Marker-Assisted Selection in Plant Breeding: From Publications to Practice

Yunbi Xu* and Jonathan H. Crouch

ABSTRACT

The volume of publications on the development and to a lesser extent the application of molecular markers in plant breeding has increased dramatically during the last decade. However, most of the publications result from investments from donors with a strategic science quality or biotech advocacy mandate leading to insufficient emphasis on applied value in plant breeding. Converting promising publications into practical applications requires the resolution of many logistical and genetical constraints that are rarely addressed in journal publications. This results in a high proportion of published markers failing at one or more of the translation steps from research arena to application domain. The rate of success is likely to increase due to developments in gene-based marker development, more efficient quantitative trait locus (QTL) mapping procedures, and lower cost genotyping systems. However, some fundamental issues remain to be resolved, particularly regarding complex traits, before marker-assisted selection realizes its full potential in public sector breeding programs. These include the development of high throughput precision phenotyping systems for QTL mapping, improved understanding of genotype by environment interaction and epistasis, and development of publicly available computational tools tailored to the needs of molecular breeding programs.

Genetic Resources and Enhancement Unit, International Maize and Wheat Improvement Center (CIMMYT), Apdo, Postal 6-641, 0660 Mexico, D.F., Mexico. Received 7 Apr. 2007. *Corresponding author (y.xu@cgiar.org).

Abbreviations: CGIAR, Consultative Group on International Agriculture Research; CIMMYT, International Maize and Wheat Improvement Center; cM, centimorgan; GCP, the Generation Challenge Program; GEI, genotype by environment interaction; ICIS, International Crop Information System; iMAS, integrated marker-assisted selection; LD, linkage disequilibrium; LIMS, Laboratory Information Management Systems; MAB, marker-assisted breeding; MABC, marker-assisted backcrossing; MAS, marker-assisted selection; MAYG, mapping-as-you-go; MTA, marker–trait association; NIL, near isogenic line; PCR, polymerase chain reaction; QEI, QTL by environment interaction; QTL, quantitative trait locus or loci; RFLP, restriction fragment length polymorphism; SNP, single nucleotide polymorphism; SSR, simple sequence repeat.

Developments in genomics have provided new tools for discovering and tagging novel alleles and genes. These tools can enhance the efficiency of breeding programs through their use in marker-assisted selection (MAS). In this way, the selection of target traits can be achieved indirectly using molecular markers that are closely linked to underlying genes or that have been developed from the actual gene sequences. More specifically, plant breeding will benefit from the use of genomics tools through (i) more effectively identifying, quantifying, and characterizing genetic variation from all available germplasm resources (Tanksley et al., 1989; Tanksley and McCouch, 1997; Gur and Zamir, 2004); (ii) tagging, cloning, and introgressing genes and/or quantitative trait loci (QTL) useful for enhancing the target trait using genetic transformation and

Published in Crop Sci. 48:391–407 (2008).
doi: 10.2135/cropsci2007.04.0191

677 S. Segoe Rd., Madison, WI 53711 USA

molecular marker technologies (Dudley, 1993; Gibson and Somerville, 1993; Paterson, 1998; Gur and Zamir, 2004; Peters et al., 2003; Peña, 2004; Holland, 2004; Salvi and Tuberosa, 2005); and (iii) manipulating (differentiating, selecting, pyramiding, and integrating) genetic variation in breeding populations (Stuber, 1992; Xu, 1997; Collard et al., 2005; Francia et al., 2005; Varshney et al., 2005a; Wang et al., 2007). Genomics techniques can also have significant utility in plant breeding programs through assisting plant variety protection as well as distinctness, uniformity, and stability testing processes (CFIA/NFS, 2005; Heckenberger et al., 2006; IBRD, 2006), but these applications are beyond the scope of this paper which focuses on the use of MAS to improve the efficiency and scope of crop improvement for specific traits.

The development of molecular markers for plant breeding applications was first popularized in the early 1980s when isozyme markers were used to speed up the introgression of monogenic traits from exotic germplasm into a cultivar background (Tanksley and Rick 1980; Tanksley 1983). A few years later, Beckmann and Soller (1986a) described the first use of restriction fragment length polymorphism (RFLP) markers in crop improvement including theoretical issues related to marker-assisted backcrossing (MABC) for improvement of qualitative traits. Lande and Thompson (1990) then pioneered the theoretical studies of MAS for quantitative traits, which triggered the publication of a series of simulation studies through the 1990s (e.g., Gimelfarb and Lande, 1994a, 1994b, 1995; Zhang and Smith, 1992, 1993; Hospital and Charcosset, 1997; Whittaker et al., 1997). More recently, additional theoretical considerations regarding the application of MAS have been addressed, including the optimization of MABC systems (for instance, Frisch and Melchinger, 2001, 2005; Hospital, 2002) and strategies to pyramid favorable alleles through recurrent crossing schemes (Hospital et al., 2000; Servin et al., 2004; Bernardo et al., 2006). These theoretical studies have greatly contributed to our understanding of many fundamental genetical issues regarding the development of MAS systems such as population type, sample size, genome size, and marker number. Similarly there are many reviews comparing the molecular genetic issues related to different types of marker assay which can also significantly affect the success of MAS (Avise, 2004; Guimarães et al., 2007). Thus, in this review we focus on the technical issues that are critically important for the successful translation of promising markers into effective MAS programs.

WHY USE MARKER-ASSISTED SELECTION IN PLANT BREEDING

Justifications for the development and use of MAS in plant breeding fall into four broad areas that are relevant to almost all target crops (Young and Tanksley, 1989; Ribaut and Hoisington, 1998; Xu, 2002, 2003; Koebner, 2004; Xu et al., 2005): (i) traits that are difficult to manage through conventional phenotypic selection—because they are expensive or time-consuming to measure, or have low penetrance or complex inheritance; (ii) traits whose selection depends on specific environments or developmental stages that influence the expression of the target phenotype; (iii) maintenance of recessive alleles during backcrossing or for speeding up backcross breeding in general; and (iv) pyramiding multiple monogenic traits (such as pest and disease resistances or quality traits) or several QTL for a single target trait with complex inheritance (such as drought tolerance or other adaptive traits).

There are many modeling and simulation studies regarding the power of markers to improve the pace and precision of backcross breeding. For most crops, over 90% of the recurrent parental genotype can be recovered within two generations when a suitable number of markers (e.g., one marker every 10 cM) and an adequate number of progeny is used for background selection (Tanksley et al., 1989). This represents a substantial saving in time compared to conventional backcross breeding. Molecular markers intended for MABC can be selected based on (i) their genome distribution; (ii) haplotype diversity and/or polymorphic information content indices; and (iii) their association with candidate genes and other agronomic traits (excluding target introgression trait) (Xu, 2003; Varshney et al., 2005b). Marker-assisted backcrossing has been shown to be especially valuable where there are many good varieties that need to be improved for just one simply inherited trait such as certain pest or disease resistances, or a component trait for enhancing adaptation or stress tolerance (Cregan et al., 1999; Cahill and Schmidt, 2004; Johnson, 2004; Niebur et al., 2004; Eathington, 2005; Crosbie et al., 2006; Ragot and Lee, 2007; reviewed by Xu, 2003; Miklas et al., 2006; Dwivedi et al., 2007). With recent advances in high-throughput genotyping systems (e.g., Gunderson et al., 2005; Syvänen, 2005; Bai et al., 2007), MABC is likely to become increasingly cost effective.

Introgression and pyramiding of multiple genes affecting the same trait is a great challenge to breeding programs. The target cropping environments of many breeding programs require a combination of diverse biotic stress resistances, agronomic and quality trait profiles, plus abiotic stress tolerances to improve performance, yield stability, and farmers' acceptance. The greatest impact from MAS will only be realized when breeding systems are adapted to make best use of large-scale genotyping for both multiple target traits and the genetic background. The greatest benefits from this type of integrated molecular breeding approach will be to achieve the same breeding progress in a much shorter time than through conventional breeding, and from pyramiding combinations of genes that could not be readily combined through other means.

As conventional breeding systems attempt to combine simultaneous selection for more and more target traits, there tends to be an overall loss of breeding gain and an increase in the number of breeding cycles required to generate a finished product. In contrast, MAS offers the potential to assemble target traits in the same genotype more precisely, with less unintentional losses and in fewer selection cycles.

The opportunities for improving more complex traits such as abiotic stress tolerances are confounded by low heritability, large number of contributing genes with unpredictable epistasis, and the effects of various environmental factors. Thus, establishing routine solutions for MAS of these traits still remains a challenge.

PRIMARY CONSTRAINTS TO SUCCESSFUL MARKER-ASSISTED SELECTION

To analyze the bottlenecks that may limit the application of MAS in plant breeding, we will first briefly overview the current state of the art. Substantial investments have been made by the private sector for the development of genomics tools for crops of greatest commercial interest including maize (*Zea mays* L.), soybean [*Glycine max* (L.) Merr.], canola [*Brassica* spp.], cotton (*Gossypium hirsutum* L.), and sunflower (*Helianthus annuus* L.). This has led to the development of holistic molecular breeding strategies for variety development aimed at generating an ideal genotype based on a mosaic of favorable chromosomal segments. This has included simultaneous MAS for multiple traits (selection based on marker information only) such as yield, biotic and abiotic stress resistance, and quality attributes (Ragot et al., 2000; Eathington, 2005), several of which are polygenic in nature. Using these approaches, commercial breeding programs have reported twice the rate of genetic gain over phenotypic selection (Eathington, 2005; Crosbie et al., 2006; Ragot and Lee, 2007). Although there is very limited specific information on these successes of molecular breeding, the first commercial products of molecular breeding (rather than limited MAS) are expected to be released to the market by all the major multinational breeding companies in the very near future. The first cultivar developed through MAS by Monsanto was released to the U.S market in 2006 and it is estimated that by 2010, over 12% of the commercial crop in the United States will be derived from molecular breeding (Fraley, 2006).

