Foundational aspects of 21st century crop improvement: Biotechnology, genomic applications, and traditional breeding practices

Rita H. Mumm[1ab,2] and Stephen P. Moose[1abc]
[1]*University of Illinois at Urbana-Champaign, 1102 South Goodwin Avenue, Urbana, Illinois, 61801, USA;* [a]*Illinois Plant Breeding Center,* [b]*Department of Crop Sciences,* [c]*Energy Biosciences Institute*
[2]*GeneMax Services, PO Box 727, Savoy, Illinois, 61874 USA*
E-mail: ritamumm@illinois.edu, rita.mumm@genemaxservices.com

Introduction

The sizeable investments in genomic studies at the DNA, RNA, protein, and metabolic pathway levels must ultimately provide return in terms of more efficiency in developing improved crop hybrids and varieties. The particular genomic application of molecular markers can be considered a 'window into the genotype'. By incorporating molecular marker information with phenotypic information as the basis for selection decisions, increases in rate of genetic gain and greater efficiencies in a plant breeding program can be achieved. A key issue is the integration of the technology into the breeding program as a vital component, and its focused application to both overall and specific objectives (Bernardo, 2008; Collard and Mackill, 2008; Xu and Crouch, 2008; Eathington *et al.*, 2007). The technology must be considered as another tool available to the breeder to meet his/her goals for plant improvement.

Philosophically, plant breeding is simple: cross the 'best' parents, and identify and recover the progeny that outperform the parents. The challenges arise in defining which are the best parental lines and in identifying the truly superior progeny. Choosing appropriate parents and selecting outstanding progeny are hampered by the environmental noise that is present in field observations for most traits controlled by many genes, challenges in screening for and accurately measuring certain traits of interest, and a lack of the knowledge about the underlying metabolics and genetics for many traits of interest.

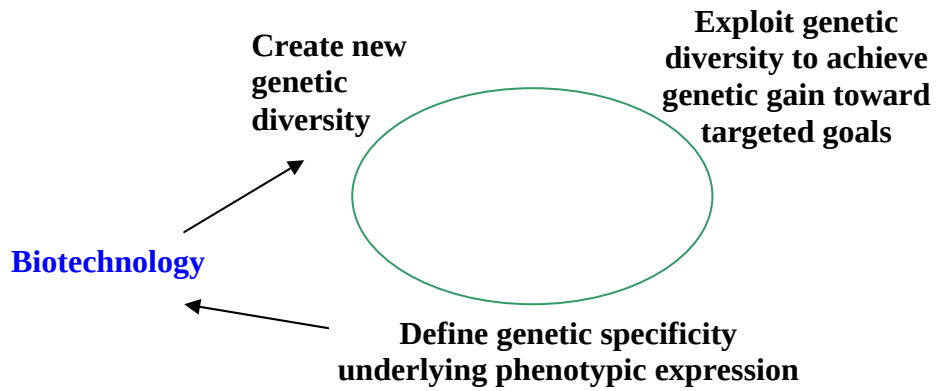
Molecular marker applications in breeding are currently focused in four primary areas: line characterization, marker-assisted selection (MAS), marker-assisted backcrossing (MABC), and gene discovery. Altogether, these can lead to a productive cycle which leverages these applications and provides further value beyond immediate breeding objectives (see Figure 1). Biotechnology can also be an integral component of such a cycle, with the creation and deployment of novel traits or modified expression of existing traits through transformation. Transgenic events become a means to expand genetic diversity. Furthermore, transformation can be used to investigate and confirm genetic specificity via over- and under-expression studies.

Creating genetic diversity

Useful genetic diversity is traditionally produced by crossing parents to create a segregating population. Choice of parents is key to achieving targeted breeding objectives. One or more parents must provide a source of favorable alleles for each of the traits of interest, and mean performance must also be a major consideration for improvement of quantitative traits.

Molecular markers can be utilized to profile prospective parent lines (i.e. proprietary commercial lines in an industrial setting) for a standard set of markers. This information is commonly used to protect intellectual property represented by the line. Furthermore, this information can be used to predict performance of progeny from particular crosses with various combinations of chromosomal segments based on their marker profiles (Eathington *et al.*, 2007; Johnson and Mumm, 1996). Effects of each marker can be estimated using phenotypic data amassed from performance evaluations, taking into account relatedness among lines for greatest precision (e.g. Bernardo, 2008; Collard *et al.*, 2008).

