COLEGIO DE POSTGRADUADOS

INSTITUCION DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO DE RECURSOS GENÉTICOS Y PRODUCTIVIDAD

GENÉTICA

DESARROLLO Y VALIDACIÓN DE MARCADORES MOLECULARES PARA IDENTIFICAR GENOTIPOS DE MAíZ DE CALIDAD PROTEíNICA

MARTHA HERNÁNDEZ RODRÍGUEZ

T E S I S

PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:

DOCTORA EN CIENCIAS

MONTECILLO, TEXCOCO, EDO. DE MEXICO

2013

La presente tesis titulada: **Desarrollo y validación de marcadores moleculares para identificar genotipos de maíz de calidad proteínica,** realizada por la alumna: **Martha Hernández Rodríguez** bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

DOCTORA EN CIENCIAS
RECURSOS GENÉTICOS Y PRODUCTIVIDAD

GENETICA

CONSEJO PARTICULAR

CONSEJERO
DR. J. JESÚS GARCÍA ZAVALA

DIRECTOR DE TESIS
DR. RAMAN BABU

ASESOR
DR. FERNANDO CASTILLO GONZÁLEZ

ASESOR
DR. AMALIO SANTACRUZ VARELA

ASESOR
DR. RICARDO LOBATO ORTIZ

Montecillo, Texcoco, Estado de México, Diciembre de 2013

DESARROLLO Y VALIDACIÓN DE MARCADORES MOLECULARES PARA IDENTIFICAR GENOTIPOS DE MAÍZ DE CALIDAD PROTEÍNICA

Martha Hernández Rodríguez, D. C.

Colegio de Postgraduados, 2013

RESUMEN

El maíz (*Zea mays* L.) de calidad proteínica (QPM) es nutricionalmente superior en comparación con el maíz normal, porque produce semillas con casi el doble de los aminoácidos (aa) lisina (Lys) y triptófano (Trp), esenciales para la síntesis endógena de la vitamina B niacina. Actualmente, la conversión de una línea de maíz normal a una QPM, se acelera mediante la selección asistida por marcadores moleculares (MAS) y métodos de detección química. En este estudio se reporta el desarrollo y validación de marcadores basados en polimorfismos de un solo nucleótido (SNPs) para los loci *opaque-2* (*o2*) y *Ask-2,* involucrados en la genética de los QPM, así como la interacción de ambos loci para el contenido de 11 aa en el endospermo de dos poblaciones F_2 de maíz, derivadas de las cruzas CML172 x CML323 y CLQRCWQ26 x La Posta Sequía. Se identificó un SNP (C>T) en el primer exón del gen *o2* después de analizar las secuencias de dos fragmentos genómicos de 16 líneas de maíz. Este SNP fue consistente en las líneas QPM examinadas, pero fue diferente en las líneas normales. Se encontró que este SNP fue 88.3% asertivo para discriminar los genotipos QPM de los no QPM en 88 líneas genotipificadas con tecnología KASP™. Por otra parte, se desarrolló un marcador basado en PCR para un SNP (A>T) reportado en la región 3´del locus *Ask-2*. Este nuevo marcador permitió explorar la variación natural de este gen en 82 líneas de maíz, empleando un nuevo procedimiento SSCP reportado también aquí. Cinco variantes SSCP se pudieron visualizar; una de ellas, identificó a cuatro líneas QPM con los contenidos más altos de Trp. Finalmente, la interacción de los loci *o2* y *Ask-2* mostró que el gen *o2* en forma recesiva incrementó la concentración de los aa en las tres clases genotípicas de *Ask-2* en la población CML172 x CML323.

Palabras clave: *Zea mays*, *opaque-2*, *Ask-2*, MAS, QPM, SNP, SSCP.

DEVELOPMENT AND VALIDATION OF MOLECULAR MARKERS TO IDENTIFY QUALITY PROTEIN MAIZE GENOTYPES

Martha Hernández Rodríguez, D. C.

Colegio de Postgraduados, 2013

ABSTRACT

Quality maize (*Zea mays* L.) protein (QPM) is nutritionally superior in comparison to normal maize, since it produces kernels with nearly twice the levels of lysine (Lys) and tryptophan (Trp), which are essentials amino acids (aa) for endogen synthesis of B vitamin niacin. Currently, the conversion of a normal line of maize into a QPM one has been accelerated by means of marker-assisted selection (MAS) and chemical detection methods. In this study we report the development and validation of single nucleotide polymorphism (SNP) markers for both *opaque-2* (*o2*) and *Ask-2* loci, involved in the genetic basis of QPM, we also report the interaction between both loci for the content of 11 aa in the endosperm of two F_2 maize populations derived from the CML172 x CML323 and CLQRCWQ26 x La Posta Sequía crosses. A SNP (C>T) in the first exon of *o2* was identified after analyzing sequences of two genomic fragments of 16 maize inbred lines. This SNP was consistent in all examined QMP lines, but it was different in the normal maize lines. After genotyping 88 maize inbred lines with KASP™ technology, it was found that this SNP was 88.3% assertive to discriminate QPM genotypes and non QPM genotypes. On the other hand, a PCR-based marker was developed from a SNP (A>T) reported in the 3´region of *Ask-2* locus. This novel marker allowed exploring the natural variation of this gen in a set of 82 maize inbred lines through a new DNA-single-strand conformational polymorphism (SSCP) procedure, also reported here. Five SSCP variants were visualized; one these variants identified four QPM lines with the highest Trp levels. Finally, the interaction between *o2* and *Ask-2* loci showed that *o2* in a recessive form increased the aa level in all three genotypic classes of *Ask-2* in the CML172 x CML323 population.

Index words: *Zea mays*, *opaque-2*, *Ask-2*, MAS, QPM, SNP, SSCP.

AGRADECIMIENTOS

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico brindado para realizar mis estudios; así como al Colegio de Postgraduados por formarme a través de los excelentes cursos que recibí de su personal académico y por las facilidades otorgadas por su personal administrativo, siempre cordial y respetuoso; agradezco también al Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT) por la flexibilidad otorgada durante los cursos del postgrado y la realización de esta investigación en sus instalaciones.

Muy especialmente, extiendo mi agradecimiento al Dr. Yunbi Xu, quien me impulsó para iniciar este camino y me respaldó generosamente en todo momento y nunca perdió la confianza en mí. También agradezco al Dr. Raman Babu por ser la parte esencial de este logro, aceptarme como su estudiante y encontrar en él siempre disposición y confianza para dirigirme durante la investigación. En el mismo sentido, impagable me resulta la confianza, paciencia y consejo del Dr. J. Jesús García Zavala, quien con una actitud siempre optimista fue incesante para que culminara este asunto. Estos atributos también los encontré en el Dr. Ricardo Lobato, quien además de la cuidadosa revisión que hizo de la tesis, le agradezco su especial insistencia para que terminara este asunto. Son también inestimables la disposición y facilidades que me brindó el Dr. Fernando Castillo, quien con sus valiosas lecciones sobre Genética y su opinión crítica, me ayudó a reconocer un nuevo universo y trabajar pensando en él. Para el Dr. Amalio Santacruz, excelente profesor de referencia, le agradezco la enorme confianza que siempre tuvo en mí, su apoyo total desde el inicio de este largo proceso

fue en extremo valioso. En deuda también estoy con el Dr. Adrián Quero y la Dra. Leonor Miranda, quienes fomentaron mi ingreso al Colegio de Postgraduados y me apoyaron en una las etapas emocionales más difíciles de este proceso. Asimismo, con orgullo extiendo mi gratitud al Dr. José D. Molina Galán, quien además de ser mi sinodal, se convirtió en mi amigo. Sus charlas sobre el México de ayer son inolvidables.

Para mi esposo Hugo Álvarez Flores y mis hijos Hugo Alejandro, Alicia Mariana y Andrea Natalia, agradezco enormemente su comprensión y paciencia para tolerar mi ausencia, no solamente en casa, sino también dentro de ella; realmente ustedes son el mejor equipo de trabajo que he tenido. Asimismo, también dirijo mi agradecimiento a mis queridos padres Alicia y Alberto, así como mi cercana familia paterna y a mis queridos sobrinos, particularmente a Juan Jesús, por hacerme sentir su apoyo para continuar adelante a pesar de la distancia.

A mis amigos y compañeros del Colegio de Postgraduados: Azucena Ibarra, José Luis Coyac, Dolores Briones, Victoria Estrada y Darío Flores, gracias por crear un ambiente cordial durante las jornadas de estudio. Por su apoyo constante, agradecida también estoy con mis amigos y compañeros del Laboratorio de Biotecnología del CIMMYT, sobre todo con Virginia García (†), Pablo Alva, Bacilisa Luna, Carlos Martínez y Raqueline de la Rosa. Finalmente, al Dr. Gregorio Alvarado y al M. en C. Aurelio Hernández les dirijo un reconocimiento especial por su apoyo técnico, al igual que a mis amigas Margarita Crosby, Gricelda Palacios y Marcelina Vélez, quienes con sus estimulantes pláticas y bromas hicieron este camino más accesible.

DEDICATORIA

Para Hugo,

Hugo Alejandro, Alicia Mariana y

Andrea Natalia:

mi aliento, mi familia. Los amo.

TABLE OF CONTENTS

RESUMEN iii

ABSTRACT iv

LIST OF TABLES x

LIST OF FIGURES xi

I. INTRODUCTION 1

1.1. General objective 6

1.2. Particular objectives 6

II. LITERATURE REVIEW 7

2.1. Role of the maize in the human nutrition 8

2.2. The kernel of maize 10

2.3. Quality protein maize for enhancing the human nutrition 11

2.4. The *opaque-2* gene 14

2.5. The aspartate kinase-2 (AK) enzyme 16

2.6. SSCP: a gel-based mutation scanning method 17

III. MATERIALS AND METHODS 21

3.1. Plant material and DNA extraction 21

3.2. Development of PCR-based markers 21

3.2.1. SNP assays for the *opaque-2* alleles 21

3.2.2. SNP assays for the *Ask-2* alleles 24

3.3. Development of a PCR-SSCP assay of moderate throughput 24

3.4. Interaction analysis between *opaque-2* and *Ask-2* genotypes 26

3.4.1. *opaque-2* and *Ask-2* genotyping 27

3.4.2. Amino acid content analysis 28

IV. RESULTS AND DISCUSSION 29

4.1. Analysis of the *opaque-2* gene sequence 29

4.2. High throughput assays for *opaque2* and *Ask-2* genes 32

4.3. SSCP analysis for allele mining of the *Ask-2* gene 39

4.4. Analysis of the free amino acid content (FAA) 41

4.5. Interaction between *opaque-2* and *Ask-2* loci 44

V. CONCLUSIONS 55

VI. REFERENCES 56

VII. APPENDIX 68

LIST OF TABLES

Page

Table 1. Maize inbred lines examined to identify single nucleotide polymorphisms and small insertion/deletions into genomic sequences of the *opaque-2* locus. 22

Table 2. Sequence and melting temperature for primers designed to amplify two genomic fragments of the *opaque-2* locus. 24

Table 3. Single nucleotide polymorphism typing for two loci involved in the genetics of QPM, *opaque-2* and *Ask-2*, using the KASP™ chemistry in a panel of QPM lines and normal maize lines. 37

Table 4. Single nucleotide polymorphisms for *opaque-2* and *Ask-2* loci, SSCP variants for *Ask-2* gene and total content of free amino acid (nmol g^{-1} FW) of ten QPM lines (italics) and three normal maize lines. 42

Table 5. Single nucleotide polymorphisms for *opaque-2* and *Ask-2* loci, SSCP variants for *Ask-2* gene and relative values of free amino acid of ten QPM lines (italics) and three normal maize lines. 43

Table 6. Total content for free amino acid (nmol g^{-1} FW) of nine genotypic classes of the F_2 of the cross CML172 x CML323. 46

LIST OF FIGURES

		Page
Figure 1.	Genomic sequence alignment of an *opaque-2* fragment showing three genetic polymorphisms.	30
Figure 2.	Small insertion and deletions (InDel) detected in the sequence of an *opaque-2* PCR fragment of seven QPM lines (in italics, on top) and eight non QPM lines.	31
Figure 3.	Sequence alignment of an *opaque-2* fragment showing a potentially discriminative SNP present in all QPM lines (C, in italics, on top) but different in non QPM lines (T).	33
Figure 4.	Cluster of SNP alleles for *opaque-2* locus in several QPM lines and normal maize lines using KASP™ chemistry.	34
Figure 5.	Cluster of SNP alleles for *Ask-2* locus in several QPM lines and normal maize lines using KASP™ chemistry.	36
Figure 6.	Gel image of PCR-SSCP variants for *Ask-2* locus in several QPM lines.	40
Figure 7.	Biplot of first two principal components showing the proportion of variation derived from the amino acid content in the three genotypes of *opaque-2* gene.	48

Figure 8. Biplot of first two principal components showing the proportion of variation derived from the amino acid content in the three genotypes of *Ask-2* gene. 49

Figure 9. Interaction between *opaque-2* and *Ask-2* loci for the accumulation of Trp (left) and Lys (right) residues in the maize F_2 population CML172 x CML323. 51

Figure 10. Interaction of *opaque-2* and *Ask-2* loci for free amino acid content in seeds of maize F_2 population CML172 x CML323. 52

I. INTRODUCTION

Maize (*Zea mays* L.) is one of the three cereals with the major production of food and agricultural commodities in the world (http://www.faostat.org). Maize is not only used for human consumption, for livestock feed, for biofuel production and other by-products, but also as an organism model exhaustively researched in genetics. As food crop, maize is widely consumed in developing countries of Africa, Latin America and Asia, where accounts for 25% of the total demand. As a staple food, its consumption reaches up to 73% in several countries of Sub-saharian Africa (Shiferaw *et al*., 2011). In these regions, maize is the main source of carbohydrates and vegetable protein, contributing with over ≥20% of total daily calories (Krivanek *et al.,* 2007). This contribution as calories is variable across regions, ranging for an average of 32% of consumed calories in Eastern and Southern Africa up to 51% in Lesotho, Malawi and Zambia (Nuss and Tanumihardjo, 2011; Cairns *et al*., 2012). Since calories provided from both carbohydrates and proteins are similar in terms of their contribution to the body, the food-dependence on this crop can have health implications once maize is taken as the major source of protein intake and other protein foods are scarce (Vivek *et al*., 2008).