Marker-assisted selection has also been used in public breeding programs for gene introgression and gene pyramiding, particularly for major gene-controlled disease resistance in primary crops but also in crops of less interest to the private sector (for a review, see Dwivedi et al., 2007). Taking wheat (*Triticum aestivum* L.) as an example, William et al. (2007) reported the extensive use of MAS in CIMMYT wheat breeding programs. Large wheat MAS programs have also been developed in Australia for around 20 genes or chromosome regions used in cultivar development (Eagles et al., 2001). During the last few years, remarkable progress in implementation of MAS strategies for cultivar development has been achieved by the MAS Wheat Consortium in the United States, including the completion of 80 MAS projects. In addition, over 300 additional backcross programs are currently attempting to incorporate 22 different disease and pest resistance genes and 21 alleles conferring favorable bread-making and pasta quality (Dubcovsky, 2004). With these and other MAS breeding programs in the public sector worldwide, it is surprising that there are still very few documented releases or registrations of new varieties resulting from MAS. Some examples available include two rice (*Oryza sativa* L.) varieties, Cadet and Jacinto, with unique cooking and processing quality traits including amylose content released in United States (Hardin, 2000). In Indonesia, two rice varieties, Angke and Conde, released possessing resistance to bacterial blight, produced 20% greater yield over IR64 (Bustamam et al., 2002). In common bean (*Phaseolus vulgaris* L.), USPT-ANT-1 was registered as an anthracnose [caused by *Colletotrichum lindemuthianum* (Sacc. and Magn.) Bri. and Cav.] resistant pinto bean line which contained the *Co-4*^{2} gene conferring resistance to all known North American races of anthracnose in the U.S. (Miklas et al., 2003). In pearl millet [*Pennisetum glaucum* (L.) R. Br.], the parental lines of the original hybrid (HHB 67) were improved for downy mildew [caused by *Sclerospora graminicola* (Sacc.) Schroet.] resistance through MAS combined with conventional backcross breeding, leading to the release in India of a new hybrid HHB 67-2 (Navarro et al., 2006).

The limited success in developing finished breeding products using MAS is further illustrated by the numbers of publications that have been generated on QTL mapping versus MAS since the discovery of the first generation of DNA markers. The term "marker-assisted selection" was first used in the literature over two decades ago (Beckmann and Soller, 1986b) in relation to potential uses. A decade later, the use of the term became increasingly associated with reports on tagging genes with molecular markers (Fig. 1). However, the first substantive article on the application of MAS in plant breeding using DNA markers is probably the one published by Concibido et al. (1996) for soybean cyst nematode (*Heterodera glycines* Ichinohe) resistance. The volume of publications on the development and to a lesser extent application of markers for assisting plant breeding has increased dramatically during the last decade. The annual number of articles containing the term "marker-assisted selection" surpassed 1000 in 2003 (Fig. 1). Although MAS has been successfully applied in cultivar development

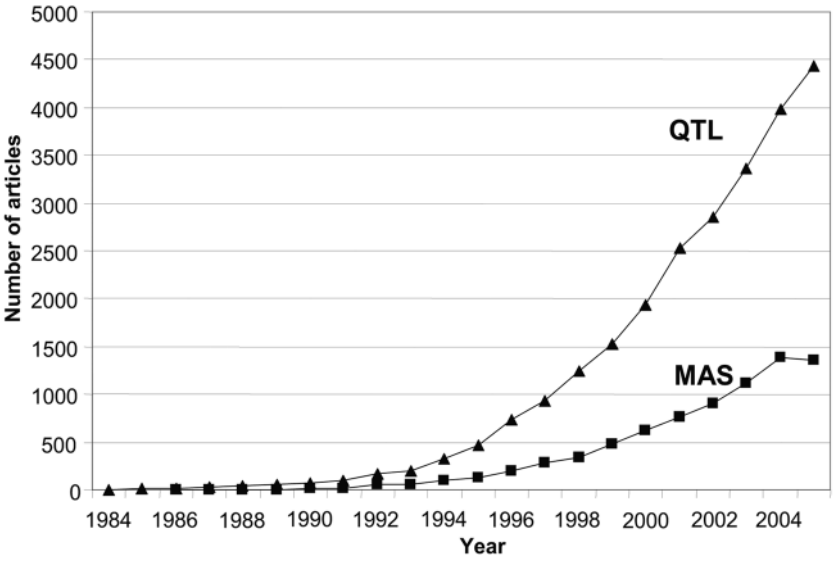

Figure 1. The numbers of articles with the terms *quantitative trait locus* or *quantitative trait loci* (QTL) and *marker-assisted selection* (MAS) by years (1984–2005) from Google Scholar (4 Aug. 2007).

in the private sector for maize (Johnson, 2004; Niebur et al., 2004; Eathington, 2005; Crosbie et al., 2006; Ragot and Lee, 2007) and soybean (Cregan et al., 1999; Cahill and Schmidt, 2004; Crosbie et al., 2006), there is limited targeted public sector funding to support the large-scale validation, refinement, and application of MAS in field breeding. This is reflected in the annual number of articles with the term "marker-assisted selection," which has consistently lagged behind the number of articles with the term "quantitative trait locus" or "quantitative trait loci" by a factor of three for the past decade, and the gap appears to be widening (Fig. 1). Moreover, it is likely that only a small proportion of papers with the words "marker-assisted selection" in their text actually report MAS applications as opposed to QTL mapping papers with discussions on the potential MAS application of their research outputs. Most articles on MAS result from either investments from donors with a scientific mandate for, or academic institutions with a specific interest in, demonstrating potential applications of MAS in plant breeding. In contrast, converting promising publications into practical large-scale applications in field breeding requires overcoming many practical, logistical, and genetical constraints. First, published markers need to be validated, in many cases, in a range of populations representative of the breeding material to be routinely screened. Next, it is necessary to develop simple, quick, and cheap technical protocols for tissue sampling, DNA extraction, genotyping, and data collection that remain reliable and precise when routinely applied in large-scale systems. Molecular breeders must also develop tailored sample and data tracking and management systems to ensure effective integration of genotyping into breeding programs. Finally, simulation analysis is required to design the optimum breeding system and powerful decision-support tools are needed to help the breeder make rapid but accurate selection decisions.

OPTIMIZING MOLECULAR MARKER SYSTEMS

Evolution of Molecular Markers

Various types of molecular marker technologies have been developed since the emergence of RFLPs in the 1980s (Phillips and Vasil, 2001). The most recent generation of molecular markers is based on direct analysis of sequence variation in each assay rather than indirect analysis using probes (RFLP) or primers (polymerase chain reaction [PCR]-based markers). Single base changes in the sequence, called single nucleotide polymorphisms (SNPs), are the most abundant source of variation in plant and animal genomes—over 31 million in humans have been databased at the time of writing and nearly 4 million in rice (http://www.ncbi.nlm.nih.gov/SNP/snp_summary.cgi). Moreover, direct sequence analysis is the most robust form of analyzing genomic variation. Thus, SNP marker analysis has many advantages over previous generations of markers including the high probability of finding a marker within the gene of interest due to the high density of SNPs across the genome (Syvänen, 2005). Although not all will be polymorphic in any given breeding population, the higher density of these markers increases the probability that at least one SNP will be polymorphic in the target gene and nearby. This provides a huge genetic advantage in MAS programs over previous markers that were at best closely linked to (but not within) the loci of interest, where this linkage could easily be lost when the marker is applied to other populations with different recombination patterns. Equally important is the ease with which SNP marker detection can be automated and thus throughput of analysis can be readily scaled up to levels appropriate for applications in plant breeding programs (Syvänen, 2005; Giancola et al., 2006). Single nucleotide polymorphism markers for candidate genes associated with virtually all target traits will become available soon in many crops following large-scale genome sequencing efforts (Lübberstedt et al., 2005). Most important perhaps, the dominance of SNP markers in human research and diagnostic applications, is driving rapid advances in SNP marker detection technologies (e.g., Bai et al., 2007), that are dramatically reducing the unit costs of detection.

Cost-Effective and High-Throughput Genotyping Systems

The current cost of DNA extraction is a rate limiting factor for many plant breeding programs, substantially inflating the overall cost per data point, especially when few assays are required on each sample. Thus, a great effort will be

needed to minimize the cost associated with each step of DNA extraction including sampling, labeling, reagents, and plastic consumables.

Polymerase chain reaction amplification is also an expensive step for all PCR-based markers. Multiplexing PCR primers can be used to significantly reduce the PCR-related cost but this requires substantial effort to optimize the protocol for suitable multiplex marker sets and is not always possible. Multiplexed PCR primers are particularly useful for genetic diversity analysis. However, for genetic mapping and MAS they often have to be optimized and even redesigned for each specific cross or population because there is no universal marker set that contains markers that are polymorphic across all crosses or populations.

Another significant cost related to MAS is the step of marker detection after PCR amplification, which varies significantly from one assay type to another. When screening PCR-based markers by agarose gel electrophoresis, which is considered more suitable for MAS of single target traits, gel preparation and electrophoresis and scoring time for a 50- to 200-sample gel can take as long as 3 to 4 h. Using microtiter plates or dot blot detection of allele-specific gene-based markers offers substantially higher throughput and lower costs than gel-based assays. However, those systems are not suitable for large-scale MAS using large numbers of markers for genetic background screens or selection of multiple target traits because the process is not scalable. Effective and efficient marker genotyping systems for large-scale MAS require a high-throughput detection system that can simultaneously deal with a large number of markers. In general, developing and optimizing such a detection system is time-consuming and expensive and requires considerable technical expertise.

Most public research and breeding institutions do not have the budget for the continuous capital investment required to maintain state-of-the-art genotyping facilities. A cost-efficient publicly available platform, including common markers, bioinformatics platforms, and analytical tools, would greatly enhance the uptake of MAS by breeding programs and help ensure access to the latest technologies. A centralized out-sourcing platform (and a series of regional service hubs) has been developed for maize through collaboration between Cornell University and CIMMYT which is now being extended to CIMMYT's partners worldwide. This will allow tropical maize scientists and breeders to have access to a platform (infrastructure and expertise) for gene-based SNP markers for both foreground selection of a target trait and background selection to rapidly recover a large proportion of the original background for accelerating the overall breeding process. This will include the development and optimization of genotyping platforms with both informative and candidate gene-specific SNPs. This public platform will encourage research and breeding groups to converge on a common set of markers, thereby facilitating greater cross-comparison and translation of results, and allowing a rapid adoption of a common set of SNP markers to rapidly replace simple sequence repeat (SSR) markers for genetic diversity analysis, mapping and molecular breeding. Most important, by establishing a cooperative platform, CIMMYT hopes to provide its stakeholders in developing countries with access to the lowest cost, highest throughput genotyping system while also assisting with data comparison, integration, and analysis. Finally, as arbitrary SNP markers in current chips are replaced by SNPs developed from within genes of interest to breeders, diversity analysis results will immediately be relevant to mapping and molecular breeding efforts. Thus, for the first time in the public sector, it will be routinely possible for anyone to effectively link diversity analysis, trait mapping, and molecular breeding through high-throughput haplotyping across the whole genome.

Effective Marker–Trait Association and Marker Validation

The volume of publications reporting the identification of new QTL has been increasing tremendously during the past two decades (as shown in Fig. 1). This now involves almost all crop plants and all types of agronomic traits (as reviewed by Dwivedi et al., 2007). However, reports of QTL mapping to date have tended to be based on individual small to moderately sized mapping populations screened with a relatively small number of markers, providing relatively low resolution of marker–trait association (MTA; Xu, 2002, 2003; Salvi and Tuberosa, 2005). Very few of the QTL reported have been used for MAS in plant breeding. Thus, it appears that the community is currently investing a large amount of time and money in generating an increasingly vast collection of publications with little impact on applied plant breeding, particularly in the public sector.