Figure 1. The cycle of creating and exploiting genetic diversity involving use of molecular markers and biotechnology in crop improvement.



Biotechnology can be utilized to introduce genetic diversity that often extends beyond species boundaries (Johnson and McCuddin, 2008; Gepts, 2002). Biotechnology enables access to genes heretofore not available through crossing and creates an essentially infinite pool of novel genetic variation. Genes may be acquired from existing genomes spanning all kingdoms of life, or designed and assembled *de novo* in the laboratory. From simple to sophisticated, the power of transgenes to introduce novel phenotypic variation can be found in the following examples of three different transgenes developed for resistance to glyphosate herbicides in maize and other crops. The first glyphosate-tolerant maize hybrids incorporated a modified version of the endogenous maize gene encoding 5-enol-pyruvylshikimate-3-phosphate synthase (EPSPS) (Spencer *et al.*, 2000), which was followed later by events produced with an EPSPS gene isolated from *Agrobacterium* (Behr *et al.*, 2004). More recently, a synthetic gene with enhanced glyphosate acetyltransferase activity was created via gene shuffling and selection in a microbial system (Castle *et al.*, 2004). These glyphosate-tolerant maize events also illustrate another benefit of biotechnology, tailored expression of the specific trait, where new combinations of regulatory sequences (*e.g.* the CaMV 35S and rice *actin1* promoters) may be used to achieve optimal trait expression with respect to overall activity and tissue distribution relative to what might be possible with endogenous genes (Heck *et al.*, 2005).

The first wave of transgenic traits focused largely on herbicide tolerances and resistances to pests, primarily insect pests, many of which represent the production of novel proteins in the plant (USDA, 2008; Castle *et al.*, 2006). The next generation of transgenic traits includes enhancements of traits basic to cultivar performance which are typically quantitative in nature *e.g*. increased yield, yield stability, tolerances to biotic and abiotic stresses, grain quality/composition (USDA, 2008). With this next generation of transgenic traits, the likelihood of the transgenic event interacting with endogenous genes associated with metabolic pathways influencing phenotypic expression for the trait of interest in one or more genetic backgrounds is significantly increased. This may necessitate new breeding approaches to line conversion that replace or modify traditional backcrossing schemes for event integration (Mumm, 2007).

Exploiting genetic diversity

Quantitative genetic principles have been particularly powerful as the theoretical basis for both population improvement and methods of selecting and stabilizing desirable genotypes (Hallauer, 2007). A foundational concept in genetic improvement is genetic gain (ΔG) (Falconer and Mackey, 1996), defined as the change in mean value of a trait within a population that occurs with selection:

$$\Delta G = h^2 \sigma_p i / L$$

where h^2 (heritability) is the probability that a trait phenotype will be transmitted from parent to offspring, σ_p is the degree of phenotypic variation present in the population expressed as the standard deviation of phenotypic values in a normal distribution, i (selection intensity) is the proportion of the population selected as parents for the next generation expressed in units of standard deviation from the mean, and L is the length of time necessary to complete a cycle. This concept provides a framework for comparing the expected effectiveness of particular breeding strategies and is often used as a guide in the judicious allocation of resources for achieving breeding objectives (Moose and Mumm, 2008).

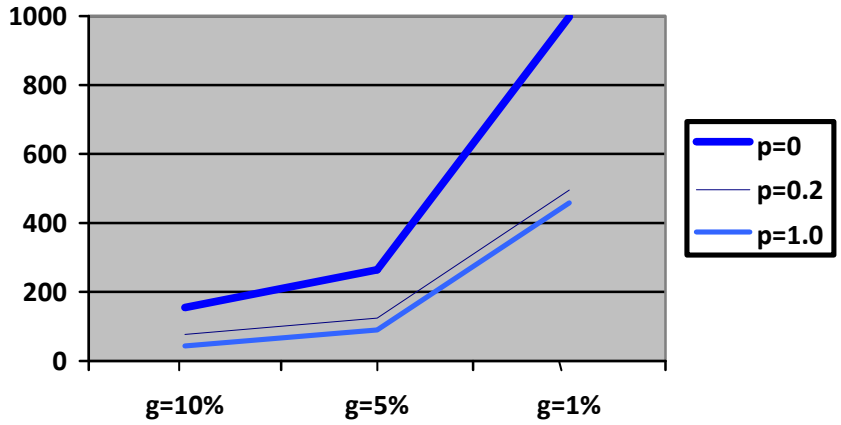
Use of molecular markers can increase breeding efficiency and genetic gains from selection relative to phenotypic selection alone (Knapp, 1998; Eathington *et al.*, 1997; Mumm, 1993; Lande and Thompson, 1990). When markers associated with QTL (quantitative trait locus/loci) for a trait of interest account for a greater proportion of the additive genetic variation that phenotype alone, the marker information has the effect of increasing heritability. The increase in breeding efficiency can be realized in a number of ways:

- By increasing the probability of identifying truly superior genotypes
- By increasing the amount of gain realized through selection *e.g.* 1000 kilograms per hectare yield increase versus 500 kilograms per hectare yield increase
- By decreasing the number of progeny that must be screened to recover a given level of gain
- By increasing the number of new cultivars commercialized/released in a given timeframe.

The interrelationship of these factors is clearly illustrated by Knapp (1998), some of which is presented in Figure 2. These data show that, when selection is based on both marker and phenotypic data, the number of progeny that must be evaluated to recover with 99% certainty at least 1 with a genotypic value in the 10%, 5%, or 1% upper tail of the phenotypic distribution is about half the number that must be evaluated when selection is based on phenotype alone, even when markers for QTL explain only about 0.2 of the proportion of genetic variation. Given population sizes of ≥200 individuals to ensure accuracy in

estimating marker effects with non-inbred populations, extrapolation of the data indicates that evaluation of a segregating population of this approximate size would provide a very high likelihood (99%) of recovering at least one line in the ~4% upper tail. In contrast, Knapp demonstrated that population sizes of 50-75 provide only an 80% probability of recovering at least 1 progeny in the upper 10% of the distribution without the use of MAS, which explains the high frequency of no superior recoveries from most breeding populations with typical conventional breeding approaches (data not shown). The table also shows the diminishing return of the marker data when the markers account for a very large proportion of the genetic variation.

Figure 2. The number of progeny that must be evaluated for a 99% certainty of recovering ≥1 progeny with a genotypic value in the *g*th percentage of the upper tail of the phenotypic distribution of the population, given 10% selection intensity and heritability = 0.2 for the trait under selection, for various values of *p* where *p* is the proportion of the genetic variation explained by markers for associated QTL (results from Knapp, 1998).



Knapp (1998) suggested theoretical increases in breeding efficiency of as much as 16.7-fold with the use of MAS but the few published reports reveal more modest increases of 2-3 fold. Monsanto Company has reported an average 2.36-fold increase in genetic gain across 248 unique maize breeding populations applying MAS for improvement of multiple quantitative traits compared to selection based on conventional selection for phenotype alone (Eathington *et al.*, 2007). Likewise, Green Giant has reported realized gains in breeding efficiency for improved agronomic and quality traits in sweet corn (Johnson, 2002).

Increase in breeding efficiency through use of molecular markers can also be realized in other ways:

- By decreasing the time over which gain is realized (as is particularly the case with MABC)
- By facilitating focus of testing resources on genotypes with greatest potential (*i.e.* early elimination of lesser genotypes).

Moreover, use of molecular markers can further enhance breeding efficiency by offering

solutions to particular challenges (Collard *et al.*, 2008; Johnson, 2004; Jansen and Nap, 2001):

- By enabling simultaneous improvement for negatively correlated traits
- By providing a means to effectively screen for difficult- or costly-to-measure traits,
- By providing a means to effectively screen for highly heritable metabolites strongly correlated with the target trait phenotype, particularly for traits defined by responses to environmental, developmental, or physiological cues
- By providing a practical means by which to evaluate and utilize exotic or elite unadapted germplasm
- By providing a starting point to investigate genetics underlying a trait of interest and for ultimately tagging genes with intrinsic markers.

Defining genetic specificity

QTLs for traits of interest can be mined to identify the genes underlying phenotypic expression for added value beyond the immediate breeding application. There are numerous examples of cloning of genes pursuant to fine mapping (Price, 2006; Varshney *et al.*, 2006) and candidate gene approaches (Hansen *et al*., 2008; Salvi and Tuberosa, 2005; Varshney *et al.*, 2005). The cloning of QTLs has yielded novel insights about the biology of quantitative traits that were not likely to be discovered from the analysis of gene knockouts and overexpression strategies, especially the impacts of regulatory variation involved in phenotypic expression and evolution (e.g. Clark *et al.*, 2006; Yan *et al.*, 2004; Cong *et al.*, 2002). Furthermore, molecular markers, genomics, and biotechnology are now applied in an iterative fashion to create new genetic diversity for crop improvement (Figure 1). The discovery of beneficial alleles and regulatory elements via QTL mapping and cloning and the use of information learned from the molecular characterization of QTL can be utilized to design optimal transgenic strategies and used to generate other novel biotechnology for deployment across species.

