Molecular Breeding for Tropical Maize Improvement

R. Babu[1], S.K. Nair [1], M.T. Vinayan[1], P.H. Zaidi [1], B.S. Vivek [1] and B.M. Prasanna[2]

[1]International Maize and Wheat Improvement Center (CIMMYT), Hyderabad, India;
[2]CIMMYT, Nairobi, Kenya
*Corresponding author; Email: r.babu@cgiar.org

Introduction

Maize is an important source of food and nutritional security for millions of people in the developing world across most continents. Yet, one-third of all malnourished children live in locations where maize is among the top-three crops grown (Hyman et al. 2008). This reinforces the need to enhance maize productivity along with nutritional quality attributes. The growth in demand for human consumption of maize in the developing world is predicted to be 1.3 percent per annum until 2020. This coupled with the rising incomes are expected to result in a doubling of consumption of meat across the developing world (Naylor et al., 2005), leading to a predicted growth in demand for feed maize of ~3 percent per annum. Unless appropriate technological and policy interventions are introduced, the current maize harvests will likely fall short of growing demand (Shiferaw et al. 2011).

Maize yields in many of the Asian countries remain low. For example, in India, Nepal and the Philippines maize productivity is about~2 tons per hectare (t/ha), Indonesia and Vietnam produce s~3.5 t/ha, Thailand produces almost 4 t/ha, and China 5 t/ha, compared to the world average of 4.7 t/ha (in 2005) and current USA average of 9.4 t/ha (Prasanna et al. 2010). Although the maize feed market is rapidly growing in Asian countries (China, India and Indonesia), maize is still an important staple food in some countries of Asia, especially in hill areas and tribal regions of Nepal, Bhutan and India. Tropical and sub-tropical maize producers face challenges from global climate change especially in Sub-Saharan Africa and South Asia.. The maize fields in these regions are increasingly exposed to rising temperatures, more frequent droughts, excess rainfall and flooding, as well as new and evolving pathogens and insect-pests. The future of maize production, and consequently, the livelihoods of several million smallholder farmers in such climate-vulnerable regions are based to a great extent on access to climate-resilient cultivars. The latest developments in maize genomics offer exciting advances that may contribute to efficient maize breeding approaches to rapidly deliver stress-resilient and nutritionally-superior maize germplasm, especially for the South Asian region.

Modern maize breeding

Maize genetics and breeding have undergone tremendous changes in the last few years. Molecular markers and doubled haploids (DH) have emerged as two of the most powerful technologies that are revolutionizing the way homozygous lines are developed in maize breeding programs (Mayor and Bernardo, 2009; Prasanna et al. 2012; Babu et al. 2013). Molecular marker-assisted breeding (referred to commonly as molecular breeding), accelerates the pace of phenotype-based breeding in a resource-efficient manner. Molecular breeding is gaining significance as more and more marker-trait associations are discovered, validated and become available for integration into product-oriented breeding pipelines. Besides marker-assisted selection (MAS) for simply inherited traits, whole genome-based 'genomic selection' (GS) has emerged as a powerful strategy for improving polygenic, complex traits such as abiotic-stress-tolerance and micronutrient concentrations. DH technology significantly reduces the time required to obtain genetically homozygous and pure lines compared to conventional inbreeding practices. Besides maximum genetic variance and increased precision in estimating the genotypic value of DH lines, this approach permits early selection of prospective hybrids, simplifies the logistics of inbred seed increase and maintenance and allows quick-fixation of favorable alleles at quantitative trait loci (QTL) (Mayor and Bernardo, 2009; Lubberstedt and Ursula, 2012; Prasanna et al. 2012). When coupled with seed DNA-based genotyping (Gao et al., 2009), especially for large effect genomic regions, conditioning nutritional quality (e.g., crtRB1-governed beta carotene content) or disease resistance traits (e.g., msv1-driven Maize streak virus resistance), DH-based molecular breeding results in enormous time-, labor-, land- and other material resource savings.

Rapid strides in high-density genotyping, high-throughput and precision phenotyping made molecular breeding in maize an attractive proposition (Prasanna et al. 2013). Ultra-high-density genotyping such as Genotyping-by-Sequencing ([GBS]; Elshire et al. 2011) and whole genome re-sequencing along with diverse and controlled association mapping panels have enabled application of genome-wide association studies (GWAS) in maize. Through this approach, several useful marker-trait associations have been

identified in the last decade, which are excellent candidates for introgression upon appropriate validation. The success of molecular breeding programs largely depends on the ability to precisely and rapidly phenotype large numbers of individuals for targeted traits. Impressive advances have been made in abiotic stress-screening methodologies and development ofstress-screening locations throughout the globe appropriately managed by international institutions such as CIMMYT, that aid in reliably distinguishing the climate-resilient tropical germplasm from the susceptible ones. Thus, modern maize breeding programs that operate with enhanced efficiency require successful coordination among different components such as doubled haploidy, high-density genotyping, high-throughput phenotyping, informatics and decision support tools and planting in year-round nurseries.