Most MTA reports to date have been based on segregating populations generated, in most cases, from two inbred lines. Genetic variation detected in the mapping population (particularly recombination patterns in the region of the target gene) may not be shared by other genetic and breeding populations because of allelic diversity. Thus, QTL markers identified using a single mapping population may not be automatically used directly in unrelated populations without marker validation and/or fine mapping (Nicholas, 2006). The MTA must be validated in representative parental lines, breeding populations, and phenotypic extremes before it can be used for routine MAS, although this process may be incorporated into genetic mapping programs. In a portion of cases, markers will lose their selective power during this validation step. In these cases, the most suitable approach is to

identify new markers (through fine mapping or candidate gene analysis) in the genomic region around the target locus to find MTAs that are shared across different breeding populations. By developing several markers within or around a single gene, it is much more likely that the parents of any breeding population will be polymorphic for at least one of them. This will then allow breeders to track the alleles donated from each parent throughout the breeding process, speeding up MAS and marker-assisted breeding (MAB) in any cross.

High-Throughput Multilocational Phenotyping

The quantity and quality of phenotyping is becoming the most significant factor affecting the accuracy of genetic mapping and thus the power of the resultant MAS, particularly for complex traits. However, precision and global phenotyping of a large number of plant samples is very expensive and time-consuming. The level of heritability of measured traits depends in part on whether the phenotyping can be repeated across different seasons, locations, and environments. Clustering target locations into mega-environments and comparing these with the success of selection at different locations has been used to understand how breeding programs can optimize their selection processes to generate germplasm with the best yield and other agronomic characters for specific target environments (e.g., Rajaram et al., 1994; Lillemo et al., 2004). Cross-population and environment comparison of phenotyping will determine how the MTAs identified under one environment can be used for selection under another. In this case, well-characterized environments and well-established selection criteria are essential prerequisites for the development of a reliable precision phenotyping system. Precision and high-throughput multilocational phenotyping, together with effective sampling and data acquisition systems being developed for many traits, provide the potential to develop a phenomics-based protocol for trait-specific breeding programs. This will not only help our understanding of the phenotypic profile a plant possesses but also improve the precision of genetic mapping and thus MAS for the target phenotype.

REDUCING COSTS AND INCREASING SCALE AND EFFICIENCY

As discussed in previous sections, highly abundant SNP-based genic markers provide great potential for increasing scale and efficiency and thus reducing the unit costs of MAS particularly because genotyping can be automated and many loci can be interrogated simultaneously. Developments of high-throughput genotyping platforms are largely driven by human and animal research and applications. Fortunately, the defining factors for cost-effective large-scale genotyping in livestock MAB and human health diagnostics will lead to important spillovers for molecular plant breeders. The feasibility of widespread uptake of marker-assisted approaches in plant breeding is heavily influenced by the relative cost (in time and money) compared with conventional breeding. Cost–benefit analysis will help us identify the bottlenecks for large-scale application and understand which components in the system need to be improved. Preliminary studies in this area have been performed in the major cereals, including maize (Dreher et al., 2003; Morris et al., 2003) and wheat (Kuchel et al., 2005). This type of analysis needs to be constantly updated as new genotyping systems become available and new optimizations are implemented in respective genotyping laboratories. Since many factors that can reduce cost may influence genetic gain, it is essential that cost–benefit analysis modules be integrated into genetic modeling and breeding simulation systems (Wang et al., 2003, 2004, 2007). In this section, we will discuss several approaches to decreasing the cost and increasing the scale and efficiency of MAS.

Single Seed–based DNA Genotyping

DNA extraction currently represents the single largest cost in most MAS pipelines and often presents the primary rate-limiting factor for scale-up of the whole process. Development and optimization of a nondestructive single seed–based DNA extraction and genotyping system will play a significant role in enhancing MAS efficiency, particularly for traits expressed late in the cropping season. Compared to MAS, using DNA extracted from leaves and other tissues, seed-based DNA for MAS has many advantages, including (i) identification of desirable genotypes and discard of undesirable genotypes before planting; (ii) increasing the speed of breeding cycles by selecting genotypes during the off-season; (iii) reducing the time-consuming and error-prone sample-collecting step that currently involves harvesting leaf tissue from plants in the field or glasshouse which then need to be retraced when the genotyping data is released; and (iv) saving land because only selected genotypes (seeds) are planted. Although DNA extraction from a single dried seed has been studied in many plant species, most reports focus on destructive protocols. Sangtong et al. (2001) developed a method for detecting a transgene and its protein product in maize endosperm that allows the kernel to be germinated after analysis. However, this method is not suitable for large-scale, low-cost, high-throughout MAS. A single seed–based DNA genotyping system that is feasible for crop species with relatively large seeds has been developed at CIMMYT for maize molecular breeding programs (Xu and Crouch, 2007; Shibin Gao, CIMMYT, personal communication, October 2007). To develop a robust and reliable system for MAS using single seed–based and

nondestructive DNA extraction, the resultant DNA must have a high quality similar to leaf-tissue DNA so as not to confound the PCR amplification and detection process. Similarly, the quantity of DNA should be sufficient for whole genome scans. Finally, the DNA extraction process must be truly high throughput, while sampled seeds should maintain a high level of germination.

Efficient Sample Tracking

A number of Laboratory Information Management Systems (LIMS) for sequencing data are available, while LIMS for genotyping data are rare with some freely available for sample tracking within the lab (e.g., Hardenbol et al., 2005; Zhao et al., 2005; Jayashree et al., 2006). Once a high-throughput system has been established for DNA extraction, the next rate-limiting factor will be the sample tracking that is required to efficiently handle a large number of plants at each step of MAS. Tracking samples from the field to the harvest bags to DNA plates for DNA extractions, PCR amplification, and marker detection and then tracing back to the field those plants selected based on the genotyping is a time-consuming and error-prone process, which also translates into a significant proportion of the overall cost for MAS. Although bar-coding systems have been widely used in the private sector for labeling and tracking samples from field to lab and from plates to databases, they are still not effective enough. As plant breeders always work with a large number of plants and populations and some crop species cannot be as easily organized in the field as others, the efficiency of the sample collecting, processing, and tracking will determine whether MAS can be processed in a high-throughput manner and thus whether MAS is practicable on a large scale.

Selective Genotyping and Pooled DNA Analysis

There have been two broad types of approaches to identifying MTA. The first is based on genotyping an entire segregating population with markers densely covering the entire genome, then testing for associations between phenotypic differences and marker genotypes. The genotyping for this approach is extensive, time-consuming, and expensive, while generating precision phenotype data at this scale may be logistically difficult or even impossible. The second approach is based on genotyping only that part of the population exhibiting extreme phenotypes for the target trait, association is then inferred by finding allelic frequency differences between the groups of plants with contrasting phenotypes (Lebowitz et al., 1987). Combining DNA pool analysis with selective genotyping (so-called bulked segregant analysis) is then the simplest and cheapest approach to identifying markers for major genes, because it requires analysis of only two DNA pools representing the two phenotypic extremes (Giovannoni et al., 1991; Michelmore et al., 1991). Pooled DNA analysis has been very successful in genetic mapping in plants using RFLP and SSR markers with numerous reports for single major genes (e.g., Barua et al., 1993; Hormaza et al., 1994; Villar et al., 1996; van Treuren 2001; Zhang et al., 2002) and even two to three major QTL (Quarrie et al., 1999). However, there have been several problems associated with most reports of pooled or bulked DNA analysis in plants: (i) insufficient marker density (e.g., 15–25 cM); (ii) low power of QTL detection due to small population sizes, resulting in phenotypic differences between pools that are sufficient only to identify large-effect genes or QTL; (iii) inaccurate estimate of allele frequencies within pools using gel-based marker systems; and (iv) a high level of false positives (MTA not actually associated with each other despite statistically significant linkage). The false positive markers have to be eliminated by a validation step involving screening the entire population with all putative markers. Southern blotting methods, though expensive and cumbersome, allow the differentiation of DNA pools with a partial difference in allele frequency at a particular locus. Other methods relying on different dyes for the alternative alleles would also allow a similar level of differentiation.

Developments in SNP genotyping technologies and methodologies recently reported in human genomics offer a vision of future possibilities for molecular plant breeding. In human research, it is possible to carry out genome-wide association mapping by using an integrated technology package including selective genotyping, pooled DNA analysis, and microarray-based SNP genotyping with 100,000 markers (Sham et al., 2002; Meaburn et al., 2006; Yang et al., 2006). This system has the power to estimate allele frequencies and identify unique alleles from a pooled DNA sample of several hundreds of individuals. When this approach can be successfully translated to plants it will resolve many of the constraints of bulked segregant analysis described above (Xu and Crouch 2007). Allele frequencies can be estimated either by collating individual genotypic scores when genotyping is based on individuals or by signal strength comparison when genotyping is based on pooled DNA samples. As high density genome-wide SNP markers will soon be available in many plant species, it is now possible to narrow down the target locus to less than 1 cM in one step using selective genotyping alone. However, optimizing SNP genotyping systems for pooled DNA analysis is considerably more complicated than for SSR markers and suffers a much higher level of redundancy. Where this has been achieved in human genomics, it required at least half a million SNPs as a starting point to finish with 100,000 optimized SNPs suitable for pooled analysis. This density of SNP markers will soon be available in rice and maize. Pooled DNA analysis will also be very useful for large-scale analysis of landraces that are highly heterogeneous and thus to date not well characterized.

