Chapter 21
Molecular Breeding for Quality Protein Maize (QPM)

Raman Babu and B. M. Prasanna

Abstract Maize endosperm protein is deficient in two essential amino acids, lysine and tryptophan. Several spontaneous and induced mutations that affect amino acid composition in maize have been discovered amongst which the *o2* gene (*opaque2*) has been used in association with endosperm and amino acid modifier genes for developing Quality Protein Maize (QPM), which contains almost double the amount of endosperm lysine and tryptophan as compared to the normal/non-QPM maize. These increases have been shown to have dramatic impacts on human and animal nutrition, growth and performance. A range of hard endosperm QPM germplasm has been developed at the International Maize and Wheat Improvement Center (CIMMYT) mostly through conventional breeding approaches to meet the requirements of various maize growing regions across the world. Microsatellite markers located within the *o2* gene provided opportunities for accelerating the pace of QPM conversion programs through marker-assisted selection (MAS). More recently, CIMMYT scientists are striving to develop reliable, easy-to-use markers for endosperm hardness and free amino acid content in the maize endosperm. Recent technological developments including high throughput, single seed-based DNA extraction, coupled with low-cost, high density SNP genotyping strategies, and breeder-ready markers for some key adaptive traits in maize, promise enhanced efficiency and cost effectiveness of MAS in QPM breeding programs. Here, we present a summary of QPM research and breeding with particular emphasis on genetic and molecular basis of *o2*, epistasis between *o2* and other high-lysine mutant genes, and the recent advances in genomics technologies that could potentially enhance the efficiency of molecular breeding for QPM in the near future.

R. Babu (✉)
International Maize and Wheat Improvement Center (CIMMYT), Km 45 Carretera, Mexico-Veracruz, Texcoco, 56130 Mexico D.F., Mexico
e-mail: r.babu@cgiar.org

B. M. Prasanna
CIMMYT, ICRAF House, United Nations Avenue, Gigiri, Nairobi 00621, Kenya

R. Tuberosa et al. (eds.), *Genomics of Plant Genetic Resources,*
DOI 10.1007/978-94-007-7575-6_21,

21.1 Introduction

Maize is one of the most important food crops in the world and together with rice and wheat, providing at least 30 % of the food calories to more than 4.5 billion people in 94 developing countries, which includes 900 million poor consumers. In parts of Africa and Mesoamerica, maize alone contributes over 20 % of food calories. The role of maize for human consumption, expressed in terms of the share of calories from all staple cereals, varies significantly across regions (Shiferaw et al. 2011). This ranges from 4 % in South Asia to 29 % in the Andean region to 61 % in Mesoamerica. In sub-Saharan Africa, maize is mainly a food crop accounting for 73 and 64 % of the total demand in eastern and southern Africa (ESA) and West and Central Africa (WCA). This makes maize particularly important to the poor in many developing regions of Africa, Latin America and Asia as a means of overcoming hunger and improving food security.

A typical kernel of a modern maize hybrid contains 73 % starch, 9 % protein, 4 % oil and 14 % other constituents (mainly fibre). The two major structures of the kernel, the endosperm and the germ (embryo), constitute approximately 80 and 10 % of the mature kernel dry weight, respectively. The endosperm is largely starch (approaching 90 %) and the germ contains high levels of oil (30 %) and protein (18 %). Bulk of the proteins in a mature maize kernel is in the endosperm (80 %) and the rest in the germ. While the germ protein is superior in both quantity and quality, the endosperm protein is deficient in lysine and tryptophan, and therefore, maize needs to be eaten with complementary protein sources such as legumes or animal products (Prasanna et al. 2001).

In the 1920s in a Connecticut (USA) maize field, a natural spontaneous mutation of maize with soft, opaque grains was discovered, which was eventually named as *o2 (opaque2)* (Singleton 1939). In 1964, Dr. Oliver Nelson's team at Purdue University, USA, discovered that the homozygous recessive *o2* allele had substantially higher lysine (+ 69 %) in grain endosperm compared to normal maize (Mertz et al. 1964). It was further determined that this mutation results in two-to-threefold increase in the level of two amino acids, lysine and tryptophan in comparison with normal genotype. The increased concentration of these two essential amino-acids (normally deficient in the maize grain endosperm) effectively doubles the biological value of maize protein (Bressani 1991) with the considerably profitable result that only half the amount of *o2* maize (relative to normal maize), needs to be consumed to obtain the same biologically usable protein (FAO 1992). In addition, other amino acids such as histidine, arginine, aspartic acid and glycine showed increase, while the decrease was observed for some amino acids such as glutamic acid, alanine and leucine. Decrease in leucine is considered desirable as it makes leucine–isoleucine ratio more balanced, which in turn helps to provide more tryptophan for niacin biosynthesis, and thus helps in combating pellagra.

Decades of efforts by researchers at CIMMYT led to the development of "Quality Protein Maize" (QPM) with enhanced nutritional value, especially through higher levels of lysine and tryptophan in the endosperm and better amino acid balance. In

Ethiopia, Tanzania and Uganda, randomized trials showed significantly improved height and weight of children consuming such varieties, particularly in Southern Ethiopia where the population relies heavily on maize (Gunaratna et al. 2008).

In this chapter, we describe the early efforts towards development of high-lysine maize, the biochemical and molecular bases of QPM, some successful examples of molecular marker-assisted breeding for QPM development, and suggest an integrated approach for enhancing the nutritional quality in maize, including low-cost and reliable markers for the QPM component traits, seed DNA-based genotyping, and genomic technologies.

