to verify the effect of a chromosomal region on phenotype across mapping populations. Our purpose in this section is not to review QTL mapping studies in detail, but rather to outline the steps taken in QTL mapping experiments, to demonstrate the general results that have been obtained, to outline some of the problems in translating these results into plant improvement, and to show the types of previously unattainable information that these results have contributed.

The first step in QTL mapping studies is to detect QTL, while minimizing the occurrence of false positives (Type I errors, that is, declaring an association between a marker and QTL when in fact one does not exist). Two distinct methods are used to detect QTL. The single marker approach, sometimes referred to as the one-way analysis of variance (ANOVA), has been used extensively, especially with isozymes (Tanksley et al., 1982; Edwards et al., 1987). The second approach, interval mapping, detects QTL by using flanking markers. This approach is more complicated analytically than the ANOVA approach and involves application of the maximum likelihood method, which requires sophisticated computer software (Lander and Botstein, 1989). Lander and Botstein (1989) have developed formulae for calculating significance levels appropriate for both methods when the genome size, number of chromosomes, number of marker intervals, and the overall false positive rate desired are given. Several statistical procedures have been developed for the application of both ANOVA and interval mapping (Soller and Brody, 1976; Edwards et al., 1987; Weller, 1987; Lander and Botstein, 1989; Knapp, 1989). When the same false positive rates are used, there are few reasons to suspect that the two methods would detect substantially different QTL. Stuber et al. (1992) compared the two methods and found that they identified basically the same QTL. Those researchers reported, however, some advantages to using the interval mapping approach. Because of the increased power associated with using flanking markers, the method gives the most likely location of the QTL under the assumption of a single QTL in the interval, and the interval mapping approach allows ambiguous or missing data.

Once QTL are detected, the next step is to estimate the genotypic effect of the QTL and to localize the QTL to a precise genomic region. The interval mapping approach seems superior to the ANOVA approach for both estimation of effects and localization of the QTL (Stuber et al., 1992). The success of both methods depends on the linkage between marker(s) and QTL, the number and type of progeny evaluated, the heritability of the trait, and the magnitude of the effects at QTL that

one desires to detect. Several methods and genetic designs have been suggested for detecting, estimating effects, and localizing QTL (Cowen, 1988; Burr et al., 1988).

Manipulating traits controlled by a few major loci

The manipulation of traits controlled by a few major loci (loci that can be studied using Mendelian genetics) may offer the greatest promise in the short term for application to plant breeding. The primary applications of this technique will be for traits controlled by a single gene (monogenic) or those controlled by at most two or three loci (oligogenic). The most successful applications will be in those species with well developed molecular marker maps. These applications will be immediately useful for "defensive breeding," that is, when a desirable genotype is available but lacks resistance to important insects and diseases. Other applications, not fitting into the category of defensive breeding may include seed modifications controlled by a few genes, restorer genes for cytoplasmic male sterility, dwarfing genes for shorter plant height, and maturity genes for adaptability.

The first requirement of using molecular markers in this context is to develop a precise molecular marker linkage map and then to use these markers to map gene(s) controlling the trait of interest. Many methods for mapping genes of interest are available, including a variety of applications suitable to most species with polymorphic markers (Melchinger, 1990). The final step is to use marker-facilitated selection to transfer the gene(s) to the genotypes desired. Two methods are available, both of which begin with inbred lines, backcross selection, and pedigree selection.

If the cost of molecular-marker technology is ignored, the primary factor affecting the design and the success of marker-facilitated selection is how tightly linked a single marker is to the gene or how tightly bracketed the gene is by two markers. The idea is to obtain a marker or a set of flanking markers linked tightly enough to the gene so that a recombination event does not occur between the marker and the gene during backcrossing or pedigree selection. Melchinger (1990) presented extensive theoretical and numerical results for the backcross method concerning the optimal family size and the number of plants per family that must be genotyped with molecular markers. The results are complicated and will have to be assessed on a case by case basis. The economics of marker assisted backcrossing will be a function of the cost of marker assays, the cost of direct screening, and the value of accelerating the backcrossing program. Results regarding sample sizes required for pedigree selection are unavailable, although marker-gene linkage is the primary consideration (Dudley, 1993).