Molecular breeding

Plant breeding in its original form was primarily based on phenotypic selection of superior individuals among segregating progenies resulting from hybridization. Although significant strides have been made, progress is limited to crop improvement through phenotypic selections for agronomically important traits through genotype x environment interactions as these technologies are expensive or unreliable. However, promising technologies with the capacity to navigate complex targeted traits (e.g. abiotic stresses) and target environments are emerging. For example, MAS involves selection of plants carrying genomic regions that are involved in the expression of traits-of-interest through molecular markers and/or selection of plants that are predicted to be superior to the rest of the population based on whole-genome marker profiles.

With the onset of relatively inexpensive and high density genotyping technologies, MAS is becoming possible irrespective of the linkage status of markers with the target genes. In general, three kinds of relationships between the markers and the target genes exist: co-location, linkage disequilibrium and linkage equilibrium (Babu et al. 2004):

1. The molecular marker is located within the gene of interest, which is an ideal situation for MAS whereby traits with genetic variance largely controlled by a single gene could ideally, be referred to as 'gene-assisted' selection. While this kind of relationship is the most preferred one however, it is difficult to find such markers and unfortunately most agronomic traits happen to be polygenic in nature. For instance, simple sequence repeat (SSR) markers have been designed for the *opaque2* allele that confers high-lysine and tryptophan content in the maize kernel. Similarly, functional markers (InDels and SNPs) have been developed for *LycE* (Lycopene Epsilon Cyclase) and *CrtRB1* (Hydroxylase) genes that govern endosperm carotenoid content in maize. This has enabled efficient means of tracking the favorable alleles of *opaque2, LycE and CrtRB1* loci in breeding for nutritionally-superior maize genotypes, since the markers are located within the gene itself and perfectly co-segregate with the target gene(s).

2. The marker is in linkage disequilibrium (LD) with the gene-of-interest throughout the population. LD is the tendency of a certain combination of alleles to be inherited together. Population-wide LD can be found when markers and genes-of-interest are physically close to each other. Selection using these markers can be called LD-MAS. Most of the markers used in the public and private plant breeding programs fall in this category and are typically located within 2 centimeters (cm) distance from the target gene.

3. The marker is in linkage equilibrium (LE) with the genes of interest throughout the population, which is the most challenging situation for applying MAS. This situation arises either due to unknown location of the target genes or because the target trait is governed by multiple minor genes that are dispersed throughout the genome. Here, MAS may serve as a potential phenotype prediction tool based on whole genome marker profiles. Though this aspect of MAS is still in infancy, preliminary results from some of the large multi-national private sector breeding programs and CIMMYT are encouraging.

Discovering marker-trait associations

Marker-trait association analysis or trait mapping can be generally defined as identifying favorable alleles at genes/genomic regions that are significantly associated with specific traits. The association can be established in several ways. Two important approaches are: a) linkage analysis using bi-parental or multi-parental populations; and b) LD analysis or association-mapping using natural populations. The LD/association mapping, popularly known as GWAS, is increasingly used in cereals like maize as a mapping strategy, because it offers several advantages compared to linkage mapping, including the time and resources saved from generating, segregating or immortal mapping populations, presence of multiple alleles in the population at higher-resolution than linkage mapping. However, there are several factors that could result in false positives in association detection, the most important of which is the population structure that can be removed through mixed-model based statistical approaches (Yu et al. 2006). Another constraint is that traits controlled by

the genes with rare alleles cannot be mapped effectively and in some cases, novel alleles sought do not exist in the population and therefore, can be only mapped using segregation of target alleles in biparental populations.A number of different approaches such as Nested Association Mapping (Yu et al. 2008), joint linkage and LD mapping (Lu et al. 2010), have been proposed to overcome some of the deficiencies of GWAS analyses.

The CIMMYT Global Maize Program carried out several GWAS analyses for a range of stress-resilient traits as well as nutritional characteristics, which resulted in the identification and subsequent validation of several marker-trait associations that are being explored for routine integration in the breeding pipeline. Salient findings of recent association mapping-based investigations for stress tolerance as well as nutritional traits include:

Drought tolerance: Water-scarcity is the leading constraint for maize production in many Asian countries. Two association panels (DTMA and CAAM), comprised of about 300 lines each, mostly from CIMMYT's tropical and subtropical germplasm, were test-crossed to CML312SR and CML451 and the hybrids evaluated for grain yield (GY) under drought-stress as well as optimal conditions across several locations in Mexico, Kenya, Zimbabwe, Thailand and India. Fifteen genomic regions were identified as consensus regions that explained 5 to 14 percent of phenotypic variance for GY under stress. The single nucleotide polymorphism (SNPs) on chromosome10 (142.6Mb) and chromosome7 (72.2Mb), had the most significant association and were located within a starch synthase and MYB family transcription-factor-related-protein-gene that has strong evidence of being associated with drought-tolerance in other species. Additionally, several 'rare alleles' that had large-positive-effects on GY under stress, were identified, mostly from LaPosta Sequia and DTP germplasm that co-located with drought candidate genes such as Annexins, ethylene insensitive-2, amino-transferases. The genomic regions identified through the GWAS analyses were subsequently validated in several bi-parental progenies such as advanced backcross-QTL populations, DH lines and recombinant inbreds derived from tolerant x elite crosses, multiple meta-QTL analyses (Almeida et al. 2012; Almeida et al. 2013, Semagn et al. 2012).