Maize provides carbohydrates, vitamins and minerals; however, is deficient in some essential amino acids and provitamin A and in micronutrients such as iron, zinc and calcium whose bioavailability is significantly lowered due to chelators as phytate indigestible by monogastric animals, including humans (Sun *et al*., 2009). Humans can

synthesize almost all the organic compounds needed for normal physiological activity, but a small number of specific molecules (fatty acids, amino acids and minerals) are required strictly in the diet (Pérez- Massot *et al*., 2013). Several reports have shown that in the most maize-dependent countries, where the access to a diversified diet is limited, the malnutrition is a prevalent problem and disorders as pellagra and anemia usually occur (FAO, 2001). Both disorders are consequence of a low intake of tryptophan and iron, respectively. Several strategies such as to increase dietary diversification, mineral supplementation and food fortification have been implemented to alleviate the malnutrition problem; however, these strategies have not always been successful in the developing countries due to cost and availability of food processing facilities, crop seasonality and infrastructure of the communities for distributing the food on time, and often they are also difficult to sustain (Vivek *et al*., 2008; Atlin *et al*., 2012, Pérez-Massot *et al*., 2013). Biofortification has been proposed as a sustainable strategy to enhance, through genetic breeding, agronomic management or genetic engineering, the nutritional value of the edible parts of staple crops as maize (Bouis and Welch, 2010). In those maize-dependent communities, this crop is a candidate for lysine (Lys) and tryptophan (Trp) biofortification. Other nutrients, revealed to be in low amounts in the maize kernel, could also be enhanced using this strategy and it has been proposed that biofortified maize cultivars could help in reducing the protein malnutrition problem in the rural environments and promote food security (Pérez-Massot *et al*., 2013).

Overall, the maize kernel contains of 8-10% of protein in the endosperm. This poor nutritional level of maize is associated with the main storage proteins in the maize grain, termed also zeins, which account for 50-70% of the total protein in the endosperm and

also influence in the texture and hardness of the grain (Holding and Larkins, 2009). Zeins are devoid in lysine (Lys) and triptophan (Trp) residues. Both amino acids are essentials for endogen synthesis of vitamin B niacin, a molecule that both humans and monogastric animals (such as pigs, poultry, and fish) depend largely on their provision on plants (Gerdes *et al*., 2012). Niacin leads to the formation of the nicotinamide adenine dinucleotide (NAD) and its 2'-phosphate ester (NADP), coenzymes involved in many redox reactions. Niacin (nicotinic acid) deficiency classically results in pellagra. Thus, a proper consumption of food rich in protein is crucial for normal human physiological activity (FAO 2001).

Eighteen mutants have been reported to alter the zein synthesis and give an opaque phenotype to the endosperm (Gibbon and Larkins, 2005). The most studied mutant is the natural mutant *opaque-2* discovered by Singleton in 1920 (Singleton, 1939). The mutant is an alteration of the *Opaque-2* locus which is located in the short arm of maize chromosome 7 and encodes a regulatory protein with a leucine zipper DNA-binding motif (Schmidt *et al*., 1990). When *opaque-2* mutant is present in the endosperm as *o2/o2/o2*, 22-kDa α-zein fraction decreases. This reduction favors the increase of other storage proteins (albumins, globulins and glutelins) richer in Trp and Lys residues (Balconi *et al*., 2007).This is the main reason why *opaque-2* maize mutants have 95% of the nutritional value of milk, an effect recognized by Mertz *et al.* since 1964. Beneficial effects of the *opaque-2* gene motivated its introgression in maize giving rise to the opaque maize. However, along with an increase in the nutritional value of the maize kernel, pleiotropic effects also occurred. The maize opaque kernels have a

chalky and lightweight endosperm, susceptible to be infested in storage conditions together with a poor agronomic performance. These issues were overcome with the development of the Quality Protein Maize (QPM) by researchers of International Maize and Wheat Improvement Center (CIMMYT) in a endeavor which began in the early 1970´s (Vasal, 2001; Gupta *et al*., 2012). QPM cultivars have been developed by researchers of CIMMYT following conventional breeding schemes; however, the advent of new approaches, combining conventional plant breeding and biotechnology tools, could help to accelerate the QPM breeding process (Babu *et al*., 2005).

The breeding of QPM involves the manipulation of three distinct genetic systems: i) the recessive mutant allele of the *opaque-2* gene, ii) *opaque-2*-modifiers that alter the soft and chalky appearance of the kernel, and iii) non *opaque-2* amino acid modifiers/enhancers influencing free amino acid (FAA) in the endosperm (Atlin *et al*., 2011). Currently, both marker assisted selection (MAS) and biochemical detection methods are being applied in the breeding programs to speed the QPM conversion (Babu *et al.,* 2005; Vivek *et al*., 2008; Gupta *et al*., 2012). Three SSR markers (umc1066, phi057 and phi112) located within the *opaque-2* gene are generally used for MAS. All three SSRs markers (phi057 and phi112 developed at Pioneer Hibred (Chin *et al*., 1996) and umc1066 (developed at Missouri Maize Project, University of Missouri, Columbia) give amplification products of about 140-160 base pairs (bp). None of these markers present a diagnostic fragment size for the recessive *opaque-2* allele compared to the dominant *Opaque-2* allele. Therefore, none of these markers will be useful if the DNA fragment amplified from the *opaque-2* and *Opaque-2* carriers are of same size.

CIMMYT´s experience shows that polymorphism detected with phi057 is very tight and needs polyacrylamide gels for proper resolution; umc1066 shows larger fragment sizes than phi057 and it can be resolved on 3-4% agarose gels in most cases; phi112 tends to be a dominant marker, that detects a null allele (no amplification product), usually, but not necessarily in the *opaque-2* sources. This marker is a bit tricky in that it will not be useful in all segregating populations such as back-crosses, but can be useful in an F_2 population if the null allele is the one of *opaque-2.* For MAS, co-dominant markers are more informative and useful than dominant markers, because co-dominant markers are capable of differentiating homozygous individual plants from heterozygous plants. Furthermore, for dominant markers such as phi112, no amplification product may also indicate that the amplification did not occur and not necessarily that it is a null product. Therefore, to make sure that the null allele of phi112 is due to genetic causes and not by PCR issues, it is convenient to perform multiplexing PCR using a housekeeping gene as internal standard of the PCR. All these operational drawbacks for each of these three microsatellites markers have restricted their routine use in the QPM breeding. In addition, it is frequently observed by breeders of CIMMYT that the simple presence of *opaque-2* in recessive form does not always ensure higher Lys and Trp levels, indicating a possible involvement of other amino acid modifier genes. Thus, to increase the efficiency of the QPM breeding process, it would be relevant to identify functional and more discriminative molecular markers that could be used to identify *opaque-2* genotypes by allowing their selection via MAS.

1.1. General objective

Develop and validate molecular markers to identify quality protein maize genotypes.

1.2. Particular objectives

i) Identify SNPs into *opaque-2* and *Ask2* sequences that can be used as discriminative SNP markers to distinguish QMP genotypes

ii) Genotype these SNPs using high throughput genotyping assays

iii) Develop and validate PCR-based SNP markers for both *opaque-2* and *Ask2* loci as well as a low tech method for their detection

iv) Analyze the interaction between *opaque-2* and *ask2* loci with respect to free amino acid content in grains of two segregating F_2 maize populations.

II. LITERATURE REVIEW

Maize (*Zea mays* ssp*. mays.)* is the representative cereal of America (Serratos, 2009) and it is one of the most important crops in the world together with rice and wheat. Maize was domesticated from teosinte (*Zea mays* spp *parviglumis*) ~9000 years ago (Matsuoka *et al.*, 2002) and was dispersed across Americas from the Balsas region (Kato *et al*., 2009). Maize belongs to the C4 grasses; it has high photosynthetic efficiency and is the model organism most exhaustively researched in plant domestication, genome evolution, developmental physiology, epigenetics, pest resistance, heterosis, quantitative inheritance, and comparative genomics (Strable and Scanlon, 2009). As cross-pollinating specie with prominent reproductive structures, maize has developed a high level of genetic diversity that has contributed to its broad morphological variability and geographical adaptation (Xu *et al.*, 2009). This diversity has been handled and conserved in the landraces by peasants, through its cultivation and selection in local environments (Castillo-Nonato and Chávez-Mejía, 2013). Maize diversity can be used to broad the genetic base of narrow breeding pools (Goodman, 1999) and contribute to higher yield, abiotic stress tolerance, disease resistance or nutritional quality improvement.

With the foundation of the modern heterosis concept from the last century (1952, by Shull), maize diversity has been exploited in the modern breeding programs to generate inbred lines, hybrids and open pollinated populations adapted to specific agroecological zones. As consequence, significant advances in the yield production have been reached (Goodman, 1999). Grain yield average has been increasing from

~2.0 ton ha^{-1} in the 1950s up to ~9.0 ton ha^{-1} in the 2000s in the United States (Lamkey and Edwards, 1999). At present, the world grain yield average was of 4.9 ton ha^{-1}, being the United States of America the primary producer with 7.74 ton ha-1. Mexico has a record of 3.18 ton ha^{-1} (FAOSTAT, 2013).

2.1. Role of the maize in the human nutrition

Maize is mainly grown for feed and food production through the world, but some specialty corns are utilized to generate byproducts for industry and fuel energy (Xu *et al.*, 2009). This crop constitutes an important source for the subsistence of communities in the rural environment, especially in large areas of Mesoamerica and sub-Saharan Africa, where the annual maize consumption ranks 80 to 130 kg *per capita* and comes up to 125 kg *per capita* in México (Atlin *et al.,* 2012; Castillo-Nonato and Chávez-Mejía, 2013). Together with wheat and rice, maize provides 30% of the food calories to more than 4.5 billion people in 94 developing countries (Shiferaw *et al*., 2011). In Eastern and Southern Africa, maize has become a food of sustenance accounting for an average of 32% of consumed calories, rising to 51% in Lesotho, Malawi and Zambia, exceeding to Mesoamerican countries where the crop had its origin (Nuss and Tanumihardjo, 2011; Cairns *et al*., 2012).

Besides provide carbohydrates, the maize kernel also supplies other micronutrients important for metabolic functions; however, maize kernel is protein deficient due to the main storage proteins of the endosperm, the prolamins or zeins, are devoid in Lys and Trp residues. Both Lys and Trp are essentials amino acids for

endogen synthesis of niacin, a type of B vitamin, which is the core molecule for nicotinamide adenine dinucleotide (NAD) and its phosphate (NADP). Both molecules are coenzymes involved in hundreds of redox reactions (Gerdes *et al*., 2012). NAD, for example, functions in intracellular respiration and with enzymes involved in the oxidation of fuel substrates such as glyceraldehyde 3-phosphate, lactate, alcohol, 3-hydroxybutyrate, and pyruvate and plays crucial roles in a variety of biological processes related to energy metabolism, aging, and calcium homeostasis (FAO, 2011; Ying, 2013). NAD is also required for non-redox adenosine diphosphate–ribose transfer reactions involved in DNA repair (Berger, 1985) and it is also a potent cytoprotective agent (Ying, 2013). NADP functions in reductive biosynthesis such as fatty acid and steroid syntheses and in the oxidation of glucose-6-phosphate to ribose-5-phosphate in the pentose phosphate pathway (FAO, 2001). All these roles for NAD and NADP molecules suggest the vital importance for a proper intake of Lys and Trp residues.

In terms of human nutrition and health, dietary deficiencies of maize (essential amino acids, and other nutrients, such as iron, zinc and calcium) can be supplied with the consumption of animal products, fruits, vegetables and legumes. However, very often in the marginal areas the access to this kind of products is limited (FAO, 2001). Human poor quality diets are characterized by high consumption of energy-staple foods and low consumption of protein-food. A protein deficient diet causes malnutrition (FAO, 2001; Vivek *et al*., 2008). A disorder caused by a protein-energy imbalance is the kwashiorkor characterized by initial growth failure, irritability, swollen abdomens, listlessness, hair change color and fatty liver. Kwashiorkor also increases susceptibility

to tuberculosis and gastroenteritis and sometimes is called the “weaning disease” because occurs with the dietary shift from breast milk to soft cereal foods in young children (Vivek *et al*., 2008; Nuss and Tanumihardjo, 2011). Communities that depend in maize as the main staple food are on risk to suffer protein malnutrition, a situation that reflects the combined impact of poverty, poor access to food, inefficient food distribution infrastructure, and an over-reliance on subsistence mono-agriculture (Pérez-Massot *et al*., 2013). These conditions have motivated that global initiatives promote strategies to deal the malnutrition problem, not only for maize but also for the main staple crops that depend the world´s poor (FAO 2001). These strategies are of three categories: a) to increase the diversity of food intake (diversification), b) to add nutrients to the diet artificially, either by providing supplements (supplementation) or by the fortification of basic food products, and c) to apply the biofortification, a strategy by which crops are modified, via conventional plant breeding or engineering, or treated to accumulate additional nutrients at source through agronomic management (Pérez-Massot *et al*., 2013). In this sense, the application of genomic, genetics and molecular tools, such as MAS in conventional breeding, could accelerate the generation of maize cultivars with a better nutritional quality.

2.2. The kernel of maize

The grain of maize is the most valuable part of the maize plant due to it is the main component of the yield and is the plant structure used for feed and food consumption. Botanically, the grain of maize is a single-seeded fruit, called caryopsis, that accumulates photosynthates and other assimilates during reproductive stage. A typical

grain or kernel of maize is composed by endosperm, pericarp and embryo. Each contributes with about 70-90%, 1.1% and 5-6% of the kernel-dry matter, respectively (Motto *et al*., 2011). In the endosperm, the principal storage proteins are prolamins (highly hydrophobic and soluble in alcoholic solutions or denaturing solvents), globulins (soluble in diluted saline solutions), albumins (soluble in pure water) and glutelins (soluble in acid or alkali). Prolamins o zeins constitute approximately 70% of the endosperm proteins, whereas glutelins account 34% and other minor proteins 3% each (Holding and Larkins, 2009). Zeins were also classified by Esen (1987) in four classes according to their structure and evolutionary relationships (α-, β-, γ-, and δ-zeins). α-zeins (19 and 22-kDa) contributes with 60% to 70% of total zeins, β-zeins (15 kDa) with 5% to 10% of total zeins, γ-zeins (16-, 27-, and 50-kDa) with 20% to 25% of total zeins, and δ -zeins (10- and 18-kDa) with <5% of total zeins), all they are encoded by distinct classes of structural genes (Holding and Larkins, 2009). The zein proteins are rich in glutamine, proline, alanine, and leucine residues and deficient in charged amino acids, especially Lys and Trp. As consequence, it has been reiteratively postulated that this deficiency is the major reason of the poor protein quality of maize and the main focus for the breeding of QPM cultivars (Vasal, 2001).