There are many practical applications of molecular markers to traits controlled by few loci (Melchinger, 1990). Most of the applications involve situations in which either screening for the trait is difficult or scoring of the trait occurs late in plant development. These applications may include pests for which natural inoculum is unreliable or artificial inoculation procedures are undeveloped or unreliable. Examples include nematodes or *Aspergillus*, both of which have broad host ranges and unreliable natural and artificial inoculation. Diseases in which resistance is influenced strongly by the environment also would be good candidates for marker facilitated selection. Unfortunately, the very situations favoring marker-facilitated selection also make it difficult to map the resistance genes precisely. Marker-facilitated selection has been advantageous

for backcrossing recessive genes, when progeny tests are needed after every backcross generation to identify heterozygotes or when resistance can be determined only after flowering. Markers in these situations could greatly reduce workload and backcrossing time. Other examples include pyramiding resistance genes, developing multilines in which many race-specific resistance genes are involved that are sometimes difficult to distinguish, and selecting for resistance to exotic or quarantined pathogens.

One of the primary advantages of marker-facilitated backcrossing has been in increasing the speed of recovery of the recurrent parent genome (Tanksley and Rick, 1980). In addition to having a tight marker-gene linkage, one or more neutral polymorphic markers will be required per chromosome arm. The idea is to screen for plants having the resistance gene and to identify those plants with the greatest proportion of markers homozygous for the recurrent parent. A possible limitation of this procedure is that there may be an increase in the number of plants needing to be assayed with markers. The procedure could be applied, however, to subsequent backcross generations to ensure the recovery of unlinked segments of recurrent parental genome.

Manipulating traits controlled by many loci

The molecular marker manipulation of traits controlled by many loci (from a plant breeding perspective many is generally greater than five) is of great interest to plant breeders and represents one of the fields greatest challenges. Plant breeders concentrate effort on breeding for quantitative traits, and breeding for qualitative traits is generally a trivial, albeit time-consuming, process. The matter is further complicated because breeders usually evaluate --simultaneously in many populations -- four or more complexly inherited traits. Obviously, any technology enhancing the breeders' ability to select for these traits would be adopted rapidly.

The molecular marker manipulation of qualitative traits is feasible precisely because so much is known about the biology of these traits. The number of loci is known; there are no questions about what a major locus is -- in fact, many of these loci have been cloned; the gene action is known precisely; epistatic interactions, if any, are relatively easy to decipher; and interactions with environment are easy to determine. In short, the amount that can be known about qualitative traits is limited only by one's desire to know. As pointed out earlier, for quantitative traits, the answers to these questions are based only upon averages over an unknown number of loci. At the outset, the manipulation of quantitative traits by molecular markers has a serious disadvantage.

Even with these limitations, there still may be many applications of marker-QTL associations. But very little theoretical or applied research has been published in this area. The most-cited application has been marker-assisted selection (MAS) although the context of application often has been ignored. In plant breeding, there are two distinct but related applications: germplasm enhancement (recurrent selection) and cultivar or hybrid development. These two applications are separated because recurrent selection usually is applied to random mating populations possibly at or near linkage equilibrium, whereas cultivar or hybrid development typically begins with populations derived by crossing elite inbred lines at or near maximum linkage disequilibrium . Clearly, two different approaches are needed for these breeding schemes.