21.1.1 Maize Protein

Maize endosperm protein is comprised of different fractions. Based on their solubility, these can be classified into albumins (water soluble), globulins (soluble in saline solution), zein or prolamine (soluble in alcohol) and glutelins (soluble in alkali). In normal maize endosperm, the average proportions of various fractions of protein are albumins 3 %, globulin 3 %, zein (prolamine) 60 % and glutelin 34 %, while the embryo protein is dominated by albumins (+ 60 %), which is superior in terms of nutritional quality. The zein in maize endosperm is low in lysine content (0.1 g/100 g of protein), which negatively affects growth of animals (Osborne and Mendel 1914). In *o2* maize, the zein fraction is markedly reduced by roughly 50 % with a concomitant increase in the relative amounts of nutritionally superior fractions such as albumins, globulins and glutelins. The endosperm of *o2* maize contains twice as much lysine and tryptophan and 30 % less leucine than normal maize. The decreased level of zein (5–27 %) in *o2* maize along with reduced leucine, leads to more tryptophan for niacin synthesis and thus helps to combat pellagra and significantly improves its nutritional quality (Prasanna et al. 2001).

21.1.2 High-Lysine Mutants in Maize

Several mutants have been detected that favorably influence maize endosperm protein quality by elevating levels of two essential amino acids, lysine and tryptophan. The discovery of *o2* (Mertz et al. 1964) was followed by recognition of the biochemical effects of *floury2* (*fl2*) (Nelson et al. 1965). Searches for new mutants continued and resulted in the discovery of several others such as *opaque7*(*o7*) (Misra et al. 1972), *opaque6*(o6) and *floury3* (*fl3*) (Ma and Nelson 1975), *mucronate* (*Mc*) (Salamini et al. 1983) and *defective endosperm* (*De-B30*) (Salamini et al. 1997). Attempts were also made to find genotypes with high-lysine genes that retained a high level of zein fraction. Two such mutants, *opaque7749* and *opaque7455* (*o11*) (Nelson 1981) are particularly interesting as they have markedly higher levels of lysine as well as a high prolamine fraction. The specific chromosomal location is known for some of

the mutants. For example, the *o2* mutant is located on chromosome7, *fl2*on chr.4, *o7* on chr.10, *fl3* on chr.8 and *de-B30* on chr.7. The genetic action of some of the mutants is also known, for example, *o2*, *o6*, *o7* and *o11* are completely recessive. The two floury mutants are semi-dominant and exhibit variable expression for kernel opacity and protein quality depending on the presence of one or more recessives in the triploid endosperm. The mutant *De-B30* is dominant and shows dosage effects on kernel opacity and zein content (Soave et al. 1982).

21.1.3 Pleiotropic and Secondary Effects of o2 and Other High-Lysine Mutants

Genes and gene combinations that bring about drastic alterations in either plant or kernel characteristics also produce several secondary or undesirable effects. The low prolamine or high-lysine mutants are no exception. In addition to influencing several biochemical traits, they adversely affect a whole array of agronomic and kernel characteristics. The *o2*and other mutants adversely affect dry matter accumulation resulting in lower grain yield due to increased endosperm size. The kernels dry slowly following physiological maturity of the grain and have a higher incidence of ear rots. Other changes generally associated with high-lysine mutants include thicker pericarp, larger germ size, reduced cob weight, increased color intensity in yellow maize grains, and reduction in kernel weight and density. Thus, despite the nutritional superiority of *o2* maize, it did not become popular with farmers as well as consumers mainly because of reduced grain yield, chalky and dull kernel appearance and susceptibility to ear rots and stored grain pests. Hence, CIMMYT undertook to improve the phenotype of *o2* kernels to facilitate greater acceptability by developing hard endosperm grain types with protein quality of chalky *o2* strains. CIMMYT received funding support beginning in 1965 from the United Nations Development Program and introduced gene modifiers that changed the soft, starchy endosperm to a vitreous type preferred by farmers and consumers whilst retaining the elevated levels of lysine and tryptophan. CIMMYT has subsequently developed a range of hard endosperm *o2* genotypes with better protein quality through genetic selection, which are popularly known as quality protein maize (QPM). Today's QPM is essentially interchangeable with normal maize in both cultivation and agronomic characteristics as well as competitive in terms of yield, lodging, disease and pest resistance, moisture level while retaining the superior lysine and tryptophan content (Vasal 2001). In 2005, QPM was planted on 695,200 ha across 24 developing countries.

21.2 QPM Development Through Conventional Breeding

There are various breeding options for developing hard endosperm, high-lysine maize that is competitive in agronomic performance and market acceptance which could be based on specific endosperm high-lysine mutants or other donor stocks. The past

approaches involving normal maize breeding populations have centered on altering germ-endosperm ratio, selection for multiple aleurone layers, and recurrent selection to exploit natural variation for high-lysine content. Altering the germ-endosperm ratio to favor selection of larger germ size will have the dual advantage of increasing both protein quantity and quality (Bjarnason and Pollmer 1972) but it is not practical to attain lysine levels approaching those of *o2* maize. Besides, increased germ size has the disadvantage of contributing to poor shelf life of maize. Recurrent selection for high lysine in normal endosperm breeding populations has been largely unsuccessful due to the narrow genetic variation and heavy dependence on laboratory facilities of this approach. Alternatively, high-lysine endosperm mutants provided two attractive options: (1) exploiting double mutants involving *o2* and (2) simultaneous use of the *o2* gene with endosperm and amino acid modifier genes. In most instances, double mutant combinations involving *o2* and other mutants associated with endosperm quality were not vitreous (Vasal 2001). The most successful and rewarding option exploited combined use of *o2* with the associated endosperm and amino acid modifier genes.