Implications

With or without the use of molecular markers in breeding, cultivar improvement is still a 'numbers game'. Recovering a superior genotype remains a matter of favorably shifting probabilities toward that result through appropriate population sizes, use of strategies and tools that promotes greater genetic gains, and prudent choice of parents. MAS can improve efficiencies but must be broadly implemented across a majority of the populations worked to improve the overall efficiency in a breeding program.

Integration of genomic applications like molecular markers is key; likewise, optimization of the use of such tools is important to maximize value of the technology. Monsanto has been quite candid about steps it has taken to integrate and optimize marker utility in its maize breeding program (Eathington *et al.*, 2007; Fraley, 2006), citing several factors including: 1) structuring of the breeding program to create specific groups focused on choice of parents, progeny performance evaluation, technology and data development, and breeder education/support in deployment of new tools; 2) increased yield trial capacity and quality e.g. equipment upgrade, uniform plant spacing, optimal planting dates, 3) upgrades to the genotyping platform, 4) establishing IT system infrastructure with a centralized database featuring uniformity across crops for pedigree nomenclature, trait definitions, and rating scales, as well software to support data collection/analysis and selection decisions to empower breeders; 5) methodology for marker utility incorporated with breeding strategies

and schemes e.g. indices for multiple trait selection, recurrent selection scheme; and 6) positioning for greater predictive use of the database information as it grows with molecular marker implementation.

The use of molecular markers has been the first genomic application to be adopted on a widespread basis in cultivar improvement, but likely will not be the last. The lessons learned in the implementation and maximization of this technology will pave the way for others to come.

Also key to the evolution of all foundational aspects of crop improvement i.e. biotechnology, genomic applications, and field-based breeding practices, is the intensified and on-going integration of the research disciplines and activities that form the core components of 21st crop improvement. These include: genetics/genomics to provide knowledge of genomic organization and the function of genes, plant biology, laboratory methods in molecular biology and translational genomics, field-based breeding practices, statistics and experimental design, and the management of large and diverse data sets. Graduate student training must facilitate active participation in the process of crop improvement if the next generation to carry the evolution forward is to be properly prepared. Since development of improved corn hybrids occurs in the private sector these days, it is imperative to have private sector engagement in providing training environments, funding, and in-kind support for education. Furthermore, the research atmosphere for graduate education must be collaborative, spanning the disciplines, demonstrating the interaction among a team of scientists focusing their collective expertise on crop improvement.

The collaborative efforts of a broad community of scientists are needed to ensure further expansion, contributions, and impact in meeting global needs for hefty and sustainable increases in maize production in the 21st century.

References

Behr C.F., Heck G.R., Hironaka C.H., You J. 2004. Corn plants comprising event PV-ZMGT32 (NK603). U.S. Patent No. 6,825,400.

Bernardo R. 2008. Molecular markers and selection for complex traits in plants: Learning from the last 20 years. Crop Science 48: 1649-1664.

Castle L.A., Siehl D.L., Gorton R., Patten P.A., Chen Y.H., Bertain S., Cho H.J., Duck N., Wong J., Kiu D., Lassner M.W. 2004. Discovery and directed evolution of a glyphosate tolerance gene. Science 304: 1151-1154.

Castle L.A., Wu G, McElroy D. 2006. Agricultural input traits: past, present, and future. Current Opinion in Biotechnology 17: 105-112.

Clark R.M., Wagner T.N., Quijada P., Doebley J. 2006. A distant upstream enhancer at the maize domestication gene tb1 has pleiotropic effects on plant and inflorescence architecture. Nature Genetics 38: 594-597.

Collard B.C.Y., Vera Cruz C.M., McNally K.L., Virk P.S., Mackill D.J. 2008. Rice molecular breeding laboratories in the genomics era: current status and future considerations. International J. of Plant Genomics 2008: 1-25.

Collard B.C.Y., Mackill D.J. 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Phil. Trans. R. Soc. B 363: 557-572.

Cong B., Liu J., Tanksley S.D. 2002. Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. Proc Natl Acad Sci USA 99: 13606-13611.

Eathington S.R., Crosbie T.M., Edwards M.D., Reiter R.S., Bull J.K. 2007. Molecular markers in a commercial breeding program. Crop Science 47(S3): S154-S163.