Lande and Thompson (1990) and Lande (1992) investigated the efficiency of MAS for both individual and mass selection in random-mating populations. There are three approaches to applying MAS to plant breeding: (1) selection on markers alone with no measurement of phenotype; (2) simultaneous selection on markers and phenotype; and (3) two-stage selection, the first stage involving use of markers to select among seedlings and the second involving phenotypic selection among surviving adults. On individuals, the efficiency of MAS relative to that of phenotypic selection of the same intensity is $(p/h^2)^{1/2}$, where p is the proportion of the additive genetic variance accounted for by markers, and h^2 is the heritability. Selection on markers alone will be more efficient than phenotypic selection only when the proportion of the genetic variance explained by markers exceeds the heritability of the trait. Therefore, selection on markers alone will be most useful for traits with low heritabilities when large proportions of their variability have been explained by markers. Lande and Thompson (1990) concluded that molecular marker loci can be used to enhance the efficiency of artificial selection for quantitative traits. The potential efficiency of MAS depends upon the heritability of the trait, the proportion of the genetic variance explained by the markers, and the selection method. A major practical problem in using MAS is that recombination will reduce linkage disequilibrium between the markers and QTL, thus diminishing selection effectiveness. The successful application of MAS will require very tight linkages between markers and QTL.

Understanding the genetic architecture of quantitative traits

Although there have been few direct applications of molecular markers in plant breeding, published research has begun to provide information on the genetic architecture of quantitative traits. Molecular marker-mapping data from several species now suggest that genetic variation for quantitative traits is controlled by a few loci with large effects and many loci with progressively smaller effects (Lande, 1992). Beavis et al. (1991) found 14 genomic regions associated with plant height in four F_2 maize populations, but few of these regions were in common across populations. All 14 regions, however, were associated with known qualitative genetic loci. Thus, circumstantial evidence supports Robertson's (1985) hypothesis that QTL have alleles with a range of effects, alleles with large effects causing recognition of the locus as qualitative, and alleles with small effects causing recognition of the locus as quantitative.

The results from QTL-mapping studies regarding gene action and epistasis are, so far, difficult to interpret. Stuber et al. (1992) reported that, with one exception, all the QTL mapped for grain yield in maize showed heterozygote superiority. Either overdominance or pseudo-overdominance, therefore, likely was important in the mapping population. These results cannot separate the two causes of heterozygote superiority, primarily because the number of QTL residing in a marker interval cannot be resolved. Although not mentioned by the authors, heterozygote superiority also could result from overestimation of effects as pointed out by Lande and Thompson (1990) and Lande (1992). Stuber et al. (1992) and Paterson et al. (1991) found little evidence for epistasis in maize and tomato (*Lycopesicon esculentum* [E]), respectively. Paterson et al. (1991) concluded that molecular marker-mapping studies conducted to date may identify preferentially QTL that function independently of unlinked genetic factors. Those researchers suggested that the role of epistasis in quantitative inheritance needs to be studied in larger populations, with more closely spaced markers, and/or with specially constructed genetic stocks carrying particular QTL.

Stuber et al. (1992) found limited evidence for G x E interaction even though a standard analysis of the data revealed significant G x E interaction for many of the traits. Thus it may be possible to detect QTL with large effects in relatively few environments. Paterson et al. (1991) reported that individual QTL showed a range of sensitivities to environments, some QTL being detected in all environments and some being detected in only one environment.

Conclusions

One of the most difficult problems in plant breeding is identifying plants or progenies containing the desired gene combinations for the traits of interest. From a plant-breeding perspective, two factors are involved in this process: selection and genetic variance generation. For most plant species of economic importance, generation of genetic variation is accomplished easily by hybridization. Identification of the desired recombinants from the pool of genetic variation is difficult and requires excellent screens for the traits of interest. The effectiveness of selection may be limited more by the reliability of the screen for the trait than by any other factor.

Quantitative genetics has provided plant breeding with the theoretical foundation of effective screening, choice of progeny type, choice of tester, relations among traits, and comparison of alternative methods. Developments in statistics have improved the effectiveness of the screen by removing bias in comparisons (randomization), determining the amount of error in comparisons (replication), and removing unwanted sources of variation from comparisons (blocking). These methods all rely on phenotypic observations, and, for many traits, the underlying genetic component of the phenotypic observations is understood poorly.

