CHAPTER ONE

1.0 INTRODUCTION

1.1 BACKGROUND TO STUDY

Vitamin A deficiency (VAD) affects millions of poor people in sub-Saharan Africa. VAD results from inadequate dietary intake of vitamin A as either preformed vitamin A (retinoic acid) from animals or as pro-vitamin A (retinol) from plants. VAD constitutes a major public health problem. About 28-35 % of children in sub-Saharan Africa are vitamin A deficient (WHO, 2002). In the developing world, it has been estimated that about 140 million preschool-aged children and more than 7 million pregnant women suffer from VAD every year. In many cases, 1.2-3 million children and significant numbers of women die, while 4.4 million children and 6.2 million women are affected by xerophthalmia (West, 2002).

Children affected by deficiency in vitamin A are most prone to diseases such as measles, malaria, diarrhoea, respiratory infections and many other childhood diseases, contributing to increased child mortality among less than five years of age (Sommer *et al.,* 1980; WHO, 1995; Aguayo *et al.,* 2004: Aguayo and Baker, 2005). Other clinical manifestations are growth retardation, depressed immune response, disturbed cellular differentiation and impaired iron mobilization in children. Women of reproductive age also suffer from VAD related diseases (Christian *et al.,* 1998; Christian *et al.,* 2000). These two risk groups have been generally affected by night blindness, corneal ulceration and keratomalacia due to VAD, which could result in total blindness (Christian *et al.,*

1998; Christian *et al.*, 2000; Sommer, 2008). VAD is prevalent where cereal-based diets are consumed.

The need to address vitamin A deficiency among pre-school aged children and young African women has been highlighted at various times by the World Health Organisation. Food fortification and supplementation with vitamin A capsule have been used as strategies to control VAD in developing countries (West, 2002; Sommer and Davidson, 2002; Lartey, 2008; UNICEF, 2009). Infants in the age group 6-36 months have been given priority to receive vitamin A supplements (preformed vitamin A) when they are likely to get insufficient vitamin A from the breast milk. Similar effort was also given to children up to 5 years of age (WHO, 1995). Women of child-bearing age also received a dose of vitamin A (200,000 I.U) around the time of birth to increase the vitamin A content of breast milk (WHO, 1995). Supplementation was a short-term strategy which had impact, however, mid and longer-term interventions were sought. Fortification was carried out by adding vitamin A in all processed foods that are commonly eaten among the people. Some of these foods include cooking oil, margarine, flour and sugar, salt and to ensure that this is achieved a logo of an eye is put on every fortified food package. However, sustaining these programmes was not achievable on a long term.

The availability of plant transformation method has offered another alternative means of producing nutritionally enhanced crops through the metabolic engineering of some important genes in the β-carotene biosynthetic pathway into some staple foods like rice (Ye *et al.*, 2000; Beyer *et al.*, 2002) and maize (Naqvi *et al.*, 2009; Aluru *et al.*, 2008).

Thus, transgenic crops that accumulate higher β-carotene have been developed. Nevertheless, some problems such as acceptability and adoption of transformed varieties as breeding materials among the sub-Saharan farmers may create hindrances for its usage. In addition, acceptance of transgenic food and safety concerns among the sub-Saharan people might limit the use of genetically modified foods. Therefore, to produce significant increase in the daily intake of vitamin A to achieve a long lasting solution to VAD, it is necessary to tap into the naturally occurring carotenoids in important food crops such as maize.

Biofortification is food-based approach that may likely contribute to the reduction of the scale and severity of VAD among people in sub-Saharan Africa (Howe and Tanumihardjo, 2006; Low *et al.,* 2007). Biofortification is an international effort to combat micronutrient deficiency in developing countries (Nestel *et al.,* 2006). It allows staple foods that are locally available to be improved for high pro-vitamin A carotenoids through conventional breeding, which can lead to the development of new pro-vitamin A enriched maize varieties (Kurilich and Juvik, 1999). Since β-carotene is the main dietary precursor of vitamin A, the source for vitamin A can be through carotenes. Increasing the level of pro-vitamin A carotenoids in maize may likely contribute to the reduction of the scale and severity of vitamin A deficiency among people in sub-Saharan Africa (Howe and Tanumihardjo, 2006; Nestel *et al.,* 2006; Low *et al.,* 2007). This approach is potentially sustainable, affordable, effective and feasible method to provide solution to nutrient deficiency in the developing countries of the world, to improve nutrition and human health (Neumann, 2007).

Maize (*Zea mays*) is one of the most important cereal crops and a staple food of the sub-Saharan African people (http://faostat.fao.org). It is a common cereal in the world after rice and wheat, with an annual production of 600 million tonnes (FAO, 2002). About 64 % of the world's maize area and 43 % of global maize production are accounted for by the developing countries (Morris, 2001). In sub-Saharan Africa, maize is widely grown for consumption and also for generating income for peasant farmers. Traditionally, it has been part of the people's diet, 85 % of the maize grown is used directly as human food and average maize consumption in Africa is 106.2 g/person/day (WHO, 2003). Maize is one of the most diverse crops in phenotype and genotype. Maize is relatively cheap compared with other staple cereals, such as rice, sorghum and millet.

Maize has become the most important staple food crop for many people in sub-Saharan Africa and it has been part of the diets of the people in different forms such as porridge and fermented food products. At present, the adapted yellow maize varieties in Africa are low in pro-vitamin A. The yellow maize endosperm contains varying amounts of pro-vitamin A carotenoids which include α-carotene, β-carotene, β-cryptoxanthin, but the concentrations are very low (Kurilich and Juvik, 1999). Therefore, researches on genetic improvements are needed for high pro-vitamin A carotenoids in maize (Toenniessen, 2002). Early research has indicated that maize carotenoid content in the grain vary considerably and breeding maize for high pro-vitamin A is possible (Hauge and Trost, 1928; Brunson and Quackenbush, 1962; Simpson, 1983). Breeding of high β-carotene maize lines (about 13-15 µg/g of pro-vitamin A) can be reached (Kurilich and Juvik,

1999; Harjes *et al*., 2008). More recently, genetic variation for specific carotenoid content has been reported in maize lines adapted to the tropics (Menkir *et al.,* 2008; Harjes *et al.,* 2008).

Menkir *et al.* (2008) measured the carotenoid contents of tropical adapted yellow maize inbred lines from different genetic background and reported a broad range of variation in β-carotene and pro-vitamin A content. These inbred lines were developed from diverse tropical adapted crosses and backcrosses at the International Institute of Tropical Agriculture (IITA), Ibadan, Oyo State, Nigeria and the diversity in concentration of carotenoids found in these lines offers the possibility to breed for high pro-vitamin A content. Among the yellow endosperm maize inbred lines, some lines were chosen with varying levels of carotenoids for the present study. These lines represent a good genetic base to breed tropical maize for higher level of pro-vitamin A. Selection of genetically diverse parents for breeding may provide a diverse array of alleles for carotenoids in tropical maize. However, the extent of genetic diversity present in this set of maize inbred lines has not been studied using molecular markers.

Knowledge of genetic diversity among genotypes of any crop is essential to estimate the potential of genetic gain in a breeding program. It may also be useful to identify parents for making crosses and establishment of heterotic groups to develop yellow endosperm maize hybrids (Russell *et al.,* 1997). Currently, various molecular markers are available for assessment of genetic diversity among genotypes (Smith and Smith, 1992; Legesse *et al.,* 2007). Such markers have been used not only for assessing the extent of genetic

diversity present in breeding materials (Livini *et al.*, 1992; Messmer *et al.*, 1992; Senior *et al.,* 1998; Reif *et al.,* 2003a) but also analysis of important quantitative traits (Melchinger *et al.,* 1990, 1991; Smith *et al.,* 1990, 1991).

Among the PCR-based markers that are available, simple sequence repeat (SSR) and amplified fragment length polymorphism (AFLP) have been extensively used for genetic diversity assessment (Akkaya *et al.,* 1992; Smith *et al.,* 1997; Heckenberger *et al.,* 2002; Warburton *et al.,* 2002; Gethi *et al.,* 2002; Reif *et al.,* 2003a; Liu *et al.,* 2003; Adetimirin *et al.,* 2008; Mohammed *et al.,* 2008). SSR motifs are 2, 3, or 4 nucleotides that are found in abundance in the genomes of eukaryotic plant species and these units are tandemly repeated many times in the DNA sequence (Hamada *et al.,* 1982; Morgante and Olivieri, 1993). AFLP are polymorphic markers generated as a result of single nucleotide changes, resulting from deletions, insertions and rearrangements in the restriction sites and/or adjacent sequences (Janssen *et al.,* 1996). The two markers will thus allow detection of polymorphisms in different regions of the genome, which can result in complementary genetic distance estimates.

Several studies have compared the use of AFLP and SSR markers for diversity assessment (Pejic *et al.,* 1998; Lubberstedt *et al.,* 2000; Heckenberger *et al.,* 2003; Garcia *et al.,* 2004; Laborda *et al.,* 2005; Allan *et al.,* 2007). Some of the comparative studies in maize showed good agreement between the patterns of diversity detected by the two markers (Pejic *et al.,* 1998; Lubberstedt *et al.,* 2000; Heckenberger *et al.,* 2003) while others did not find significant correlation between the genetic distance matrix generated by the two markers systems (Laborda *et al.,* 2005; Garcia *et al.,* 2004). The genetic relationships revealed by the two molecular markers have been consistent with known

pedigree data in maize lines (Melchinger *et al.,* 1990; Senior *et al.,* 1998; Pejic *et al.,* 1998). Kernel carotenoid composition of yellow endosperm maize inbred of tropical-adapted lines can also be used to assess the genetic relationship among the lines. Only some studies have used molecular markers and their comparison with biochemical based data for genetic diversity study (Babu *et al.,* 2009).

Single nucleotide polymorphism (SNP) such as insertions/deletions (indels) has been used in the genetic diversity studies of loci controlling important traits of agronomic importance in plants (Bhatttramakki *et al.,* 2002). SNPs are highly informative, abundant and have potential to provide the variability that can be used to distinguish alleles. They are thought to contribute to phenotypic differences (Bhatttramakki *et al.,* 2002). SNPs have been used to study diversity in loci such as *y1* (Phelps *et al.,* 1996), *y1, psy2* (Palaisa *et al.,* 2003), *Adh1* locus (Jung *et al.,* 2004), six major genes involved in starch metabolism (Whitt *et al.,* 2002) and 21 loci along chromosome 1 (Tenaillon *et al.,* 2002) in maize. *PSY1* gene (*Y1* gene) represented the first gene that was shown to possess sequence variation for β-carotene accumulation in maize (Palaisa *et al.,* 2003).

Most agronomic traits are genetically controlled by multiple loci and their effects have been estimated using linkage analysis or quantitative trait loci (QTL) studies in plant species (Wong *et al.,* 2004; Chander *et al.,* 2008). Natural allelic variation within genes may bring about the phenotypic differences observed among genotypes of species. The discoveries of candidate genes in the maize carotenoid biosynthesis pathway capable of enhancing accumulation of beta-carotene in maize grain have been achieved. Recently, by

screening natural variability, allelic variations in two candidate genes (*LCYE* and *crtRB1*) related to pro-vitamin A accumulation in endosperm of maize have been studied (Harjes *et al.,* 2008; Vallabhaneni *et al.,* 2009; Yan *et al.,* 2010). This resulted into further characterisation of favourable and unfavourable alleles in the genes involved in variation of β-carotene content. The *LCYE* expression was identified to have a partitioning effect on the two branches (β, Ɛ- leading to synthesis of lutein and β, β- leading to synthesis of zeaxanthin) of the carotenoids and it has four principal functional polymorphism sites (Harjes *et al.,* 2008). Further, rare genetic variations were also found in β-carotene hydroxylase 1 gene (*crtRB1*), also known as *HYD3* and have been shown to affect the increase in accumulation of carotenes (Yan *et al.,* 2010). Three polymorphisms have been identified for the (*crtRB1*) and they have been shown to affect conversion of β-carotene to zeaxanthin. Moreover, other alleles in *HYD3* gene have been shown to have significant correlation with β-carotene content in the maize endosperm and a polymorphism was identified.

The β-carotene genes (*LCYE* and *crtRB1*) have been completely sequenced from some maize inbred lines and nucleotide sequence polymorphisms have been elucidated. Allelic sequence differences between inbred lines were shown to exist and are due to polymorphisms which include single nucleotide polymorphisms (SNPs) and insertions/deletions (INS/DEL). SNP markers or gene-based polymerase chain reaction (PCR) marker sets linked to β-carotene genes have been designed (Harjes *et al.,* 2008; Vallabhaneni *et al.,* 2009; Yan *et al.,* 2010) to amplify and score SNPs or INS/DEL in *LCYE* and *crtRB1* in maize inbred lines. These markers are powerful tools for screening

and selection of tropical maize inbred lines having favourable *LCYE* and *crtRB1* alleles useful for marker-assisted selection (MAS) and introgression in maize breeding program. This will further be used in breeding of biofortified tropical inbred lines with increased level of pro-vitamin A. Thus, new maize inbred lines with higher level of pro-vitamin A can be bred. For this to be achieved, it is important to have the knowledge of the genetic diversity among the tropical-adapted yellow endosperm maize inbred lines and to identify the best inbred lines having favourable alleles at the two major genes (*LCYE* and *crtRB1)* for β-carotene trait.

1.2 STATEMENT OF PROBLEM

Vitamin A deficiency is widespread in the sub-Saharan continent in the world where maize is a staple food. Majority of the poor sub-Saharan Africans subsists on large amounts of white endosperm maize varieties which contain no pro-vitamin A, while the current varieties of tropical-adapted yellow endosperm maize varieties contain low amounts of vitamin A precursors [β-carotene (0.5 to 1.5 µg per g), *α*-carotene and β-cryptoxanthin]. These situations have contributed to VAD among pre-school aged children, pregnant and lactating women in these parts of the world (Harjes *et al.*, 2008). Many pre-school aged children have developed eye problems like xerophthalmia and many young women have night blindness disorder and corneal lesions among others. It is estimated that 50 % of diagnosed children with blindness die within 12 months of losing their sight because of insufficient intake of vitamin A. The problem of VAD has been recognised, and over the past years, various vitamin A control programs have been used in order to reduce VAD in sub-Saharan countries.

Most people living in the developing countries do not have the financial capability to afford plant foods (green vegetables, palm kernels) and animal derived products (fish, eggs, oils, milk) that contain natural vitamin A precursors. They depend on cereals like yellow maize grains that contain naturally occurring carotenoids. It is only recently that research has been done to determine the carotenoid contents among tropical-adapted yellow endosperm maize inbred lines (Menkir *et al.,* 2008) and sufficient genetic variability that would facilitate breeding for higher the beta-carotene was found. However, increase in the pro-vitamin A content of maize grain through an approach referred to as biofortification has not been employed for tropical adapted yellow endosperm maize inbred lines. For this to be achieved molecular markers are needed to assess the genetic variability among the lines.

The genetic diversity present among yellow endosperm maize inbred lines adapted to the tropics have not been assessed using DNA markers. Also, validation of functional allelic variant markers at the candidate genes for β-carotene genes required for breeding increased β-carotene level and characterisation of the allelic variants of the genes in these inbred lines have not been carried out. Because of the widespread prevalence of VAD in the sub-Saharan Africa (WHO, 2009), this study is necessary to enable breeders to effectively produce maize grain with higher pro-vitamin A.

1.3 SIGNIFICANCE OF STUDY

1. It will be useful to determine the level of genetic diversity that exists within tropical adapted yellow endosperm maize inbred lines and to examine genetic relationship among them for their efficient utilisation as breeding materials.
2. The present research will facilitate the selection of superior maize lines for crossing to improve the pro-vitamin A content in maize.
3. The knowledge gained in this work will be incorporated in breeding programs to develop new maize inbred lines with enhanced beta-carotene level.
4. This study is important because it will potentially help to reduce death and diseases caused by vitamin A deficiency among the pre-school aged children and young African women in sub-Saharan Africa and other developing tropical countries.

1.4 PURPOSE/OBJECTIVES OF STUDY

1.4.1 Purpose of study

The purpose of this research was to apply molecular marker-based techniques to understand the extent of genetic diversity and relationship among tropical adapted yellow endosperm maize inbred lines. Also, to use PCR-based markers to validate marker trait for β-carotene among lines for selecting favourable alleles of *LCYE* and *crtRB1* genes which are candidate genes in the carotenoid pathway for breeding of maize genotypes with higher level of β-carotene.

1.4.2 OBJECTIVES OF STUDY

1. Assessment of extent of DNA marker-based (Amplified Fragment Length Polymorphism, Simple Sequence Repeat and Single Nucleotide Polymorphism) genetic diversity.
2. Assessment of relationship between marker based grouping and carotenoid concentration based grouping.
3. Validation of alleles and characterisation of allelic variants of candidate genes (*LCYE* and *HYD3* (*crtRB1*)).
4. Determination of allelic diversity of genes involved in carotenoid biosynthesis.
5. Use the SSR-based grouping among yellow endosperm maize inbred lines as the basis to assess the allelic diversity of genes involved in carotenoid biosynthesis.

1.5 RESEARCH QUESTIONS

1. What is the extent of genetic diversity among yellow tropical-adapted yellow endosperm maize inbred lines?
2. Is there any correlation between grouping based on DNA markers and carotenoid grouping?
3. Can the gene-based PCR markers be validated in tropical maize inbred lines?
4. How many yellow tropical-adapted yellow endosperm maize inbred lines have favourable alleles necessary for breeding increased beta-carotene content?
5. Can the SSR-based grouping among 122 yellow endosperm maize inbred lines be used as basis to assess the allelic diversity of genes involved in carotenoid biosynthesis?

1.6 OPERATIONAL DEFINITION OF TERMS AND ABBREVIATIONS

1.6.1 Terms	Biological definitions
ABI 310 Genetic Analyzer:	A capillary electrophoresis instrument sold by Applied Biosystems.
Allele:	An alternative form of a gene or a section of DNA at a particular genetic location (locus), typically multiple alleles.
Amplification:	An increase in the number of copies of a specific DNA fragment.
Dendrogram:	Any branching diagram that shows, by means of lines shaped like U's a hierarchy of categories or objects based on the degree of similarity or number of shared characters. Often, the length of each U represents the distance between the two objects being connected.
Diversity analysis:	A study undertaken to classify an individual or population or species

	compared to other individual or populations or species.
DNA sequence:	The relative order of base pairs, whether in a fragment of DNA, a gene, a chromosome, or an entire genome.
Electrophoresis:	A technique for separating the components of a mixture of molecules (proteins, DNA or RNA) by size as a result of an electric field within a support gel.
Gene:	The basic physical and functional unit of heredity, which passes information from one generation to the next.
Genetic distance:	The difference between two entities that can be described by allelic variation or extent of gene differences between populations or species that is measured by some numerical quantity or any quantitative measure of genetic difference be it at the sequence level or allele frequency level, which is

	calculated between individuals, populations or species.
Genotype:	The specific allelic composition for a certain gene or set of genes.
Haplotype:	A specific allelic constitution at a number of loci within a defined linkage block.
Heterosis	The tendency of a crossbred individual to show qualities superior to those of both parents.
Insertion:	A type of chromosomal abnormality in which a DNA sequence is inserted into another DNA sequence.
Kilobase:	Kb a unit of length with 1000 bases in DNA or RNA.
Marker assisted selection (MAS):	A tool in modern plant breeding programs. It is an indirect selection using molecular genotyping methods to enable detection of desired alleles and haplotypes early in the plant life cycle.

Marker:	A polymorphic genetic character used to follow the transmission of a chromosomal segment in a family or population.
Microsatellite:	A DNA variant due to tandem repetition of a short DNA sequence usually two to four nucleotides.
Multiplex PCR:	Co-amplification of multiple regions of a genome with more than one set of primers; enables information from the different target sequences to be collected simultaneously.
Nucleotide:	A unit of nucleic acid composed of phosphate, ribose or deoxyribose, and a purine or pyrimidine base.
Oligonucleotide:	A short segment of DNA that is synthesised artificially.
Pedigree:	A conventional representation of a family tree constructed using a standardized set of symbols.
Polymerase chain reaction (PCR):	An in vitro process that yields millions of copies desired DNA through

	repeated cycling of a reaction involving the DNA polymerase enzyme.
Primer:	A short DNA or RNA fragment annealed to a single-stranded DNA and to which further nucleotides can be added by DNA polymerase.
Single nucleotide polymorphism (SNP):	Any polymorphic variation at a single nucleotide; most SNPs are biallelic (e.g., either C or T).
Template:	A molecule that serves as the pattern for synthesising another molecule, e.g. a single-stranded DNA molecule can be used as a template to synthesise the complementary nucleotide strand.
Validation:	The process by which a sample, measurement method, or a piece of data is deemed useful for a specified purpose; the process of extensive and rigorous evaluation of DNA methods before acceptance for routine use.

1.6.2 ABBREVIATIONS

AFLP	Amplified fragment length polymorphism
BC	Back-cross
Bp	Base pairs
CAPS	Cleaved Amplified Polymorphic Sequences
DNA	Deoxyribonucleic acid
DEL	Deletion
°C	Degree centigrade
crtRB1	Beta-hydroxylase 1 gene
dNTP	Deoxynucleotide 5’ triphosphate
HYD3	Hydroxylase 3 gene
INS	Insertion
kb	Kilo base pair
LCYE	Lycopene epsilon cyclase gene
M	Molar
MAS	Marker-assisted selection
mg	Microgram
mg	Microgram
min	Minute
µl	Microlitre
ml	Millilitre
µM	Micromolar

mM	Millimolar
ng	Nanogram
PCR	Polymerase chain reaction
PIC	Polymorphic Information content
PSY	Phytoene synthase gene
QTL	Quantitative trait loci
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
rpm	Revolution Per Minute
s	Second
SSR	Simple sequence repeat
STS	Sequence-Tagged Site
TBE	Tris-borate-EDTA buffer
TE	Tris-EDTA buffer
UPGMA	Unweighted pair-group method with arithmetic averages

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Yellow maize crop: A good dietary source of pro-vitamin A

Maize is a staple food crop for many in sub-Saharan Africans including children and adults. The sub-Saharan region contributes an estimated 328 kilocalories per capita per day. Of the 140 million hectares cultivated for maize production across the world, approximately 96 million hectares are in the developing world. The production of maize is likely to increase globally by 50 %, with up to 79 % in developing countries and up to 93 % in Asia and Saharan Africa until 2020 (http://www.cimmyt.org/Research/Economics/map/facts_trends/maizeft9900/pdfs/maizeft9900/pdf). There is now more demand for maize as food among the people living in the tropical and sub-tropical parts of Africa because the level of poverty is not abating while the population is increasing.

In sub-Saharan countries, maize grains are often used to prepare delicious meals that are acceptable to many people, while a large number of the African population depend predominantly on white maize; some preferably consume white or yellow maize. White maize endosperm contains essentially no carotenoids. Yellow maize grain colour ranges from light yellow to dark orange and they are said to contain pigments called carotenoids which are responsible for this colouration. Carotenoid pigments are of two major classes: carotenes and xanthophylls; however yellow maize stores more of xanthophylls than carotene in the grain (Grogan and Blessin, 1968).

Maize endosperm of the kernel has been shown to contain carotenoids (Blessin *et al.*, 1963). The horny endosperm contains 74–86 % of the carotenoids, floury endosperm has 9–23 %, and the rest is present in the germ and bran of the kernel (Blessin *et al.*, 1963). Yellow maize kernels contain different isoforms of carotenes (β-cryptoxanthin, β-carotene and α-carotene) which are precursors for vitamin A and they have long been known to play important roles in human nutrition and disease (Mangelsdorf and Fraps, 1931; Simpson, 1983). The β-cryptoxanthin, β-carotene and α-carotene are said to possess vitamin A activity and are referred to as pro-vitamin A carotenoids, however, beta-carotene is essentially two molecules of vitamin A having maximal pro-vitamin A activity while β-cryptoxanthin and α-carotene, with one ring contributing 50 % of the activity of β-carotene. The pro-vitamin carotenoids are capable of being transformed via biological activities into vitamin A in the body. Therefore they are valuable source of vitamin A in human diet and its human health benefit is enormous (Bartley and Scolnik, 1995; West *et al.*, 2002). The vitamin A activity of β-carotene in grains seems to be more effective than those from the dark-green leafy vegetables (Brown *et al.*, 1989). The importance of adequate vitamin A to influence growth, pregnancy and lactation, alleviate chronic diseases has since been recognised.

Yellow maize varieties have become a target for breeding purposes since the early 20th century (Blessin *et al.*, 1963). In early work on heritability studies for maize carotenoids, Brunson and Quackenbush (1962) concluded that maize hybrids with higher pro-vitamin A content can be bred by using genetic approaches. Also, association between vitamin A content and the inheritance of the yellow endosperm kernel has been studied long time

ago by Hauge and Trost (1928). They reported that vitamin A was transmitted, exclusively with yellow endosperm and there is a close physiological association between vitamin A and yellow endosperm kernel character in dent corn.

Studies of genetic variability for carotenoid concentrations of different yellow maize genotypes and hybrids adapted to various environments have been reported and it was indicated that the content vary among grain samples for all carotenoid components (Weber, 1987; Blessin *et al.,* 1963). Recently, HPLC method has been used to determine diverse carotenoid content and composition of kernels from different maize inbred lines (Harjes *et al.,* 2008; Menkir *et al.,* 2008).

Presently, the available maize genotypes in sub-Saharan Africa contain extremely low pro-vitamin A levels that may not supply adequate amounts needed for healthy living for the people. Harjes *et al.* (2008) reported that most yellow/orange maize grains grown and consumed throughout the world contain only 0.5-1.5 ug/g β-carotene as compared to xanthophyll carotenoids. This has contributed to VAD in this region. VAD has taken a great toll particularly on sub-Saharan children and women where maize diets are primary sources of calories. To combat this deficiency, maize genotypes with enhanced pro-vitamin A carotenoids, primarily as β-carotene (Kurilich and Juvik, 1999) is needed to increase the consumption of vitamin A. Furthermore, in a review, Sommer (2008) recently stated that solving the problem of VAD seems to require genetic improvement of food crops.

It would be desirable if maize inbred lines with varying β-carotene content and consisting of different genetic backgrounds are used as resources for nutritional improvement. Currently, there is breeding efforts through biofortification strategy to develop new cultivars of maize inbred lines that are adapted to tropical and sub-tropical regions. Therefore, high priority in the areas of research is urgently being pursued to enhance β-carotene level of staple food crops to meet dietary adequate intake of vitamin A (DellaPenna, 1999).

2.2 Health benefits of carotenoids

Pro-vitamin A carotenoids are important nutrients to human health because they can be converted to retinols which are imperative for eye vision, and later to retinoic acids which are needed for development and cellular differentiation (Fraser and Bramley, 2004). They are fat soluble and preformed retinal precursors (vitamin A). The efficiency of the carotenes in the formation of retinol is directly related to the nature of the enzymatic cleavage. The all-*trans* configuration of pro-vitamin A carotenoids gives the greatest pro-vitamin A activity while the isomerised *cis* configuration compromises activity which may influence enzymatic interactions responsible for conversion to vitamin A. β-carotene yields two molecules of retinol upon enzymatic cleavage. Also in humans, β-carotene prevents cataracts of the eyes and protects against free radicals (Johnson *et al.*, 2003).

The reduction in the risk of chronic diseases, like cancer in major organs of the body and cardiovascular disease has been attributed to the antioxidant activities (as free radical scavengers) of carotenoids (Hirschberg, 2001). Lycopene and carotene have actually been

known to prevent prostate, breast and other cancers (Agarwal and Rao, 2000). Carotenoids have also been involved in the delay of aging process due to their antioxidant properties (Bartley and Scolnik, 1995). Lutein and zeaxanthin play a protective role in macular degenerative processes (Semba and Dagnelie, 2003). They have been shown to provide immune system enhancement. In addition, carotenoids can be extracted from plants and used as micronutrient supplements. The importance of β-carotene in humans is quite enormous; therefore, if tropical-adapted yellow endosperm maize genotypes are developed through genetic improvement, they can be used to reduce the scourge of VAD among the people of sub-Saharan Africa.

2.3 Genetic diversity and relationship assessment

Genetic diversity assessment is critical to the successful achievement of crop genetic improvement. Genetic diversity analysis reveals genetic variability, genetic relationships and genetic backgrounds among inbreds in maize and can be determined by specific or combined methods (Mohammadi and Prasanna, 2003). The knowledge of genetic variation present in breeding lines is important. It will allow for selection of genetically diverse parental lines to make efficient crosses in breeding programs (Melchinger, 1999). Assessment of breeding lines is also necessary because of the narrowing genetic base (Goodman, 1990), for future breeding progress to ascertain the genetic diversity present for effective planning and management of breeding lines and may help to sort probable mix up in breeding materials. Also, when new inbred lines are derived from different origins and are to be used for hybrid breeding, to maximize efficient combinations of parents, genetic diversity is important for clear characterisation (Xia *et al.,* 2005).

Development of new lines through the exploitation of heterosis requires detailed information about genetic diversity and relationship among breeding lines (Smith and Smith, 1992; Legesse *et al.,* 2007).

A set of lines having an interesting agronomic trait are usually chosen from breeding pools and evaluated for genetic diversity using various methodologies such as, morphological, phenotypic characteristics, pedigree relationship, heterosis, isoenyzme markers and molecular markers. Each method has both advantages and disadvantages. However many researchers prefer to use both morphological data and molecular markers. Some depend mainly on the accuracy of DNA markers since they are not affected by environmental conditions (Bernardo, 1992), not time consuming and they need not carry out evaluation in different environments (Gerdes and Tracy, 1994). Nowadays, the vast application of molecular markers is well known for genetic variability analysis among breeding materials such as maize inbred lines as well as to estimate their genetic relationships.

Previous methods of measurement of diversity rely mostly on differences in morphological characters. Morphological markers could help classify maize lines through the use specific of combining ability with some line, pedigree information and or field hybrid-yield information for heterotic groupings (Fan *et al.*, 2009). Pedigree analysis and heterosis data (Smith and Smith, 1989) have been applied in genetic characterisation of maize germplasm. Qualitative and quantitative traits have been used to study many kinds of phenotypic variations in maize inbred lines (Menkir *et al.,* 2004). The use of

morphological traits to distinguish genotypes is expensive and time consuming because of many replication trials are involved. Similarly, because of their low heritability, low polymorphism and the significant interaction with environmental factors, their usefulness for genetic diversity studies is limited.

Isoenzyme markers were regarded as the first biochemical/protein markers and have been used for evolutionary studies (Hamrick and Godt, 1997) and for the assay of genetic diversity within and among populations (Doebley *et al.*, 1985). However, isoenzyme polymorphism is commonly low and may not give accurate classification of breeding lines. Subsequently, DNA/molecular markers were introduced for estimating genetic diversity (Pinto *et al.*, 2003; Reif *et al.*, 2003a, b; Olivera *et al.*, 2004). DNA markers are not affected by genotype X environment interaction and are capable of revealing differences among genotypes at the molecular level because they are abundant in the genome. Each type of marker has features that make it distinct and efficient. Combination of different markers can also help to establish precise genetic relationships among closely related genotypes.

The application of molecular markers in maize has become tremendous and many data have been generated and utilised in genetic mapping studies to characterise the maize genome (Gardiner *et al.*, 1993), studying quantitative trait locus analysis of important agronomic traits (Xia *et al.*, 2005; Ajmone-Marson *et al.*, 1998), variation of DNA fingerprints (Heckenberger *et al.*, 2003) genetic diversity studies (Reif *et al.*, 2003b; Li *et al.*, 2004), genetic characterisation of germplasm collections (Smith *et al.*, 1997),

molecular assisted breeding, studies of population genetic structure (Zhang *et al.,* 1992; Yang *et al.,*1994, 2003; Palaisa *et al.,* 2003). In the genetic characterisation of maize inbred lines, molecular tools have been used to study genetic diversity and to assign lines into heterotic groupings (Smith *et al.,* 1990; Livini *et al.,* 1992; Messmer *et al.,* 1993; Dubreuil *et al.,* 1996; Senior *et al.,* 1998; Benchimol *et al.,* 2000; Menkir *et al.,* 2004). Various molecular marker types are available to investigate relationships among maize inbred lines and the most commonly used are RFLP, RAPD, SSR, AFLP and SNPs. DNA markers generated are usually used to generate genetic distance from where the relationships and genetic diversity are inferred.

2.4 Types of Molecular markers used for genetic diversity assessments

Assessment of genetic diversity in plants can be evaluated using a variety of different molecular techniques (Zhang *et al.,* 2003). Molecular marker offers an efficient and reliable method to study variation at the DNA level. The high technology and automation of DNA marker methodology have allowed for their wide application to various researches in plants and animals. DNA markers can detect different kinds of DNA polymorphisms. They differ in methodology for detecting genetic differences, the type of data generated, cost-effectiveness, efficiency and complexity (Yang *et al.,* 1996; Pejic *et al.,* 1998).

DNA markers when compared with morphological analysis are not affected by physiological and environmental influences. The first ever molecular marker available to researchers was restriction fragment length polymorphism (RFLP) (Botstein *et al.,* 1980).

RFLPs are hybridisation-based markers, codominant in nature, therefore can distinguish heterozygosity. It has been applied to maize crop for various genetic characterisation studies (Messmer *et al.,* 1993; Yuan *et al.,* 2000; Reif *et al.,* 2003a, b) and it is still being used. Recently, Warburton *et al.* (2005) characterized 218 elite CIMMYT maize inbred lines by using 32 RFLP markers. RFLP analysis is based on restriction endonuclease digestion of genomic DNA. Differences in DNA sequence between individuals of a species are detected by separation on gel, nitrocellulose membrane blotting and using randomly selected genomic or cDNA cloned sequences as hybridization probes. Differences among individuals are caused by variation in fragment lengths arising either when mutations alter restriction sites, or resulting insertions/deletion between them. This non PCR-based assay is time consuming, labour intensive and difficult to automate.

The development of a powerful technique called the polymerase chain reaction (PCR) technique has led to the provision of new molecular methodologies. It allows the amplification of specific areas of genomes from extracted DNA of organisms. The invention of PCR was first used in DNA markers such as, dominant RAPD (Williams *et al.,* 1990). The RAPD method uses a single short 'arbitrary primers', usually 8-10 bp, to amplify several PCR fragments in non-coding as well as coding regions of DNA template. They are difficult to reproduce; however, they are still recognised and used for genotyping to distinguish individuals, cultivars or accessions and studying population.

In recent years, DNA techniques such as microsatellites (or simple sequence repeats, SSRs), amplified fragment length polymorphism (AFLP), single nucleotide polymorphisms (SNPs) have been used extensively to study genetic diversity in plant

species. In this study, the usefulness of SSRs, AFLP, and SNPs markers to assess polymorphisms in maize for a range of studies will be reviewed.

2.4.1 Simple sequence repeats

Simple Sequence Repeats (SSR), also known as microsatellites markers are based on repetitive short DNA sequence or short tandem repeats (SSRs/STRs) of short (2-6 bp) DNA sequence per unit. They are the most ideal markers due to their multi-allelic capacity, high polymorphic information content, reproducibility and locus specificity (Morgante and Olivieri, 1993). They allow heterozygotes to be detected, providing an accurate means of detecting variation in many plant species (Powell *et al.,* 1996; Wang *et al.,* 2006; Wei *et al.,* 2009).

Development of SSRs for a particular crop is carried out by screening DNA libraries for repeat motifs via hybridization and DNA sequencing of candidate clones (Taramino and Tingey 1996; Li *et al.,* 2000). Several SSRs have been isolated and mapped in maize. For detection of polymorphism in SSR analysis, primers are designed to the flanking sequence, the flanking sequence of these sites being unique (Jones *et al.,* 1997) then followed by the use of the primers for amplification of target sequence from total genomic DNA using the PCR method. Recently, SSR primers from related species are being applied (Li *et al.,* 2000). Amplified products (fragments) are separated according to size by gel electrophoresis (metaphor or superfine agarose/denaturing polyacylamide) allowing for the detection of different amplified alleles. SSR markers are highly variable loci which may be present at many sites and widely distributed in eukaryotic genomes but

absent in pro-karyotic genomes. The number of tandem repeats in a simple sequence repeat is highly variable among genotypes in species and is the basis for length variability of the marker. A high allelic variation per locus revealed by these markers comes from differences in the number of repeat units caused by slippage of DNA polymerase during replication.

In maize researches, the usefulness of SSR markers are quite many, they have been used for genetic diversity assessments (Pinto *et al.,* 2003; Laborda *et al.,* 2005; Wei *et al.,* 2009) assigning of lines into heterotic groups, population and genetic conservation (Powell *et al.,* 1996; Aguiar *et al.,* 2008), marker assisted selection (Weising *et al.,*1998; Bouchez *et al.,* 2002), genome mapping studies (Taramino and Tingey, 1996), prediction of hybrid performance (Mohammed *et al.,* 2008) and population genetic diversity studies (Labate *et al.,* 1997).

To reveal genetic diversity in maize inbred lines, SSR markers have been efficiently applied for estimation of accurate relationship (Smith *et al.*, 1997; Legesse *et al.*, 2007). Liu *et al.* (2003) studied the genetic diversity among 260 maize inbred lines consisting of essentially temperate breeding lines, tropical, and sub-tropical lines using 94 polymorphic SSR loci. The DNA markers produced 2039 alleles which placed the lines into five clusters that agreed with major breeding groups while some lines showed evidence of mixed origins. It was also revealed that the tropical and subtropical lines exhibited a greater number of alleles and greater gene diversity than those that are adapted to the temperate region.

Jambrovic *et al.* (2008) used SSR analysis to assess genetic diversity and relatedness among maize inbred lines relevant for breeding in Eastern Croatia, which consisted of B73, Mo17 and other 13 lines. Allelic diversity varied from 0.13 to 0.86, inbred lines were clearly placed according to their genetic background into two population varieties of Reid Yellow Dent and Lancaster Sure Crop and four subsequent maize families.

Xia *et al.* (2005) reported the characterisation of genetic diversity among CIMMYT maize inbred lines with 79 SSR markers, 566 alleles were amplified which generated a range of 0.45 to 0.93 genetic distances. The cluster analysis gave no defined grouping, indicating a mixed composition of CIMMYT subtropical, tropical mid-altitude and highland maize population pools. A large amount of variation was found to be present in CIMMYT germplasm.

Adetimirin *et al.* (2008) evaluated seventeen elite maize inbred lines of West Africa and Central Africa adaptation with tropical and temperate x tropical origin, alongside two temperate inbred lines for diversity using 18 SSR markers. A mean of 9.7 per SSR locus was detected and the SSR data grouped the lines based on their origin while the SSR markers did not consistently produce groups that correspond to the heterotic groupings.

For efficient analysis of SSR markers for genetic diversity assessment, markers with adequate genome coverage of the maize genome are usually chosen (Gethi *et al.*, 2002; Warburton *et al.*, 2002; 2003a). Hoxha *et al.* (2003) used 20 SSR loci that were uniformly

distributed across the maize genome to study genetic diversity of 20 Albanian maize populations. The average Polymorphic Information Content (PIC) value (0.71) was sufficient to assess genetic variation within and between local populations. Furthermore, Warburton *et al.* (2002) suggested that a minimum 53 core SSR markers can remarkably be used for genotyping CIMMYT inbred maize lines and open pollinated pollinations and clustering of maize genotypes for identification of related lines. These markers were chosen from 85 repeatable and easily automated markers.

2.4.2 Amplified fragment length polymorphism

Nuclear DNA polymorphism can also be studied using a DNA fingerprinting technique called Amplified Fragment Length Polymorphism (AFLP). This methodology was developed by Vos *et al.* (1995) and it is a relatively cheap, easy, fast and reliable for generating hundreds of informative genetic marker loci at once. They have become the marker of choice because they are reliable and efficient DNA marker method (Pejic *et al.*, 1998). It combines the advantage of PCR with the reliability of RFLP, thus it has the capacity for the simultaneous screening of many different DNA regions distributed randomly throughout the genome (Mueller and Wolfenbarger, 1999). Polymorphisms from AFLP markers are based on variation in restriction sites and length polymorphisms to estimate genetic variations across a genome.

It is a dominant marker whereby only a single allele can be scored for each generated locus, however markers can be scored quantitatively and used to deduce heterozygosity. AFLP markers are generated through three steps. The first step involves that a highly

purified genomic DNA is cut with two specific restriction enzymes, one frequent cutter (3 bp recognition site) and one rare cutter (bp recognition site). AFLP oligonucleotide adaptors are ligated (joined) to the ends of a subset of resultant fragment representing many loci; one end with a complementary sequence for the rare cutter and the other with the complementary sequence of the frequent cutter. This is designed so that ligation of a fragment to an adaptor does not reconstitute the restriction site. This way only fragments which have been cut by the frequent cutter and the rare cutter will be amplified. These known end sequences serve as priming sites in the subsequent PCR amplification.

Depending on genome size, restriction-ligation generates thousands of adapted fragments. Two steps of amplifications are usually required: (i) pre amplification- this is done with a single bp extension. (ii) selective amplification- followed by a more selective primer with up to a 3 bp extension, amplifying only those fragments in which the primer extensions match the nucleotides flanking the restriction sites. Sequences not matching selective nucleotides in the primer will not amplified.

Amplified fragments are resolved using denaturing polyacrylamide sequencing gel analysis (it allows the specific co-amplification of high numbers of restriction fragments) and visualized with silver staining technique. Typical results give 50-100 restriction fragments (bands) are amplified, this high number of fragments makes this technique the ability to provide more polymorphisms.

AFLP data have been used in population genetics and quantitative trait loci (QTL) mapping studies in plants (Mueller & Wolfenbarger, 1999). In maize, high-density AFLP linkage maps have been generated (Vuylsteke *et al.*, 1999). AFLP analysis is reproducible and gives high resolution (Mueller and Wolfenbarger, 1999; Garcia *et al.*, 2004). These reasons have made this analysis increasingly popular for fingerprinting. Many researchers have demonstrated AFLP as a powerful and highly accurate technique for studying genetic diversity.

In maize, AFLP analyses have been used to study the relationship between genetic distances and hybrid performance or heterosis grouping (Ajmone-Marsan *et al.*, 1998; Lubberstedt *et al.*, 2000). Genetic variability and genetic relationship estimation studies in germplasm of different breeding programs have been investigated based on AFLP information data (Pejic *et al.*, 1998; Lubberstedt *et al.*, 2000; Menkir *et al.*, 2006). AFLP is particularly valuable for estimating relatedness to maximise genetic gains during selections, which is an important aspect for successful hybrid breeding. Oliveira *et al.* (2004) used AFLP to investigate the genetic relationships among 96 tropical maize inbred lines from two different origins. Genetic similarities (GS), determined by Jaccard's similarity coefficient, varied from 0.345 to 0.891, with an average of 0.543. The dendrogram based on the GS and on the UPGMA cluster method did not separate the inbred lines in well-defined groups. However, when Tocher's optimization procedure was carried out, 17 groups were identified.

Genetic relationship among selected tropical and mid altitude maize inbred lines was estimated by Menkir *et al.* (2004) based on AFLP data. The markers generated was able to group the lines into two main groups with subdivision into subgroups consistent with breeding history, origin and parentage of the lines. Also, Menkir *et al.* (2006) determined the genetic diversity of 46 BC yellow maize lines derived from adapted X exotic backcross using AFLP data.

2.4.3 Single nucleotide polymorphism

Single nucleotide polymorphism (SNP) markers have become an efficient methodology used by many researchers around the world because large data can be generated to study genetic variations. SNP is a single nucleotide (base) change, which is thought to be dispersed in all plant and animals in numerous numbers. They are far more common DNA polymorphism in the genome (Bhattramakki and Raflaski, 2001). The low mutation rate of SNPs and the genetic information provided from SNP data are exact nature of the allelic variants (co-dominant biallelic markers) made them to be superior over markers like RAPD and RFLP, allowing them to decipher complex genetic traits (Syvanen, 2001).

Identification in organisms depends on precise generation of sequence data using DNA genetic analyzer. SNPs are being considered for genotyping because of automation, relatively easy to detect, and they are thought to contribute more to the variation in phenotype. In human genetics, detection of alleles associated with multifactorial genetic diseases has been achieved through the SNP tool. Maize is one of the many plant species that exhibit high SNP polymorphism (Bhattramakki *et al.*, 2002; Buckler and

Thornsberry, 2002; Vroh-Bi *et al.,* 2005). More than one million SNPs are presently available in maize databases (www.panzea.org).

SNPs can be categorised according to nucleotide substitution (change of one base to another). Mutation mechanisms result either in transitions: purine-purine (A ↔ G) or pyrimidine-pyrimidine (C ↔ T) or transversions: purine-pyrimidine or pyrimidine-purine (A ↔ C, A ↔ T, G ↔ C, G ↔ T) and single base of insertion/deletion ("INDELS") are also SNPs (Batley *et al.,* 2003). INDELS are generated by another mechanism. Exploring DNA polymorphism using single nucleotide polymorphisms (SNPs) and small insertion/deletion polymorphisms (In/Dels) as DNA markers for genetic analysis has recently received more attention (Hamblin *et al.,* 2007). SNPs occur frequently in crop plants, they have been found in genes, they affect the phenotype directly (Thornsberry *et al.,* 2001) or they can associate with a phenotype (LD analysis and genome-wide association studies) as a result of linkage disequilibrium (Tenaillion *et al.,* 2002), thus, also making them potentially useful for association studies of important genes (Rafalski, 2002).

SNPs have been used to construct genetic maps (high density linkage map). Maps are widely used in plant breeding to identify the regions of the genome that are controlling traits of interest, that is, for understanding the genetic basis of a trait (QTL mapping). SNPs are also useful in the construction of individual's haplotype (nucleotide base occurs at each position of these common SNPs for each chromosome). This has helped

researchers to better understand complex disease by correlating an individual's haplotypes with the presence of a disease in humans (Halperin and Eskin, 2004).

SNP discovery approaches are usually achieved through resequencing or data mining while SNP genotyping assays are carried out by several automated technologies ranging from allele-specific PCR, allele-specific hybridisation fluorescence resonance energy transfer based methods, pyrosequencing, array-based technology to a lot of others which an overview of their technical issues have been written by Bhattramakki and Raflaski (2001). Appropriately designed PCR primers can be used to discriminate SNP alleles. Recently, Sequenom-based quantitative SNP-typing assays were developed to identify 1,359 maize SNPs in phenotypic mutants via comparative next-generation transcriptomic sequencing. Approximately 75 % of these SNPs were successfully converted into genetic markers for genetic mapping (Liu *et al.,* 2010).

A maize Illumina GoldenGate assay with 1536 SNPs from 582 loci was developed to genotype a highly diverse maize collection of 632 inbred lines from temperate, tropical, and sub tropical public breeding programs. About 1229 informative SNPs and 1749 haplotypes within 327 loci were used to estimate the genetic diversity, population structure, and familial relatedness (Yan *et al.,* 2009). This genetic characterisation based on SNPs genotyping and linkage disequilibrium (LD) measurement defined a core set of inbreds based on haplotypes and 60 lines, captured 90 % of the haplotypes diversity among the genotyped lines. The LD decay distance differed among chromosomes and it was found out that it ranged between 1 to 10 kb, and it was much higher in temperate

than in tropical and subtropical lines, because tropical and subtropical lines are more diverse and contain more rare alleles than temperate lines.