Segregation and analysis of kernels with a range of endosperm modification began at CIMMYT as early as in 1969 by John Lonnquist and V. L. Asnani. Initial efforts towards development of QPM donor stocks with good kernel phenotypes as well as good protein quality proved to be highly challenging. Two effective approaches, i.e., intra-population selection for genetic modifiers in *o2* backgrounds exhibiting a higher frequency of modified *o2*kernels and recombination of superior hard endosperm *o2* families, resulted in development of good quality QPM donor stocks with a high degree of endosperm modification. This was followed by the large-scale development of QPM germplasm with a wide range of genetic backgrounds, representing tropical, subtropical and highland maize germplasm and involving different maturities, grain color and texture. A summary of characteristics of promising QPM gene pools and populations developed at CIMMYT is provided in Table 21.1. An innovative breeding procedure designated as 'modified backcross cum recurrent selection' was designed to enable rapid and efficient conversion programs (Vasal et al. 2001). More recently pedigree back crossing schemes have been used to convert elite QPM lines to maize streak virus (MSV) resistance for deployment in Africa as well as conversion of elite African lines to QPM (Krivanek et al. 2007).

A QPM hybrid breeding program was initiated at CIMMYT in 1985 as the QPM hybrid product has several advantages over open pollinated QPM cultivars: (1) higher yield potential comparable to the best normal hybrids, (2) assured seed purity, (3) more uniform and stable endosperm modification and (4) less monitoring of protein quality required during seed production. Several QPM hybrid combinations were derived and tested through international trial series at multiple CIMMYT and NARS locations in Asia, Africa and Latin America. Current QPM breeding strategies at CIMMYT focus on pedigree breeding wherein the best performing inbred lines with complementary traits are crossed to establish new segregating families. Both QPM × QPM and QPM × non-QPM crosses are made depending upon the specific requirements of the breeding project. In addition, backcross conversion is also followed to develop QPM versions of parental lines of popular hybrid cultivars that

Table 21.1 Characteristics of QPM gene pools and populations developed at CIMMYT (using *o2* and associated modifiers) including protein, tryptophan and lysine contents in the whole grain. (Based on Vasal 2001)

QPM pop/pool	Adaptation	Maturity	Color	Texture	Protein (%)	Tryptophan in protein (%)	Lysine in protein (%)	Quality index
Population 61	Tropical	Early	Y	Flint	9.2	0.98	4.2	3.8
Population 62	Tropical	Late	W	Semi-flint	9.9	0.92	3.9	4.4
Population 63	Tropical	Late	W	Dent	9.1	0.97	4.3	4.3
Population 64	Tropical	Late	W	Dent	9.6	1.00	3.8	4.3
Population 65	Tropical	Late	Y	Flint	9.2	0.96	4.2	4.4
Population 66	Tropical	Late	Y	Dent	9.3	1.01	4.3	4.3
Population 67	Subtropical	Medium	W	Flint	9.9	1.04	3.9	4.8
Population 68	Subtropical	Medium	W	Dent	9.5	1.01	4.0	4.3
Population 69	Subtropical	Medium	Y	Flint	10.0	0.98	4.2	4.4
Population 70	Subtropical	Medium	Y	Dent	9.3	1.10	4.3	4.7
Pool 15 QPM	Tropical	Early	W	Flint-Dent	9.1	0.94	4.2	4.6
Pool 17 QPM	Tropical	Early	Y	Flint	8.9	1.04	4.5	4.5
Pool 18 QPM	Tropical	Early	Y	Dent	9.9	0.93	4.0	4.6
Pool 23 QPM	Tropical	Late	W	Flint	9.1	1.03	3.8	4.2
Pool 24 QPM	Tropical	Late	W	Dent	9.4	0.92	3.8	4.0
Pool 25 QPM	Tropical	Late	Y	Flint	9.8	0.94	4.0	4.0
Pool 26 QPM	Tropical	Late	Y	Dent	9.5	0.90	4.1	4.3
Pool 27 QPM	Subtropical	Early	W	Flint-Dent	9.5	1.05	4.2	4.8
Pool 29 QPM	Subtropical	Early	Y	Flint-Dent	9.2	1.06	4.3	4.8
Pool 31 QPM	Subtropical	Medium	W	Flint	10.2	0.96	4.1	4.5
Pool 32 QPM	Subtropical	Medium	W	Dent	8.9	1.04	4.2	4.5
Pool 33 QPM	Subtropical	Medium	Y	Flint	9.3	1.05	–	4.2
Pool 34 QPM	Subtropical	Medium	Y	Dent	9.1	1.10	4.1	4.5

are widely grown in CIMMYT's target regions. Inbred lines developed through this process are then used in formation of QPM hybrids and QPM synthetics (Krivanek et al. 2007).

The 2000 World Food Prize jointly honored two CIMMYT Scientists, Dr. S. K. Vasal (Plant Breeder) and Dr. Evangelina Villegas (Cereal Biochemist) for their combined efforts and stellar achievements in developing and promoting QPM varieties to improve productivity and nutrition in malnourished and poverty-stricken areas of the world.

The genetic make-up of the QPM necessitates their cultivation in isolation from normal maize, as any contamination with *O2* allele will be apparent in the form of normal transparent kernels in contrast to the marble-like appearance of *o2*kernels. As isolation distance of 300–400 m is adequate; with the increase in the number of border rows, this distance can be suitably reduced. Even in the absence of isolation, the farmers planting 2–4 ha can save the seed from the middle of the field, whereas the rest of the crop can be used as nutritionally superior grain.

