Review of Literature

Maize occupies an important position in world economy and trade as a food, feed and industrial grain crop and is one of the most important cereal crops after wheat and rice. The total production area for maize is 157 million hectare (FAO STAT, 2007) and millions of people worldwide are dependent on maize as a staple food through economic necessity, and derive their basic needs of protein and carbohydrates from it. Both in the developed and developing world, maize is an important source of human and animal nutrition.

The maize kernel, like that of other cereal grains, includes pericarp (6%), endosperm (82%) and germ (12%). The main structural component of the endosperm is starch, a complex carbohydrate that constitutes on an average 71% of the grain and is a source of concentrated energy. Bulk of the proteins in a mature maize kernel is in the endosperm and germ; but, the germ protein is superior in both quantity and quality. Normal maize protein, as a point of comparison, has a biological nutritional value of 40% of that of milk (Bressani, 1991) and therefore, needs to be eaten with complementary protein sources such as legumes or animal products.

**Amino Acids and their Importance**

Amino acids are the main nitrogen storage compounds in plants and the basis for the synthesis of proteins (Medici et al., 2004). Twenty different amino acids are usually incorporated into proteins (Lea and Azevedo, 2003) which are further classified into essential and non-essential amino acids on the basis of their requirement. Humans and monogastric animals are not able to synthesize nine of the amino acids that are found in proteins. These nine amino acids (lysine, threonine, methionine, phenylalanine, tryptophan, isoleucine, leucine, valine, and histidine) are designated as essential amino acids and must be acquired through the diet (Galili et al., 2002).

Plant proteins supply nearly 65% of the total global ingested proteins, with cereal grains representing 47% of it. Plant proteins in developed countries
constitute a low proportion compared with animal sources. However, in the developing countries plant proteins are the main and, in many cases, the only source of protein (Millward, 1999). However, cereal endosperm proteins are usually deficient in lysine, threonine and tryptophan, while legumes are deficient in methionine. These nutritional deficiencies have given rise to a great deal of interest in research into breeding for improved nutrition quality.

Tryptophan is the one of the least abundant yet, in terms of energy, one of the most expensive to produce among the standard amino acids (Hrazdina and Jensen, 1992). The low level of soluble tryptophan present in plants (1 to 15 μM) signifies the importance of this amino acid and the pathway that produces it (Gilchrist and Kosuge, 1980). Animals, and some eubacteria, lack the ability to synthesize tryptophan and must obtain it from plant and microbial sources (Crawford, 1989; Bentley, 1990; Herrmann et al., 1992). This essential amino acid is required by animals for protein synthesis as well as for the production of other compounds, including the neuro-hormone serotonin and the vitamin nicotinic acid.

Storage Protein Synthesis in Maize Endosperm

The mature maize kernel consists of embryo and a larger endosperm both of which are surrounded by the seed coat. The outermost layer, the aleurone is composed of specialized cells that secrete hydrolytic enzyme during germination. Beneath the aleurone are starchy endosperm cells that are filled with starchy endosperm. Typically the endosperm is ~90% starch and 10% protein.

Nearly 70% of endosperm protein is composed of several types of prolamin, known as ‘zeins’ (Gibbon and Larkin, 2005). Zeins are hydrophobic and are soluble in high concentrations of alcohols such as ethanol or butanol. Zeins have never been detected in any part of the plant other than seed (Boston et al., 1986), where in it is most abundant in endosperm than in embryo (Tsai, 1979).
The proportion of various endosperm storage protein fractions in normal maize, on an average are albumin (3%), globulins (3%), zeins (60%) and glutelins (34%) (Salamini and Soave et al., 1982). The α-zeins are the most abundant among the zeins and further sub classified as 22- and 19-kD proteins. There is a significant decrease in the level of 22-kD α-zein fraction, and to a lesser extent reduction in the 19-kD α-zein found in o2 genotype. The abundance of other zein fraction is less affected by opaque2 mutation (Habben et al., 1993). The non zein fraction comprise mainly of enzyme structural polypeptides and membrane associated proteins, are heterogeneous and do not vary much between the wild type maize and opaque2 mutant.

**Discovery and Analysis of opaque2 Mutation**

In the 1920s in a Connecticut USA maize field, a natural spontaneous mutation of maize with soft, opaque grains was discovered and the maize mutant was eventually named as opaque2 (o2) by a Connecticut researcher (Singleton, 1939). In 1964, Dr. Oliver Nelson’s team discovered that the homozygous recessive o2 allele had substantially higher lysine (+69%) in grain endosperm compared to normal maize (Mertz et al., 1964). Very soon, another mutation floursy2 was discovered, which also has the ability to alter the endosperm protein quality. It was further determined that these mutation result in two-fold increase in the level of two amino acids content, 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, the 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 help to liberate more tryptophan for niacin biosynthesis, and thus helps in combating pellagra.

The other additional mutation reported were opaque7 (McWhirter, 1971, Misra et al., 1972), opaque6, floury3 (Ma and Nelson, 1975), defective B30 (Salamani et al., 1979), mucronate (Salamani et al., 1983). None of the mutant has proved their practical utility in QPM breeding program except opaque7 which showed up to some extent (Bjarnason and Vasal, 1992), although they help in other way about the understanding of storage protein synthesis in the maize endosperm.

The Opaque2 (O2) gene was cloned using a transposon tagging strategy with the maize mobile genetic elements, Spm (Schmidt et al., 1987) and Ac (Moto et al., 1988). The O2 gene encodes a transcription factor required mainly for the expression of 22 kDa α-zein-coding genes and a 32-kDa albumin gene b-32, and is necessary for their expression (Schmidt et al., 1990; Lohmer et al., 1991; Ueda et al., 1992). The O2 protein contains a basic domain/ leucine zipper (bZIP) motif identified in DNA-binding proteins of animal proto-oncogenes and in transcriptional regulators of yeast (Harting et al., 1989). 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). Transcription factors of bZIP type frequently function as heterodimers and heterodimerization between O2 and OHP1 (another bZIP protein), has been demonstrated (Pysh et al., 1993), suggesting the involvement of multiple bZIP proteins in transcriptional control of zein genes (Or et al., 1993).