With recent advances in genomics, the bottleneck in MTA analysis is increasingly the phenotyping and not the genotyping. Thus, it may be more efficient to genotype the whole population first to identify the most informative subset of individuals in terms of minimum level of relatedness between individuals plus optimum subpopulation structure and allele representativeness; then, carry out precision phenotyping of this subset, particularly for the traits that are difficult or expensive to evaluate. A final refinement to achieve maximum efficiency would be to combine selective phenotyping with selective genotyping thereby focusing on a very small proportion of the total number of genotypes for the full analysis but ensuring that these are the most informative plants or families. In this approach, the total number of individuals to be phenotyped and genotyped may not change, but the power of the analysis will be dramatically increased. This would be most effective for traits where phenotypic extremes can be easily identified using a simple screening method. For example abiotic stress tolerance, where a large number of plants or families can be eliminated easily under stress conditions through visual scoring. Although the level of stress and the selection threshold need to be carefully optimized to maximize the probability of finding genes that confer economic tolerance (with little or no yield penalty) as opposed to survival traits (with a substantial negative effect on yield). Following selective genotyping of these individuals with extreme phenotypes, precision phenotyping of the resultant subset of individuals can be performed using physiological component and surrogate traits. High-density planting and selection at early stages of plant development, combined with selective phenotyping and genotyping should also be investigated as a potential option for some traits to allow one to work with more plants or families at the same cost (see Fig. 2). Where the target trait is influenced by planting density or strong selection pressure this will clearly confound the ability to make genetic gain. However, many major gene-controlled traits can be investigated in this way without much disturbance. It can be inferred that phenotypic extremes or extremely stress-tolerant plants are those with an accumulation of favorable alleles from multiple loci, each with small to large effects, so that genetic mapping, particularly with relatively small populations, will identify the genetic regions with relatively large accumulative effect on the target trait. This is supported to some extent by results on drought tolerance in rice (Li et al., 2005). When allele frequencies can be estimated from the selected individuals or the DNA pools on which genotyping is based, the putative gene locations can be identified based on the comparison of allele frequencies, even if markers have distorted segregation ratios. However, when only one extreme (only survivors or most resistant individuals) is available, which might be true in the case that no susceptible individuals survive, or reliable estimation of allele frequencies is not possible, a control population is needed to distinguish selection effects from segregation distortion (see Fig. 2B).

Genome-wide linkage disequilibrium (LD)-based association mapping may provide a shortcut to discovering functional alleles and allelic variations that contribute to agronomic traits of interest. The Generation Challenge Program (GCP) has funded the identification of minicomposite germplasm collections in many crops. Maize is being screened via resequencing and precision phenotyping to test the feasibility of the whole "population" based approach for simply inherited traits. Thus, the selective genotyping and pooled DNA analysis discussed here can be extended to using extremes of inbreds selected from the GCP minicomposite collections, which is in principle similar to LD-based association mapping but using selected phenotypic extremes. For association mapping of quantitative traits governed by a large number of minor genes which interact with each other and the environment, through either whole population analysis or selective genotyping, we will face the same challenges as experienced with linkage-based QTL mapping.

Trait-specific genetic and breeding materials, with novel properties including phenotype extremes, eternal or fixed segregating populations (e.g., recombinant inbred lines, doubled haploids, near isogenic lines [NILs], introgression lines), genetic stocks (e.g., single segment substitution lines), and mutant libraries, have been developed and maintained by many groups across the world spanning most crops. These are valuable directly for the purpose they were developed but also offer a powerful opportunity when used collectively. In many cases, they have been phenotyped in multiple environments by taking advantage of their permanent property and fixed segregation pattern. By collecting phenotypic extremes from currently available genetic and breeding materials, and using selective genotyping and pooled DNA analysis (once it is developed), one 384 plate could be used for genetic mapping of almost all important major gene- or QTL-controlled traits. This assumes that each trait needs only two wells and that in one plate it is possible to carry out 192 pairwise DNA pool-based comparisons. When such existing materials are used collectively and combined with LD mapping (as reviewed by Mackay and Powell, 2007), they will also provide a shortcut toward one-step genome-wide association mapping for all target traits.

Integrating Diversity Analysis, Genetic Mapping, and Marker-Assisted Selection

Genetic mapping and MAS usually involve multiple consecutive steps from development of mapping populations, genetic mapping, and marker validation to MAS application. New multipurpose methodologies are emerging that facilitate the integration of genetic diversity analysis,

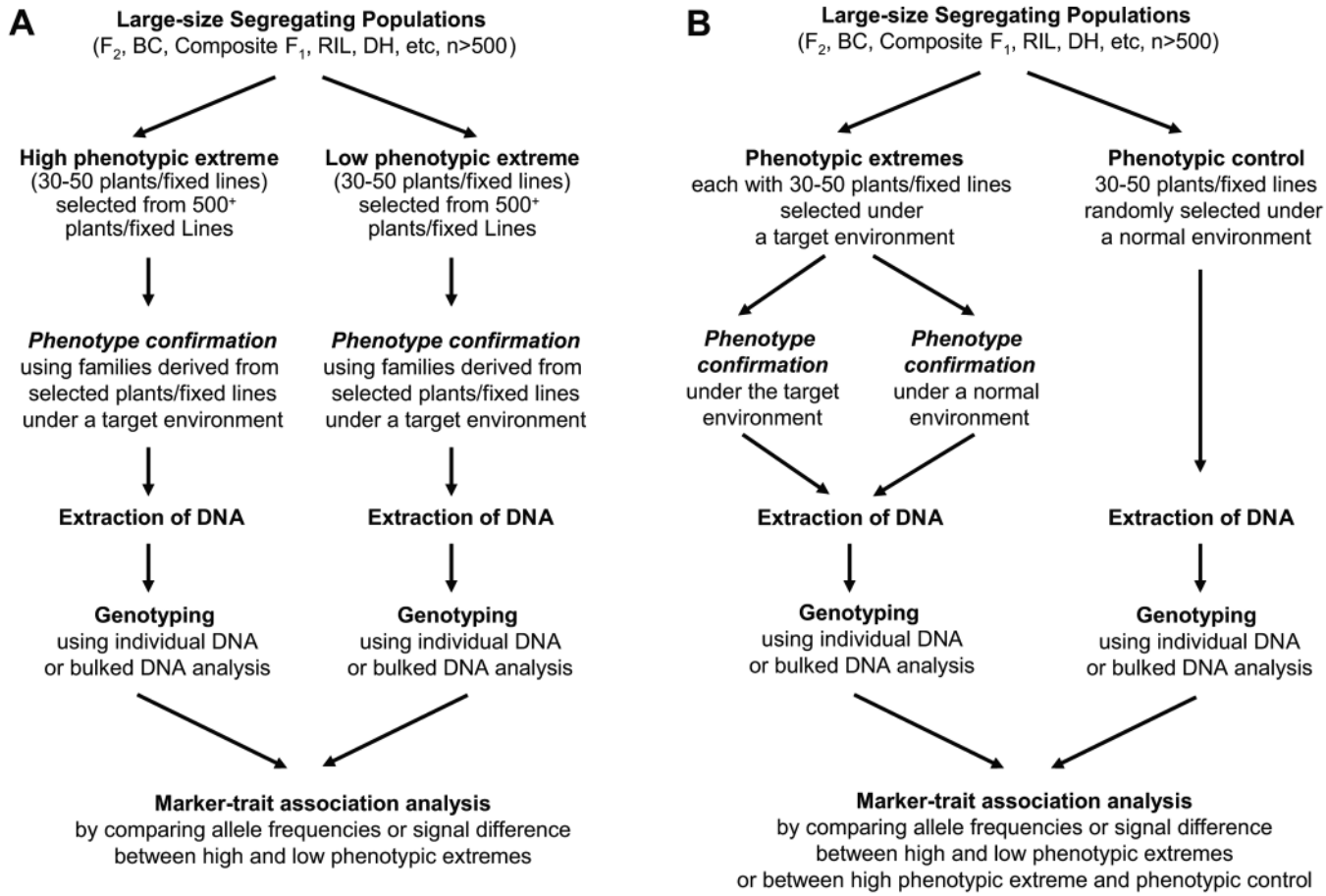

Figure 2. Flowchart for large-scale selective genotyping and genetic mapping, including selection of phenotypic extremes from large-size segregating populations, phenotype confirmation, DNA extraction, genotyping, and marker–trait association analysis. (A) A procedure for most target traits which can be scored phenotypically for all individuals or fixed lines, and then high- and low-phenotypic extremes are selected for further analysis. (B) A procedure particularly suitable for abiotic and biotic stress tolerance where only the phenotypic extreme for tolerance is available under a target environment and comparison is made between the extreme and the phenotypic control that is randomly selected from the individuals or fixed lines under a normal environment.

MTA analysis, MAS validation, and application within a single breeding program context. Achieving this integration would rely on combining multiple approaches such as LD analysis of diverse genotypes, advanced backcross mapping (Tanksley and Nelson, 1996), and "mapping-as-you-go" (MAYG) (Podlich et al., 2004). In the MAYG approach, estimates of QTL allele effects are continually revised by remapping new elite germplasm generated during cycles of MAS, thus ensuring that QTL estimates remain relevant to the current set of germplasm in the breeding program. The integration of genetic mapping and MAS offers two major advantages: (i) ability to carry out MTA analysis using breeding populations directly rather than having to follow time-consuming development of genetic populations, and, (ii) combining MTA development and validation. This saves time, both in the process itself but also in the generation of the necessary genetic materials. However, perhaps most important, the common use of end-user relevant genetic materials throughout the process is likely to dramatically reduce the level of redundancy that is commonly experienced when taking outputs from genetic studies and validating them in breeding populations.

Developing Breeding Strategies for Simultaneous Improvement of Multiple Traits

Developments required for multiple trait MAS-based improvement strategies will include understanding the correlation between different traits (including the interaction between components of a very complex trait such as drought tolerance); genetic dissection of the developmental correlation between multiple traits; understanding of genetic networks for correlated traits; and construction of selection indices across multiple traits. Much progress has already been made in this area which is relevant to drought tolerant crops, for example, in maize (Edmeades et al., 2000; Bänziger et al., 2006) and wheat (Babar et al., 2006a, 2006b). A MAS kit can be developed to include markers associated with a set of key major gene-controlled traits plus markers evenly spread over the whole genome required for marker-assisted background selection. Several thousands of well-selected SNP markers can be included

in a single chip and they can be updated and ultimately replaced by gene-based and functional markers as more and more genes are identified and functionally characterized for traits of economic importance. In theory, simultaneous selection for many traits can be completed in one step as long as the population is large enough to facilitate the identification of individuals with the targeted permutation of alleles. However, in practice the number of traits that can be manipulated in one step is limited as the population size required to provide the necessary recombinants increases exponentially with the increase in the number of traits. To manipulate numbers of multiple genes or traits that are beyond the population sizes that are amenable to routine breeding programs, a two-stage two-generation selection strategy has been proposed by Bonnett et al. (2005) and simulated by Wang et al. (2007). This approach requires that individuals are selected across all target markers for both homozygous and heterozygous forms to obtain a subset population that contains higher frequencies of the target alleles so that a much smaller population size is required in the following generation to recover individuals homozygous at the target loci.

UNDERSTANDING THE GENETIC BASIS OF COMPLEX TRAITS

There are two important genetic factors, epistasis and genotype by environment interaction (GEI), that have confounded our understanding of complex traits and the use of MAS for their improvement. Most major gene-controlled traits are less affected by these two factors while most quantitative traits are greatly affected by either one of them or both. In addition, the breeding process for complex traits, particularly with respect to hybrid performance, could be too complicated to be managed in an MAS program. As the number of genes involved and the environmental complexity increase, understanding of both epistasis and GEI and their related breeding processes may increasingly depend on computer simulation and modeling, which may be the only way in which all complex scenarios can be taken into consideration.