2.3. Quality protein maize for enhancing the human nutrition

As mentioned previously, several approaches have been implemented to alleviate the malnutrition (food fortification, diet diversification and supplemetation); however, their coverage and dissemination are limited by the infrastructure of the communities, resources, crop seasonality, among others (Vivek *et al*., 2008; Atlin *et al*., 2012; Pérez-

Massot *et al*., 2013). To convert a staple food in a more nutritious cultivar, locally adaptable and cost-effective, and face the malnutrition, a sustainable approach is the biofortification (FAO, 2001, Bouis and Welch, 2010; Pérez-Massot *et al*., 2013). Some plant breeders have already realized on this strategy to ameliorate the quality of the maize kernel in maize-dependent communities. Two researchers, Dr. S. K. Vasal and Dr. Evangelina Villegas, focused their work in developing the Quality Protein Maize (QPM) at the International Maize and Wheat Improvement Center (CIMMYT) in a long journey that began in the early 1970´s. This achievement, besides to diminish the child malnutrition in those maize-dependent communities, led them to share the 2000 World Food Prize. QPM cultivars have the great potential in alleviating protein malnutrition in people dependent on maize for their sustenance. At present, CIMMYT breeders have released 59 QPM lines together with some QPM hybrids and open pollination varieties in parts of Asia, Central America, and Africa (Vasal 2001; Atlin *et al*., 2012).

QPM cultivars have the same grain appearance and quantity of protein with respect to common maize. The main difference is in the Lys and Trp content of the endosperm tissue, which is the nearly double of normal maize cultivars (Vivek *et al*., 2008). This higher concentration of both amino acids in the grain endosperm duplicates its biological value (amount of N retained in the body for its different metabolic functions) as it was reported by Bressani (1992). The causative gene of this increment, and also of the beneficial effects of the opaque maize genotypes, is the natural mutant *opaque-2*. When *opaque-2* is present in the endosperm as *o2/o2/o2,* opaque maize protein rises to 95% of the nutritional value of milk compared of that of 45% of normal maize protein. Such effect was recognized by Mertz and co-workers since 1964

(Prassana *et al*., 2001). Besides increasing the Lys and Trp content, *opaque-2* also entailed inconvenient effects as soft and starchy appearance to the endosperm, features which are associated with low seed density, brittleness, pest susceptibility and inferior agronomic traits. In a large endeavor carried out by the maize breeding program of CIMMYT, these pleiotropic effects associated to *opaque-2* were corrected by the selection of quantitative trait loci, called *opaque-2* modifiers (*mo2*), which confer a hard and vitreous texture to the endosperm (Vasal, 2001).

Endosperm texture modification is another important and complex requisite in the conversion of normal maize to QPM genotype. Kernel vitreousness is another trait for the successful adoption of the QPM genotypes. The kernel modification is a trait quantitatively inherited and several reports have indicated preponderance of additive gene action in kernel modification. Hossain *et al*. (2008) analyzed in detail different attributes of kernel modification (endosperm modification, crown opaqueness and ear appearance) in QPM inbred lines and a set of experimental crosses (7 x 7 full diallel). Their diallel analysis indicated almost equal contribution of additive and non-additive effects for endosperm modification and predominance of non-additive gene effects on crown modification and ear appearance. Additional studies conducted to associate genes to this trait found disperse quantitative trait loci (QTL). One of QTLs has been linked to the 27-kDa γ-zein locus on chromosome 7S. Using RNAi technology, Wu *et al*. (2010) eliminate the expression of both 27- and 16-kDa γ-zein disrupting endosperm modifications. They introduced a dominant RNAi transgene into a QPM line (CM105Mo2) to eliminate expression of both 27- and 16-kDa γ-zein. They found

abnormalities in protein body structure and their interaction with starch granules in the F_1 with Mo2/+; o2/o2; γRNAi/+ genotype suggesting that γ-zeins are essential for restoring protein body density in the endosperm.

Opaque-2 (*o2*) gene is not the only one that can affect the texture of the maize endosperm. Eighteen mutants altering the timing and the rate of zein synthesis have been described and also positioned on the maize genetic map (Gibbon and Larkins, 2005).These mutants vary in mode of inheritance and value for protein quality improvement. Besides *opaque-2*, the most studied mutants are *opaque-1, opaque-5, opaque-6, opaque-9, opaque-7, opaque-11, floury-1, floury-2, floury-3, Mucronate and Defective Endosperm B30* (Vasal, 2001). These mutants have regulatory functions and some of them modify the storage proteins of endosperm (Gibbon and Larkins, 2005) affecting the size, shape or number of subcellular structures (called protein bodies) that determine the opaque phenotype (Holding and Larkins, 2009). Although, many advances has been reported for knowing loci and molecular mechanisms involved in the maize kernel modification, many questions remain and more work is required to better understand the concrete mechanism implicated in the modification of the vitreous endosperm.

2.4. The *opaque-2* gene

Opaque-2 gene is a transcriptional factor that is involved in the expression of amino acid synthesis, in carbon metabolism (trichloroacetic acid and glycolysis), in storage protein and starch metabolism, in gene transcription and translation processes,

in signal transduction, and in protein, fatty acid, and lipid synthesis as it was revealed by protein, amino acid and transcriptome-wide analyses (Hartings *et al.,* 2011). It is located in the short arm of chromosome 7 of maize, bin 7.01, between 10,793,452 and 10,796,233 positions (http://www.maizegdb.org) of the reference genome B73 and segregates as simple recessive gene. *Opaque-2* gene encodes a protein with a leucine zipper motif (bZIP) that is specifically expressed in the endosperm (Schmidt *et al*., 1990). It has a DNA-binding motif characteristic of proto-oncogenes and transcriptional factors of yeast (Hartings *et al.,* 1989). The sequence of *Opaque-2* deposited in the GeneBank, the X15544 sequence, has been reported to contain 4761 bp. This sequence is broken in six exons and five introns. After splicing, *Opaque-2* sequence has 1353 bp to give a total of 451 aa (Maddaloni *et al*., 1989). In particular, Opaque-2 protein regulates the expression of a subset of prolamin genes, specifically 22-kDa α–zein and 15-kDa β–zein genes reducing their expression level and increasing the content of minor storage proteins (Gibbon and Larkins, 2005; Wu *et al*., 2010; Motto *et al*., 2011) and non-zeins proteins as the elongation factor 1-α (*ef1-α*), which is highly correlated with the high lysine content in the opaque endosperm (Habben *et al*., 1995). Reduction in the transcription of 22-kDa α–zein fraction in the endosperm, without altering the contribution of other richer Lys and Trp storage proteins, could be a feasible approach to improve the amino acid balance of maize grain (Prassana *et al*., 2001). Opaque-2-prolamin modulation occurs by direct DNA-binding in conserved 7-bp sequence element 5´-TGTAAAG-3´, called prolamin box, found in the promoter region of seed storage protein genes. Opaque-2 protein can also interact with a DNA-binding factor in those prolamin genes have no the prolamin box (Vicente-Carbajosa *et al*.,

1997). Furthermore, opaque-2 protein modulates other non-storage proteins such as type I ribosome-inactivating protein *b-32* (Lohmer *et al*., 1991), one of cytosolic isoform of the pyruvate orthophosphate dikinase gene (cyPPDK1) (Manicacci *et al*., 2009), a heat shock protein *b-70*, and the aspartate pathway enzymes lysine-ketoglutarate reductase, aspartate kinase1 and aspartate kinase2 (Wang *et al.*, 2007).

2.5. The aspartate kinase-2 (AK) enzyme

Aspartate kinase (AK, EC 2.7.2.4) is an enzyme in the aspartate-derived amino acid biosynthetic pathway [the aspartate pathway directs the synthesis of Lys, Thr, Met, and Ile (Gaziola *et al*., 1999)] which catalyzes the phosphorylation of aspartate to β-aspartyl phosphate. In plants, AK and homoserine dehydrogenase (HSDH, another enzyme in the same pathway) are feedback inhibited by pathway end products (Muehlbauer *et al*., 1994). In maize, there are at least five genes encoding two or more isoforms of AK, based on their feedback inhibition properties. Two genes, *Ask1* and *Ask2*, encoding monofunctional AKs have been reported (Wang *et al*., 2001). They have been mapped to the short arm of chromosome 7 and the long arm of chromosome 2, respectively (Muehlbauer *et al*., 1994). Two mutants, *ask1 and ask2*, identified by genetic screening, encode enzymes are less sensitive to lysine resulting in a phenotype with over-production of free lysine, threonine, methionine, and isoleucine (Muehlbauer *et al*., 1994, Wang *et al*., 2007). In a cross between inbred maize lines Oh545o2 and Oh51Ao2, Wang *et al*. (2001) identified several QTLs linked with this phenotype. One of these QTLs was located on the long arm of chromosome 2 and was linked with loci encoding aspartate (Asp) kinase 2 enzyme. Wang *et al*. (2007) also studied the nature

of the monofunctional aspartate kinase genes in the parental inbreeds, Oh545o2 and Oh51Ao2. They found that the higher level of free amino acids derived from the aspartate pathway in Oh545o2 endosperm results from a single amino acid change in the ASK2 enzyme. Their results also showed that the carboxy-terminus of the enzyme plays an important role as regulatory mechanism for monofunctional AK. It would be interesting to determine if this single amino acid change is also present in other genotypes different to Oh545o2 as well as its relationship with other amino acid involved in the same pathway.

2.6. SSCP: a gel-based mutation scanning method

Single strand conformation polymorphism (SSCP) is a gel-based method for detecting differences in the sequence of amplified DNA fragments through polymerase chain reaction (PCR). The method was first described by Orita *et al*. in 1989 and it is based on the observation that single stranded DNA is unstable and re-anneals to itself to make conformations that vary depending on the sequence pattern (Hodgkinson *et al*., 2002). Published mutations that have been detected using this method include base substitutions, small insertions and deletions, and rearrangements (Hayashi and Yandell 1993). SSCP is considered a screening technique rather than a definitive means of identifying mutations as it does not provide information on the specific change in DNA sequence. Although, DNA sequencing has become cheaper and faster than in the past, using SSCP method is possible to identify only those individuals harboring the mutation in the sequence, such that the number of samples to be sequenced is greatly reduced (Sunnucks *et al*., 2000).

The SSCP method has been extensively applied to detect rare point mutations especially in biomedical research, but also it has been applied, in a lesser extension, in molecular phylogeography and systematics through study of microsatellite allele length homoplasy (alleles with different evolutionary histories but having the same length), as well as in molecular biology of populations by examining the mtDNA sequence variation by mentioning some fields of study (Hayashi, 1991; Sunnucks *et al*., 2000; Lia *et al*., 2007). Many scores of adaptations and refinements for SSCP method have been reported. The method is reputed to approach 100% if more than one experimental condition is employed (Kourkine *et al*., 2002).

The principle of SSCP is based on that electrophoretic mobility of a particle in a gel is sensitive to both its size and shape (Hayashi, 1991), so that an alteration in DNA nucleotide sequence caused by a mutation will affect the single-stranded DNA (ssDNA) folding and, hence, the electrophoretic mobility of a DNA fragment analyzed under nondenaturing conditions (Kourkine *et al*., 2002). In SSCP, PCR fragments of double strand normal DNA (wild type) and mutant DNA looked very similar in agarose gel, are denatured to form single stranded molecules of equal length. These molecules re-anneal onto themselves according to degree of intra strand base pairing. The newly formed structures will show differences in electrophoretic mobility and can be separated on a chilled non-denaturing polyacrylamide gel. For applying SSCP method, no specialized equipment or reagent is needed, minimal optimization is required and PCR product can be analyzed for length and sequence variation, reamplified or isolated for sequencing (Sunnucks *et al*., 2000).

Despite SSCP is a sensitive, inexpensive and easily performed method for haplotype detection, there are several electrophoretic variables (running temperature, pH and buffer composition, denaturants, DNA concentration, DNA size and sequence, and gel polyacrylamide concentration) that influence in the degree of strand separation (Hongyo *et al*., 1993; Sunnucks et al., 2000; Kourkine *et al*., 2002). Much of the literature on SSCP focuses on technical refinements to optimize the sensitivity of the method. Some electrophoretic variables have been studied very systematically and with relative thoroughness (Hongyo *et al*., 1993), however, these reports indicate only technical adjustments for a particular fragment or specific sequence, so that there is no a global SSCP procedure and the optimization to this method continues being highly empirical (Hayashi and Yandell, 1993).

Unlike SSCP method, other non-gel-based SNP genotyping methods are available either for use in MAS or for allelic information (as KASP™ chemistry). These methods found impulse with the publication of genome sequence of model organisms such as maize (Schnable *et al*., 2009; Vielle-Calzada *et al*., 2009) and the development of next-generation sequencing technologies (NGS). These technologies are based on hybridization methods, allele-specific PCR, primer extension, oligonucleotide ligation, and endonuclease cleavage (Gupta *et al.,* 2001; Semagn *et al.,* 2013), and they open the opportunity to increase the marker density for discovery applications, the genomic efficiency of genotyping, and do cost-effective and of high throughput. In that point, single nucleotide polymorphisms (SNP) are becoming the markers of choice (Hamblin *et al*., 2007). Both SNPs and small insertion and deletions (InDels) are the most frequent form of naturally occurring genetic variation in eukaryote genomes. SNPs are

widespread dispersed in the genome, are abundant, and are subjected to low mutation rates (Syvänen, 2001). However, SNPs are bi-allelic with a polymorphism lower than SSRs (multi-allelic). This drawback is compensated by using a higher number of SNP in the screenings through next-generation sequencing (NGS) technologies (Jones *et al*., 2007). These ultra-high throughput genotyping technologies facilitate the SNP screening; besides provide a low rate of missing data (Semagn *et al.,* 2013). NGS technologies can also provide information about copy number aberrations, allelic information, somatic rearrangements and base pair mutations in a single experiment (Reis-Filho, 2009). However, most of these NSG technologies are based on known sequence information and rely on detection methods that require highly sophisticated equipment. Furthermore, in spite of NSG technologies are highly accurate; their prices are prohibitive for most laboratories since the initial cost for investment is considerable. In addition, to perform SNP assays is required a provision of expensive consumables as specific probes, reagents and particular conditions that increase the cost per data point beyond the budget that research allows. In that point, SSCP is an alternative for the detection of genetic variations when the genome sequence of the organism under study is not available and NSG technologies for SNP discover cannot be applied. As SSCP focuses only in the sequencing of those amplicons that showed shifts in the electrophoretic mobility, the technique helps to discover those variants implied in the expression of the trait. The method is simple, versatile and cost-effective for genetic screening. As SSCP is a gel-based method, it has a low initial investment for its implementation and can work well with the stuff usually found in most laboratories (Sunnucks *et al*., 2000).