With the discovery of the nutritional benefits of the o2 mutation, breeding programmes in several countries began to incorporate this gene through backcrossing, leading to o2 cultivars. However, the direct use of the
o2 mutation in breeding programs soon receded after the discovery of serious negative secondary (pleiotropic) effects of this mutation (Prasanna et al., 2001). The soft endosperm texture caused by the opaque2 gene was not favoured by the farming community in many developing countries where translucent vitreous grain types are the most liked ones. Besides this, the soft endosperm of o2 genotypes caused up to a 25% yield loss due to the lower density of the opaque grains, as well as due to increased susceptibility to fungal ear rots and storage pests (Vasal, 2000). Such negative secondary effects severely limited practical use of the mutation in the field.

Endosperm Modifiers for QPM Development
During the process of converting normal maize populations to o2 versions, partially hard endosperm (i.e. vitreous) or “modified” grains was observed by many researchers including breeders at CIMMYT in Mexico. Separations of such grains when encountered began as early as 1969 by Dr. John Lonnquist and provided a new way to development of opaque2 varieties with hard kernels (Bajarnson and Vasal, 1992; Vasal, 2000; Prasanna et al., 2001).

At CIMMYT, the modified kernels were classified into different categories and laboratory analysis was carried out to study the effect of the degree of modification on biochemical characteristics (Villegas et al., 1984). A few o2-converted population were then identified which had usually a high frequency of modified kernels. The breeding procedures followed for the incorporation of genetic endosperm modifiers along with o2 gene were reviewed in detail by Prasanna et al. (2001).

Glover and Mertz (1987) indicated that the modified endosperm texture is polygenically controlled with additive type of genetic variation playing an important role, although in some materials a few major genes may contribute significantly to kernel modification. The genetic background of the material and its kernel texture could also influence kernel modification and frequencies of various modification classes. Pixley and Bjarnason (2002) demonstrated that
endosperm modification is less influenced by genotype x environment interaction and therefore is stable character.

The mechanism(s) by which the modifier genes convert the starchy endosperm of o2 to a normal phenotype is still poorly understood, but some important clues have been obtained through analysis of biochemical changes in modified o2 endosperm. QPM genotypes appear to have levels of α-zein comparable to unmodified o2 lines, but the level of γ-zein is increased by 2-3 folds. 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 (Boyer and Larkin, 1993). Mutations that reduce α-zein synthesis, such as opaque2 (o2), result in small, unexpanded protein bodies (Geetha et al., 1991), whereas those that reduce γ-zein synthesis, such as o15 (Dannenhoffer et al., 1995), lead to a smaller number of protein bodies. Conversely, the overproduction of γ-zein appears to enhance protein body number and result in the formation of more vitreous endosperm (Lopes et al., 1995; Moro et al., 1995). Other opaque mutants, such as floury2 (fl2), Mucuronate (Mc), and Defective endosperm B30 (DeB30), are associated with irregularly shaped protein bodies (Fontes et al., 1991; Coleman et al., 1997b). At least 18 mutations have been described that cause a soft and starchy endosperm phenotype (Thompson and Larkins, 1994). Only in the case of o2 and fl2 is the molecular basis for the mutation is well understood. Genetic and molecular analysis using RFLP (Restriction Fragment Length Polymorphism) markers revealed two major loci involved in o2 modification; one locus maps near the centromere of chromosome 7 and the second maps near the telomere on the long arm of chromosome 7 (Lopez et al., 1995).

**QPM Cultivars**

Initial QPM breeding efforts at CIMMYT focused on conversion of a range of subtropical and tropical lowland adapted, normal endosperm populations to o2 versions through backcross-recurrent selection procedures, with a focus of
accumulating the hard endosperm phenotype, maintaining protein quality and increasing yield and resistance to ear rot (Villegas et al., 1992, Bajarnson and Vasal, 1992 Vasal, 2001; Prasanna et al., 2001). The resulting genotypes with elevated lysine and tryptophan content relative to normal maize but without negative soft endosperm phenotype were termed by CIMMYT as “Quality Protein Maize” (QPM) (Vasal et al., 1984b; Bajarnson and Vasal, 1992). The term QPM now refers to maize homozygous for the o2 allele, with increased lysine and tryptophan content without the secondary effects of soft endosperm (Vasal, 2001).

QPM looks and performs like normal maize and can be reliably differentiated only through laboratory test (Villegas et al., 1992). Several QPM population and pools possessing different ecological adaptation, maturity, grain colour and texture were developed (Vasal et al., 1984; CIMMYT, 1985; Vasal, 2001). A number of advanced maize populations in CIMMYT maize programme were successfully converted to QPM populations using this procedure.

Several advantages were advocated for QPM hybrids over the open pollinated varieties: (i) improving yield potential through exploitation of heterosis; (ii) facilitating maintenance of seed purity of inbred progenitor with respect to the agronomic trait, the genetic modifiers and the protein quality; (iii) reduced dependence on laboratory facilities for monitoring the protein quality provided the lines are fixed and kept genetically pure; (iv) the hybrids exhibit more uniformity and stability with respect to kernel modification and; (v) attracting involvement of private seed industry in the QPM efforts (Krivanek et al., 2007).