Epistasis

The importance of epistasis in plant breeding has been recognized for a long time (Holland, 2001); however, the extent to which quantitative traits are affected has remained controversial. QTL mapping results have revealed significantly different contributions of epistasis to genetic variation in different crop species and across different traits in the same species. Different results may be due to differences in experimental design, population structure, population size, and statistical method used in QTL mapping and genetic analysis.

In almost all cases, statistical methods currently available for epistasis detection are designed to detect two locus interactions. This is partly because including higher order interactions requires too many parameters in the genetic model, which would be difficult to estimate properly except in extremely large populations. For example, a three-locus model may need a population size of over 1000 to enable reasonably reliable estimations for all parameters. As a result, it would be very difficult to work with an epistatic model with more than three loci involved. Thus, understanding complicated epistatic effects is likely to depend on the use of specific genetic materials such as NILs to reduce the complexity of genetic effects, populations of large sample sizes, and suitable statistical methods.

Genotype by Environment Interaction

The importance of GEI, as a constraint to MAB, has been recognized because it affects both the power of QTL detection and the response to MAS. To evaluate QTL by environment interaction (QEI), precision phenotyping of multiple location or environment trials is required. Selection of suitable locations for phenotyping and accurate estimation of QTL effects across environments are two factors that determine whether the QTL identified can be effectively used for MAS. Also either through linkage mapping or LD mapping, QEI effects should also be incorporated into the statistical model for MTA analysis.

In QTL mapping, it was found that when the same mapping population was phenotypically evaluated at different environments, some QTL could be detected in all tested environments but others could be detected only in some of them (e.g., Paterson et al., 1991; Stuber et al., 1992; Lu et al., 1996). However, even in the absence of QEI, a QTL may be detected in one environment but not in others due to sampling error or experimental error. As indicated by Jansen et al. (1995), the chance for simultaneous detection of QTL in multiple environments is small. On the other hand, QEI may exist even when QTL can be detected in multiple environments (Yan et al., 1999).

To determine genetic factors responsible for GEI, QEI can be evaluated on the basis of agronomic data collected on a mapping population in multiple environment trials and comparison of QTL detection across environments by analysis of variance to test marker locus × environment interactions (Sari-Gorla et al., 1997). Quantitative trait loci by environment interaction can also be evaluated by the regression of marker genotype mean on an environmental index to discern if their linear regression coefficients are significantly different (Campbell et al., 2003). When two or more environments are involved, QEI can be estimated from a complete analysis of variance (ANOVA), $Q_i + E_j + QE_{ij}$, where significant QEIs are assessed from the significance or lack of the QE_{ij} interaction terms.

Traits that have a low heritability and are controlled by many QTL of small effects may arise from complex interaction networks which limit the possibility to detect

and clone the QTL. Marker-assisted selection can be inefficient if the effects of GEI and epistasis cannot be anticipated (Openshaw and Frascaroli, 1997; Moreau et al., 2004). When GEI and epistasis are important, we may have to regularly re-estimate QTL effects within the breeding program (Podlich et al., 2004). Bernardo and Charcosset (2006) showed that poor estimates of the allele effects within the selected population, even when genes are known, reduced MAS efficiency. As a result, both the marker set or chip developed for MAS and the genetic effects associated with the complex target trait will need to be updated as more markers become available and breeding progress provides new insights into underlying genetics.

The incorporation of environmental variables and molecular markers into statistical models facilitates the identification of the causes of GEI and therefore helps explain QEI. This allows interpreting, understanding, and exploiting QEI and also detecting the regions of the chromosomes affecting a trait that are highly influenced by external environmental conditions. These are critically important factors for complex traits such as many abiotic stress tolerances. Vargas et al. (2006) and Malosetti et al. (2004) developed factorial regression methods for mapping QTL and QEI using external environmental variables such as maximum and minimum temperatures.

To capture all the issues regarding the genetic basis of complex traits and to optimize the necessary breeding processes, computer simulation and modeling are becoming increasingly essential tools. In addition, crop modeling can be a powerful tool to resolve GEIs and to dissect complex traits into component characters that might be under simpler genetic control. Based on the complementary aspects of crop modeling and QTL mapping, for example, Yin et al. (2003) proposed an approach that integrates MAS into a model-based ideotype framework to support breeding for high crop yields. For this approach to be effective, there is a need to develop crop models that are capable of predicting yield differences among genotypes in a population under various environmental conditions.

INCREASING THE ACCESSIBILITY OF MAS TO PUBLIC SECTOR BREEDERS

The prerequisite for increasing accessibility of MAS to breeders is developing more efficiency systems that can be easily integrated into large-scale breeding programs. This is a particularly important issue for resource-limited plant breeding programs in developing countries. Several strategies can be used to establish such systems, including selection at early breeding stages to eliminate most segregants, particularly for highly heritable traits; selection at early developmental stages using high-selection pressure and an optimized selection rate, particularly for large-size plants; one-step selection for multiple traits using high-throughput genotyping; use of cost-effective genotyping systems; highly efficient phenotyping, sample tracking, and data acquisition; development and use of quick fixation and stabilization approaches; and genotyping once and phenotyping multiple times. To increase accessibility of MAS to breeders in developing countries, the most important aspect is to build skills and capacity and to develop decision support tools to help improve the efficiency of MAB programs.

Building Skills and Capacity in Developing Countries

Many additional factors will affect the application of MAS in developing countries. Building the necessary skills among national program staff and ensuring those breeding programs have direct or indirect access to sufficient genotyping capacity are essential prerequisites.

Continual improvement in the capability of laboratories to generate molecular data has come through the development of new types of markers that allow increased automation. However, this has tended to come with the negative consequence of an increase in the cost of equipment required to achieve high-throughput low-cost genotyping and in turn, the capacity to see molecular genotyping achieve impacts at the scale of modern plant breeding programs. In advanced laboratories, particularly in animal and human research, this has led to an increased tendency toward centralization including a shift to an out-sourcing mode of operation. Therefore, for molecular plant breeding, the actual genotyping might also be most efficiently and effectively performed through regional hubs and/or out-sourcing services. However, it should be realized that out-sourcing of genotyping is not a replacement for adequate training in all areas of molecular breeding. As a minimum the national programs must have facilities for high-throughput DNA extraction, precision phenotyping (controlled and field environments for trait evaluation), sample collection and tracking, plus data management and analysis.

Several crop-specific biotechnology networks have been established in Asia, Africa, and Latin America during the 1980s and 1990s. Many of these covered a wide range of activities including upstream research and capacity building. Unfortunately, in some cases major donors have pulled out from further funding of such networks. However, all these networks still present an excellent basis for the development of molecular breeding communities of practice that can be used to validate, refine, and apply new technologies in national breeding programs. Conversely in other crops, conventional breeding networks have sufficiently matured to become prime candidates for the introduction of MAS systems and other molecular breeding approaches. However, many of these breeding programs are not receiving international development assistance or are significantly underfunded, which seriously threatens their long-term impact.

Capacity building will upgrade the skills of participating plant breeders, and improve the understanding

of plant breeding and associated molecular technologies among the broader community. As many molecular techniques become sufficiently routine, there are great opportunities for scientists to profitably shift their attention to experimental design, analysis, and interpretation—as opposed to their current predominant time contribution to data generation.

Developing Molecular Breeding Decision Support Tools

The journey from the phenotyping and genotyping of individuals in genetic or breeding populations, to the identification and validation of MTAs, and onto the application of markers in molecular breeding depends on the sequential use of a number of data management, quality control, analysis, and interpretation tools that facilitate communication between genomics scientists, geneticists, bioinformaticians, trait specialists, and breeders. Decision support tools are needed for assisting germplasm evaluation, breeding population management, GEI analysis, genetic map construction, marker–trait linkage and association analysis, MAS, breeding system design and simulation, information management, and other integrated tools required for rapid and efficient decisions in molecular breeding programs (reviewed by Dwivedi et al., 2007).

A huge amount of data will be generated through large-scale MAS programs and these datasets have to be analyzed and integrated with other types of data to make selection decisions in a short time window; commonly four weeks during vegetative to flowering stages, or harvest to planting the next season. Thus, decision support tools are essential to accelerate this process while maintaining accuracy and precision. Although many tools have been developed for assisting germplasm evaluation, genetic mapping, and parts of the MAS process, they either function independently, depend on different operating systems, or require different data formats, which makes it impossible to complete a comprehensive data analysis to make the results available to breeders for decision making in a short time frame.

To address the paucity of public domain tools to assist in the application of MAS programs, the GCP has supported one such initiative (iMAS, the integrated MAS system) to integrate freely available software used for experimental design, phenotype and genotype data analysis, and the identification of MTAs (http://www.icrisat.org/gt-bt/Imas.htm). The system is specifically designed to be used with minimal hardware and uses open-source software packages or no-cost licenses to distribute the software to all users. Although iMAS includes software associated with some steps in MAS application, the next big challenge is to integrate iMAS with the International Crop Information System (ICIS; http://www.icis.cgiar.org), and molecular breeding modeling and simulation tools.

Crop informatics is very important in modern plant breeding, particularly when MAS is involved. ICIS has been developing over many years through a collaborative effort between CGIAR centers, advanced research institutions, and private sector breeding companies to resolve this challenge. ICIS has some fundamental components required for breeding programs but also has modules for genetic resources and genebanks plus increasing functionality for molecular breeding. There are several specific data management needs for public breeding programs: (i) databasing for all breeding-related information such as climate, soil, and phenotype data for selection and target environments; (ii) data mining for specific breeding purposes such as environment classification, GEI analysis, and identification of novel alleles and genetic variation; (iii) modeling breeding processes and selection schemes using multiple sources of breeding information to eliminate some field and lab tests required for making selection decision, which may be especially critical for complex traits; and (iv) extracting useful information by an integrated exploration of the information created in a specific breeding program with all related information in public databases.

Conventional plant breeding programs largely depend on phenotypic selection, breeders' experience, and knowledge of plant genetics for traits of agronomic importance. A large amount of biological data is available from genetic studies related to important target traits in crop breeding, which may in turn directly assist genotypic selection. However, gene information has not been effectively used in most breeding programs due to the fact that for most complex agronomic traits the relationship between gene information and phenotypic variation is insufficiently well defined and also due to the lack of appropriate tools for using such information even if the relationship is known. The next natural evolution of such tools is their incorporation into simulation and decision-support software that can empower breeders to easily synthesize complex multidimensional datasets to make rapid, precise, and effective breeding decisions. These computational tools will help fully integrate genomics into specific breeding programs and collectively increase the scale, efficiency, and impact of MAS. The genetics and breeding simulation tool QuLine/QuCim developed by CSIRO (the Commonwealth Scientific and Industrial Research Organization), University of Queensland and CIMMYT has the potential to use vast and varied sources of genetic information, then predict the cross performance and compare different selection methods. The best crosses and breeding strategies can then be identified. QuLine is a computer tool that is capable of defining genetic models ranging from simple to complex. Based on the results from simulation experiments, breeders may optimize their breeding programs and therefore greatly improve breeding efficiency (Wang et al., 2003). Most recently the functionality has

been expanded to include decision-support for molecular breeding programs including optimizing MAS for efficiently pyramiding multiple genes (Kuchel et al., 2005; Wang et al., 2007).