Molecular markers may complement plant breeding in three general ways. First, molecular markers provide a reliable genetic-diversity measure that can be used for determining relations among inbred lines and cultivars, assessing changes in genetic diversity over time, protecting intellectual property rights, registering germplasm in countries that have ratified rules of the UPOV convention, evaluating new germplasm for its potential to increase genetic diversity, and selecting parents to hybridize in a breeding program. Second, molecular markers through their linkage with alleles with large effects (qualitative traits) and alleles with small effects (quantitative traits) may improve screens for many traits. Third, molecular markers will provide the first understanding of the biology and the architecture of many traits, particularly of quantitative traits.

Adaptation and application of molecular markers to plant improvement will be unique for each species and breeding program. Theoretically, many of the proposed applications of molecular markers are viable. The question is whether they will improve the efficiency and the cost effectiveness of a breeding program. This question can be answered only on a case-by-case basis. Factors such as cost of the molecular marker technology, turnaround time in the lab, cost of measuring a trait, etc. all will determine if and how markers are used in breeding programs. Species in which traits are measured by processing through a commercial factory or species with very long generation times clearly will benefit from applications of molecular-marker technology. For species, such as annual grains and cereals, the situation is ambiguous. One of the primary contributions of molecular markers will be an expansion of our knowledge of genetics and of

genome organization (Cox, 1992). This type of knowledge obviously will improve our scientific understanding of many plant breeding problems, but the direct effect on plant improvement will be intangible and difficult to measure.

Acknowledgements

Joint contribution from the Field Crops Research Unit, USDA, Agricultural Research Service, Dep. of Agronomy, Iowa State Univ. and Journal Paper No. J-15223 of the Iowa Agricultural and Home Economics Experiment Station. Project Nos. 3082 and 3134.

Bibliography

Barton, N. H. 1990. Pleiotropic models of quantitative variation. Genetics 124:773-782.

Barton, N. H., and M. Turelli. 1989. Evolutionary quantitative genetics: How little do we know? Annu. Rev. Genet. 23:337-370.

Beavis, W. D., D. Grant, M. Albertsen, and R. Fincher. 1991. Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. Theor. Appl. Genet. 83:141-145.

Beckman, J. S., and M. Soller. 1986. Restriction fragment length polymorphisms and genetic improvement of agricultural species. Euphytica 35:111-124.

Bernardo, R. 1992. Relationship between single-cross performance and molecular marker heterozygosity. Theor. Appl. Genet. 83:628-634.

Burr, B., F. A. Burr, K. H. Thompson, M. C. Albertson, and C. W. Stuber. 1988. Gene mapping with recombinant inbreds in maize. Genetics 118:519-526.

Burr, B., S. V. Evola, F. A. Burr, and J. S. Beckman. 1983. The application of restriction fragment length polymorphisms to plant breeding. p. 45-59. In: J. K. Setlow and A. Hollander (eds.) Genetic engineering: Principles and methods, Vol. 5. Plenum Publishing Corp., New York, NY.

Coe, E. H. Jr., M. G. Neuffer, and D. A. Hoisington. 1988. The genetics of corn. p. 83-258. In: G. F. Sprague and J. W. Dudley (eds.) Corn and Corn Improvement. Am. Soc. Agron., Madison, WI.

Cowen, N. M. 1988. The use of replicated progenies in marker-based mapping of QTL's. Theor. Appl. Genet. 75:857-865.

Cox, T. S. 1992. Use of molecular markers for analyzing quantitative traits in wheat. In: Proc. 19th Hard Red Winter Wheat Workers Conf.

Crossa, J. 1990. Statistical analyses of multilocation trials. Adv. Agron. 44:55-85.

Darlington, C. D. 1937. The biology of crossing-over. Nature 140:759-761.

Dudley, J. W. 1993. Molecular markers in plant improvement: Manipulation of genes affecting quantitative traits. Crop Sci. 33:(in press).

Dudley, J. W. and R. J. Lambert. 1992. Ninety generations of selection for oil and protein in maize. Maydica 37:81-87.