In India, a nutritionally superior o2 composite with hard kernel texture, designated as ‘Shakti-1’ was released in 1997. Since 1998, intensive efforts were made in different breeding centres in the country, resulting in the release of a series of QPM hybrids. The first white grained QPM hybrid ‘Shaktimaan-1’ (a three-way cross hybrid using CIMMYT QPM lines) was released during 2001 by RAU, Pusa, Bihar. ‘Shaktimaan-2’ (a single-cross hybrid with white
grain) was released by the same institution during 2004. The first yellow-grained single cross hybrid ‘HQPM-1’ was released by CCSHAIU, Karnal, during 2005, followed by the release of ‘Shaktimaan-3’ and ‘Shaktimaan-4’ during 2006 and ‘HQPM-5’ during 2007. Current QPM breeding strategies at CIMMYT and National breeding programs in India is focused on introducing and testing of QPM developed elsewhere, conversion of existing adapted genotypes to QPM either by conventional as well as marker-assisted selection (MAS) approach.

**Nutritional Significance and Economic Benefits of QPM**

The QPM offers tremendous benefits in the nutrition of monogastric animals including humans, since the two limiting amino acids (lysine and tryptophan) cannot be synthesized through metabolism in these organisms. In animal nutrition, QPM can provide a cheaper way of obtaining a balanced animal feed. Its potential impact can be found by comparing the price of optimal feed ratios, typically calculated using Linear Programming, with QPM and with normal maize. Results for poultry and pig ratios in the US, using USDA feed requirement and average US prices for maize, soybean, sorghum and synthetic lysine and methionine, showed relatively modest cost reductions, higher for pigs (3.4% for meat pigs and 3.0% for sows) than for poultry (2.8% for broilers and 2.6% for layers) (López-Pereira, 1993). A similar study in Kenya found a 5% cost reduction from substituting QPM for normal maize in broiler ratios (Nyanamaba et al., 2003).

The protein quality of QPM was evaluated using nitrogen balance technique in several studies (Bressani et al., 1969; Luna-Jaspe et al., 1971; Pradilla et al., 1973; Graham et al., 1980, 1989), where o2 maize has proved its superiority over the normal maize with respect to apparent nitrogen retention and biological value. These studies concluded that nitrogen balance and retention were higher with o2 maize and QPM, especially at lower levels of total protein intake. Moreover, unlike normal maize, young children can consume o2 maize in amounts needed for the positive nitrogen balance that is
required for growth (Bressani et al., 1969). Recovering malnourished children fed QPM further showed the same growth as those fed modified cow milk formula (Graham et al., 1990). Earlier studies was done in controlled environment, to check the impact of QPM in natural environment community level studies have been conducted in at least six countries including India, Guatemala, Brazil, Ghana, Mexico, and Ethiopia. The first of these, a six-month feeding trial conducted in 1975-1976 by the Indian Agricultural Research Institute (IARI), involved 134 children aged 18-30 months (Singh, 1977). Various measures of growth suggested some positive benefit of o2 to normal maize, but statistical analyses were not conducted to establish significant differences between these two groups. Lysine, tryptophan and protein superiority of QPM over normal maize has been studied by several researchers (Zarkadas et al., 1995; 2000; Bhatnagar et al., 2003; Kassahun and Prasanna, 2004). These finding indicate that QPM has added advantage of being superior in protein quality and higher in food and feed efficiency (Krivanek et al., 2007).

Besides, QPM protein contains, in general, 55% more tryptophan, 30% more lysine and 38% less leucine than that of normal maize. The biological value of normal maize protein is 45%, while that of o2 maize is 80%. Only 37% of common maize protein intake is utilized compared to 74% of the same amount of o2 maize protein. The nitrogen balance index for skim milk and o2 maize protein is 0.80 and 0.72, respectively, which indicates that the protein quality of QPM is 90% of that of milk (Graham et al., 1980; Bresanni, 1995). The other nutritional benefits of QPM include higher niacin availability due to a higher tryptophan and lower leucine content, higher calcium and carbohydrate, and carotene utilization (De Bosque et al., 1988). Further, high quality protein maize can be transformed into edible products without deterioration of its quality or acceptability, and can be used in conventional and new food products. The nutritional and biological superiority of QPM has also been amply demonstrated in model systems such as rats, pigs, infants and small children as well as adults (Prasanna et al., 2001). In Guatemala, it was
demonstrated that o2 maize has 90% of the nutritive value of milk protein in young children. Children in Colombia suffering from Kwashiorkar, a severe protein deficiency disease, were brought back to normalcy on a diet containing only o2 (Bresanni, 1990; Prasanna et al., 2001).

Molecular Marker-assisted Selection in QPM Breeding

‘Marker-assisted selection’ (MAS) refers to the manipulation of genomic regions that are involved in the expression of traits of interest through molecular markers. The recent developments in plant biotechnology including molecular mapping and MAS offer a choice of options for introgression of the target gene(s) in the genetic background of elite varieties of major crops. MAS involving the use of PCR based molecular markers helps in targeted selection of segregating/backcross progenies possessing the desired gene(s), besides shortening of the breeding cycle significantly.

Since 1980s, DNA-based markers have become important tools for genetic analyses. DNA markers, which are phenotypically neutral and literally unlimited in number, have allowed scanning of the whole genome and assigning landmarks in high density on every chromosome in many plant species, including tomato. During the past two decades, different types of molecular markers have been developed and evolved, including, but not limited to, restriction fragment length polymorphisms (RFLPs), random amplified polymorphic DNAs (RAPDs), amplified fragment length polymorphisms (AFLPs), simple sequence repeats (SSRs) or microsatellites, cleaved amplified polymorphic sequences (CAPS), sequence characterized amplified regions (SCARs), expressed sequence tags (ESTs), single nucleotide polymorphisms (SNPs) etc. (Prasanna and Hoisington, 2003).