21.3 Molecular Basis of QPM: O2 and Modifier Gene Action

The breeding of QPM involves manipulation of three distinct genetic systems (Bjarnason and Vasal 1992; Krivanek et al. 2007): (1) the recessive mutant allele of the *O2* gene, (2) the endosperm hardness modifier genes and (3) the amino acid modifiers/genes influencing free amino acid content in the endosperm. The *O2*gene was cloned using a transposon tagging strategy with the maize mobile genetic elements, *Spm*(Schmidt et al. 1990) and *Ac*(Motto et al. 1988). The *O2* gene encodes a leucine-zipper class transcription factor required mainly for the expression of 22 kDa α-zein-coding genes and a gene encoding a ribosomal inactivating protein (Lohmer et al. 1991; Bass et al. 1992). Lower α-zein content in *o2* endosperm results in protein bodies that are about one-fifth to one tenth the normal size, which is presumed to alter packing of starch grains during seed desiccation, thereby conferring a characteristic soft texture to the kernel. With the reduction of α-zeins in the endosperm due to *o2* mutation, there is a usually concomitant increase in the level of γ-zeins (Habben et al. 1993). The homozygous recessive allele causes a decrease of the production of these zeins resulting in a corresponding increase in non-zein proteins, rich in lysine and tryptophan (Gibbon and Larkins 2005). Additionally, the recessive allele of the *o2* transcription factor also reduces the production of the enzyme, lysine keto-glutarate reductase, involved in free lysine degradation resulting in enhanced free lysine in the endosperm of *o2* maize. In the segregating generations, this recessive allele is selected either visually (identifying mosaic ears on F_2 harvests) or using molecular markers.

The mechanism(s) by which the endosperm modifier genes convert the starchy endosperm of *o2* to a normal phenotype is still not completely understood, but some important clues have been obtained through analysis of biochemical changes in modified *o2* endosperm. Studies suggest that the products of the modifier genes interact

with γ-zein mRNA transcripts and enhance their transport from the nucleus or increase their stability and translation. The overproduction of γ-zein appears to enhance protein body number and result in the formation of more vitreous endosperm. The endosperm hardness modifier genes, which convert the soft/opaque endosperm to a hard/vitreous endosperm without much loss of protein quality, are selected through a low-cost but effective method of light-box screening, where light is projected through the vitreous grains or blocked by the opaque grains. Endosperm modification is polygenically controlled. However, genetic and molecular analyses revealed some major loci involved in *o2* modification; for example, one locus maps near the centromere of chromosome 7 and the second maps near the telomere on the long arm of chromosome 7 (Lopes et al. 1995).

Despite the presence of *o2* and associated endosperm hardness modifier genes, the lysine and tryptophan levels in segregating families vary widely indicating the existence of third set of genes that modify the amino acid content, which necessitates systematic biochemical evaluation of lysine and/or tryptophan levels in each breeding generation. The lysine content of normal maize is around 2 %, whereas it is approximately 4 % (of the total protein) in QPM, with a range 1.6–2.6 % in normal maize and 2.7–4.5 % in QPM. Three genes associated with lysine level have been mapped to locations on chromosome 2, 4 and 7, besides several major *o*2 modifier-QTLs on chromosomes 1, 7 and 9 (Gibbon and Larkins 2005). Therefore, it is possible to get favorable responses to selection for endosperm texture modification as well as relative content of the essential amino acids, if they are monitored efficiently, during the QPM breeding programs.

21.4 Molecular Marker-Assisted Selection (MAS) in QPM Breeding

The *o2* gene is recessive and the modifiers are polygenic. Their introgression into elite inbred lines is not straight forward because of three major factors, (1) each conventional backcross generation needs to be selfed to identify the *o2* recessive gene and a minimum of six backcross generations are required to recover satisfactory levels of the recurrent parent genome, (2) in addition to maintaining the homozygous *o2* gene, multiple endosperm modifiers must also be selected, and (3) rigorous biochemical tests to ensure enhanced lysine and tryptophan levels in the selected materials in each breeding generation require enormous labor, time and financial resources. Although conventional breeding procedures have been used to convert commercial lines to QPM forms, these procedures are tedious and time consuming. Rapid advances in genomics research and technologies has led to the use of MAS which holds promise in enhancing selection efficiency and expediting the development of new cultivars with higher yield potential (Ribaut and Hoisington 1998; Xu and Crouch 2011). While marker-assisted foreground selection (Tanksley 1983; Melchinger 1990) helps in identifying the gene of interest without extensive phenotypic assays, marker-assisted background selection (Young and Tanksley 2005; Hospital et al. 1992; Frisch et al. 1999a, b) significantly expedites the rate of genetic

gain/recovery of recurrent parent genome in a backcross breeding program. With the development and access to reliable PCR-based allele-specific markers such as simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs), MAS is becoming an attractive option, particularly for oligogenic traits such as QPM (Babu et al. 2004).

A rapid line conversion strategy for QPM has been developed (Babu et al. 2005), consisting of a two-generation backcross program that employs foreground selection for the *o2* gene in both backcross (BC) generations, background selection at non-target loci in the BC_2 generation, and phenotypic selection for kernel modification and other desirable agronomic traits in two subsequent selfed generations. The rapid line conversion strategy outlined in this investigation brings together the salient features of both marker-aided and phenotypic-based selection approaches such as fixing the large segregating generation for the target locus (*o2*), reducing the linkage drag by selection of flanking markers for recipient allele type, recovering maximum amount of recurrent parent genome within two BC generations and providing scope for precise phenotypic selection for desirable agronomic and biochemical traits on a reduced number of progeny.

21.4.1 Low-Cost and Reliable Markers for o2 and Modifier Genes

The cloning and characterization of the *o2* gene, followed by detection of three SSR markers (*phi057*, *phi112* and *umc1066*) within the gene (Lin et al. 1997; Yang et al. 2008), led to effective differentiation of the *O2* and *o2* alleles (Bantte and Prasanna 2003; Babu et al. 2005). These *o2*-specific SSR markers provide an excellent foundation for MAS but this alone is not sufficient to bring to bear the full effectiveness of molecular breeding for QPM genotypes. Each of the microsatellite markers located within the *o2* gene are associated with factors that challenge their routine use in MAS programs. *umc1066* is easily visualized on agarose gels but is commonly not polymorphic in CIMMYT breeding populations; *phi057* is difficult to visualize on agarose gels, usually requiring the use of polyacrylamide gels; *phi112* is a dominant marker and hence cannot be used in the identification of heterozygotes in F2/BC populations. However, *phi112*, which is based on a deletion in the promoter region, has the advantage of being a widely conserved marker, consistent with the phenotype in QPM germplasm tested. In order to overcome these difficulties, CIMMYT is in the process of identifying functional and more discriminative SNP markers that could be used in high throughput genotyping platforms for selection of the *o2* allele.