Clark *et al.* (2004) in order to understand the impact of selection at the *tb1* gene that largely controls the increase in apical dominance in maize relative to its wild ancestor (teosinte) sequenced for SNPs characterisation in the upstream *tb1* genomic region and systematically detected nucleotide diversity for sites located as far as 163 kb upstream to *tb1*. They defined a selective sweep of ~60–90 kb 5' to the *tb1* transcribed sequence, which harbours a mixture of unique sequences and large repetitive elements, but it contains no predicted genes. The pattern of diversity at the nearest 5' gene to *tb1* is typical of that for neutral maize loci, indicating that selection at *tb1* has had a minimal impact on the surrounding chromosomal region.

Tenaillon *et al.* (2002) reported that maize has an average of one single nucleotide polymorphism (SNP) every 104 bp between two randomly sampled sequences when sequence diversity in 21 loci distributed along chromosome 1 of maize was measured among 25 individuals representing 16 exotic landraces and nine US inbred lines.

Palaisa *et al.* (2003), while investigating sequence diversity in two maize genes, the *Y1* phytoene synthase and *PSY2*, a putative second phytoene synthase among 75 white and yellow maize inbred lines. Many polymorphic sites showed strong association with the endosperm color phenotype at *Y1*, but no detectable association was found at *PSY2*. The sequenced regions of the *Y1* gene contained 32 insertions/deletions (indels) of varying

sizes, 85 non-coding single nucleotide polymorphisms (SNPs), and 21 coding SNPs. Of the 21 coding SNPs, 17 were informative and 10 produced amino acid changes.

Bhattramakki *et al.* (2002) resequenced 502 maize (*Zea mays*) loci across 8 maize inbreds (selected for their high allelic variation) and they identified 655 indels; 433 were polymorphic loci, with indels found in 215 loci, single-nucleotide indels accounted for more than half (54.8 %) followed by two- and three-nucleotide indels. However, a high frequency of 6-base (3.4 %) and 8-base (2.3 %) indels were also observed in their analysis. All the studies reviewed above have pointed out the successful applications of molecular markers in assessing genetic diversity in crops.

2.5 Comparison of different molecular markers for genetic diversity assessment

Recently, different molecular marker-based assays are available to assess genetic diversity. DNA markers detect DNA polymorphism at different parts of the nuclear genome. Therefore, for genetic diversity analysis of a crop, one or combination of methods can be used to understand fundamental questions in genetic studies (Mohammadi and Prasanna, 2003; Heckenberger *et al.,* 2003). Comparative analysis of markers involves the combination of molecular marker technologies and it is carried out for various reasons such as understanding the discriminatory power of various marker systems, efficiency, usefulness for genetic relationships and correlation analysis. Also, the use of two markers will allow their utility as tools in genetic analysis to be evaluated as well as for direct comparison of relationship between inbred lines at the molecular

level (Xia *et al.,* 2004). The first of its kind was the comparative between RFLPs, RAPDs, AFLPs, SSRs analyses by Powell *et al.* (1996). Similar study was investigated by Pejic *et al.* (1998), Lubberstedt *et al.* (2000) in inbred lines of maize, Montemurro *et al.* (2005) in olive plants, Menz *et al.* (2004) and Bohn *et al.* (1999) in winter wheat cultivars.

Powell *et al.* (1996) studied the utility of the AFLP, RAPD, RFLP, and SSR markers in both cultivated and wild soybean accessions. They reported low correlation of RAPD markers data with those obtained using other marker systems. However, a high correlation between AFLP, RFLP, and SSR data were reported, concluding congruence among measures of diversity. Pejic *et al.* (1998) utilized four molecular markers (AFLPs, RAPDs, SSRs, and RFLPs) to classify 33 maize inbred lines and SSRs were found to be more efficient. Garcia *et al.* (2004) compared the efficiencies of AFLP, RAPD, RFLP and SSR markers to find the most suitable for maize diversity studies.

RFLP and SSR markers have been used to group temperate lines into known heterotic groups (Messmer *et al.,* 1992; Dubreuil *et al.,* 1996). Comparative studies of RAPD and SSR markers in many crop species including maize (Souza *et al.,* 2008) and wheat (Jones *et al.,* 1997) have generally shown good congruence between the two genetic marker measurements. Laborda *et al.* (2005) used 569 AFLP and 50 SSR markers to genotype and group 85 tropical maize inbred lines. They found out that there was no clear defined groupings, suggesting that tropical maize studied is not well organised as temperate maize. Regardless of genetic similarity coefficient used, the AFLPs data were poorly correlated with the SSR data.

In the RAPD and SSR markers analysis to compare the genetic diversity among the 16 maize inbred lines, the similarity based on Dice coefficient for the RAPD data ranged from 53 to 84 % and for the SSR from 11 to 82 %. The dendrogram obtained by the RAPD showed five groups, while the SSR dendrogram showed three groups and one isolated line. The correlation between the two data using Mantel test was of a moderate value 0.54 (Souza *et al.,* 2008).

Hamblin *et al.* (2007) in their study on relatedness and evaluation of genetic diversity compared analyses based on 89 SSRs (primarily dinucleotide repeats) and based on 847 SNPs in individuals from the same 259 inbred maize lines. They indicated that SSRs performed better at clustering germplasm into populations than did a set of 847 SNPs or 554 SNP haplotypes, and SSRs provided more resolution in measuring genetic distance based on allele-sharing. Comparison at the individual level, measures of distance based on SSRs were only weakly correlated with measures of distance based on SNPs but not the case at the population level.

2.6 Correlation of molecular markers and phenotypic data

Genetic diversity within and between populations can be screened using morphological traits or by DNA techniques (Franco *et al.,* 2001). In the past, selection of parents for breeding was based on morphological evaluations that can be measured for a particular trait. Morphological markers have the disadvantage that they may not be totally genetic but can be influenced by some environmental factors and measuring them is time consuming. They may not be as informative as molecular markers for the purpose of

revealing detailed genetic relationship. Molecular markers can be used to reveal more genetic variation among very close genotypes and for better classification of the genotypes. However, evaluation of genetic relationships among breeding lines can also be based on morphological data (Mohammadi and Prasanna, 2003). The biochemical composition of the maize endosperm of tropical adapted lines has been characterised for the carotenoid variability. Few studies have used combinations of molecular markers and their comparison with morphological information for genetic diversity studies (Ramakrishnan *et al.,* 2004; Geleta *et al.,* 2006). Morphological traits based data have been variously combined with various molecular markers for genetic diversity studies using maize inbred lines or landraces (Smith and Smith, 1989; Heckenberger *et al.,* 2005; Wei *et al.,* 2009)

Beyene *et al.* (2005) used 15 morphological traits, eight AFLP primer combinations and 20 SSR primers to classify 62 selected highland maize accessions into groups based on molecular data and morphological traits. There was a low correlation between the morphological dissimilarity matrix and the matrices of genetic dissimilarity based on SSR and AFLP markers. The combination of molecular markers with phenotypic data is necessary to describe accurate relationship between genotypes. When phenotypic data are compared with molecular markers, there may be low or high correlations in their estimates of relatedness among breeding lines. In case of reported low correlations, various reasons have been attributed. This could be due to some lines possessing alleles at low or high frequency in comparison with other lines in another study, number of loci used, number of samples and the types of markers.

2.7 Carotenoids

Carotenoids are lipid-soluble isoprenoid pigmented molecules of 40 carbons, widely synthesized by organisms such as plants, algae, bacteria and some fungi and they have roles in both plants and animals. Carotenoid pigments include red, orange, deep yellow, which are present in different parts of the plants modification (Bartley and Scolnik, 1995). The colours of the carotenoids are as a result of large amounts of conjugated carbon-carbon double bonds present in them. The formation of carotenoids is from the isoprenoid biosynthetic pathway involving enzymatic catalyzing steps.

In plants, all the enzymes are encoded by nuclear genes and imported into the chloroplasts and chromoplasts for translational or post-translational (Bartley and Scolink, 1995; Fraser *et al.*, 1994; 1999; Cunningham and Gantt, 1998; Hirschberg, 2001) where biosynthesis of carotenoid takes place (Cunningham *et al.*, 1996). The carotenoids accumulate in large quantities in the chloroplast and chromoplast plastids (Howitt and Pogson, 2006). Many of the genes that catalyse carotenoid synthesis have been cloned and studied from several plants like Arabidopsis, tomato, pepper, maize (Dellapenna and Pogson, 2006) and fruits (Alquezar *et al.*, 2008).

Carotenoids are of two types, the carotenes (pure hydrocarbons) and the xanthophylls which contain at least one oxygen molecule in addition to the hydrocarbons and hydrogen molecule. They have been known to play many roles in plants. Carotenoid in plants possesses different functions such as accessory pigments in light-harvesting antennae during photosynthesis, as precursors of some scents, flavours, defence compounds, and

growth regulator (hormone) like abscisic acid (ABA) that modulate plant developmental stress processes, and as photo-oxidants to prevent the formation of reactive oxygen species (ROS) and photo oxidation of the organelle, enhancement of pollen and seed dispersal and plant development (Giuliano *et al.,* 2003; Fraser and Bramley, 2004; Howitt and Pogson, 2006). In addition to the several roles carotenoids play in plant, they are also important to animals in many ways, however, they are not able to synthesize it; rather they have to acquire them from plant-based dietary sources to meet their health nutritional needs. Their functions in human health have been discussed earlier (Section 2.2).

Carotenoid composition and accumulation differ in various organisms depending on the regulatory mechanisms of the genes underlying carotenoid biosynthesis and this will directly be linked with phenotypic variability in colours of the plant tissues. At certain times minimal or higher contents of expression of some genes are required for accumulation of high levels of carotenoids. The carotenoid biosynthesis has been well characterized (Fraser and Bramley, 2004; Hirschberg, 2001; Howitt and Pogson, 2006). Cloned carotenoid genes have facilitated the extensive characterization of the biosynthetic pathway and this in turn has led to the identification of the respective gene functions. This has also allowed the possibility of production of transgenic crops with enhanced level of carotenoids and in other cases provided opportunities for crop improvement through conventional breeding.

2.8 Characterisation of carotenoid biosynthetic genes in maize

The carotenoid biosynthetic pathway in higher plants (Bartley and Scolnik, 1995; Cunningham & Gantt 1998) is highly conserved and has been a subject of much extensive review. Figure 2.1 shows the carotenoid pathway (Hirschberg, 2001; Yan *et al.*, 2010). Biosynthesis of carotenoid begins with the formation of phytoene from geranylgeranyl diphosphate by phytoene synthase. The maize grain biosynthetic pathway in diverse maize genotypes has been studied to elucidate the biosynthetic steps that control carotenoids accumulation in the endosperm tissue (Vallabhaneni and Wurtzel, 2009). The relevant genes and cDNA encoding nearly all the enzymes involved in the biosynthesis of carotenoid pigments in the maize endosperm have been cloned, sequenced, and mapped. The steps and enzymatic reactions leading to the formation of coloured carotenoid compounds and the corresponding genes are discussed below.

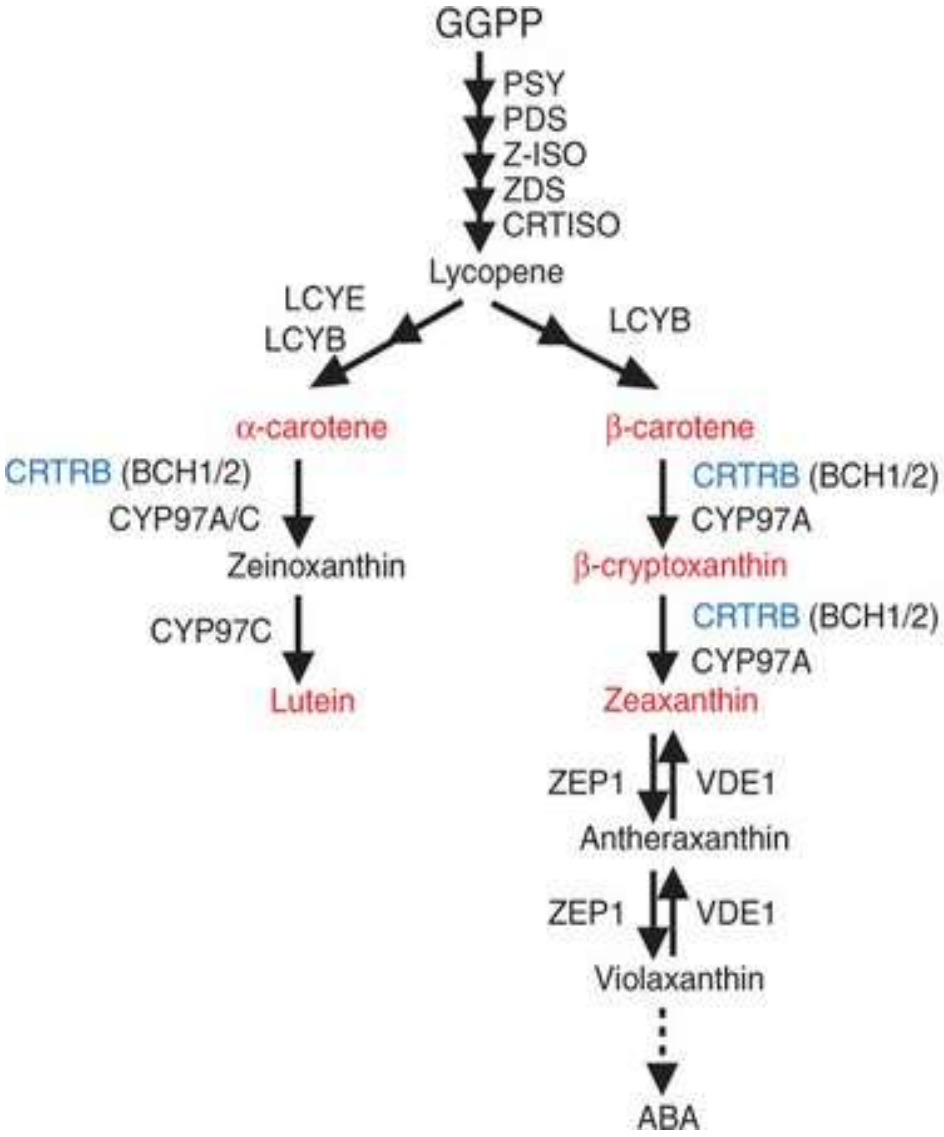


Figure 2.1: Simplified carotenoid biosynthetic pathway in Maize and Arabidopsis (Yan *et al.*, 2010). Carotenoids intermediates are in red. GGPP, geranylgeranyl pyrophospate; PSY, Phytoene synthase; PDS, phytoene desaturase; Z-ISO, ζ-carotene, LCYE, lycopene Ɛ-cyclase; LCYB, lycopene beta-cyclase; CRTRB, beta-carotene hydroxylase family; BCH1 and BCH2 are CRTRB, orthologous family in Arabidopsis; CYP97A, beta-carotene hydroxylase (P450); CYP97C, Ɛ-carotene hydroxylase (P450) ZEP1, zeaxanthin epoxidase; VDE1, violaxanthin de-epoxidase; ABA, abscisic acid.

2.8.1 Phytoene synthase and synthesis of phytoene

The first enzyme involved in the carotenoid biosynthetic pathway, phytoene synthase (y1/*PSY1*) is known to control flux to carotenoids in seeds (Gallagher *et al.,* 2004). Previous study has shown that *PSY* genes were said to be encoded by two paralogs, *PSY1* and *PSY2,* in 12 species across eight Sub-families of the Poaceae (grasses) (Gallagher *et al.,* 2004). Both genes have been tested for encoding functional enzymes active in maize. The *PSY* (*Y1*) gene was cloned by transposon tagging (Buckner *et al.,* 1990) and mapped to 6.01bin. *PSY1* transcripts correlate with carotenoid content of endosperm in maize (Buckler *et al.,* 1999; Li *et al.,* 2009), therefore, *PSY1* gene is critical for carotenoid accumulation.

PSY1 gene has been said to have undergone a selective sweep during the domestication process, leading to yellow maize lines accumulating carotenoids in their kernels (Palaisa *et al.,* 2003). *PSY2* (phytoene synthase 2), having significant protein similarity to the phytoene synthase 1 (*Y1*) is expressed in high mRNA levels in leaves but limited in the endosperm tissue, thus, it will not have influence on its level of carotenoids. It has been mapped on chromosome 8 (8.07) near marker umc1268 (Chander *et al.,* 2008).

Later on, *PSY3* which is another phytoene synthase gene was stumbled upon, identified and isolated in the maize and rice genomes (Li *et al.,* 2008). The full length of cDNA *PSY3* was later cloned by using reverse transcription (RT)-PCR for maize and rice. These three genes were thought to have been duplicated in the carotenoid biosynthetic pathway before evolution of the grasses (Gallagher *et al.,* 2004). Genomic and cDNA analysis

showed that conserved gene structure in *PSY3* as compared to *PSY1* and *PSY2* in the grass family, as well as with Arabidopsis *PSY* (Li *et al.,* 2008). *PSY3* expression has been suggested to be mainly accumulated in the other carotenogenic tissues such as root and embryo tissues of maize (Li *et al.,* 2008). *PSY* genes have been shown to be up regulated during plastid development in Arabdopsis and tomato plants, leading to the accumulation of carotenoids in many plant organs (Bartley and Scolinik, 1995; Giuliano *et al.,* 1993).

The gene product of *Y1* has been associated with production of carotenoids in the endosperm tissue, yielding yellow endosperm phenotype (Palaisa *et al.,* 2003; Li *et al.,* 2009). The first committed step in the carotenoid pathway is the condensation of two molecules of geranyl geranyl pyrophosphate (C_{20}GPP), a precursor, catalysed by the enzyme phytoene synthase (*PSY1*) to form a colourless product named phytoene. Phytoene's basic structure can either be 15 cis or all trans; depending on the stereochemistry, different types and levels of conversion of this C_{40} will result in the synthesis of other carotenoids by the other enzymes in the pathway.

2.8.2 Desaturases and phytoene desaturation

The carotenoid desaturases known in maize are phytoene desaturase (*pds*), mapped to 1.02 bin (Li *et al.,* 1996), zeta carotene desaturase (zds), mapped to 7.02 bin (Luo and Wurtzel, 1999), they primarily catalyze four sequential steps. The biochemical studies of maize desaturase mutant varieties, viviparous5 (*vp5*), *vp2* and white3 (*w3*) were the basis of establishing the function of *PDS* by the accumulation of phytoene and *vp9* and *y9* are associated with *ZDS* by the accumulation of ζ-carotene (Matthews *et al.,* 2003). A

similar work to elucidate the desaturase pathway in tomato was carried out using tomato mutants (Fraser *et al.,* 1994).

In maize, the desaturation steps are involved in producing phytofluene, neurosporene and zetacarotene (ζ-carotene) as intermediates, and finally to lycopene. Phytoene desaturase (*pds*) is the second enzyme in the biosynthetic pathway and it is responsible for two-step desaturation, it synthesizes zeta (ζ)-carotene from phytoene. Zeta-carotene desaturase is the third enzyme; it synthesizes all-trans-lycopene from zeta (ζ)-carotene by two-step saturation. This is the step at which visible coloured carotenoids are formed. The desaturation reactions introduce a series of carbon-carbon double bonds that constitute the chromophore in carotenoid pigments leading to a coloured compound referred to as lycopene. In fungi and non-photosynthetic bacteria, the four carotene desaturations are catalysed by a single enzyme known as the *CRITI.*

Plant desaturation steps also requires another reaction, isomerisation by carotenoid isomerases (*CRITO*) which will bring about acceptable geometrical isomer substrates for the action of desaturases and later for the lycopene cyclization steps. Two *cis-trans* isomerases of Z-ISO and carotenoid isomerase (*CRTISO*) (Park *et al.,* 2002) are used to convert poly-*cis*-configured phytoene into the all-*trans* form lycopene. This isomerisation is critical for metabolic engineering of carotene accumulation in food crops (Matthews *et al.,* 2003).

2.8.3 Lycopene cyclases and lycopene cyclization

Following the formation of lycopene is the cyclization step, which involves the rearrangement of the ends of the straight chain carotenoid lycopene to have a cyclic end group (to form, β, ε, or γ). The cyclized step represents the first branch point in maize carotenoids biosynthesis. For cyclization to take place, two enzymes are needed, namely lycopene β-cyclase (*LCYB*) and lycopene Ɛ-cylase (*LCYE*).

The beta lycopene beta cyclase (*LCYB*) was mapped to 5.04 bin (Singh *et al.,* 2003). The lycopene epsilon cyclase (*LCYE*) locus has recently been cloned and found to be responsible for variation in the content of vitamin A precursors in maize grains by association study of the candidate genes combined with QTL analysis and chemical mutagenesis by Harjes *et al.* (2008). *LCYE* and *LCYB* genes are single copy in maize and other grasses (Buckner *et al.,* 1990; Matthews *et al.,* 2003; Singh *et al.,* 2003; Li *et al.,* 2008).

The straight chain lycopene is converted into beta (β)–carotene by adding double beta (β, β) – rings or into γ-carotene by addition of a β-ionone ring at one end using the lycopene beta (β)-cyclase (*LCYB*), while the other cyclase (*LCYE*) converts lycopene into α-carotene by introducing two different rings of an ε- and β- rings (Cunningham *et al.,* 1996). β and ε- ionone rings differ at a double bond position within the ring structure. The carotenoid accumulation patterns in maize embryo and endosperm tissues were hypothesized to be due to the differential expression of genes encoding lycopene beta cyclase and lycopene epsilon cyclase (Bai *et al.,* 2009). Another recent study by Harjes *et*

al. (2008) reported that *LCYE* plays a key role in determination of the *β*-carotene/*α*-carotene ratio.

The functional analysis of the carotenoid cyclases of Arabidopsis has shown that the beta cyclase are linked to the addition of rings to both ends of lycopene while the epsilon cyclase will add only a single ε-ring to the symmetrical straight carotenoid (Cunningham *et al.,* 1996). It has also been presumed that the α-carotene synthesis is due to the co-action of the two cyclases (Hirschberg, 2001).

2.8.4 Hydroxylases and xanthophylls synthesis

Hydroxylation of cyclic carotenes takes place after the cyclization step leading to the production of the xanthophylls carotenoids in the carotenoid biosynthetic pathway; this step involves two types of hydoxylases, α and β hydroxylases (*HYDs*). Non-heme di-iron monooxygenases (non-heme) hydoxylases have been found to be present in most carotenoid-containing organisms and they can bring about the hydroxylation of carotenes (Tian and Dellapenna, 2004).

The β-carotene hydroxylases (non-heme di-iron type *HYDB*) are ferredoxin dependent and require iron, they convert carotene to cryptoxanthin and finally to zeaxanthin. The ε-carotene hydroxylase is associated mainly with the other second half pathway leading to the synthesis of lutein. However, lutein which is synthesized through the hydroxylation of C-3 of both the beta and the ε-rings requires the action of β-ring and ε-ring hydroxlyases. Lutein is a dihyroxy carotene and the most abundant carotenoid in plant tissues. Another

β-ring carotene hydroxylase cytochrome P450 type (*CYP93A*) and *ε*-ring carotene hydroxylase cytochrome P450 type (*CYP93C*) have also being identified in plants and they can also catalyze *α*-carotene and *β*-carotene to produce lutein and zeaxanthin, respectively (Tian *et al.,* 2004; Kim and DellaPenna, 2006).

There are six unlinked paralogs encoding non heme di-iron β-hydroxylases (*HYD*1-6) in maize, *HYDB1*, *HYD2*, *HYD3*, *HYD4*, *HYD5* and *HYD6*. Also, found present in maize are the P450 heme-thiolate *CYP97A* and *CYP97C* enzymes, each having a copy of the gene (Matthews and Wurtzel, 2007). *HYDBI* (*HYD*3/ *CrtRB1*) is linked to accumulation of β-carotene in maize endosperm.

The two types of hydroxylases are found to function in *Arabidopsis* (Hirschberg, 2001; Tian *et al.,* 2004). Two non-heme di-iron monooxygenases carotenoid hydroxylases (the *B1* and *B2* loci) that most times catalyze hydroxylation of the beta-ring of β,β-carotenoids and one heme-containing monooxygenase carotenoid hydroxylase (*CYP97C1,* the *LUT1* locus) that catalyzes hydroxylation of the ε-ring of ε, β-carotenoids have since been identified in Arabidopsis (Tian *et al.,* 2004). Recently, Kim and Dellapenna (2006) in their study, discovered that Arabidopsis *CYP97A3* (the *LUT5* locus) encodes a fourth carotenoid hydroxylase which has maximal activity for hydroxylation of the ε-ring of ε, β-carotenoids and less on the beta-ring of β, β-carotenoids.

In summary, based on the elucidated genetic studies of the maize carotenoid biosynthetic pathway, major quantitative trait loci have been identified for accumulation of carotenoids in maize.

2.9 Quantitative trait loci controlling carotenoid content in yellow maize

An important step elucidating the molecular basis underlying the carotenoid trait is the rapid discovery of important genes in the carotenoid pathway through various genetic analyses. Quantitative trait loci (QTLs) mapping was used to identify major QTLs for accumulation of carotenoids in maize. Loci controlling much of the phenotypic variation in contents of carotenoids have been identified to be caused by two loci (*y1* and *y9*). These loci were identified from a genetic linkage map constructed from 79 SSR, 8 CAPS, and 14 STS genotypic data using recombinant inbred lines (RILs), population derived from a cross between two parents By804 and B73 (Chander *et al.*, 2008). A total of four QTLs, for each β-carotene and β-cryptoxanthin and five for α-carotene were identified. Most of the QTL were located only on chromosome(s) 6 and 10. A major QTL was located only on chromosome 6 which accounted for 6.6-27.2 % genetic variation (largest of the variations) on levels of individual as well as total carotenoids (Chander *et al.*, 2008). The *y9* locus mapped on chromosome 10 was previously mapped to the same location and its role in the biosynthetic pathway for Z-ISO activity was equally identified. Chander *et al.* (2008) in their study suggested that some carotenoids biosynthesis pathway genes such as phytoene desaturase (*PDS*), lycopene β-cylase (*LCYB*) and phytoene synthase 2 (*PSY2*) in maize did not always correlate with the quantitative variations for

carotenoids in maize. Some QTL mapping results of *Psy1* have been mentioned in section 2.8.1.

The study by Wong *et al.* (2004) used two sets of segregating families; a set of $F_{2:3}$ lines derived from a cross of W64a x A632, and their testcross progeny with AE335 for mapping the chromosomal region associated with carotenoids accumulation in maize kernels. Almost similar results were reported as stated in the recent study above, therefore, Chander *et al.* (2008) confirmed *y1* and *y9* QTLs to be important in carotenoids accumulation in maize grain.

Harjes *et al.* (2008) showed that natural genetic variation (accumulation of mutations during evolution) at the lycopene epsilon cylase (*LCYE)* alters flux down α-carotene and β-carotene branches of the carotenoid pathway and was exploited for QTL mapping. The diverse panel of maize lines from different genetic background provided good basis for genetic analysis of the control of beta-carotene accumulation. Association mapping panels used for this study included the following three panels [P1-; U.S. (diversity) with 281 lines (77 white lines), P2-; CIMMYT (tropical) with 241 lines (61 white lines), P3-; Chinese (temperate) with 155 lines (22 white lines)]. Also, for linkage mapping work three mapping populations consisted- B73 × By804 RIL population; with 233 lines, A619 × SC55 $F_{2:3}$ population; with 181 lines and DEexp × CI7 $F_{2:3}$ population; with 102 lines. The study was investigated through association mapping, linkage mappings, expression analysis and mutagenesis for dissecting phenotypic diversity. They mapped (*LCYE*) to

chromosome 8 bin 5, near marker bnlg 1599. The complete sequence of *LCYE* has been obtained and also annotated (Figure 2.2).

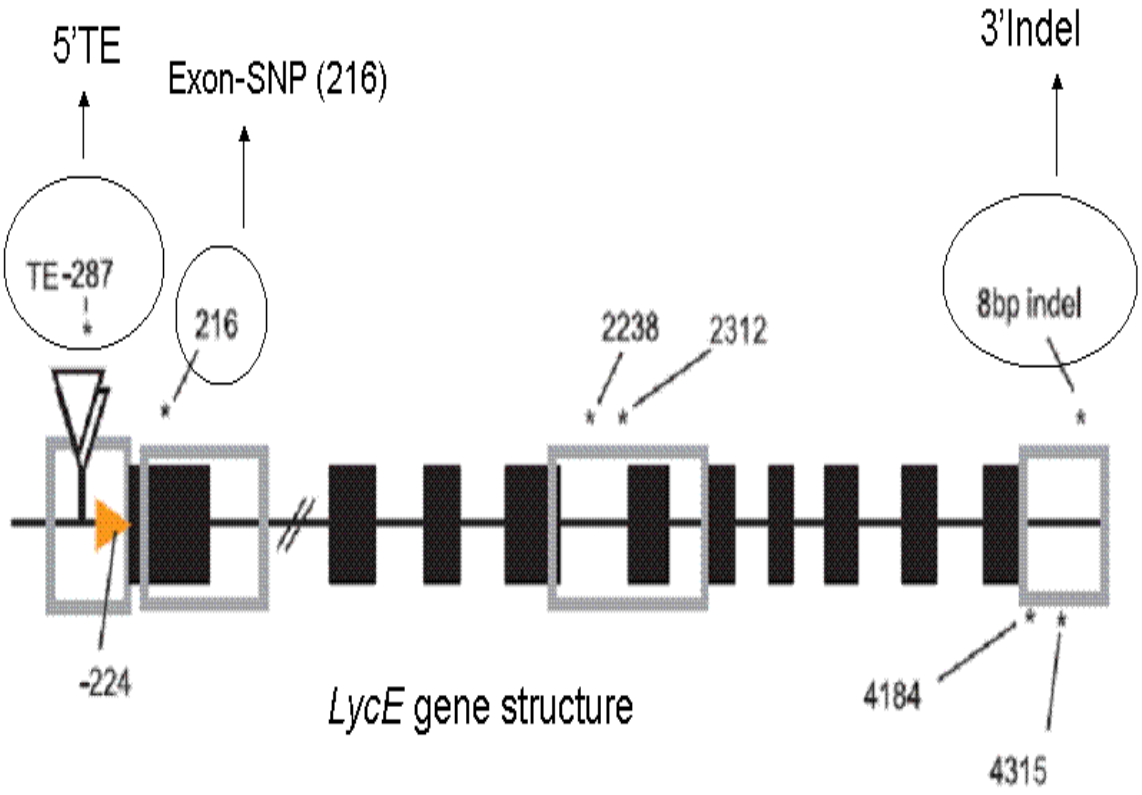


Figure 2.2: ***LCYE*** **gene structure showing different functional polymorphisms influencing alpha to beta carotene branching of the maize carotenoid pathway identified based on association analysis, linkage mapping and expression profiling (Harjes *et al.*, 2008). The regions in circle highlight the polymorphism significant for accumulation of β-carotene in maize.**

In another genetic analysis, carotene beta-hydroxylase 1 gene (*crtRB1*) has been demonstrated as significant in the maize carotenoid pathway and found to have a large effect on beta-carotene accumulation in the endosperm (Yan *et al.*, 2010). *crtRB1 was* identified via association mapping with three independent population panels (P1- 281 maize inbred lines grown in Urbana, Illinois (USA), P2- 245 diverse maize inbred from tropical and subtropical adapted maize germplasm and P3- 155 maize inbred lines derived from temperate-adapted maize germplasm, grown in China), and QTL mapping with five independent populations consisting of recombinant inbred lines or $F_{2:3}$ progenies (B73 X BY804RIL, DEexp X C17, A619 X Sc55 $F_{2:3,}$ K13 X SC55 K13 X B77 $F_{2:3)}$), *E. coli* activity assays and quantitative reverse transcription PCR (qRT-PCR) transcript profiling. *CrtRB1* has been cloned, sequenced and analysed (Figure 2.3).

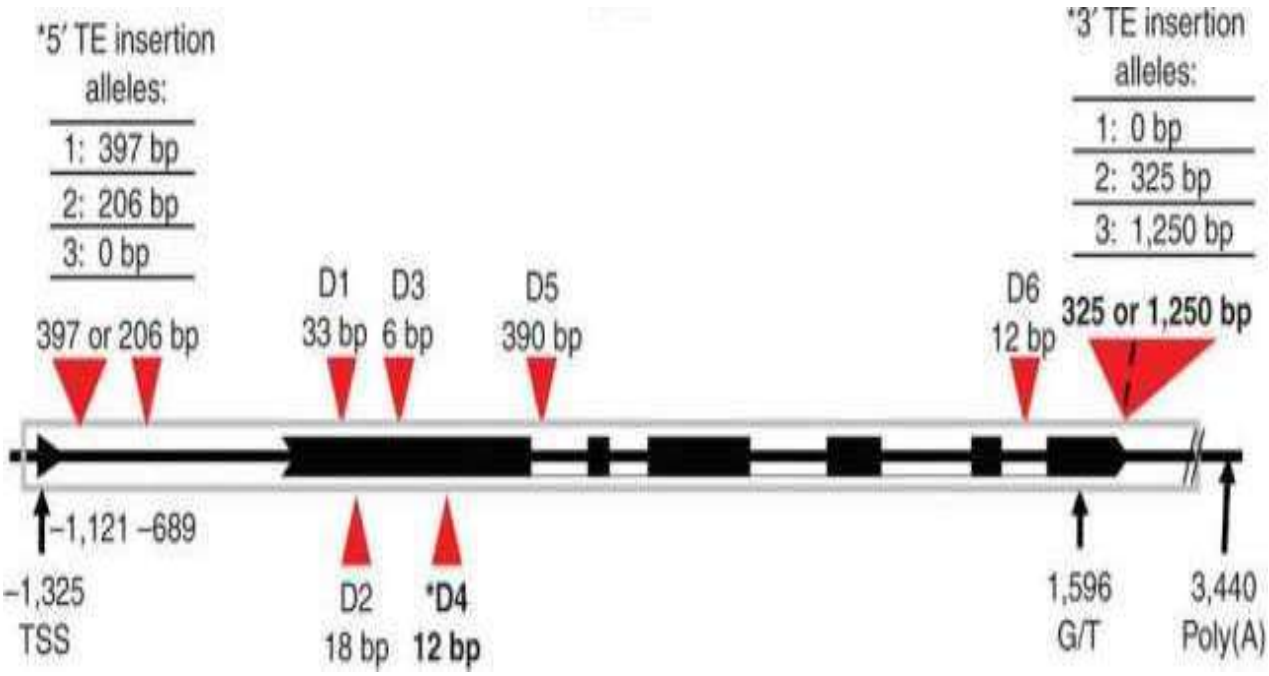


Figure 2.3: Annotated Maize *CritRB1* sequence showing different functional polymorphisms influencing β-carotene content (Yan *et al.*, 2010). The sequenced region is framed in gray, translated exons are depicted as black boxes and the putative start of transcription (TSS) and poly(A) sites are indicated. Polymorphism found in original P1 (population 1) sequence alignments are marked in the diagram, and those that are significantly associated with changes in beta-carotene are labelled with asterisks. TE, transposon element insertion; D, deletion.

Moreover in another study, through transcript profiling quantification, *crtRB1* (*HYD3*) locus was implied in the involvement of carotenoid accumulation (Vallabhaneni *et al.*, 2009). They observed the location of a QTL for endosperm β-carotene content that maps together with *HYD3*, between markers *umc1506* and *bnlg1028* on maize chromosome 10 as was previously shown by (Chander *et al.*, 2008). The study was on 10 maize inbred lines (A619, B73, B37, CI.7, C131A, DE3, KUI2007, NC300, SD44 and TZI18) that were selected through metabolite data sorting method. They represent genetically diverse subset possessing biochemical extremes of maize kernel carotenoid content and composition.

Vallabhaneni *et al.* (2009) investigated the maize duplicated non-heme *HYD* paralogs; six unlinked *HYD* genes and *CYP97A* and *CYP97C*. The maize *HYD1* and *HYD2* were found to be pseudogenes, while *HYD*3-6 genes were pointed out to be containing hydroxylase domains and plastid targeting signals, this suggests that when translated they will be functional. Subsequently, five carotene β-ring hydroxylase genes, *HYD3-6* and *CYP97A* and one ε-ring hydroxylase gene, *CYP97C* were considered for further work. Consequently, the study demonstrated that *HYD3* was the only gene for which transcripts were abundantly present and found to be statistically correlated with carotenoids when the quantitative abundance of the transcripts produced during a period of endosperm developmental stages were measured.

In conclusion, genetic linkage mapping, association analysis, mutagenesis and expression analysis (quantitative RT-PCR) strategies have helped identify chromosomal regions and

polymorphisms (alleles) that explain the variation in carotenoid accumulation in the maize. The *LCYE* and *crtRB1* (*HYD3*) genes/enzymes of the maize carotenoid pathway represent functional candidate genes for increasing the level of β-carotene in maize endosperm. The identification and characterisation of these QTLs have also led to their utility for functional diversity studies and development of functional PCR gene-based markers to help detect polymorphism that may affect the trait's variation. The PCR markers have been used to genotype maize lines with different genetic background. Validation of functional alleles in candidate genes in other genetic backgrounds such as the tropical adapted maize inbred lines will further assist breeders to incorporate marker-assisted selection programs that will finally facilitate the enhancement of pro-vitamin A in maize grains.

2.10 Functional markers for polymorphisms in β-carotene candidate genes

2.10.1 Gene expression patterns and detection of allelic variation in maize endosperm tissues

It has been suggested that DNA polymorphisms in the protein-regions of the genome is likely to be the cause of observable differences in the phenotypes both between and within a species (Mackay, 2001; Guo *et al.,* 2004). The allelic sequence differences between two maize genotypes can be characterised by analysis of SNPs or Indels and they can alter gene function (Vroh-Broh *et al.,* 2005; Bartley *et al.,* 2003). SNP marker analysis is the most recent marker system used in generating large abundance of SNPs in a genome or gene of interest. Sequencing analysis is one of the best methods for

analysing SNP variations in DNA. However, many available high-throughput technologies have given insight into the detection of functionally important alleles that influence variations in a quantitative trait.

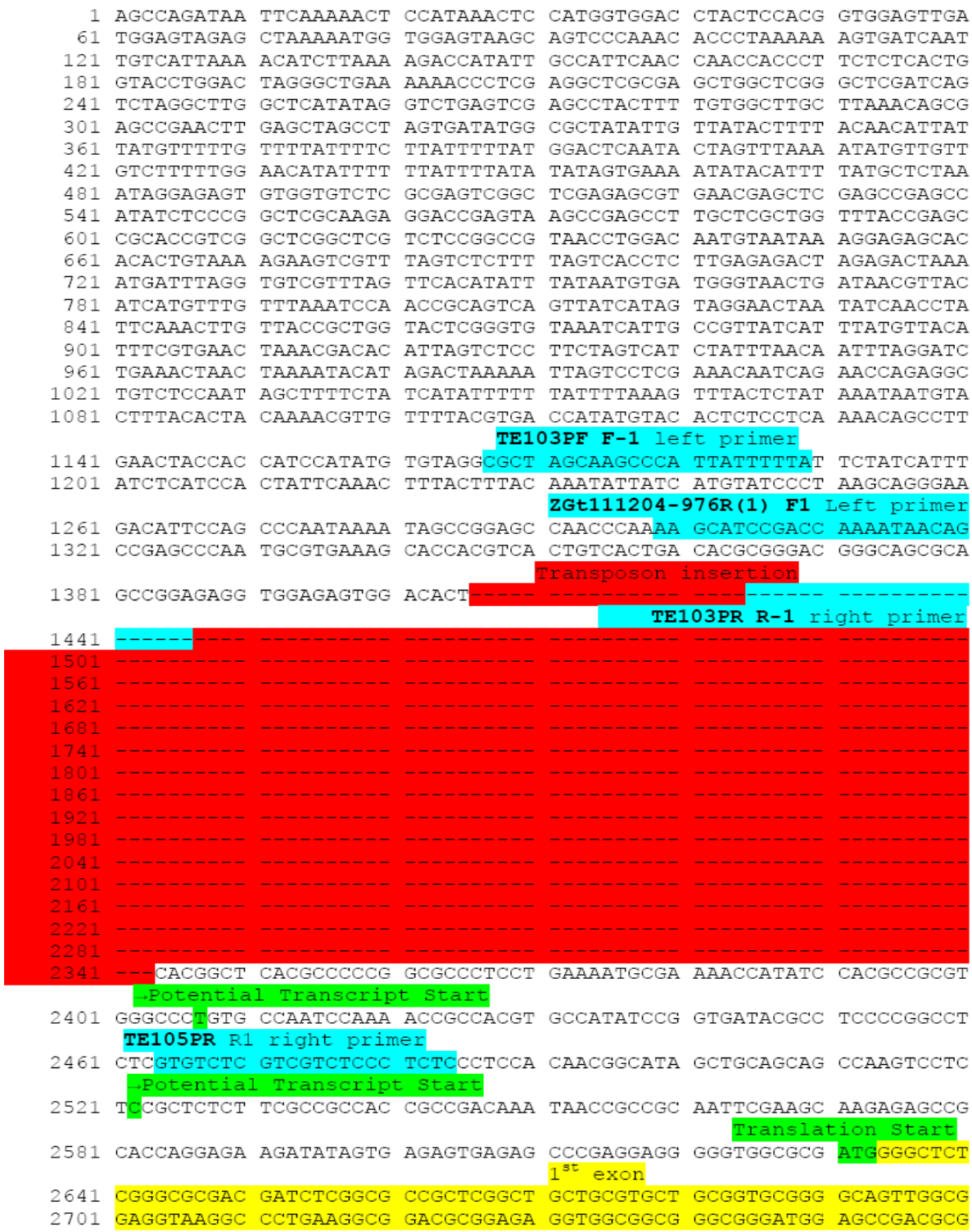
The use of resequencing and qRT-PCR for large scale expression analysis, expression of carotenoid genes in transformed *Esherichia coli*, molecular marker technology and various statistical approaches like linkage and association mapping studies have allowed for the identification of allelic variants of the β-carotene candidate genes. Quantitative mRNA expression levels of genes has been used to reveal, associate functionality of certain genes as well as causative genetic variations (SNPs-eQTLs), for example, changes in the gene expression or transcription factors. Also, statistical correlation between transcript level and a relevant trait have been achieved by using genetic linkage mapping or association mapping to dissect the allele-phenotype variation in known genes (Thorsnberry *et al.,* 2001; Jung *et al.,* 2004Yamasaki *et al.,* 2005).

Beyond the identification of functional genes related to β-carotene accumulation in endosperm of maize (Harjes *et al.,* 2008; Vallabhaneni *et al.,* 2009; Yan *et al.,* 2010), allelic variations at the transcript level in yellow maize have also been studied and analysed for two candidate genes (*LCYE* and *crtRB1*) (Harjes *et al.,* 2008; Vallabhaneni *et al.,* 2009; Yan *et al.,* 2010). This has resulted into further characterisation of favourable and unfavourable alleles of various polymorphisms involved in variation of β-carotene content that will facilitate the enhancement of β-carotene level. The functional gene-

based markers are linked to β-carotene functional diversity and are powerful for detecting polymorphic alleles.

2.10.2 Polymorphisms and effects of allelic variation

The *LCYE* expression was pointed out to have had variations having a partitioning effect on the two branches of the carotenoids and it has four principal functional polymorphism sites (Harjes *et al.,* 2008). A single nucleotide polymorphism at position 216 was located in the first exon of *LcyE* {*LCYE*-Exn-SNP (216)} with two alleles, T and G. The T allele is the unfavourable allele which may be replaced by the G allele which is referred to as the favourable allele. This polymorphism has been shown to have significant effect on the phenotype in association with another region showing a large promoter indel in the 5' end (*LCYE5*'TE polymorphism) with a 5.2-fold effect. The second indel in the 3'UTR region of *LCYE* gene significantly influence the ratio of alpha to beta-carotene in the carotenoid pathway, which is different from the variation caused by the promoter polymorphism, it also has 3.3-fold effect on the phenotype. The fourth significant polymorphism at position 2238 is in intron 4 and it is associated with a 2.5-fold effect. About 58 % of the overall polymorphisms have been found to explain the variation in maize endosperm colour. *LCYE* gene has been annotated for the sequence polymorphisms as shown in Figure 2.4.