Waterlogging tolerance: Waterlogging is an emerging abiotic stress constraint that causes significant yield losses in maize grown throughout South- and Southeast Asia. Two association panels (DTMA and CAAM) were evaluated under waterlogged conditions, for a range of traits assocated with waterlogging tolerance such as brace root, surface root, percent mortality and other grain-yield-associated traits and multiple genomic regions have been identified. Subsequently, a set of recombinant inbred lines, derived from tolerant x susceptible cross, were evaluated for performance and waterlogging conditions and a set of five QTL for GY on chromosomes 1, 3, 5, 7 and 10, together, explained approximately 30 percent of phenotypic variance for GY based on per se evaluation with effects ranging from 520 to 640 kilograms per hectare (kg/ha) for individual genomic regions (Zaidi et al. 2014). Additionally, 12 QTL for secondary traits associated with waterlogging tolerance were identified, each individually explaining from 3 percent to 14 percent of phenotypic variance. In order to validate these QTL regions, a set of ~800 F2:3 families derived from eight bi-parental populations, were evaluated for their per se performance under waterlogging, drought and optimal conditions, in Hyderabad, in 2013 and 2014. The eight populations were organized into two heterotic groups and within each group, a drought donor was used as the common female parent, while the rest of the lines were waterlogging tolerant. Individual QTL mapping carried out using ~ 870 polymorphic SNPs in the eight populations revealed a total of 20 QTL for GY and other associated traits of waterlogging and drought tolerance, of which one region explained more than 10 percent of the phenotypic variance. Based on multiple criterion such as RIL-QTL overlaps with the previously reported waterlogging tolerance QTL, RIL-QTL overlaps with bi-parental family identified genomic regions, the CAAM-GWAS identified genomic regions with RIL-QTL and top-most significant associations in each of the GWAS analyses for waterlogging and drought tolerance. As a result, a set of 18 large-effect genomic regions have been identified and these are strong candidates for marker-assisted introgression in the Asia-based breeding programs.

Heat tolerance: The CIMMYT-Asia maize team initiated research on flowering-stage, heat-stress tolerance by multi-location screening of a large collection of elite inbred lines (including lines with drought tolerance developed in Mexico, Africa and Asia). This clearly revealed that the majority of the tropical maize germplasm was highly susceptible to reproductive stage heat stress. Some promising lines with tolerance to high-temperature stress and/or drought tolerance have been identified for further evaluation and utilization. Recently, a USAID-funded initiative to develop and deploy heat stress- tolerant maize for Asia (HTMA) was launched, which brings together public and private institutions. Expertise from CIMMYT, Purdue University, Pioneer Hi-Bred, NARS and SME partners in South Asia, complementary strengths in heat-stress phenotyping, advanced germplasm technologies such as doubled haploids, high-density genotyping, genome-wide association analysis as well as genomic selection and

effective seed delivery mechanisms for sustainable delivery of climate-resilient cultivars.

An association panel (HTAM) comprised of lines from all of the partners was evaluated for *per se* and test-cross-performance, under severe heat-stress conditions, at multiple hot-spot locations of South Asia. A set of 10 genomic regions were identified for heat tolerance *per se*, as well as other associated secondary traits such as tassel blasting, tassel sterility and leaf firing. *Acc1* was identified in multiple germplasm backgrounds as one of the significant regions on chromosome 2 in the HTAM panel. Besides *acc1*, *sig2B* on chromosome 1, *bzip1* on chromosome 3 and *rho6* on chromosome 5 were found to be significant for GY under heat stress. GWAS analyses for individual heat-associated symptoms such as tassel blasting, identified additional candidate genes such as *hsp26* (heat-shock protein) on chromosome 1 and lipoxygenase/grf-1 on chromosome 10, which have been selected for validation in independent experiments.

Biotic stress resistance: *Turcicum* leaf blight (TLB) caused by *Exerohilum turcicum* is among the most devastating diseases of maize crops grown in the Asian tropics. The second part of the association panels, involved around 700 inbreds /breeding lines from CIMMYT-Asia and different tropical and sub-tropical breeding programs of CIMMYT-Africa and CIMMYT-Latin America which were evaluated for TLB under artificial inoculation conditions in Mandya, India. The frequency of tolerance was relatively lower in the CIMMYT-Asia panel compared to the lines adapted to Africa and Latin America indicating the potential for untapped elite germplasm as a source of TLB resistance for the Asian tropics. A mixed-model GWAS analysis was conducted using ~900,000 SNPs generated by GBS platform, which identified 26 significant, genomic regions (P<1.00E-04) associated with TLB resistance, of which, five were common between the two panels. The proportion of variance explained, ranged from 4 percent to 10.6 percent with an effect-size that ranged from 0.24 to 1.1. One of the most important regions identified on 8.05/06 in the current study, overlapped with the known locations of *ht2*, *htn1* and many other minor QTL previously reported, thereby indicating the effectiveness of these major- and/or minor-loci and their potential deployment in Asian germplasm. Characterization of resistant haplotypes in the identified donor lines is currently underway to enable the integration of MAS for developing enhanced TLB resistance in routine breeding pipelines (Sudha Nair et al. 2014).