Effective markers associated with modifying loci for both endosperm hardness and amino acid levels need to be identified. Unfortunately relatively little is known about the number, chromosomal location and mechanism of action of these modifier genes. A complex system of genetic control of these modifier loci with dosage effects, cytoplasmic effects, incomplete and unstable penetrance in different QPM germplasm

creates a major bottleneck to the accelerated development of QPM germplasm. Using a limited set of restriction fragment length polymorphism (RFLP) markers and bulked segregant analysis (Lopes et al. 1995), two chromosomal regions on the long arm of chromosome 7 that are associated with *o2* endosperm hardness modification were identified. The locus near the centromere is linked with the gene encoding the 27-kDa gamma zein. More recently, the analysis of two different QPM genotypes, Ko326Y and CM105Mo2 (derived from CIMMYT's Pool 33 QPM) corroborated the existence of a common quantitative trait locus (QTL) near the centromere of chromosome 7 that appears to have a major effect (30 % of the phenotypic variance) on *o2* endosperm modification, in addition to a QTL on 9.04/9.05 (Holding et al. 2008). In this specific F_2 population segregating for kernel vitreousness, these two loci accounted for 40 % of the phenotypic variation and thus may prove to be strong candidates for MAS for QPM breeding.

Precise information on genes controlling the level of amino acid modification especially with respect to lysine and tryptophan is relatively scarce and studies to date have found several quantitative trait loci (QTL) on many of the maize chromosomes (Wang and Larkins 2001; Wu et al. 2007). The free amino acid (FAA) content in Oh545o2 is 12 times greater than its wild-type counterpart, and three and 10 times greater than in Oh51Ao2 and W64Ao2, respectively. QTL mapping in these lines identified four significant loci that account for about 46 % of the phenotypic variance for FAA (Wang et al. 2001). One locus on the long arm of chromosome 2 is coincident with genes encoding a monofunctional *Aspartate kinase 2* (*Ask2*) whereas another locus on the short arm of chromosome 3 is linked with a cytosolic triose phosphate isomerase 4. Subsequent feedback inhibition analysis has suggested that *Ask2* is the candidate gene associated with the QTL on 2S and that a single amino acid substitution in the C-terminal region of the *Ask2* allele of Oh545o2 is responsible for altered basal activity of the enzyme (Wang et al. 2001). Using a RIL population from the cross between B73o2 (an *o2* conversion of B73) and a QPM line (CML161), it was possible to identify three QTL for lysine content and six QTL for tryptophan content, which explained 32.9 and 49.1 % of the observed variation, respectively (Gutierrez-rojas 2007). Thus a series of molecular markers for manipulation of different genetic components of QPM is available and hence their validation and fine mapping in appropriate breeding populations should now be carried out in order to establish a single cost-effective MAS assay for molecular breeding of QPM. Concerted research efforts to quantify the effect of these loci affecting endosperm hardness and amino acid levels coupled with marker development and validation will also accelerate the pace and precision of QPM development programs.

The previously cited studies involved very diverse germplasm, including tropical x temperate crosses, but current research at CIMMYT is seeking useful markers associated with kernel hardness and high amino acid levels in elite QPM $\times$ QPM crosses. A combination of bulked segregant analysis (BSA) and genome-wide SNP scan (using Illumina's GoldenGate assay) in phenotypically contrasting (opaque vs. modified and high vs. low tryptophan) progenies of seven QPM $\times$ QPM populations have identified several genomic regions putatively associated with kernel hardness and high tryptophan concentration. BSA of class5 (opaque) and class1 (completely

modified) bulks identified five regions [1.04/1.05 (1S), 7.02 (7S), 8.06 (8L), 9.04/9.05 (9L) and 10.04 (10S)] associated with kernel hardness modification, each indicated by more than three SNP markers in at least four populations. The region on 7.02, in which a major QTL has been previously reported to be linked to the 27 kDa gamma-zein gene responsible for kernel hardness modification in QPM, was consistently associated with kernel modification for all seven populations. Genome-wide SNP scan of the backcross-derived 'high-low' tryptophan lines revealed several polymorphic regions, notably 2.07, 5.03. and 10.03, which are currently being validated at CIMMYT using SSR and SNP markers and additional segregating populations. As a cross validation, whole genome SNP profiles of 12 isoline pairs of QPM and their normal versions were generated, which confirmed the importance of the above-mentioned eight genomic regions associated with kernel modification and tryptophan content (Babu et al. 2009). These genomic regions and their diagnostic markers may be useful for designing a comprehensive system for cost-effective marker-assisted QPM breeding.

21.4.2 Seed DNA-Based Genotyping in MAS for QPM Development

Leaf collection from the field, labeling and tracking back to the source plants after genotyping are rate limiting steps in leaf DNA-based genotyping. Recently, an optimized genotyping method using endosperm DNA sampled from single maize seeds was developed at CIMMYT, which has the potential to replace leaf DNA-based genotyping for marker-assisted QPM breeding (Gao et al. 2008). This method is suitable for various types of maize seeds, produces high quality and quantity of DNA and has minimal effects on subsequent germination and establishment. A substantial advantage of this approach is that it can be used to select desirable genotypes before planting, which can bring about dramatic enhancements in efficiency by planting only the plants containing *o2* gene in recessive form in BC_nF_2 generations of non-QPM × QPM crosses, and also by minimizing the labor costs and scoring error associated with light box screening of a large number of grains for endosperm hardness. Over several breeding cycles, this is likely to lead to cumulative and accelerated gains in selection pressure (such as light box screening for endosperm hardness modification and systematic biochemical evaluation of lysine and/or tryptophan) and improvements in overall QPM breeding efficiency.