III. MATERIALS AND METHODS

3.1. Plant material and DNA extraction

All plant material was obtained from the Germplasm Bank of CIMMYT and consisted primarily of maize inbreed lines to assay polymorphisms (Table 1) and two segregating F_2 populations to measure the free amino acid content in the seeds. The genomic DNA was extracted from leaves of four individual seedlings of three to two-week-old following the CTAB protocol as it has been described in CIMMYT´s laboratory protocols (2005). Genomic DNA from endosperm tissue of three seeds per maize line was extracted according to the CTAB-lauril sarcosyl procedure described by Gao *et al*. (2008) with minor modifications consisting in avoiding the soaking of the seeds and use scissors for clipping dog´s claws to facilitate the collection of the tissue. All DNA for SNP typing and polymerase-chain reactions (PCR) was checked by spectrophotometry using a NanoDrop 8000 (Thermo Fisher Scientific, USA) and when it was required, samples were pooled manually in equimolecular concentrations.

3.2. Development of PCR-based markers

3.2.1. SNP assays for the *opaque-2* alleles

In this step, Dr. Debra Skinner´s collaboration (former scientific staff of CIMMYT) was essential in analyzing the genomic sequence of the *opaque-2* gene. The survey for identifying SNPs as a source for developing PCR-based markers was conducted in seven QPM lines and nine normal maize lines (Table 1). QPM lines, from diverse

agroecological zones, were chosen based on their excellent kernel modifying capacity and because they have been extensively used in the conversion activities at CIMMYT, while normal maize materials were chosen based on their adaptability and superior performance in Asia, Latin America and East African highland regions.

Table 1. Maize inbred lines examined to identify single nucleotide polymorphisms and small insertion/deletions into genomic sequences of the *opaque-2* locus.

Grain color and adaptation	QPM line	Non QPM line
white tropical lowland	CML159	CML343
	CML491	CLRCW01(309)
	CML502	CLRCW01(310)
	CML492	CL RCW22
yellow tropical lowland	CML161	CML348(G26 SEQC3)
	CML165	
white sub-tropical	CML176	
white highland		CML244
		CML349
Ethiopian white		F-7287
		A-7018

Five sets of primers were designed to amplify large fragments of the *opaque-2* locus for sequence comparisons; the objective was to identify regions, in exons, with

high degree of conservation. Their design was done after determining positions of introns and exons of six cDNA sequences of *opaque-2* gene. Such sequences were of the next accessions: AJ41297, AJ491298, AJ491299, AJ491300, M29411 and X16618. The reference sequence was the genomic sequence of the zein regulatory gene *opaque-2* (O2) of *Zea mays* (accession X15544). All accessions were downloaded from the GenBank (http://www.ncbi.nlm.nih.gov/). From all designed primers, only a pair set (Table 2) was used to amplify and clone a total of 1 kb of the *opaque-2* locus. The fragments were obtained in duplo by PCR and separated by submarine electrophoresis in 5% MetaPhor™: Seakem® (4:1) agarose gels with 1.0X TAE buffer (0.04M Tris-acetate, 0.001M EDTA, pH 8.0). Bands were excised from the gel under UV light and cloned into pGEM®-T Easy Vector (Promega, USA) according to manufacturer's recommendations. Heat shock-bacterial transformations were performed in competent cells of *Escherichia coli* DH5α strain following CIMMYT´s laboratory protocols (2005) and plated onto S-Gal medium (Sigma S-9811) plus ampicillin (100 ug ml^{-1}). Colonies were recovered in three ml of LB broth plus ampicillin (100 ug ml^{-1}) and incubated at 37°C overnight at 250 rpm. Plasmid DNA was extracted using Wizard® *Plus* SV Minipreps DNA Purification System (Promega, USA). Once plasmids were tested for the presence of the insert by restriction digestion with *Eco*R1, two plasmids of each genotype were sequenced at CINVESTAV facilities (Irapuato, México). Resulting sequences were aligned using Clustal X version 1.8 (Thompson *et al*., 1997) and refined by hand. Only a consistent SNP (hereafter referred as SNP1(315)-C/T) in the examined QPM lines but different in the normal lines, was chosen to be typed in 90 maize lines at Kbiosciences (Hoddesdon, UK) using the KASPar® chemistry (Table 3).

The allele calls were visualized with the SNPviewer PC tool. Missing data were removed from analysis.

3.2.2. SNP assays for the *Ask-2* alleles

In a similar way to the *opaque-2* assay, a reported SNP involved in the single amino acid substitution in the C-terminal region of the ASK2 protein of Oh545o2 maize line, and also responsible for the altered basal activity of the aspartate kinase enzyme (Wang *et al.,* 2007), was typed using the same panel of genotypes with the KASPar® chemistry as occurred with the SNP1(315)-C/T (Table 3).

Table 2. Sequence and melting temperature for primers designed to amplify two genomic fragments of the *opaque-2* locus.

Primer name	Length	Tm	sequence (5´- 3´)	Product size
O2F1	22	60.0	CTAGTGTTTGCTTCTCCCTTCC	543
O2R1	23	60.0	TCCTCAGTATGGCATTGTACTCC	
O2F3	21	60.5	TGTTGTGACCTCAGATCAACG	529
O2R3	21	59.5	TTCCAGTTCTTTCAGGTGAGC	

3.3. Development of a PCR-SSCP assay of moderate throughput

To carry out a study for allele mining of *Ask-2* and assess also the utility of PCR-based markers, a gel-based PCR-SSCP assay was designed. The applicability of this

assay was only focused in the polymorphism for *Ask-2* since all efforts with the SNP1(315)-C/T polymorphism were not large enough to effectively score the haplotypes at the established SNP for *opaque-2* locus. The PCR-SSCP assay consisted in analyzing the conformational polymorphism of a fragment of *Ask-2* obtained via PCR in a set of 62 QMP lines and 20 normal maize lines. Primers for PCR were 5´-CTTGGTGGTCCATGACAGTG-3´ (forward) and 5´-CTAATGGCTGTGGATTGTGC-3´ (reverse) and they were synthesized by Integrated DNA Technologies (Coralville, IA, USA). PCR amplification was accomplished in 15 µl reaction volume containing 50 ng of genomic DNA, 0.1 µM of each primer, 166 µM of each dNTP, 1.0X PCR *Taq* buffer [100 mM of Tris-HCl (pH 9.0), 500 mM of KCl, 1% Triton X-100], 2.0 mM $MgCl_2$ and 1.0 Unit of *Taq* polymerase and PCR-grade water. After an initial denaturation at 94°C for 2 min, 30 amplification cycles were performed at 94°C for one min, 58°C for 40 sec and 72°C for 1 min, followed by a final extension step at 72°C for 5 min. Five µl of loading buffer (10 mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol) was incorporated to each PCR product. To check amplification, 10 µl of each PCR product (including loading buffer) were visualized by electrophoresis in 2% Seakem® LE Agarose (Lonza Rockland Inc., USA) gels using 1.0X TBE buffer (89 mM Tris base, 89 M boric acid, 2 mM $Na_2EDTA\ 2H_2O$, pH 8.3).

To perform polyacrylamide gel electrophoresis (PAGE), chemicals employed were of molecular grade. A nondenaturing vertical discontinuous gel system based on Laemmli´s method (1970) was used to separate the PCR products. The formulations for this system are included in the Appendix A.1. Discontinuous polyacrylamide gels

consisted of resolving or separating (lower) gels and stacking (upper) gels. The final concentration of resolving gel was 16.34% (w/v) acrylamide-bisacrylamide (40%, 29:1), 0.384 M Tris base pH 8.8, 0.04% (w/v) ammonium persulfate (APS) and 0.08% (w/v) 1,2-bis-dimethylamino-ethane (TEMED), while stacking gel consisted of 5% (w/v) acrylamide-bisacrylamide (40%, 29:1), 0.133 M Tris base pH 6.8, 0.096% (w/v) APS and 0.12% (w/v) TEMED. Two vertical electrophoresis systems were employed according to interest of the study. For allele mining, non-denaturing gels of 13 x 14 cm, one mm thick were prepared with the ATTO AE-6220 system (ATTO Co., Tokyo, Japan), while for mass genotyping purposes, a MGV-216-33 Dual Triple Wide Mini-Vertical Electrophoresis System (C.B.S. Scientific, USA) was employed. Gels and running buffer (1X Tris glycine buffer composed by 24.7 mM Tris base, 191.8 mM glycine, pH 3.0) were kept at 8°C before loading the PCR products. The samples were denatured for 10 min at 94°C and plunged into ice. Aliquots of 5 µl were loaded onto gel and subjected to electrophoresis for 15-16 h at 150 V constant and 8°C. Gel was stained in a 1x SYBR®Gold Nucleic Acid Gel Stain (Life Technologies, USA) staining solution during 20 min prepared with 1x TBE buffer. Images were scanned using a Typhoon 8600 system (Amersham Biosciences Ltd., Piscataway, NJ).

3.4. Interaction analysis between *opaque-2* and *Ask-2* genotypes

In order to validate if there was changes in the free amino acid content in the kernel and to study the interaction between opaque-2 and *Ask-2* genotypes, two di-hybrid F_2 populations, CML172 x CML323 and CLQRCWQ26 x La Posta Sequía (C7-F64-2-6-2-1-B-B)-B), were evaluated. Each population was a cross of a QPM line with a

normal maize line. Four hundred individuals in each population were genotyped (single seed based) for both *opaque-2* and *Ask-2* loci. After genotyping, nine genotypic classes were determined. These classes were the result of the segregation of two loci in a F_2 population.

3.4.1. *opaque-2* and *Ask-2* genotyping

To perform the genotyping, the phi057 SSR marker was chosen for the *opaque-2* gene while the newly developed marker for *Ask-2* was used for genotyping the locus of the same name. For phi057 marker, the primer sequences were 5´-CTCATCAGTGCCGTCGTCCAT-3´ (forward) and 5´- CAGTCGCAAGAAACCGTTGCC-3´(reverse). PCR amplification used a 15-µL reaction mix containing 2 µl of 20 ng template DNA, 1.0X PCR *Taq* buffer [100 mM of Tris-HCl (pH 9.0), 500 mM of KCl, 1% Triton X-100], 2.5 mM $MgCl_2$, 150 µM of each dNTP, 0.3 µM each of forward and reverse primers, 1.0 unit of *Taq* DNA polymerase and PCR-grade water to reach the final volume. PCR cycling conditions were 2 min DNA denaturation step at 94°C, followed by 30 cycles consisting in 30 s at 94°C, 1 min at 58°C, and a 1 min at 72°C, and a final step of 5 min at 72°C on a MJ Tetrad thermocycler (MJ Research, MA, USA). After PCR, 5 µl of 5X loading buffer (50 mM Tris base pH 8.0, 5 mM EDTA pH 8.0, 25% sucrose, 0.2% bromophenol blue, and 0.2% xylene cyanol) were added to each reaction. To resolve PCR products, 3.0 µl per reaction were loaded in PAGE and stained with silver nitrate according to CIMMYT´s laboratory protocols (2005). Gel matrices included 10% polyacrylamide (40%, 29:1) using a CBS system. Electrophoresis was carried out in 1.0X TG buffer (24.7 mM Tris base, 191.8 mM

glycine, pH 3.0) at 250 V constant and room temperature for 3 h. Finally, *Ask2* genotypes were resolved following the recently developed protocol of non-denaturing vertical discontinuous gel described here in the previous section.

3.4.2. Amino acid content analysis

All nine different genotypic classes were subjected to high-performance liquid chromatography (HPLC) analysis to measure free amino acid content (FAA) in the endosperm. Eleven amino acids were considered for the analysis, six of them are involved in the aspartic acid metabolic pathway [asparagine (Asn), aspartic acid (Asp) , threonine (Thr), lysine (Lys), isoleucine (Ile) and methionine (Met)] and five amino acids are from carbon metabolism [histidine (His), glutamine (Gln), leucine (Leu), phenylalanine (Phe) and tryptophan Trp)]. Individual analyses of variance per genotype in each population were conducted using PROC GLM of SAS and principal component analysis was applied using PROC PRINCOMP of SAS to construct biplots graphs and examine the relationships between the genotypes and the accumulation of free amino acid content. Also, a rapid examination of the data was done to verify for the presence of interaction at those two locus following Cockerham (1954). Once it was verified, graphs were constructed with R software version 2.11.1 (15-07-2013).

IV. RESULTS AND DISCUSSION

4.1. Analysis of the *opaque-2* gene sequence

The sequence variation of two large PCR fragments of the *opaque-2* gene was examined to discover polymorphisms as potential marker-alleles. A total of 31.34 kb were sequenced from 62 PCR products. Sixty-five polymorphisms were detected in the panel of 16 examined genotypes which provided one polymorphism each 30.13 bp. This value is similar to those found by Ching *et al*. (2002), who reported one polymorphism per 31 bp in noncoding regions and one polymorphism per 124 bp in coding regions. The detected polymorphisms were irrespective to the nature of the line. In most cases, the polymorphisms were present in a particular set of genotypes (QPM) and were also present in some lines of another group of genotypes (non QPM). For example, the nucleotides changes detected in the intron 2 of the genomic sequence of *opaque-2* make possible to define the haplotype GT for the QPM lines set. However, this haplotype was also identified both in CML348 and CLRCW01 310, which are two normal lines where the prevailing haplotype was AG for the same polymorphic sites (Figure 1). A similar observation was found for other two SNPs which defined the haplotype AT in the QPM lines and TC in most of the set of normal lines; in this position a non QPM line CML348 had the same haplotype (AT) as QPM lines whereas another non QPM line (CLRCW01 310) revealed a haplotype different (AC) with respect to the haplotype registered in most of normal lines (TC). These observations indicate a high variation at sequence level and reinforces the concept of that maize is a specie highly polymorphic.