Provitamin A enrichment: Micronutrient malnutrition alone affects more than two billion people, mostly among low-income families in developing countries. More than 300 million people in India suffer micronutrient deficiencies and 35 percent of the world's malnourished children live in India. Though maize is principally a feed crop in Asia, there are several areas in the country (especially the tribal and hill regions) where maize is considered an important food crop. Maize cultivars that combine high-grain-yield with good amino acid composition, increased levels of pro-vitamin A and kernel-micronutrient (iron and zinc) concentrations could enhance production while improving nutrition, health and the quality of life, in areas where poverty and low-incomes limit access to diversified diets, dietary supplements or fortified foods. Yellow maize germplasm, in general, contains on an average, 1-2 parts per million (ppm) of kernel provitamin A. Under the HarvestPlus Generation Challenge Program of CGIAR, efforts focus on nutritional improvement of cereals and legumes), through an intensive program for enhancing provitamin A levels beyond 10-15 ppm in maize. This goal has been successfully achieved and is expected to make a significant impact on human nutrition. Two key genes governing carotenoid metabolism have been discovered, LcyE (lycopene epsilon cyclase) and CrtRB1 (Carotenoid β–hydroxylase) (Harjes et al. 2008; Yan et al. 2010) that promise very high-levels of provitamin A in maize endosperm. CIMMYT's recent marker validation studies (Babu et al. 2013) confirmed 2-10 fold effects of a favorable allele, identified in CrtRB1 across 26 tropical diverse genetic backgrounds, aided in developing maize genotypes by as much as 15-20 ppm of provitamin A. Provitamin A deficiency is one of the serious micronutrient malnutrition issues in several parts of South Asia and the CrtRB1-based functional markers can potentially accelerate the pace of Asian breeding programs for rapid delivery of provitamin A enriched maize.

MAS for simply inherited traits

If the quantitative variation for a given trait is controlled by fewer genes with larger effects, the breeding strategy should be to find the genes/QTLs either through standard bi-parental QTL mapping or through association- or comparative- mapping and introgress them into elite germplasm lacking them (Bernardo, 2009). Once genes/QTLs with large-effect on the target trait (explaining significant portion of phenotypic variance are identified), several marker-based methods could be applied to ensure their effective introgression into elite germplasm in a time- and cost-efficient manner.

BC-MAS and line conversion

For a long time, the introgression of qualitative traits such as pathotype-specific disease resistance or certain quality characteristics, which are typically governed by single dominant/recessive genes,

backcross breeding has been used. For transfer of a single dominant gene, a minimum of 5-6 backcrosses are required to recover 99 percent of the recurrent parent genome, which would almost double, if the target gene happens to be recessive. Rapid advances in genome research and molecular technology have led breeders to the use of DNA marker-assisted selection which holds promise in enhancing selection-efficiency and expediting the development of new cultivars with higher-yield potential (Ribaut and Hoisington 1998), especially in a back-cross context (Babu et al. 2004). As recognized by Holland (2004) and reviewed by Collard and Mackill (2008), three levels of MAS could be exercised within a marker-enabled BC breeding program: i) foreground selection for the target locus/loci using either linked markers or markers located within the target locus/loci based on the functional polymorphism information (this has significant advantages especially when the phenotypic assays are expensive and/or time-consuming/laborious), if reproductive stage traits are to be identified in the seedling stage itself and if the target gene is recessive; ii) selecting the critical recombinant through flanking markers so as to minimize the linkage drag around the target locus/loci - this would be an essential pre-requirement if the donor-parent happens to be from wild/unadapted germplasm or very different from the recipient line in terms of maturity, quality or other agronomy characteristics; iii) background or whole genome selection at non-target loci, to recover rapidly maximum recurrent parent genome (RPG) with minimum number of back-crosses (this is based on the fact that any given individual, in a particular BC generation, may possess greater or lesser RPG than the population average and hence provide an opportunity to identify desirable individuals through genome-wide marker profiling).

One of the earliest successes reported in maize (Stuber et al. 1992) using the BC-MAS approach, relates to the transfer of six-chromosomal segments each, in Tx303 and Oh43, using three-backcross generations into the target lines B73 and Mo17. The single-cross hybrids derived by crossing the 'enhanced B73' × 'enhanced Mo17' exceeded the hybrid 'checks' by 12 percent to 15 percent. This study was one of the pioneering demonstrations that marker-facilitated backcrossing can be successfully employed to improve even complex traits such as grain-yield in maize. An especially powerful BC-MAS approach for enhancing the QTL mapping in tandem with MAS was proposed by Tanksley and Nelson (1996). This provided for simultaneous discovery as well as transfer of significant, trait-enhancing, QTLs. Subsequently, Ho et al. 2002 demonstrated that the advanced backcross QTL method can be applied to identify and manipulate useful QTLs in heterotic inbreds of elite maize. More recently, BC-MAS incorporated all three features of marker-assisted selection and were reported to achieve faster line-conversion, especially for recovering QPM characteristics (Babu et al. 2005). This study reported a rapid line conversion strategy with a two-generation backcross program that employed foreground selection for the *opaque2* gene in both the backcross generations, background selection at non-target loci (only in the BC2 generation), and phenotypic selection for kernel modification and other desirable agronomic traits in two subsequent selfed generations. An early-maturing QPM hybrid, 'Vivek QPM9,' was recently developed through this process and has been released for commercial cultivation as the first MAS product in maize to be recommended for farmer adoption in India (Gupta et al. 2009). A number of studies have reported successful use of the BC-MAS approach for rapid transgene introgression, once successful transgenic events have been identified, especially in private sector maize breeding programs (Hoisington and Melchinger, 2005).