TE103PF F-1 left primer
ZGt111204-976R(1) F1 Left primer
Transposon insertion
TE103PR R-1 right primer
Potential Transcript Start
TE105PR R1 right primer
Potential Transcript Start
Translation Start
1st exon

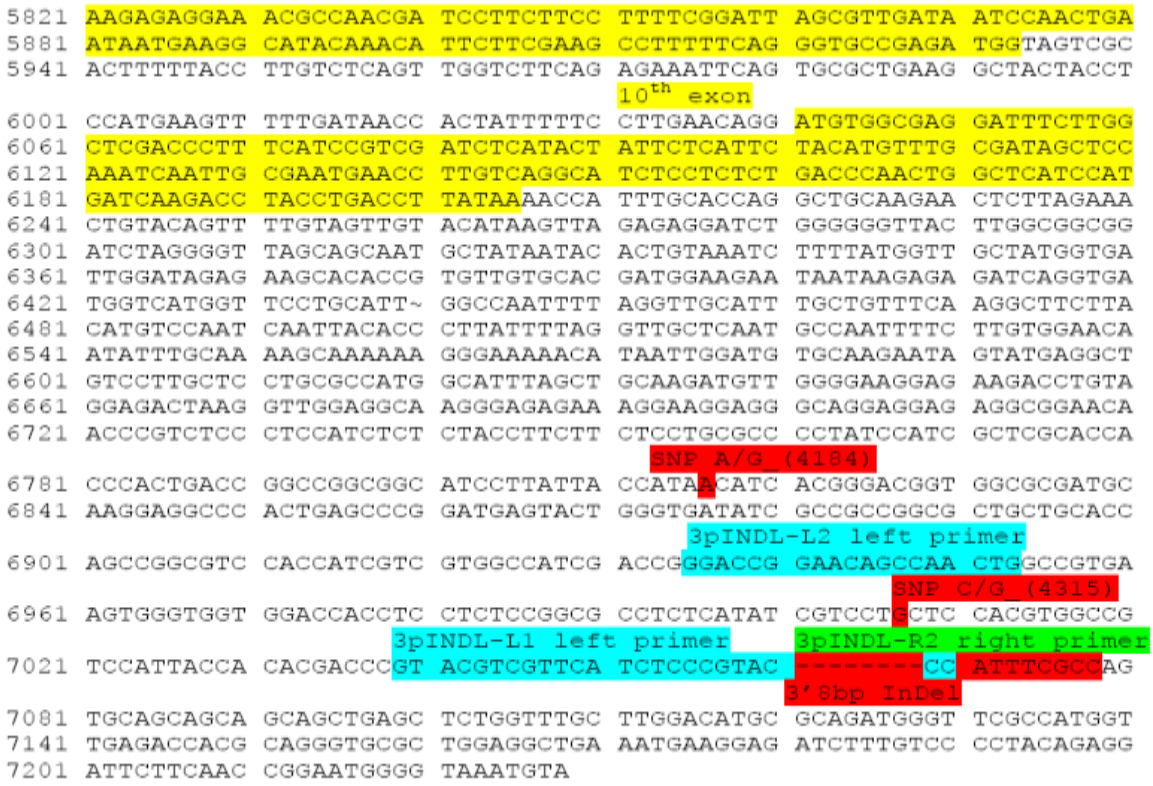
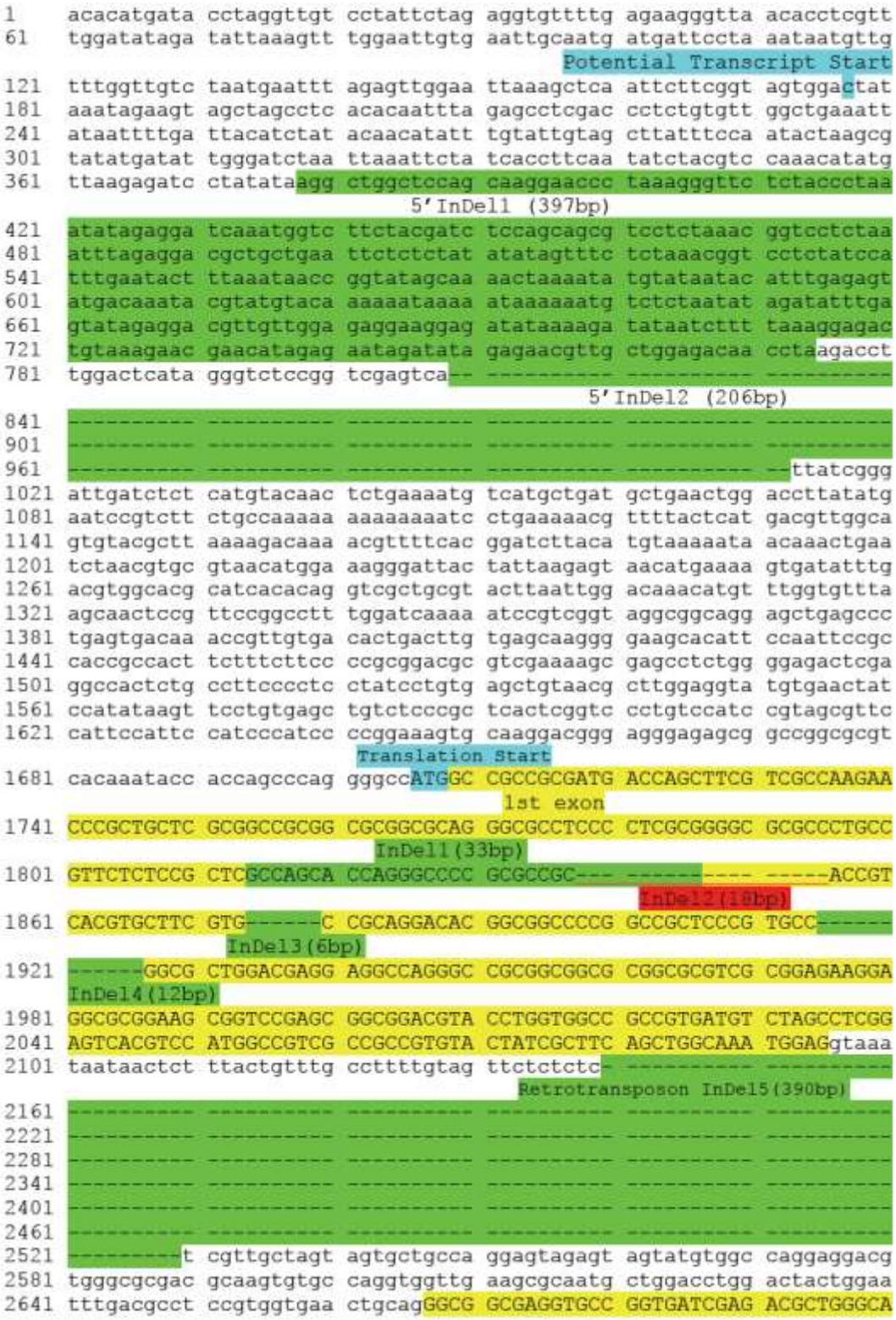


Figure 2.4: *LCYE* gene structure based on maize ZmGSStruc 11-12-04.976.1 according to Harjes *et al.* (2008). Exons are in yellow, other gene features are in green, polymorphisms of interest in red and primers to score key polymorphisms in blue.

Furthermore, rare genetic variations were also found in beta carotene hydroxylase 1 gene (*crtRB1*), also known as *HYD3* and have been shown to affect the increase in accumulation of carotenes (Yan *et al.,* 2010). Three polymorphisms have been identified for *crtRB1* and they have been shown to affect the conversion of beta carotene to zeaxanthin. A number of polymorphisms in *crtRB1* have been identified, in particular, 5'TE, InDel4, and 3' TE were found to be significantly associated with carotenoid variation. The 5'TE represents a TE insertion with a 206-bp insertion allele leading to higher β-carotene concentrations while the second is a 12 bp deletion in the first exon of crtRB1, representing and influencing the beta carotene content which represents deletion number 4 and it influences beta carotene content (InDel4).

The absence of a large TE insertion also influences beta carotene content at the 3' region of *crtRB1* gene**,** the third polymorphism (*crtRB1* 3'TE). c*rtRB1* is suggested to be involved in reducing the conversion of β-carotene to β-cryptoxanthin. The three polymorphisms are associated with different degrees of phenotypic variation for β-carotene and for β-carotene relative to total carotenoids in the three population panels used for association studies. The 5'TE explains 32 % phenotypic variation in β-carotene, and as regards β-carotene to β-cryptoxanthin (60 %), β-carotene to zeaxanthin (42 %), β-carotene to β- cryptoxanthin + zeaxanthin +α-carotene + lutein (42 %) based on the P1-281 maize inbred lines grown in Urbana, Illinois (USA). InDel4 and 3' TE account for 7-27 % of phenotypic variation in β-carotene, β-carotene to β-cryptoxanthin, β-carotene to zeaxanthin, β-carotene to β-cryptoxanthin + zeaxanthin + α-carotene + lutein. *CrtRB1* gene has been annotated for the sequence polymorphisms as shown in figure 2.5.


Potential Transcript Start
5'InDel1 (397bp)
5'InDel2 (206bp)
Translation Start
1st exon
InDel1(33bp)
InDel2(10bp)
InDel3(6bp)
InDel4(12bp)
Retrotransposon InDel5(390bp)

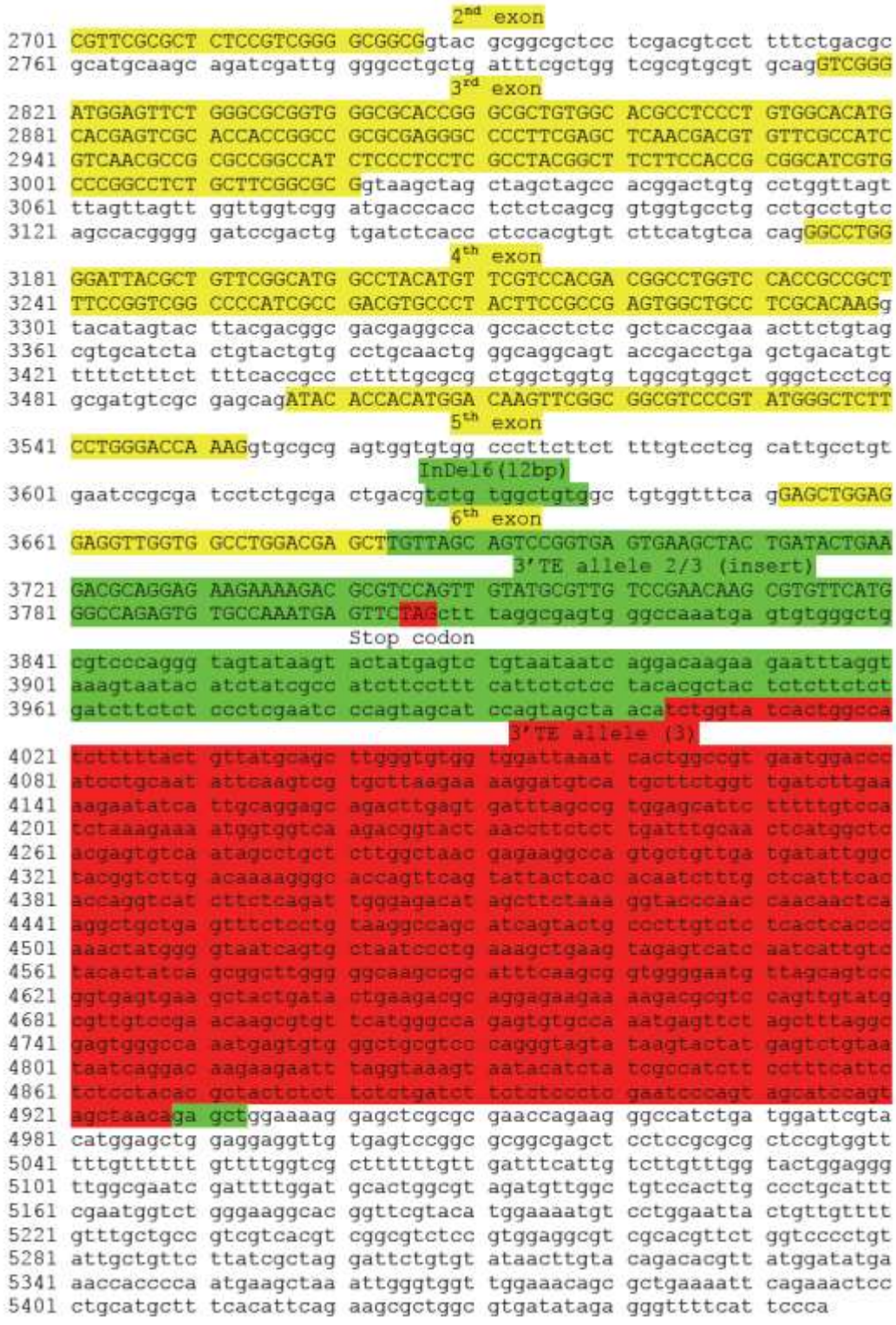


Figure 2.5: *CrtRB1* gene structure based on maize inbred B73 genomic sequence, according to Yan *et al.* (2010). Exons are in yellow, other features in blue, polymorphism of interest in green and red.

Moreover, other alleles in *HYD3* gene have been shown to have significant correlation with beta carotene content in the maize endosperm and a polymorphism was identified. *HYD3* in some maize lines has been sequenced and studied for the detection of additional polymorphism. Sequence variation was found in a ~ 40 bp region adjacent to the transcript start site and it was highly expressed and consistently seen in the high β-carotene lines. A conserved transcript start site and first ATG were mapped by aligning available paralog-specific ESTs and genomic DNA for maize. It was concluded that the *HYD3* alleles (polymorphism) brought about 78 % of variation and ~11-fold difference in β-carotene relative to beta-cryptoxanthin and 36 % of the variation and 4-fold difference in absolute levels of β-carotene. Part of *HYD3* gene showing the sequence polymorphisms is shown in figure 2.6.

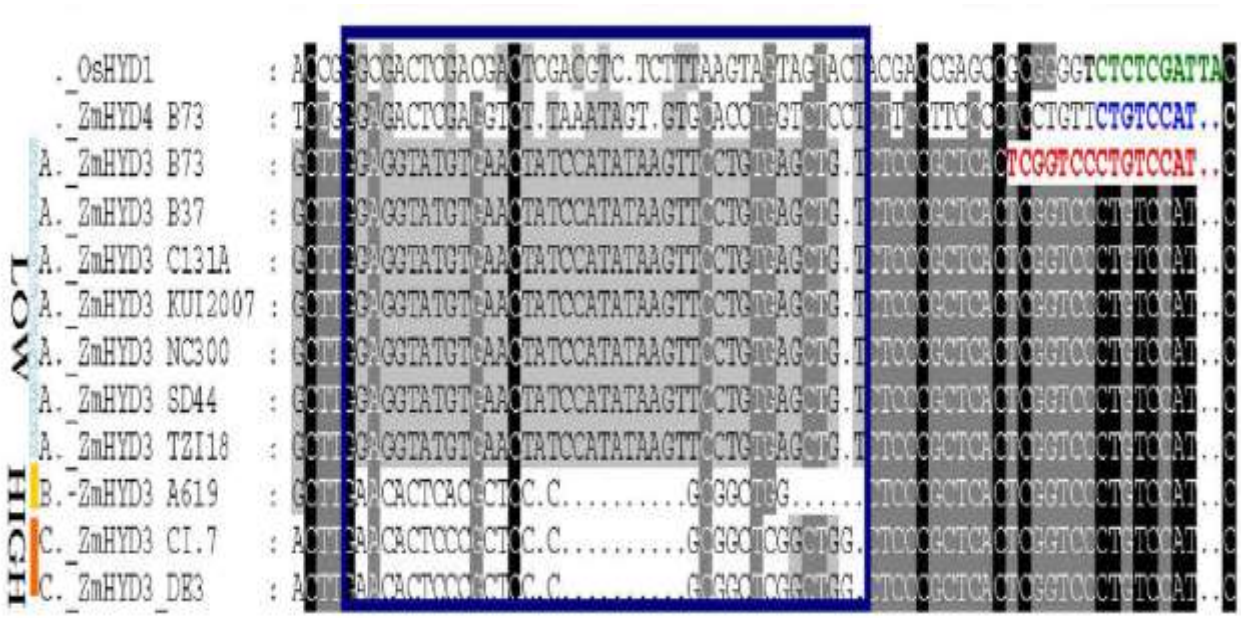


Figure 2.6: Part of *HYD3* sequence showing the variant regions that distinguish the high and low beta-carotene lines, according to Vallabhaneni *et al.* (2009). Black, dark grey, and light grey shading indicates degree of conserved nucleotides where black is the highest match.

2.10.3 Allele specific PCR-based molecular markers for amplifying variation (functional alleles) in β-carotene content

Alleles are naturally occurring and are said to be associated with genetic contribution to phenotypic variation in plants and they can easily be evaluated at each genetic locus for a range of them (Guo *et al.,* 2004; Pressoir *et al.,* 2009). Allele-specific primers for the genes of interest are possible to be designed according to the DNA sequence and sensitive enough for amplification and detection of allele variants (functional markers).

Following the detection of functional alleles that could provide added value in terms of higher β-carotene, various efficient PCR-based primers or polymorphic markers based on sequence differences for the two important genes (*LCYE* and *HYD3*) have been designed. They were designed around all the detected polymorphisms predicted to affect maize carotene biosynthesis (Harjes *et al.,* 2008; Vallabhaneni *et al.,* 2009; Yan *et al.,* 2010). The molecular markers will be useful for accurate characterisation of alleles and tracking of variation in beta carotene content for marker assisted selection. The polymorphic markers would be helpful for the selection of genotypes having the favourable alleles among tropical-adapted maize inbred lines and will facilitate effective application of MAS in breeding programs.

Primer sets have been designed for multiplex PCR assays for simultaneous amplification and identification of all specific allelic variants of *LCYE* and *crtRB1* polymorphisms, still enabling their separate detection in each genotype (Harjes *et al.,* 2008; Vallabhaneni *et al.,* 2009; Yan *et al.,* 2010). The specificity of these markers has been tested and validated

in some important maize germplasm resources (cultivars and breeding lines). These markers confirmed the presence of the predicted alleles. These alleles have been detected by agarose and or polyacrylamide gel electrophoresis. The sequences of the primers for each target gene have been made available (Harjes *et al.*, 2008; Vallabhaneni *et al.*, 2009; Yan *et al.*, 2010). The published markers (Harjes *et al.*, 2008; Vallabhaneni *et al.*, 2009; Yan *et al.*, 2010) need to be validated in the tropical adapted yellow endosperm maize inbred lines.

2.11 Breeding strategy for genetic improvement of β-carotene content in maize

Dietary deficiencies of the micronutrients such as vitamin A have led to serious health problems among children and women in the sub-Saharan Africa. Eye related problems such as night blindness, xeropthalmia, keratomalacia are as a result of VAD and these symptoms if not taken care of could lead to total blindness. Infections such as measles and diarrhoea which are leading cause of deaths among children have been attributed to VAD. Recently, the prevention of VAD may be achieved through biofortification.

HarvestPlus Global Challenge is a biofortification program of the Consultative Group on International Agricultural Research (CGIAR). It is coordinated by the International Centre for Tropical Agriculture (CIAT) and the International Food Policy Research Institute (IFPRI). International Institute of Tropical Agriculture (IITA), Nigeria is one of the research groups of HarvestPlus Maize Alliance. The development of this strategy is to biofortify breeding staple food crops such as rice, wheat, barley, maize, cassava, potatoes and beans local germplasms. Biofortification is relying upon the crop's biosynthetic

capability to accumulate the micronutrient desired (Mayer *et al.,* 2008), presence of potential genetic variation, availability of molecular markers and the use of traditional breeding method to assisted the breeders in marker-assisted breeding for crop improvement (Breseghello and Sorrells, 2006).

Apparently, maize is an important staple crop among the sub-Saharans and the existing lines have been shown to contain small amount of pro-vitamin A carotenoids which could be used as sources of vitamin A. Maize breeders have always been concerned with improvement of this food crop (Fernie *et al.*, 2006).

2.12 Marker assisted selection (MAS) for crop improvement

Traditional plant breeding involves the transfer of alleles from a donor line into a recipient genotype and followed by subsequent successive generations of backcrossing of F_1 to recipient to remove undesirable traits from the donor plant. In modern day plant breeding program, marker assisted selection (MAS) is an efficient tool for crop improvement (Ribaut and Ragot, 2007). It is an indirect selection of genotypes having alleles and haplotypes of interest for a given trait using molecular markers that are tightly linked to target gene. This assay can be carried out at any developing stage of the plant, allowing selection before the adult stage. In addition, selection will not be influenced by environmental factors, no problem of low heritability, saves time and cost (Collard *et al.*, 2005). Marker assisted backcrossing has several advantages over conventional back crossing which include effective selection of target loci, minimizes linkage drag and allows for quick recovery of recurrent parent. It provides an opportunity to select

individuals after conventional crosses have been made on the basis of their molecular profiles.

MAS has been shown to be valuable in the development of new varieties via marker assisted backcrossing (Collard *et al.*,2005) and introgression of genes through successive backcrosses between favourable allele carrying-genotype and recipient line, meaning introducing favourable allele from a donor genotype to recipient line (Bouchez *et al.*, 2002). Bouchez *et al.* (2002) have demonstrated the introgression of favourable alleles at three QTL underlying earliness and yield traits between maize elite lines based on marker assisted backcrossing. Recently, the use and application of MAS to improve its efficiency in plant breeding for crop improvement have been reviewed extensively by Xu and Crouch, (2008). Markers can also facilitate the process of pyramiding QTLs (genes) associated with the phenotype of interest in the same genotype.

There is a potential application of MAS in maize for improving the pro-vitamin A level of tropically-adapted yellow endosperm maize inbred lines. This will help reduce time and cost involved in phenotypic assays. MAS can allow for an efficient and more accurate selection of specific genotypes as parent lines for recombining favourable alleles for both *LCYE* and *crtRB1* loci using molecular markers. This accelerates breeding for new maize lines with enhanced nutritional quality. The validation of the favourable alleles responsible for carotenoid trait variation in this study is of great importance for application of MAS in maize breeding programs

CHAPTER THREE

3.0 MATERIALS AND METHODS

3.1 Plant materials

In this study, yellow endosperm maize inbred lines were obtained from the International Institute of Tropical Agriculture (IITA), Ibadan, Oyo State, Nigeria. Thirty-eight (38) maize lines were derived from different sources and contain varying levels of pro-vitamin A (Table 3.1) and one hundred and twenty-two (122) maize inbred lines with varying carotenoid content (Table 3.2) were included in this study. The 38 maize inbred lines were taken from existing lines. The 122 inbred lines were developed purposely as high provitamin A lines and represent a wide range of genetic backgrounds because they were derived from adapted X adapted crosses of lines with high pro-vitamin A content and backcrosses containing temperate germplasm as sources of high pro-vitamin A. The lines were developed by pedigree and are adapted to tropical and subtropical environments. These lines were chosen because they possess genes for adaptation and superior agronomic performance.

Table 3.1 Pedigree for the 38 yellow endosperm tropical-adapted maize inbred lines (Source, IITA Maize breeding Unit, July, 2008).

Inbred	Pedigree
PVL01	9450xKI 21-7-3-1-1-1-B-B-B
PVL02	9450xKI 21-7-3-1-1-3-B-B-B
PVL03	9450xKI 21-7-3-1-2-4-B-B-B
PVL04	9450xKI 21-7-3-1-2-5-B-B-B
PVL05	(9450xCM 116x9450)-5-2-2-2-B-B-B
PVL06	(9450 x KI 28)-1-2-1-2-B-B-B
PVL07	9450xKI 21-7-2-1-1-B-B-B
PVL08	9450xKI 21-7-2-1-2-B-B-B
PVL09	4001 x B73LPA x 4001-33-2-1-B-B-B
PVL10	1368 x GT-MAS-Gk-10-3-1-2-B-B-B
PVL11	(9450xCM 116x9450)-3-3-1-2-1-B-B-B-B
PVL12	9450xKI 21-3-2-2-1-3-B-B-B-B
PVL13	9450xKI 21-1-5-3-2-2-B-B-B-B
PVL14	9450xKI 21-1-5-3-2-1-B-B-B-B
PVL15	SYN-Y-STR-34-1-1-1-1-2-1-B-B-B-B-B-B-B-B
PVL16	9450xKI 21-1-4-1-1-2-B-B-B-B
PVL17	9450xKI 21-5-2-3-1-B-B-B
PVL18	ACR97TZL-CCOMP1-Y-S3-12-2-B-B-B-B-B
PVL19	ACR97TZL-CCOMP1-Y-S3-33-6-B-B-B-B-B
PVL20	ACR97TZL-CCOMP1-Y-S3-40-3-B-B-B-B-B
PVL21	KU1414-SR/NC350-4-1-B-B-B
PVL22	KU1414-SR/NC350-1-1-B-B-B
PVL23	(9450 x KI 28)-1-2-1-1-B-B-B-B
PVL24	KU1414-SR/KVI43-6-4-B-B-B
PVL25	KU1414-SR/KVI43-6-1-B-B-B
PVL26	KU1414-SR/KVI11-7-2-B-B-B
PVL27	KU1414-SR/KVI11-7-1-B-B-B
PVL28	(9450xCM 116x9450)-5-1-3-3-1-B-B-B-B
PVL29	9450xKI 21-4-2-3-1-1-B-B-B-B
PVL30	Taraba-14-2-2-4-2-B-B-B-B-B
PVL31	Z.Diplo.BC4-467-4-1-2-1-1-B-1-B-B-B-B-B-B-B
PVL32	TZE-COMP5-Y-C7-S3-61-B-B-B-B-B-B-B
PVL33	(9450 x KI 21)-8-2-1-1-B-B-B
PVL34	(9450 x KI 28)-5-1-2-1-1-B-B-B
PVL35	9450xKI 21-7-2-2-1-1-B-B-B
PVL36	9450xKI 21-7-2-4-2-1-B-B-B
PVL37	9450
PVL38	KU1414-SR

Table 3.2 Pedigree for the 122 yellow endosperm tropical- adapted maize inbred lines (Source, IITA Maize Breeding Unit, November, 2009)

Inbred	Pedigree
PV001	(9450xCM 116x9450)-3-3-1-B-B
PV002	9450xKI 21-7-2-4-2-1-B-B-B-B
PV003	(KU1409/KU1414-SR/KVI11)-S2-2-B-B
PV004	(KU1409/KU1414-SR/KVI3)-S2-5-B-B
PV005	(KU1409/KU1414-SR/KVI3)-S2-8-B-B
PV006	(KU1409/KU1414-SR/M162W)-S2-2-B-B
PV007	(KU1409/KU1414-SR/M162W)-S2-4-B-B
PV008	(KU1409/KU1414-SR/A619)-S2-1-B-B
PV009	(KU1409/KU1414-SR/A619)-S2-2-B-B
PV010	(KU1409/KU1414-SR/A619)-S2-3-B-B
PV011	(KU1409/KU1414-SR/A619)-S2-4-B-B
PV012	(KU1409/KU1414-SR/A619)-S2-5-B-B
PV013	(KU1409/KU1414-SR/A619)-S2-6-B-B
PV014	(KU1409/KU1414-SR/A619)-S2-7-B-B
PV015	(KU1409/KU1414-SR/A619)-S2-8-B-B
PV016	(KU1409/KU1414-SR/A619)-S2-9-B-B
PV017	(KU1409/KU1414-SR/NC298)-S2-2-B-B
PV018	(KU1409/KU1414-SR/NC298)-S2-5-B-B
PV019	(KU1409/KU1414-SR/NC298)-S2-6-B-B
PV020	(KU1409/KU1414-SR/NC298)-S2-9-B-B
PV021	(9450xCM 116x9450)-3-3-1-2-1-B-B-B-B-B-B
PV022	(KU1409/KU1414-SR/KVI3)-S2-5-1-B-B
PV023	(KU1409/KU1414-SR/M162W)-S2-2-1-B-B
PV024	(KU1409/KU1414-SR/A619)-S2-2-1-B-B
PV025	(KU1409/KU1414-SR/A619)-S2-6-1-B-B
PV026	(KU1409/DE3/KU1409)S2-1-B-B-B
PV027	(KU1409/DE3/KU1409)S2-2-B-B-B
PV028	(KU1409/DE3/KU1409)S2-3-B-B-B
PV029	(KU1409/DE3/KU1409)S2-4-B-B-B
PV030	(KU1409/DE3/KU1409)S2-5-B-B-B
PV031	(KU1409/DE3/KU1409)S2-6-B-B-B
PV032	(KU1409/DE3/KU1409)S2-7-B-B-B
PV033	(KU1409/DE3/KU1409)S2-9-B-B-B
PV034	(KU1409/DE3/KU1409)S2-10-B-B-B
PV035	(KU1409/DE3/KU1409)S2-11-B-B-B
PV036	(KU1409/DE3/KU1409)S2-13-B-B-B
PV037	(KU1409/DE3/KU1409)S2-14-B-B-B
PV038	(KU1409/DE3/KU1409)S2-15-B-B-B
PV039	(KU1409/DE3/KU1409)S2-18-B-B-B
PV040	(KU1409/DE3/KU1409)S2-20-B-B-B

Table 3.2 Pedigree for the 122 yellow endosperm tropical- adapted maize inbred lines (Source, IITA Maize Breeding Unit, November, 2009)

Inbred	Pedigree
PV041	(KU1409/DE3/KU1409)S2-21-B-B-B
PV042	(KU1409/DE3/KU1409)S2-26-B-B-B
PV043	(KU1409/DE3/KU1409)S2-27-B-B-B
PV044	(KU1409/DE3/KU1409)S2-28-B-B-B
PV045	(KU1409/DE3/KU1409)S2-31-B-B-B
PV046	(KU1409/DE3/KU1409)S2-32-B-B-B
PV047	(KU1409/DE3/KU1409)S2-35-B-B-B
PV048	(KU1409/DE3/KU1409)S2-36-B-B-B
PV049	(KU1409/KU1414-SR/KVI3)-S2-2-2-BB-B-B
PV050	(KU1409/KU1414-SR/KVI3)-S2-3-1-BB-B-B
PV051	(KU1409/KU1414-SR/KVI3)-S2-3-2-BB-B-B
PV052	(KU1409/KU1414-SR/KVI3)-S2-3-3-BB-B-B
PV053	(KU1409/KU1414-SR/KVI3)-S2-3-4-BB-B-B
PV054	(KU1409/KU1414-SR/KVI3)-S2-4-1-BB-B-B
PV055	(KU1409/KU1414-SR/KVI3)-S2-5-1-BB-B-B
PV056	(KU1409/KU1414-SR/KVI3)-S2-8-1-BB-B-B
PV057	(KU1409/KU1414-SR/KVI3)-S2-8-2-BB-B-B
PV058	(KU1409/KU1414-SR/M162W)-S2-4-2-BB-B-B
PV059	(KU1409/KU1414-SR/A619)-S2-1-1-BB-B-B
PV060	(KU1409/KU1414-SR/A619)-S2-3-1-BB-B-B
PV061	(KU1409/KU1414-SR/A619)-S2-5-1-BB-B-B
PV062	(KU1409/KU1414-SR/A619)-S2-5-2-BB-B-B
PV063	(KU1409/KU1414-SR/A619)-S2-7-1-BB-B-B
PV064	(KU1409/KU1414-SR/A619)-S2-8-1-BB-B-B
PV065	(KU1409/KU1414-SR/A619)-S2-9-1-BB-B-B
PV066	(KU1409/KU1414-SR/NC298)-S2-4-1-BB-B-B
PV067	(KU1409/KU1414-SR/NC298)-S2-4-2-BB-B-B
PV068	(KU1409/KU1414-SR/NC298)-S2-5-1-BB-B-B
PV069	(KU1409/KU1414-SR/NC298)-S2-5-2-BB-B-B
PV070	(KU1409/KU1414-SR/NC298)-S2-6-2-BB-B-B
PV071	(KU1409/KU1414-SR/NC298)-S2-14-1-BB-B-B
PV072	(KU1409/KU1414-SR/SC55)-S2-12-1-BB-B-B
PV073	(KU1409/KU1414-SR/SC55)-S2-13-2-BB-B-B
PV074	(KU1409/KU1414-SR/NC350)-S2-1-1-BB-B-B
PV075	(KU1409/KU1414-SR/NC350)-S2-1-2-BB-B-B
PV076	(KU1409/KU1414-SR/NC350)-S2-5-1-BB-B-B
PV077	(KU1409/KU1414-SR/NC350)-S2-6-1-BB-B-B
PV078	(KU1409/KU1414-SR/NC350)-S2-9-2-BB-B-B
PV079	(KU1409/KU1414-SR/NC350)-S2-11-1-BB-B-B
PV080	(KU1409/KU1414-SR/NC350)-S2-13-1-BB-B-B

Table 3.2 Pedigree for the 122 yellow endosperm tropical- adapted maize inbred lines (Source, IITA Maize Breeding Unit, November, 2009)

Inbred	Pedigree
PV081	(KU1409/KU1414-SR/NC350)-S2-13-2-BB-B-B
PV082	(KU1409/KU1414-SR/NC350)-S2-16-1-BB-B-B
PV083	(KU1409/KU1414-SR/NC350)-S2-16-2-BB-B-B
PV084	(KU1409/KU1414-SR/NC350)-S2-16-3-BB-B-B
PV085	(KU1409/KU1414-SR/NC350)-S2-19-1-BB-B-B
PV086	(KU1409/KU1414-SR/NC350)-S2-20-1-BB-B-B
PV087	(KU1409/KU1414-SR/NC350)-S2-21-1-BB-B-B
PV088	(SYN-Y-STR-34-1-1-1-1-2-1-B*5/NC354/SYN-Y-STR-34-1-1-1-1-2-1-B*5)-S2-7-5-BB-B-B
PV089	(KU1414-SR/CML328/KU1414-SR)-S2-5-2-BB-B-B
PV090	(KU1409/DE3/KU1409)S2-28-1-BB-B-B
PV091	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-1XSYN-Y-STR-34-1-1-1-1-2-1-BBB)S2-1-BB-B-B
PV092	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-1XSYN-Y-STR-34-1-1-1-1-2-1-BBB)S2-4-BB-B-B
PV093	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-4X4001xKI21-4-1-1-1-1)S2-2-BB-B-B
PV094	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-4X4001xKI21-4-1-1-1-1)S2-3-BB-B-B
PV095	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-6X(MP420x4001xMP420)-3-1-3-1-B)S2-1-BB-B-B
PV096	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-6X(MP420x4001xMP420)-3-1-3-1-B)S2-5-BB-B-B
PV097	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-6X(MP420x4001xMP420)-3-1-3-1-B)S2-10-BB-B-B
PV098	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-8XPOP61-SR-11-2-3-3-1-B)S2-3-BB-B-B
PV099	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-8XPOP61-SR-11-2-3-3-1-B)S2-4-BB-B-B
PV100	(POP61-SR-11-2-3-3-1-BB/9450xKI21-3-2-2-1-3)S2-1-BB-B-B
PV101	(POP61-SR-11-2-3-3-1-BB/9450xKI21-3-2-2-1-3)S2-6-BB-B-B
PV102	(KU1409/SC55/KU1409)-S2-8-1-B-B
PV103	(KU1409/SC55/KU1409)-S2-19-1-B-B
PV104	(KU1409/SC55/KU1409)-S2-38-1-B-B
PV105	(KU1409/DE3/KU1409)S2-18-2-B-B
PV106	(KU1409/DE3/KU1409)S2-30-2-B-B
PV107	SC55/KU1414-SR/KU1414-SR-6-BB-B-B
PV108	DE3/KU1414-SR/KU1414-SR-6-BB-B-B
PV109	KU1409/DE3/KU1414-SR-9-BB-B-B
PV110	KU1409/DE3/KU1414-SR-12-BB-B-B
PV111	KU1409/NC358/KU1409-2-BB-B-B
PV112	KU1409/NC358/KU1409-8-BB-B-B
PV113	KU1409/NC358/KU1409-17-BB-B-B
PV114	KU1409/SC55/KU1409-4-B-B
PV115	DE3/KU1414-SR/KU1414-SR-2-B-B
PV116	DE3/KU1414-SR/KU1414-SR-7-B-B
PV117	KU1409/DE3/KU1414-SR-10-B-B
PV118	KU1409/NC358/KU1409-14-B-B
PV119	KU1409/NC358/KU1409-16-B-B
PV120	KU1409
PV121	4001
PV122	KU1414-SR

3.2 DNA extraction

Seedlings of the 38 and 122 inbred lines were raised in the green house at IITA, Ibadan. The land was prepared by weeding and rows were made. The rows were sown with 20 seeds per inbred line and the plants were wetted every three days. Leaves of each of the inbred lines were harvested from three weeks old plants. Fresh leaf tissue of 5-6 seedlings of each inbred line was pooled together for total genomic DNA extraction and stored in Eppendorf tubes at -80°C freezer. The modified Dellaporta *et al.* (1983) method was used with some modifications to extract genomic DNA. One gram of this pooled leaf tissue was ground into a fine powder in liquid nitrogen with Konte pestle in 1.5 ml Eppendorf tube and used for mini-preparation of DNA. Eight hundred micro litres (800 µl) of extraction buffer [100 mM Tris-HCl (pH 8.0), 50 mM EDTA (pH 8.0), 500 mM NaCl, 1.0 % polyvinyl pyrrolidone and β-mercaptoethanol] was added to the ground leaf material and mixed to disperse all the tissue in the buffer. SDS (20 % w/v, 100 µl) was added and mixed thoroughly for another 1 min. Incubation of samples was carried out in water bath for 15 min at 65 °C with 5-6 intermittent mixing, after which they were removed and allowed to cool to room temperature for 2 min. Ice-cold 5 M potassium acetate (300 µl) was added and mixed gently inverting 5-6 times and incubated on ice for 20 min.

Cell debris (protein and polysaccharides in form of insoluble potassium-SDS complex) was removed and precipitated by centrifugation at 12000 rpm for 10 min. The supernatant was carefully transferred into two new labelled 1.5 ml Eppendorf tubes and was precipitated with equal volume of ice-cold iso-propanol (approx.700 µl) and mixed gently

by inverting 8-10 times and incubated at -80 °C for 1 h and centrifuged at 12,000 rpm for 10 min. The supernatant was decanted and the last drops were removed by facing down the tubes on paper towels. For removal of RNA, the DNA was re-dissolved in 250 µl of high salt TE buffer (1 M Tris and 0.5 M EDTA) and incubated with 4 µl of RNase A solution [(10 mg/ml) DNase free] at 37 °C for 1 h. For precipitation of DNA, 500 ml of isopropanol was added and incubated at -80 °C for 1 h and centrifuged at 12,000 rpm for 10 min. The supernatant was removed and the DNA pellet was washed twice with 1 ml of 70 % ethanol with centrifugation for 3 min at 12,000 rpm. The DNA was air-dried completely at room temperature for 1 h, re-suspended in 100-200 µl of sterile water and stored at 4 °C overnight for total dissolution. The extracted DNA was stored at -20 °C until ready for use.

3.2.1 Checking the quality and quantity of extracted DNA

The DNA concentrations were estimated using NanoDrop spectrophotometer (ND-100 Technologies, Wilmington, Delaware, USA) machine by dropping 2 µl of extracted DNA on the NanoDrop spectrophotometer. Sterile water was used for calibration and as well as for blanking. DNA quality was determined by two methods. (i) Measuring the 260:280 with NanoDrop which varied between 1.8 and 2.0 (ii) By agarose gel electrophoresis; each genomic DNA sample (3.0 µl + 1.5 µl gel loading dye) was loaded onto the agarose gel (1.0 %, 0.5X TBE) and stained with ethidium bromide solution for 3-8 min and destained in water for about 5-10 min. The gels were photographed under UV light attached to a gel documentation system.

3.2.2 DNA Dilution

Extracted DNA was diluted to working concentration for PCR using $M_1V_1 = M_2V_2$. M_1 is the stock DNA concentration, V_1 is the volume of stock to be diluted, and M_2 is the concentration of working concentration and V_2 is the volume of working solution to be prepared. Dilution was prepared by taking appropriate volume from stock solution into a new Eppenddorf tube and adding sterile water to give the final volume of working solution. The working dilution was stored in -20 °C.

3.3 Genetic diversity assessment among yellow endosperm maize inbred lines

3.3.1 AFLP-based assessment

AFLP assay was carried out as described by Vos *et al.* (1995). One and a quarter (1.25) units each of a mixture of EcoRI and MseI (Invitrogen) restriction enzymes in 5X reaction buffer (50 mM Tris-HCl (pH 7.5), 50 mM Mg-acetate, 250 mM K-acetate) was used for DNA digestion (200 ng/µl) of the maize inbreds; 8 µl of DNA and 2.5 µl of distilled water was added to give a final volume of 12.2 µl. The reaction mixture was incubated at 37 °C for 3 h. Eight and a quarter (8.25) µl of a mixture containing EcoRI and MseI adapters, 0.4 mM ATP, 10 mM Tris-HCl (pH 7.5), 10 mM Mg-acetate, 50 mM K-acetate and 0.25 units of T4 DNA ligase were added to the digested DNA and the reactions were incubated at 18 °C overnight. Samples were subsequently diluted 1:10 with water (3 µl from digestion and ligation PCR products and 27 µl of water). PCR was performed in two consecutive reactions. The pre-selective PCR amplifications were

performed in 13.25 µl reaction volumes each of which contained 10 µl pre-amplification primer mix 1, 5 µl diluted adapter-ligated genomic DNA, 0.25 unit *Taq* DNA polymerase (Bioline), 10X PCR buffer (100 mM Tris-HCl (pH 8.0), 500 mM KCl) and 50 mM $MgCl_2$. Pre-selective PCR amplifications (28 cycles) were performed as follows; 30 sec at 94 °C, 30 sec at 56 °C, and 60 sec at 72 °C. A total of 17 EcoR1and Mse1 primer combinations with three nucleotides were used for selective amplification. The selective PCR amplifications were performed in 10 µl reaction volumes each of which contained 2.5 µl diluted (1:50) pre-amplified DNA template, 2.5 µl mix of MseI primer and EcoRI primer, dNTPs (200 µM each), 0.1 unit of *Taq* DNA polymerase, in 1.0 µl 10X PCR buffer and 0.4 µl of 50 mM $MgCl_2$. PCR program consisted of two segments. Firstly, 12 cycles, with annealing temperature decreased from 65 °C by 1 °C in each cycle: 30 sec at 94 °C, 30 sec at 65 °C to 56 °C and 1 min at 72 °C. This was followed by 28 cycles of 30 sec at 94 °C, 30 sec at 56 °C, and 60 sec at 72 °C. All PCR reactions were conducted in a MJ Research (PTC-200 Peltier) thermal cycler. The amplified fragments were then separated by electrophoresis on a denaturing 6 % polyacrylamide gel and visualized by silver staining procedures. The gels were allowed to dry and gel images were captured by a scanner system.

3.3.1.1 Separation of amplified fragments on denaturing polyacrylamide gels (PAGE)

The short plate was cleaned using detergent, rinsed with de-ionised water and a final wash to remove detergent residues was performed with 100 % ethanol three times. To prepare a fresh binding solution, 3 µl of bind silane was added to 1.0 ml of 95 % ethanol,

0.5 % glacial acetic acid (95 ml of absolute ethanol, 500 µl of glacial acetic acid, + 4.5 ml of distilled water). The whole surface of the cleaned plates was wiped with 1.0 ml of binding solution with Kim wipes, after which the plate was allowed to dry for 10 min. Excess binding solution was wiped with Kim wipes briefly. The long plate was cleaned using detergent and rinsed thoroughly with de-ionised water. The long plate was treated with 1ml of silica gel solution by spreading it on the plate, and the wiped evenly with Kim wipes so that the plate was completely covered. After 10 min, excess silica gel was removed with Kim wipes and allowed to dry. The caster base was placed flat on the bench with the long glass facing upward, clean and dry spacers (0.4 mm) was aligned along the edge of the long plate and the short plate (coated surface facing down) was carefully placed on the long plate. The two plates were sandwiched securely with clamps and the plates at the comb end were also clipped. The fit of the comb was checked by trying to place it between the plates (slight resistance to being placed between the glass plates without damaging the comb). To prepare PAGE gel, chilled 6 % polyacryl/bisacrylamide solution (100 ml) was poured into a beaker, 500 µl of 10 % Ammonium persulphate (APS) [(APS) 0.1 g of APS in 1ml of distilled water] and 50 µl of N, N, N1, N1 tetramethylenediamine (TEMED) were also added. The solution was stirred with pipette tip. The gel was poured immediately into the plate sandwich and all bubbles were removed by gently tapping the short plate. The comb (48 wells) was inserted in position between the plates to the desired depth. The gel was allowed to polymerise for 1 h.

The comb and clamps were removed once the gel has polymerised. The plates were mounted onto the electrophoretic unit, with the longer plate outermost. TBE (1X) buffer (800 ml) was poured into the top and bottom reservoirs. After pre-running for 1 h, the preformed well area was flushed using a syringe to remove any loose/un-polymerised acrylamide and urea tucked along the edge of the small plate to prevent interference with sample loading. The comb was placed again into a position so that the points of comb were at the edge of the longer plate.

The DNA samples (5 µl) and 2 µl of PAGE loading dye (98 % formamide, 10 mM EDTA pH 8.0, 0.25 % bromophenol blue and 0.25 % Xylene cyanol) were prepared for denaturing by heating at 94 $^{\circ}$C for 5 min in Pelkin Elmer MJ machine. The PCR tubes containing the samples were quickly placed on ice. Each sample (5 µl) was loaded into the well of the gel with DNA standard. The machine runs at 2500 V, 60 mA and 50 W for 2-3 h and after the first dye until the darker blue (bromophenol blue) dye runs off into the buffer (1X TBE). The plates were removed, the tape was removed and the plastic wonder wedge was used to separate the two plates. The gel was completely in contact with the short bind silane plate.

3.3.1.2 Silver staining procedure

The gel/plate was immersed in a shallow tray containing 2 L of fixing solution for 20-30 min with gentle shaking and washed 3 times in fresh distilled water for 2 min with agitation on a shaker. With each rinsing, the gel was lifted out of the wash and allowed to drain 10-20 sec before transferring it to the next wash. Immediately after the last wash the

silver stain solution was added and left on a shaker for half an hour. The gel was later rinsed in 2 L of distilled water for 5-10 sec to remove excess silver stain from the gel surface, while the gel was drained. The developing solution was poured into a tray and the gel was placed in it. The tray was tipped to ensure that the developer covered the gel evenly. After about 10 min, when the developer was turning black and when the bands (fragments) began to develop, the top end of the plate was immersed up. This way the middle / smaller sized fragments remain in the developer longer. The gel was rinsed twice in a tray containing 2 L of distilled water for 2 min each and then removed when all the bands appeared and the reaction was stopped by adding the 2 L of fixing solution saved from earlier fixation. The gel was rinsed twice in a tray containing distilled water for 2 min each. It was hanged up to air dry.

3.3.1.3 Gel scoring of AFLP gels

After drying, image of plates were scanned and polymorphic fragments (bands) were coded manually in binary form as 1or 0 for their presence or absence of band respectively in each genotype.

3.3.2 SSR-based assessment

3.3.2.1 Oligonucleotide primers

Oligonucleotide primers were synthesized at nanomole concentration by Integrated DNA Technologies Leuven, Belgium (IDT) Primer Company. All oligonucleotide primers were diluted to a working concentration of 5 µM with sterile water and stored at -20 $^{\circ}$C.

3.3.2.2 Preliminary screening of SSR primers

Before proceeding to detailed evaluation of the markers, primers were verified to amplify the desired region, and optimization of polymerase chain reaction (PCR) was carried out for 152 maize SSR primer pairs; they were screened using DNA of two genotypes. A total of 87 SSR maize primers were used for PCR amplification of the 38 inbred lines; 35 SSR markers used in this study were chosen from the MaizeGDB database (http://nucleus.agron.missouri.edu/cgi-bin/ssr_bin.pl), previous studies (Smith *et al.*, 1997; Senior *et al.*, 1998) and the 52 core SSR primers described by Warbuton *et al.* (2002). They were all selected based on the bin locations, which provides a uniform coverage of all the ten chromosomes in the maize genome. A total of sixty two SSR markers were selected from the MAIZEGDB database (http://www.maizegdb.org) and the 52 core SSR primers described by Warburton *et al.* (2002) and based on bin locations in order to provide a uniform coverage of the genome were used for PCR reactions for the 122 inbred lines.

3.3.2.3 Polymerase chain reaction (PCR) amplification

Amplification PCR reactions were performed in a MJ Research (PTC-200) Peltier thermal cycler in a 25 µl reaction mixture. Each reaction contained 2.0 mM $MgCl_2$, 200 mM dNTPs, 200 mM of each forward and reverse primer, 1 µl of DMSO, 0.15U of *Taq* polymerase, and 2.5 µl (100 ng) of genomic DNA template. The amplification consisted of a denaturation step of 1 min at 95 $^{\circ}$C followed by a "touch down" PCR profile as described by Senior *et al.* (1998)**.** This profile annealing temperature was varied from 65-55 $^{\circ}$C for the PCR amplification conditions with an initial denaturation at 94 $^{\circ}$C for 2 min,

followed by 9 cycles of 93 °C for 15 sec; and 65°C for 20 sec. The annealing temperature was then reduced by 1 °C per each cycle and followed by 24 cycles (15 sec at 93 °C; 55 °C for 20 sec and 72 °C for 15 min) of denaturation, annealing and extension steps respectively. A final extension step of one cycle of 72 °C for 5 min was carried out. Variable "touch down" annealing temperatures (70°C-63 °C and 60°C-50 °C) were used for different SSR primers.