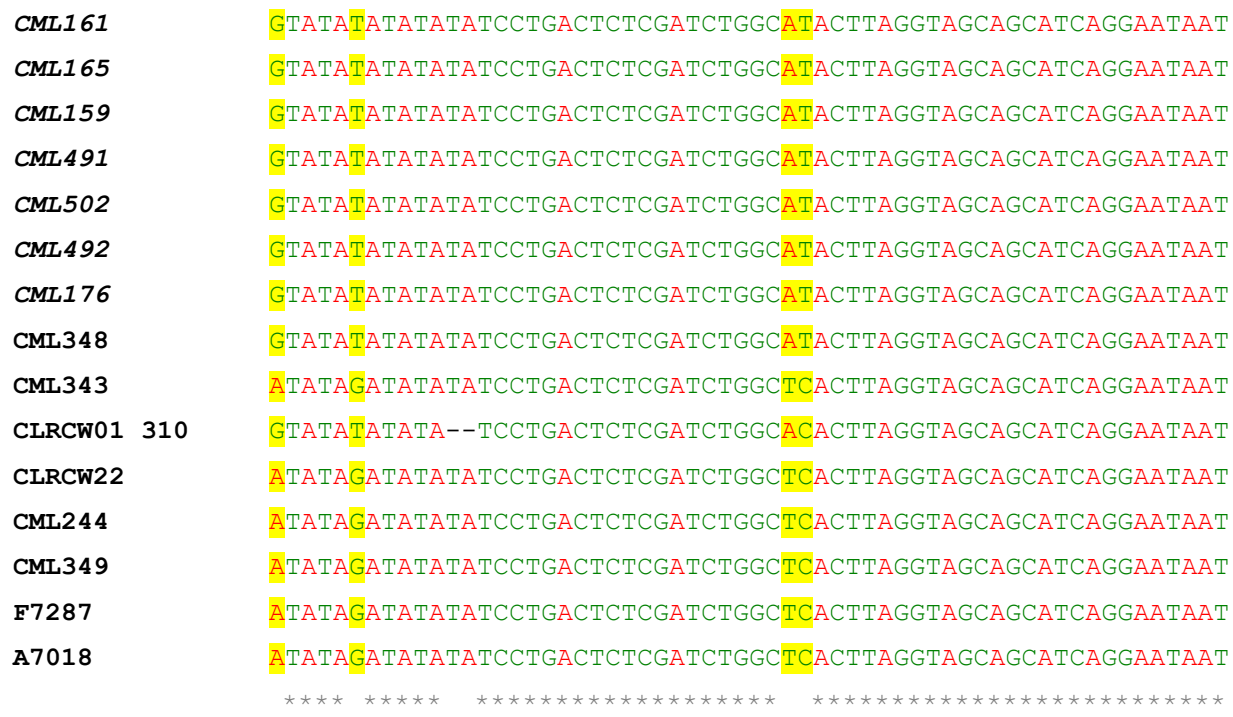


Figure 1. Genomic sequence alignment of an *opaque-2* fragment showing three genetic polymorphisms. Most of polymorphisms predominated in a particular set of genotypes but occasionally were observed in another group of genotypes.

Two small insertion and deletion (InDel) were also observed (Figure 2). One InDel was present in all QPM set but also in two normal lines, CML348 and CLRCW01 310. Another InDel was detected in the same genotypes as occurred with the first InDel, with the annotation that this sequence variation in CLRCW01 310 was three nucleotides larger than the first one. Ching *et al*. (2002) also reported that InDels are frequent in non-coding regions (1 per 85 bp), but rare in coding regions. Although, the extent of polymorphism varies between species and genome regions; clearly, the DNA sequences of the panel of maize lines under examination were highly polymorphic,

especially in the normal lines and are a good representation of the diversity at nucleotide level. Underlying causes of this variation could come of the type of cross-pollinating of this specie and for transposons. It has been suggested that the active transposons systems promotes diversity (Gaut *et al*., 2000). In fact, in maize nearly 85% of the genome is composed of hundreds of families of transposable elements dispersed non-uniformly across the genome (Schnable *et al*., 2009).

Figure 2. Small insertion and deletions (InDel) detected in the sequence of an *opaque-2* PCR fragment of seven QPM lines (in italics, on top) and eight non QPM lines.

On the basis of this variation, two regions with a minimum score of polymorphisms were chosen to examine polymorphisms present in the set of QPM lines but different in the set of normal lines. The selected regions were located in the first exon of *opaque-2*. These regions shared the same sequence in a long track. Over them, five SNPs and two InDels were distinguished after aligning sequence data. Three single-base changes were identified as being transitions (C→T or G→A), whereas two of them were transversions (T→G or A→T). Only a single substitution C→T in the exon 1, position 315 [referred as SNP1(315)-C/T in Materials and Methods] was consistently observed among all QPM genotypes but different in the set of normal maize inbred lines. This single nucleotide substitution was C in all QPM material whereas in the normal lines was T (Figure 3). This polymorphism was the source for the SNP assays development. Accordingly, it was further investigated if this polymorphism SNP1 (315)-C/T could be helpful as molecular marker to discriminate QPM genotypes and non-QPM genotypes.

4.2. High throughput assays for *opaque2* and *Ask-2* genes

KASPar chemistry as SNP detection method was suitable for the high-throughput screening for SNP alleles of both *opaque-2 and Aks-2* loci. This automated, non-gel based protocol allowed discriminate alleles for the accessions included in the panel under testing (60 QPM lines and 28 normal maize lines). For the candidate SNP detected in the *opaque-2* locus (C or T), the clustering pattern of the SNP calls (Figure 4) clearly differentiated two groups of haplotypes. One of them showed a strong

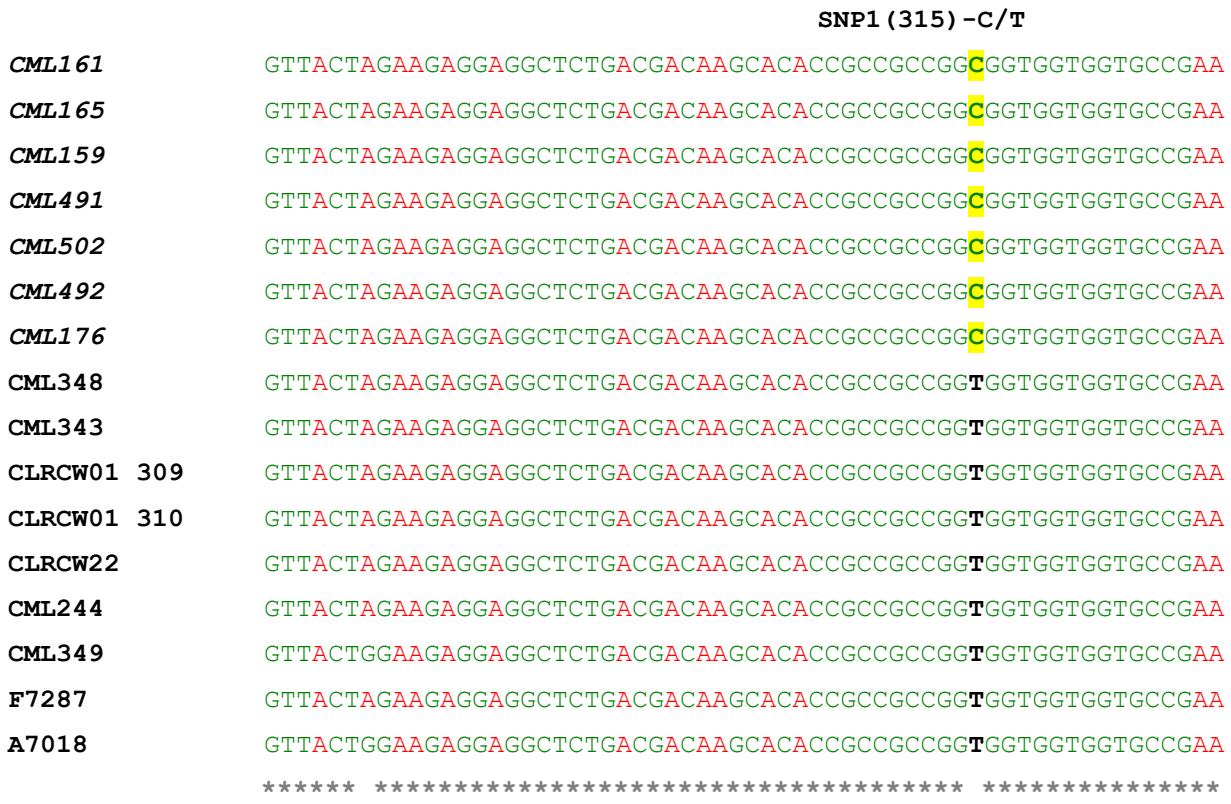


Figure 3. Sequence alignment of an *opaque-2* fragment showing a potentially discriminative SNP present in all QPM lines (C, in italics, on top) but different in non QPM lines (T).

clustering while another group had a slightly disperse clustering but with the enough signal to be correctly scored. Only there was one miscalling data. Hence, the high throughput assay showed a 99% of effectiveness in the number of valid data. The first group was scored with the allele C (53 QPM lines) whereas the second group was scored with the allele T (35 maize lines). Intriguingly, in this group seven QPM lines (CML145, CML167, CML168, CML177, CML180, CML184 and CML188) were scored to have the allele T as the normal lines (Table 3). Unlike, none of the normal lines scored

had the allele C. Therefore, the polymorphism SNP1(315)-C/T marker was not perfectly associated with the *opaque-2* gene and was 88.3% assertive to discriminate QPM lines from normal lines. However, this allele could be used for MAS only in the cases when the QPM progenitor is harboring the allele C and the non QPM progenitor is harboring the allele T. Regrettably, these results stop the intention to develop a PCR-based marker for *opaque-2* locus using the polymorphism SNP1(315)-C/T.

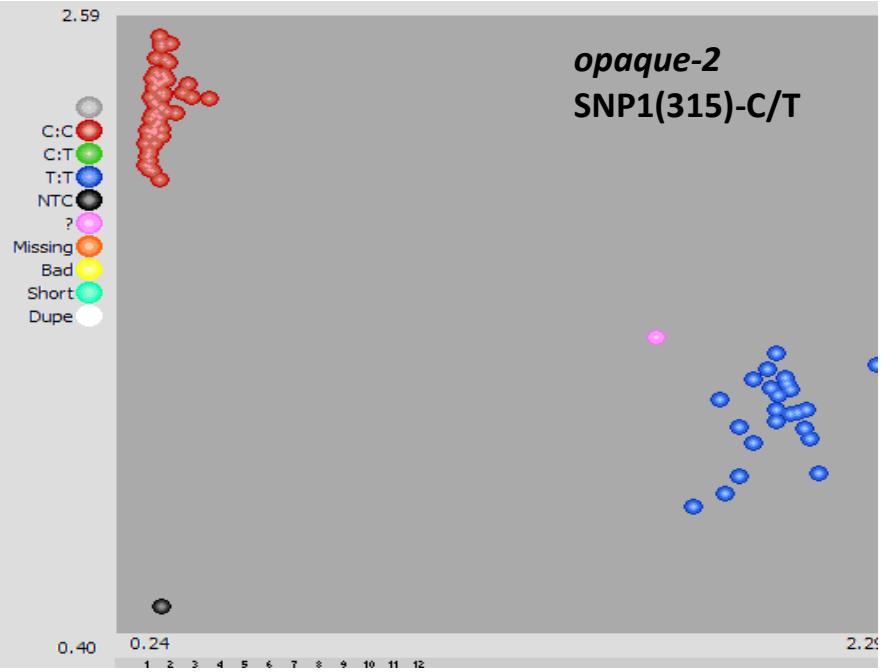


Figure 4. Cluster of SNP alleles for *opaque-2* locus in several QPM lines and normal maize lines using KASP™ chemistry. Genotyping data was plotted for the polymorphism SNP1 (315)-C/T of the *opaque-2* gene. Dots in red represent samples with the allele C prevailing in most of the QPM lines. Dots in blue are samples with the allele T identified in the maize normal lines and in six QPM lines.

Despite the results, is relevant to mention that a high throughput method enables to simultaneously genotyping of several samples in one go and has some operational advantages over gel electrophoresis. A high throughput method avoids manual gel tracking and health exposition to moderately hazardous reagents, as well as reduces potential data point-scoring errors, labor and time. Respect to quality score, a high throughput technology is based in the analysis of cluster data. Each cluster is composed by homozygote and heterozygote individuals depending of the SNP allele. Clusters should be clearly separated and each SNP call is scored according to the cluster where it falls. As this technology uses two different oligonucleotides specific for each SNP allele, labeled with different fluorescence dyes, even if the signal is weak and the SNP call no falls in a cluster, it will be possible see the presumable cluster belonging to, and therefore, determine the type of SNP allele. This differs using a gel-based system where sometimes is necessary amplify again the sample, and follow all post-PCR steps to know the polymorphism of the marker, which implies cost, labor and time. Therefore, if the polymorphism is known for a particular locus, the high throughput method could be a good alternative for MAS applications.

Parallel to validation of SNP1(315)-C/T, another high throughput assay was conducted to validate the feasibility (as a marker-allele) of a SNP which had been reported previously in the C-terminal end of the Ask-2 locus responsible of high Lys content in the Oh545o2 maize line (Wang *et al.,* 2007). The non-gel high throughput assay demonstrated to be effective for allelic discrimination again (Figure 5). However, the assayed polymorphism for *Ask-2* locus demonstrated to be less discriminative than

SNP1(315)-C/T to distinguish QPM haplotypes with respect to non QPM haplotypes. The screening of this SNP in a range of diverse germplasm of CIMMYT, revealed that 92.0% of genotypes were represented by the allele T (Table 3). This percentage included all 26 normal maize lines and 55 QPM lines. The other allele A was detected only in five QPM lines (CML148, CML161, CML162, CML179 and CML188) and two normal lines (CML328 and CML 370). Instead of discarding this information due to outcomes for *Ask-2*, the data was compared with those generated by SSCP analysis.