Dudley, J. W., M. A. Saghai Maroof, and G. K. Rufener. 1992. Molecular marker information and selection of parents in corn breeding programs. Crop Sci. 32:301-304.

Duvick, D. N. 1984. Genetic diversity in major farm crops on the farm and in reserve. Econ. Bot. 32:161-178.

Edwards, M. D., C. W. Stuber, and J. F. Wendel. 1987. Molecular marker-facilitated investigations of quantitative trait loci in maize. I. Numbers, distribution, and types of gene action. Genetics 116:113-125.

Falconer, D. S. 1981. Introduction to quantitative genetics. 2nd ed. Longman, New York.

Fehr, W. R. (ed.) 1984. Genetic contributions to yield gains of five major crop plants. Special publ. No. 7. Crop Sci. Soc. of Am., Madison, WI.

Fehr, W. R. (ed.). 1987. Principles of cultivar development. Vol. 2. Crop Species. Macmillan Publ. Co., New York.

Fisher, R. A. 1918. The correlation between relatives under the supposition of mendelian inheritance. Trans. R. Soc. Edinburgh 52:399-433.

Haldane, J. B. S. 1932. The causes of evolution. Longmans, Green, London.

Hallauer, A. R. 1990. Methods used in developing maize inbreds. Maydica 35:1-16.

Hallauer, A. R. and J. B. Miranda, FO. 1988. Quantitative genetics in maize breeding. 2nd ed. Iowa State University Press, Ames.

Helentjaris, T. G. 1992. RFLP analyses for manipulating agronomic traits in plants. p. 357-372. In: H. R. Stalker and J. P. Murphy (eds.) Plant breeding in the 1990s. CAB International, Wallingford, UK.

Helentjaris, T. G., G. King, M. Slocum, C. Siedenstrang, and S. Wegman. 1985. Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. Plant Mol. Biol. 5:109-118.

Horner, E. S., E. Magloire, and J. A. Morera. 1989. Comparison of selection for S_2 progeny vs.

testcross performance for population improvement in maize. Crop Sci. 29:868-874.

Kempthorne, O. 1977. Status of quantitative genetic theory. p. 719-760. In: E. Pollak, O. Kempthorne, and T. B. Bailey, Jr. (eds.) Proc. of the Int. Conf. on Quantitative Genetics, 16-21 Aug. 1976. Iowa State University Press, Ames.

Kempthorne, O. 1988. An overview of the field of quantitative genetics. p. 47-56. In: B. S. Weir, E. J. Eisen, M. M. Goodman, and G. Namkoong (eds.) Proc. of the Second Int. Conf. on Quantitative Genetics, 31 May - 5 June 1987, Sinauer Associates Inc., Sunderland, Mass.

Knapp, S. J. 1989. Quasi-Mendelian analyses of quantitative traits using molecular marker linkage maps. p. 51-67. In: G. Robbelen (ed.) Science for plant breeding, Proceedings of the XII congress of EUCARPIA. Paul Parey Scientific Publishers, Berlin.

Lande, R. 1981. The minimum number of genes contributing to quantitative variation between and within populations. Genetics 99:541-553.

Lande, R. 1992. Marker-assisted selection in relation to traditional methods of plant breeding. p. 437-451. In: H. R. Stalker and J. P. Murphy (eds.) Plant breeding in the 1990s. CAB International, Wallingford, UK.

Lande, R. and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743-756.

Lander, E. S. and D. Botstein. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 124:743-756.

Lewontin, R. C. 1977. The relevance of molecular biology to plant and animal breeding. p. 55-62. In: E. Pollak, O. Kempthorne, and T. B. Bailey, Jr. (eds.) Proc. of the Int. Conf. on Quantitative Genetics, 16-21 Aug. 1976. Iowa State University Press, Ames.

Mackay, T. F. C., R. F. Lyman, and M. S. Jackson. 1992. Effects of *P* element insertions on quantitative traits in *Drosophila melanogaster*. Genetics 130:315-332.

Mather, K. 1941. Variation and selection of polygenic characters. J. Genet. 41:159-193.