Markers tightly linked to the target genes can greatly aid in MAS programmes, by tracking the gene under selection through generations rather than waiting for the phenotypic expression. DNA marker technology has been used in commercial plant breeding programmes since the early 1990s, and has proved helpful for the rapid and efficient transfer of useful traits into
agronomically desirable genotypes (Prasanna and Hoisington, 2003). DNA-based markers can be effectively utilized for three basic purposes: (i) identification of suitable parental lines for improvement with respect to a specific target trait; (ii) tracing favorable allele(s) (dominant or recessive) across generations; and (iii) identifying the most suitable individual(s) among the segregating progeny, based on allelic composition across a part or the entire genome.

‘Marker-assisted background selection’, a term coined by Hospital and Charcosset (1997), was proposed by Young and Tanksley (1989), and experimented by various researchers (e.g., Hospital et al., 1992; Ragot et al., 1995; Frisch et al., 1999). The optimal distance between the target gene and flanking markers governs the selection intensity that can be exerted. The equations given by Hospital et al. (1992) and Frisch et al. (1999) are helpful in determining the number of BC plants that need to be generated and genotyped with a special set of flanking markers. Microsatellite or Simple Sequence Repeat (SSR) markers are particularly useful in undertaking marker-assisted selection in crop plants like maize.

**Microsatellite or Simple Sequence Repeat (SSR) markers**

Details of various types of molecular markers have been provided by several workers (e.g., Karp et al., 1997; Prasanna and Hoisington, 2003). The emergence of PCR-based microsatellite assay in many crop plants, including maize, can be considered as an important milestone. Microsatellites or SSRs are stretches of tandemly arranged short sequence motifs (ranging from two to six nucleotides), which are abundant and highly polymorphic in several eukaryotic genomes, including maize. SSRs appear to be ubiquitous in higher organisms, although the frequency of microsatellites varies between species. They are abundant, dispersed throughout the genome and show higher level of polymorphism than other genetic markers. Studies on SSR patterns, distribution in the coding, non-coding or intergenic regions reveal species-
specific pattern or length dependent association (Powell et al., 1996; Morgante et al., 2002).

Microsatellites are genetically codominant and usually single locus. Because of high mutation rates, the SSR loci are more often highly polymorphic, leading to the detection of several alleles per locus if a high resolution system like PAGE or Sequencers are used. SSR markers are highly reproducible, robust, reliable, easy to use, and can minimize many of the difficulties associated with other marker types (Powell et al., 1996). However, the main drawback of this marker system is the effort and cost required to develop specific primer pairs for each polymorphic locus, which requires cloning and sequencing of a large number of genomic DNA fragments containing SSRs. DNA sequences flanking SSRs are known to be conserved, leading to the design of suitable primers for amplification of the SSR loci using PCR (Gupta et al., 1996). When SSR primers are used to amplify a particular SSR locus in a number of genotypes, polymorphism may be revealed in the form of differences in the length of amplified product, each length representing an allele at that locus. The length differences are largely attributed to the variation in the number of repeat units at a particular locus, which may vary from one genotype to another. Due to varying repeat number at a given microsatellite locus, the elements frequently change their length by slipped strand mispairing (Tautz, 1986). Mutational mechanisms such as unequal crossing over, replication slippage, transposition and DNA repair during double stranded breaks can leads to the genesis of SSRs.

A major difference between plant and animal systems is in the relative abundance of various kinds of microsatellites. In animals, microsatellites with CA/GT repeats are more common, while in plants AT/TA repeats are prominent followed by GA/CT repeats. In plants, SSRs with repeat units of 3-4 bases are generally not common and are clustered in specific regions of genome like centromeres and telomeres. Plant SSRs have been demonstrated to be a powerful tool in genotype identification and plant variety protection, seed purity evaluation and germplasm conservation (e.g., Brown
and Kresovich, 1996), diversity studies (e.g., Warburton et al., 2002), pedigree analysis and marker-assisted selection (e.g., Yang et al., 1994). About 1200 microsatellites have been so far mapped in maize and primers for nearly 600 loci are available in public domain. Information about these markers is available from the electronic database of MaizeGDB (http://www.maizegdb.org).

SSR markers have also been found useful in analysis of genetic diversity in QPM lines (Kassahun and Prasanna, 2003) and for facilitating the development of QPM lines (Lin et al., 1997; Prasanna and Hoisington, 2003). Kassahun and Prasanna (2003) analyzed a set of 23 QPM lines, including 13 lines developed in India and 10 lines at CIMMYT, Mexico using 36 polymorphic SSR markers. An opaque2-specific microsatellite marker, phi057, also facilitated differentiation of opaque2-carrying QPM inbreds from the non-opaque genotypes. Cluster analysis using SSR data, followed by canonical discriminant analysis, clearly distinguished the Indian QPM inbreds from those developed at CIMMYT. The cluster patterns were largely in congruence with the available pedigree information of the QPM inbreds studied.

One of the successful examples of MAS for maize improvement is the utilization of opaque2-specific SSR markers in conversion of the normal maize lines into QPM lines with enhanced nutritional quality (Prasanna et al., 2001; Morris et al., 2003; Babu et al., 2004, 2005). MAS-derived QPM hybrid, ‘Vivek QPM Hybrid 9’, developed by VPKAS (Almora), has been released in India in the year 2008. This QPM hybrid was developed through transfer of o2 gene and endosperm modifiers in two parental lines (CM145 and CM212) of ‘Vivek Hybrid 9’.