3.3.2.4 Superfine agarose gel electrophoresis

The SSR loci amplified were separated on 2 % (w/v) superfine agarose gel (Amresco). Superfine agarose gels (3 g) (SFR) was dissolved by melting in a microwave oven in 150 ml of 0.5X TBE (Trisma-base-boric acid-0.5 M EDTA) by strictly following the manufacturer's (AMRESCO) detailed protocol for high resolution. After electrophoresis, the gel was stained with ethidium bromide solution and fragments were visualised under UV light attached to a gel documentation system. Simple Sequence Repeat (SSR) alleles using this protocol were effectively resolved with a resolution of about three base pairs as revealed by the molecular ladders run on each gel. Allele sizes of amplified fragments were scored on the basis of size in comparison with DNA molecular weight markers.

3.3.2.5 Gel scoring of SSR gels

Data were scored based on the presence (1) or absence (0) of each allele with columns representing the inbred lines and rows the different SSR markers to give a binary matrix.

3.3.3 SNP-based assessment

3.3.3.1 Primer design and preliminary screening of SNP primers

The *PSY1* gene used for this study was selected based on the location in the carotenoid pathway. The SNP characterization of these maize lines was prompted by the study of Palaisa *et al.* (2003). In order to have an insight into the nature of SNP polymorphism in tropical maize, two regions were selected from this sequence. Two primer pairs were designed from the published *PSY1* sequence deposited in the NCBI/GenBank database under accession number U32636 (Buckner et *al.*, 1990). Primers were designed using the Primer 3.0 program software with appropriate conditions (primer size of approx. 18 bases, product size between 400-600 base pairs, 60-70 $^{\circ}$C of T_m, ideal GC content of 50 %, no more than three consecutive identical nucleotides and a 2-base GC clamp). The primers are Y1-4F and Y1-4R in the region of 840 bp and Y5P-F and Y5P-R in the region of 1,040 bp (Table 3.3) and synthesized at nanomole concentration by Integrated DNA Technologies Leuven, Belgium (IDT) Primer Company. All oligonucleotide primers were diluted to a working concentration of 5 µM with sterile water and stored at -20 $^{\circ}$C. Polymerase chain reaction (PCR) for each primer set was used to amplify genomic DNA of the two genotypes for optimisation.

Table 3.3: Designed primer sequences used for *PSY1* Genotyping

Gene symbol	Primer sequence (5'→3')
Y1-4F	TCCACCACAAGAAGATGC
Y1-4R	GACGTCGTAGACCTTCTGC
Y5P-F	GTCTCTCCCGCCTTCTTTCT
Y5P-R	CCGCTCCTCTGTCATCAAC

3.3.3.2 Polymerase chain reaction (PCR) amplification and SNP genotyping

PCR reaction was performed in a total volume of 25 µl containing: 100 ng DNA template, 2.0 mM $MgCl_2$, 200 mM dNTPs, 200 mM of each forward and reverse primer, 1 µl of DMSO, and 0.15 U of Taq DNA polymerase (Bioline). The PCR reactions were carried out using a MJ Research PTC-200 Peltier thermal cycler. The cycling conditions for a touch down PCR programme for Y1-4F and Y1-4R primer pair were as follows: 94 °C for 2 min, [93 °C for 15 sec and 65 °C for 20 sec -reduction of 1 °C per each cycle] for 10 cycles, [93 °C for 15 sec; 55 °C for 20 sec and 72 °C for 15 min] for 24 cycles and a final extension step of one cycle at 72 °C for 5 min. A touch down PCR programme TD58-48 for the primer pairs Y5P-F and Y5P-R was used. Amplified products were separated using 2.0 % agarose gel electrophoresis in TBE buffer (Trisma base, boric acid and 0.5 M EDTA, pH 8.0) and stained with ethidium bromide to confirm the presence of products. Following the testing, the designed primer sets were used to amplify all the thirty eight maize genotypes using the same conditions described above.

The PCR products were purified with ethanol precipitation for DNA sequencing as follows: two volumes of 100 % ethanol (40 µl) was added to each product, carefully mixed and incubated for 15 min at room temperature. Thereafter it was centrifuged for 15 min at room temp at 10,000 rpm, washed with 70 % ethanol and again centrifuged at 10,000 rpm for 10 min. The ethanol was discarded, the pellet was allowed to dry at room temperature and the DNA was re-suspended in 20 µl of ultra pure water. The quantification of the amount and quality of purified DNA were carried out using

NanoDrop spectrophotometer. The purified DNA products were visualized after electrophoresis on 1.5 % (w/v) agarose.

3.3.3.3 Sequencing reaction and purification of extension products

Sequence analysis was performed for all 38 maize inbred lines using ABI BigDye Terminator Cycle Sequencing (Applied Biosystems, Forster City, CA, USA) at IOWA State University, USA. For direct sequencing, the following reagents were added to each tube: BigDye Terminator v3.1 (1.0 µl), 2.5X buffer (3 µl), primer (1.5 µl), template (1 µl) and deionised water (3.5 µl) to make a total volume of 10 µl. The forward PCR primer was used as sequencing primer.

The PCR cycle sequencing was carried out in an AB thermal cycler. An initial denaturation at 96 $^{\circ}$C for 1 min (rapid thermal ramp to 96 $^{\circ}$C -1 $^{\circ}$C /second, followed by 25 cycles (96 $^{\circ}$C for 10 sec, 50 $^{\circ}$C for 5 sec and 60 $^{\circ}$C for 4 min) of denaturation, annealing and extension steps respectively and holding at 4 $^{\circ}$C until ready to purify. Ethanol/EDTA/Sodium acetate precipitation method was used for purification of extension products. To each reaction 125 mM EDTA (1.0 µl), 1.0 µl of 3 M sodium acetate and 25 µl 100 % ethanol was added. The plate was sealed with aluminium foil, followed by mixing by inverting four times and incubated for 15 min at room temperature (32 $^{\circ}$C). Cold centrifugation was performed at 3700 rpm for 30 min at room temperature (32 $^{\circ}$C). The ethanol was discarded by inverting on paper towel and spinning was carried out for 1 min at 1000 rpm. Ethanol (70 % v/v, 70 µl) was added to each reaction and

centrifuged for 5 min at 3700 rpm. The ethanol was discarded and the pellet was dried at room temperature for about 5 min. DNA was re-suspended in 10 µl ultra pure water.

3.3.3.4 Electrophoresis on ABI genetic analyzer, alignment of sequence

Formamide (9 µl) was added to the cleaned DNA (1.0 µl), denatured at 94 $^{\circ}$C for 5 min in a PCR machine and chilled immediately on ice and later sequenced on an ABI DNA analyser. The sequences for the two primer pairs designed from the *PSY* sequence were analysed across a diverse set of 38 tropical-adapted lines.

3.4 Carotenoid composition based diversity

Carotenoid concentrations (trans-β-carotene and cis-β-carotene, zeaxanthin, total pro-vitamin A, lutein, β-cryptoxanthin, α-carotene) that have been measured from seed samples of yellow inbred lines harvested from a trial grown in one location using HPLC by Menkir et *al.* (2008) was used for the carotenoid composition-based diversity assessment. The 38 and 122 inbred lines with contrasting concentrations of pro-vitamin A were selected from this trial for this study (Tables 3.4 and 3.5). The carotenoid data was used for grouping the lines so that the resulting clusters will be compared with clusters generated by the two molecular markers.

Table 3.4: Carotenoid contents of the 38 tropical-adapted yellow endosperm maize inbred lines (Taken from Menkir *et al*., 2008).

Inbred	Total β-carotene (mg/g)	Total pro-vitamin A (mg/g)	Lutein (mg/g)	Zeaxanthin (mg/g)	β-cryptoxanthin (mg/g)	α-carotene (mg/g)	Trans β-Carotene (mg/g)
PVL01	3282	6058	10847	16295	4843	709	2200
PVL02	3571	5557	3809	11387	3803	168	2500
PVL03	2610	5207	10972	11872	4225	969	1700
PVL04	3071	5085	5514	10386	3783	246	2100
PVL05	3896	6114	9614	13504	3698	738	2500
PVL06	3774	6862	3603	9799	5985	191	2400
PVL07	3838	5877	7011	13648	3656	420	2700
PVL08	4190	6129	5845	13491	3661	218	3000
PVL09	3965	6179	5981	16054	4183	245	2800
PVL10	538	872	1196	902	601	96	126
PVL11	4749	7752	18172	24501	5504	501	2647
PVL12	3898	5511	8887	6462	2679	545	1372
PVL13	3101	5951	8210	9906	5374	327	1537
PVL14	3006	5484	7474	10744	4607	348	1541
PVL15	8277	9792	5410	13864	2686	344	6300
PVL16	2385	5317	6979	13751	5330	533	1260
PVL17	1028	1845	2893	10040	1385	250	475
PVL18	607	863	2324	6025	381	133	258
PVL19	953	1698	3874	4666	1304	186	448
PVL20	1096	1802	5058	9312	1222	189	564
PVL21	2563	5003	4759	6877	4559	322	1469
PVL22	2877	5017	3390	11617	4021	260	1683
PVL23	2624	5114	3509	10805	4750	230	1397
PVL24	3019	5287	7410	18600	4266	270	1725
PVL25	2854	5051	8973	19281	4075	318	1655
PVL26	3161	5421	3537	28081	4349	172	1664
PVL27	3243	5736	3514	32395	4815	171	1769
PVL28	637	1689	7900	6007	716	1388	400
PVL29	692	1223	4368	6125	924	138	280
PVL30	764	1024	3775	1496	425	95	319
PVL31	372	433	1830	285	70	82	88
PVL32	1670	1998	13699	2455	199	653	303
PVL33	4579	7872	5531	14340	6281	305	3100
PVL34	3399	5496	3371	11744	4050	144	2300
PVL35	4249	6432	8053	13511	3955	411	2900
PVL36	3471	5402	6860	11718	3573	290	2200
PVL37	1927	3842	6175	6408	3385	446	1027
PVL38	2561	4544	6576	20878	3712	253	1405

Table 3.5 Carotenoid contents of the 122 tropical-adapted yellow endosperm maize inbred lines (Taken from Menkir *et al.*, 2008)

Inbred	Total β-carotene (mg/g)	Total pro-vitamin A (mg/g)	Lutein (mg/g)	Zeaxanthin (mg/g)	β-cryptoxanthin (mg/g)	α-carotene (mg/g)	Trans β-Carotene (mg/g)
PV001	4749	7752	18172	24501	5504	717	2647
PV002	3471	5402	6860	11718	3573	290	2200
PV003	3707	5564	4419	12980	3504	210	2380
PV004	3513	6046	6339	12952	4763	304	2253
PV005	3565	5958	5020	11258	4519	268	2310
PV006	3580	6116	3186	5656	4919	154	2235
PV007	3474	5963	3474	6512	4819	158	2147
PV008	5239	6614	5911	6583	2503	247	3851
PV009	5526	6839	5774	8499	2394	231	4141
PV010	5754	7038	5547	7932	2341	227	4315
PV011	5423	6557	4955	6705	2041	228	4157
PV012	5816	7142	5276	7376	2416	237	4408
PV013	4987	6144	4486	6338	2095	219	3832
PV014	5229	6506	5114	6398	2350	203	3977
PV015	5515	6730	4894	6415	2209	219	4237
PV016	4498	5826	4737	8279	2436	220	3355
PV017	3590	5712	4993	11650	4083	161	2235
PV018	3805	5689	5470	12360	3603	164	2414
PV019	3670	5493	4812	11704	3460	186	2419
PV020	3569	5779	5534	14128	4241	179	2279
PV021	4749	7752	18172	24501	5504	717	2647
PV022	3513	6046	6339	12952	4763	304	2253
PV023	3580	6116	3186	5656	4919	154	2235
PV024	5526	6839	5774	8499	2394	231	4141
PV025	4987	6144	4486	6338	2095	219	3832
PV026	4580	5919	10217	6741	1669	1010	3116
PV027	9094	10268	10239	1423	719	1631	6493
PV028	3708	5470	8051	26034	2759	766	2604
PV029	8324	9325	7128	2102	336	1666	6471
PV030	4541	6012	11685	14751	2009	934	3243
PV031	8543	10271	3473	12398	1923	1535	6893
PV032	8998	10103	1881	594	363	1848	7316
PV033	3942	5848	8915	9529	2989	824	2577
PV034	6844	7796	7908	3014	616	1288	5105
PV035	6986	8678	3507	17163	1945	1440	5346
PV036	6425	7465	7563	2546	937	1142	4950
PV037	7530	8419	6807	71	332	1447	5795
PV038	7080	8000	17168	4780	379	1462	5238

Table 3.5 Carotenoid contents of the 122 tropical-adapted yellow endosperm maize inbred lines (Taken from Menkir *et al.*, 2008).

Inbred	Total β-carotene (mg/g)	Total pro-vitamin A (mg/g)	Lutein (mg/g)	Zeaxanthin (mg/g)	β-cryptoxanthin (mg/g)	α-carotene (mg/g)	Trans β-Carotene (mg/g)
PV039	5508	6205	11515	1887	363	1030	4077
PV040	3430	5263	3096	7991	2551	1115	1951
PV041	4926	7075	8276	22256	3302	996	3260
PV042	5968	7000	8164	3206	1205	860	4350
PV043	7448	8821	4230	6123	1697	1049	5942
PV044	5772	7019	6124	7136	1409	1086	4328
PV045	6893	8127	26299	4447	1219	1250	5001
PV046	5895	6889	2109	2196	323	1665	4147
PV047	4929	5725	6998	7610	688	904	3774
PV048	4568	5882	17751	8577	1544	1085	3036
PV049	3036	5298	3788	8583	4323	203	1953
PV050	2943	5206	4262	9746	4328	198	1941
PV051	2943	5206	4262	9746	4328	198	1941
PV052	2943	5206	4262	9746	4328	198	1941
PV053	2943	5206	4262	9746	4328	198	1941
PV054	3082	5270	4363	9979	4148	230	2036
PV055	3513	6046	6339	12952	4763	304	2253
PV056	3565	5958	5020	11258	4519	268	2310
PV057	3565	5958	5020	11258	4519	268	2310
PV058	3474	5963	3474	6512	4819	158	2147
PV059	5239	6614	5911	6583	2503	247	3851
PV060	5754	7038	5547	7932	2341	227	4315
PV061	5816	7142	5276	7376	2416	237	4408
PV062	5816	7142	5276	7376	2416	237	4408
PV063	5229	6506	5114	6398	2350	203	3977
PV064	5515	6730	4894	6415	2209	219	4237
PV065	4498	5826	4737	8279	2436	220	3355
PV066	3421	5547	5256	13067	4065	187	2169
PV067	3421	5547	5256	13067	4065	187	2169
PV068	3805	5689	5470	12360	3603	164	2414
PV069	3805	5689	5470	12360	3603	164	2414
PV070	3670	5493	4812	11704	3460	186	2419
PV071	3227	5111	4666	12039	3628	139	2030
PV072	2654	5000	3957	12861	4559	133	1562
PV073	2718	5110	4105	14577	4658	125	1622
PV074	3050	5456	2521	4745	4663	149	1927
PV075	3050	5456	2521	4745	4663	149	1927
PV076	2919	5187	3194	5340	4398	138	1892

Table 3.5 Carotenoid contents of the 122 tropical-adapted yellow endosperm maize inbred lines (Taken from Menkir *et al.*, 2008).

Inbred	Total β-carotene (mg/g)	Total pro-vitamin A (mg/g)	Lutein (mg/g)	Zeaxanthin (mg/g)	β-cryptoxanthin (mg/g)	α-carotene (mg/g)	Trans β-Carotene (mg/g)
PV077	2844	5044	2704	5177	4271	130	1784
PV078	3667	6309	3109	4565	5115	169	2415
PV079	3327	5707	3396	5420	4578	183	2235
PV080	2768	5083	3258	6086	4487	142	1750
PV081	2768	5083	3258	6086	4487	142	1750
PV082	3311	5770	2806	4102	4764	153	2244
PV083	3311	5770	2806	4102	4764	153	2244
PV084	3311	5770	2806	4102	4764	153	2244
PV085	2979	5181	3467	5454	4234	169	1912
PV086	3320	5694	2828	4788	4599	150	2273
PV087	2892	5125	2657	5070	4338	129	1885
PV088	6328	7322	4220	7487	1575	413	4431
PV089	3709	5258	3474	12699	2948	149	2245
PV090	3957	6303	4796	15196	4494	197	2258
PV091	4011	5292	6084	13467	2369	194	2283
PV092	3905	5089	5780	12233	2168	200	2238
PV093	3942	5751	9114	11633	3339	279	2272
PV094	4522	6632	5669	10753	3981	239	2470
PV095	3969	5861	6455	21777	3431	352	2371
PV096	5700	8697	10519	18658	5744	251	3400
PV097	3641	5580	8667	21079	3719	159	2076
PV098	3837	6460	5743	13077	5050	196	2368
PV099	4277	7237	6428	12908	5702	216	2817
PV100	3554	5574	4289	12066	3918	121	1741
PV101	3700	5317	3995	10784	3070	165	1729
PV102	3092	5032	6527	13434	3646	234	1686
PV103	4577	7641	9037	26199	5852	278	2592
PV104	4008	6143	4004	6552	4088	182	2125
PV105	8669	9163	9111	2013	312	675	6274
PV106	4169	6163	11851	7514	3530	457	2319
PV107	7764	8339	9502	3474	583	567	5783
PV108	3038	5100	16621	10237	3284	839	1404
PV109	4491	6822	5579	16374	4420	241	2346
PV110	3239	5332	6650	11186	3832	354	1463
PV111	4264	6335	2049	4465	3985	158	2382
PV112	4320	5020	1107	734	1127	272	2728
PV113	6608	7086	435	366	791	166	5054
PV114	7065	8648	3828	5123	2924	244	4627

Table 3.5 Carotenoid contents of the 122 tropical-adapted yellow endosperm maize inbred lines (Taken from Menkir *et al.*, 2008).

Inbred	Total β-carotene (mg/g)	Total pro-vitamin A (mg/g)	Lutein (mg/g)	Zeaxanthin (mg/g)	β-cryptoxanthin (mg/g)	α-carotene (mg/g)	Trans β-Carotene (mg/g)
PV115	3520	5631	6423	15168	4011	209	1526
PV116	3413	5228	7274	13186	3209	421	1824
PV117	4087	6376	5022	15265	4347	231	1910
PV118	6194	7661	2494	2424	2707	228	3851
PV119	4835	6904	2938	5154	3962	175	2787
PV120	3234	5104	2984	7065	3567	174	2100
PV121	2755	4377	5926	17309	3092	151	1900
PV122	2605	3995	5538	15843	2587	194	1600

3.5 Molecular analysis of allele-specific PCR markers of β-carotene candidate genes

The two candidate carotenoid genes known to be important in the high accumulation of β-carotene in yellow maize endosperm were chosen for this study based on previously reported studies (Harjes *et al.,* 2008; Yan *et al.,* 2010; Vallabhaneni *et al.,* 2009). Allele-specific primer pair (markers) representing DNA sequence variants i.e polymorphisms developed by Harjes *et al.* (2008), Yan *et al,* (2010), and Vallabhaneni *et al.* (2009) were used for amplification of alleles at the *LCYE* and *crtRB1/HYD3* loci (Table 3.6). Information on primer sequences for the polymorphisms were obtained from published supporting online material (Harjes *et al.,* 2008; Yan *et al.,* 2010; Vallabhaneni *et al.,* 2009) and synthesized by Integrated DNA Technologies Leuven, Belgium (IDT) Primer Company.

Table 3.6 : Sets of gene-specific markers for genotyping assays of single nucleotide polymorphisms in *LCYE* and *CrtRB1 (HYD3)* genes of maize carotenoid pathway.

Primer set	Genetic polymorpism types	Forward and reverse primer sequences (5'→3')	References
S216-L1	*LCYE*-Exn-216	F:GCGGCAGTGGGCGTGGAT	Harjes *et al.*, 2008
S216-R1		R:TGAAGTACGGCTGCAGGACAACG	
3pINDL-L1	*LCYE3*' INDEL	F:GTACGTCGTTCATCTCCCGTACCC	Harjes *et al.*, 2008
3pINDL-R1		R:CTTGGTGAACGCATTTCTGTTGG	
3pINDL-L2		FGGACCGGAACAGCCAACTG	
3pINDL-R2		R:GGCGAAATGGGTACGGCC	
TE103PF F-1	*LCYE5*' INDEL/TE	F:CGCTAGCAAGCCCATTATTTTTA	Harjes *et al.*, 2008
TE103PR R-1		R:CGGTATGGTTTTTGGTATACGG	
TE105PR R1		F:GAGAGGGAGACGACGAGACAC	
ZGt111204-976R(1) F1		R:AGCATCCGACCAAAATAACAGA	
***CrtRB1* H1UF**	*CrtRB15*'TE	F:TTAGAGCCTCGACCCTCTGTG	Yan *et al.*, 2010
***CrtRB1* H1UR**		R:AATCCCTTTCCATGTTACGC	
***CrtRB1* D4F**	*CrtRB1*D4	F:ACCGTCACGTGCTTCGTGCC	Yan *et al.*,.2010
***CrtRB1* D4R**		R:CTTCCGCGCCTCCTTCTC	
***CrtRB1*65F**	*CrtRB13*'INDEL/TE	F:ACACCACATGGACAAGTTCG	Yan *et al.*, 2010
***CrtRB1*62R**		R:ACACTCTGGCCCATGAACAC	
***CrtRB1*66R**		R:ACAGCAATACAGGGGACCAG	
***HYD3* (Cc) F**	HYD3-Duplicated Seq	F:GACTTGTGAGCAAGGGGAAG	Vallabhaneni *et al.*, 2009
***HYD3* (Cc) R**	transcript start site	R:GACGTGACTCCGAGGCTAGA	
***HYD3* (C) F**		F:AACACTCCCGCTCCCGCGGCTCG	
***HYD3* (A) R**		R:TTATATGGATAGTTCACATACCTC	
***HYD3*(B) F**		F:AACACTCACGCTCCCGCG	

3.5.1 Lycopene epsilon cyclase (*LCYE*) gene

3.5.1.1 Primers for PCR assays for SNP216, 5'TE, 3'TE polymorphisms

PCR-based agarose markers developed by Harjes *et al.* (2008) based on the sequence information of lycopene epsilon cyclase (*LCYE*) for three important functional polymorphisms (*LCYE*-Exn-SNP216, *LCYE* 5'TE, *LCYE* 3'TE) were used to detect genetic variations. To detect a single nucleotide polymorphism, the T (unfavourable allele) is replaced by G (favourable allele)] at 216^{th} base position in the first exon of *LCYE*; the forward and the reverse primers S216-L1 and S216-R1 respectively were used. To determine the TE insertion or deletion in the 5' promoter, four primers (TE103PF F-1, TE103PR R-1, TE105PR R and ZGT111204-976R (1) F1) were used. Four (4) allele specific markers (3pINDL-1, 3pINDL-R1, 3pINDL-L2 and 3pINDL-R2) were used to determine 8-base pair insertion/deletion in the 3'UTR region of *LCYE* gene. The same sets of maize inbred lines, thirty-eight (set 1) and one hundred and twenty-two (set 2) were studied.

3.5.1.2 Optimization of *LCYE* primers

Optimization of polymerase chain reaction (PCR) was carried out for all the combinations of *LCYE* primers using various PCR reaction modifications, all to obtain reliable, high specificity and efficiency of amplification of each primer pair. Optimization of PCR conditions such as primer concentrations (2.5-5 µM), $MgCl_2$ (1-2 mM) and annealing temperature (54-64 $^{\circ}$C) was always carried out with each pair-specific primer to amplify expected different allele sizes for the polymorphisms. Primer sets were tested on four

maize genotypes and PCR conditions optimized before genotyping was performed on the full maize inbred line sets.

3.5.1.3 PCR amplification

For all *LCYE* PCR assays, unless otherwise specified, 25 µl reaction volume contained 2.0 mM $MgCl_2$, 200 mM dNTPs, 200 mM of each forward and reverse primer, 1 µl of DMSO, 0.15 U of *Taq* DNA polymerase, and 2.5 µl 100 ng of genomic DNA template. The PCR reactions were performed using Applied Biosystems Veriti 96 well PCR thermal cycler (AB, USA). Following Harjes *et al.* (2008), PCR was programmed for an initial denaturation of 3 min at 94 °C, 5 cycles of 94 °C for 1 min and 64 °C for 1 min-reduction of 1 °C per each cycle and at 72 °C for 1.5 min. This was followed by 28 cycles (1 min at 94 °C, 58 °C for 1 min and 72 °C for 1.5 min) of denaturation, annealing and extension steps respectively. A final extension step of one cycle of 72 °C for 5 sec was followed by a hold step at 4 °C. This program was used for the amplification of the following primers: S216-L1, S216-R1; 3pINDL-L1, 3pINDL-R1; 3pINDL-L2, 3pINDL-R2; TE103PF F-1(F1), TE103PR R-1 (R1); TE105PR R-1- (R2), ZGT111204-976R (1) F1(F2); TE103PF F-1(F1), TE105PR R-1(R2); TE103PR R-1-(R1), ZGT111204-976R (1) F1(F2). The PCR profile used for 3pINDL-L2, 3pINDL-R1 primers consist of initial denaturation of 5 min at 94 °C, 19 cycles of 94 °C for 1 min and 64 °C for 1 min reduction of 0.5 °C per each cycle and at 72°C for 1.5 min and followed by 19 cycles (1 min at 95 °C, 54 °C for 1 min and 72 °C for 1.5 min) of denaturation, annealing and extension steps respectively (Harjes *et al.,* 2008). A final extension step of one cycle of 72 °C for 10 min was followed by a hold step at 4 °C. Primer concentration is similar for S216-L1, S216-

R1 (100 mM of each forward and reverse primer using 1.75 mM $MgCl_2$; 3pINDL-L, 3pINDL-R (1.50 mM $MgCl_2$).

3.5.2 β-carotene hydroxylase 1 *(HYDb1) or (crtRB1*) gene

3.5.2.1 Primers for PCR-based assays for *CrtRB1* InDel4, 3'TE, 5'TE polymorphisms

Molecular markers described by Yan *et al.* (2010) for detecting the polymorphisms in *crtRB1* gene derived from the sequence information of *crtRB1* were used. To detect a 12 bp deletion in the first exon of *crtRB1,* four primers (HYDB77F, HYDB76R, HYDB55F and HYDB58R) were used. They were combined as follows for amplification: HYDB77F, HYDB76R; and HYDB55F, HYDB58R. To determine the insertion of TE in the 3' region of *crtRB1* in the first exon, three primers (HYDB65F, HYDB62R and HYDB66R) were used as 3 primer allele specific PCR system to amplify alleles for the polymorphism. Two primers (*crtRB1* H1UF and *crtRB1* H1UR) were used for genotyping assay to amplify the TE alleles in the 5' region of *crtRB1*. The same sets of maize inbred lines, thirty-eight (set 1) and one hundred and twenty-two (set 2) were studied.

3.5.2.2 Optimization of *crtRB1* primers

The same procedures for *LCYE* primers optimisation were followed.

3.5.2.3 PCR amplification

Amplification reactions were carried out in 25 µl reaction mixture as described above with adjustments in primer and magnesium chloride concentrations in *HYD* primers;

Varying $MgCl_2$ concentrations were used for different primers; HYDB77F, HYDB76R (1.25 mM), HYDB55F and HYDB58R (1.5 mM), *crtRB1* H1UF and *crtRB1* H1UR (2 mM) and cocktail of HYDB66R, HYDB65F and HYDB62R primers (1.25 mM).The primer concentration for primer HYDB66R was 100 mM. The PCR reactions were performed using Applied Biosystems Veriti 96 well PCR thermal cycler (AB, USA). For crtRB1 3'TE, PCR protocol followed was according to Yan *et al.* (2010); 5 min at 94 °C, 34 cycles of 94 °C for 1 min and 57 °C and followed by 72 °C for 4 min and a final 72 °C for 10 min. PCR protocol for crtRB1 5'TE consist 5 min at 94 °C, 9 cycles of 94 °C for 1 min and 64 °C for 1 min decreasing of 0.5 °C per each cycle and at 72 °C for 1.5 min. This was followed by 26 cycles (1 min at 95 °C, 54 °C for 1 min and 72 °C for 1.5 min) and a final extension step of 72 °C for 10 min (Yan *et al.*, 2010) . For *crtRB1* InDel4, PCR protocol followed 3 min at 94 °C, 10 cycles of 94 °C for 1 min and 68 °C for 1 min reduction of 1 °C per each cycle and at 72 °C for 1.5 min. This was followed by 30 cycles (1 min at 95 °C, 58 °C for 1 min and 72 °C for 1.5 min) and a final extension step of 72 °C for 10 min (Yan *et al.*, 2010).

3.5.3 Hydroxylase 3 *(HYD3)* gene

3.5.3.1 Primers for PCR-based assays for *HYD3* polymorphism

To amplify three different *HYD3* variations (A, Band C alleles) found in a ~ 40 bp region adjacent to the transcript start site, five primers (*HYD3* (P1) F, *HYD3* (P2) *R*, *HYD3* (C) F, (*HYD3* (A) R and *HYD3* (B) F) were used to distinguish between A and C or between A and B alleles and a common *HYD3* paralog-specific product. The multiplex PCR assay for amplification of *HYD3* alleles as advocated by Vallabhaneni *et al*. (2009) did not

work. Consequently, the following modification was adopted. The primers were combined in pairs according to amplification direction to amplify allele-specific regions. For allele A; P1 (F) and allele A specific primers (R) were used. For allele B; allele B specific primer (F) and P2 (R) and for allele C; allele C specific primer (F) and P2 (R) were used.

3.5.3.2 Optimization of *HYD3* primers

The primer optimisations were followed as above.

3.5.3.3 PCR amplification

Amplification reactions were carried out in 25 µl reaction mixture as described above. Amplification PCR reactions were performed using Applied Biosystems Veriti 96 well PCR thermal cycler (AB, USA). The PCR reactions used 1.75 mM $MgCl_2$ using the following program: a denaturation step at 94 °C for 3 min, 40 cycles of 94 °C for 45 sec, 54 °C for 45 sec and 72 °C for 1 min and a final step of 72 °C for 10 min.

3.5.4 Gel electrophoresis

PCR products were separated by electrophoresis for 4 h in a 2 % agarose gel run in 0.5X TBE buffer. Gels were later stained with ethidium bromide, followed by destaining in deionized water and exposed to ultra violet light (UV) to visualize DNA fragments. Photographs were taken using gel documentation system. Size of each band was estimated by means of a 50 bp DNA Ladder. Scoring of bands was based on the allele sizes and / classes (haplotypes) of target polymorphisms for each maize inbred genotype

in the two sets. Classes of alleles were scored accordingly (Harjes *et al.*, 2008; Yan *et al.*, 2010; Vallabhaneni *et al.*, 2009).

3.6 Data analysis

3.6.1 Molecular markers

For the AFLP marker data, total number of alleles in all loci and mean number of alleles were determined. To determine the genetic variability among the 38 inbred lines, the following indicators were considered for the polymorphic SSR bi-allelic loci; allelic richness, total number of alleles in all the SSR loci and the mean number of alleles. Polymorphic information content (PIC) values were calculated at each locus using the formula according to Weir (1996)

$$PIC = 1 - \sum_{i=1}^{n} f_i^2$$

where f_i is the allele frequency for i^{th} allele and n is the number of alleles analysed. PIC values give an estimate of the discriminatory power of a marker by taking into account not only the number of alleles at the locus but also the relative frequencies of these alleles.

3.6.2 Estimates of genetic distances between inbred lines

The genetic diversity between pairs of inbred lines i and j was computed based on Modified Roger's Distance (Rogers, 1972) for the SSR data and based on Jaccard's coefficient of similarity (Jaccard, 1908) for the AFLP data. For all molecular data, genetic distances were computed with Winboot software.

3.6.3 Carotenoid data

Euclidean distance estimates among inbreds were calculated from carotenoid data. We first standardized all mean carotenoid values of the lines using SAS (version 8, Cary, North Carolina, USA) before using them to calculate Euclidean distance (Sneath and Sokal, 1973).

3.6.4 Clustering analyses

Genetic distance matrices generated from the SSR, AFLP and carotenoid composition data sets were then subjected to cluster analysis using a sequential agglomerative hierarchical nested clustering method (SAHN), based on the unweighted pair-group method with arithmetic averages (UPGMA), as suggested by Sneath and Sokal (1973). In addition, the cophenetic correlation coefficient, which measures the correspondence between the original distances and the distances defined in the dendrogram (Sneath and Sokal, 1973), was calculated to test if the data actually contain clusters. When the two matrices show the same clustering patterns, they will produce high cophenetic correlation, indicating low distortion (Rohlf and Fisher, 1968). In this study the cophenetic values were determined for the SSR, AFLP and carotenoid based clusters using the MXCOMP procedure of NTSYSpc 2.01 (Rohlf, 1997). The significance of cophenetic correlation was determined by the Mantel test (Mantel, 1967) based on 1000 permutations. Furthermore, associations among maize inbred lines were determined using principal-coordinates analysis (Gower, 1966). The UPGMA and PCoA analyses and Mantel's test were performed with version 2.01 of the NTSYS-pc package (Rohlf, 1997; Exeter Software. Setauket, USA).

3.6.5 Correspondence analysis of matrices

Mantel correlation statistic was used to ascertain correspondence between pairs of distance matrices for estimating a linear relationship between matrices (Mantel, 1967). To compare the genetic distance matrices for all data sets, it was assumed that genetic distance values computed by these methods were normally distributed. The relationships between the Euclidean distance matrix based on carotenoid concentrations and the genetic distance matrices obtained with SSR and AFLP markers were analyzed. Significance of Z was determined by comparing the observed Z values with a critical Z value obtained from its permutation distribution. This distribution was derived by calculating Z values for one matrix with 1000 permutations of a second matrix.

3.6.6 SNP analysis in *PSY1* gene

Raw sequences were edited by using CodonCode aligner. Also, for proper alignments, all the sequences were trimmed at the same point in the alignments on the 5' and 3' end and used for analysis. Further editing, alignment and other manipulations were done with the BioEdit software program. Single nucleotide polymorphism survey was carried out in the *PSY1* gene after aligning edited sequences. SNP variations in the sequence were estimated by directly looking for SNPs [single DNA base differences between homologous DNA fragments plus small insertions and deletions (InDels)] that distinguish the lines.

3.6.7 SNP analysis of allelic variations of β-carotene candidate genes

Characterisation of favourable and unfavourable alleles revealing gene polymorphisms was computed for each class of marker or primer pairs and their frequency of occurrence evaluated for all inbred lines in each set.

Class allele frequency study of favourable allele

Class allele frequencies were calculated for favourable allele for each polymorphism as:

$$\text{Class allele frequency} = \frac{\text{Number of inbred lines in which a favourable allele amplified}}{\text{Total number of inbred lines}} \times 100$$

Haplotype scoring in genes

The best combinations of two genes were evaluated in the sets.

CHAPTER FOUR

4.0 RESULTS

4.1 GENETIC DIVERSITY IN YELLOW ENDOSPERM TROPICAL MAIZE INBRED LINES

4.1.1 Extent of polymorphism detected with AFLP and SSR markers

The AFLP and SSR markers were used to reveal the genetic diversity among the 38 yellow endosperm tropical maize inbred lines. The 17 AFLP primer combinations produced 716 polymorphic bands for the 38 inbred lines. The number of polymorphic bands detected ranged from 25 for primer combination ACC/CAC to 61 for primer combination AAC/CAG, with a mean of 42.1 (Table 4.1). The number of alleles detected among the 38 inbred lines using SSR markers and PIC values for polymorphic loci are presented in Table 4.2. Among the 87 SSR loci used for genotyping, 75 were polymorphic across the 38 inbred lines and produced a total of 297 alleles. The number of polymorphic alleles per locus varied from 2 to 11 with a mean of 3.96. The PIC values obtained for the polymorphic SSR markers varied from 0.17 (phi032) to 0.84 (bnlg2162) with an average of 0.56 (Table 4.2).

Table 4.1 Number of fragments generated with 17 AFLP primer pairs applied to 38 tropical-adapted yellow maize inbred lines.

Primer Combination	Number of polymorphic markers
ACC+CAA	48
AAC+CAC	47
AAC+CTG	47
AAC+CAT	33
AAC+CAG	61
AAG+CAC	42
AAG+CTC	57
ACA+CAC	25
ACA+CAG	33
ACC+CTC	48
ACC+CAC	25
ACG+CAT	39
ACT+CAA	48
ACT+CTA	54
AGC+CAA	30
AGG+CAG	31
AGG+CTA	48
Mean number of polymorphic markers	**42.1**

Table 4 2: Bin numbers, allele numbers and PIC values for SSR loci used to genotype 38 yellow endosperm maize inbred lines

SSR locus	Bin no.	No. of alleles	PIC value	SSR locus	Bin no.	No. of alleles	PIC value
umc1593	7.04	5	0.63	**phi064**	1.11	6	0.75
phi077	6.01	6	0.77	**phi072**	4.01	4	0.74
phi011	1	4	0.73	**phi059**	10.02	3	0.53
phi065	9.03	5	0.76	**nc133**	2.05	3	0.42
phi126	6	6	0.77	**phi015**	8.09	5	0.71
bnlg2162	4.08	11	0.84	**phi032**	9.04	2	0.17
phi075	6	5	0.67	**phi056**	1.01	4	0.32
bnlg1014	1	4	0.7	**phi070**	6.07	3	0.43
umc1154	7	4	0.63	**phi087**	5.06	4	0.6
bnlg118	5.07	5	0.79	**phi089**	6.08	3	0.51
nc004	4.03	4	0.52	**phi093**	4.08	2	0.21
umc1136	3.1	3	0.66	**phi109188**	5	3	0.56
phi002	1	5	0.79	**phi109275**	1	3	0.23
phi021	10.02	3	0.34	**phi109642**	2	3	0.59
phi042	4.02	3	0.52	**phi123**	6.07	3	0.47
phi047	3.09	5	0.69	**phi053**	3.05	3	0.48
umc1008	4.01	6	0.68	**phi227562**	1.12	5	0.71
umc1792	5.08	4	0.69	**phi233376**	8.03	4	0.68
phi014	8.04	4	0.56	**phi328175**	7.04	3	0.44
umc1312	10	3	0.36	**phi374118**	3.03	3	0.64
umc1426	7	7	0.77	**phi420701**	8.01	3	0.39
phi076	4.11	4	0.7	**phi423796**	6.02	2	0.23
phi025	10.04	5	0.75	**Phi448880**	9.05	2	0.18
phi001	1.03	7	0.59	**phi452693**	6.06	2	0.31
bnlg1721	2.08	10	0.82	**phi453121**	3	4	0.51
umc1403	1.03	3	0.6	**phi96100**	2	5	0.54
umc1592	8.01	4	0.71	**umc1061**	10.06	4	0.51
umc1530	8	2	0.39	**umc1109**	4.1	4	0.53
phi029	3.04	2	0.43	**umc1143**	3.1	3	0.59
phi033	9.01	3	0.49	**umc1152**	10.01	7	0.82
phi082	7.05	3	0.63	**umc1153**	5.09	4	0.69
phi034	7.02	4	0.71	**umc1161**	8.06	3	0.38
phi127	2.08	3	0.31	**umc1196**	10.07	4	0.74
umc1231	9.04	5	0.66	**umc1277**	9.08	3	0.56
umc1304	8.02	2	0.48	**umc1279**	9	4	0.56
umc1026	2.04	2	0.24	**umc1399**	3.07	4	0.65
phi041	10	3	0.47	**phi046**	3.08	3	0.61
phi084	10.04	3	0.46				
Total number of alleles						**297**	
Mean						**3.96**	**0.56**

The SSR markers were used to reveal the genetic diversity among the 122 yellow endosperm tropical maize inbred lines. A total of 190 alleles were identified among the 122 inbred lines with an average of 3.72 alleles per locus for the 51 polymorphic SSR out of a total of 62 SSRs that were evaluated. The number of alleles detected among the 122 inbred lines using SSR markers and PIC values for polymorphic loci are presented in Table 4.3. The range in allele number was 2 to 6 with three markers showing the highest number of alleles (phi064, phi328175, phi96100). PIC value on average was 0.43 for the polymorphic SSR markers varied from 0.12 (phi032) to 0.74 (umc1399). Nearly 63 % of these markers had more than 0.4 PIC values.

Table 4.3: Bin numbers, allele numbers and PIC values for SSR loci used to genotype 122 yellow endosperm tropical-adapted maize inbred lines

SSR locus	Bin no.	No. of alleles	PIC value	SSR locus	Bin no.	No. of alleles	PIC value
nc133	2.05	3	0.46	**phi374118**	3.02	4	0.25
phi011	1.09	3	0.54	**phi420701**	8	3	0.28
phi015	8.09	4	0.41	**phi423796**	6.01	4	0.14
phi029	3.04	3	0.53	**phi452693**	6.04	4	0.69
phi032	9.04	3	0.12	**phi453121**	3.01	3	0.37
phi034	7.02	4	0.55	**phi96100**	2	6	0.43
phi041	10	5	0.32	**umc1061**	10.06	2	0.5
phi046	3.08	2	0.37	**umc1109**	4.1	3	0.49
phi053	3.05	3	0.57	**umc1136**	3.09	4	0.55
phi056	1.01	4	0.45	**umc1152**	10.01	4	0.54
phi059	10.02	3	0.27	**umc1153**	5.09	2	0.37
phi064	1.11	6	0.72	**umc1161**	8.06	3	0.52
phi070	6.07	4	0.35	**umc1196**	10.07	4	0.39
phi072	4.01	5	0.32	**umc1277**	9.08	3	0.51
phi076	4.11	4	0.53	**umc1279**	9	3	0.23
phi084	10.04	2	0.19	**umc1399**	3.07	4	0.74
phi087	5.06	4	0.34	**phi065**	9.03	3	0.46
phi093	4.08	4	0.52	**phi077**	6.01	5	0.56
phi109188	5.03	3	0.48	**bnlg 1445**	-	5	0.42
phi109275	1	3	0.14	**bnlg 143**	5.01	2	0.23
phi109642	2	4	0.5	**bnlg2037**	8.01	5	0.2
phi112	7.01	3	0.4	**blng2162**	4.08	5	0.59
phi123	6.07	4	0.35	**dupssr24**	2.08	5	0.69
phi227562	1.11	4	0.52	**umc1008**	4.01	4	0.49
phi233376	8.09	3	0.49	**blng 1016**	1.04	4	0.53
phi328175	7.04	6	0.44				
Total number of alleles						**190**	
Mean						**3.72**	**0.43**

4.1.2 Diversity and cluster analysis AFLP-based assessment

The AFLP-based genetic distance (GD) between pairs of yellow endosperm maize inbred lines varied from 0.08 to 0.64 with an average of 0.48. The dendrograms obtained from cluster analysis of AFLP-based genetic distance estimates separated the 38 yellow endosperm maize inbred lines into two major groups consistent with pedigrees (Figure 4.1). Sixteen of the 19 inbred lines included in group I was derived from crosses having one or two common parents. The remaining 19 inbred lines included in group II (Table 3.1) were derived from crosses containing different exotic germplasm (CM 116, KI 28, B73LPA, GT-MAS-Gk, KVI43, KVI11 and KI21) as well as from broad-based populations (SYN-Y-STR, ACR97TZL-COMP1-Y, Taraba, Z.DiploBC4 and TZE-COMP5-Y). Cluster analysis showed a good fit with distance estimates as reflected in cophenetic coefficient (r=0.97). The genetic relationships among the 38 yellow endosperm maize inbred lines were represented using principal coordinate analysis of AFLP-based GD estimates. The PCoA analysis showed a clearer differentiation of group I from group II as expected, group II contains diverse lines divided into sub-groups (Figure 4.2).