Melchinger, A. E. 1990. Use of molecular markers in breeding for oligogenic disease resistance. Plant Breeding 104:1-19.

Melchinger, A.E. 1993. Use of RFLP markers for analyses of genetic relationships among breeding materials and prediction of hybrid performance. Proc. of the First Int. Crop Sci. Cong.

Melchinger, A.E., M. Messmer, M. Lee, W. L. Woodman, and K. R. Lamkey. 1991. Diversity and relationships among U.S. maize inbreds revealed by restriction fragment length polymorphisms. Crop Sci. 31:669-678.

Mukai, T. and C. C. Cockerham. 1977. Spontaneous mutation rates at enzyme loci in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 74:2514-2517.

Paterson, A. H., S. D. Tanksley, and M. E. Sorrells. 1991. DNA markers in plant improvement. Adv. Agron. 46:39-90.

Pollak, E., O. Kempthorne, and T. B. Bailey, Jr. 1977. Proceedings of the first international conference on quantitative genetics. Iowa State University Press, Ames.

Robertson, D. S. 1985. A possible technique for isolating genic DNA for quantitative traits in plants. J. Theor. Biol. 117:1-10.

Shull, G. 1908. The composition of a field of maize. Am. Breeder's Assoc. Rep. 4:296-301.

Shull, G. 1909. A pure-line method of corn breeding. Am. Breeder's Assoc. Rep. 5:51-59.

Simmonds, N. W. 1991. Bandwagons I have known. Trop. Agric. Newsl. December 7-10.

Smith, O. S. and J. S. C. Smith. 1992. Measurement of genetic diversity among maize hybrids; a comparison of isozymic, RFLP, pedigree, and heterosis data. Maydica 37:53-60.

Smith, O. S., J. S. C. Smith, S. L. Bowen, R. A. Tenborg, and S. J. Walls. 1990. Similarities among a group of elite maize inbreds as measured by pedigree, F_1 grain yield, grain yield heterosis, and RFLP's. Theor. Appl. Genet. 80:833-840.

Soller, M., and T. Brody. 1976. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theor. Appl. Genet. 47:35-39.

Sprague, G. F. 1966. Quantitative genetics in plant improvement. p. 315-354. In: K. J. Frey (ed.) Plant Breeding, Iowa State University Press, Ames.

Stuber, C. W., S. E. Lincoln, D. W. Wolff, T. Helentjaris, and E. S. Lander. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbreds using molecular markers. Genetics 132:823-839.

Tanksley, S. D., H. Medina-Filho, and C. M. Rick. 1982. Use of naturally-occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity 49:11-25.

Tanksley, S. D. and C. M. Rick. 1980. Isozymic gene linkage map of the tomato: applications in genetics and breeding. Theor. Appl. Genet. 57:161-170.

Thompson, J. N., Jr. 1975. Quantitative variation and gene number. Nature 258:665-668.

Weir, B. S. 1990. Genetic data analysis. Sinauer Associates, Inc., Suderland, Massachusetts.

Weir, B. S. and C. C. Cockerham. 1977. Two-locus theory in quantitative genetics. p. 247-269. In: E. Pollak, O. Kempthorne, and T. B. Bailey, Jr. (eds.) Proc. of the Int. Conf. on Quantitative Genetics, 16-21 Aug. 1976. Iowa State University Press, Ames.

Weir, B. S., E. J. Eisen, M. M. Goodman, and G. Namkoong (eds.). 1988. Proceedings of the second international conference on quantitative genetics. Sinauer Assoc. Inc., Sunderland, MA.

Weller, J. I. 1987. Mapping and analysis of quantitative trait loci in *Lycopersicon* (tomato) with the aid of genetic markers using approximate maximum likelihood methods. Heredity 39:413-421.

Wright, S. 1921. Systems of mating. Genetics 6:111-178.

Wright, S. 1968. Evolution and genetics of populations. Vol. I. Genetic and biometric foundations. University of Chicago Press, Chicago.