Maize Starch and Its Uses
Starch is the major component in the human diet, for which the Food and Agriculture Organization (FAO) and the World Health Organization (WHO) jointly recommend that 55-75% of daily food intake should come from carbohydrates (http://apps.fao.org). Cereal grains are major source of starch;
maize, in particular, is the third largest quantity of cereal grains grown worldwide. Progress in cereal starch production is especially important because these starches comprise 55 to 75% of daily human food intake and are the main source of food for domestic animals (Pan, 2000). Maize starch is a white or creamy powder having various direct and indirect applications in industries. Starch can be modified to suit individual requirement of the users and such starches are known as chemical starches/dextrin. The maize starch finds application in different industries as a binder, stabilizers, yarn sizing, thickening and suspending agent, viz. textile, paints, detergent, paper, ceramics, pharmaceuticals etc. Besides such direct industrial application of starch, it is also converted into a range of starch derivatives, viz., liquid glucose, dextrose monohydrate, dextrose anhydrous, maltodextrin, sorbitol etc., all of which find application in the food processing and pharmaceuticals industries. Other traditional industrial uses for starch include adhesives and papers (Hallauer, 2001). The possibility of starch-based degradable plastics and super-absorbent polymers are promising (Johnson et al., 2001).

Today’s food industry relies heavily on maize and its modified starches for thickening properties and low-temperature stability in freeze-thaw cycles of convenience foods. Furthermore, new and modified starches that do not require additional chemical modifications are highly desirable (White, 2001). Additionally, maize contains numerous mutants that provide a unique source of specialty starches (e.g. amylose-free waxy starches), which are important for many commercial applications (Hallauer, 2001), or the sugary mutants responsible for sweet corn production. The maize kernels are also important as raw material for biofuel or ethanol. Due to continuously increasing demands for biofuel, the share of grain maize used for starch production advanced significantly during the last couple of years, marked by recent breeding attempts to enhance efficiency in fermentation to ethanol (e.g., Dien et al., 2002).
**Starch Biosynthesis**

Starch is the most significant form of carbon reserve in plants in terms of the amount made. It consists of different glucose polymers arranged into a three dimensional, semi crystalline structure-the starch granule. The biosynthesis of starch involves not only the production of the composite glucans but also their arrangement into an organized form within the starch granule (Smith and Martin, 1993).

Starch is synthesized in leaves during the day from photosynthetically fixed carbon and is mobilized at night. It is also synthesized transiently in other organs, such as meristems and root cap cells, but its major site of accumulation is in storage organs, including seeds, fruits, tubers, and storage roots. Starch is synthesized in plastids which are called amyloplasts. These develop directly from proplastids and have little internal lamellar structure. Starch may also be synthesized in plastids that have other specialized functions, such as chloroplasts (photosynthetic carbon fixation), plastids of oilseed (fatty acid biosynthesis), and chromoplasts of roots such as carrot (carotenoid biosynthesis).

Starch consists of two types of glucan polymers: amylose and amylopectin. Amylose consists of predominantly linear chains of α (1-4)-linked glucose residues, each ~1000 residues long. Amylose is usually branched at a low level (approximately one branch per 1000 residues) by α-(1-6) linkages and makes up to 30% of starch. This proportion however, may vary considerably with the plant species and also with the plant organ, the developmental age of that organ, and, to some extent, the growth conditions of the plant (Shannon and Garwood, 1984). Detherage et al. (1954) found a range of 11 to 35% in a survey of 51 species and a range of 20-36% in a survey of 399 maize varieties. Amylopectin on the other hand consists of highly branched glucan chains, and makes up ~70% of starch. Chains of roughly 20 α(1-4)-linked glucose residues are joined by α(1-6) linkages to other branches. The branches themselves form an organized structure. The branches are not randomly arranged but are clustered at 7 to 10-nm intervals.
An average amyllopectin molecule is 200 to 400nm long (20 to 40 clusters) and 15 nm wide (Kalnuma, 1988; Smith and Martin, 1993). Within the starch granule, which may vary in size from <1 pm to >100 pm, the amyllopectin molecules are arranged radially, and adjacent branches within the branch clusters may form double helices that can be packed regularly, giving a crystallinity to the starch granule (Martin and Smith, 1995). The degree of branching and consequently the crystallinity of starch granules may vary considerably, even between different organs of the same plant.

The biosynthetic steps required for starch biosynthesis are relatively simple, involving three committed enzymes: ADPglucose pyrophosphorylase (ADPGPPase), starch synthase (SS), and starch branching enzyme (SBE). Both amyllose and amyllopectin are synthesized from ADP glucose, which is synthesized from glucose-1-phosphate and ATP in a reaction that is catalyzed by ADPGPPase and that liberates pyrophosphate. The glucose-1-phosphate can be supplied by the reductive pentose phosphate pathway in chloroplasts via phosphoglucoisomerase and phosphoglucomutase (Smith and Martin, 1993). In non-photosynthetic tissues, it may be imported directly from the cytosol (Tyson and Rees, 1988) or synthesized in the plastid from glucose-phosphate via the action of a plastidial phosphoglucomutase (Hill and Smith, 1991). In the next step of starch synthesis, SS catalyzes the synthesis of $\alpha$-(1-4) linkage between the non-reducing end of a preexisting glucan chain and the glucosyl moiety of ADPglucose, causing the release of ADP. The $\alpha$-(1-6) branches in starch polymers are made by SBE, which hydrolyzes an $\alpha$-(1-4) linkage within a chain and then catalyzes the formation of an $\alpha$-(1-6) linkage between the reducing end of the “cut” glucan chain and another glucose residue, probably one from the hydrolyzed chain. Branches are not created randomly, but show an average periodicity of 20 glucan residues.