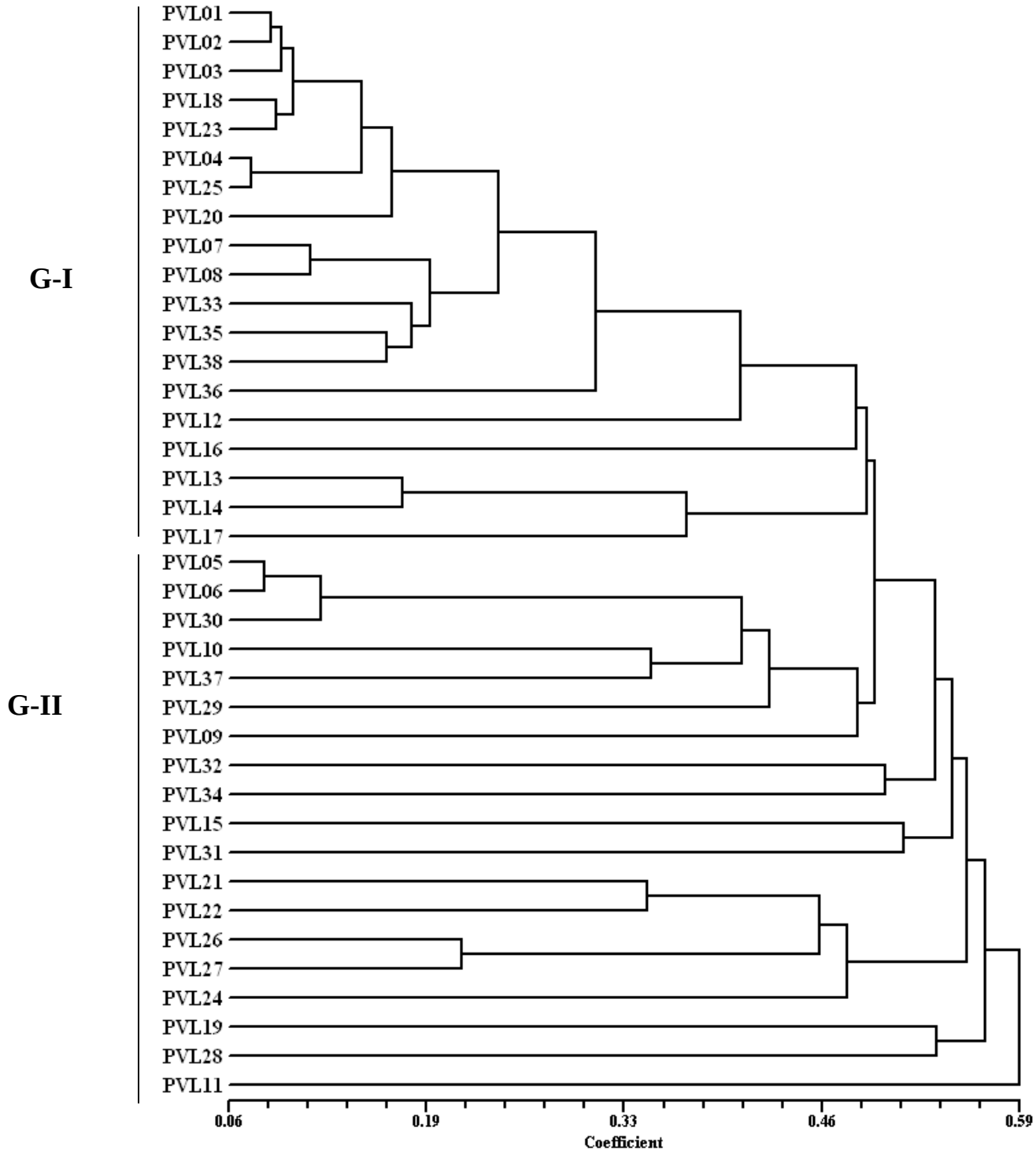


Figure 4.1 Dendrogram of 38 tropical-adapted yellow maize inbred lines obtained using AFLP markers

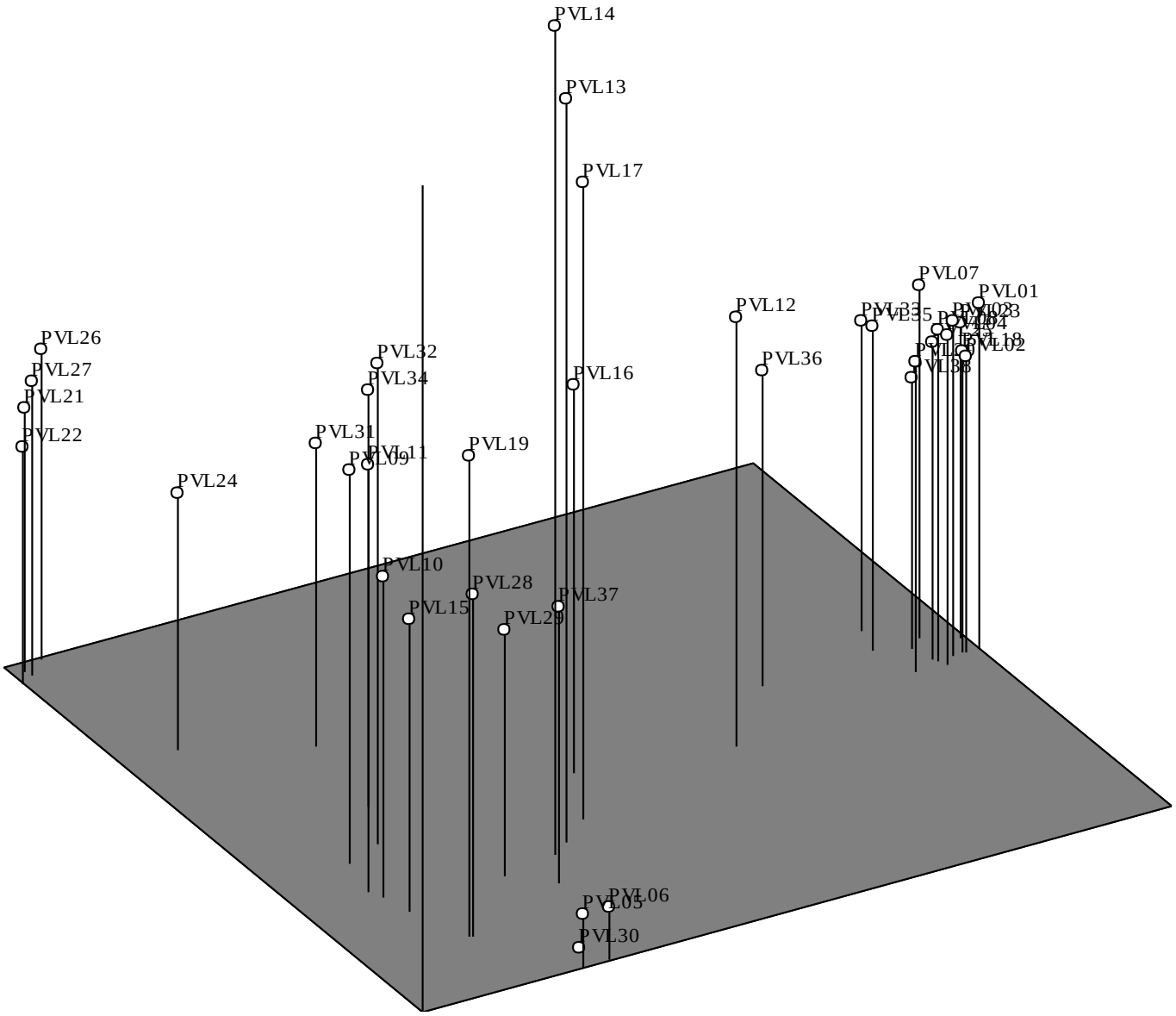


Figure 4.2: 3D plot of tropical-adapted yellow maize determined on the basis of principal coordinate analysis of AFLP-based estimates of Jaccard's (1908) distance.

4.1.3 Diversity and cluster analysis SSR-based assessment

Estimates of Modified Roger's distance (MRD) between all pairs of 38 lines varied from 0.007 to 0.59, with an average of 0.45. Cluster analysis based on genetic distance from SSR data showed two well defined main groups of the 38 inbred lines, consistent with their pedigrees (Figure 4.3). Among the 23 inbred lines included in group I, 19 had one or two common parent in their pedigrees (Figure 4.3). The remaining 15 inbred lines clustered in group II with SSR markers were the same set of lines included in group II with AFLP markers, except PVL 23 and PVL 38 (Figure 4.3). The correlation coefficient (r) between the original genetic distance matrix and the cophenetic matrix generated from the dendrogram was 0.89, showing a moderately good fit. The principal coordinate analysis also grouped the lines into two distinct groups with some defined within each main group (Figure 4.4). Gel photographs showing amplification of SSR alleles for 38 maize inbred lines using two polymorphic SSR primers are shown in figures 4.5 and 4.6.

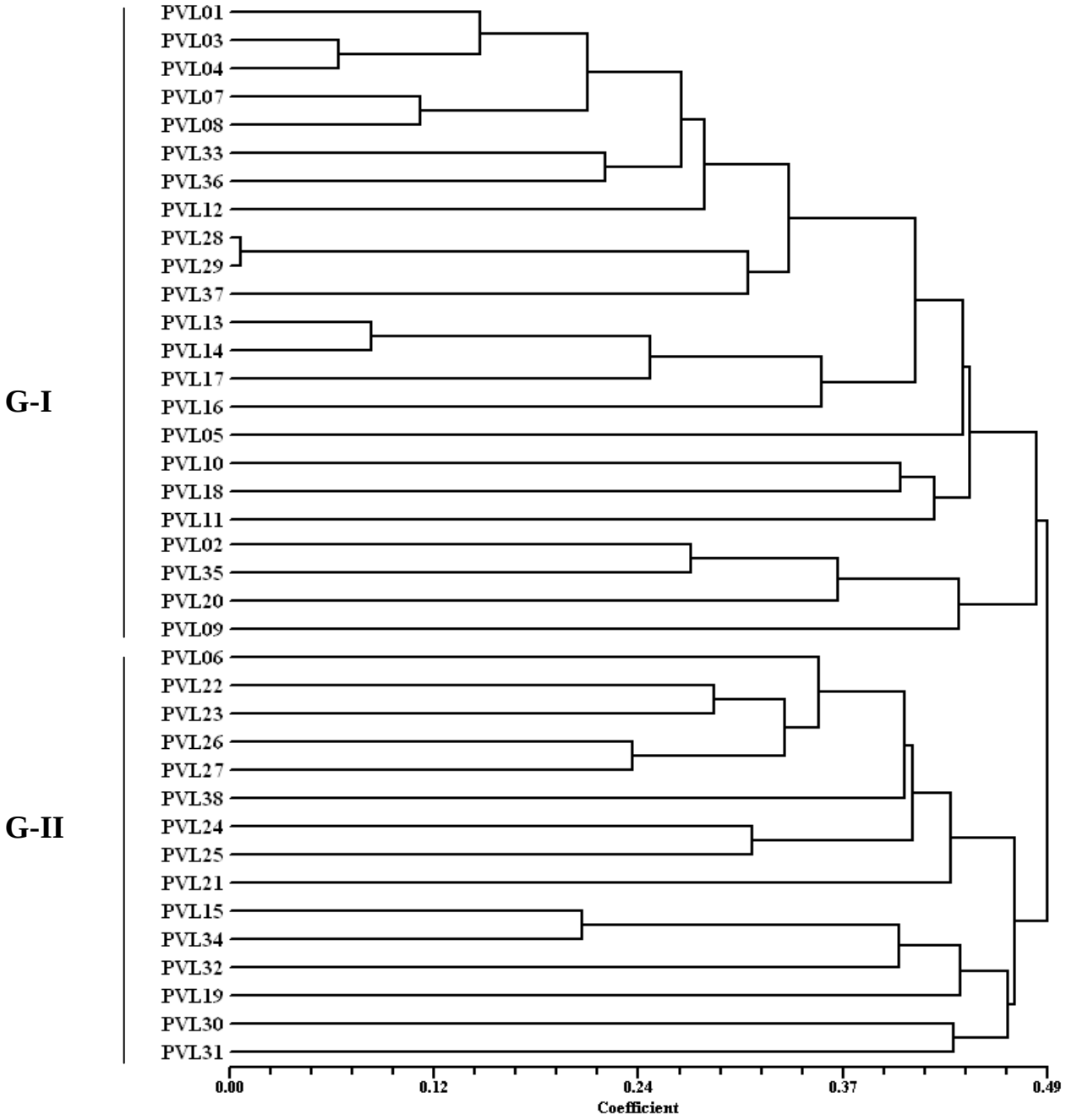


Figure 4.3: Dendrogram of 38 tropical-adapted yellow maize inbred lines obtained using SSR markers

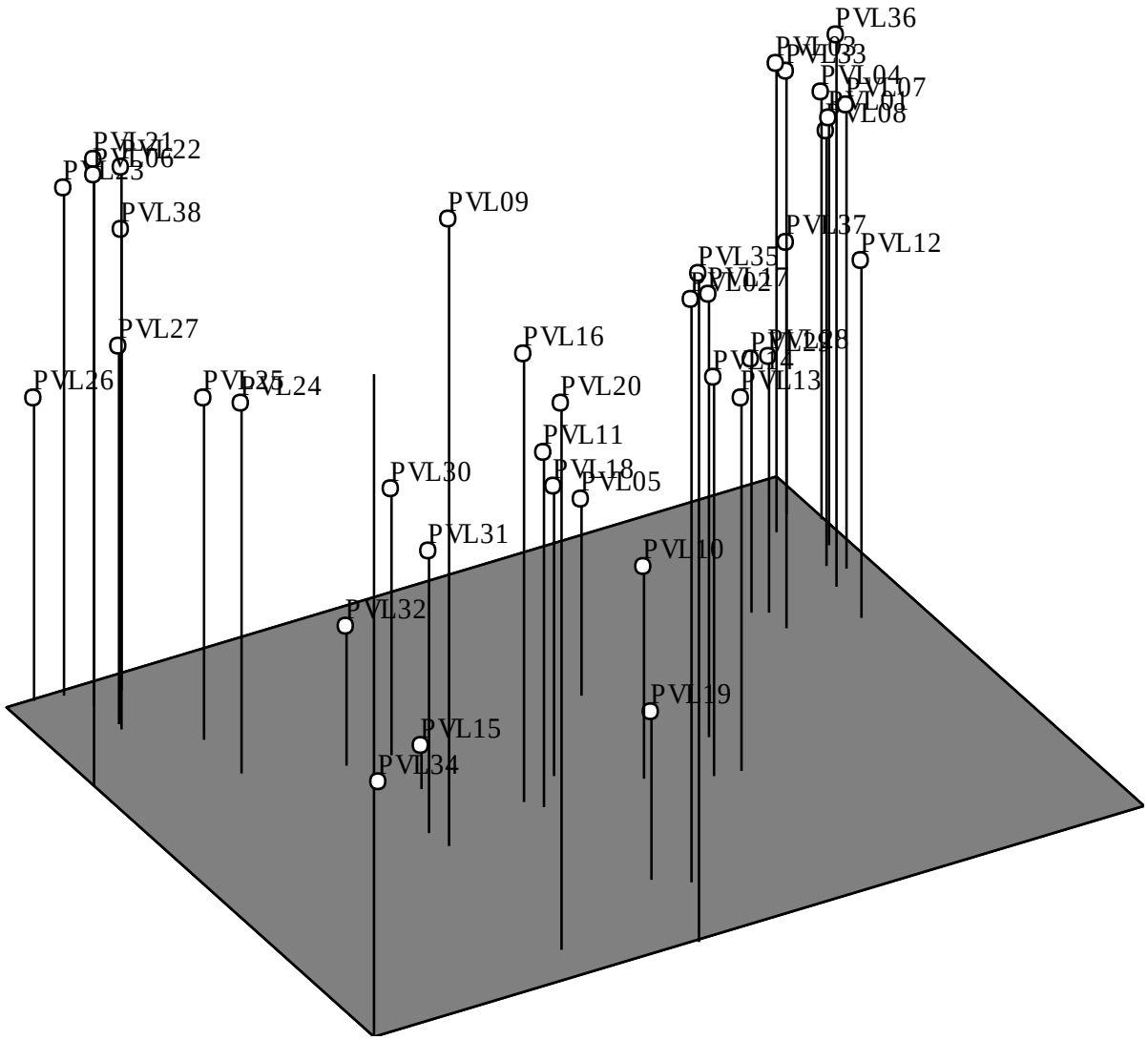


Figure 4.4: 3D plot of 38 tropical-adapted yellow maize determined on the basis of principal coordinate analysis of SSR-based estimates of Modified Roger's (1972) distance.

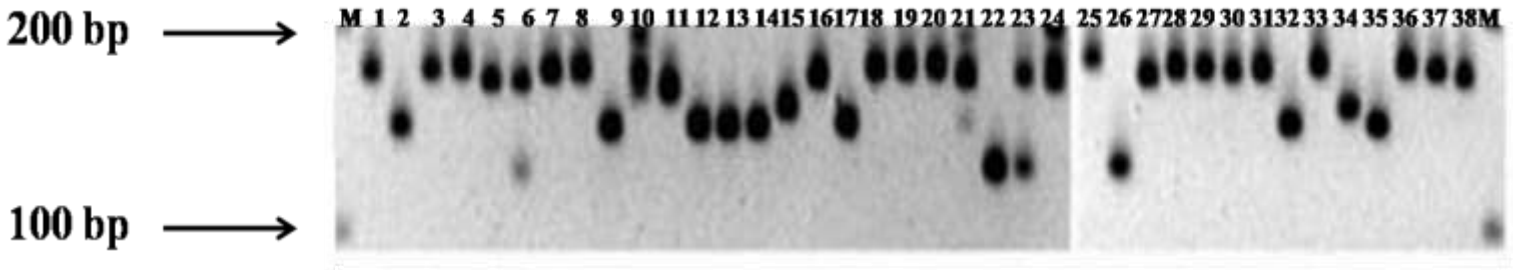


Figure 4.5: Gel photograph of amplification of SSR alleles for SSR primer phi 126 amplified in 38 yellow endosperm tropical maize inbred lines.

M – Molecular mass marker.

Figure 4.6: Gel photograph of amplification of SSR alleles for SSR primer bnlg 1721 amplified in 38 yellow endosperm tropical maize inbred lines.

M – Molecular mass marker.

Estimates of MRD between all pairs of 122 inbred lines varied from 0.02 to 0.61, with an average of 0.41. Four groups denoted by G-I, G-II, G-III and G-IV were formed using UPGMA cluster analysis of SSR-based GD estimates (Figure 4.7). The first group included lines derived from 9450 or pop66SR. The second group involved inbred lines containing KU1414 or 4001. The third group which is the largest main group consists of five subgroups and it contained lines derived from backcrosses involving two typical parents (KU1409 and KU1414) and with eight exotic non-recurrent parents. When each of the subgroup was examined, the maize inbreds with similar exotic parent tended to group together within G-III. The fourth group contained lines, a broad-based population (SYN-Y-STR) belongs to this group and intermixed with lines coming from pop66SR. The cophenetic correlation coefficient between the dendrogram constructed using UPGMA and the calculated GD (r = 0.78) indicated that the clusters accurately represented the GD estimates. In the principal coordinate analysis of SSR-based GD estimates, it also separated the lines into four main groups consistent with clustering using UPGMA with clear demarcation of groups (Figure 4.8). Gel photographs showing amplification of SSR alleles for some of the 122 maize inbred lines using two polymorphic SSR primers are shown in figures 4.9 and 4.10.

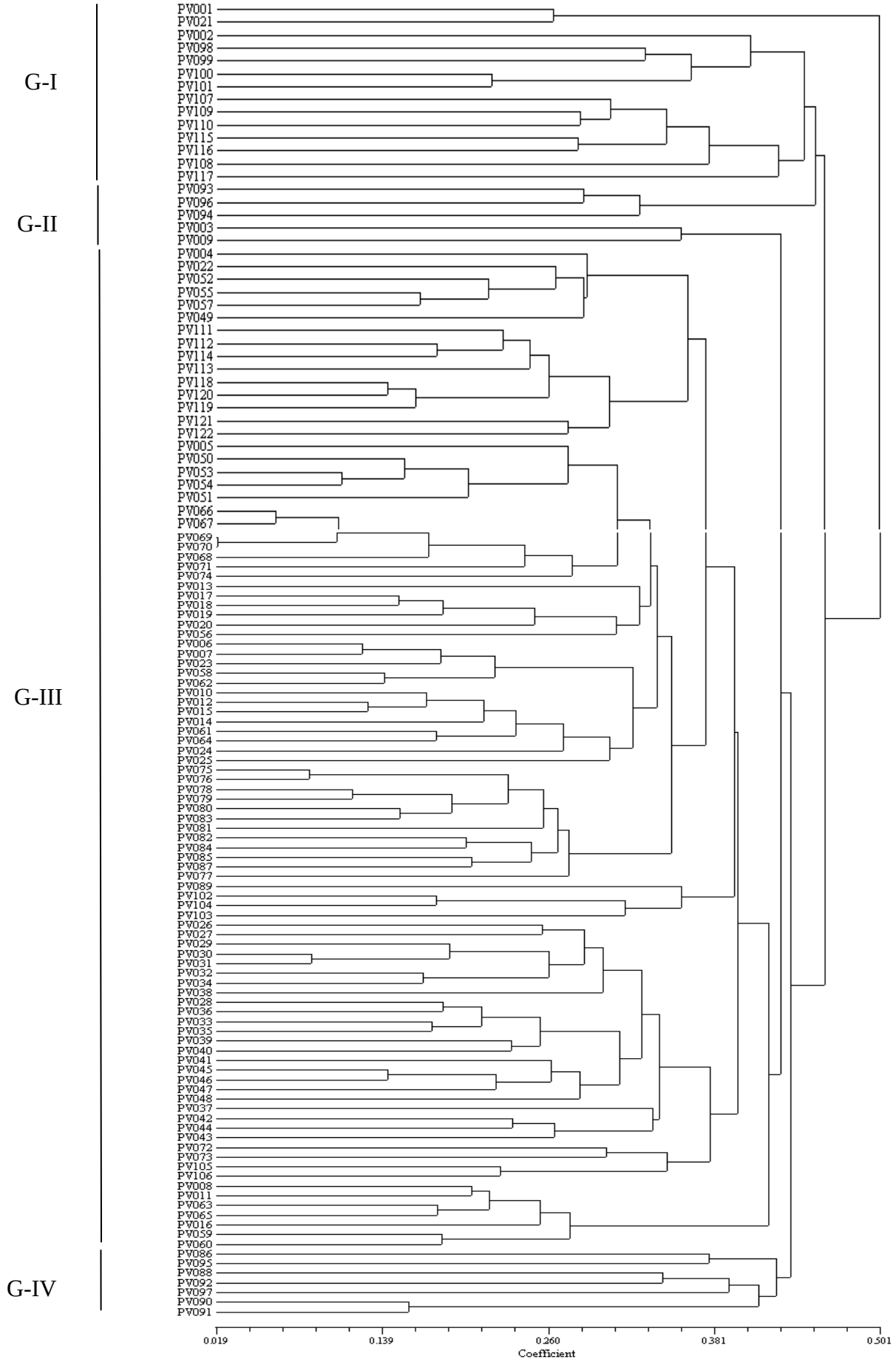


Figure 4.7: Dendrogram of 122 tropical-adapted yellow maize inbred lines obtained using SSR markers

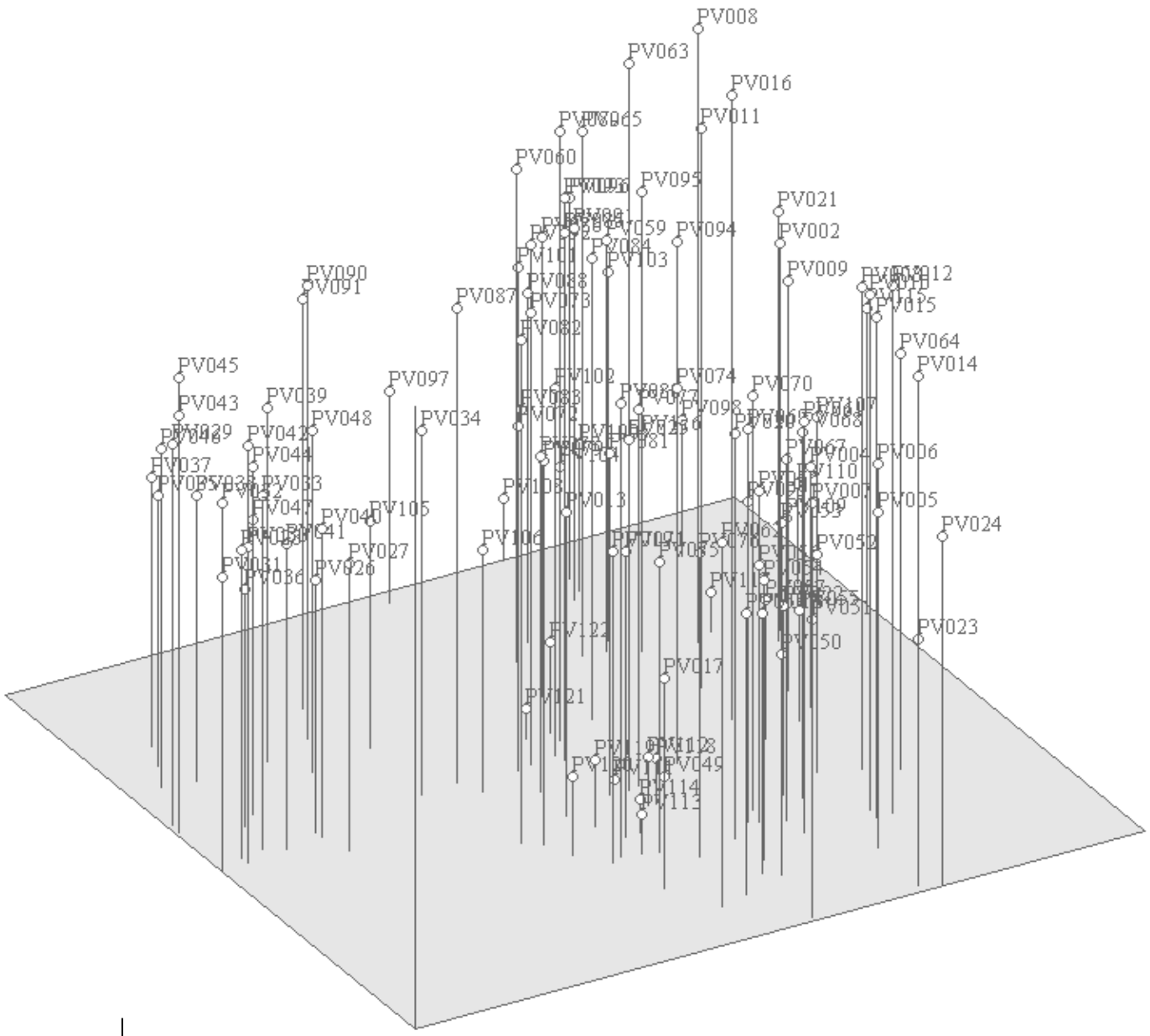


Figure 4.8: 3D plot of 122 tropical-adapted yellow maize determined on the basis of principal coordinate analysis of SSR-based estimates of Modified Roger's (1972) distance.

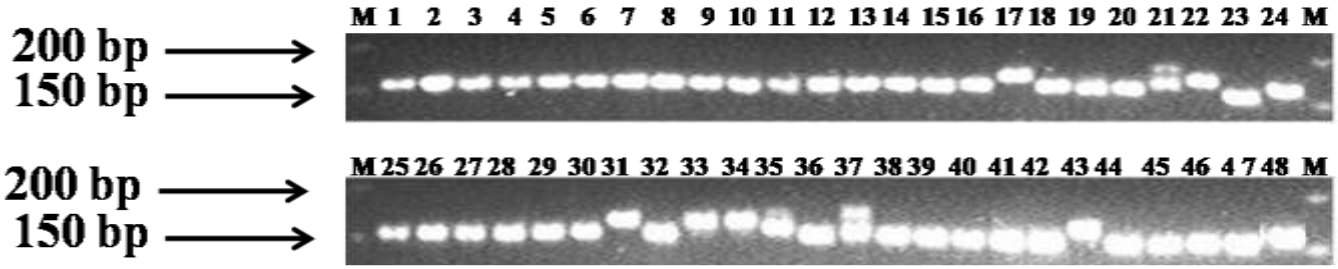


Figure 4.9: Gel photograph of amplification of SSR alleles for SSR primer phi96100 amplified in 48 of 122 yellow endosperm tropical maize inbred lines.

M – Molecular mass marker.

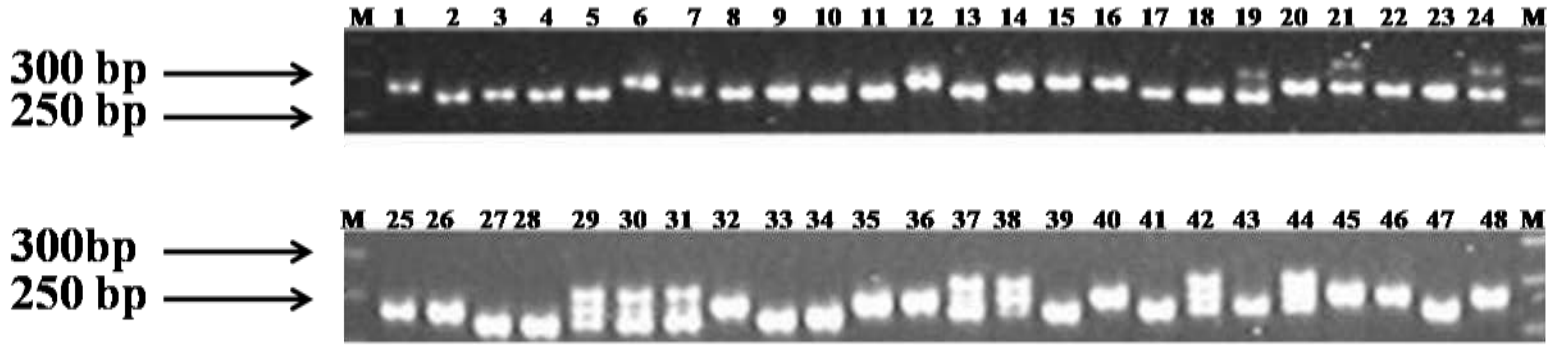


Figure 4.10: Gel photograph of amplification of SSR alleles for SSR primer phi076 amplified in 48 of 122 yellow endosperm tropical maize inbred lines.

M – Molecular mass marker.

4.1.4 DNA sequence in *PSY*1 gene

Nucleotide variation in *PSY1* among the 38 inbred lines after amplifying and sequencing the PCR product with primer pairs Y1-4-F and Y1-4-R did not show any polymorphism in the inbred lines. Also, pair of primers designed to amplify and sequence the upstream of the regulatory region of the *PSY1* gene by using primer pairs Y5P-F and Y5P-R revealed a total of 18 heterozgotes in the genomic sequences analysed.

4.1.5 Comparison between AFLP-based grouping and SSR-based grouping

The correlation between the genetic distance matrices obtained with AFLP and SSR molecular markers was very low ($r = 0.26$, $P < 0.001$). However, the dendrogram constructed using both AFLP and SSR markers showed similar grouping of the inbred lines. Seventeen of the nineteen inbred lines that clustered together in group I based on AFLP-based GD estimates were included in group I based on SSR-based GD estimates. Similarly, 13 of the 15 inbreds found in group II based SSR data were also included in group II based AFLP data. Only 8 inbred lines clustered differently based on AFLP and SSR data. Lines PVL11 and PVL28 have the same pedigree but were included in group II based on AFLP data but in group I based on SSR data. PVL23 and PVL34 shared a common genetic background with PVL06 and were grouped together based on the SSR data, but not with the AFLP data. The remaining four inbred lines included in different groups had diverse genetic backgrounds. Another notable difference between the genetic relationships revealed by SSR and AFLP relates to the classification of the two major parental lines PVL37 and PVL38. These parents grouped very closely with their progenies based on the SSR data but not based on the AFLP data. SSR data grouped 17

lines having PVL37 as a parent and all the line having PVL38 as a parent into two separate groups. However, grouping of PVL 37 in relation to its progeny was less clear with AFLP markers.

4.2 Carotenoid composition-based assessment

The Euclidean distances between pairs of 38 inbred lines varied from 0.30 to 9.00 with an average of 3.27. The highest genetic distance was detected between PVL38 and PVL015, which have diverse genetic backgrounds. The dendrogram generated from the similarity coefficients of carotenoid composition based data separated the 38 lines into two distinct groups (Figure 4.11). The 24 yellow maize inbred lines included in group I had higher concentrations of all carotenoids. Twelve lines having low concentrations of α-carotene, total β-carotene, lutein, zeaxanthin and total pro-vitamin A were included in group II. Two lines were separated from the two groups because they had concentrations of β-carotene and pro-vitamin A having very high content. Thus the resulting dendrogram clearly demonstrated that grouping was based not only on carotenoid concentration but also on carotenoid profile. The cophenetic correlation coefficient was strong (r=0.87), suggesting a positive representation between the dendrogram and genetic distance matrix. The genetic relationships among the 38 yellow endosperm inbred lines was determined using principal coordinate analysis of carotenoid-based GD estimates as shown in Figure 4.12. PCoA separated the inbred lines into two distinct groups which were in good agreement with the results of cluster analysis.

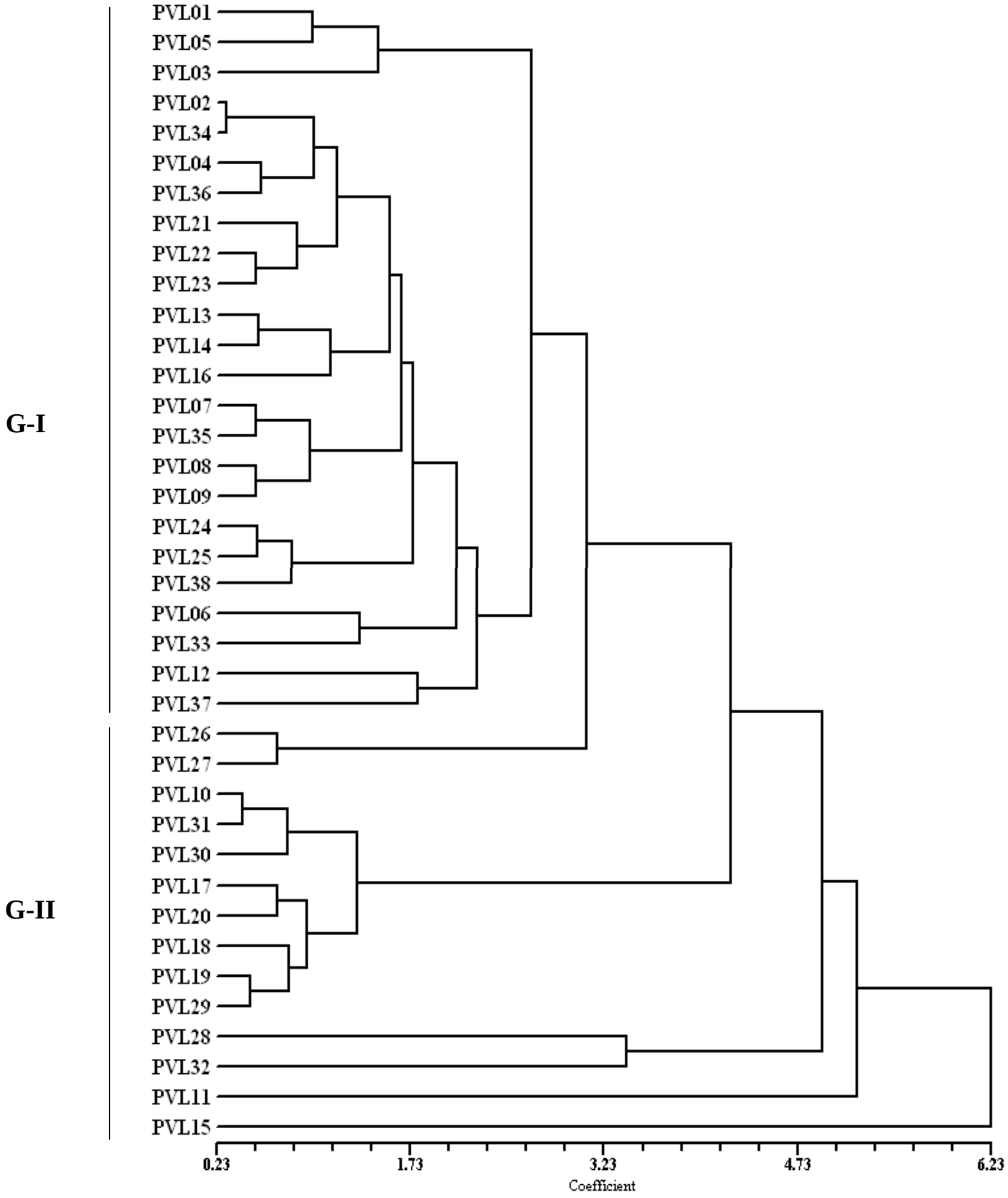


Figure 4.11: Dendrogram of 38 tropical-adapted yellow maize inbred lines obtained using carotenoid data.

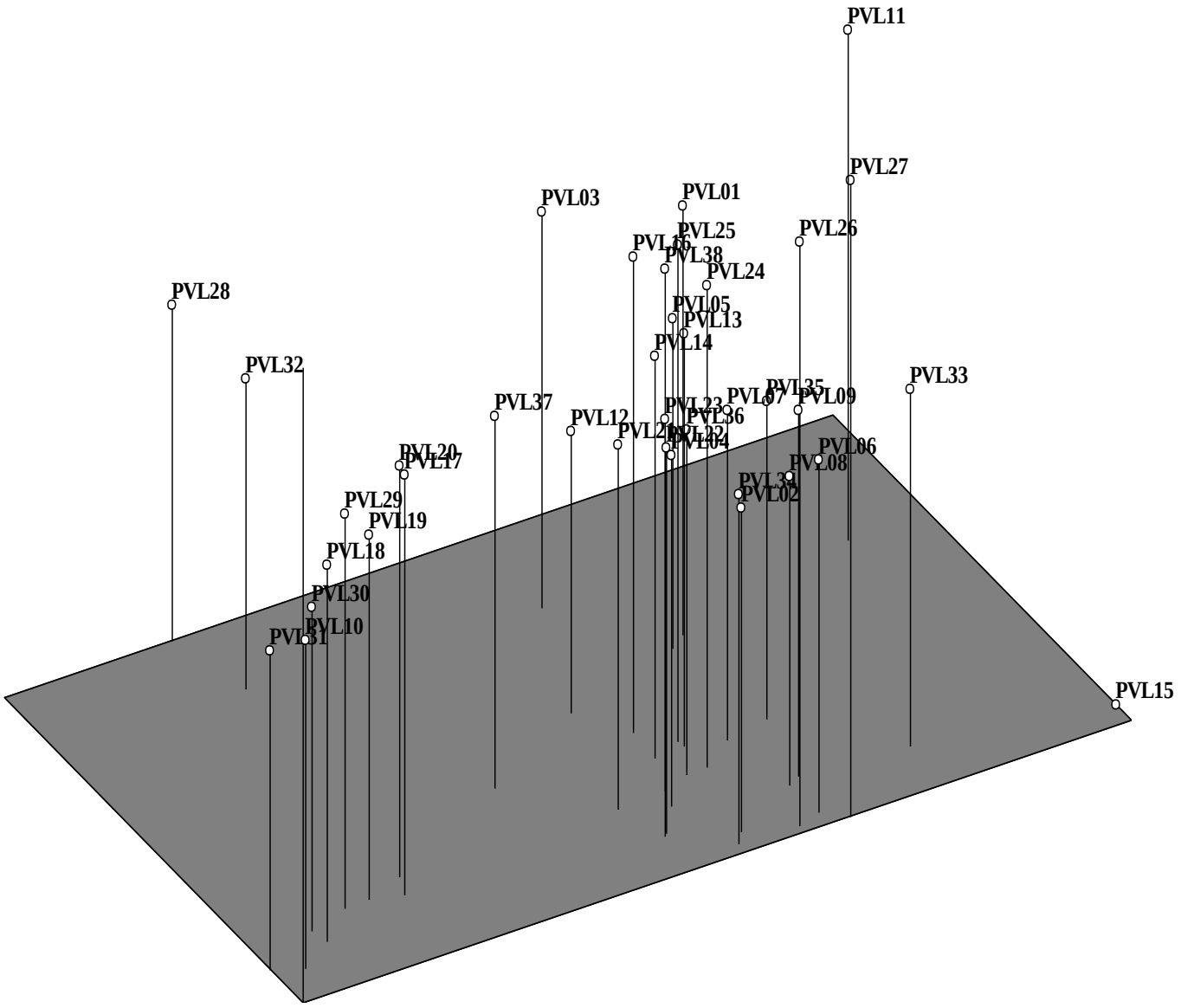


Figure 4.12: 3D plot of 38 tropical-adapted yellow maize determined on the basis of principa coordinate analysis of carotenoid-based estimates of Euclidean's (Sneath and Sokal, 1973) distance

The Euclidean distances between pairs of 122 inbred lines varied from 1.00 to 9.97 with an average of 3.81. The dendrogram generated four main groups (Figure 4.13). The 5 maize inbred lines included in group I had higher concentrations of zeaxanthin. The lines included in group II which had 6 subgroups, grouping into subgroups was based on pro-vitamin A concentration with each subgroup having inbred lines having similar range. Lines in group III were sub-grouped into 2 based on pro-vitamin A concentration. Fourteen lines were found in subgroup III because they had differing concentrations for all carotenoids. Group IV included 11 lines possessing highest concentration of pro-vitamin A. Thus the dendrogram showing genetic relationships amongst 122 maize inbreds based on carotenoid data clearly showed that grouping was based on both the carotenoid concentration and carotenoid profile. The cophenetic correlation coefficient was strong (r=0.87), showing a positive representation between the dendrogram and genetic distance matrix. The genetic relationship among the 122 yellow endosperm inbred lines was determined using principal coordinate analysis of carotenoid-based GD estimates (Figure 4.14). PCoA separated the inbred lines into four distinct groups which were in concordance with the results of cluster analysis (Figure 4.13).

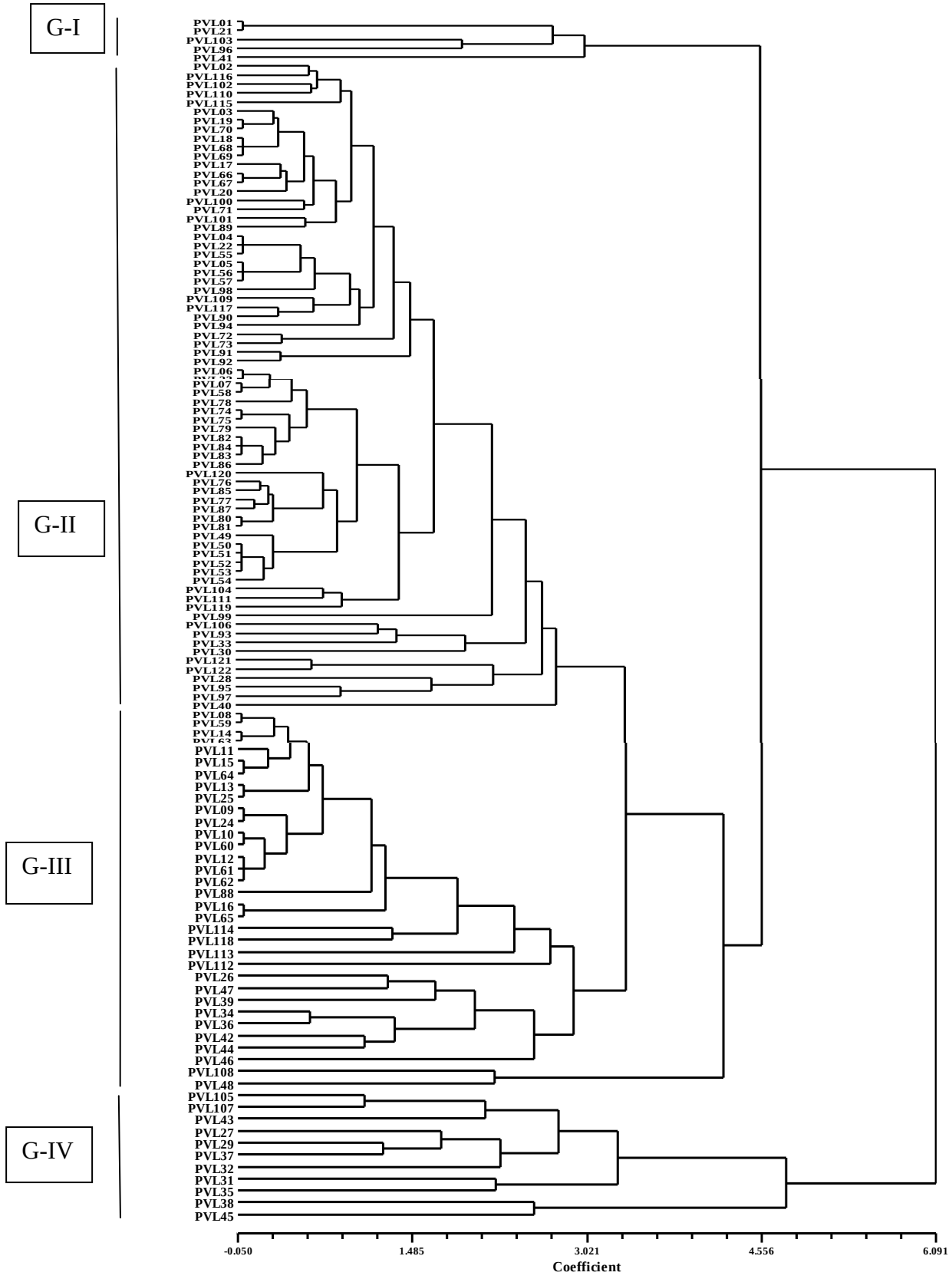


Figure 4.13: Dendrogram of 122 tropical-adapted yellow maize inbred lines obtained using Carotenoid data.

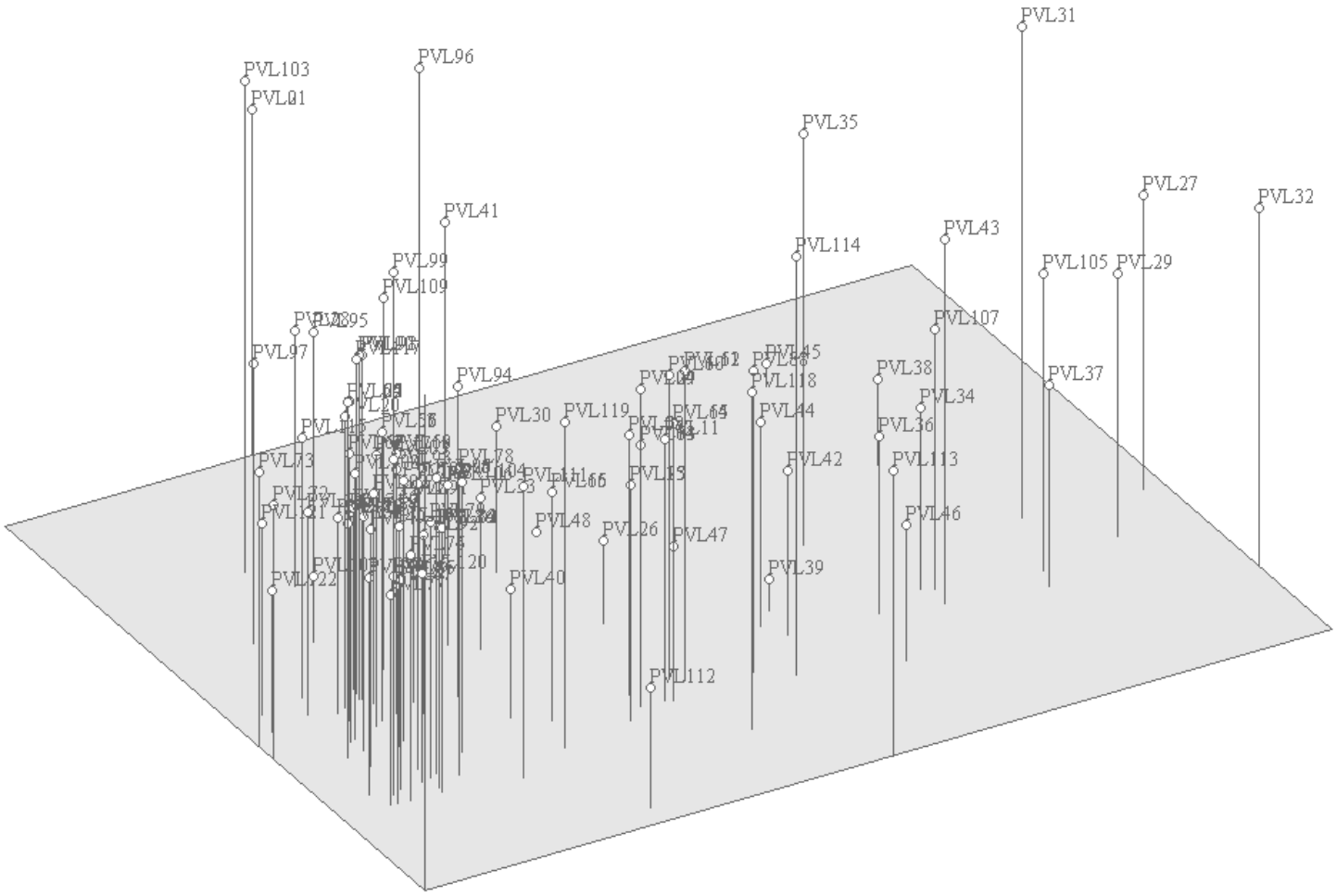


Figure 4.14: 3D plot of 122 tropical-adapted yellow maize determined on the basis of principal coordinate analysis of carotenoid-based estimates of Euclidean's distance (Sneath and Sokal, 1973)

4.2.1 Comparison of AFLP- and SSR-based groupings with carotenoid based grouping of lines

Cophenetic correlation of AFLP-based and SSR-based groupings with carotenoid-based grouping of 38 inbred lines was 0.24 and 0.05, respectively. The number of lines included in group I using AFLP data with high concentrations of β-carotene, α-carotene and pro-vitamin A were 15 lines (Table 4.4) while those included in group I using SSR data with high concentration of β-carotene, α-carotene and pro-vitamin A were 17 (Table 4.5). The number of lines with low concentrations of β-carotene, α-carotene and pro-vitamin A included in group II were 7 for the AFLP data and 4 for SSR data. In carotenoid-based grouping, group I consisted of 24 lines with high concentration of lutein, zeaxanthin, β-cryptoxanthin, β-carotene, α-carotene and pro-vitamin A. The 24 lines descended from two different parentages. Group II was composed of lines with low concentration of lutein, zeaxanthin, β-cryptoxanthin, β-carotene, α-carotene and pro-vitamin A (Table 4.6).