Pedigree-MAS and line development

Although BC conversion is an effective and low- cost and low-risk approach in a breeding program, it is unlikely to be as effective as pedigree breeding in generating elite hybrids that are competitive with the current generation materials, in terms of yield and other agronomic traits (Atlin et al. 2011). This is especially true in competitive modern 'hybrid maize' breeding programs, where turnover time for new lines and hybrids is fast. Pedigree breeding as practiced in large- scale breeding programs including the commercial sector, and is close to reciprocal recurrent selection (Lee and Tracy, 2009), which allows maize breeders to improve both additive- and non-additive-genetic effects, leading to greater overall genetic gains. Simply, pedigree breeding involves generating a F2 bulk from a bi-parental cross of the same heterotic group and advancing the further generations through continuous inbreeding with top-cross evaluations at early- (S2) and/or late- (S4/S5) generations. Molecular markers could be used in various ways and to varying degree within a pedigree breeding program aimed at developing lines with superior combining ability and *per se* performance. One way of using the markers for specific genes/QTL is to screen the individual F_2 plants and fix the loci with the desirable allelic constitution. This method is routinely used in CIMMYT for fixing the desirable alleles of CrtRB1, LycE and opaque2 loci in a homozygous condition, at an early (F_2/F_3) stage in the carotenoid- and quality- protein-breeding programs. Similarly, this approach could be useful in retaining desirable alleles for any of major loci conditioning disease resistance, or other quality characteristics.

Ribaut and Betran (1999) suggested single large-scale MAS (SLS-MAS) to overcome the limitation of BC-MAS breeding, which tends to result in a loss of genetic variability in non-target loci. SLS-MAS involves screening up to 3 QTL as a single-step in an early generation (F2/F3) and no pressure of selection is applied outside the targeted regions, which assured good allelic variability in the rest of the genome for future line development under various conditions and environments. An alternative strategy is to 'enrich' rather than fix alleles by selecting homozygotes and heterozygotes, for the target loci, in order to reduce the size of breeding populations required (Bonnet et al. 2005; Collard et al. 2005). Peleman and Voort (2003) proposed 'breeding by design,' a theoretical concept which aimed at controlling all allelic variations, at all loci of agronomic importance, so as to be able to design superior genotypes '*in silico,*' thereby suggesting that markers could not only be used to enhance the efficiency of the selection process, but also to aid in creating novel genotypes. However, this would require understanding and unraveling the genetic basis of quantitative agronomic traits in their entirety, which does not appear to be possible in the immediate future.

MAS for complex trait improvement

As observed by Bernardo (2008), when much of the quantitative variation is controlled by many QTL that mostly have smaller effects, the 'find-and-introgress-QTL' approach has limited applicability due to a plethora of QTLs for any given agronomic trait and/or their inconsistently estimated effects. The conventional breeding wisdom has operated a unique procedure since the 1950s, called 'recurrent selection' based on observed phenotypes which increases the frequency of favorable alleles in a population, thereby aiding in the development of 'improved source populations' from which superior recombinant inbreds were derived with higher probability of success. With the advent of molecular markers and low-cost whole-genome marker profiling techniques in the post-genomic era, marker-based selection methods are gaining momentum to increase QTL allele frequencies in a time- and resource-efficient manner. Two important marker-based selection approaches increase desirable QTL allele frequencies in a population improvement context either utilizing or not utilizing the QTL information.

F_2 enrichment and marker-assisted recurrent selection (MARS)

Both of these approaches require prior QTL identification through any of the standard mapping procedures in a representative germplasm. Markers that are either linked to the QTLs., or located within the QTLs. In F_2-enrichment, the individual F_2 plants are screened with informative markers and the unfavorable homozygotes are culled, in order to ensure that all the remaining plants are carriers of desirable alleles of QTLs (Howes et al. 1998; Bonnet et al. 2005; and Wang et al. 2007) either in homozygous or heterozygous conditions. This increased the probability of success in deriving a superior recombinant inbred with smaller populations but was not very effective, as only one generation of marker-based selection is performed in a typical F_2 enrichment exercise, with an additional round in the latter stages not being efficient either. MARS can overcome this limitation, through multiple rounds of marker-based selections made and each cycle consists of selecting the selfed progenies of each marker-selected individual and recombining them to form the next generation material. F_2 enrichment can target up to 9-to-12 unlinked QTL whereas MARS can target a larger number of marker loci (as many as 30). However, recombinant inbreds eventually developed from MARS, might not be fixed for the favorable allele at all target loci (Bernardo, 2008).