**Maize Mutants affecting Starch Biosynthesis**

Maize mutants with abnormal endosperm phenotypes have contributed greatly to the understanding of starch synthesis (reviewed in Shannon and Garwood,
1984; Nelson and Pan, 1995) and facilitated the identification of many genes coding for starch biosynthetic enzymes. Maize genes involved in starch biosynthesis are *waxy* (wx), coding for granule-bound starch synthase I (GBSSI) (Shure et al., 1983; Klösgen et al., 1986), *amylose extender* (ae), coding for SBEIIb (Fisher et al., 1993; Stinard et al., 1993), *shrunken2* (sh2) and *brittle2* (bt2), coding for the large and small subunits of AGPase, respectively (Bae et al., 1990; Bhave et al., 1990), *brittle1* (bt1), thought to code for an adenylate translocator (Sullivan et al., 1991; Shannon et al., 1996; Cao and Shannon, 1997), and *sugary1* (su1), coding for the SDBE SU1 (James et al., 1995) (maize genetic nomenclature is according to Beavis et al., 1995). The details of function of the *sugary1* (su1) gene, analyzed in the present study, are presented below.

**sugary1 (su1) gene in maize:** The *sugary1* (su1) gene codes for the starch debranching enzyme (DBE) which leads to the accumulation of soluble sugars and water soluble polysaccharide (WSP) called phytoglycogen, ultimately resulting in a ‘sweet corn’ phenotype (James et al., 1995; Morris and Morris, 1939; Pan and Nelson, 1984). The *su1* mutant has pleiotropic effects on other starch biosynthetic enzymes (Doehlert et al., 1993) and results in high sucrose content in these mutants. Kernels homozygous for *su1* have a glassy, translucent and shrunken appearance at maturity resulting from their altered carbohydrate composition. The total amount of phytoglycogen reported to be present in *su1* kernel has range from 25-35% of total glcan (Creech, 1968; Shannon and Garwood, 1984).

Amylopectin in normal maize endosperm constitutes ~75% of starch, which comprises many linear chains of monomer joined by \(\alpha(1\rightarrow4)\) linkages these chains are joined to each other by \(\alpha(1\rightarrow6)\) glycosidic bonds catalyzed by starch debranching enzyme (DBE) (Kossman and Liyod, 2000; Myers et al., 2000). The non-random placement of branch linkages and determination of chain length results in a complex, hierarchical structure and renders the molecule both crystalline and insoluble. Two types of DBE are found in maize:
isoamylases, which hydrolyze amylopectin, glycogen, and phytoglycogen; and pullulanases, which hydrolyze pullulan, which is composed of α(1→6) maltotriosyl units (Doehlert and Knutson, 1991) and differ primarily in their substrate specificities. Both the types of DBEs directly hydrolyze α-(1-6) branch linkages, but differ in their activities toward specific polysaccharides.

Along with other hydrolytic enzyme DBEs plays a critical role in the germination of seed by breaking down starch into soluble sugars for developing seedling. However the main role of DBE is still amylopectin synthesis (Martins and da Silva, 1998; Revilla et al., 2000; Rahman et al., 1998). The role of pullulanase or isoamylase is yet to be made clear; however, both the DBEs are believed to be involved in degradation of endosperm starch after seed germination (Manners and Rowe, 1969; Toguri, 1991).

**Sweet Corn and Its Importance**
Sweet corn is one of the most popular vegetables and is gaining importance throughout the globe, including countries like India. Nearly all commercial sweet corn genotypes are based on one or more simple recessive alleles that alter carbohydrates content of the endosperm and having higher degree of sweetness. Until 1961, the sugary1 (su1) allele on chromosome 4 was extensively utilized for ‘sweet corn’ development. However, there are at least eight identified genes which affect carbohydrate synthesis affecting eating quality, appearance of ear and seed viability. Sweet corn having su1 gene existed in central and South America in the pre-Columbian period (Hendry, 1930). Despite its occurrence in Latin America, sweet corn was not widely grown, nor was it used for roasted ears. Sweet corn was considered too gummy when consumed in this fashion while it was highly prized for other uses.

Sweet corn quality is determined by the flavor, aroma and texture of the endosperm and tenderness of pericarp, which is affected by amount of sugar and starch in the endosperm. Starch synthesis mutants may be divided into two classes based on their effects on endosperm composition (Boyer and
Shannon, 1984). The Class I mutants - *brittle1 (bt1), brittle2 (bt2) and shrunken2 (sh2)* accumulate sugars at the expense of starch and have greatly decreased total carbohydrates at the mature seed stages. Due to higher sugar levels, Class I mutants may be used independently of other carbohydrates mutants in sweet corn varieties. Due to elevated sugar levels, varieties of these genotypes are often called ‘super sweet’ or ‘extra sweet corn’. The Class II mutants are *amylose extender1 (ae1), dull1 (du1), sugary1 (su1) and waxy1 (wx1)* which alter the types and amounts of polysaccharides produced. These alleles lead to slightly less starch in the mature kernel than non-mutant types (Boyer and Shannon, 1984) with smaller increase in total sugar content at 21 days after pollination relative to class I mutants and do not make acceptable sweet corn when used singly. Not only do Class I mutants start out with significantly higher levels of sugars at harvest than does *su1*, but the decline in sugar level is much slower even without refrigeration. Due to extended shelf life, varieties based on Class I mutants are better suited for long distance shipping as well as the habits of modern consumers.

Many combinations of endosperm genes are possible and some have been used commercially in sweet corn development. The most common type of gene combination is partial modification based on two recessive genes. Kernels in which both mutations are expressed usually have higher sugar content. The *su1su1 Sh2sh2* and *su1su1 Se1se1* combinations are most common although there are commercial hybrids in which the F₁ is *Su1su1 sh2sh2*. Super sweet corn hybrids, in general, have poorer germination than do *su1* hybrids, and seeds of super sweets hybrids are more difficult to produce (Marshal, 1987). However, the high sugar mutants generally have reduced field emergence and seedling vigour, Acceptance of *sh2* varieties was delayed due to these problems as well as differences in texture and tenderness.