Table 4.4 Minimum, maximum and means carotenoid values for 38 tropical-adapted yellow maize inbred lines separated into two groups based on AFLP markers

	Group I			Group II		
Carotenoids	Min (µg/g)	Max (µg/g)	Mean (µg/g)	Min (µg/g)	Max (µg/g)	Mean (µg/g)
Lutein	2.32	10.9	6.59	1.19	18.1	5.87
Zeaxanthin	6.03	20.87	12.3	0.28	32.4	11.33
β-cryptoxanthin	0.38	6.28	3.75	0.07	5.98	2.93
α-carotene	0.13	0.96	0.36	0.08	1.38	0.34
Trans β-carotene	0.25	3.10	1.78	0.08	6.30	1.59
Total β-carotene	0.61	4.58	2.94	0.37	8.28	2.66
Pro-vitamin A	0.86	7.87	5.00	0.43	9.79	4.29

Table 4.5 Minimum, maximum and means carotenoid values for 38 tropical-adapted yellow maize inbred lines separated into two groups based on SSR markers

	Group I			Group II		
Carotenoids	Min (µg/g)	Max (µg/g)	Mean (µg/g)	Min (µg/g)	Max (µg/g)	Mean (µg/g)
Lutein	1.2	18.1	6.94	1.83	13.7	5.15
Zeaxanthin	0.90	24.5	11.15	0.28	32.95	12.85
β-cryptoxanthin	0.38	6.28	3.38	0.07	5.98	3.28
α-carotene	0.96	1.38	0.42	0.08	0.65	0.24
Trans β-carotene	0.13	3.10	1.70	0.08	6.30	1.66
Total β-carotene	0.54	4.75	2.80	0.37	8.27	2.81
Pro-vitamin A	0.86	7.87	4.70	0.43	9.79	2.58

Table 4.6 Minimum, maximum and means carotenoid values for 38 tropical-adapted yellow maize inbred lines separated into two groups based on Carotenoid Composition

	Group I			Group II		
Carotenoids	Min (µg/g)	Max (µg/g)	Mean (µg/g)	Min (µg/g)	Max (µg/g)	Mean (µg/g)
Lutein	3.37	6.64	6.63	1.20	18.17	5.54
Zeaxanthin	6.41	20.8	12.6	0.29	32.40	10.40
β-cryptoxanthin	2.67	6.28	4.27	0.07	5.50	1.75
α-carotene	0.14	0.96	0.37	0.09	1.38	0.32
Trans β-carotene	1.03	3.10	2.19	0.08	6.30	1.12
Total β-carotene	1.93	4.57	3.70	0.37	8.27	1.98
Pro-vitamin A	3.84	7.87	5.60	0.43	9.79	2.58

4.2.2 Comparison of SSR-based groupings with carotenoid based grouping of lines

Cophenetic correlation of SSR-based groupings with carotenoid-based groupings for 122 inbred lines was -0.06. The number of lines included in groups I and II using SSR data with high concentrations of β-carotene, α-carotene and pro-vitamin A were 17 lines (Table 4.7). The number of lines with low concentrations of β-carotene, α-carotene and pro-vitamin A included in group II were 14 for SSR data. Group IV had 7 lines with high level of pro-vitamin A, trans β-carotene and α-carotene. In carotenoid-based grouping group I consisted of lines with high concentration of lutein and zeaxanthin. Group III was composed of lines with both high and low concentration of lutein, zeaxanthin, β-cryptoxanthin, β-carotene, α-carotene and pro-vitamin A (Table 4.8) while group IV had lines with very high concentrations of pro-vitamin A and trans β-carotene

Table 4.7: Minimum, maximum and means carotenoid values for 122 tropical-adapted yellow endosperm maize inbred lines separated into four groups based on SSR markers.

CAROTENOIDS	GROUP I				GROUP II		GROUP III			GROUP IV		
	MIN (µg/g)	MAX (µg/g)	MEAN (µg/g)	MIN (µg/g)	MAX (µg/g)	MEAN (µg/g)	MIN (µg/g)	MAX (µg/g)	MEAN (µg/g)	MIN (µg/g)	MAX (µg/g)	MEAN (µg/g)
Lutein	3.99	18.17	8.62	4.41	10.52	7.09	0.44	26.29	5.29	2.28	8.67	5.55
Zeaxanthin	3.47	24.50	13.88	8.49	18.67	12.51	0.71	26.19	6.25	4.78	21.77	13.72
β-Cryptoxanthin	0.58	5.10	4.00	2.39	5.14	3.79	0.31	5.85	3.08	1.57	4.59	3.19
α-Carotene	0.12	0.84	0.37	0.21	0.28	0.24	0.13	1.85	0.45	0.15	0.41	0.24
Trans β-carotene	1.40	5.78	2.31	2.27	4.14	2.93	1.56	7.32	3.17	2.08	4.43	2.56
Total β-carotene	3.04	7.76	4.13	3.71	5.70	4.56	5.56	8.68	6.69	3.32	6.33	4.16
Pro-vitamin	5.10	8.33	6.31	5.56	8.69	6.61	2.06	9.09	4.56	5.07	7.32	5.87

Table 4.8: Minimum, maximum and means carotenoid values for 122 tropical-adapted yellow endosperm maize inbred lines separated into four groups based on Carotenoid data.

CAROTENOIDS	GROUP I				GROUP II		GROUP III			GROUP IV		
	MIN (µg/g)	MAX (µg/g)	MEAN (µg/g)	MIN (µg/g)	MAX (µg/g)	MEAN (µg/g)	MIN (µg/g)	MAX (µg/g)	MEAN (µg/g)	MIN (µg/g)	MAX (µg/g)	MEAN (µg/g)
Lutein	8.27	18.17	12.84	2.04	11.9	4.93	0.44	17.75	6.08	1.88	26.3	9.03
Zeaxanthin	18.65	26.2	23.22	4.1	26.03	10.62	0.37	10.24	6.03	0.71	17.60	4.96
β-Cryptoxanthin	3.50	5.85	5.22	2.00	5.70	4.03	0.32	3.28	1.91	0.31	19.45	0.89
α-Carotene	0.25	0.97	0.59	0.12	1.11	0.24	0.16	1.67	0.49	0.57	1.85	1.32
Trans β-carotene	2.59	3.40	2.91	1.46	3.24	2.15	1.40	5.11	3.99	5.00	7.32	6.05
Total β-carotene	4.58	5.70	4.94	2.61	4.84	3.50	3.04	7.07	5.47	6.89	9.09	7.93
Pro-vitamin	7.07	8.69	7.78	3.99	7.24	5.63	5.02	8.65	6.67	8	10.27	9.05

4.3 Allelic variation of carotenoid biosynthetic genes in yellow maize inbred lines

4.3.1 PCR amplifications for SNP polymorphisms

PCR-based DNA markers targeting naturally occurring allelic variations in different regions of *LCYE* and *crtRB1/HYD3* genes were amplified and used to analyse the two sets,S1 (n=38) and S2 (n=122) of yellow endosperm tropical-adapted maize inbred lines.

4.3.2 Genes that significantly affect flux in the first branch (lycopene – α-carotene – lutein) of the carotenoid biosynthesis

Gene-specific PCR markers based on the *LCYE* gene were assayed to detect alleles for three *LCYE* polymorphisms in panels S1 and S2 which consist of different sets of inbred lines differing in *LCYE* allele classes. The 4-primer cocktail as advocated by Harjes *et al.* (2008) for *LCYE* 5'TE polymorphism gave poor amplifications of all the expected allele sizes. However good amplifications were obtained when used in pairs (one forward and one reverse primer): TE103PF F-1(F1), TE103PR R-1(R1); TE105PR R-1-(R2), ZGT111204-976R(1) F1(F2); TE103PF F-1(F1),TE105PR R-1(R2); TE103PR R-1-(R1), ZGT111204-976R (1) F1(F2) after a careful study of the orientation of these primers. Similarly, the 4-primer cocktail as reported previously for the *LCYE* 3'TE primers by Harjes *et al.* (2008) produced poor amplification of all the expected alleles, however good amplifications were obtained when used in pairs (one forward and one reverse primer): 3pINDL-L1, 3pINDL-R1; 3pINDL-L2, 3pINDL-R2; 3pINDL-L2, 3pINDL-R1**.** PCR reactions gave expected classes of allele in both sets and in some cases genetic variability in some loci were observed (Table 4.9).

Table 4.9: Summary of SNPs in Lycopene Epsilon Cyclase (*LCYE)* in 38 and 122 tropical-adapted yellow endosperm maize inbred lines

POLYMORPHISMS		**CLASSES OF ALLELES**	
Polymorphism site	Types of SNP	Expected size of DNA fragment (bp)	Observed size of DNA fragment (bp)
*LCYE*5'	INDEL/TE	993 (best)	150 + 280 +1700 (best + 2nd best)[ab]
		150 + 280 (2nd best)	Null [ab]
		250 (3rd best)	250+380 (Worst)[b]
		250+380 (worst)	280 + 250 + 1700+ 150[b]
			280[b]
*LCYE*3'	INDEL (DNA Seq)	144+502 (best)	144+502 (best)[ab]
		399+502 (worst)	399+502 (best)[a]
		Null	502[a]
			399+502+144[ab]
LCYE-Exn-216	Base	G (best)	G (best)[ab]
		T (worst)	T(worst)[a]
		Null	Null[b]

[a] Only in 38 inbreds.

[b] Only in 122 inbreds.

[ab] Only in 38 and 122 inbreds.

INDEL: Insertiom/Deletion, TE: Transposable Element and Seq: Sequence.

The functional markers allowed for detection of the favourable/best alleles in the lines analysed by SNP-based PCR assay. At the 5'TE, PCR products gave expected classes of allele in both sets and additional class of alleles were also amplified among tropical maize inbred lines in set 2. The observed classes of alleles in the two sets are shown in Table 4.9. In 3'TE, expected classes of alleles were amplified and two additional classes of alleles were observed; 502 bp allele class PCR products in some inbreds in set 1 and 399+502+104 bp class was found in both sets. The two classes of alleles at the SNP216 locus were amplified with the G allele as the predominant allele. Only one maize inbred had the T allele in set 1 and some with null allele. Favourable alleles for the *LCYE* polymorphisms were detected in both sets with different frequencies. Gel photographs showing genetic variations for the three polymorphisms at *LCYE* loci are shown in Figures 4.15-4.18.

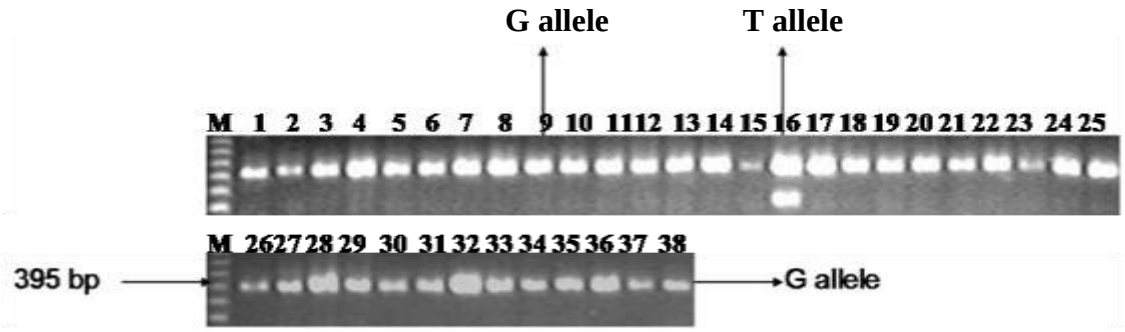


Figure 4.15: *LCYE* PCR assay SNP216 showing the maize inbred lines with best allele (G) and worst allele (T) on 2 % agarose gel. The first lanes are molecular mass marker and the others are DNA samples for the 38 tropical yellow maize inbred lines.

.

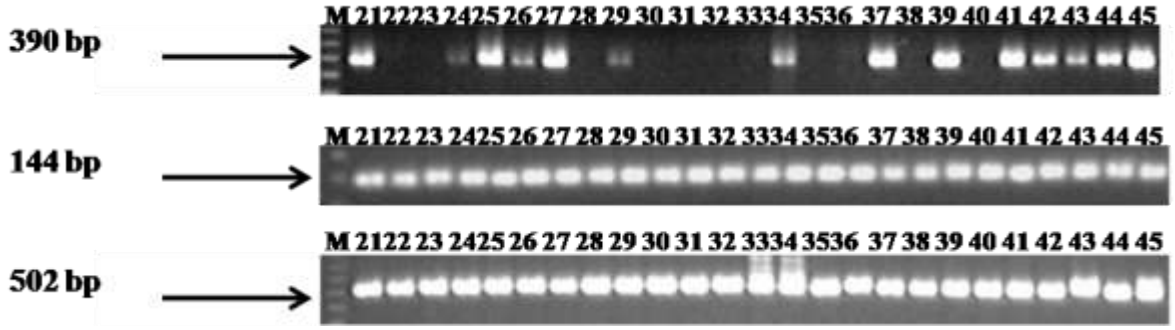


Figure 4.16: *LCYE* PCR assay 3' INDEL gel showing the lines with best allele (144+502 bp) representing 8 bp insertion and worst allele (399+144+502 bp). The first lanes are molecular mass marker and the others are DNA samples for 24 of the 122 tropical yellow maize inbred lines.

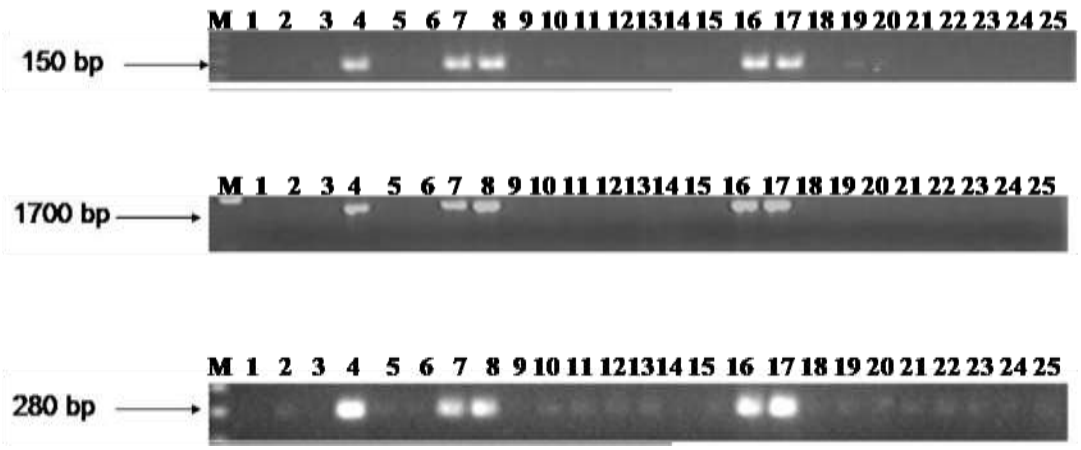


Figure 4.17: *LCYE* PCR assay 5'INDEL showing the lines with best allele (150+280 + 1700 bp) and worst allele (null) on 2 % agarose gel. The first lanes are molecular mass marker and the others are DNA samples for 25 of the 38 tropical yellow maize inbred lines.

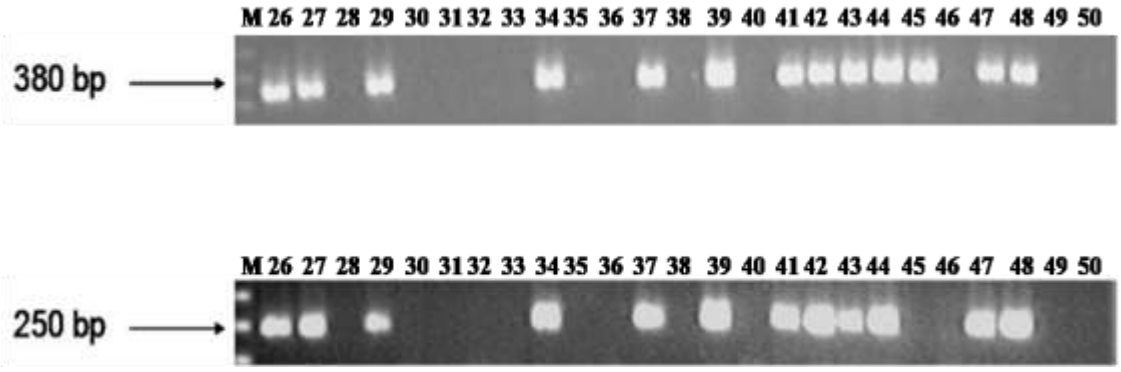


Figure 4.18: LCYE PCR assay 5'INDEL showing the lines with worst allele (250+380 bp) on 2 % agarose gel. The first lanes are molecular mass marker and the others are DNA samples for 25 of the 122 tropical yellow maize inbred lines.

The frequencies of the favourable alleles for the three *LCYE* polymorphisms varied considerably within the two sets. The most favourable (1700 bp class and 150+280 bp class, with promoter transposon insertion) of 5'TE was present at frequencies of 24 % in S1 and 3.3 % in S2 (Table 4.10). The two classes (the first and second best classes) occurred in the same lines as they are functionally similar. These two classes collapsed together for many of the haplotype statistical analysis (Harjes *et al.,* 2008). Frequency of the lines containing favourable allele (144+502 bp haplotype class), with 8 bp deletion in the genomic sequence) of 3'TE is 29 % in the S1 and 67 % in S2. Frequencies were equal for the best allele (G) at the SNP216 polymorphism in the two sets, 97 % of the maize inbred lines in S1 and 98 % in S2.

The 4 maize inbred lines in S2 possessing favourable allele for *LCYE* 5'TE having POP66R or 9450XKI21 in their genetic background however, were found to possess lower α-carotene concentration between 0.12 and 0.35 µg with an average of 0.23 µg than other inbred lines. This shows negative α-carotene/SNP correlation. The remaining lines (118) analyzed in S2 having unfavourable allele for *LCYE* 5'TE have their lutein concentration ranging from 0.44 to 26.3 µg with an average of 5.92 µg, they belonged to different genetic backgrounds. In contrast, for *LCYE* 3'TE polymorphism, 82 lines have favourable allele and the range for α-carotene concentration is 0.13 to1.85 µg with an average of 0.33 µg, majority of the lines belong to KU1409 or KU1414 pedigree.

Table 4.10: Comparison of favourable *LCYE* allele frequency in the 38 and 122 tropical adapted yellow endosperm maize inbred lines in the first branch of carotenoid biosynthesis

Gene polymorphism /haplotype	Population (no of genotypes)	Population (no of genotypes)		Frequency of favourable alleles		
	38	122	Combined	38	122	Combined
***LCYE*5'INDEL/TE**						
A1: 150 + 280 +1700	**9**	**4**	**13**	24	3.3	**8.1**
A2: Null	29	95				
A3: 250+380	0	18				
A4: 280 + 250 + 1700+ 150	0	1				
A5: 280	0	4				
***LCYE*3'INDEL**						
B1: 144+502	**11**	**82**	**93**	29	67	**58**
B2: 399+502	18	0				
B3: 502	7	0				
B4: 399+502+144	2	40				
***LCYE*-Exn-216**						
C1: G	**37**	**120**	**157**	97	98	**98**
C2: T	1	0				
C3: Null	0	2				

The 40 lines with unfavourable allele have lutein concentration as low as 1.11 µg and as high as 26.3 µg with an average of 8.20 µg. This revealed strong association between unfavourable allele and lutein concentration among the inbred lines, though they belonged to varying genetic backgrounds. On the other hand, α-carotene concentration among the 11 lines in S1 with favourable allele for *LCYE* 3'TE was low, it ranged from 0.13 to 0.35 µg with a mean of 0.25 µg and the lines belong to different genetic background. The lines with favourable allele for *LCYE* 5'TE belonged to similar genetic origin except for 2 lines and had α-carotene concentration between 0.095 and 0.53 µg with an average of 0.31 µg. Lutein content for lines having unfavourable allele for *LCYE* 3'TE and *LCYE* 5'TE polymorphism has a range of 1.19 to 18.2 µg and 1.20 to 18.2 µg respectively. Also, these lines belong to diverse genetic background which revealed that the absence of a functional *LCYE* allele may lead to accumulation of lutein.

The 11 lines in S1 possessing favourable alleles for *LCYE* 3'TE polymorphism have either 9450 or KU1414-SR as a parent with the total β-carotene (TBC) ranging from 2.56 to 3.77 µg except for line (9450 x KI 28)-1-2-1-2-B-B-B that has very low β-carotene concentration (0.60 µg). In S2, 67 % of the inbred lines had favourable alleles for *LCYE* 3'TE polymorphism, with one parent line belonging to one exotic germplasm background or the other. The maize inbred lines in sets 1 and 2 with favourable alleles for 3'TE and 5'TE polymorphisms are listed in Tables 4.11 and 4.12 respectively. The total β-carotene of these inbred lines varies between 2.65 and 8.99 µg. For *LCYE* 5'TE polymorphism, 9 lines with favourable allele have TBC variying between 0.76 and 4.5 µg while in S2 the TBC ranged from 3.47-3.96 µg. None of the inbred line had optimal alleles for both

polymorphisms in the two sets. *LCYE* SNP216 (G) polymorphism was found to be very common in the two panels irrespective of their β-carotene content.

Table 4.11: Genotypes with favourable allele for *LCYE* polymorphisms in 38 tropical adapted yellow endosperm maize inbred lines

5' *LCYE* TE IN/DEL polymorphism		3' *LCYE* TE IN/DEL polymorphism	
Inbred	**Pedigree**	**Inbred**	**Pedigree**
PVL04	9450xKI 21-7-3-1-2-5-B-B-B	**PVL06**	(9450 x KI 28)-1-2-1-2-B-B-B
PVL07	9450xKI 21-7-2-1-1-B-B-B	**PVL13**	9450xKI 21-1-5-3-2-2-B-B-B-B
PVL08	9450xKI 21-7-2-1-2-B-B-B	**PVL14**	9450xKI 21-1-5-3-2-1-B-B-B-B
PVL16	9450xKI 21-1-4-1-1-2-B-B-B-B	**PVL18**	ACR97TZL-CCOMP1-Y-S3-12-2-B-B-B-B-B
PVL17	9450xKI 21-5-2-3-1-B-B-B	**PVL21**	KU1414-SR/NC350-4-1-B-B-B
PVL30	Taraba-14-2-2-4-2-B-B-B-B-B	**PVL22**	KU1414-SR/NC350-1-1-B-B-B
PVL33	(9450x KI 21)-8-2-1-1-B-B-B	**PVL23**	(9450 x KI 28)-1-2-1-1-B-B-B-B
PVL36	9450xKI 21-7-2-4-2-1-B-B-B	**PVL24**	KU1414-SR/KVI43-6-4-B-B-B
PVL37	9450	**PVL25**	KU1414-SR/KVI43-6-1-B-B-B
		PVL26	KU1414-SR/KVI11-7-2-B-B-B
		PVL27	KU1414-SR/KVI11-7-1-B-B-B

Table 4.12: Genotypes with favourable allele for *LCYE* polymorphisms in 122 tropical adapted yellow endosperm maize inbred lines

	5' *LCYE* TE IN/DEL polymorphism
Inbred	**Pedigree**
PVL02	9450xKI 21-7-2-4-2-1-B-B-B-B
PVL95	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-6X(MP420x4001xMP420)-3-1-3-1-B)S2-1-BB-B-B
PVL97	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-6X(MP420x4001xMP420)-3-1-3-1-B)S2-10-BB-B-B
PVL100	(POP61-SR-11-2-3-3-1-BB/9450xKI21-3-2-2-1-3)S2-1-BB-B-B
	3' *LCYE* TE IN/DEL polymorphism
Inbred	**Pedigree**
PVL03	(KU1409/KU1414-SR/KVI11)-S2-2-B-B
PVL04	(KU1409/KU1414-SR/KVI3)-S2-5-B-B
PVL05	(KU1409/KU1414-SR/KVI3)-S2-8-B-B
PVL06	(KU1409/KU1414-SR/M162W)-S2-2-B-B
PVL07	(KU1409/KU1414-SR/M162W)-S2-4-B-B
PVL08	(KU1409/KU1414-SR/A619)-S2-1-B-B
PVL09	(KU1409/KU1414-SR/A619)-S2-2-B-B
PVL11	(KU1409/KU1414-SR/A619)-S2-4-B-B
PVL12	(KU1409/KU1414-SR/A619)-S2-5-B-B
PVL13	(KU1409/KU1414-SR/A619)-S2-6-B-B
PVL14	(KU1409/KU1414-SR/A619)-S2-7-B-B
PVL15	(KU1409/KU1414-SR/A619)-S2-8-B-B
PVL16	(KU1409/KU1414-SR/A619)-S2-9-B-B
PVL17	(KU1409/KU1414-SR/NC298)-S2-2-B-B
PVL18	(KU1409/KU1414-SR/NC298)-S2-5-B-B
PVL19	(KU1409/KU1414-SR/NC298)-S2-6-B-B
PVL20	(KU1409/KU1414-SR/NC298)-S2-9-B-B
PVL22	(KU1409/KU1414-SR/KVI3)-S2-5-1-B-B
PVL23	(KU1409/KU1414-SR/M162W)-S2-2-1-B-B
PVL28	(KU1409/DE3/KU1409)S2-3-B-B-B
PVL30	(KU1409/DE3/KU1409)S2-5-B-B-B
PVL31	(KU1409/DE3/KU1409)S2-6-B-B-B
PVL32	(KU1409/DE3/KU1409)S2-7-B-B-B
PVL33	(KU1409/DE3/KU1409)S2-9-B-B-B
PVL35	(KU1409/DE3/KU1409)S2-11-B-B-B
PVL36	(KU1409/DE3/KU1409)S2-13-B-B-B
PVL38	(KU1409/DE3/KU1409)S2-15-B-B-B
PVL40	(KU1409/DE3/KU1409)S2-20-B-B-B
PVL46	(KU1409/DE3/KU1409)S2-32-B-B-B
PVL49	(KU1409/KU1414-SR/KVI3)-S2-2-2-BB-B-B
PVL50	(KU1409/KU1414-SR/KVI3)-S2-3-1-BB-B-B
PVL51	(KU1409/KU1414-SR/KVI3)-S2-3-2-BB-B-B
PVL52	(KU1409/KU1414-SR/KVI3)-S2-3-3-BB-B-B
PVL53	(KU1409/KU1414-SR/KVI3)-S2-3-4-BB-B-B

Table 4.12: Genotypes with favourable allele for *LCYE* polymorphisms in 122 tropical adapted yellow endosperm maize inbred lines

	3' *LCYE* TE IN/DEL polymorphism
Inbred	**Pedigree**
PVL55	(KU1409/KU1414-SR/KVI3)-S2-5-1-BB-B-B
PVL56	(KU1409/KU1414-SR/KVI3)-S2-8-1-BB-B-B
PVL57	(KU1409/KU1414-SR/KVI3)-S2-8-2-BB-B-B
PVL58	(KU1409/KU1414-SR/M162W)-S2-4-2-BB-B-B
PVL59	(KU1409/KU1414-SR/A619)-S2-1-1-BB-B-B
PVL62	(KU1409/KU1414-SR/A619)-S2-5-2-BB-B-B
PVL63	(KU1409/KU1414-SR/A619)-S2-7-1-BB-B-B
PVL65	(KU1409/KU1414-SR/A619)-S2-9-1-BB-B-B
PVL66	(KU1409/KU1414-SR/NC298)-S2-4-1-BB-B-B
PVL67	(KU1409/KU1414-SR/NC298)-S2-4-2-BB-B-B
PVL68	(KU1409/KU1414-SR/NC298)-S2-5-1-BB-B-B
PVL69	(KU1409/KU1414-SR/NC298)-S2-5-2-BB-B-B
PVL70	(KU1409/KU1414-SR/NC298)-S2-6-2-BB-B-B
PVL71	(KU1409/KU1414-SR/NC298)-S2-14-1-BB-B-B
PVL72	(KU1409/KU1414-SR/SC55)-S2-12-1-BB-B-B
PVL73	(KU1409/KU1414-SR/SC55)-S2-13-2-BB-B-B
PVL74	(KU1409/KU1414-SR/NC350)-S2-1-1-BB-B-B
PVL75	(KU1409/KU1414-SR/NC350)-S2-1-2-BB-B-B
PVL76	(KU1409/KU1414-SR/NC350)-S2-5-1-BB-B-B
PVL77	(KU1409/KU1414-SR/NC350)-S2-6-1-BB-B-B
PVL78	(KU1409/KU1414-SR/NC350)-S2-9-2-BB-B-B
PVL79	(KU1409/KU1414-SR/NC350)-S2-11-1-BB-B-B
PVL80	(KU1409/KU1414-SR/NC350)-S2-13-1-BB-B-B
PVL81	(KU1409/KU1414-SR/NC350)-S2-13-2-BB-B-B
PVL82	(KU1409/KU1414-SR/NC350)-S2-16-1-BB-B-B
PVL83	(KU1409/KU1414-SR/NC350)-S2-16-2-BB-B-B
PVL84	(KU1409/KU1414-SR/NC350)-S2-16-3-BB-B-B
PVL85	(KU1409/KU1414-SR/NC350)-S2-19-1-BB-B-B
PVL86	(KU1409/KU1414-SR/NC350)-S2-20-1-BB-B-B
PVL87	(KU1409/KU1414-SR/NC350)-S2-21-1-BB-B-B
PVL88	(SYN-Y-STR-34-1-1-1-1-2-1-B*5/NC354/SYN-Y-STR-34-1-1-1-1-2-1-B*5)-S2-7-5-BB-B-B
PVL89	(KU1414-SR/CML328/KU1414-SR)-S2-5-2-BB-B-B
PVL90	(KU1409/DE3/KU1409)S2-28-1-BB-B-B

Table 4.12: Genotypes with favourable allele for *LCYE* polymorphisms in 122 tropical adapted yellow endosperm maize inbred lines

	3' *LCYE* TE IN/DEL polymorphism
Inbred	**Pedigree**
PVL92	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-1XSYN-Y-STR-34-1-1-1-1-2-1-BBB)S2-4-BB-B-B
PVL93	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-4X4001xKI21-4-1-1-1-1)S2-2-BB-B-B
PVL96	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-6X(MP420x4001xMP420)-3-1-3-1-B)S2-5-BB-B-B
PVL98	(POP66SR/ACR91SUWAN1-SRC1/ACR91SUWAN1-SRC1-8XPOP61-SR-11-2-3-3-1-B)S2-3-BB-B-B
PVL103	(KU1409/SC55/KU1409)-S2-19-1-B-B
PVL111	KU1409/NC358/KU1409-2-BB-B-B
PVL113	KU1409/NC358/KU1409-17-BB-B-B
PVL114	KU1409/SC55/KU1409-4-B-B
PVL115	DE3/KU1414-SR/KU1414-SR-2-B-B
PVL117	KU1409/DE3/KU1414-SR-10-B-B
PVL119	KU1409/NC358/KU1409-16-B-B
PVL120	KU1409
PVL121	4001

4.3.3 Genes that significantly affect flux in the second branch (Lycopene - β-carotene - β-cryptoxanthin - zeaxanthin) of the carotenoid biosynthesis

Allele-specific PCR markers based on the *crtRB1* gene for the three SNP polymorphisms were used to genotype the maize inbred line sets carrying different alleles. The PCR markers allowed for screening of polymorphisms in the lines analysed by SNP-based assay. The observed classes of alleles are shown in Table 4.13. The 5'TE polymerase chain reactions amplified lines with 397 and 206 bp insertions, but no line with any insertion were found, instead lines with null alleles were observed in some maize inbred lines in S2. Lines in S1 had alleles with 397 bp insertions. High allelic variation was observed for the *crtRB1* 3'TE, as the expected classes of alleles were found with additional seven different classes at this polymorphic site in S2. For *crtRB1* Del4 polymorphism, the expected allele classes were amplified. Gel photographs showing genetic variations for the three polymorphisms at *crtRB1* are shown in Figures 4.19- 4.21.

Table 4.13: Summary of SNPS in β-carotene hydroxlase 1 (*CrtRB1*) in 38 and 122 tropical-adapted yellow endosperm maize inbred lines.

POLYMORPHISMS		CLASSES OF ALLELES	
Polymorphism site	Types of SNP	Expected size of DNA fragment (bp)	Observed size of DNA fragment (bp)
CrtRB15'	INDEL/TE	800 (with 397 INS)-worst	600[b]
		600 (with 206 INS)- best	800[ab]
		400 (with no INS)-2nd best	Null[ab]
			600+800[b]
CrtRB13'	INDEL/TE	543(without INS)-best	543[b]
		296 + 875 (with 325 INS)-worst	296+1221+1800[ab]
		296 + 1221+1800 (with 1250 INS)-2nd best	296+1221[b]
			296+1800[b]
			543+296[b]
			296+875+1221[b]
			543+296+1221[b]
			543+1221[b]
			296+1221[b]
			296[b]
*CrtRB1*D4	DEL (bp)	129 (with 12 INS)-best	129[b]
		117 (with 12 DEL)-worst	117[ab]
HYD3	Duplicated Seq transcript start site	163-(worst)	163[ab]
		473-(best)	163+608[b]
		476-(best)	
		608-(common)	

[a] Only in 38 inbreds.

[b] Only in 122 inbreds,
[ab] Only in 38 and 122 inbreds
INDEL: Insertion/ Deletion, TE: Transposable Element and Seq: Sequence.

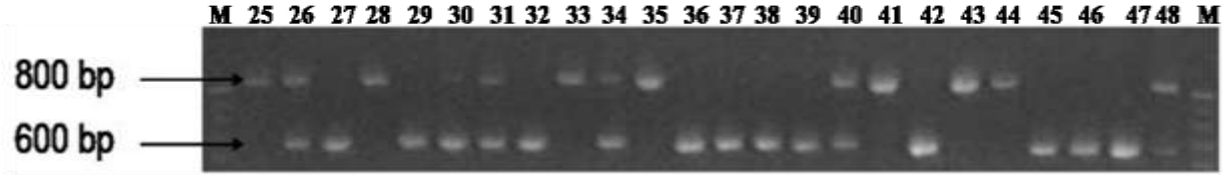


Figure 4.19: *HYDB1/crtRB1* PCR assay 5'INDEL showing the lines with best allele (600 bp) and worst allele (800 bp) on 2 % agarose gel. The first lanes are molecular mass marker and the others are DNA samples for 24 of the 122 tropical yellow maize inbred lines.

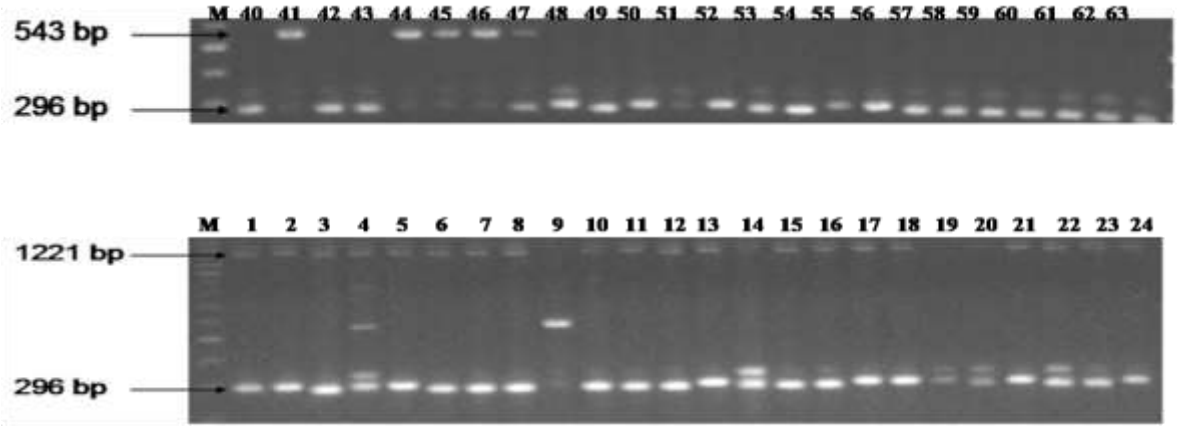


Figure 4.20: *HYDB1/crtRB1* PCR assay 3' TE/INDEL showing the lines with best allele (543 bp), worst allele (296 bp) and (296+1221 bp) on 2 % agarose gel. The first lanes are molecular mass marker and the others are DNA samples for 24 of the 122 tropical yellow maize inbred lines.

.

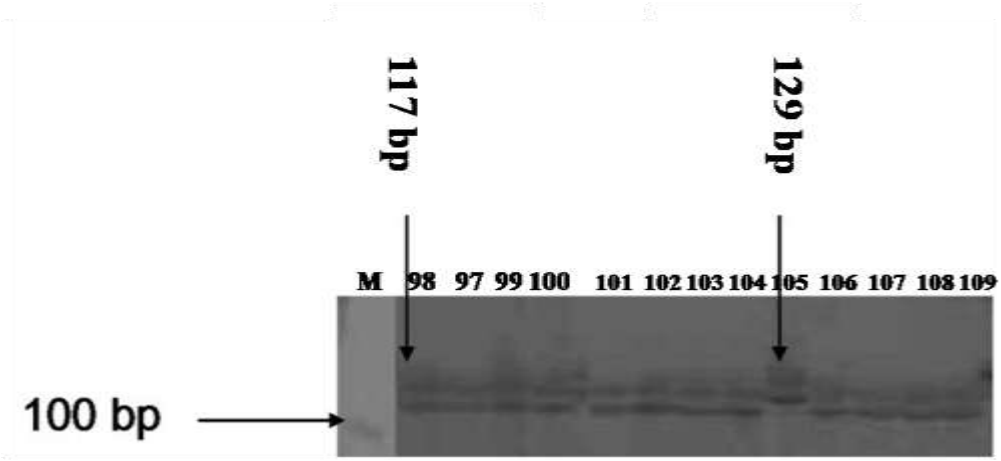


Figure 4.21: *HYDB1/crtRB1* INDEL 4 PCR assay showing the lines with worst allele 117 bp (with 12 bp deletion at InDel4) and best allele129 bp (with 12 bp insertion at InDel 4) on 6 % denaturing polyacrylamide gel. The first lane is the molecular mass marker and the others are DNA samples for 12 of the 122 tropical yellow maize inbred lines.

The frequencies of the favourable alleles for the three *crtRB1* polymorphisms within the two sets were extremely low. The allele 2 of 5'TE (TE insertion of 206 bp) was present at 14 % frequency in set 2, while the allele-frequencies for 3'TE (allele 1, no insertion) and InDel4 (best allele, 12 bp deletion) polymorphisms were 11 % and 0.8 % respectively in set 2 (Table 4.14). In S1, no inbred line with best favourable alleles were found for all the polymorphic sites of the *crtRB1* gene, however, 87 % of the maze inbred lines had the second best class (296 +1221 + 1800 bp haplotype) for the 3'TE. Table 4.15 shows maize inbred lines in set 2 with favourable alleles for 3'TE, 5'TE and InDel4 polymorphisms. Twelve maize inbred lines (9.8 %) had favourable alleles for 3'TE and 5'TE polymorphisms and lower (0.8 %) for 3'TE and InDel4 favourable allele combination. No line was observed carrying favourable alleles for the three polymorphisms (Table 4.13).

The best allele for *crtRB1* was not found among the inbred lines belonging to S1. For *crtRB1* 5'TE polymorphism, the lines with favourable allele had a range of 4.54-9.09 µg for β-carotene content and the lines have a common exotic parent line (DE3) in their genetic background. The 17 lines possessing favourable allele for *crtRB1* 3'TE polymorphism have the TBC range to be 4.92-9.09 µg and they also have DE3 as a parent. The only inbred line **(**KU1409/DE3/KU1409) S2-18-2-B-B) having favourable allele for the InDel4 polymorphism has TBC of 8.66 µg. There was positive correlation between β-carotene concentration and favourable alleles for all *crtRB1* polymorphisms because lines with very high β-carotene concentration have the best alleles.

Table 4.14: Comparison of favourable *CrtRB1* allele frequency in the 38 and 122 tropical adapted yellow endosperm maize inbred lines in the second branch of carotenoid biosynthesis.

Gene polymorphism /haplotype	Population (no of genotypes)	Population (no of genotypes)		Frequency of favourable alleles		
	38	122	Combined	38	122	Combined
***CrtRB1*5'INDEL/TE**						
D1: 800	29	65				
D2: 600	**0**	**17**	**17**	0	14	**11**
D3: 400	0	0				
D4: Null	9	28				
D5: 800+600	0	2				
***CrtRB1*3'INDEL/TE**						
E1:543	0	**14**	**14**	0	11	**9**
E2:296+1221+1800	0	33				
E3:296+1221	0	40				
E4:296+1800	38	4				
E5:543+296	0	6				
E6296+875+1221	0	1				
E7:543+296+1221	0	4				
E8:543+1221	0	1				
E9:296	0	19				
***CrtRB1*DEL4**						
F1: 129	0	**1**	**1**	0	0.8	**0.6**
F2: 117	38	121				
***HYD3*TSS**						
G1: 163	38	102				
G2: 608	0	0				
G3: 163+608	0	20				
G4:473/476+608	0	0	**0**	0	0	**0**

Table 4.15: Genotypes with favourable allele for *CrtRB*1 polymorphisms in 122 tropical adapted yellow maize inbred lines

5'TE IN/DEL polymorphism	
Inbred	**Pedigree**
PVL09	(KU1409/KU1414-SR/A619)-S2-2-B-B
PVL27	(KU1409/DE3/KU1409)S2-2-B-B-B
PVL29	(KU1409/DE3/KU1409)S2-4-B-B-B
PVL30	(KU1409/DE3/KU1409)S2-5-B-B-B
PVL31	(KU1409/DE3/KU1409)S2-6-B-B-B
PVL32	(KU1409/DE3/KU1409)S2-7-B-B-B
PVL34	(KU1409/DE3/KU1409)S2-10-B-B-B
PVL35	(KU1409/DE3/KU1409)S2-11-B-B-B
PVL36	(KU1409/DE3/KU1409)S2-13-B-B-B
PVL37	(KU1409/DE3/KU1409)S2-14-B-B-B
PVL38	(KU1409/DE3/KU1409)S2-15-B-B-B
PVL39	(KU1409/DE3/KU1409)S2-18-B-B-B
PVL42	(KU1409/DE3/KU1409)S2-26-B-B-B
PVL45	(KU1409/DE3/KU1409)S2-31-B-B-B
PVL46	(KU1409/DE3/KU1409)S2-32-B-B-B
PVL47	(KU1409/DE3/KU1409)S2-35-B-B-B
PVL114	KU1409/SC55/KU1409-4-B-B
	3'TE IN/DEL polymorphism
Inbred	**Pedigree**
PVL09	(KU1409/KU1414-SR/A619)-S2-2-B-B
PVL27	(KU1409/DE3/KU1409)S2-2-B-B-B
PVL29	(KU1409/DE3/KU1409)S2-4-B-B-B
PVL32	(KU1409/DE3/KU1409)S2-7-B-B-B
PVL37	(KU1409/DE3/KU1409)S2-14-B-B-B
PVL38	(KU1409/DE3/KU1409)S2-15-B-B-B
PVL39	(KU1409/DE3/KU1409)S2-18-B-B-B
PVL42	(KU1409/DE3/KU1409)S2-26-B-B-B
PVL45	(KU1409/DE3/KU1409)S2-31-B-B-B
PVL46	(KU1409/DE3/KU1409)S2-32-B-B-B
PVL47	(KU1409/DE3/KU1409)S2-35-B-B-B
PVL105	(KU1409/DE3/KU1409)S2-18-2-B-B
PVL113	KU1409/NC358/KU1409-17-BB-B-B
PVL114	KU1409/SC55/KU1409-4-B-B
	Del 4 polymorphism
Inbred	**Pedigree**
PVL105	(KU1409/DE3/KU1409)S2-18-2-B-B

The 105 lines in S2 without the unfavourable allele for *crtRB1* 5'TE polymorphism have β-cryptoxanthin concentration to be between 0.31 and 5.85 µg with an average of 3.56 µg as well as lowest (0.36 µg) to highest (26.2 µg), average of 10 µg for zeaxanthin. In comparison, 108 lines analysed had unfavourable allele for *crtRB1* 3'TE and concentrations of β-cryptoxanthin to be 0.58-5.85 µg with a mean of 3.52 µg and zeaxanthin, 0.73-26.2 µg with a mean of 10.2 µg. All the lines in S2 having unfavourable alleles for *crtRB1* gene have diverse genetic backgrounds and there is strong association of high carotenoid xanthophylls content with unfavourable allele variant.

The gene based PCR markers to detect SNP polymorphism at *HYD3* transcript site in the two sets were assayed. S1 and S2 primer pair did not produce PCR product for the 608 allele in all genotypes. In total, 20 genotypes among the 122 lines had the common *HYD3* paralog-specific product. The unfavourable allele A with 168 bp was present in all the analysed materials. Gel photographs showing genetic variations for the *HYD3* polymorphism is shown in Figure 4.22**.** The combined group of 160 inbred lines had no favourable sequence variants of *HYD3* "C" (473 bp) or "B" (476 bp) in the transcript start site (duplicated sequence variant polymorphism).

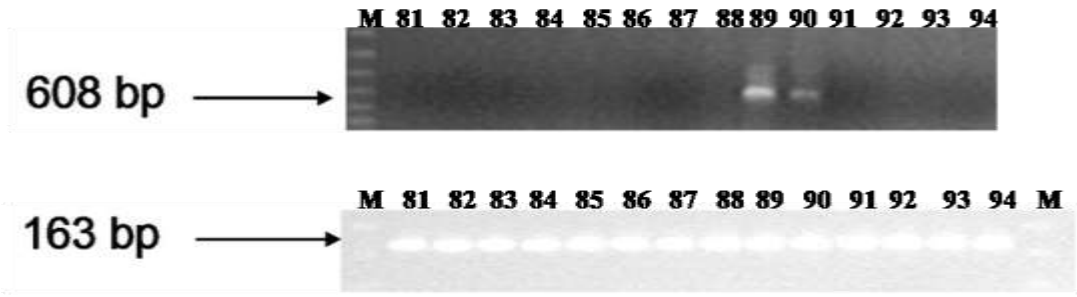


Figure 4.22: *HYD3* PCR assay showing the lines with worst allele (163 bp) and a common allele (608 bp) on 2 % agarose gel. The first lane is the molecular mass marker and the others are DNA samples for 14 of the 122 tropical yellow maize inbred lines.