21.4.3 MAS and Development of QPM Cultivars: Successful Examples

To hasten the pace of progress of QPM cultivar development, and most importantly, to diversify the genetic base of QPM cultivars for any targeted agro-ecology, it is

important to convert some of the highly diverse and agronomically elite non-QPM inbred lines into QPM versions and derive more heterotic QPM hybrid combinations. Conversion of a normal maize line to a QPM version through conventional backcross breeding requires at least 6–7 years, since the desired *o2* allele has to be in a homozygous recessive state, thereby warranting progeny testing after every backcross. In addition, as discussed previously, the introgression of the polygenic endosperm modifiers into elite inbreds is not a straight-forward procedure as this is presently based on phenotypic selection alone and rigorous biochemical tests have to be undertaken to ensure enhanced lysine and tryptophan levels in the selected genotypes. Consequently, such a breeding programme would involve significant land, labour and financial resources (Dreher et al. 2003).

The recent developments in plant biotechnology, including molecular mapping and marker-assisted selection (MAS) offer a choice of options for introgression of the target gene(s) in the genetic background of elite varieties of major crops. MAS refers to the manipulation of genomic regions that are involved in the expression of traits of interest using molecular markers (Babu et al. 2004). The PCR based molecular markers help in targeted 'foreground selection' of segregating/backcross progenies possessing the desired gene(s), besides shortening of the breeding cycle significantly through rapid recovery of recurrent parent genome using 'background selection'. Microsatellite or Simple Sequence Repeat (SSR) markers are particularly useful in undertaking MAS in crop plants like maize.

Identification and utilization of *o2*-specific SSR markers (Lin et al. 1997; Bantte and Prasanna 2003) offered tremendous advantage in molecular marker-assisted conversion of non-QPM lines into their QPM versions. 'Foreground selection' for the *o2* allele using SSR markers and 'background selection' (using markers polymorphic between the donor and recurrent parents) aid in recovering individuals with desired genotype at the target locus, besides high levels of recovery of recurrent parent genome, within two to three backcross generations. This program can be thus implemented in a cost- and time-effective manner as compared to that based on phenotypic selection alone (Dreher et al. 2003).

There are a few successful examples of MAS for maize improvement using *o2*-specific SSR markers (Babu et al. 2005; Gupta et al. 2009; Prasanna et al. 2010). The parental lines of 'Vivek Hybrid 9' (CM145 and CM212), developed at Vivekananda Parvatiya Krishi Anusandhan Sansthan (VPKAS), Almora, were converted into QPM versions through transfer of *o2* gene using MAS and phenotypic screening for endosperm modifiers. The MAS-derived QPM hybrid, 'Vivek QPM 9' has been released in the year 2008 for commercial cultivation in zones I and IV in India. Vivek QPM 9 shows 41 % increase in tryptophan, 30 % in lysine, 23 % in histidine, 3.4 % in methionine, coupled with 12 % reduction in leucine, as compared to Vivek Hybrid 9. Domestic consumption of such QPM grains will help in reducing protein malnutrition in the hills and mountains. In view of this, F_1 hybrid seeds of Vivek QPM 9 are being produced in large scale for distribution in Uttarakhand and other parts of the country. Few villages in Uttarakhand have also been identified for converting them into QPM villages. Vivek QPM 9 can potentially replace Vivek Hybrid 9 as well as composites without any yield penalty in these areas. The approach outlined above

was also used to develop QPM versions of several elite, early maturing inbred lines adapted to the hill regions of India (Gupta et al. 2009, 2013).

QPM versions of six elite inbred lines, which are the parents of three single-cross hybrids, PEHM2, Parkash and PEEHM5 have been recently developed by the Maize Genetics Unit, IARI, New Delhi, through the ICAR Network Project on Molecular Breeding (Prasanna et al. 2010). A Network Project on molecular breeding for QPM is also being implemented in India, funded by the Department of Biotechnology (DBT), Govt. of India, for conversion of several important Indian maize inbred lines into QPM versions.

21.5 Need for an Integrated Approach for Enhancing Maize Protein Quality

Research at CIMMYT is currently focused on developing a package of molecular markers for cost-effective large scale, marker-aided QPM breeding program. CIMMYT recently developed gene-based SNP markers and high throughput KASPar chemistry-based genotyping assay for the *o2* gene. This system is likely to be useful especially for NARS programs with limited or no lab facilities, wherein NARS researchers could effectively outsource the SNP genotyping process to cost-effective service providers in the region. In addition, intensive efforts are being made to develop and validate new as well as existing markers for the endosperm and amino acid modifier genes across a wide range of populations and improved pools. Recent investigations into the improved protein quality of the *o2* mutant and the genetic mechanisms that can suppress its starchy kernel phenotype provide new insights to support the continued improvement of QPM. Chief among these developments are the use of transgenic approaches to improve nutritional quality and the discovery that an important component of modified endosperm texture in QPM is related to altered starch granule structure (Gibbon and Larkins 2005).

Another possible opportunity to further improve the protein quality conferred by *o2* is to pyramid it with second high-lysine mutant, *o16*, which was selected from Robertson's *Mutator* (*Mu*) stock where the lysine content in the F_3 *o2o2o16o16* families derived from recombination of both *o2* and *o16*, was about 30 % higher than that of *o2o2* or *o16o16* F3 families (Yang et al. 2004). Using both genes together could lessen the requirement for phenotypic screening for amino acid content and genetic screening for amino acid modifiers.