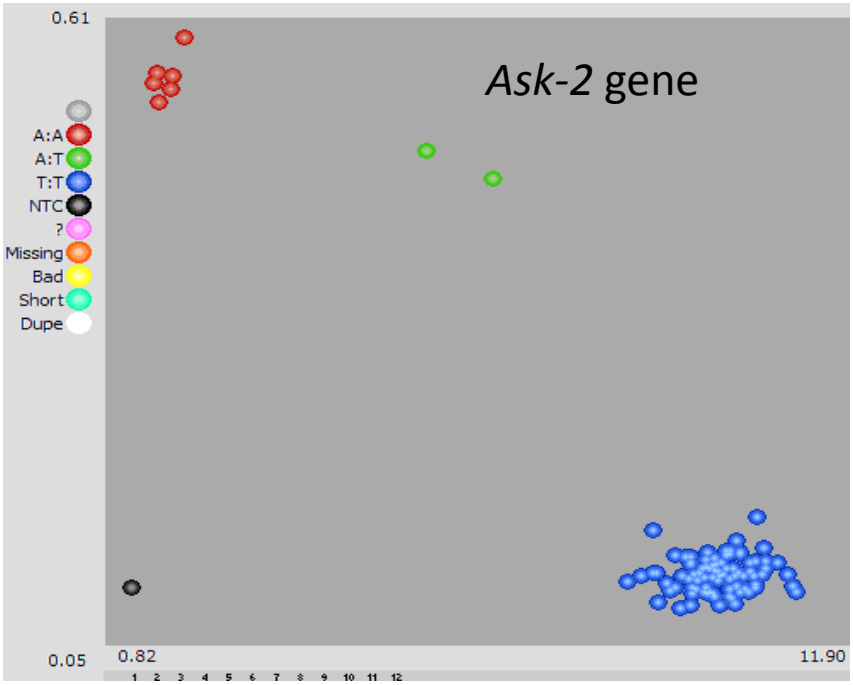


Figure 5. Cluster of SNP alleles for *Ask-2* locus in several QPM lines and normal maize lines using KASP™ chemistry. SNP calls were plotted for the polymorphism of *Ask-2* gene. Dots in red are samples with the allele A detected in five QPM lines plus the control. Dots in blue identify 83 samples with the allele T. Dots marked green are heterozygous samples.

Table 3. Single nucleotide polymorphism typing for two loci involved in the genetics of QPM, *opaque-2* and *Ask-2*, using the KASP™ chemistry in a panel of QPM lines and normal maize lines.

Maize Line	type	*Opaque-2*	*Ask-2*
CML-010	normal	T	T
CML-091	normal	T	T
CML-108	normal	T	T
CML-140	QPM	C	T
CML-142	QPM	C	T
CML-143	QPM	C	T
CML-144	QPM	C	T
CML-145	QPM	T	T
CML-146	QPM	C	T
CML-147	QPM	C	T
CML-148	QPM	C	A
CML-149	QPM	C	T
CML-150	QPM	C	T
CML-151	QPM	C	T
CML-152	QPM	C	T
CML-153	QPM	C	T
CML-154	QPM	C	T
CML-155	QPM	C	T
CML-156	QPM	C	T
CML-157	QPM	C	T
CML-158	QPM	C	T
CML-159	QPM	C	T
CML-160	QPM	C	T
CML-161	QPM	C	A
CML-162	QPM	C	A
CML-163	QPM	C	T
CML-164	QPM	C	T
CML-165	QPM	C	T
CML-166	QPM	C	T
CML-167	QPM	T	T
CML-168	QPM	T	T
CML-169	QPM	C	T
CML-170	QPM	C	T
CML-172	QPM	C	T
CML-173	QPM	C	T

CML-174	QPM	C	T
CML-175	QPM	C	T
CML-176	QPM	C	T
CML-177	QPM	T	T
CML-178	QPM	C	T
CML-179	QPM	C	A
CML-180	QPM	T	T
CML-181	QPM	C	T
CML-182	QPM	C	T
CML-183	QPM	C	T
CML-184	QPM	T	T
CML-185	QPM	C	T
CML-186	QPM	C	T
CML-187	QPM	C	T
CML-188	QPM	T	A
CML-189	QPM	C	T
CML-190	QPM	C	T
CML-191	QPM	C	T
CML-192	QPM	C	T
CML-193	QPM	C	T
CML-194	QPM	C	T
CML-228	normal	T	T
CML-247	normal	T	T
CML-254	normal	T	T
CML-261	normal	T	T
CML-297	normal	T	T
CML-300	normal	T	T
CML-311	normal	T	T
CML-312	normal	T	T
CML-312 SR	normal	T	T
CML-323	normal	T	T
CML-325	normal	T	T
CML-326	normal	T	T
CML-328	normal	T	A
CML-332	normal	T	T
CML-370	normal	T	A
CML-371	normal	T	T
CML-374	normal	T	T
CML-376	normal	T	T
CML-447	normal	T	T
CML-483	normal	T	T

CML-484	normal	T	T
CML-486	normal	T	T
CML-490	QPM	C	T
CML-491	QPM	C	T
CML-492	QPM	C	T
CML-493	QPM	C	T
CML-495	normal	T	T
CLQ-RCWQ38	QPM	C	T
CLQ-RCWQ50	QPM	C	T
CLQ-RCWQ83	QPM	C	T
B73	normal	T	T
Mo17	normal	T	T

4.3. SSCP analysis for allele mining of the *Ask-2* gene

Considering that SNPs are the most common polymorphisms (Syvänen, 2001), SSCP method was chosen to migrate PCR products and identify possible functional variants for *Ask-2* locus. A new SSCP technique based on Laemmli´s discontinuos system was developed for such purpose. This technique was straightforward to implement and optimize in the laboratory. Its sensitivity to detect sequence variation without sequencing was also helpful because of the high resolution of the PCR products in PAGE. Although it was not able to provide information on the nature and position of the genetic variation (Sunnucks *et al*., 2000; Kourkine *et al*., 2002), SSCP technique provided further insight about the natural allele variation for *Ask-2* locus in the germplasm under study (82 maize inbred lines). After screening in PAGE using a modified SSCP technique, five different types of conformers, or SSCP variants, were visualized. Sixty lines showed the allelic variant 1 (46 QPM lines and 14 non QPM

lines), eleven lines evidenced to have the allelic variant 2 (10 QPM lines and one normal line), five lines had the allelic variant 3 (all QPM lines), five showed to have the allelic variant 4 (all maize normal lines from IIlinois University, USA), and only one QPM line (CIMMYT, Africa) had the SSCP variant 5 (Figure 6). It was particularly striking that all five QPM lines that showed to have the SSCP variant 3, were the same group of lines with the allele A for *Ask-2* as was determined with the KASPar™ chemistry. These lines were CML148, CML161, CML162, CML179 and CML 188. From these results, it was interesting to investigate if these different allelic variants had a relationship with the free amino acid content of the seeds.

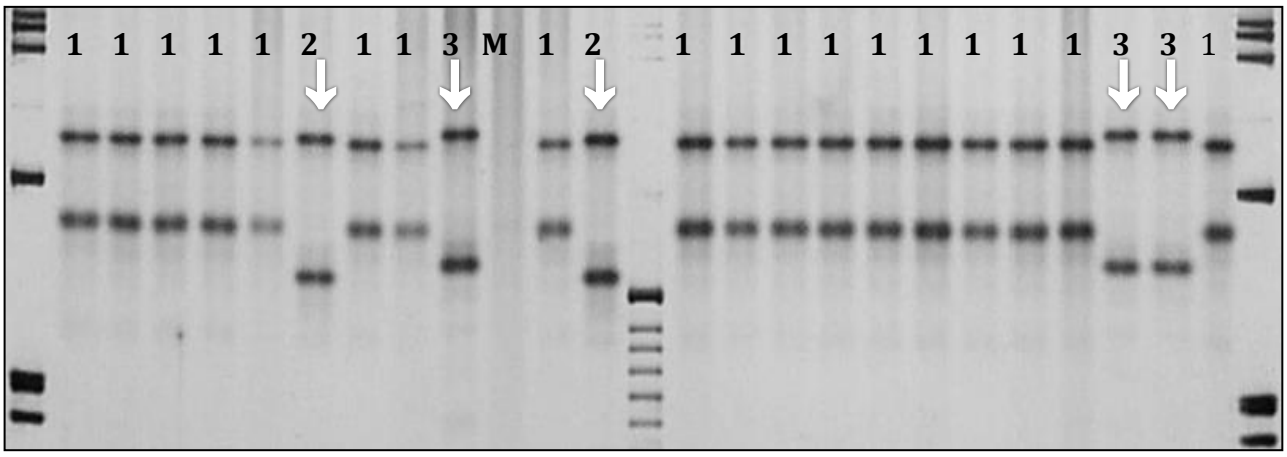


Figure 6. Gel image of PCR-SSCP variants for *Ask-2* locus in several QPM lines. The figure shows three of five types of SSCP variants (conformers) that could be distinguished in 90 assayed maize lines using two-layer PAGE. 1. SSCP variant prevailing in the set of maize lines carrying the allele T. 2. SSCP variant detected in nine QPM lines carrying the allele T. 3. SSCP variant detected in five QPM lines carrying the allele A.

4.4. Analysis of the free amino acid content (FAA)

In order to evaluate if the SNPs chosen as markers for both genes and the type of SSCP conformer had an effect in the FAA content in QPM lines, several amino acids were determined. Values for eleven amino acids are listed in the Table 4. Results indicated that there was a relation among all allelic variants of *Ask-2* and the FAA content. In particular, the effect for the total FAA content was more noticeable for those QPM lines bearing the haplotypes C:A:3 or C:T:2 (*opaque-2* allele:*Ask-2* allele: type of conformer, in that order). Altogether, these genotypes had 79.7% and 75.1% higher total FAA content with respect to the lines characterized with the haplotype C:T:1 or T:T:1, *i. e*., 4-5 fold more. These last lines had a similar but lower value of total FAA content and appeared to be irrespective to the type of *opaque-2* allele (C or T). Although this evidence should not be conclusive, it seemed that there was a liaison between the type of allele for *Ask-2* (A or T) along with the type of conformer (2 or 3) when the data of FAA content were converted in relative values.

Using relative values, those early differences annotated for allelic variants of *Ask-2* with respect to total FAA were inverted, especially for the Trp and Lys content (Table 5). Now, the lines typed as C:T:1 have the highest values for Trp and Lys residues (1.7% and 6.8%, respectively) compared with those lines typed as C:T:2 (with 0.8% and 4.6% for Trp and Lys, respectively), C:A:3 (0.8% and 4.0%) or T:T:1 (1.2% and 3.6%). Clearly, the conformer 1, in combination with the SNP alleles C of *opaque-2* had a strong effect in the Trp and Lys content and, in a lesser extension; the conformer 1 in combination with the allele T of *Ask-2* had also effect in the Trp content.

Table 4. Single nucleotide polymorphisms for *opaque-2* and *Ask-2* loci, SSCP variants for *Ask-2* gene and total content of free amino acid (nmol g^{-1} FW) of ten QPM lines (italics) and three normal maize lines.

Line	Alleles			Asp	Glu	Asn	Thr	His	Met	Trp	Phe	Ile	Leu	Lys
CML150	C†	T††	1¶	993.8	424.2	586.5	102.8	92.8	42.0	45.7	68.6	66.1	99.7	149.8
CML491	C	T	1	1617.9	582.4	1972.5	93.9	173.9	57.1	88.6	61.9	50.8	78.5	410.8
CML151	C	T	2	5574.4	5755.6	4068.2	1204.4	366.9	397.0	168.6	509.6	598.9	546.9	1001.1
CML172	C	T	2	4969.7	2759.5	4631.1	286.3	61.1	28.0	53.1	210.1	159.8	268.8	719.1
CML176	C	T	2	5431.1	3733.3	4304.8	493.6	249.8	79.1	132.5	219.5	230.4	398.2	437.6
CML492	C	T	2	4027.6	1686.0	2527.7	400.8	274.9	339.9	137.5	195.9	299.0	307.9	911.1
CML148	C	A	3	5806.1	4908.8	3874.7	732.9	227.6	48.7	184.8	386.0	260.1	447.5	1304.9
CML161	C	A	3	3246.6	2799.7	3018.5	395.9	173.6	70.0	79.9	133.0	0.0	178.2	245.5
CML162	C	A	3	2027.3	1524.5	3171.1	141.8	41.9	44.1	55.3	91.5	129.8	138.9	136.9
CML179	C	A	3	9036.5	4734.7	8193.6	677.6	321.0	185.7	184.3	697.6	891.0	978.0	1290.4
CML254	T	T	1	495.0	303.2	1041.2	68.1	54.7	32.1	40.4	51.5	64.2	114.1	88.0
CML312	T	T	1	1165.6	1106.8	934.1	294.1	102.6	67.1	28.6	70.2	93.0	269.0	137.3
CML447	T	T	1	860.5.9	669.9	552.0	167.5	88.0	65.7	38.0	113.3	219.1	132.9	118.2

† Allele for *opaque-2*. ††Allele for *Ask2*. ¶Type of *Ask2*-SSCP conformer.

Table 5. Single nucleotide polymorphisms for *opaque-2* and *Ask-2* loci, SSCP variants for *Ask-2* gene and relative values of free amino acid of ten QPM lines (italics) and three normal maize lines.