4.3.4 SNP allelic haplotype diversity at *LCYE* and *crtRB1* loci

The number of SNP allelic combinations or diversity at two gene loci present among the 122 and 38 maize inbred lines was determined (Table 4.16). Fifteen lines combined different favourable alleles at *LCYE* and *crtRB1* loci with five favourable haplotype combination classes. The maize inbred lines that contained three or more SNPs classes in the two loci are shown in Table 4.17. In one class, five maize inbred lines combined favourable alleles at *LCYE* 3'TE, *crtRB1* 3'TE and 5'TE. In only two maize inbred lines were found having *LCYE* 3'TE and *crtRB1* 5'TE combining favourable alleles. The 7 inbred lines having the favourable alleles for *LCYE* (SNP216, 3'TE) and *crtRB1* (3'TE, 5'TE) polymorphisms have a TBC range of 5.52 to 8.99 µg, while the 7 maize inbred lines possessing the *LCYE* (SNP216) and *crtRB1* (3'TE, 5'TE) favourable alleles have their TBC between 4.93 and 9.09 µg. Line PVL 105 has favourable alleles at *LCYE* (SNP216) and *crtRB1* (3'TE, indel4) and the TBC is 8.67 µg.

Table 4.16: Frequency of lines with *LCYE and crtRB1* haplotype combinations in the 38 and 122 tropical-adapted yellow maize inbred lines

LCYE					
Panel	N	**SNP216 + 3' TE**	**SNP216 + 5' TE**	**5' TE + 3' TE**	**SNP216 + 3' TE + 5' TE**
P1	38	11	9	0	0
P2	122	4	81	0	0
crtRB1					
Panel	N	**5' TE + 3' TE**	**InDel 4+3' TE**	**InDel4 + 5' TE**	**5' TE + 3' TE + InDel4**
P1	38	0	0	0	0
P2	122	18	1	0	0

Haplotypes are shown as combinations of polymorphism sites. They are in bold face type, and only number of lines having them are counted.

Table 4.17: Tropical adapted yellow endosperm maize inbred lines with *LCYE* and *crtRB1* haplotype combination for favourable alleles in 122 set

Inbreds	*LCYE* 3' IN/DEL	*LCYE* SNP 216	*crtRB1* 3' TE	*crtRB1* Del 4	*crtRB1* 5' INDEl/TE
PVL09	**144+502**	**G**	**543**		**600**
PVL32	**144+502**	**G**	**543**		**600**
PVL38	**144+502**	**G**	**543**		**600**
PVL46	**144+502**	**G**	**543**		**600**
PVL114	**144+502**	**G**	**543**		**600**
PVL27		**G**	**543**		**600**
PVL29		**G**	**543**		**600**
PVL35	**144+502**	**G**			**600**
PVL36	**144+502**	**G**			**600**
PVL37		**G**	**543**		**600**
PVL39		**G**	**543**		**600**
PVL42		**G**	**543**		**600**
PVL45		**G**	**543**		**600**
PVL47		**G**	**543**		**600**
PVL105		**G**	**543**	**129+350**	

Haplotypes are shown as combinatons of favourable alleles of *LCYE* (144 + 502 -bp, G), favourable *crtRB1* (543 -bp, 129 + 350 -bp, 600 -bp). Alleles are in color and only lines with the observed haplotypes are listed.

4.4 SSR-based diversity among 122 maize inbred lines with the allelic diversity based β-carotene candidate genes

Two of the 4 inbred lines possessing favourable allele for *LCYE* 5'TE polymorphism were grouped in group I while the other two were found in group IV based on dendrogram resulting from UPGMA cluster analysis of SSR data and groupings were consistent with respect to pedigree data (POP66R or 9450XKI21). The 82 lines having favourable allele for *LCYE* 3'TE were found in the 4 groups based on SSR data. The majority of the lines developed from KU1414 or KU1409 were contained in group III and 5 lines derived from POP66R genetic background belonged to either group I, II, or IV. All inbred lines with favourable alleles for *crtRB1* (3'TE, 5'TE and InDel4) polymorphisms belonged to group III in accordance with the genetic background (KU1409/DE3) based on SSR data. The maize inbred lines with *LCYE* and *crtRB1* haplotype combination for favourable alleles in 122 set constituted into group III based on SSR data according to similar genetic background. Inbred lines having unfavourable alleles for both genes were not restricted to particular groupings but were grouped in any of the four groups identified by the SSR data.

CHAPTER FIVE

5.0 DISCUSSION

Hybrid maize breeding programs depend on selection of diverse lines to be able to get potential hybrid vigour as a means of improving crop productivity (Melchinger, 1999). Understanding the genetic diversity among important breeding materials is generally considered a critical first step to achieve successful goals in breeding programs. In the present study PCR-based molecular techniques were used to determine the genetic relationships among 38 and 122 maize inbred lines. The AFLP and SSR loci used in this study are well distributed across the entire maize genome, but they measure differences in different regions of the genome. AFLP markers generated a high level of polymorphism in the lines analysed, while the SSR polymorphism was low in spite of using a core set of highly informative markers. These markers detected considerable level of genetic diversity among the 38 yellow endosperm maize inbred lines. The two markers revealed similar average genetic distance estimates in spite of measuring different DNA polymorphisms in the genome.

The level of allelic richness obtained among the 38 maize inbreds was moderate, possibly due to high level of relatedness among most of the lines. The mean allele per SSR marker in this study was lower than the average of 4.9 to 7.4 alleles reported in previous SSR studies (Senior *et al.,* 1998; Lu and Bernardo, 2001; Warburton *et al.,* 2002; Xia *et al.,* 2004). However, the average value found in this study was higher than the 3.85 reported by Legesse *et al.* (2007). Such differences in detecting number of alleles can result from

differences in diversity present in the lines, the number of genotyped lines and genotyping methods used to reveal differences in allele sizes. The results showed that the two markers were able to detect considerable level of genetic diversity among the yellow endosperm inbred lines although most of them share at least one parent in their pedigrees.

The lines also exhibited marked differences in carotenoid concentrations. The lines included in this study were derived from different source germplasm that can be used as parents to develop new lines having higher level of pro-vitamin A. The genetic relationships portrayed by cluster analysis of SSR data were to some extent similar to the output of AFLP cluster analysis, as both dendrograms placed most of the lines into similar groups. Nearly 80 % of the yellow endosperm inbred lines were included in two major groups created with both markers. However, some inbred lines clustered with different sets of lines when the AFLP and SSR markers were used, possibly due to differences in the capacity of the markers to detect variations in different parts of the maize genome. Gerdes and Tracy (1994) pointed out that pedigree relationship can be used as a benchmark to test the effectiveness of markers in determining relationships among breeding lines.

AFLP and SSR markers efficiently separated the tropical-adapted yellow maize inbred lines into groups, consistent with pedigree. However, SSR markers were found to be more effective in separating closely related lines, which was similar to some previous findings (Smith *et al.,* 1997; Legesse *et al.,* 2007; Reif *et al.,* 2003; Romero-Serverson *et al.,* 2001). This disparity in separating the inbred lines on the basis of pedigree records

may be related to the extent of genome coverage of SSR loci to detect differences in the origin of the lines. The complex directional mutations that produce SSR loci (Morgante and Olivieri, 1993) and its high mutation rates (Vigouroux *et al.,* 2005) may have enhanced their power to detect genetic diversity among the diverged lines, used in this study. The inability of AFLPs markers to differentiate heterozygotes from homozygotes may have rendered them to be less effective to detect the genetic diversity among closely related inbred lines compared to SSR markers. The SNP analysis of *PSY1* shows the absence of genetic polymorphism in the regions of the gene studied among the lines evaluated. In a variability study within *Zea* species, Palaisa *et al.* (2003) sequenced a small region of *Y1* (nucleotides 1331 to 2185), known to contain the insertion site of *Ins2*, the CCA microsatellite, and SNPs in the coding region; the studied regions were found to be heterozygous in several *Zea* species.

The genetic relationships among 122 tropical maize inbred lines were elucidated using SSR markers. In the 122 maize inbred lines, the average number of SSR alleles per locus for the 51 polymorphic loci was 3.72, which is comparable to the 3.96 alleles per locus for 75 polymorphic SSR loci in 38 yellow tropical inbred lines. The average number of alleles per locus (4.44) for 43 SSR primers in the sample of 15 Chinese maize inbreds reported by Xu *et al.* (2004) was higher than for these lines. Lu and Bernado (2001) found an average of 4.9 alleles per locus in 40 U.S maize inbreds across 83 markers. The homozygousity present in the various inbred lines may have accounted for this range of allelic diversity. Furthermore, differences of average number of alleles reported for each study may be due to types of SSR markers used and composition of the inbreds. The

discriminatory power of the markers measured as polymorphic information content (PIC) showed an average value of 0.43 which strongly suggests that SSR markers are highly informative. This also showed a significant level of genetic diversity detected in these maize inbreds. It was observed that many of the maize inbreds had common parents in their genetic background; this might be a possible reason for the low PIC. In the UPGMA dendrogram analysis, most of the maize inbred lines were grouped with respect to their common origin as well as parental relationship. Thus, the reason for grouping can be explained based on good agreement with the known pedigree and genetic background information. Again, these results support SSR markers to be generally efficient in detecting genetic diversity among closely related genotypes (Senior *et al.,* 1998; Smith *et al.,* 1997). SSR markers are highly reproducible, reliable and have been promisingly consistent in the discrimination of maize inbred lines into groups based on pedigree data.

This reflection of pedigree information in groupings based on SSR markers makes this DNA marker efficient (Legesse *et al.,* 2007). Overall genetic diversity (0.41) among all pairs of tropical maize inbred lines using the SSR markers revealed substantial levels of genetic diversity assessments for this present study. The knowledge of the level of genetic diversity among 122 tropical maize inbreds suggests their possible use in selection programs as important source of diverse alleles for increasing the content of β-carotene in the endosperm of yellow maize.

Although the correlations between the genetic distance matrices of the two markers were weak, AFLP and SSR markers detected comparable level of genetic relationships among

the 38 lines. Poor correlation of diversity estimates from different DNA marker systems have also been reported in other studies (Allan *et al.,* 2007; Schut *et al.,* 1997; Laborda *et al.,* 2005; Manifesto *et al.,* 2001). Garcia *et al.* (2004) observed low correlation of 0.33 between RAPD and SSR markers. Strong associations between genetic distances generated by different markers have also been reported for various molecular markers (Smith *et al.,* 1997; Pejic *et al.,* 1998; Lubbersted *et al.,* 2000; Garcia *et al.,* 2004). Both molecular markers were more effective in grouping the lines along their genetic backgrounds than the carotenoid data based grouping possibly because the underlying cause of the variation in carotenoid content could be independent of the variation observed at the DNA level.

The carotenoid-based cluster analysis separated the 38 lines into groups based on differences in concentrations of carotenoids in the grain. Inbred lines originating from diverse source germplasm with either higher or lower concentrations of the different carotenoids were thus grouped together. On the other hand the two markers separated the lines into groups based on their origin and parentage. To some extent both the AFLP and SSR markers also grouped the lines on the basis of differences in carotenoid concentrations. Overall, AFLP and SSR markers separated the yellow endosperm maize inbred lines into well defined groups. The correlations of molecular marker-based clustering with caroteniod-based clustering were not strong, suggesting that the molecular analysis could be used to identify diverse lines with high pro-vitamin A that can be used as parents for making bi-parental crosses or hybrids to increase the level of pro-vitamin A

in tropical maize. Phenotypic data like carotenoid content are affected by environment and by genotype interaction

Nutritionally, the 122 maize inbreds vary considerably for carotenoid data. The separation of the 122 lines into groups for carotenoid-based cluster analysis was based mainly on differences in carotenoid concentrations in the grain and this also suggests the variability among the inbreds. The maize inbred lines found in each group of the carotenoid based UPGMA dendrogram have either higher or lower concentrations of the different carotenoids belonged to diverse origin and genetic backgrounds. When compared with the genetic diversity determined by SSR markers, it was apparent that the grouping patterns between the two data based clustering were somewhat different. SSR markers separated the 122 inbred lines into groups based on their common origin and parentage more efficiently than the carotenoid based data.

Hybrid maize development has been successful over years and its manifestation depends on the genetic divergence of the two parental lines (Moll *et al.,* 1965). Inbred lines from different heterotic groups generally contribute to superior hybrid performance in breeding practice. Generally, SSR-based groupings depended mostly on the pedigree information than the caotenoid-based groupings. The groupings of maize inbred lines based on SSR data would facilitate the assigning of inbred lines into heterotic groups for optimum selection gains and the development of new maize hybrids. This study has identified a substantial level of genetic diversity in the 122 tropical maize inbreds based on SSR data. The findings garnered about the genetic relationship among the yellow maize inbred lines

from this study will provide maize breeders more information for effective selection of different inbred lines as potential parent lines.

Marker-trait association analyses based on two important candidate genes encoding enzymes in carotenoid biosynthesis in maize endosperm have been evaluated and reported to have large effects on the β-carotene accumulation in a diverse set of yellow maize inbred lines with different genetic backgrounds (Harjes *et al*., 2008; Yan *et al.*, 2010; Vallabhaneni *et al*., 2009). SNP markers for amplification of functional or non-functional alleles for the two genes have also been designed from polymorphic regions that affect carotenoid variation which may be useful as genetic markers for marker assisted selection (MAS) in maize breeding (Xu and Crouch, 2008). In this study, single nucleotide polymorphism of candidate genes in the β-carotene synthesis pathway was validated in tropically adapted yellow endosperm maize inbred lines. Thus the robustness of the SNP-based PCR based primers was established to determine the specificity to amplify either insertion or deletions (indels) and single nucleotide polymorphisms for the two genes as well as characterisation of allelic variants at these loci in tropical maize inbred lines. The introgression of favourable alleles into other breeding lines will produce maize varieties and hybrids containing sufficient levels of β-carotene in their endosperm.

The 38 and 122 locally adapted maize inbred lines showed variation based on SNPs present which were detected in both genes revealing the presence or absence of favourable polymorphisms. This indicates that variation in DNA sequence is the underlying cause of phenotypic variation in carotenoid concentration of maize grains and is due to regulation at the level of gene expression within the targeted genes. The PCR

genotyping method presented in this study showed that a simple primer pair had higher accuracy than primer multiplexing which was used in other studies (Harjes *et al*., 2008; Yan *et al*., 2010; Vallabhaneni *et al*., 2009) to determine the presence or absence of SNPs except for the 3' region of *crtRB1* polymorphism for which 3 primers were multiplexed. In the multiplex PCR, many times there was interference with the assay. Using single pair PCR assay, PCR products had very good amplification showing strong and visible bands and differences in allele sizes were easily resolved on agarose/polyacrylamide gel electrophoresis. Therefore, alleles produced in the PCR amplification were correctly scored according to appropriate fragment size. Hence, functional markers were effective to genotype and readily distinguished individual inbred for naturally occurring allelic variants in the two genes involved in phenotypic variation in β-carotene content of the maize endosperm. The optimized methodology for SNP detection has been successfully utilized to verify *LCYE* and *crtRB1* gene polymorphisms in the tropical maize inbred lines. This suggests the applicability of the SNPs for use in marker assisted selection (MAS) in a breeding program to improve pro-vitamin A content in maize for allelic selection to predict β-carotene phenotype at seedling stage.

The two sets differ in SNP allele frequency for both genes. The difference between the two sets may be due to inbred lines belonging to diverse genetic background. For the *LCYE* gene, several lines carried the favourable allele G in both panels for the SNP216 polymorphism. This suggests that this is the most common polymorphism that would frequently occur in any yellow endosperm tropical maize inbreds and the effect of this polymorphism may likely not be crucial for α-carotene and β-carotene accumulation in

these set of maize inbreds. Frequency of the lines containing favourable allele (144+502 bp haplotype class), with 8 bp deletion in the genomic sequence of 3'TE differs in the two panels. The occurrence of this polymorphism was highest in S2 (67 %) compared with S1 (29 %). In contrast, the most favourable class with promoter transposon insertion of 5'TE occurred more in S1 than in S2 (3.3 %). This suggests that there is a significant association of *LCYE* 5'TE polymorphism in some lines. The reason for the differences observed in frequency of occurrences could possibly be due to the differing in genetic backgrounds among maize inbred lines, though some maize inbred lines with low α-carotene or β-carotene concentration possessed favourable alleles for *crtRB1* and *LCYE* genes. The possible explanation for this could be due to environmental conditions which can influence the accumulation of pro-vitamin A carotenoids in maize grains.

The frequency results for the favourable alleles for the *crtRB1* gene was not as high in the two sets of inbred lines as was found for *LCYE*. This result supports the previous reports that favourable alleles for *crtRB1* gene are more common in temperate germplasm than in tropical germplasm (Harjes *et al*., 2008; Yan *et al*., 2010). High allelic variation was observed for the *crtRB1* 3'TE, as the expected classes of alleles were found with additional seven different classes at this polymorphic site in S2, which were not reported by Yan *et al.* (2010). This may be as a result that most of the lines are crosses of tropical and temperate parent lines. Perhaps the absence of *crtRB1* functional alleles might likely account for the very low accumulation of β-carotene in maize lines adapted to the tropics because *crtRB1* is related to accumulation of more β-carotene in maize endosperm at the expense of significant reduction in zeaxanthin. It has been reported that β-carotene is present in highest concentration than α-carotene in maize

(Wong *et al*., 2004). The absence of favourable allele among all the tropical lines analysed for *HYD3* (*crtRB1*) promoter variation (B, C) may be indicative of the rare occurrence of genetic variation in *crtRB1* gene. However, this probably suggests that this polymorphism is weakly associated with β-carotene (Yan *et al*., 2010).

Overall the maize inbreds in the two sets that contain functional alleles at both *LCYE* 3'TE and *LCYE* 5'TE have TBC as low as 0.60 µg and as high as 8.99 µg, while the TBC for inbred lines in S2 with *crtRB1* favourable alleles is as low as 4.54 µg and as high as 9.09 µg. This suggests that inbred lines carrying *crtRB1* favourable alleles have higher TBC, though, the presence of more than one favourable allele may have accounted for their high TBC content in these inbred lines and thus there is a correlation. The genetic joint effects of functional gene variants (*LCYE* and *crtRB1*) may be involved in synthesis of high β-carotene where maize inbred lines have favourable allele at both loci. This could have resulted from the fact that the inbred lines used in this study were derived from crosses and backcrosses between the tropical and temperate germplasms of varying carotenoid composition and diversity in genetic backgrounds.

Favourable haplotypes of *LCYE* and *crtRB1* were noted to occur together naturally in this study, although previous studies (Harjes *et al*., 2008; Yan *et al*., 2010) found none. The lines possessing more than one favourable allele are good start for introgression, they will facilitate maize breeding. The presence of *crtRB1* favourable alleles in maize inbred lines having high β-carotene concentration analysed in this study supports that this locus is a key gene essential in breeding for enhanced level of β-carotene (Yan *et al*., 2010). It can be concluded that the presence of the favourable allele for *crtRB1* had a major influence on β-carotene accumulation.

In addition, the PVL105 maize inbred line does not have favourable allele for the significant *LCYE* polymorphisms but has two favourable alleles for *crtRB1*, this may also indicate that the *LCYE* contribution to β-carotene concentration is minimal. It was observed that some genotypes with high TBC did not show the presence of the most significant favourable allelic variations responsible for the high TBC in the lines. The favourable allele for SNP216 polymorphism that is present in all the inbred lines may have contributed to the β-carotene content. Perhaps, additional research work needs to be done to identify other genes or polymorphisms responsible for the accumulation of high β-carotene in maize inbred lines. Lycopene cyclization is controlled in both branches by two enzymes, *LCYB* (involved in both branches) and *LCYE* (involved only in the β branch, so the possibility that genetic variation influencing the activity of *LCYB* needs to be established. This study found that the absence of the favourable alleles for *crtRB1* gene might have resulted into the accumulation of large amounts of xanthophylls (lutein and zeaxanthin are derived from α-carotene and β-carotene pathways respectively) in these maize inbred lines as well as in maize genotypes generally. β-carotene hydroxylase 1 gene specific to the ε-ring and β-ring catalyze the double hydroxylation of α-carotene and β-carotene (Cunning and Gantt, 1998) resulting in the formation of lutein and zeaxanthin respectively. Maize accumulates large concentration of xanthophylls in their endosperm (Harjes *et al*., 2008; Menkir *et al*., 2008).

The gene specific functional markers for *LCYE* and *crtRB1* polymorphisms are reliable and efficient. This study has shown that some tropical adapted maize inbred lines harbour most favourable alleles at the two loci. Therefore, potential parental genotypes can be used to breed for higher beta carotene content at the *LCYE* and *crtRB1* favourable alleles

using the allelic-specific marker assisted selection. They will be useful for selection, recombination and as well as introgression of alleles in maize breeding programs for increased level of pro-vitamin A concentration in maize.

The groupings of 122 inbred lines generated from SSR-based data demonstrated that the lines belonging to the same groupings varied for the gene variants. This further suggests that genetic variations in candidate genes for β-carotene exist among the inbred lines irrespective of their genetic backgrounds. In addition, this goes on to show that when different parts of the genome are assayed, they may not correlate. Although, the SSR data grouped the 122 inbred lines into four groups, lines with the same favourable alleles for *LCYE* were not restricted to a particular group of the SSR cluster analysis. However, for *crtRB1* gene, lines with favourable alleles belonged to a specific group while the lines with unfavourable allele for both loci were found in any of the groups formed using SSR data. Overall, the grouping of lines based on SSR data and the presence of favourable alleles of β-carotene genes in the carotenoid biosynthesis pathway should be considered to make informed decisions.

Information from the current study could be used for tropical-adapted yellow endosperm maize improvement.

Conclusion

Genetic diversity assessments are critical to the successful selection of promising parent combinations for genetic improvement of any important trait. The study has provided insights to the genetic potential that exists among the inbred lines that will facilitate the effective utilization of lines and exploitation of hybrid vigour for development of new maize varieties having high content of β-carotene. The gene specific functional markers are reliable and efficient and they were used to identify inbred lines with favourable allelic states for *LCYE* and *crtRB1*. These lines will be useful for selection, recombination and as well as introgression of alleles in MAS breeding programs. Also, these diverse sets of tropical maize inbred lines are useful sources for selection of promising lines for development of new varieties in maize biofortification.

Contribution to knowledge

1. The GD estimates can be used as the basis for effective utilization of the yellow endosperm inbred lines with diverse genetic backgrounds.
2. Assigning of maize inbred into groupings for increased efficiency for designing accurate crosses to maximise genetic gains.
3. Validation of allele-specific markers for amplifying functional alleles associated with accumulation of β-carotene content has provided the opportunity for carotenoid enhancement efforts in tropical maize.
4. Identification of tropical inbred lines having most favourable alleles for *LCYE* and *crtRB1* will be used for efficient introgression and selection in the national maize breeding programs.

REFERENCES

Adetimirin, V.O., Vroh-BI, I., The C., Menkir, A., Mitchell, S.E. and Kresovich, S. (2008). Diversity analysis of elite maize inbred lines adapted to west and central Africa using SSR markers. *Maydica.* 53: 143-149.

Agarwal, S. and Rao, A.V. (2000). Tomato lycopene and its role in human health and chronic diseases. *Canadian Medical Association*. 163(6): 739-44.

Aguayo, V.M. and Baker, S.K. (2005). Vitamin A deficiency and child survival in Saharan Africa: A reappraisal of challenges and opportunities. *Food and Bulletin*. vol. 26, no. 4.

Aguayo, V.M., Kahn. S., Ismael. C. and Meershoek, S. (2004). Vitamin A deficiency and child mortality in Mozambique. *Public Health Nutrition*. 8(1): 29–31.

Aguiar, C.G., Schuster, I., Amaral Jr. A.T., Scapim, C.A. and Vieira, E.S.N. (2008). Heterotic groups in tropical maize germplasm by test crosses and simple sequence repeat markers. *Genetics and Moecular Reearch.* 7: 1233-1244.

Ajmone-Marsan, P., Castiglioni, P., Fusari, F., Kuiper, M. and Motto, M. (1998). Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. *Theoretical and Applied Genetics.* 96: 219-227.

Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. (1992). Length Polymorphisms of Simple Sequence Repeat DNA in Soybean. *Genetics.* 132: 1131-1139.

Allan, G., Williams, A., Rabinowcz, P.D., Chan, A.P., Ravel, J. and Keim, P. (2007). Worldwide genotyping of castor bean germplasm (*Ricinus communis L.*) using AFLPs and SSRs. *Genetic Resources and Crop Evolution*. 55(3): 365-378.

Alquezar, B., Rodrigo, M.J. and Zacarías, L. (2008). Regulation of carotenoid biosynthesis during fruit maturation in the red-fleshed orange mutant Cara Cara. *Phytochemistry.* 69: 1997-2007.

Aluru, M., Xu, Y., Guo, R., Wang, Z., Li, S., White, W., Wang, K. and Rodermel, S. (2008) Generation of transgenic maize with enhanced pro-vitamin A content. *Journal of Experimental Botany.* 59: 3551-3562.

Babu, B.K., Agrawal, P.K., Mahajan, V. and Gupta, H.S. (2009). Molecular and Biochemical Characterization of Short Duration Quality Protein Maize. *Journal of Plant Biochemistry and Biotechnology*: 18(1).

Bai, L., Kim, E.H., Dellapenna, D. and Brutnell, T.P. (2009). Novel lycopene epsilon cyclase activities in maize revealed through perturbation of carotenoid biosynthesis. *Plant Journal.* 59: 588-599.

Bartley, G.E. and Scolnik, P.A. (1995). Plant carotenoids: pigments for photoprotection, visual attraction, and human health. *Plant Cell.* 7: 1027-1038.

Batley, J., Barker, G., O'Sullivan, H., Edwards, K.J. and Edwards, D. (2003). Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. *Plant Physiology.* 132: 84–91.

Benchimol, L.L., Souza, C.L., Jr, Garcia, A.A.F., Kono, P.M.S., Mangolim, C.A., Barbosa, A.M.M., Coelho, A.S.G. and Souza, A.P. (2000). Genetic diversity in tropical maize inbred lines: Heterotic group assignment and hybrid performance determined by RFLP marker. *Plant Breeding.* 119: 491-496.

Bernardo, R. (1992). Relationship between single-cross performance and molecular marker heterozygosity. *Theoretical and Applied Genetics.* 83: 628-634.

Beyene, Y., Botha, A-M and Myburg, A.A. (2005). A comparative study of molecular and morphological methods of describing genetic relationships in traditional Ethiopian highland maize. *African Journal of Biotechnology*. 4 (7): 586-595.

Beyer, P., Al-Babili, S., Ye, X., Lucca, P., Schaub, P., Welsch, R., and Potrykus, I. (2002). Golden rice: introducing the β-carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat vitamin A deficiency. *Journal of Nutrition*. 132, 506S-510S.

Bhattramakki, D. and Rafalski, A. (2001). Discovery and application of single nucleotide polymorphism markers in plants. In: HenryRJ (ed) Plant genotyping: the DNA Fingerprinting of plants. CAB International, Lismore, pp 179–191.

Bhattramakki, D., Dolan, M., Hanafey, M., Wineland, R., Vaske, D., Register, J.C., III, Tingey, S.V. and Rafalski, A. (2002). Insertion deletion polymorphisms in 39 regions of maize genes occur frequently and can be used as highly informative genetic markers. *Plant Molecular Biology*. 48, 539–547.

Blessin, C.W., Brechner J.D and Dimler, R.J. (1963). Carotenoids of corn and sorghumV. Distribution of xanthophylls and carotenes in hand-dissected and dry milled fractions of Yellow Dent Corn. *Cereal Chemistry*. 40: 582-586.

Botstein, D., White, R.L., Skolnick, M and Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics.* 32: 314-331.

Bohn, M., Utz, H. F. and Melchinger. A. E. (1999). Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their use for pre-dicting progeny variance. *Crop Science.* 39: 228-237.

Bouchez, A., Hospital, F. Causse, M., Gallais, A. and Charcosset, A. (2002). Marker-assisted introgression of favorable alleles at quantitative trait loci between maize elite lines. *Genetics*. 162(4): 1945–1959.

Brown, E. D., Micozzi, M. S., Craft, N. E., Bieri, J. G., Beecher, J., Edwards, B. K., Rose, A., Taylor, P. R. and Smith, J. C., Jr. (1989). Plasma carotenoids in normal men after a single ingestion of vegetables or purified ß-carotene. *American Journal Clinical Nutrition.* 49: 1258-1265.

Breseghello, F. and Sorrells, M.E. (2006). Association analysis as a Strategy for Improvement of Quantitative Traits in Plants. *Crop Science Society of America.* 46:1323-1330.

Brunson, A.M. and Quackenbush, F.W. (1962). Breeding corn with high pro-vitamin A in the grain. *Crop Science*. 2:344–347.

Buckler, E.S. and Thornsberry, J.M. (2002). Plant molecular diversity and applications to genomics. *Current Opinion in Plant Biology*. 5: 107–111.

Buckler. E.S., Miguel, P.S., Janick-Buckner, D. and Bennet J.L. (1999). The y1 gene of maize codes for phytoene synthase. *Genetics*.143: 479-488.

Buckner, B., Kelson, T.L. and Robertson, D.S. (1990). Cloning of the y1 locus of maize, a gene involved in the biosynthesis of carotenoids. *Plant Cell*. 2:867–876.

Chander, S., Guo, Y.Q., Yang, X.H., Zhang, J., Lu, X.Q., Yan, J.B., Song, T.M., Rocheford, T.R., Li, J.S. (2008). Using molecular markers to identify two major loci controlling carotenoid contents in maize grain. *Theoretical and Applied Genetics.* 116: 223-233.

Chavarriaga-Aguirre, P., Maya, M.M., Tohme, J., Duque, M.C., Iglesias, C., Bonierbale, M.W., Kresovich, S. and Kochert, G. (1999). Using microsatellates, isoenzymes and AFLP to evaluate genetic diversity and redundancy in the cassava core collection and to assess the usefulness of DNA-based markers to maintain germplasm collections. *Molecular Breeding*. 5:263-273.

Christian, P., West, K.P. Jr, Khatry, S.K., Katz, J., LeClerq, S. C., Kimbrough-Pradhan , E., Dali, S. M. and Shrestha, S.R. (2001). Vitamin A or β-carotene supplementation reduces symptoms of illness in pregnant and lactating Nepali women. *Journal of Nutrition*. 130: 2675–2682.

Christian, P., West, K.P. Jr, Khatry, S.K., Katz,.J., Shrestha, S.R., Pradhan, E.K., LeClerq, S.C and Pokhrel, R.P. (1998). Night blindness of pregnancy in rural Nepal—nutritional and health risks. *International Journal of Epidemiology*. 17: 231-237.

Christian, P., West, K. P. Jr., Khatry, S.K., Kimbrough-Pradhan, .E., LeClerq, S.C., Katz, J., Shrestha, S.R., Dali, .S.M. and Sommer, .A. (2000). Night blindness due to vitamin A deficiency is common during pregnancy among women in Nepal. *American Journal Epidemiology*. 152: 542–447.

Clark, R. M., Linton, E., Messing, J. and Doebley, J. F. (2004). Patternof diversity in the genomic region near the maize domestication gene *tb1*. *Proceedings of the National Academy of Sciences of USA*. 101: 700–707.

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

Crouch, J.H. and Ortiz, R. (2004). Applied genomics in the improvement of crops grown in Africa. *African Journal of Biotechnology*. 39(10): 489-496.

Cunningham, F.X., Pogson, B., Sun, Z., McDonald, K.A., DellaPenna, D. and Gantt, E. (1996). Functional analysis of the b- and e- lycopene cyclase enzymes of Arabidopsis reveals a mechanism for control of cyclic carotenoid formation. *Plant Cell*. 8: 1613–1626.

Cunningham, F.X. (2002). Regulation of carotenoid biosynthesis and accumulation in plants. *Pure and Applied Chemistry*. 74: 1409–1417.

Cunningham, F.X. and Gantt E. (1998). Genes and enzymes of carotenoid biosynthesis in plants. *Annual Review of Plant Physiology and Plant Molecular Biology*. 49: 557–583.

DellaPenna, D. and Pogson, B.J. (2006). Vitamin synthesis in plants: tocopherols and carotenoids. *Annual Review of Plant Physiology and Plant Molecular Biology*. 57: 711–738.

DellaPenna, D. (1999). Nutritional genomics: manipulating plant micronutrients to improve human health. *Science*. 85 (5426): 375–379.

Dellaporta, S.L., Wood, J and Hicks, J.B. (1983). A plant DNA minipreparation: Version II. *Plant Molecular Biology Reporter*. 1: 19-21.

Diretto, G., Tavazza, R., Welsch, R., Pizzichini, D., Mourgues, F., Papacchioli, V., Beyer, P. and Giuliano, G. (2006). Metabolic engineering of potato tuber carotenoids

through tuber-specific silencing of lycopene epsilon cyclase. *BioMedCentral Plant Biology*. 6:13-24.

Doebley, J.F., Goodman, M.M. and Stuber, C.W. (1985). Isozyme variation in the races of maize from Mexico. *American Journal of Botany*. 72: 629-639.

Dubreuil, P., Dufour, P., Drejci, E., Causse, M., de Vienne, D., Gallais, A. and Charcosset. A. (1996). Organization of RFLP diversity among inbred lines of maize representing the most significant heterotic groups. *Crop Science.* 36:790–799.

Eathington, S.R., Crosbie, T.M., Edwards, M.D., Reifer, R.S. and Bull, J.K. (2007). Molecular Markers in a commercial Breeding Program. *Crop. Science*. 47(5): 154 – 163.

Fan, X.M., Zhang, Y.M., Yao, W.H., Chen, H.M., Tan, J., Xu, C.X., Han, X.L., Luo, L.M., Kang, M.S. (2009). Classifying maize inbred lines into heterotic groups using a factorial mating design. *Agronomy Journal.* 101: 102-106.

FAOSTATS- Agriculture. (2008). http://faostat.fao.org/default.aspx Accessed 10th June, 2009

Fernie, A.R., Tadmor, Y. and Zamir, D. (2006). Natural genetic variation for improving crop quality. *Current Opinion in Plant Biology*. 9: 196 – 202.

Food and Agricultural Organization of the United Nations. (2009). Grassland species profiles: *Zea mays L.* Available from:http://www.fao.org/ag/AGP/AGPC/doc/Gbase/DATA/Pf000342.htm.

Food and Agriculture Organization of the United Nations (FAO). (2002). Database. http://www.fao.org. Accessed 10th June, 2009

Franco, J., Crossa, J., Ribaut, J. M., Betran, J., Warburton, M. L. and Khairallah, M. (2001). A method for combining molecular markers and phenotypic attributes for classifying plant genotypes. *Theoretical and Applied Genetics.* 103: 944–952.

Fraser, P.D and Bramley, P.M. (2004). The biosynthesis and nutritional uses of carotenoids. *Progress in Lipid Research*. 43: 228–265.

Fraser, P.D., Kiano, J.W., Truesdale, M.R., Schuch, W. and Bramley, P.M. (1999). Phytoene synthase-2 enzyme activity in tomato does not contribute to carotenoid synthesis in ripening fruit. *Plant Molecular Biology*. 40: 687–698.

Fraser, P.D., Truesdale, M.R., Bird, C.R., Schuch, W. and Bramley, P.M. (1994). Carotenoid biosynthesis during tomato fruit development (evidence for tissue-specific gene expression). *Plant Physiology*. 105: 405–413.

Gallagher, C.E., Matthews, P.D., Li, F. and Wurtzel, E.T. (2004). Gene duplication in the carotenoid biosynthetic pathway preceded evolution of the grasses (Poaceae). *Plant Physiology*. 135:1776-1783.

Garcia, A.A.F., Benchimol, L.L., Barbosa, A.M.M., Geraldi, I.O., Souza, C.L. and de Souza, A.P. (2004). Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. *Genetics and Molecular Biology.* 27(4): 579-588.

Garcia-Mas, Oliver, J.M., Gomez-Paniagua, H. and de Vicente. M.C. (2000). Comparing AFLP, RAPD and RFLP markers for measuring genetic diversity in melon. *Theoretical and Applied Genetics.* 101: 860-864.

Gardiner, J. M., Coe, E. H., Melia-Hancock, S., Hoisington, D. A. and Chao, S. (1993). Development of a core RFLP map in maize using an immortalized F2 population. *Genetics*. 134 (3): 917–930.

Geleta, L.F., Labuschagne, M.T. and Viljoen, C.D. (2004). Relationship between heterosis and genetic distance based on morphological traits and AFLP markers in pepper. *Plant Breeding*. 123: 467-473.

Geleta, N., Labuschagne, M.T. and Viljoen, C.D. (2006). Genetic diversity analysis in sorghum germplasm as estimated by AFLP, SSR and morpho-agronomical markers. *Biodiversity and Conservation.* 15: 3251-3265.

Gerdes, J.T. and Tracy, W.F. (1994). Diversity of historically important sweet corn inbreds as estimated by RFLPs, morphology, isozymes and pedigree. *Crop Science.* 34: 26-33.

Gethi, J.G., Labate, J.A., Lamkey, K.R., Smith, M.E. and Kreesovich, S. (2002). SSR variation in important US maize inbred lines. *Crop Science.* 42: 951-957.

Giovannucci, E. (2002). A review of epidemiologic studies of tomatoes, lycopene, and prostate cancer. *Experimental Biology and Medicine (Maywood).* 227(10): 852–859.

Giuliano, G., Al-Babili, S. and von Lintig, J. (2003). Carotenoid oxygenases: cleave it or leave it. *Trends in Plant Science*. 8:145–149.

Goodman, M.M. (1990). Broadening the genetic diversity in breeding by use of exotic germplasm. pp. 139-148. *In:* J.G. Coors, S. Pandey (Eds.), Genetics and Exploitation of Heterosis in Crops. Crop Science Society of America, Madison, WI.

Gower, J.C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*. 53: 325–338.

Grogan, C. O and Blessin, C.W. (1968). Characterization of major carotenoids in yellow maize lines of differing pigment concentration. *Crop Science* .8: 730-732.

Gupta, P.K., Roy, J.K. and Prasad, M. (2001). Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Current Science.* 80(4): 524-535.

Guo, M., Rupe, M.A, Zinselmeier, C., Habben, J., Bowen . B.A. and Smith O.S. (2004). Allelic Variation of Gene Expression in Maize Hybrids. *The Plant Cell.* 16: 1707-1716.

Halperin, E. and Eskin, E. (2004). Haplotype reconstruction from genotype data using Imperfect Phylogeny. *Bioinformatics.* 20(12): 1842–1849.

Hamada, H., Petrino, M.G. and Kakunaga, T. (1982). A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proceedings of the National Academy of Sciences of USA.* 79: 6465-6469.

Hamblin, M.T., Warburton, M.L., Buckler, E.S. (2007). Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. *PLoS ONE* 2: e1367.

Hamrick, J.L. and Godt, M.J.W. (1997). Allozyme diversity in cultivated crops. *Crop Science*. 37: 26-30.

Harjes, C.E., Rocheford, T.R., Bai, L., Brutnell, T.P., Kandianis, C.B., Savinski, S.G., Stapleton, A.E, Vallabhaneni, R., Williams, M., Wurtzel, E.T., Yan, J., and Buckler, E.S. (2008). Natural Genetic Variation in Lycopene Tapped for Maize Biofortification. *Science*. 319: 330-333.

Hauge, M. and Trost. J.F. (1928). An inheritance study of the distribution of vitamin A in maize. *Journal of Biological Chemistry*. 80: 107-114.

Heckenberger, M., Bohn, M., Ziegle, J.S., Joe, L.K., Hauser, J., Hutton, M. and Melchinger, A. (2002). Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties I Genetic and technical sources of variation in SSR data. *Molecular Breeding.* 10: 81–191.

Heckenberger, M., van der Voor, T. J.R.., Melchinger, A.E., Peleman, J. and Bohn, M. (2003). Variation of DNA fingerprints amog accessions within maize inbred lines and implications for identification of essentially derived varieties: II Genetic and technical sources of variation in AFLP data and comparison with SSR data. *Molecular Breeding*. 12: 97-106.

Hirschberg, J. (2001). Carotenoid biosynthesis in flowering plants. *Current Opinion in Plant Biology.* 4: 210-218.

Howe, J.A. and Tanumihardjo, S.A. (2006). Carotenoid-Biofortified maize maintains adequate vitamin A Status in Mongolian Gerbils. *Journal of Nutrition* 136: 2562-2567.

Howitt, C.A. and Pogson, B.J. (2006). Carotenoid accumulation and function in seeds and non-green tissues. *Plant Cell Environment.* 29: 435–445.

Hoxha, S., Sharifluo, M.R. and Sharp. P. (2003). Evaluation of genetic diversity in Albanian maize using SSR markers. *Maydica*. 49: 97-103.

Innan, H., Terauchi, R., Kahl, G. and Tajima, F. (1999). A Method for Estimating Nucleotide Diversity From AFLP Data. *Genetics*. 151: 1157–1164.

Jaccard, P. (1908). Nouvelles recherches sur la distribution florale Bul Soc Vaudoise. *Science Nature.* 44:223–270.

Jambrovic, A., Šimic, D., Ledencan, T., Zdunic, Z. and Brkic, I. (2008): Genetic diversity among maize (*Zea mays, L.*) inbred lines in Eastern Croatia. *Period. Biol.* 110(3): 251–255.

Janssen, P., Coopman, R., Huys, G., Swings, J., Bleeker, M., Vos, P., Zabeau, M. and Kersters, K. (1996). Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. *Microbiology.* 142: 1881-1893.

Johnson, L.J., Meacham, S.L. and Kruskall, L.J. (2003). The antioxidants – vitamin C, vitamin E, selenium, and carotenoids. *Journal of Agromedicine.* 9: 65-80.

Jones, C.J., Edwards, K.J., Castaglione, S., Winfield, M.O., Sala, F., Van de Wiel, C., Bredemeijer, G., Vosman, B., Matthes, M., Maly, A., Brettschneider, R., Bettini, P., Buiatti, E. Maestri, M., Malcevschi, A., Marmiroli, N., Aert, R., Volckaert, G., Rueda, J., Linaacero, R., Vazque, A. and Karp, A. (1997). Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding* 3: 381-390.

Jung, M., Ching, A., Bhattramakki, D., Dolan, M., Tingey, S., Morgante, M. and Rafalski, A. (2004). Linkage disequilibrium and sequence diversity in a 500-kbp region around the adh1 locus in elite maize germplasm. *Theoretical and Applied Genetics.* 109**:** 681–689.

Kim, J. and DellaPenna, D. (2006). Defining the primary route for lutein synthesis in plants: the role of Arabidopsis carotenoid beta-ring hydroxylase CYP97A3. *Proceedings of the National Academy of Sciences of USA.* 103: 3474–3479.

Kurilich, A.C. and Juvik, J.A. (1999). Simultaneous quantification of carotenoids and tocopherol in corn kernels extracts by HPLC. *Journal Liquid Chromatography and Related Technology*. 19: 2925–2934.

Labate, J. A., Lamkey, K. R., Lee M. and Woodman, W. (2000). Hardy-Weinberg and linkage equilibrium estimates in the BSSS and BSCB1 random mated populations. *Maydica.* 45(3): 243-256.

Labate, J., Lamkey, K.R., Lee, M and Woodman. W.L. (1997). Molecular genetic diversity after reciprocal recurrent selection in BSSS and BSCB1 maize populations. *Crop Science.* 37:416–423.

Laborda, P.R., Oliveira, K.M., Garcia, A.A.F., Paterniani, M.E.A. and de Souza, A.P. (2005). Tropical maize germplasm: what can we say about its genetic diversity in the light of molecular markers. *Theoretical and Applied Genetics*. 111: 1288–1299.

Lartey, A. (2008). Maternal and child nutrition in Sub-Saharan Africa: challenges and interventions. Proceedings of the Nutrition Society. 67: 105–108.

Legesse, B.W., Myburg, A.A., Pixley, K.V. and Botha, A.M. (2007). Genetic diversity of African maize inbred lines revealed by SSR markers. *Hereditas*. 144: 10-17.

Li, F., Tzfadia, O. and Wurtzel, E.T. (2009). The *Phytoene Synthase* gene family in the Grasses: Subfunctionalization provides tissue-specific control of carotenogenesis. *Plant Signaling and Behavior*. 4(3): 208-211.

Li, F., Vallabhaneni, R., Yu, J., Rocheford, T. and Wurtzel, E.T. (2008). The maize phytoene synthase gene family: overlapping roles for carotenogenesis in endosperm, photomorphogenesis, and thermal stress-tolerance. *Plant Physiology.* 147: 1334-1346

Li, X.H., FU, J.H., Zhang, S.H.H., Yaun, L.X. and Li, M.S. (2000). Genetic variation of inbred lines of maize detected by SSR markers. *Science Agriculture Sinica.* 33: 1-9.

Li, Y., Shi, Y.S., Song, Y.C., Du, J.Y., Tuberosa, R. and Wang, T.Y. (2004). Analysis of genetic diversity in maize inbred lines based on AFLP markers. *Maydica*. 49(2): 89-95.

Li, Z.H., Matthews, P.D., Burr, B. and Wurtzel, E.T. (1996). Cloning and characterization of a maize cDNA encoding phytoene desaturase, an enzyme of the carotenoid biosynthetic pathway. *Plant Molecular Biology.* 30: 269-279.

Liu, K., Goodman, M., Muse, S., Smith, J. S., Buckler. Ed. and Doebley. J. (2003). Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics*. 165: 2117-2128.

Liu, S., Chen, H.D., Makarevitch, I., Shirmer, R, Emrich, S.J., Dietrich, C.R Barbazuk, W.B., Springer, N.M. and Schnable, P.S. (2010). High-throughput genetic mapping of mutants via quantitative single nucleotide polymorphism typing. *Genetics*. 184(1): 19–26.

Livini, C., Ajmone-Marsan, P., Melchinger, A.E., Messmer, M.M. and Motto, M. (1992). Genetic diversity of maize inbred lines within and among heterotic groups revealed by RFLPs. *Theoretical and Applied Genetics*. 84: 17-25.

Low, J.W., Arimond, M., Osman, N., Cunguara, B., Zano, F. and Tschirley, D. (2007). A food-based Approach Introducing orange-fleshed sweet potatoes increased vitamin A intake and serum retinol concentrations in young children in rural Mozambique *Journal of Nutrition.*137: 1320-1327.

Lu, H, and Bernardo, R. (2001). Molecular marker diversity among current and historical maize inbreds. *Theoritical and Applied Genetics.* 103: 613–617.