Genome-wide selection/Genomic selection

Though QTL mapping experiments successfully identified a number of small- effect, genomic regions did not translate into tangible germplasm products, especially for complex traits such as abiotic-stress-tolerance or polygenic-bioticstress-resistance (Bernardo, 2010). A more recent breakthrough is "genomic selection" (Meeuwissesn, 2001 and Hamblin et al. 2011), which takes into account genome-wide marker polymorphisms rather than specific genomic regions (or a priori QTL information), and promises to overcome many constraints associated with conventional QTL-mapping-based-MAS. The GS approach met with significant success in a number of animal breeding programs and is currently being actively researched in the plant breeding realm especially for complex/polygenic trait improvement. This approach does not require prior QTL information and focuses entirely on prediction of performance. Selections are typically performed based on GEBVs (Genomic estimated breeding values), which are calculated for each individual in the population, by fitting all the polymorphic markers as random effects in a linear model. Simulation studies demonstrated that across different numbers of QTL (20, 40, and 100) and levels of heritability, responses to genome-wide selection were 18 percent to 43 percent greater than the corresponding responses to MARS (Bernardo and Yu, 2007). Genome-wide selection is likely to be useful for complex traits, governed by numerous minor QTLs with a low heritability.

Integrating molecular breeding in tropical maize improvement: Challenges and opportunities

There are only a few reports of successful application of molecular markers in tropical maize breeding programs and one of the major factors responsible for this is the extent of LD, which refers to the combination of alleles that are inherited together. LD that is a result of the physical linkage between two loci and helps in indirect selection of target locus (QTL/gene) through a linked molecular marker. LD decay is a phenomenon which results in partial or complete loss of such marker-trait associations and is considered generally-variable between species and germplasm pools (Flint-Garcia et al. 2003; Remington et al. 2001). Tropical maize germplasm is known for its tremendous genetic diversity, which favors the discovery of causal genes with high-mapping resolution through approaches such as genome-wide association analysis. However, this diversity restricts applicability of marker-trait associations beyond the population/germplasm group in which such associations were discovered.

Generally, self-pollinated crops such as rice and wheat have larger LD blocks as compared to cross-pollinated crops including maize. In rice, LD extends up to 75 Kb in indica group, 150 Kb in tropical japonica and more than 500 Kb in temperate japonica group (Mather et al., 2007). In barley, up to 212 Kb of LD is reported in elite lines and 98 Kb in land races (Caldwell et al. 2006). Such large LD blocks favor marker-assisted selection in relatively distant germplasm. However, in maize, LD decays within 1 kb in landraces (Tenaillon et al., 2001), within 2 kb in diverse inbred lines (Remington et al., 2001) and ~500 kb in commercial elite inbred lines derived from temperate germplasm (Rafalski, 2002). Recently, large-scale analysis using multiple-association-mapping panels comprised of more than 1500 inbred lines, derived from tropical maize germplasm at CIMMYT, revealed genome-wide LD decay of ~1.4kb, which is in stark contrast to that of temperate maize germplasm (Willy et al. unpublished). Such rapid LD decay occurs even within genes. Babu et al (2013) reported dramatic effects of two-fold to 10-fold increase in beta-carotene levels, across diverse genetic backgrounds, for one of the three functional polymorphisms within the CrtRB1 locus. However, selection for the same favorable allele in an independent tropical germplasm of India, didn't enhance beta-carotene levels (Vignesh et al. 2013). This calls for not only extensive characterization of key representative tropical germplasm of Asia using high-density, genotyping/re-sequencing platforms and cataloging useful variations and associations, but also fine mapping of the large-effect causal loci (for instance, downy mildew QTL on chromosome 6 for biotic-stress-resistance as well as some of the key nutritional quality traits, so that reliable marker screening could be employed across diverse genetic backgrounds.

Globally, the field of maize genomics research has been fortunate to have received impressive investment in terms of the development of a wide-array of reverse genetic resources, high-density genotyping, whole-genome re-sequencing, hapmap construction and variety of mapping populations (Gore et al. 2009). However, most of these developments have occurred in the temperate germplasm background, leaving a vast scope of research to be carried out (and huge amount of genetic resources to be built out) of the tropical germplasm, which is far more diverse, wide-spread and possesses greater significance as both food and feed, in the developing countries of Asia, Sub-Saharan Africa and Latin America. In the Asian region, molecular breeding is gradually gaining momentum and a number of public sector institutions and universities have already built state-of-the-art facilities to facilitate integration of molecular breeding into product pipeline. In 2008, as part of a government-funded initiative, an institution of the Indian Council for Agricultural Research (ICAR) located in the northwestern hills of India (Vivekananda Parvatiya Krishi Anusandhan Sansthan [VPKAS], Almora), developed and released 'Vivek QPM-9,' a nutritionally-superior version of high-yielding, short-duration, hybrid, which incidentally, was the first MAS product to be released for commercial cultivation in the region (Gupta et al. 2009). Despite such isolated successful efforts, integration of molecular breeding tools in the routine product development pipeline, across the Asian region, remains elusive. Some of the obstacles include lack of high-throughput genotyping hubs, high-genotyping and DNA-extraction costs, long turn-around times especially for high-density genotype generation and inadequate bioinformatics capacity. For instance, a number of SNP markers have been identified in the recent past that are tightly linked to traits of interest in tropical maize germplasm, but lack of efficient service labs that could quickly and inexpensively type such markers across a large number of breeding samples and therefore, stifles their routine integration into breeding programs.