Line	Alleles			Asp	Glu	Asn	Thr	His	Met	Trp	Phe	Ile	Leu	Lys
CML150	C†	T††	1¶	37.2	15.9	22.0	3.8	3.5	1.6	1.7	2.6	2.5	3.7	5.6
CML491	C	T	1	31.2	11.2	38.0	1.8	3.4	1.1	1.7	1.2	1.0	1.5	7.9
CML151	C	T	2	27.6	28.5	20.1	6.0	1.8	2.0	0.8	2.5	3.0	2.7	5.0
CML172	C	T	2	35.1	19.5	32.7	2.0	0.4	0.2	0.4	1.5	1.1	1.9	5.1
CML176	C	T	2	34.6	23.8	27.4	3.1	1.6	0.5	0.8	1.4	1.5	2.5	2.8
CML492	C	T	2	36.3	15.2	22.8	3.6	2.5	3.1	1.2	1.8	2.7	2.8	8.2
CML148	C	A	3	31.9	27.0	21.3	4.0	1.3	0.3	1.0	2.1	1.4	2.5	7.2
CML161	C	A	3	31.4	27.1	29.2	3.8	1.7	0.7	0.8	1.3	0.0	1.7	2.4
CML162	C	A	3	27.0	20.3	42.3	1.9	0.6	0.6	0.7	1.2	1.7	1.9	1.8
CML179	C	A	3	33.2	17.4	30.1	2.5	1.2	0.7	0.7	2.6	3.3	3.6	4.7
CML254	T	T	1	21.0	12.9	44.3	2.9	2.3	1.4	1.7	2.2	2.7	4.8	3.7
CML312	T	T	1	27.3	25.9	21.9	6.9	2.4	1.6	0.7	1.6	2.2	6.3	3.2
CML447	T	T	1	28.5	22.1	18.2	5.5	2.9	2.2	1.3	3.7	7.2	4.4	3.9

† Allele for *opaque-2*. ††Allele for *Ask2*. ¶Type of *Ask2*-SSCP conformer.

Therefore, the data revealed a combined effect of both *opaque-2* and *Ask-2* alleles for the total and relative content of amino acids, such that while a particular combination of alleles increased the total FAA content (C:A:3 and C:T:2), the another combination (C:T:1 and T:T:1) was remarkable in increasing the relative content of amino acids such as Trp and Lys. Although, these observations were obtained when data were analyzed together for a concrete combination of alleles and SSCP varianst, it was evident that there was variation for this trait independently of allelic combination. For instance, CML148 and CML492 had the highest relative content of Trp and Lys with a SSCP variant different to the type 1. These all outcomes demonstrate how different combinations of nucleotide diversity can affect the expression of a trait and reveal also the potential of this diversity to be exploited in the modern breeding programs. As the recessive condition of *opaque-2* is a hallmark of the QPM material, it was further investigated the interaction between *opaque-2* and *Ask-2* for enhanced free amino acid content.

4.5. Interaction between *opaque-2* and *Ask-2* loci

Two F2 populations (CML172 X CML323 and CLQRCWQ26 X La Posta Seq C7-F64-2-6-2-1-B-B)-B) identified as segregating ones for *opaque-2* and *Ask-2* were utilized to study the interaction between both genes. Values of the total FAA content for both populations are reported in the Table 6. As it is shown, significant differences ($P<0.01$) among *opaque-2* genotypes (*o2/o2*, *o2/O2* and *O2/O2*) in the population CML172 X CML323 were observed but not in the population CLQRCWQ26 X La

Posta Seq C7-F64-2-6-2-1-B-B)-B). On average, *opaque-2* in the form of homozygous recessive genotype (*o2/o2)* had the highest values for the content of ten amino acids. Clearly, the differences were higher for the content of Asp, Glu, Asn, Thr, His, Trp, Phe, Ile, Leu and Lys residues whereas the content for Met were non-significant. In proportion, the amount for the free content of aspartic acid accumulated by the homozygous recessive genotype was 83.3% and 87.2% highest when it was compared to the other two genotypes (homozygous dominant and heterozygous, respectively). Similar differences were observed for the accumulation of other amino acids. For example, in the case of Trp, the difference for the total amount by the recessive genotype was 52.9% and 59.7% higher with respect to the accumulation obtained in the other genotypes. Lys accumulation was 67.1% and 69.5% higher in a similar way as well. Particularly, the higher accumulation of the homozygous recessive genotype for other amino acids derived from the aspartate pathway was 74% and 80% for Asp, 39% and 50% for Thr and 58% and 68% for Ile. In the case of the amino acids derived from the carbon metabolism the accumulation was alike to the previous one. This was: 64% and 75% for Glu, 47% and 49% for His, 49% and 50% for Met, 59% and 67% for Phe and 34% and 52% for Leu of the recessive genotype with respect to other genotypes.

To graphically visualize a putative effect between the genotypes and the amino acids content, biplots graphs were constructed. It should be noted, that in spite of the differences could not be large to be significant neither for *opaque-2* nor *Ask-2* in the CLQRCWQ26 X La Posta Seq C7-F64-2-6-2-1-B-B)-B population, this population

Table 6. Total content for free amino acid (nmol g^{-1} FW) of nine genotypic classes of the F_2 of the cross CML172 x CML323.

Class[††]	Genotypes		Asp	Glu	Asn	Thr	His	Met	Trp	Phe	Ile	Leu	Lys
	o2	*Ask-2*											
P1	1[¶]	1	4969.7	2759.5	4631.1	286.3	61.1	28.0	53.1	210.1	159.8	268.8	719.1
P2	2[¶¶]	2	268.0	329.0	656.8	64.0	41.1	39.1	45.9	36.2	41.8	55.0	81.8
G1,G2,G3	1	-	2819.4 a	1202.3 a	1657.0 a	317.8 a	170.7 a	69.6	94.8 a	141.7 a	187.5 a	241.5 a	355.5 a
G4,G5,G6	2	-	470.8 b	439.1 b	438.2 b	193.9 b	90.3 b	35.3	44.7 b	58.1 b	79.3 b	159.4 b	117.0 b
G7,G8,G9	H[x]	-	361.9 b	297.1 b	331.8 b	160.0 b	85.6 b	34.2	38.3 b	46.8 b	58.8 b	115.5 c	108.6 b
G1,G4,G7	-	1	1363.8	614.9	709.7	209.1	124.2	54.6	60.0	86.6	109.1	166.4	209.6
G2,G5,G8	-	2	1369.5	731.7	1095.9	247.3	117.9	41.1	65.8	91.6	119.8	185.0	176.3
G3,G6,G9	-	H	918.7	592.0	621.4	215.4	104.4	43.4	51.9	68.5	96.7	165.0	195.3
Mean[xx]			1217.3**	646**	809**	223**	115**	46 ns	59**	82**	108**	172**	193**

[†] FAA total content was determined in complete seeds. [††] Class is the genotype. P1 (QPM parent) and P2 (normal parent) were not considered for the statistical analysis. [¶]1. Homozygous recessive. [¶¶] 2. Homozygous dominant. [x]H: Heterozygous. [xx]Multiple comparisons were tested with Tukey, 0.05. **Indicates significance at p<0.01. ns not significant.

was considered into analysis due to data recorded for amino acids such as Thr, Ile and Lys. As it was evidenced, the effect of the *opaque-2* genotypes over the accumulation of total FAA revealed that the two principal components axes explained the 96.98% of the total variation (Figure 7). Four groups were distinguished. The recessive genotype had a sharp effect in the accumulation of all amino acids. Particularly, for the CML172 X CML323 population the homozygous recessive genotype affected the accumulation of Thr, Glu, His, Leu, Met, Phe, Ile and Asn residues. The Trp accumulation was also affected by the recessive genotype as well as the Lys content. In a lesser extent, the accumulation of Asp was also affected by the recessive genotype, but for this amino acid, the heterozygous genotype of both populations turned out to affect more its accumulation. Dominant genotypes of *opaque-2* appeared no affect the accumulation of amino acids on both populations.

By the other hand, the effect of the *Ask-2* genotypes over the amino acid content was also graphed. Here, the two principal components explained the 86.51 % of total variation (Figure 8). For CML172 X CML323 population, the recessive genotype had an effect in the accumulation of Trp and Met, but also with the Phe, Leu and Ile content whereas the homozygous dominant genotype had also a putative association with the accumulation of Asn, Thr, Lys, Glu, and His. The heterozygous genotype apparently had not relationship with the accumulation of any amino acid. In contrast, the homozygous recessive genotype in the CLQRCWQ26 X La Posta Seq C7-F64-2-6-2-1-B-B)-B population appeared to be associated with Asn, and a lesser extent with Thr, Lys, Glu and His, the dominant genotype did not affect the amino

acid content and the heterozygous genotype appeared to affect the accumulation of Asp as occurred with its *opaque-2* counterpart.

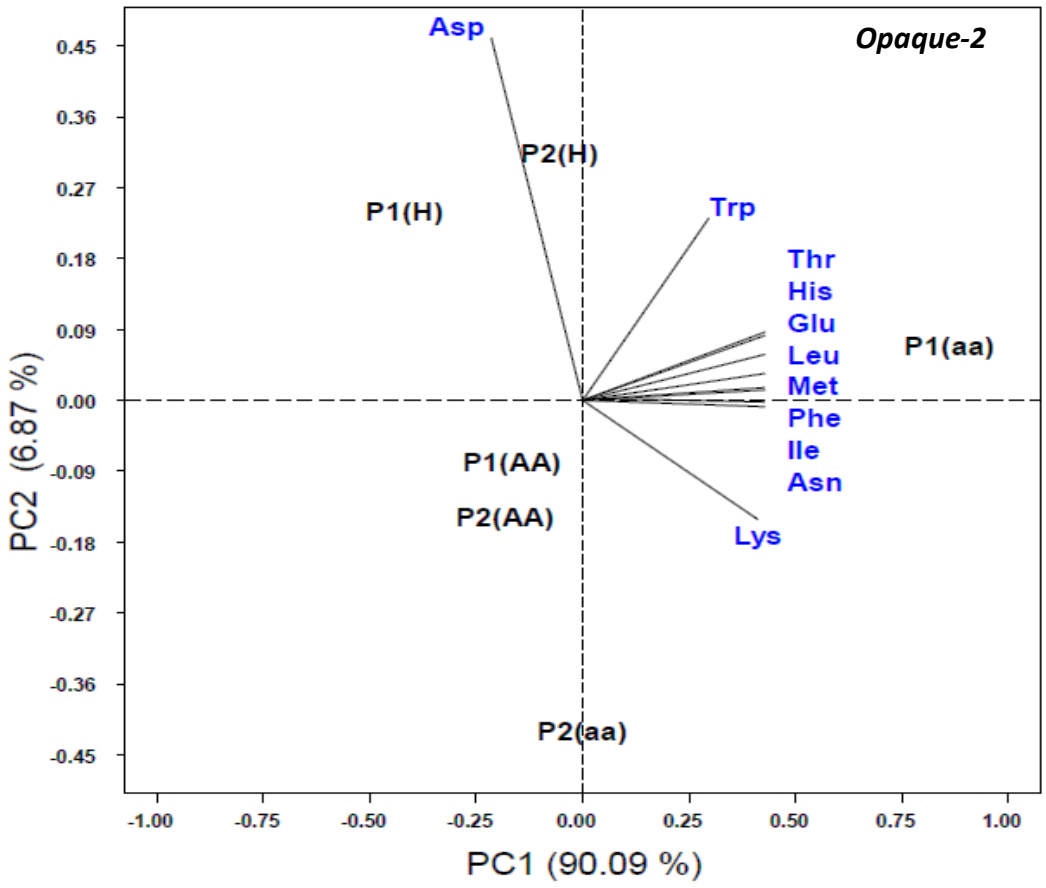


Figure 7. Biplot of first two principal components showing the proportion of variation derived from the amino acid content in the three genotypes of *opaque-2* gene. P1 is the CML172 X CML323 population. P2 is the CLQRCWQ26 X La Posta Seq C7-F64-2-6-2-1-B-B)-B) population. AA is the dominant genotype. Aa is the heterozygous and aa is the homozygous recessive genotype.

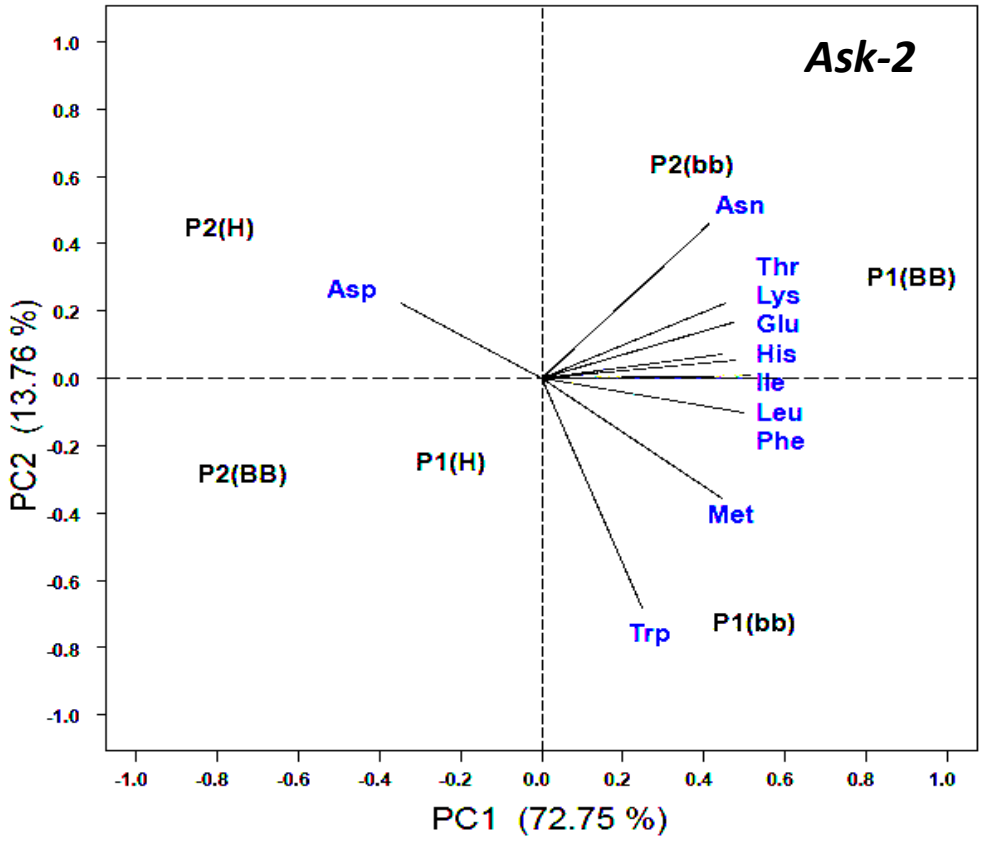
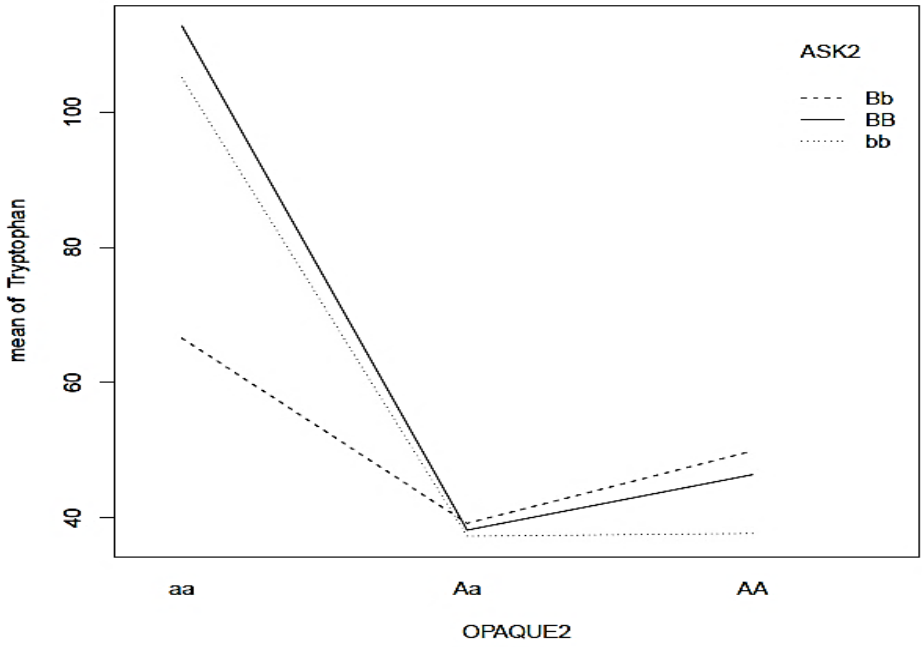