Lubberstedt, T., Melchinger, A.E., Duble, C., Vuylsteke, M. and Kuiper, M. (2000). Relationships among early European maize inbreds: IV Genetic diversity revealed with AFLP markers and comparison with RFLP, RAPD and pedigree data. *Crop Science.* 40: 783-791.

Luo, R. and Wurtzel, E.T. (1999). A maize cdNA encoding zeta carotene desaturase (Accession No. AF047490) (PGR99-118). *Plant Physiology* 120: 1205-1207

Mackay, T.F.C. (2001). The genetic architecture of quantitative traits. *Annual Review of Genetics*. 35: 303–339.

Mangelsdorf, P.C. and Fraps, G.S. (1931). A direct quantitative relationship between vitamin A in corn and the number of genes for yellow pigmentation. *Science*. 73: 241–242.

Manifesto, M.M., Schlatter, A.R., Hopp, H.E., Sua´rez, E.Y, and Dubcovsky, J. (2001). Quantitative evaluation of genetic diversity in wheat germplasm using molecular markers. *Crop Science*. 41: 682–690.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*. 27: 209-220.

Matthews, P.D, Luo, R. and Wurtzel, E.T. (2003). Maize phytoene desaturase and {zeta}-carotene desaturase catalyse a poly-Z desaturation pathway: implications for genetic engineering of carotenoid content among cereal crops. *Journal of Experimental Botany*. 54: 2215-2230.

Matthews, P.D. and Wurtzel, E.T. (2007). Biotechnology of food colorant production. *In* C Socaciu, ed, Food Colorants: Chemical and Functional Properties. CRC Press, Boca Raton, pp 347-398.

Mayer, J.E., Pfeiffer, W.H. and Beyer, P. (2008). Biofortified crops to alleviate micronutrient malnutrition. Curent Opinion in Plant Biology. 11: 166–170.

Maziya-Dixon, B., Kling, J.G., Menkir, A. and, Dixon, A. (2000). Genetic variation in total carotene, iron, and zinc contents of maize and cassava genotypes, and ascorbic acid in cassava roots. *Food and Nutrition Bulletin*. 21: 419–422.

Melchinger, A.E. (1999). Genetic diversity and heterosis In: JG Coors and S Pandey (eds) The Genetics and Exploitation of Heterosis in Crops. ASA, CSS, and SSSA, Madison, Wisconsin pp 99-118.

Melchinger, A.E., Lee, M., Lamkey, K.R. and Woodman, W.L. (1990). Genetic diversity for restriction fragment length polymorphisms: Relation to estimated genetic effects in maize inbreds. *Crop Science.* 30: 1033-1040.

Melchinger, A.E., Messmer, M.M., Lee, M., Woodman, W.L. and Lamkey, K.R. (1991). Diversity and relationships among US maize inbreds revealed by restriction length polymorphisms. *Crop Science*. 31: 669–678.

Menkir, A., Olowolafe, M. O., Ingelbrecht , I., Fawole, I., Badu-Apraku, B. and Vroh. B.I. (2006). Assessment of testcross performance and genetic diversity of yellow endosperm maize lines derived from adapted x exotic backcrosses. *Theoretical and Applied Genetics* 113(1): 90-99

Menkir, A., Kling J. G. , Badu-Apraku, B. and Ingelbrecht, I. (2005). Molecular marker-based genetic diversity assessment of *Striga*-resistant maize inbred lines. *Theoretical and Applied Genetics*. 110 (6)0040-5752

Menkir, A., Liu, W., White, W.S., Maziya-Dixon, B. and Rocheford, T. (2008). Carotenoid diversity in tropical-adapted yellow maize inbred lines. *Food Chemistry*. 109: 521 – 529.

Menkir, A., Melake-Berhan, A., Ingelbrecht, C.I. and Adepoju, A. (2004). Grouping of tropical mid-altitude maize inbred lines on the basis of yield data and molecular markers. *Theoretical and Applied Genetics* 108: 1582-1590.

Menz, M.A., Klein, R.R., Unruh, N.C., Rooney, W.L., Klein, P.E. and Mullet. J.E. (2004). Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. *Crop Science.* 44: 1236-1244.

Messmer, M.M., Melchinger, A.E., Boppenmaier, J., Herrmann, R. and Brunklaus-Jung, E. (1992). RFLP analyses of early-maturing European maize germplasm: I Genetic diversity among flint and dent inbreds. *Theoretical and Applied Genetics* 83: 1003–1012.

Messmer, M.M., Melchinger, A.E., Hermann, R. and Bopprnnmaier, J. (1993). Relationships among early European maize inbreds: II Comparison of pedigree and RFLP data. *Crop Sci*ence 33: 944-950.

Mohammadi, S.A and Prasanna, B.M. (2003). Analysis of Genetic Diversity in Crop Plants—Salient Statistical Tools and Considerations. *Crop Science*. 43: 1235-1248.

Mohammed, S.A., Prasanna, B.M., Sudan, C. and Singh, N.N. (2008). SSR heterogenic patterns of maize parental lines and prediction of hybrid performance. *Biotechnology and Biotechnology. 2*2: 541-547.

Moll, R.H., Lonquist, J.H., Foreuno, J.V., Johnson, E.C. (1965). The relationship of heterosis and genetic divergence in maize. *Genetics* 52: 139-144.

Montemurro, C., Simeone, R., Pasqualone, A., Ferrara, E. and Blanco, A. (2005). Genetic relationships and cultivar identification using AFLP and SSR markers. The Journal of Horticuitural Science and Biotechnology. 80: 105–110.

Morgante, M. and Olivieri, A.M. (1993). PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal*. 3(1): 175-182.

Morris, M.L. (2001). Assessing the benefits of international maize breeding research: An overview of the global maize impacts study. pp. 25-34. *In:* P.L. Pingali (Ed.), CIMMYT 1999-2000 world maize facts and trends. Meeting world maize needs: Technological opportunities and priorities for the public sector. CIMMYT, Mexico D.F., Mexico.

Mueller, U. G. and Wolfenbarger, L. L. (1999). AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution.* 14: 389–394

Muhammad, F.A., Muhammad, I., Muhammad, S.M., Malik, A.R., Muhammad, M. (2010). Assessment of genetic diversity among Pakistani wheat (*Triticum aestivum* L.) advanced breeding lines using RAPD and SDS-PAGE. *Electronic Journal of Biotechnology*. 13(3):

Naqvi, S., Zhu, C., Farre, G., Ramessar, K., Bassie, L., Breitenbach, J., Conesa, D.P., Ros, G., Sandmann, G., Capell, T. and Christou, P. (2009). Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *Proceedings of the National Academy of Sciences of USA*. 106(10): 7762–7767.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of USA*. 70: 3321-3323.

Nei, M. and Li W.H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of USA*.76: 5269–5273.

Nestel, P., Bouis, H.E., Meenakshi, J.V. and Pfeiffer, W. (2006). Biofortification of staple food crops. *Journal of Nutrition*. 136: 1064-1067.

Neumann, C.G. (2007). Food-based approaches to combating micronutrient deficiencies in children of developing countries. *Journal of Nutrition.* 137: 1091–1092

Oliveira, M.K., Laborda, P., Garcia, A.F., Me, Z.P. and de Souza, A.P. (2004). Evaluating genetic relationships between tropical maize inbred lines by means of AFLP profiling. *Hereditas*. 149: 24-33.

Park, H., Kreunen, S.S., Cuttriss, A.J., DellaPenna, D. and Pogson, B.J. (2002). Identification of the carotenoid isomerase provides insight into carotenoid biosynthesis, prolamellar body formation and photomorphogenesis. *The Plant Cell.* 14: 321–332.

Palaisa, K.A., Morgante, M., Williams, M. and Rafalski, A. (2003). Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. *Plant Cell*. 15: 1795-1806.

Pejic, I., Ajmone-Marsan, P., Morgante, M., Kovumplick, V., Castiglioni, P., Taramino, G., and Motto, M. (1998). Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. *Theoretical and Applied Genetics*. 97: 1248–1255.

Pfeiffer, W.H and McClafferty, B. (2007). HarvestPlus: breeding crops for better nutrition. *Crop Science*. 47: S88- S105.

Phelps, T.L., Hall, A.E. and Buckner. B. (1996). Microsatellite repeat variation within the y1 of maize and teosinate. *Journal of Heredity.* 87: 396-399.

Pinto, L.R., Vieira, M.L.C., de Souza, C.L. and de Souza, A.P. (2003). Genetic-diversity assessed by microsatellites in tropical maize populations submitted to a high-intensity reciprocal recurrent selection. *Euphytica*. 134: 277-286.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. and Rafalski, A. (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding*. 2(3): 225-238.

Pressoir, G., Brown, P.J., Zhu, W., Upadyayula, N., Rocheford, T., Buckler, E.S.and Kresovich1, S. (2009). Natural variation in maize architecture is mediated by allelic differences at the PINOID co-ortholog barren inflorescence. *The Plant Journal.* 58: 618–628.

Quackenbush, F.W., Firch, J.G., Brunson, A.M. and House, L.R. (1963). Carotenoid, oil, and tocopherol content of corn inbreds. *Cereal Chemistry*. 40:250–259.

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

Ramakrishnan, A., Meyer, S.E., Waters, J., Stevens. M.R., Coleman, C.E, and Fairbanks, D. J. (2004). Correlation between molecular markers and adaptively significant genetic variation in *Bromus tectorum* (Poaceaae), an inbreeding annual grass. *American Journal of Botany*. 91(6): 797-803.

Ranatunga, M.A.B., Meenakshisundaram, P., Arumugachamy, S. and Maheswaran, M. (2009). Genetic diversity analysis of maize (*Zea mays L*) inbreds determined with morphometric traits and simple sequence repeat markers. *Maydica*. 54: 113-121.

Rao, A.V. and Agarwal. S. (2000). Role of Antioxidant Lycopene in Cancer and Heart Disease. *Journal of the American College of Nutrition.* 19: 563-569.

Reif, J.C., Melchinger, A.E., Xia, X.C., Warburton, M.L., Hoisington, S.A., Vasal, S.K., Srinivasan, G., Bohn, M. and Frisch, M. (2003a). Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. *Crop Science*. 43: 1275-1282.

Reif, J.C., Melchinger, A.E., Xia, X.C., Warburton, M.L., Hoisington, D.A., Vasal, S.K., Beck, D., Bohn, M. and Frisch, M. (2003b). Use of SSRs for establishing heterotic groups in subtropical maize. *Theoretical and Applied Geneics*. 107: 947-957.

Reif, J. C., Melchinger A.E. and Frisch, M. (2005). Genetical and mathematical properties of similarity and dissimilarity coefficients. *Crop science*. 45: 1–7.

Reif, J.C, Hamrit, S. and Melchinger, A.E. (2006). Genetic Diversity Trends in Central European heterotic groups. *Acta Agronomica Hungarica*. 54 (3): 315-320.

Remington, D.L., Thornsberry, J.M., Matsuoka, Y., Wilson, L.M., Whitt, S.R., Doebely, J., Kresovich, S., Goodman, M.M., Buckler, E.S IV. (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences of USA*. 98: 11479–84.

Ribaut, J.M. and Ragot, M. (2007). Marker-assisted selection to improve drought adaptations in maize: the backcross approach, perspectives, limitations, and alternatives. *American Journal of Botany*. 58: 351–360.

Rogers, J.S. (1972). Measures of genetic similarity and genetic distance. pp 145–153 *In* Studies in Genetics VII Publ 7213 Univ of Texas, Austin.

Rohlf, F.J. (1997). NTSYSpc, Numerical taxonomy and multivariate analysis system, Version 202. Applied Biostatistics Inc, New York.

Rohlf, F.J. and Fisher, D.R. (1968). Tests for hierarchical structure in random data sets. *Systematic Zoology*. 17: 407-412.

Romero-Serveson, J., Smith, J.S.C., Ziege, J., Hauser, J. and Hooksta, G. (2001). Pedigree analysis and haplotype sharing with divers groups of *Zea mays L* inbreds. *Theoretical and Applied Genetics.* 103:567-574.

Russell, J.R., Fuller, J.D., Macaulay, M., Hatz, B.G., Jahoor, A. Powell, W. and Waugh, R. (1997). Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, SSRs, and RAPD. *Theoretical and Applied Genetics.* 95: 714-722.

SAS Institute (1999) SAS/STA T user's guide, version 8.

Schut, J.W., Qi, X. and Stam, P. (1997). Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. *Theoretical and Applied Genetics.* 95: 1161-1168.

Semba, R.D. and Dagnelie, G. (2003). Are lutein and zeaxanthin conditionally essential nutrients for eye health? *Medical Hypotheses.* 61: 465-472.

Senior, M.L., Murphy, J.P., Goodman, M.M. and Stuber C.W. (1998). Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. *Crop Science*. 38: 1088-1098.

Simpson, K. (1983). Relative value of carotenoids as precursors of vitamin A. *Proceedings of the Nutrition Society.* 42: 7-17.

Singh, M., Lewis, P.E., Hardeman, K., Rose. J.K.C., Azourek. M.., Chomt, P. and Brutnell.T.P. (2003). Activator mutagenesis of the pink scutellum/viviparous7 locus of maize. *Plant Cell.* 15(4): 874-884.

Smith, J.S.C. and Smith, O.S. (1989). The description and assessment of distances between inbred lines of maize: The utility of morphological, biochemical and genetic descriptors and a scheme for the testing of distinctiveness between inbred lines. *Maydica*. 34: 151-161.

Smith, J.S.C. and Smith. O.S. (1992). Fingerprinting crop varieties. *Advances in Agronomy*. 47: 85 -140.

Smith, J.S.C., Chin, E.C.L., Shu, H., Smith, O.S., Wall, S.J., Senior, M.L., Mitchell, S.E., Kresovich, S. and Ziegle, J. (1997). An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L): Comparisons with data from RFLPs and pedigree. *Theoretical and Applied Genetics*. 95: 163–173.

Smith, J.S.C., Smith, O.S., Bowen, S.L., Tenborg, R.A. and Wall S.J. (1991). The description and assessment of distances between inbred lines of maize III A revised scheme for the testing of distinctiveness between inbred lines utilizing DNA RFLPs. *Maydica*. 36:213–226.

Smith, O.S., Smith, J.S.C., Bowen, S.L., Tenborg, R.A. and Walls, S.J. (1990). Similarities among a group of elite maize inbreds as measured by pedigree, F1 grain yield, grain yield heterosis, and RFLPs. *Theoretical and Applied Genetics*. 80: 833-840.

Sneath, P.H.A. and Sokal, R.R. (1973). Numerical taxonomy. Freeman, San Francisco.

Sommer, A. (1993). Vitamin A, infectious disease, and childhood mortality: a 2 solution? *Journal of Infectious Diseases.* 167: 1003–1007.

Sommer, A. (2008). Vitamin A Deficiency and Clinical Disease: An Historical Overview. *Journal of Nutrition*. 138: 1835–1839.

Sommer, A. and Davidson, F.R. (2002). Assessment and Control of Vitamin A Deficiency: The Annecy Accords. *Journal of Nutrition*. 132: 2845S–2850S.

Sommer, A., Hussaini, G., Muhila, H., Tarwotjo, I., Susanto, D. and Saroso, J.S. (1980). History of night blindness: a simple tool for xerophthalmia screening. *The American Journal of Clinical Nutrition*. 33: 887-891.

Souza, I.R. P., Schuelter, A.R., Guimara, C.T., Schuster, I., Oliveira, E. and Medinbaugh, E. (2008). Clustering of QTL conferring SCMV resistance in tropical maize. *Hereditas*. **000**: 1-7.

Syvänen, A.C. (2001). Assessing genetic variation: genotyping SNPs. *Nature Review Genetics.* 12: 930-942.

Taramino, G. and Tingey, S. (1996). Simple sequence repeats for germplasm analysis and mapping in maize. *Genome.* 39: 277-287.

Tenaillon, M., Sawkins, M., Anderson, L., Stack, S., Doebley, J., Gaut, B. (2002). Patterns of diversity and recombination along chromosome 1 of maize (*Zea mays* ssp *mays* L.). *Genetics*. 162: 1401-1413.

Thornsberry, J. M., Goodman, M. M., Doebley, Kresovitch, J. F., Nielson .S. and Buckler, IV E. S. (2001). *Dwarf8* polymorphisms associate with variation in flowering time. *Nature Genetics.* 28: 286-289.

Tian, L. and DellaPenna, D. (2004). Progress in understanding the origin and functions of carotenoid hydroxylases in plants. *Archives of Biochemistry and Biophysics*. 430: 22–29.

Tian, L., DellaPenna, D., Zeevaart, J.A.D. (2004). Effect of hydroxylated carotenoid deficiency on ABA accumulation in Arabidopsis. *Plant Physiology.* 122: 314-320.

Toenniessen, G.H. (2002). Crop genetic improvement for enhanced human nutrition. *Journal of Nutrition*. 132: 2493S-2946S.

UNICEF, (2009). Vitamin A deficiency Child info-Monitoring the situation of children and women.

Unnevehr, L., Pray, C. and Paarlberg, R. (2007). Addressing micronutient deficiencies: alternative interventions and technologies. *Agro-Biotechnology Forum* .10(3): 124-134.

Vallabhaneni, R. and Wurtzel, E.T. (2009). Timing and biosynthetic potential for carotenoid accumulation in genetically diverse germplasm of maize. *Plant Physiology*. 150: 562-572.

Vallabhaneni, R., Gallagher, C.E., Licciardello, N., Cuttriss, A.J., Quinlan, R.E. and Wurtzel, E.T. (2009). Metabolite sorting of a germplasm collection reveals the*hydroxylase 3* locus as a new target for maize pro-vitamin A biofortification *Plant Physiology*. 151: 1635–1645.

Varshney, R.K., Graner, A. and Sorrells, M.E. (2005). Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology.* 23: 48-55.

Vigouroux, Y., Jaqueth, S. J., Matsuoka, Y, Smith, O. S., Beavis, J.S.C and Doebley, J. (2002). Rate and pattern of mutation at microsatellite loci in maize. *Molecular Biology and Evolution.* 19: 1251-1260.

Vigouroux, Y., Mitchell, S., Matsuoka, Y., Hamblin, M., Kresovich, S., Smith, J.S.C., Jaqueth, J., Smith, O.S. and Doebley, J. (2005) An analysis of genetic diversity across the maize genome using microsatellites. *Genetics*. 169: 1617-1630.

Vos, P., Hogers, R., Bleeker, M., Reeijans, M., va de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Juiper, M. and Zabeau, M. (1995). AFLP: a new technique for DNA finger printing. *Nucleic Acids Research.* 23: 4407- 4414.

Vroh-Bi, I., McMullen, M.D., Sanchez, Villeda, H., Schroeder, S., Gardiner, J., Polacco, M., Soderlund, C., Wing, R., Fang, Z. and Coe, E.H. Jr. (2005). Single Nucleotide Polymorphisms and Insertion–Deletions for Genetic Markers and Anchoring the Maize Fingerprint Contig Physical Map. *Crop Science*. 46: 12–21.

Vuylsteke, M., Mank, R., Antonise, R., Bastiaans, E., M. L. Senior, Stuber, C.W., Melchinger, A.E., Lubberstedt, T., Xia, X.C., Stam, P., Zabeau, M. and Luiper. M. (1999). Two high-density AFLP linkage maps of *Zea mays L.*: analysis of distribution of AFLP markers. *Theoretical and Applied Genetics* 99: 921-935.

Wang, L., Guan, R., Zhangxiong, L., Chang, R. and Qiu L. (2006). Genetic Diversity of Chinese Cultivated Soybean Revealed by SSR Markers. *Crop Science*. 46: 1032–1038.

Wang, L.X., Guan, R.X., Liu, Z.X., Chang, R.Z. and Qiu, L.J. (2006). Genetic diversity of Chinese cultivated soybean revealed by SSR markers. *Crop Science*. 46: 1032–1038.

Warburton, M.L., Xia, X.C., Crossa, J., Franco, J., Melchinger, A.E., Frisch, M., Bohn, M. and Hoisington, D. (2002). Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. *Crop Science*. 42: 1832-1840.

Warburton, M.L. (2005). Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers: II Subtropical, tropical midaltitude, and highland maize inbred lines and their relationships with elite US and European maize. *Crop Science.* 45: 2573-2582.

Weber, E.J. (1987). Carotenoids and tocols of corn grain determined by HPLC. *Journal of the American Oil Chemists' Society*. 47: 337–339.

Wei, K., Zhang, H., Xu, X., Du, H., Huang, Y., and Zhang, Z. (2009). Evaluation of phenotype and genetic diversity of maize landraces from Hubei province, Southwest China. *Frontiers of Agriculture in China*. 394: 374-382.

Weir, B.S. (1996). Genetic data analysis II Sinauer Press, Sunderland, Massachusetts.

Weir, B.S. and Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*. 38:1358–1370.

Weising, K., Winter, P. and Huttel, P.B. (1998). Microsatellite markers for molecular breeding. *Journal of Crop Production* 1: 113-143.

Welsh, J. and McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research.* 18: 7213-7218.

West, C.E, Eilander, A. and van Lieshout, M. (2002). Consequences of revised estimnates of carotenoid bioefficacy for dietary control of vitamin A deficiency in developing countries. *Journal of Nutrition*. 132:2920S–2926S.

West, K.P. Jr. (2002). Extent of vitamin A deficiency among preschool children and women of reproductive age. *Journal of Nutrition.* 132 (9 suppl): 2857S–66S.

Whitt, S. R., Wilson, L. M., Tenaillon, M. I., Gaut, B. S and Buckler, IV E. S. (2002).Genetic diversity and selection in the maize starch pathway. *Proceedings of the National Academy of Sciences of USA.* 99*(20):12959-12962*

WHO, (2003). Micronutrient Deficiency: Combating Vitamin A Deficiency. World Health Organization, Geneva.

WHO, (2009). World Health Organization Global Database on Vitamin A Deficiency. 2009. Global prevalence of vitamin A deficiency in populations at risk 1995–2005. Available from: http://www.who.int/vmnis/vitamina/en/

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey. S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research.* 18: 6531-6535.

World Health Organisation (2002). Quantifiying selected major risks to health: In Reducing Risks, Promoting Healthy Life Geneva, Switzerland Pp 47-49.

World Health Organization. (1995). Global Prevalence of Vitamin A Deficiency. Micronutrient Deficiency Information System Working Paper No 2 WHO/NUT/953 World Health Organization, Geneva.

Wong, J.C., R.J. Lambert, E.T. Wurtzel, and T.R. Rocheford. (2004). QTL and candidate genes phytoene synthase and β-carotene desaturase associated with the accumulation of carotenoids in maize. *Theoretical and Applied Genetics.* 108: 349–359.

Wright, A.J. (1971). The analysis and prediction of some two factor interactions in grass breeding. *Journal of Agricultural Science*. 76: 301-306.

Wright, S. (1943). Isolation by distance. *Genetics.* 28: 114-138
.

Xia, X.C., Reif, J.C., Hoisington, D.A., Melchinger, A.E., Frisch, M. and Warburton, M.L. (2004). Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers: I Lowland tropical maize. *Crop Science.* 44: 2230–2237.

Xia, X.C., Reif, J.C., Melchinger, A.E., Frisch, M., Hoisington, D.A., Beck, D., Pixley, K. and Warburton, M.L. (2005). Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers: II Subtropical, tropical midaltitude, and highland maize inbred lines and their relationships with elite US and European maize. *Crop Science*. 45: 2573-2582.

Xu, S., Liu, J., and Liu, G. (2004). The use of SSRs for predicting the hybrid yield and yield heterosis in 15 key inbred lines of Chinese maize. *Hereditas.* 141: 207-215.

Xu, Y. and Crouch, J. H. (2008). Marker-assisted selection in plant breeding: from publications to practice. *Crop Science.* 48(2): 391–407.

Yamasaki, M., Tenaillon, M.I., Bi, I.V., Schroeder, S.G., Sanchez-Villeda, H., Doebley, J.F., Gaut, B.S. and McMullen, M.D. (2005). A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *Plant Cell*. 17: 2859-2872.

Yan, J., Shah, T., Warburton, M.L., Buckler, E.S., McMullen, M.D. Crouch, J. (2009). Genetic Characterization and Linkage Disequilibrium Estimation of a Global Maize Collection Using SNP Markers. *PLoS ONE* 4: e8451

Yan, J., Kandianis, C.B., Harjes, C.E., Bai, L., Kim, E-H., Yang, X., Skinner, D.J., Fu, Z., Mitchell, S., Li, Q., Fernandez, M.G.S., Zaharieva, M., Babu, .R., Fu, Y., Palacios, N., Li. J., DellaPenna, D., Brutnell, T., Buckler, E.S., Warburton, M.L., Rocheford. T. (2010). Rare genetic variation at *Zea mays crtRB1* increases β-carotene in maize grain. *Nature Genetics.* 42: 322–327.

Yang, G.P., Saghai Maroof, M.A., Xu, C.G., Zhang, Q. and Biyashev, R.M. (1994). Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. *Molecular Genetics and Genomics.* 245: 187-194.

Yang, W., de Oliveira.A.C., Godwin.I., Schertz.K., and Bennetzen, J.L. (1996). Comparison of DNA Marker Technologies in Characterizing Plant Genome Diversity: Variability in Chinese Sorghums. *Crop Science* 36: 1669-1676.

Ye, X., Al-Babili, S., Kloti, A., Zhang, J., Lucca, P., Beyer, P. and Potrykus, I. (2000). Engineering the pro-vitamin A (β-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science.* 287: 303-305.

Yuan, Y; SanMiguel, P; Bennetzen, J, 2003. High-Cot sequence analysis of the maize genome. *Plant Journal*. 34: 249-255.

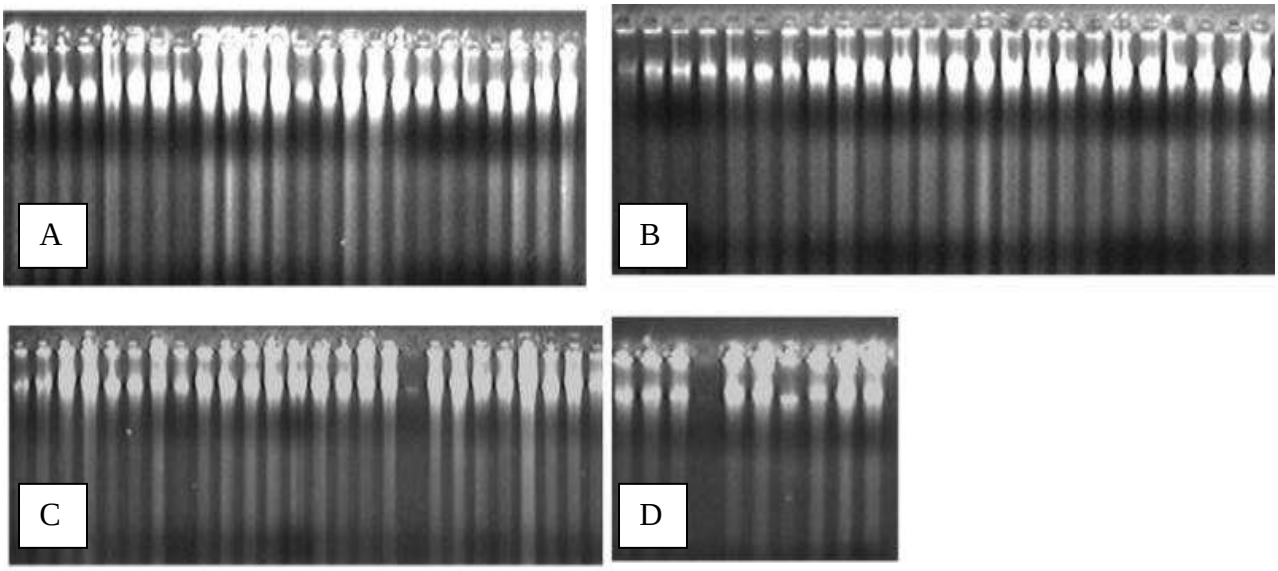
Yuan, L.X., Fu, J.H., Warburton, M., Li, X.H., Zhang, S.H. and Khairallah, M. (2000). Comparison of genetic diversity among maize inbred lines based on RFLPs, SSRs, AFLPs and RAPDs. *Yi Chuan Xue Bao.* 27: 725–733.

Zhang, Q., Saghai Maroof, M.A., Lu, T.Y.and Shen, B.Z. (1992). Genetic diversity and differentiation of indica and japonica rice detected by RFLP analysis. *Theoretical and Applied Genetics*. 83: 495-499.

Zhang, D.X. and Godfrey, M.H. (2003). Nuclear DNA analyses in genetic studies of populations: Practice, problems and prospects. *Molecular Ecology*.12: 563-584.

http://www.cimmyt.org/Research/Economics/map/facts_trends/maizeft9900/pdfs/maizeft9900.pdf
http://www.maizegdb.org
http://nucleus.agron.missouri.edu/cgi-bin/ssr_bin.pl
www.panzea.org
http://faostat.fao.org)

Appendix I



Photographs of extracted DNA for some maize inbred lines showing high quality.
A- Maize inbred lines, 1-24 from 38, B to D- Maize inbred lines, 1-58 from 122.

Appendix II

Polyacrylamide gel preparation:

40 % Acrylamide solution (19:1)

500 ml of 40 % acrylamide solution; 190 g of acrylamide and 10 g bisacrylamide were added in distilled water to make a final volume of 500 ml.

6 % Acrylamide solution

500 ml of 6 % of acrylamide; 75 ml of 40 % acrylamide solution, 210 g of urea, 50 ml of 10X TBE were added in distilled water, heated and to make a final volume of 500 ml. The solution was filtered using a filter paper.

Preparation of staining solution

Silver staining solution was prepared by adding 2 g of silver nitrate powder in 2 L of distilled water and 3 ml formaldehyde (40 % solution) and mix with stirrer. The solution was stored in a dark cupboard until ready for use.

Preparation of developing solution

Sodium carbonate (60 g) was dissolved in 2 L of distilled water and 300 µl of sodium thiosulphate solution (0.1001 N) and 3 ml of formaldehyde (40 % solution) were also added to prepare the developing solution. The solution was refrigerated at 4 $^{\circ}$C (pre-cooling of the sodium carbonate also greatly slows down the rate of development).

Preparation of fixing solution

The 10 % acetic acid fixing solution was prepared by adding 200 ml of glacial acetic acid into 1.8 L of distilled water.

Appendix III

***LCYE* and *critRB1* alleles obtained in 38 yellow endosperm maize inbred lines evaluated for single nucleotide polymorphism using allele specific primers**

Inbreds	*LCYE*	*LCYE*	*LCYE*	*CritRB1*	*CritRB1*	*CritRB1*	*HYD3*
	3' IN/DEL	SNP 216	5' INDEl/TE	Del 4	3'TE	5'InDEl/TE	TSS
PVL01	399+502	G	null	117+350	296+1800	800	163
PVL02	399+502	G	null	117+350	296+1800	800	163
PVL03	399+502	G	null	117+350	296+1800	800	163
PVL04	399+502	G	150+280+1700	117+350	296+1800	800	163
PVL05	502	G	null	117+350	296+1800	800	163
PVL06	502+144	G	null	117+350	296+1800	800	163
PVL07	399+502	G	150+280+1700	117+350	296+1800	800	163
PVL08	399+502	G	150+280+1700	117+350	296+1800	800	163
PVL09	399+502+144	G	null	117+350	296+1800	800	163
PVL10	399+502	G	null	117+350	296+1800	800	163
PVL11	399+502	G	null	117+350	296+1800	800	163
PVL12	399+502	G	null	117+350	296+1800	800	163
PVL13	502+144	G	null	117+350	296+1800	800	163
PVL14	502+144	G	null	117+350	296+1800	800	163
PVL15	502	G	null	117+350	296+1800	800	163
PVL16	502	T	150+280+1700	117+350	296+1800	800	163
PVL17	502	G	150+280+1700	117+350	296+1800	800	163
PVL18	502+144	G	null	117+350	296+1800	800	163
PVL19	399+502	G	null	117+350	296+1800	800	163
PVL20	399+502	G	null	117+350	296+1800	800	163
PVL21	502+144	G	null	117+350	296+1800	800	163
PVL22	502+144	G	null	117+350	296+1800	800	163
PVL23	502+144	G	null	117+350	296+1800	800	163
PVL24	502+144	G	null	117+350	296+1800	800	163
PVL25	502+144	G	null	117+350	296+1800	800	163
PVL26	502+144	G	null	117+350	296+1800	800	163
PVL27	502+144	G	null	117+350	296+1800	800	163
PVL28	399+502	G	null	117+350	296+1800	800	163
PVL29	399+502	G	null	117+350	296+1800	800	163
PVL30	399+502	G	150+280+1700	117+350	296+1800	800	163
PVL31	399+502	G	null	117+350	296+1800	800	163
PVL32	399+502	G	null	117+350	296+1800	800	163
PVL33	502	G	150+280+1700	117+350	296+1800	800	163
PVL34	399+502	G	null	117+350	296+1800	800	163
PVL35	399+502	G	null	117+350	296+1800	800	163
PVL36	502	G	150+280+1700	117+350	296+1800	800	163
PVL37	502	G	150+280+1700	117+350	296+1800	800	163
PVL38	399+502+144	G	null	117+350	296+1800	800	163

LCYE and *critRB1* alleles obtained in 122 yellow endosperm maize inbred lines evaluated for single nucleotide polymorphism using allele specific primers

Inbred	*LCYE* 3' IN/DEL	*LCYE* SNP 216	*LCYE* 5' InDEl/TE	*critRB1* 3' TE	*critRB1* Del 4	*critRB1* 5'INDEl/TE	*HYD3* TTS
PVL01	144+502+399	G	280	296+1221+1800	117+350	800	163
PVL02	144+502+399	G	280+1700+150	296 +1221	117+350	800	163+608
PVL03	144+502	G	280	296+1221+1800	117+350	800	163+608
PVL04	144+502	G	0	296+1221+1800	117+350	800	163
PVL05	144+502	G	0	296 +1221	117+350	800	163
PVL06	144+502	G	0	296+1221+1800	117+350	800	163
PVL07	144+502	G	280	296+1221+1800	117+350	800	163
PVL08	144+502	G	0	296+1221+1800	117+350	800	163
PVL09	144+502	G	0	543	117+350	600	163
PVL10	144+502+399	G	0	296+ 875+1221	117+350	800	163+608
PVL11	144+502	G	0	296+1221+1800	117+350	800	163+608
PVL12	144+502	G	0	296+1221+1800	117+350	800	163+608
PVL13	144+502	G	0	296+1221+1800	117+350	800	163
PVL14	144+502	G	0	296+1221+1800	117+350	800	163
PVL15	144+502	G	0	296+1221+1800	117+350	800	163
PVL16	144+502	G	0	296 +1221	117+350	800	163
PVL17	144+502	G	0	296+1221+1800	117+350	800	163
PVL18	144+502	G	0	296 +1800	117+350	800	163
PVL19	144+502	G	0	296	117+350	0	163
PVL20	144+502	G	0	296+1221+1800	117+350	0	163
PVL21	144+502+399	G	0	296+1221+1800	117+350	800	163
PVL22	144+502	G	0	296+1221+1800	117+350	800	163
PVL23	144+502	G	0	296+1221+1800	117+350	800	163
PVL24	144+502+399	G	0	296 +1800	117+350	800	163
PVL25	144+502+399	G	0	296 +1800	117+350	800	163
PVL26	144+502+399	G	250+380	543+296	117+350	800+600	163+608
PVL27	144+502+399	G	250+380	543	117+350	600	163
PVL28	144+502	G	0	296	117+350	800	163
PVL29	144+502+399	G	250+380	543	117+350	600	163
PVL30	144+502	G	0	543+296	117+350	600	163
PVL31	144+502	G	0	543+296	117+350	600	163
PVL32	144+502	G	0	543	117+350	600	163
PVL33	144+502	G	0	296 +1800	117+350	800	163
PVL34	144+502+399	G	250+380	543+296	117+350	600	163+608
PVL35	144+502	G	0	296+1221+1800	117+350	600	163+608
PVL36	144+502	G	0	543+296	117+350	600	163+608

LCYE and *critRB1* alleles obtained in 122 yellow endosperm maize inbred lines evaluated for single nucleotide polymorphism using allele specific primers

Inbred	*LCYE*	*LCYE*	*LCYE*	*critRB1*	*critRB1*	*critRB1*	*HYD3*
	3' IN/DEL	SNP 216	5' InDEl/TE	3' TE	Del 4	5'INDEl/TE	TTS
PVL37	144+502+399	G	250+380	543	117+350	600	163
PVL38	144+502	G	0	543	117+350	600	163
PVL39	144+502+399	G	250+380	543	117+350	600	163
PVL40	144+502	G	0	543+296+1222	117+350	800+600	163
PVL41	144+502+399	G	250+380	296+1221+1800	117+350	800	163+608
PVL42	144+502+399	G	250+380	543	117+350	600	163
PVL43	144+502+399	G	250+380	296+1221+1800	117+350	800	163
PVL44	144+502+399	G	250+380	296+1221+1800	117+350	800	163
PVL45	144+502+399	G	250+380	543	117+350	600	163
PVL46	144+502	G	0	543	117+350	600	163
PVL47	144+502+399	G	250+380	543	117+350	600	163
PVL48	144+502+399	G	250+380	543+1221	117+350	800	163+608
PVL49	144+502	G	0	296+1221+1800	117+350	800	163
PVL50	144+502	G	0	296+1221	117+350	800	163
PVL51	144+502	G	0	296+1221+1800	117+350	800	163
PVL52	144+502	G	0	296+1221	117+350	800	163
PVL53	144+502	G	0	296+1221+1800	117+350	800	163
PVL54	144+502	G	0	296+1221+1800	117+350	800	163
PVL55	144+502	G	0	296 +1221	117+350	800	163
PVL56	144+502	0	0	296 +1221	117+350	800	163
PVL57	144+502	G	0	296+1221+1800	117+350	800	163
PVL58	144+502	G	0	296+1221+1800	117+350	800	163
PVL59	144+502	G	0	296+1221+1800	117+350	800	163
PVL60	144+502+399	G	0	296+1221+1800	117+350	800	163
PVL61	144+502+399	G	0	296+1221+1800	117+350	800	163
PVL62	144+502	G	0	296+1221+1800	117+350	800	163
PVL63	144+502	G	0	296 +1221	117+350	800	163
PVL64	144+502+399	G	0	296+1221+1800	117+350	800	163
PVL65	144+502	G	0	296+1221+1800	117+350	800	163
PVL66	144+502	G	0	296 +1221	117+350	800	163
PVL67	144+502	G	0	296 +1221	117+350	800	163
PVL68	144+502	G	0	296 +1221	117+350	800	163
PVL69	144+502	G	0	296+1221+1800	117+350	800	163
PVL70	144+502	0	0	296 +1221	117+350	800	163
PVL71	144+502	G	0	296 +1221	117+350	800	163
PVL72	144+502	G	0	296 +1221	117+350	800	163

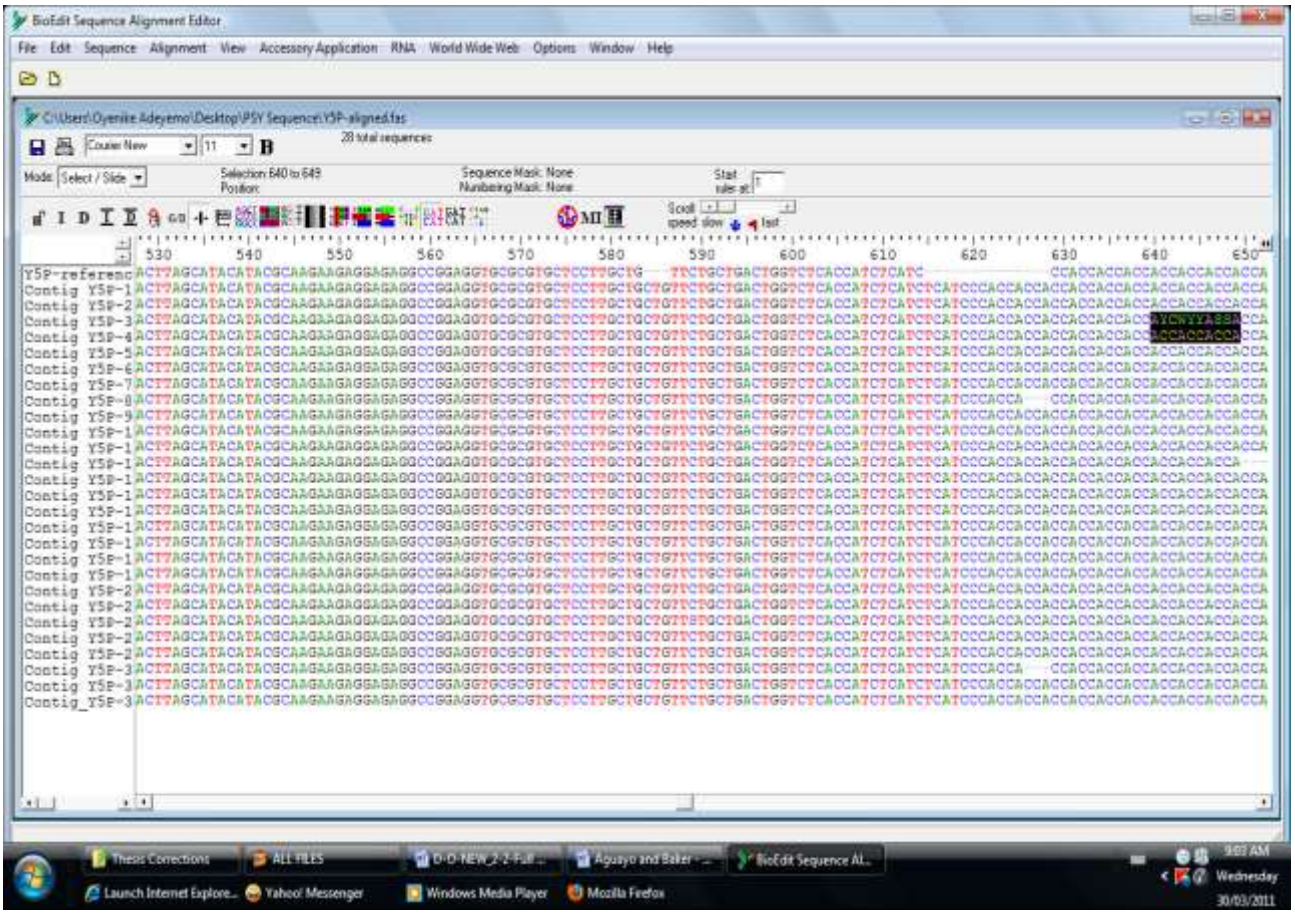
LCYE and *critRB1* alleles obtained in 122 yellow endosperm maize inbred lines evaluated for single nucleotide polymorphism using allele specific primers

Inbreds	***LCYE***	***LCYE***	***LCYE***	***critRB1***	***critRB1***	***critRB1***	***HYD3***
	3' IN/DEL	**SNP 216**	**5' InDEl/TE**	**3' TE**	**Del 4**	**5'INDEl/TE**	**TTS**
PVL73	144+502	G	0	296 +1221	117+350	800	163
PVL74	144+502	G	0	296 +1221	117+350	800	163
PVL75	144+502	G	0	296 +1221	117+350	800	163
PVL76	144+502	G	0	296	117+350	0	163
PVL77	144+502	G	0	296	117+350	800	163
PVL78	144+502	G	0	296 +1221	117+350	0	163
PVL79	144+502	G	0	296 +1221	117+350	800	163
PVL80	144+502	G	0	296 +1221	117+350	800	163
PVL81	144+502	G	0	296	117+350	0	163
PVL82	144+502	G	0	296	117+350	0	163+608
PVL83	144+502	G	0	296	117+350	0	163+608
PVL84	144+502	G	0	296 +1221	117+350	0	163
PVL85	144+502	G	0	296 +1221	117+350	800	163
PVL86	144+502	G	0	296 +1221	117+350	800	163
PVL87	144+502	G	0	296 +1221	117+350	0	163
PVL88	144+502	G	0	296	117+350	0	163
PVL89	144+502	G	0	296 +1221	117+350	800	163+608
PVL90	144+502	G	0	296 +1221	117+350	800	163+608
PVL91	144+502	G	0	296 +1221	117+350	0	163+608
PVL92	144+502	G	0	296 +1221	117+350	0	163
PVL93	144+502	G	0	296 +1221	117+350	800	163
PVL94	144+502+399	G	0	296	117+350	0	163
PVL95	144+502+399	G	**280+1700+150**	296 +1221	117+350	0	163
PVL96	144+502	G	280	296	117+350	0	163
PVL97	144+502+399	G	**280+1700+150**	296	117+350	0	163+608
PVL98	144+502	G	0	296	117+350	0	163
PVL99	144+502+399	G	0	296	117+350	0	163
PVL100	144+502+399	G	**280+1700+150**	296	117+350	0	163
PVL101	144+502+399	G	280+1700+250+150	296	117+350	0	163
PVL102	144+502+399	G	0	296 +1221	117+350	0	163
PVL103	144+502	G	0	296	117+350	0	163
PVL104	144+502+399	G	0	296 +1221	129+350	0	163
PVL105	144+502+399	G	250+380	543	117+350	0	163
PVL106	144+502+399	G	250+380	296 +1221	117+350	0	163+608
PVL107	144+502+399	G	0	543+296+1222	117+350	0	163
PVL108	144+502+399	G	250+380	296 +1221	117+350	0	163

LCYE and *critRB1* alleles obtained in 122 yellow endosperm maize inbred lines evaluated for single nucleotide polymorphism using allele specific primers

Inbreds	***LCYE***	***LCYE***	***LCYE***	***critRB1***	***critRB1***	***critRB1***	***HYD3***
	3' IN/DEL	**SNP 216**	**5' InDEl/TE**	**3' TE**	**Del 4**	**5'INDEl/TE**	**TTS**
PVL109	144+502+399	G	0	296 +1221	117+350	0	163
PVL110	144+502+399	G	0	296 +1221	117+350	0	163
PVL111	144+502	G	0	543+296	117+350	0	163
PVL112	144+502+399	G	0	296 +1221	117+350	0	163
PVL113	144+502	G	0	543	117+350	0	163
PVL114	144+502	G	0	543	117+350	600	163
PVL115	144+502	G	250+380	296	117+350	0	163
PVL116	144+502+399	G	250+380	296 +1221	117+350	0	163
PVL117	144+502	G	0	296 +1221	117+350	0	163+608
PVL118	144+502+399	G	0	543+296+1221	117+350	0	163
PVL119	144+502	G	0	543+296+1222	117+350	0	163
PVL120	144+502	G	0	296 +1221	117+350	800	163
PVL121	144+502	G	0	296	117+350	800	163+608
PVL122	144+502+399	G	0	296	117+350	0	163

Appendix IV



PSY 1 Sequence nucleotides for primers Y5P-R and Y5P-F primers for 27 of 38 yellow endosperm tropical adapted maize inbred lines.

Highlighted are some heterozygotes

Legend:
M = A/C
Y = C/T
R = A/G
W = A/T
S = C/G
K = G/T