Maize molecular breeding initiatives in Asia can gain from greater synergies from institutional expertise, germplasm exchange, capacity building and continued delivery of enhanced products, andstronger interface with the international and regional organizations in the region. The importance and utility of such a synergistic network was amply demonstrated by the success of the Asian Maize Biotechnology Network (AMBIONET), an Asian Development Bank-funded and CIMMYT-led project(1998-2005), which

generated critical breeding information on tropical germplasm with respect to genetic diversity, heterotic groups, disease resistance and nutritional quality. Clearly, there is a need to bring together the active maize molecular breeding units across the Asian region, on a common platform, so that each team can benefit from the other's and the network acts as a conduit for continued supply of value-added tropical maize germplasm, developed through state-of-the -art marker technologies such as MAS or GS. This would make effective use of other complementary technologies such as doubled haploids, high-density genotyping, high-throughput and precision phenotyping, and breeding informatics.

References

Almeida G, Makumbi D, Magorokosho C, Nair SK, Borem A, Ribaut JM, Banziger M, Prasanna BM, Crossa J, and Babu R (2013) QTL mapping in three tropical maize populations reveals a set of constitutive and adaptive genomic regions for drought tolerance. Theor. App. Genet. 126 (3):583-600

Almeida G, Nair SK, Borem A, Cairns J, Trachsel S, Ribaut JM, Banziger M, Prasanna BM, Crossa J and Babu R (2014) Molecular mapping across three populations reveals a QTL hotspot region on chromosome 3 for secondary traits associated with drought tolerance in tropical maize. Mol. Breed. DOI: 10.1007/s11032-014-0068-5

Atlin GN, Palacios N, Babu R, Das B, Twumasi-Afriyie S, Friesen D, De Groote H, Vivek B and Pixley KV (2010) Quality Protein Maize: Progress and Prospects. Plant Breeding Reviews. 34: 83-131

Babu R, Nair SK, Kumar A, Venkatesh S, Sekhar JC, Singh NN, Srinivasan G, Gupta HS (2005) Two-generation marker aided backcrossing for rapid conversion of normal maize lines to Quality Protein Maize (QPM). Theor. Appl. Genet. 111: 888-897

Babu R, Nair SK, Prasanna BM and Gupta HS (2004) Integrating marker –assisted selection (MAS) in crop breeding: prospects and challenges. Curr Sci. 87(5): 607-619

Babu R, Nair SK, Vivek B, SanVicente F and Prasanna BM (2013) Integrating marker-assisted selection in DH-based breeding pipeline for rapid advancement and delivery of superior parental lines and cultivars. In: BM Prasanna, Vijay Chaikam and George Mahuku (eds). 2012. Doubled haploid (DH) Technology in Maize Breeding: Theory and Practice. Mexico, D.F.: CIMMYT

Babu R, Palacios N, Gao S, and Yan J, and Pixley K (2013) Validation of the effects of molecular marker polymorphisms in *LcyE* and *CrtRB1* on provitamin A concentrations for 26 tropical maize populations. Theor. Appl. Genet. (Accepted)

Bernardo R (2009). Should maize doubled haploids be induced among F1 or F2 plants? Theor. Appl. Genet. 119: 255–262

Bernardo R and Yu J (2007) Prospects for genome-wide selection for quantitative traits in maize. Crop Sci. 47:1082–1090

Bonnet D, Rebetzke GJ and Spielmeyer W (2005) Strategies for efficient implementation of molecular markers in wheat breeding. Mol. Breed. 15:75–85

Caldwell SK, Peter L and Powell W (2006) Extreme Population-Dependent Linkage Disequilibrium Detected in an Inbreeding Plant Species, Hordeum vulgare. Genetics. 172(1): 557–567

Collard BCY, Jahufer MZZ, Brouwer JB and Pang ECK (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica 142, 169–196

Elshire RJ, Glaubitz JC, Sun Q, Poland J a, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6:e19379. doi: 10.1371/journal.pone.0019379

Flint-Garcia SA, Thuillet A-C, Yu J, Pressoir G, Romero SM, et al. (2005) Maize association population: a high-resolution platform for quantitative trait locus dissection. Plant J 44: 1054–1064

Gao S, Martinez C, Debra J, Krivanek AF, Crouch JH, Xu Y (2009) Development of a seed DNA-based genotyping system for marker-assisted selection in maize. Mol. Breed. 22: 477–494

Gore MA, Chia JM, Elshire RJ, Sun Q, Ersoz ES, Hurwitz BL, Peiffer JA, McMullen MD, Grills GS, Ross-Ibarra J, Ware DH and Buckler ES (2009) A first-generation haplotype map of maize. Science 326:1115–1117

Gupta HS, Agrawal PK, Mahajan V et al (2009) Quality protein maize for nutritional security: rapid development of short duration hybrids through molecular marker assisted breeding. Curr Sci 96:230-237

Hamblin MT, Buckler ES and Jannik, JL (2011) Population genetics of genomics-based crop improvement methods. Trends Genet. 27: 98-106

Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, Yan J, Buckler ES (2008) Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. Science 319:330–333.

Hoisington DA and Melchinger AE (2005) Molecular Marker Systems in Plant Breeding and Crop Improvement. Biotechnology in Agriculture and Forestry, 55: 335-352.

Holland JB (2004) Implementation of molecular markers for quantitative traits in breeding programs—challenges and opportunities. In Proc. 4th Int. Crop Sci. Congress, Brisbane, Australia, 26 September—1 October.