While single genes, such as *su1, sh2*, have major effects on carbohydrate composition in the endosperm, genetic back ground also affects carbohydrate composition. Wong et al. (1994) examined chemical components
of quality of 24 commercial sh2 hybrids and found extensive variation for sucrose, total sugar levels and other traits. They concluded the variation was due to allelic variation at loci other than sh2. Flora and Wiley (1974) found seven compounds that produced aromas; many of these compounds are repulsive in pure state but in combination produced a pleasant aroma. Dimethyl sulfide seems to be a dominant compound determining the aroma of cooked corn. Tenderness is determined by the resistance of the pericarp to chewing and is negatively correlated with the pericarp thickness (Bailey and Bailey, 1938, Ito and Brewbaker, 1981). The number of gene determining the pericarp thickness is unknown, although in some of the crosses, a major gene appears to condition thinness with modification for thickness by a number of minor genes.

There are two main theories proposed for origin of modern sweet corn (Erwin, 1934, 1942). The first proposes that modern sweet corn is descendent from maize landraces 'Dulce' and 'Chullpi', while the second suggests that North American sweet corn is the recent origin resulting from a mutation to su1 in the field corn.

**Combining Ability and Heterosis in Specialty Corn**

The capacity or ability of a genotype to transmit superior performance to its crosses is referred to as ‘combining ability’. The combining ability values of an inbred line depend on its ability to produce superior hybrids in combination with other inbreds (Sprague and Tatum, 1942). Analysis of combining ability assumes importance in inbred line evaluation and population improvement, particularly in maize hybrid breeding (Hallauer and Mirinda, 1988; Crossa et al., 1990).

The total combining ability of lines was portioned into general combining ability (GCA) and specific combining ability (SCA) by Sprague and Tatum (1942). While GCA was described as the average performance of the line in hybrid combinations, SCA is the parameter to judge which specific hybrid combinations are superior or inferior with respect to the average
performance of other tested lines. GCA was interpreted as an indication of
genes having largely additive effects and SCA as indicative of dominance
effect. The information about the components of genetic variance is provided
by the estimates of combining ability, and thus helps in selection of desirable
parents for hybrid breeding.

The Line x Tester (L x T) design is one of the most important mating
designs that facilitates estimation of the combining ability effects and also
partitioning of the genetic variance. It is a modified form of top cross where
only one tester is used. While in the L x T system, several testers are used as
male parents and crossed with a set of inbreds as females (Sharma, 1998). L
x T is useful in analyzing the relative ability of number of female and male
inbreds to produce desirable hybrid combinations. It also provides information
about the usefulness of male and female inbreds as parents for hybridization
and to generate segregating population. However, very little work has been
published on combining ability of sweet corn lines for total sugar content as
well as estimation of heterosis.

Widstrom and Davis (1967) evaluated a diallel set involving five inbred
lines of sweet corn for earworm injury at Lafayette, Indiana. Diallel analysis for
kernel carbohydrates was undertaken by Rosenbrook and Andrew (1971),
revealing highly significant GCA variances and non significant SCA variances
for reducing sugars, sucrose, and water-soluble polysaccharides and highly
additive gene action for carbohydrate fractions.

Younes and Andrew (1978) analyzed the inheritance of prolificacy with
a set of diallel series of market sweet corn hybrids and determined its
association with productivity; additive gene action was found to play an
important role. Andrew and Elbe (1979) evaluated a set of hybrids of high
sugar corn to determine the processing potential of sh1 sh2 germplasm.
Highly significant GCA and SCA mean squares indicated the importance of
both additive and non additive gene action and suggested the incorporation of
sh1 and sh2 gene combination in sweet corn breeding to increase their
processing potential and germination viability.
Analyzing ten hybrids from a diallel cross of five shrunk2 sweet corn inbreds, Michaels and Andrew (1986) reported that the accumulation of sucrose and total sugar in each hybrid was depended on the combination of inbreds involved in the cross. It was also reported that hybrid sharing a common inbred had similar levels of total sugar accumulation. Hence both the characters were influenced by the general combining effects.

Kumari et al. (2008) estimated the combining ability and heterosis for field emergence, kernel quality traits and yield component in sweet corn genotypes. A line x tester mating design was used involving seven field corn and four sweet corn inbred lines which led to the identification of inbred lines as well as experimental cross combinations for grain yield, total soluble solids (TSS), total as well as non-reducing sugars.

Hybrid vigour is one of the greatest practical contributions of genetics to the agricultural world. Maize has been at the forefront in the basic, strategic and applied research on heterosis and hybrid breeding. The term ‘heterosis’ was coined by Shull (1914) and was initially defined as the superiority of F1 hybrids over its parents. Shull (1952) further elucidated heterosis concept as the increase in vigour, size, fruitfulness, speed of development of resistance to disease and insects, or to climatic vagaries of any kind manifested by crossbred individuals as a result of unlikeness in the constitution of uniting parental gametes.

The manifestation of heterosis usually depends on genetic divergence of the parental types. Establishment of heterotic patterns among parental lines in a crop like maize has important implications for selecting inbreds for hybrid development (Hallauer and Mirinda 1988). The phenomenon of heterosis has been exploited extensively in maize breeding starting with the single-cross, double-cross, three-way cross, and coming back to the single-cross hybrid development (Dhillon and Prasanna, 2001). The development and wide spread cultivation of maize hybrids can justifiably be called one of the greatest accomplishments of plant breeding.
In India, sweet corn composites such as Madhuri, Priya and Win Orange have been bred and released by the public sector institution and are popular among the breeder and farming community. However, these populations like other composites are low in terms of yield potential as compared to hybrids. Therefore, it is important to develop sweet corn hybrids with high yielding potential and high degree of sweetness.