There are several practical implementation issues associated with breeders' accessibility to data management and decision support tools for MAS that should receive more attention in the future. Communications and training required to combine modeling and simulation with real breeding program data through involvement of other scientists including trait specialists, agronomists, and geneticists. Standardization and documentation of data collection for phenotypic, environmental, and genomic information needs to be enforced throughout breeding programs, particularly for developing countries. Unexpected and great variation between selection and target environments for breeding abiotic stresses may require much more comprehensive data collection, compared to the selection environments for breeding other traits. When more and more factors are involved in modeling and simulation, data generation and collection required for model tests should be performed with more data dimensions including more locations, genotypes, and replications, which in turn will determine the range and degree of the final application of results from modeling and simulation.

CONCLUSIONS AND FUTURE PROSPECTS

Marker-assisted selection has been successful for introgressing and pyramiding major-effect genes, however many challenges remain to be resolved before MAS can routinely provide added value for breeding very complex traits (Holland, 2004). The rate, scale, and scope of uptake of MAS in public crop breeding programs has continually lagged behind expectations. There are many technical and logistical factors that have hindered the speed and scope of MAS uptake. These include the unit cost and scalability of DNA extraction systems, the capital costs associated with high throughput genotyping equipment, the prolonged and labor-intensive methods for identifying MTAs, and the absence of freely available software tailored to the needs of molecular breeding programs.

The uptake of MAS in the private sector has been much more dramatic, but it continues to be dominated by transgene introgression programs and to a lesser extent backcross conversion programs for simple traits. However, there are clear signs from the multinational breeding companies that simultaneous MAS for a range of simple and complex traits plus background genome provides substantial increases in selective gain that have both time and cost advantages. It is likely that the lag in seeing products from holistic molecular breeding programs is related to the time required for the development of large-scale high throughput low unit cost SNP genotyping platforms plus sufficient gene-based markers and cloned QTL for traits of interest, and to the long product development cycle in plant breeding.

In the short term it is expected that the greatest growth in public sector MAS will be for mono- and oligogenic traits that are difficult or expensive to screen using conventional phenotyping methods. In the medium term, advances in structural genomics will provide huge amounts of sequence information while advances in functional genomics, transcriptomics, and metabolomics will help establish MTA for complex agronomic traits. This will drive the generation of large numbers of gene-based SNP markers required to facilitate a gradual shift from MAS for individual simply inherited traits to more holistic molecular breeding strategies. This transition will be hastened by new methodologies that allow the identification and validation of MTA in breeding populations. Over the next decade, MAS technologies will become substantially cheaper and easier to apply at large scale, and knowledge from genomics research will become more readily translated from publications into breeding tools and thus more routinely used in breeding systems. Perhaps most important, MAS will facilitate more efficient utilization of new genetic variation from exotic sources which will provide considerable added value.

Plants exhibit massive changes in gene expression during morpho-physiological and reproductive development as well as when exposed to a range of biotic and abiotic stresses. A new field of genetics focusing on global gene expression has emerged based on extrapolating traditional techniques of linkage and association analysis to the thousands of transcripts measured by microarrays. Dissecting the architecture of quantitative traits in this way connects DNA sequence variation with phenotypic variation, and is improving our understanding of transcriptional regulation and regulatory variation (Rockman and Kruglyak, 2006). Dynamic mapping has been proposed as a method to understand genetic expression at different developmental stages (Xu and Zhu, 1994; Xu, 1997). As more information about the dynamic properties of QTL across different stages and expression QTL revealed by whole transcriptome profiling becomes available, strategies for phenology-specific MAS and overall life cycle MAS can be developed, which will not only allow MAS at any developmental stage but also provide opportunities for breeders to optimize phenotypic selection procedures. Advances in these areas are likely to have substantial impacts on our ability to deal with the effects of GEI. However, rapid developments in these areas will continue to be constrained until high-throughput precision phenotyping systems are routinely available for evaluation of target traits in replicate multilocational trials.

Finally, the genetic basis of complex traits and the interaction between all related traits will become much

better understood. This will allow accurate modeling of gene networks and the development of robust simulation tools for designing target genomic ideotypes. With the availability of such knowledge and tools, the early stages of plant breeding programs will become much more efficient in a design-led way. However, there will continue to be no substitute for multilocational replicated evaluation trials for screening elite breeding lines for the selection and validation of finished products before distribution to local breeding companies and farmers' fields.

Acknowledgments

We acknowledge beneficial discussions with our colleagues at CIMMYT, including Drs. Marilyn Warburton, Manilal Williams, and Debra Skinner during preparation of this manuscript. We thank Dr. James Holland and two anonymous reviewers for their helpful comments and suggestions. Genomics and molecular breeding research at CIMMYT is supported by the Rockefeller Foundation, Bill and Melinda Gates Foundation, European Community, Generation Challenge Program, and United States Agency for International Development, including major collaborations with Cornell University for maize and the Molecular Plant Breeding-Cooperative Research Center for wheat.

References

Avise, J.C. 2004. Molecular markers, natural history, and evolution. 2nd ed. Sinauer Associates, Sunderland, MA.

Babar, M.A., M.P. Reynolds, M. van Ginkel, A.R. Klatt, W.R. Raun, and M.L. Stone. 2006a. Spectral reflectance to estimate genetic variation for in-season biomass, leaf chlorophyll, and canopy temperature in wheat. Crop Sci. 46:1046–1057.

Babar, M.A., M. van Ginkel, A.R. Klatt, B. Prasad, and M.P. Reynolds. 2006b. The potential of using spectral reflectance indices to estimate yield in wheat grown under reduced irrigation. Euphytica 150:155–172.

Bai, S.-L., X. Zhong, L. Ma, W. Zheng, L.-M. Fan, N. Wei, and X.W. Deng. 2007. A simple and reliable assay for detecting specific nucleotide sequences in plants using optical thin-film biosensor chips. Plant J. 49:354–366.

Bänziger, M., P.S. Setimela, D. Hodson, and B. Vivek. 2006. Breeding for improved drought tolerance in maize adapted to southern Africa. Agric. Water Manage. 80:212–224.

Barua, U.M., K.J. Chalmers, C.A. Hackett, W.T. Thomas, W. Powell, and R. Waugh. 1993. Identification of RAPD markers linked to a *Rhynchosporium secalis* resistance locus in barley using near-isogenic lines and bulked segregant analysis. Heredity 71:177–184.

Beckmann, J.S., and M. Soller. 1986a. Restriction fragment length polymorphisms in plant genetic improvement. Oxford Surv. Plant Mol. Cell Biol. 3:196–250.

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

Bernardo, R., and A. Charcosset. 2006. Usefulness of gene information in marker-assisted recurrent selection: A simulation appraisal. Crop Sci. 46:614–621.

Bernardo, R., L. Moreau, and A. Charcosset. 2006. Number and fitness of selected individuals in marker-assisted and phenotypic recurrent selection. Crop Sci. 46:1972–1980.

Bonnett, D.G., G.J. Rebetzke, and W. Spielmeyer. 2005. Strategies for efficient implementation of molecular markers in wheat breeding. Mol. Breed. 15:75–85.

Bustamam, M., R.E. Tabien, A. Suwarmo, M.C. Abalos, T.S. Kadir, I. Ona, M. Bernardo, C.M. VeraCruz, and H. Leung. 2002. Asian rice biotechnology network: Improving popular cultivars through marker-assisted backcrossing by the NARES. *In* Abstract of Int. Rice Congr., Beijing. 16–22 Sept. 2002. Int. Rice Research Inst., Manila, Philippines, and Chinese Academy of Agricultural Sciences, Beijing.

Cahill, D.J., and D.H. Schmidt. 2004. Use of marker-assisted selection in a product development breeding program. *In* T. Fischer (ed.) New directions for a diverse planet. Proc. 4th Int. Crop Sci. Congress, Brisbane, QLD, Australia. 26 Sept.–1 Oct. 2004. Available at www.cropscience.org.au/icsc2004/ (verified 4 Feb. 2008). The Regional Institute Ltd., Gosford, NSW, Australia.

Campbell, B.T., P.S. Baenziger, K.S. Gill, K.M. Eskridge, H. Budak, M. Erayman, and Y. Yen. 2003. Identification of QTLs and environmental interactions associated with agronomic traits on chromosome 3A of wheat. Crop Sci. 43:1493–1505.

CFIA/NFS. 2005. Seminar on the use of molecular techniques for plant variety protection. Available at www.inspection.gc.ca/english/plaveg/pbrpov/molece.shtml (verified 4 Feb. 2008). Canadian Food Inspection Agency/National Forum on Seed, Ottawa, ON.

Collard, B.C.Y., M.Z.Z. Jahufer, J.B. Brouwer, and E.C.R. Pang. 2005. An introduction to markers, quantitative trait loci (QTL) mapping, and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142:169–196.

Concibido, V.C., R.L. Denny, D.A. Lange, J.H. Orf, and N.D. Young. 1996. RFLP mapping and molecular marker-assisted selection of soybean cyst nematode resistance in PI 209332. Crop Sci. 36:1643–1650.

Cregan, P.B., J. Mudge, E.W. Fickus, D. Danesh, R. Denny, and N.D. Young. 1999. Two simple sequence repeat markers to select for soybean cyst nematode resistance conditioned by the *rhg1* locus. Theor. Appl. Genet. 99:811–818.

Crosbie, T.M., S.R. Eathington, G.R. Johnson, M. Edwards, R. Reiter, S. Stark, R.G. Mohanty, M. Oyervides, R.E. Buehler, A.K. Walker, R. Dobert, X. Delannay, J.C. Pershing, M.A. Hall, and K.R. Lamkey. 2006. Plant breeding: Past, present, and future. p. 3–50. *In* K.R. Lamkey and M. Lee (ed.) Plant breeding: The Arnel R. Hallauer international symposium. Blackwell, Ames, IA.

Dreher, K., M. Khairallah, J.-M. Ribaut, and M. Morris. 2003. Money matters (I): Cost of field and laboratory procedures associated with conventional and marker-assisted maize breeding at CIMMYT. Mol. Breed. 11:221–234.

Dubcovsky, J. 2004. Marker-assisted selection in public breeding programs: The wheat experience. Crop Sci. 44:1895–1898.