Howes NK, Woods SM, Townley-Smith TF (1998) Simulations and practical problems of applying multiple marker assisted selection and doubled haploids to wheat breeding programs. Euphytica 100:225–230

Hyman G, Fujisaka S, Jones P, Wood S, de Vincente C, Dixon J (2008) Strategic approaches to targeting technology generation: Assessing the coincidence of poverty and drought-prone crop production. Agricultural Systems 98, 50-61

Jannink J-L, Lorenz AJ, Iwata H (2010) Genomic selection in plant breeding: from theory to practice. Briefings in Functional Genomics 9: 166–177

Lee EA and Tracy WF (2009) Modern Maize Breeding. In. JL Bennetezen, S Hake (eds). Maize Handbook Volume II: Genetics and Genomics. Springer Science + Business Media LLC, pp. 141-160.

Lu XW, Brewbaker JL, Nourse SM, Moon HG, Kim SK, Khairallah M (1999) Mapping of quantitative trait loci

conferring resistance to maize streak virus. Maydica 44: 313–318

Lu Y, Yan J, Guimares CT et al. (2009) Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. Theor. Appl. Genet. 120: 93-115

Lu Y, Zhang S, Shah T et al. (2010) Joint linkage–linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. Proc Natl Acad Sci USA 107:19585–19590.

Lubberstedt T, Ursula K (2012) Application of doubled haploids for target gene fixation in backcross programmes of Maize. Plant Breed. doi:10.1111/j.1439-0523.2011.01948.x

Mather KA, Caicedo AL, Polato NR, Olsen KM, McCouch S, Purugganan MD (2007) The extent of LD in rice. Genetics. 177(4):2223-32

Mayor P, Bernardo R (2009) Doubled haploids in commercial maize breeding: One-stage and two-stage selection versus marker-assisted recurrent selection. Maydica 54: 439–448

Meuwissen THE, Hayes BJ and Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829

Nair SK, Zerka M, Pradeep K, Babu R (2014) Genome-wide scan reveals candidate regions for resistance to Turcicum Leaf Blight in Asian tropical maize germplasm. In: Book of Abstracts: 12th Asian Maize Conference and Expert Consultation on "Maize for Food, Feed, Nutrition and Environmental Security". 30th Oct-01st Nov, 2014, Bangkok

Naylor R, Steinfeld H, Falcon W et al. (2005) Losing the links between livestock and land. Science 310: 1621-1622

Peleman JD, van der Voort, JR (2003) Breeding by design. Trends Plant Sci. 8, 330–334

Prasanna BM, Araus JL, Crossa J, Cairns JE, Palacios N, Das B, Magorokosho C (2013) High-throughput and precision phenotyping for cereal breeding programs. In: Cereal Genomics-II (eds. PK Gupta, RK Varshney). Springer-Verlag, Dordrecht, pp. 341-374.

Prasanna BM, Chaikam V, Mahuku G (2012) Doubled haploid technology in maize breeding: theory and practice. Mexico D.F.: CIMMYT, 50 pages.

Prasanna BM, Pixley KV, Warburton M, Xie C (2010) Molecular marker-assisted breeding for maize improvement in Asia. Mol. Breed. 26: 339-356.

Prasanna BM, Vasal SK, Kassahun B, Singh NN (2001) Quality protein maize. Curr. Sci. 81: 1308–1319.

Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. Current Opinion in Plant Biology 5:94-100

Remington DL, Thornsberry JM et al. (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. 98:11479–11484

Ribaut JM and Betran J (1999) Single large-scale marker-assisted selection (SLS–MAS). Mol. Breed. 5, 531–541

Ribaut JM, Hoisington DA (1998) Marker assisted selection: new tools and strategies. Trends in Pl Sci 3:236–239

Semagn K, Babu R, Hearne S and Olsen M (2014) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Mol. Breed. 33:1-14

Shiferaw B, Prasanna B, Hellin J, Banziger M (2011) Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. Food Security 3: 307-327.

Stuber CW, Lincoln SE, Wolff DW, Helentjaris T and Lander ES (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132: 823-839

Tanksley SD, Nelson JC (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor. Appl. Genet. 92: 191–203

Vignesh M, Firoz H, Nepolean T, Supradip S, Agrawal PK, Guleria SK, Prasanna BM and Gupta HS (2012) Genetic variability for kernel B-carotene and utilization of *crtRB1* 3'TE gene for biofortification in maize (*Zea mays* L.). Indian J. Genet. 72(2): 189-194

Wang J, Chapman SC, Bonnett DG, Rebetzke GJ and Crouch J (2007) Application of population genetic theory and simulation models to efficiently pyramid multiple genes via marker-assisted selection. Crop Sci. 47:582–588

Yan J, Kandianis C, Harjes CE et al. (2010) Rare genetic variation at *Zea mays crtRB1* increases beta-carotene in maize grain. Nat. Genet. 42: 322–327.

Yu J, Holland JB, McMullen MD, Buckler, ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178:539–551

Yu J, Pressoir G et al. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38:203–208

Zaidi PH, Zerka R, Vinayan MT, Almeida GD, Ramesh KP and Babu R (2014) QTL mapping of agronomic waterlogging tolerance using recombinant inbred lines derived from tropical maize germplasm. PLOS ONE. Manuscript in Review.