**Nucleotide Diversity Analysis of sugary1 Gene in Maize**

The diverse germplasm collections available with national and international research institution are ‘gold mines’ for analysis of allelic diversity (Hamilton and McNally, 2005). Using such germplasm accessions, researchers can discover a wide array of novel favorable alleles. ‘Allele mining’ involves PCR-based amplification and sequencing of different versions of gene found in inbred line or pure lines, varieties, landraces and wild relatives. Variation in gene sequence is then correlated with trait or performance of the accession, to enable identification of the favorable alleles for future experiments. The whole strategy involves an intensive series of genetic analysis of carefully selected set of genotypes (preferably ‘core collections’) and putative candidate genes that are expected to contribute to phenotypic diversity (Prasanna, 2007).

Different worker used different strategy for mining allele, one of the most powerful strategy for allele mining which was recently developed is based on TILLING (Targeting Induced Local Lesion in Genome) a reverse genetic approach developed at University of Washington and Fred Hutchinson Cancer Research Center by group of Drs. Steven Henikoff and Luca Comai (Colbert et al., 2001; Till et al., 2003) This process allows detection of an allelic series of point mutations at any gene of interest in a genome. An extension to TILLING is EcoTILLING, which allows natural alleles at a locus to be characterized across naturally occurring germplasm accessions, enabling both SNP discovery and haplotyping at these loci (Comai et al., 2004; Gilchrist and Haughn, 2005)
The accessibility to the genome sequences as well as sequences of specific genes in several organisms enabled the study of sequence variations between individuals, cultivars, and subspecies. These studies revealed that single nucleotide polymorphisms (SNPs) and insertions and deletions (InDels) are highly abundant and distributed throughout the genome in various species including plants (Garg et al., 1999; Drenkard et al., 2000; Nasu et al., 2002; Batley et al., 2003a). The abundance of these polymorphisms in plant genomes makes the SNP marker system an attractive tool for mapping, marker-assisted breeding and map-based cloning (Gupta et al., 2001; Rafalski, 2002a; Batley et al., 2003b). As suggested by the acronym, a SNP marker is just a single base change in a DNA sequence, with a usual alternative of two possible nucleotides at a given position. All methods for SNP genotyping combine two elements: first, the generation of an allele-specific product, and second the analysis thereof. SNPs are increasingly becoming the marker of choice in genetic analysis and are used routinely as markers in plant breeding programmes (Gupta et al., 2001). The low mutation rate of SNPs also make them excellent marker for studying complex genetic trait and as tool for understanding of genome evolution (Syvanen, 2001).

In maize, SNPs have been found more frequently, with one SNP approximately every 48 bp and every 130 bp in 3'-untranslated region respectively (Remington et al., 2002; Tenaillon et al., 2001). Mogg et al. (2002) amplified and sequenced the flanking regions of 97 previously characterized microsatellite primer set in 11 maize inbred lines. The sequencing results indicated that the flanking region of maize microsatellite loci show increased levels of polymorphisms when compared to other regions of genome, with SNPs in these regions found on an average every 40 bp. Therefore, essential and useful variation for the grain quality (starch composition) needs to be found out in crops like maize.

Starch production is critical to both the yield and the quality of the grain. So far, more than 20 genes were identified to affect starch production in maize. Among these, nucleotide sequence variations in six candidate genes,
amylase extender (ae1), brittle2 (bt2), shrunken1 (sh1), shrunken2 (sh2), sugary1 (su1), waxy (wx1) have been recently analyzed in a set of American inbred lines by Whitt et al. (2001) for association with starch concentration and starch composition quality. The nucleotide diversity surveyed across six major loci involved in starch metabolism demonstrated usually low genetic diversity and strong evidence of selection. Strong selection has been reported at starch branching and debranching enzymes (coded by ae1 and su1, respectively). This suggested that the amylopectin structure, and thereby pasting properties, were the key targets of selection (Whitt et al., 2001). Candidate gene association analysis revealed that bt2, sh1 and sh2 showed significant association for kernel composition traits, whereas ae1 and sh2 showed significant associations for starch pasting properties. Haplotype analysis of sh2 suggested that this gene is involved in starch viscosity properties and amylase content (Whitt et al., 2004).

‘Association genetic analysis’ or Association mapping’ makes use of genomic surveys of linkage disequilibrium (LD). Originally developed for human genetics, statistical methods and their derivatives for detection of LD are now increasingly being applied to crop plants like maize, leading to analyses of population genetic structure and QTL detection. The candidate gene approach, using a Bayesian model-based probabilistic clustering, implemented through the STRUCTURE software, was first utilized for associating Dwarf8 polymorphisms with flowering time variation in maize (Thornsberry et al., 2001).

Candidate gene association mapping aims at relating polymorphism in the selected gene that have purported roles in controlling phenotypic variation for specific traits. While genome-wide association mapping approach surveys genetic variation in the whole genome to find signals of association for various complex traits (Risch and Merikangas, 1996), the candidate gene association mapping requires the identification of SNPs between lines and within specific genes. Therefore, the most straightforward method of identifying candidate gene SNPs relies on the re-sequencing of amplicons from several genetically
distinct individuals of a larger association population. Fewer diverse individuals in the SNP discovery panel are needed to identify common SNPs, whereas many more are needed to identify rarer SNPs. In a recent candidate-gene association mapping study, lycopene epsilon cyclase (lcyE) locus has been identified to alter flux down alpha-carotene versus beta-carotene branches of the carotenoid pathway among diverse maize inbred lines (Harjes et al., 2008).

Advances in high-throughput genotyping and sequencing technologies have markedly reduced the cost per data point of molecular markers, particularly single nucleotide polymorphisms (SNPs) (Hirschhorn and Daly, 2005; Syvanen, 2005). Obviously, association mapping is one approach that heavily leverages these emerging genomic technologies, with sequencing, re-sequencing, and genotyping as the intermediate steps to the final goal of linking functional polymorphisms to complex trait variation.