Recent efforts in genetic transformation are focused on developing a dominant *o2* trait in maize. RNA interference (RNAi) technology has also been used to reduce 22-kDa (Segal et al. 2003) and 19-kDa alpha zeins (Huang et al. 2004, 2005) using antisense transformation constructs, which resulted in moderate increases (15–20 %) in lysine content. In a recent study, using an improved double strand RNA (dsRNA) suppression construct, Huang et al. (2006) reported lysine and tryptophan levels similar to conventionally bred QPM genotypes. While the dominant nature of the anti-sense transgene is a definite advantage as compared to recessive allele of *o2*, the opaque

endosperm still needs to be modified by endosperm modifier genes whose epistasis with the transgene has not yet been tested. Very recently, Wu and Messing (2011) reported a potential accelerated QPM selection scheme, which is based on RNA interference construct that is directed against 22 and 19 kDa zeins, fused with visible green fluorescent protein (GFP) marker gene. When such RNAi lines were crossed with QPM lines, carrying *o2* kernel hardness modifier genes, green and vitreous progenies could be selected in the segregating generations thereby demonstrating that high-lysine content and hard endosperm traits could be selected in dominant fashion. Although RNAi technology has emerged as a powerful tool to overcome the pleiotropic and secondary effects of the desirable mutant genes, social acceptance and biosafety concerns regarding genetically modified food crops still exist in some countries for large scale adoption.

21.6 Conclusion

The opportunities for implementing breeding for improved nutritional quality in crops like maize have increased tremendously in the recent years. Significant strides have also been made, particularly with regard to MAS for generating QPM versions for elite inbred lines (Babu et al. 2005; Gupta et al. 2013) and identification of genes/QTLs influencing diverse quality traits in maize (Babu et al. 2013). Yet, the application of molecular tools to accelerate breeding for improved nutritional quality in maize has barely begun, and there is vast potential and need to expand the scope and impact of such operations. Breeders will want to avail molecular tools to more efficiently add value to new maize cultivars, for example, by enhancing their nutritional or biochemical qualities for use as food, feed, and industrial material.

Biofortification of maize grains is an important area of research for which *o2* provides an ideal platform upon which a number of nutritionally important traits such as enhanced Fe and Zn, low phytate (for increased bioavailability of nutrients), high provitamin A and high methionine could be combined to derive multiple benefits. Considering the pace of technological developments in genome research, molecular breeding is likely to be the leading option in future for stacking a range of nutritionally important specialty traits.

References

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

Babu R, Nair SK, Kumar A et al (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, Atlin G, Vivek B et al (2009) Bulked segregant analysis using the high throughput maize GoldenGate SNP genotyping assay reveals multiple genomic regions associated with kernel hardness and tryptophan content in quality protein maize. Maize Genet Conf Abstracts 51:P178

Babu R, Rojas NP, Gao S, 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 126:389–399

Bantte K, Prasanna BM (2003) Simple sequence repeat polymorphism in Quality Protein Maize (QPM) lines. Euphytica 129:337–344

Bass HW, Webster C, Obrian GR et al (1992) A maize ribosome-inactivating protein is controlled by the transcriptional activator *Opaque-2*. Plant Cell 4:225–234

Bjarnason M, Pollmer WG (1972) The maize germ: Its role as contributing factor to protein quantity and quality. Zeitschrift Pflanzenzuchtung 68:83–89

Bjarnason M, Vasal SK (1992) Breeding of quality protein maize (QPM). Plant Breed Rev 9:181–216

Bressani R (1991) Protein quality of high lysine maize for humans. Cereal Food World 36:806–811

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

FAO (1992) Maize in human nutrition. Food and nutrition series, no. 25. Food and Agriculture Organization (FAO), Rome, Italy

Frisch M, Bohn M, Melchinger AE (1999a) Comparison of selection strategies for marker assisted back crossing of a gene. Crop Sci 39:1295–1301

Frisch M, Bohn M, Melchinger AE (1999b) Minimum sample size and optimum positioning of flanking markers in marker assisted back crossing for transfer of a target gene. Crop Sci 39: 967–975

Gao S, Martinez C, Skinner DJ et al (2008) Development of a seed DNA-based genotyping system for marker-assisted selection in maize. Mol Breed 22:477–494

Gibbon BC, Larkins BA (2005) Molecular genetic approaches to developing quality protein maize. Trends Genet 21:227–233

Graham GG, Lembcke J, Morales E (1990) Quality protein maize as the sole source of dietary protein and fat for rapidly growing young children. Pediatrics 85:85–91

Gunaratna NS, De Groote H, Nestel P, Pixley KV, McCabe, CP (2010). A meta-analysis of community-based studies on quality protein maize. Food Policy. 35:202–210.

Gupta HS, Agarwal 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

Gupta HS, Babu R, Agarwal PK, Mahajan V et al (2013) Accelerated development of quality protein maize hybrid through marker-assisted introgression of *opaque-2* allele. Plant Breed 132:77–82

Gutierrez-rojas LA (2007) Quantitative trait loci analysis to identify modifier genes of the gene *o2* in maize endosperm. PhD Dissertation, Texas A & M University

Habben IE, Kirleis AW, Larkins BA (1993) The origin of lysine-containing proteins in *opaque-2* maize endosperm. Plant Mol Biol 23:825–838

Holding DR, Hunter BG, Chung T et al (2008) Genetic analysis of *o2* modifier loci in quality protein maize. Theor Appl Genet 117:157–170

Huang SS, Adams WR, Zhou Q et al (2004) Improving nutritional quality of maize proteins by expressing sense and antisense zein genes. J Agric Food Chem 52:1958–1964

Huang SS, Kruger DE, Frizzi A et al (2005) High-lysine corn produced by the combination of enhanced lysine biosynthesis and reduced zein accumulation. Plant Biotech J 3:555–569

Huang SS, Frizzi A, Florida CA et al (2006) High lysine and high tryptophan transgenic maize resulting from the reduction of both 19- and 22-kD alpha-zeins. Plant Mol Biol 61:525–535