Dudley, J.W. 1993. Molecular markers in plant improvement: Manipulation of genes affecting quantitative traits. Crop Sci. 33:660–668.

Dwivedi, S.L., J.H. Crouch, D.J. Mackill, Y. Xu, M.W. Blair, M. Ragot, H.D. Upadhyaya, and R. Ortiz. 2007. The molecularization of public sector crop breeding: Progress, problems, and prospects. Adv. Agron. 95:163–318.

Eagles, H.A., H.S. Bariana, F.C. Ogbonnaya, G.J. Rebetzke, G.J. Hollamby, R.J. Henry, P.H. Henschke, and M. Carter. 2001. Implementation of markers in Australian wheat breeding. Aust. J. Agric. Res. 52:1349–1356.

Eathington, S.R. 2005. Practical applications of molecular technology in the development of commercial maize hybrids. *In* Proc. of the 60th Annual Corn and Sorghum Seed Res. Conf., Chicago [CD-ROM]. 7–9 Dec. 2005. Am. Seed Trade Assoc., Washington, DC.

Edmeades, G.O., M. Bänziger, and J.M. Ribaut. 2000. Maize improvement for drought-limited environments. p. 75–111. *In* M.E. Otegui and G.A. Slafer (ed.) Physiological bases for maize improvement. Food Products Press, New York.

Fraley, R. 2006. Presentation at Monsanto European Investor Day, 10 November 2006. Available at www.monsanto.com/investors/presentations.asp (verified 4 Feb. 2008).

Francia, E., G. Tacconi, C. Crosatti, D. Barabaschi, D. Bulgarelli, E. Dall'Aglio, and G. Vale. 2005. Marker assisted selection in crop plants. Plant Cell Tissue Organ Cult. 82:317–342.

Frisch, M., and A.E. Melchinger. 2001. Marker-assisted backcrossing for simultaneous introgression of two genes. Crop Sci. 41:1716–1725.

Frisch, M., and A.E. Melchinger. 2005. Selection theory for marker-assisted backcrossing. Genetics 170:909–917.

Giancola, S., H.I. McKhann, A. Bérard, C. Camilleri, S. Durand, P. Libeau, F. Roux, X. Reboud, I.G. Gut, and D. Brunel. 2006. Utilization of the three high-throughput SNP genotyping methods, the GOOD assay, Amplifluor and TaqMan, in diploid and polyploidy plants. Theor. Appl. Genet. 112:1115–1124.

Gibson, S., and C. Somerville. 1993. Isolating plant genes. Trends Biotechnol. 11:306–313.

Gimelfarb, A., and R. Lande. 1994a. Simulation of marker-assisted selection for non-additive traits. Genet. Res. 64:127–136.

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

Gimelfarb, A., and R. Lande. 1995. Marker-assisted selection and marker-QTL associations in hybrid populations. Theor. Appl. Genet. 91:522–528.

Giovannoni, J.J., R.A. Wing, M.W. Ganal, and S.D. Tanksley. 1991. Isolation of molecular markers from specific chromosomal interval using DNA pools from existing mapping populations. Nucleic Acids Res. 19:6553–6558.

Guimarães, E.P., J. Ruane, B.D. Scherf, A. Sonnino, and J.D. Dargie. 2007. Marker-assisted selection, current status, and future perspectives in crops, livestock, forestry, and fish. FAO, Rome.

Gunderson, K.L., F.J. Steemers, G. Lee, L.G. Mendoza, and M.S. Chee. 2005. A genome-wide scalable SNP genotyping assay using microarray technology. Nat. Genet. 37:549–554.

Gur, A., and D. Zamir. 2004. Unused natural variation can lift yield barriers in plant breeding. PLoS Biol. 2:e245.

Hardenbol, P., F. Yu, J. Belmont, J. MacKenzie, C. Bruckner, T. Brundage, A. Boudreau, S. Chow, J. Eberle, A. Erbilgin, M. Falkowski, R. Fitzgerald, S. Ghose, O. Iartchouk, M. Jain, G. Karlin-Neumann, X. Lu, X. Miao, B. Moore, M. Moorhead, E. Namsaraev, S. Pasternak, E. Prakash, K. Tran, Z. Wang, H.B. Jones, R.W. Davis, T.D. Willis, and R.A. Gibbs. 2005. Highly multiplexed molecular inversion probe genotyping: Over 10,000 targeted SNPs genotyped in a single tube assay. Genome Res. 15:269–275.

Hardin, B. 2000. Rice breeding gets marker assists. Available at www.ars.usda.gov/is/AR/archive/dec00/rice1200.pdf (verified 2 Feb. 2008). Agric. Res. Dec.: 11.

Heckenberger, M., J. Muminovic, J. Rouppe van der Voort, J. Peleman, M. Bohn, and A.E. Melchinger. 2006. Identification of essentially derived varieties from biparental crosses of homogenous lines: III. AFLP data from maize inbreds and comparison with SSR data. Mol. Breed. 17:111–125.

Holland, J.B. 2001. Epistasis and plant breeding. Plant Breed. Rev. 21:27–92.

Holland, J.B. 2004. Implementation of molecular markers for quantitative traits in breeding programs—Challenges and opportunities. *In* T. Fischer (ed.) New directions for a diverse planet. Proc. of the 4th Int. Crop Science Congress, Brisbane, QLD, Australia. 26 Sept.–1 Oct. 2004. Available at www.cropscience.org.au/icsc2004/ (verified 4 Feb. 2008). The Regional Institute Ltd., Gosford, NSW, Australia.

Hormaza, J.I., L. Dollo, and V.S. Polito. 1994. Identification of a RAPD marker linked to sex determination in *Pistacia vera* using bulked segregant analysis. Theor. Appl. Genet. 89:9–13.

Hospital, F. 2002. Marker-assisted backcross breeding: A case study in genotype building theory. p. 135–141. *In* M.S. Kang (ed.) Quantitative genetics, genomics, and plant breeding. CAB Int., New York and Oxon.

Hospital, F., and A. Charcosset. 1997. Marker-assisted introgression of quantitative trait loci. Genetics 147:1469–1485.

Hospital, F., I. Goldringer, and S. Openshaw. 2000. Efficient marker-based recurrent selection for multiple quantitative trait loci. Genet. Res. 75:1181–1189.

IBRD. 2006. Intellectual property rights: Designing regimes to support plant breeding in developing countries. The International Bank for Reconstruction and Development, The World Bank, Washington, DC.

Jansen, R.C., J.W. Van-Ooijen, P. Stam, C. Lister, and C. Dean. 1995. Genotype by environment interaction in genetic mapping of multiple quantitative trait loci. Theor. Appl. Genet. 91:33–37.

Jayashree, B., P.T. Reddy, Y. Leeladevi, J.H. Crouch, V. Mahalakshmi, H.K. Buhariwalla, K.E. Eshwar, E. Mace, R. Folksterma, S. Senthilvel, R.K. Varshney, K. Seetha, R. Rajalakshmi, V.P. Prasanth, S. Chandra, L. Swarupa, P. SriKalyani, and D.A. Hoisington. 2006. Laboratory information management software for genotyping workflows: Applications in high throughput crop genotyping. BMC Bioinformatics 7:383.

Johnson, G.R. 2004. Marker-assisted selection. Plant Breed. Rev. 24:293–301.

Koebner, R.M. 2004. Marker assisted selection in the cereals: The dream and the reality. p. 317–329. *In* P.K. Gupta and R.K. Varshney (ed.) Cereal genomics. Kluwer Academic, Dordrecht, the Netherlands.

Kuchel, H., G. Ye, R. Fox, and S. Jefferies. 2005. Genetic and economic analysis of a targeted marker-assisted wheat breeding strategy. Mol. Breed. 16:67–78.

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

Lebowitz, R.L., M. Soller, and J.S. Beckmann. 1987. Trait-based analysis for the detection of linkage between marker loci and quantitative trait loci in cross between inbred lines. Theor. Appl. Genet. 73:556–562.

Li, Z.K., B.Y. Fu, Y.M. Gao, J.L. Xu, J. Ali, H.R. Lafitte, Y.Z. Jiang, J.D. Rey, C.H.M. Vijayakumar, R. Maghirang, T.Q. Zheng, and L.H. Zhu. 2005. Genome-wide introgression lines and their use in genetic and molecular dissection of complex phenotypes in rice (*Oryza sativa* L.). Plant Mol. Biol. 59:33–52.

Lillemo, M., M. van Ginkel, R.M. Trethowan, E. Hernández, and S. Rajaram. 2004. Associations among international CIMMYT

bread wheat yield testing locations in high rainfall areas and their implications for wheat breeding. Crop Sci. 44:1163–1169.

Lu, C., L. Shen, Z. Tan, Y. Xu, P. He, Y. Chen, and L. Zhu. 1996. Comparative mapping of QTL for agronomic traits of rice across environments using a double haploid population. Theor. Appl. Genet. 93:1211–1217.

Lübberstedt, T., I. Zein, J.R. Andersen, G. Wenzel, B. Krützfeldt, J. Eder, M. Ouzunova, and S. Chun. 2005. Development and application of functional markers in maize. Euphytica 146:101–108.

Mackay, I., and W. Powell. 2007. Methods for linkage disequilibrium mapping in crops. Trends Plant Sci. 12:57–63.

Malosetti, M., J. Voltas, I. Romagosa, S.E. Ullrich, and F.A. van Eeuwijk. 2004. Mixed models including environmental variables for studying QTL by environment interaction. Euphytica 137:139–145.

Meaburn, E., L.M. Butcher, L.C. Schalkwyk, and R. Plomin. 2006. Genotyping pooled DNA using 100K SNP microarrays: A step towards genomewide association scans. Nucleic Acids Res. 34(4):e28.

Michelmore, R.W., I. Paran, and R.V. Kesseli. 1991. Identification of markers linked to disease resistance gene by bulked segregant analysis: A rapid method to detect markers in specific genomic regions using segregating populations. Proc. Natl. Acad. Sci. USA 88:9828–9832.

Miklas, P.N., J.D. Kelly, S.E. Beebe, and M.W. Blair. 2006. Common bean breeding for resistance against biotic and abiotic stresses: From classical to MAS breeding. Euphytica 147:105–131.

Miklas, P.N., J.D. Kelly, and S.P. Singh. 2003. Registration of anthracnose-resistant pinto bean germplasm line USPT-ANT-1. Crop Sci. 43:1889–1890.

Moreau, L., A. Charcosset, and A. Gallais. 2004. Use of trial clustering to study QTL × environment effects for grain yield and related traits in maize. Theor. Appl. Genet. 110:92–105.

Morris, M., K. Dreher, J.-M. Ribaut, and M. Khairallah. 2003. Money matters (II): Cost of maize inbred line conversion schemes at CIMMYT using conventional and marker-assisted selection. Mol. Breed. 11:235–247.