Eathington S.R., Dudley J.W., Rufener G.K. II. 1997. Usefulness of marker-QTL associations in early generation selection. Crop Science 37: 1686-1693.

Falconer D.S., Mackay T.F.C. 1996. Introduction to quantitative genetics. Longman A.W. (Ed.) Benjamin Cummings, Essex, UK.

Fraley R. 2006. Monsanto "Whistle Stop Tour" Investor Presentations, Corn Breeding. 31 July 2006. http://www.monsanto.com/pdf/investors/2006/07-31-06a.pdf

Gepts P. 2002. A comparison between crop domestication, classical plant breeding, and genetic engineering. Crop Science 42: 1780-1790.

Hallauer A.R. 2007. History, contribution, and future of quantitative genetics in plant breeding: lessons from maize. Crop Science 43: 1938-1944.

Hansen B.G., Halkier B.A., Kliebenstein D.J. (2008): Identifying the molecular basis of QTLs: eQTLs add a new dimension. Trends in Plant Science, 13(2), 72-77.

Heck G.R., Armstrong C.L., Astwood J.D., Behr C.F., Bookout J.T., Brown S.M., Cavato T.A., DeBoer D.L., Deng M.Y., George C., Hillyard J.R., Hironaka C.M., Howe A.M., Jaske E.H., Ledesma B.E., Lee T.C., Lirette R.P., Mangano M.L., Mutz J.N., Qi Y., Rodriguez R.E., Sidhu S.R., Silvanovich A., Stoecker M.A., Yingling R.A., You J. 2005. Development and

characterization of a CP4 EPSPS-based glyphosate-tolerant corn event. Crop Science 44: 329-339.

Jansen R.C., Nap J.P. 2001. Genetical genomics: the added value from segregation. Trends in Genetics 17: 388-391.

Johnson G.R., McCuddin Z.P. 2008. Maize and the biotech industry. In Bennetzen J., Hake S. (Eds.) The Maize Handbook, Vol. 2. Springer (in press).

Johnson G.R. 2004. Marker assisted selection. p293-309. In: Janick J. (Ed.) Plant Breeding Reviews, Long Term Selection: Maize, Vol. 24. John Wiley and Sons, Inc., Hoboken, NJ.

Johnson G.R., Mumm R.H. 1996. Marker assisted maize breeding. Proceedings of the 51st Corn and Sorghum Conference, Chicago, Illinois, USA. 11-12 December 1996. American Seed Trade Association, Washington, D.C.

Johnson L. 2002. A dozen years of marker assisted sweet corn breeding. Presented at the International Plant and Animal Genome Meeting, San Diego, California, USA, 12 January 2002.

Knapp S. 1998. Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. Crop Science 38: 1164-1174.

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

Moose S.P., Mumm R.H. 2008. Molecular plant breeding as the foundation for 21st century crop improvement. Plant Physiology 147: 969-977.

Mumm R.H. 2007. Backcross versus forward breeding in the development of transgenic maize hybrids: theory and practice. Crop Science 47(S3): S164-S171.

Price A.H. 2006. Believe it or not, QTLs are accurate! Trends in Plant Science 11(5): 213-216.

Salvi S., Tuberosa R. 2006. To clone or not to clone plant QTLs: present and future challenges. Trends in Plant Science 10(6): 297-304.

Spencer M., Mumm R., Gwyn J. 2000. Glyphosate resistant maize lines. U.S. Patent No. 6,040,497.

United States Department of Agriculture (USDA). 2008. Resource on regulated field releases in the USA and internationally; GMOs deregulated in the USA and internationally. Available at www.isb.vt.edu/ (verified 11 Feb. 2009). Virginia Polytechnic Inst. And State Univ., Blacksburg, Virginia, USA.

Varshney R.K., Hoisington D.A., Tyagi A.K. 2006. Advances in cereal genomics and applications in crop breeding. Trends in Biotechnology 24(11): 490-499.

Varshney R.K., Graner A., Sorrells M.E. 2005. Genomics-assisted breeding for crop improvement. Trends in Plant Science 10(12): 621-630.

Xu Y., Crouch J.H. 2008. Marker-assisted selection in plant breeding: from publications to practice. Crop Science 48: 391-407.

Yan L., Loukoianov A., Blechl A., Tranquilli G., Ramakrishna W., San Miguel P., Bennetzen J.L., Echenique V., Dubcovsky J. 2004. The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. Science 303: 1640-1644.