Krivanek AF, De Groote H, Gunaratna NS et al (2007) Breeding and disseminating quality protein maize (QPM) for Africa. African J Biotech 6:312–324

Lin KR, Bockoit AJ, Smith JD (1997) Utilization of molecular probes to facilitate development of Quality Protein Maize. Maize Genet Coop News 71:22–23

Lohmer S, Maddaloni M, Motto M et al (1991) The maize regulatory locus *Opaque-2* encodes a DNA-binding protein which activates the transcription of the *b-32* gene. EMBO J 10:617–624

Lopes MA, Takasaki K, Botswick DE et al (1995) Identifications of two *o2* modifier loci in Quality Protein Maize. Mol Gen Genet 247:603–613
Ma Y, Nelson OE (1975) Amino acid composition and storage proteins in two new high lysine mutants in maize. Cereal Chem 52:412–419
Melchinger AE (1990) Use of molecular markers in plant breeding. Plant Breed 104:1–19
Mertz ET, Bates LS, Nelson OE (1964) Mutant genes that change protein composition and increase lysine content of maize endosperm. Science 145:279–280
Misra PS, Jambunathan R, Mertz ET et al (1972) Endosperm protein synthesis in maize mutants with increased lysine content. Science 176:1426
Motto M, Maddolini M, Panziani G et al (1988) Molecular cloning of the o2-*m5* allele of *Zea mays*, using transposon tagging. Mol Gen Genet 121:488–494
Nelson OE (1981) The mutants *opaque9* through *opaque13*. Maize Genetics Coop. Newsletter 55:68
Nelson OE, Mertz ET, Bates LS (1965) Second mutant gene affecting the amino acid pattern of maize endosperm proteins. Science 150:1469–1470
Osborne TB, Mendel LB (1914) Nutritive properties of protein of the maize kernel. J Biol Chem 18:1–16
Prasanna BM, Pixley K, Warburton ML, Xie C (2010) Molecular marker-assisted breeding for maize improvement in Asia. Molecular Breeding. 26:339–356
Prasanna BM, Vasal SK, Kassahun B, Singh NN (2001) Quality protein maize. Curr Sci 81: 1308–1319
Ribaut JM, Hoisington DA (1998) Marker assisted selection: new tools and strategies. Trends Plant Sci 3:236–239
Salamini F, Di Fonzo N, Fornasari E et al (1983) *Mucronate*, *mc*, a dominant gene of maize which interacts with *o2* to suppress zein synthesis. Theor Appl Genet 65:123–128
Salamini F, Di Fonzo N, Gentinetta E, Soave C (1997) A dominant mutation interfering with protein accumulation in maize seeds. In: Seed protein improvement in cereals and grain legumes, IAEA, Vienna, p. 97
Schmidt RJ, Burr FA, Aukerman MJ, Burr B (1990) Maize regulatory gene *opaque-2* encodes a protein with a "leucine-zipper" motif that binds to zein DNA. Proc Natl Acad Sci USA 87:46–50
Segal G, Song RT, Messing J (2003) A new opaque variant of maize by a single dominant RNA-interference-inducing transgene. Genetics 165:387–397
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
Singleton WR (1939) Recent linkage studies in maize: V. *opaque endosperm-2* (*o2*). Genetics 24:61–63
Soave C, Reggiani R, Di Fonzo N, Salamini F (1982) Clustering of genes for 20 kd zein sub units in the short arm of maize chromosome 7. Genetics 97:363–377
Tanksley SD (1983) Molecular markers in plant breeding. Plant Mol Bio Rep 1:1–3
Vasal SK (2000) The quality protein maize story. Food Nutr Bull 21:445-450
Vasal SK (2000) The quality protein maize story. Food Nutr Bull 21:445–450
Vasal SK (2001) High quality protein corn. In: Hallauer A (ed) Specialty corn, 2nd edn. CRC, Boca Raton, FL, pp 85–129
Vasal SK, Villegas E, Bjarnason M et al (1980) Genetic modifiers and breeding strategies in developing hard endosperm *o2* materials. In: Pollmer WG and Phipps RH (eds) Improvement of quality traits of maize grain and silage use
Villegas E, Ortega E, Bauer R (1984) Chemical methods used at CIMMYT for determining protein quality in corn. CIMMYT, Mexico
Wang X, Larkins BA (2001) Genetic analysis of amino acid accumulation in *o2* maize endosperm. Plant Physiol 12:1766–1777
Wang X, Stumpf DK, Larkins BA (2001) Aspartate kinase 2-a candidate gene of quantitative trait locus influencing free amino acid content in maize endosperm. Plant Physiol 125:778–787

Wang X, Lopez-Valenzuela JA, Gibbon BC et al (2007) Characterization of monofunctional aspartate kinase genes in maize and their relationship with free amino acid content in the endosperm. J Exp Bot 58:2653–2660

Wu RL, Lou XY, Ma CX et al (2002) An improved genetic model generates high-resolution mapping of QTL for protein quality in maize endosperm. Proc Natl Acad Sci USA 99:11281–11286

Wu Y, Messing J (2011) Novel genetic selection system for Quantitative Trait Loci of Quality Protein Maize. Genetics 188:1019–1022

Xu Y, Crouch JH (2008) Marker assisted selection in plant breeding: From publications to practice. Crop Sci 48:391–407

Yang W, Zheng Y, Ni S, Wu J (2004) Recessive allelic variation of three microsatellite sites within the *o2* gene in maize. Plant Mol Biol Rep 22:361–374

Yang W, Zheng Y, Zheng W, Feng R (2005) Molecular genetic mapping of a high-lysine mutant gene (opaque-16) and the double recessive effect with *opaque-2* in maize. Mol Breed 15:257–269

Young ND, Tanksley SD (1989) RFLP analysis of the size of chromosomal segments retained around *Tm-2* locus of tomato during backcross breeding. Theor Appl Genet 77:353–359