Navarro, R.L., G.S. Warrier, and C.C. Maslog. 2006. Genes are gems: Reporting agri-biotechnology—A sourcebook for journalists. Int. Crops and Research Inst. for the Semi-Arid Tropics, Patancheru, Andhra Pradesh, India.

Nicholas, F.W. 2006. Discovery, validation, and delivery of DNA markers. Aust. J. Exp. Agric. 46:155–158.

Niebur, W.S., J.A. Rafalski, O.S. Smith, and M. Cooper. 2004. Applications of genomics technologies to enhance rate of genetic progress for yield of maize within a commercial breeding program. *In* T. Fischer (ed.) New directions for a diverse planet. Proc. 4th Int. Crop Sci. Congress, Brisbane, QLD, Australia. 26 Sept.–1 Oct. 2004. Available at www.cropscience.org.au/icsc2004/ (verified 4 Feb. 2008). The Regional Institute Ltd., Gosford, NSW, Australia.

Openshaw, S., and E. Frascaroli. 1997. QTL detection and marker-assisted selection for complex traits in maize. Annu. Corn Sorghum Research Conf. Proc. 52:44–53.

Paterson, A.H. (ed.). 1998. Molecular dissection of complex traits. CRC Press, Boca Raton, FL.

Paterson, A.H., S. Damon, J.D. Hewitt, D. Zamir, H.D. Rabinowitch, S.E. Lincoln, E.C. Lander, and S.D. Tanksley. 1991. Mendelian factors underlying quantitative traits in tomato: Comparison across species, generation, and environments. Genetics 127:181–197.

Peña, L. (ed.). 2004. Methods in molecular biology. Vol. 286. Transgenic plants: Methods and protocols. Humana Press, Totowa, NJ.

Peters, J.L., F. Cnudde, and T. Gerats. 2003. Forward genetics and map-based cloning approaches. Trends Plant Sci. 8:484–491.

Phillips, R.L., and I.K. Vasil (ed.). 2001. DNA-based markers in plants. Kluwer Academic, Dordrecht, the Netherlands.

Podlich, D.W., C.R. Winkler, and M. Cooper. 2004. Mapping as you go: An effective approach for marker-assisted selection of complex traits. Crop Sci. 44:1560–1571.

Quarrie, S.A., V. Lazić-Jančić, D. Kovačević, A. Steed, and S. Pekić. 1999. Bulk segregant analysis with molecular markers and its use for improving drought resistance in maize. J. Exp. Bot. 50:1299–1306.

Ragot, M., G. Gay, J.-P. Muller, and J. Durovray. 2000. Efficient selection for the adaptation to the environment through QTL mapping and manipulation in maize. p. 128–130. *In* J.-M. Ribaut and D. Poland (ed.) Molecular approaches for the genetic improvement of cereals for stable production in water-limited environments. CIMMYT, Mexico, D.F.

Ragot, M., and M. Lee. 2007. Marker-assisted selection in maize: Current status, potential, limitations and perspectives from the private and public sectors. p. 117–150. *In* E.P. Guimarães et al. (ed.) Marker-assisted selection, current status and future perspectives in crops, livestock, forestry, and fish. FAO, Rome.

Rajaram, S., M. van Ginkel, and R.A. Fischer. 1994. CIMMYT's wheat breeding mega-environments (ME). p. 1101–1106. *In* Proc. of the 8th Int. Wheat Genetics Symp., Beijing. 20–25 July 1993. Institute of Genetics, Chinese Academy of Sciences, Beijing.

Ribaut, J.M., and D. Hoisington. 1998. Marker-assisted selection: New tools and strategies. Trends Plant Sci. 3:236–239.

Rockman, M.V., and L. Kruglyak. 2006. Genetics of global gene expression. Nat. Rev. Genet. 7:862–872.

Salvi, S., and R. Tuberosa. 2005. To clone or not to clone plant QTLs: Present and future challenges. Trends Plant Sci. 10:297–304.

Sangtong, V., E.C. Mottl, M.J. Long, M. Lee, and M.P. Scott. 2001. Serial extraction of endosperm drillings (seed): A method for detecting transgenes and proteins in single viable maize kernels. Plant Mol. Biol. Rep. 19:151–158.

Sari-Gorla, M., T. Calinski, Z. Kaczmarek, and P. Krajewski. 1997. Detection of QTL × environment interaction in maize by a least squares interval mapping method. Heredity 78:146–157.

Servin, B., O.C. Martin, M. Mézard, and F. Hospital. 2004. Toward a theory of marker-assisted gene pyramiding. Genetics 168:513–523.

Sham, P., J.S. Bader, I. Craig, M. O'Donovan, and M. Owen. 2002. DNA pooling: A tool for large-scale association studies. Nat. Rev. Genet. 3:862–871.

Stuber, C.W. 1992. Biochemical and molecular markers in plant breeding. Plant Breed. Rev. 9:37–61.

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

Syvänen, A.C. 2005. Toward genome-wide SNP genotyping. Nat. Genet. 37:S5–S10.

Tanksley, S.D. 1983. Molecular markers in plant breeding. Plant Mol. Biol. Rep. 1:1–3.

Tanksley, S.D., and S.R. McCouch. 1997. Seed banks and molecular maps: Unlocking genetic potential from the wild. Science 277:1063–1066.

Tanksley, S.D., and J.C. Nelson. 1996. Advanced backcross QTL analysis: A method for simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor. Appl. Genet. 92:191–203.

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

Tanksley, S.D., N.D. Young, A.H. Paterson, and M.W. Bonierbale. 1989. RFLP mapping in plant breeding: New tools for an old science. Biotechnology (N. Y.) 7:257–263.

van Treuren, R. 2001. Efficiency of reduced primer selectivity and bulked DNA analysis for the rapid detection of AFLP polymorphisms in a range of crop species. Euphytica 117:27–37.

Vargas, M., F. van Eeuwijk, J. Crossa, and J.-M. Ribaut. 2006. Mapping QTL and QTL × environment interaction in CIMMYT maize drought stress program using factorial regression and partial least squares method. Theor. Appl. Genet. 112:1009–1023.

Varshney, R.K., A. Graner, and M.E. Sorrells. 2005a. Genomics-assisted breeding for crop improvement. Trends Plant Sci. 10:621–630.

Varshney, R.K., A. Graner, and M.E. Sorrells. 2005b. Genic microsatellite markers in plants: Features and applications. Trends Biotechnol. 23:48–55.

Villar, M., F. Lefevre, H.D. Bradshaw, Jr., and E.T. du-Cros. 1996. Molecular genetics of rust resistance in poplars (*Melampsora larici-populina* Kleb/*Populus* sp.) by bulked segregant analysis in a 2 × 2 factorial mating design. Genetics 143:531–536.

Wang, J., S.C. Chapman, D.G. Bonnett, G.J. Rebetzke, and J. Crouch. 2007. Application of population genetic theory and simulation models to efficiently pyramid multiple genes via marker-assisted selection. Crop Sci. 47:582–588.

Wang, J., M. van Ginkel, D. Podlich, G. Ye, R. Trethowan, W. Pfeiffer, I.H. DeLacy, M. Cooper, and S. Rajaram. 2003. Comparison of two breeding strategies by computer simulation. Crop Sci. 43:1764–1773.

Wang, J., M. van Ginkel, R. Trethowan, G. Ye, I. DeLacy, D. Podlich, and M. Cooper. 2004. Simulating the effects of dominance and epistasis on selection response in the CIMMYT Wheat Breeding Program using QuCim. Crop Sci. 44:2006–2018.

Whittaker, J.C., C.S. Haley, and R. Thompson. 1997. Optimal weighting of information in marker-assisted selection. Genet. Res. 69:137–144.

William, H.M., R. Trethowan, and E.M. Crosby-Galvan. 2007. Wheat breeding assisted by markers: CIMMYT's experience. Euphytica 157:307–319.

Xu, Y. 1997. Quantitative trait loci: Separating, pyramiding, and cloning. Plant Breed. Rev. 15:85–139.

Xu, Y. 2002. Global view of QTL: Rice as a model. p. 109–134. *In* M.S. Kang (ed.) Quantitative genetics, genomics, and plant breeding. CABI Publishing, Wallingford, UK.

Xu, Y. 2003. Developing marker-assisted selection strategies for breeding hybrid rice. Plant Breed. Rev. 23:73–174.

Xu, Y., and J.H. Crouch. 2007. Marker-assisted selection in plant breeding: From publications to practice. p. 46. *In* Z.K. Li and Q.F. Zhang (eds.) From genomics to plant improvement. Proceedings of the 2nd Int. Conf. on Plant Molecular Breeding, Sanya, Hainan, China. 23–27 March 2007. Beijing Kexin Printing House, Beijing.

Xu, Y., S.R. McCouch, and Q. Zhang. 2005. How can we use genomics to improve cereals with rice as a reference genome? Plant Mol. Biol. 59:7–26.

Xu, Y., and L. Zhu. 1994. Molecular quantitative genetics. (In Chinese.) China Agriculture Press, Beijing.

Yan, J., J. Zhu, C. He, M. Benmoussa, and P. Wu. 1999. Molecular marker-assisted dissection of genotype × environment interaction for plant type traits in rice (*Oryza sativa* L.) Crop Sci. 39:538–544.

Yang, H.-C., Y.-J. Liang, M.-C. Huang, L.-H. Li, C.-H. Lin, J.-Y. Wu, Y.-T. Chen, and C.S.J. Fann. 2006. A genome-wide study of preferential amplification/hybridization in microarray-based pooled DNA experiments. Nucleic Acids Res. 34:e106.

Yin, X., P. Stam, M.J. Kropff, and A.H.C.M. Schapendonk. 2003. Crop modeling, QTL mapping, and their complementary role in plant breeding. Agron. J. 95:90–98.

Young, N.D., and S.D. Tanksley. 1989. Restriction fragment length polymorphism maps and the concept of graphical genotypes. Theor. Appl. Genet. 77:95–101.

Zhang, J., Y. Xu, X. Wu, and L. Zhu. 2002. A bentazon and sulfonylurea sensitive mutant: Breeding, genetics, and potential application in seed production of hybrid rice. Theor. Appl. Genet. 105:16–22.

Zhang, W., and C. Smith. 1992. Computer simulation of marker-assisted selection utilizing linkage disequilibrium. Theor. Appl. Genet. 83:813–820.

Zhang, W., and C. Smith. 1993. Simulation of marker-assisted selection utilizing linkage disequilibrium: The effects of several additional factors. Theor. Appl. Genet. 86:492–496.

Zhao, L.-J., M.-X. Li, Y.-F. Guo, F.-H. Xu, J.-L. Li, and H.-W. Deng. 2005. SNPP: Automating large-scale SNP genotyping data management. Bioinformatics 21:266–268.