Figure 8. Biplot of first two principal components showing the proportion of variation derived from the amino acid content in the three genotypes of *Ask-2* gene. P1 is the CML172 X CML323 population. P2 is the CLQRCWQ26 X La Posta Seq C7-F64-2-6-2-1-B-B)-B) population. BB is the dominant genotype. Bb is the heterozygous and bb is the homozygous recessive genotype.

Once evidenced that separately a particular genotype of *opaque-2* or *Ask-2* had a relationship with the FAA accumulation, the next step was to verify the relationship between the two genes together. Figure 9 shows the interaction between *opaque-2*

and *Ask-2* loci for the accumulation of Trp and Lys in the CML172 X CML323 population. It was evident the effect of the homozygous recessive genotype *opaque-2* over the accumulation of both amino acids when it was compared with its other genotypes. The accumulation dropped drastically with the heterozygous genotype and then a slight increase with the homozygous dominant was observed but not to the accumulation level detected with the homozygous recessive genotype. This same trend was also stayed fairly constant for Asp, Glut, Ile and Leu residues (Figure 10) but it was particularly striking for Met and His. For other amino acids as Asn, Thr and Phe, although there was a decrease in the amino acid accumulation when there was a change in the genotype, the interaction was not so evident and apparently they were independent. These results indicated that there was an effect of *o2*/*o2* in the accumulation of FAA. This effect was, apparently, dependent of the genetic background where *opaque-2* has its expression. Particularly, in the CML172 X CML323 population the amount of FAA was nearly always duplicated for all amino acids when *opaque-2* was as double recessive. Conversely, this effect was not so evident in the population CLQRCWQ26 X La Posta Seq C7-F64-2-6-2-1-B-B)-B). A double increase in the content of only Lys and Trp residues for QPM genotypes has been commonly referred to occur (Vasal, 2001; Babu *et al*., 2005; Vivek *et al*, 2008). In fact, most of research carried out to measure the amino acid content in *opaque-2* endosperm usually determines Lys and Trp content. Underlying causes for this unique determination are due to the main storage proteins in the maize endosperm, zeins, are devoid in Lys and Trp residues and their estimation is an indirect way to know the effect of o*paque-2* gene during the QPM conversion.



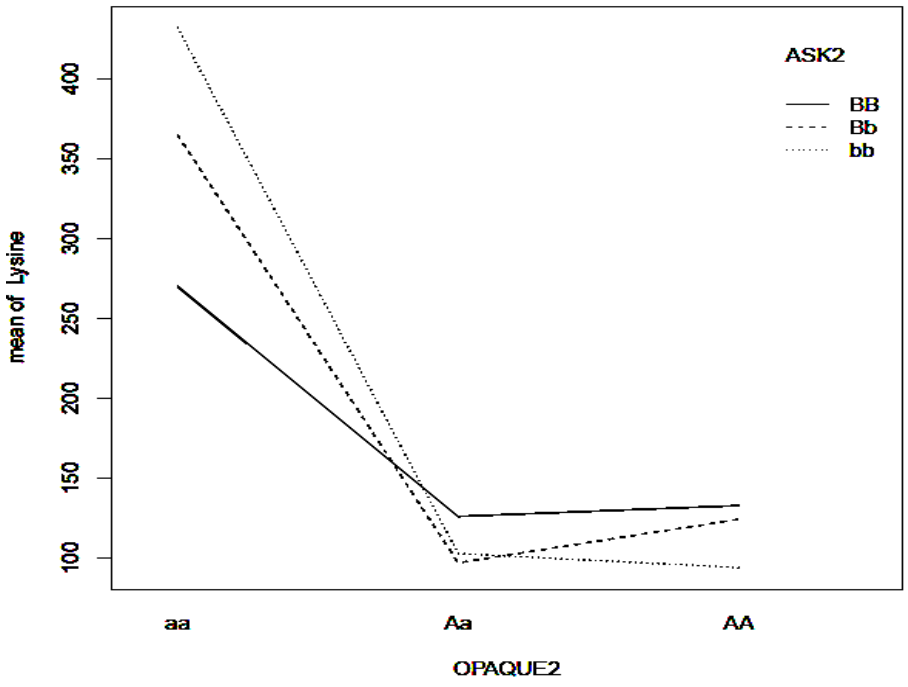


Figure 9. Interaction between *opaque-2* and *Ask-2* loci for the accumulation of Trp (left) and Lys (right) residues in the maize F_2 population CML172 x CML323.

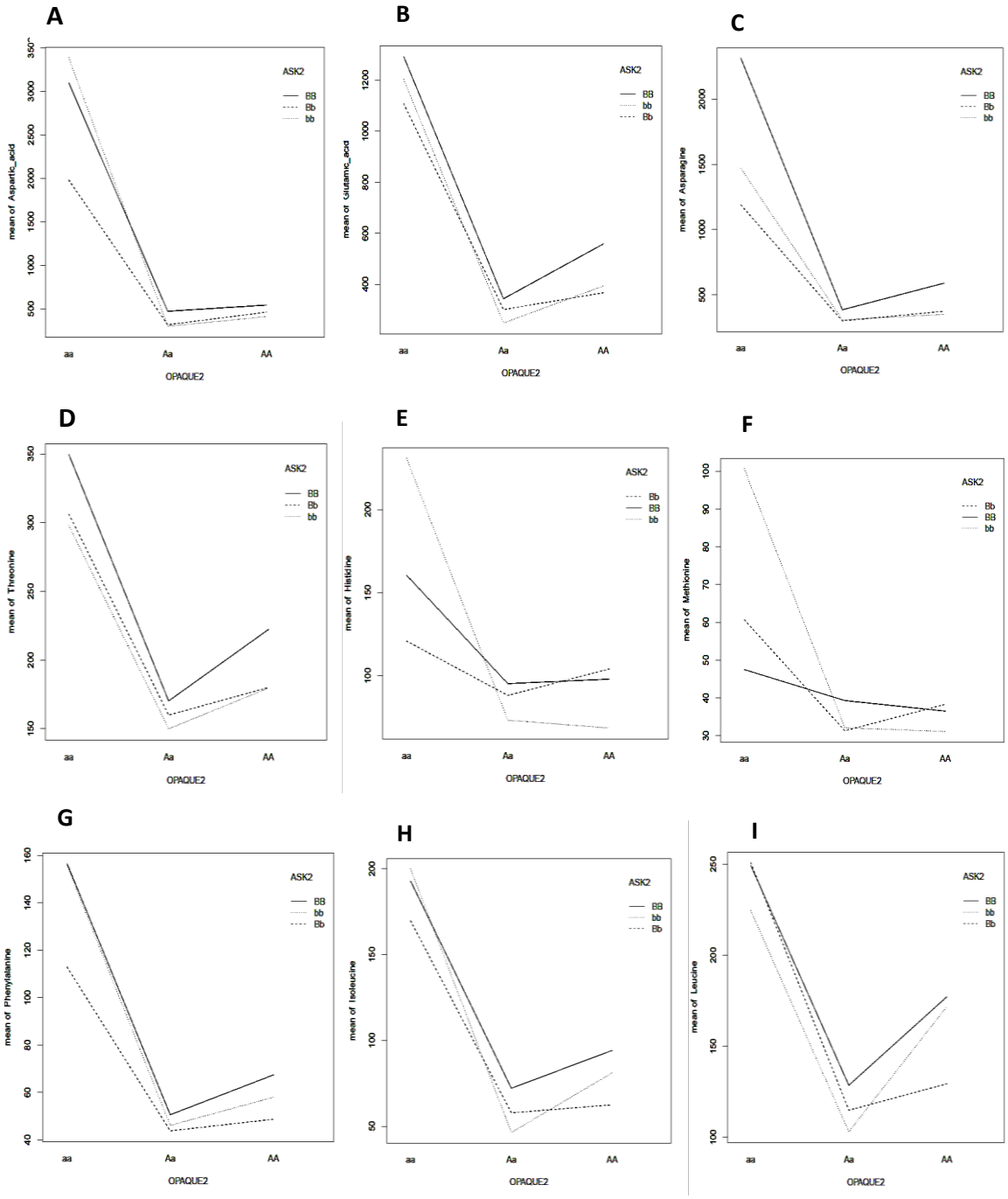


Figure 10. Interaction of *opaque-2* and *Ask-2* loci for free amino acid content in seeds of maize F_2 population CML172 x CML323. A. Aspartic acid. B. Glutamic acid. C. Asparagine. D. Threonine. E. Histidine. F. Methionine. G. Phenylalanine. H. Isoleucine. I. Leucine.

Opaque-2 mutant (*o2/o2*) exerts its effect specifically in the modulation of 22-kDa α–zein and 15-kDa β–zein genes reducing their expression level and increasing the content of minor storage proteins (Gibbon and Larkins, 2005; Wu *et al*., 2010; Motto *et al*., 2011) and non-zeins proteins as the elongation factor 1-α (*ef1-α*), which is highly correlated with the high lysine content in the opaque endosperm (Habben *et al*., 1995). It has been proposed that manipulation of the reduction of zein transcription in the endosperm, without altering the contribution of other richer Lys and Trp storage proteins, could be a feasible approach to improve the amino acid balance of maize grain (Prassana *et al*., 2001).

A proper amino acid balance of maize grain is essential for a proper human physiological activity. Maize is deficient in Lys and Trp and its single consumption may cause protein malnutrition in those maize-dependent communities. This is the reason of the efforts in the QPM breeding: increase the level of Lys and Trp for duplicating the biological value of maize grain. These efforts not only had been focused for breeding maize for human consumption. Since Lys is the principal amino acid limiting in cereals as maize (Ufaz and Galili, 2008) and maize is used also for feed for livestock, Lys has been more studied in animal nutrition than any other amino acid. It has been estimated that 90% of the total Lys production is used to supplement animal diets (Baker, 2007). In fact, breeding to develop high-lysine maize cultivars is the focus in some breeding programs (Azevedo and Arruda, 2010). Biochemical studies for enzymes involved in the synthesis of Lys have been also intensive. It has been reported that the increase free Lys in the opaque endosperm is the loss of Lys

ketoglutarate reductase (LKR) activity, an enzyme that degrades Lys as the endosperm matures (Arruda *et al*., 2000).

Emphasis in the maize breeding for Lys and Trp content for human nutrition began 49 years ago with the demonstration by Mertz *et al.* (1964) that opaque-2 endosperm had 69% more Lys content than normal seeds and the major reason to this change was the synthesis of proteins with a greater content of basic amino acids. Results of this study showed a similar Lys accumulation which it was of 67.1% and 69.5% higher in the genotypic classes with recessive *opaque-2* gene. Although lower, Trp content also recorded a parallel content: 52.9% and 59.7%. The effect of *opaque-2* over the relationship between the amino acid content and the recessive *opaque-2* was evidenced in the biplot analysis. Two principal axes explained 96.98% of the total variation (Figure 7) meaning that *opaque-2* locus had affected in a high degree the zein genes. More additional evidence for the effect of *opaque-2* in the FAA was obtained from the interaction between *opaque-2* and *Ask-2* genotypes. As was shown per each amino acid, a sharp change occurred in the accumulation of a particular amino acid when a change occurred in the *opaque-2* genotype from recessive genotype to heterozygous or dominant genotype (Figures 9 and 10). These results indicated that presumably *opaque-2* was involved in the higher accumulation of all amino acids evaluated in the CML172 X CML323 population. Unlike previous results with the CML172 X CML323 population, in the CLQRCWQ26 X La Posta Seq C7-F64-2-6-2-1-B-B)-B) population, the presence of *opaque-2* did not ensure higher Lys and Trp contents, indicating a possible role of the type of genetic background of this population for that opaque-2 performed its action.

V. CONCLUSIONS

Sequence analysis of two fragments of *opaque-2* gene revealed a remarkable level of genetic variation in the set of 16 examined genotypes. In total, sixty-five polymorphism were detected in 31.34 kb, one polymorphism each 30.13 nucleotides. One of these single nucleotide polymorphisms was genotyped using the KASP™ technology and it was found that this polymorphism (C→T) was not perfectly associated with the *opaque-2* gene, such that it was 88.3% assertive to discriminate QPM lines (C) from normal lines (T) in a panel of 88 maize lines. However, this polymorphism could be used for MAS only in the cases when the QPM progenitor is harboring the allele C and the non QPM progenitor is harboring the allele T.

This study also showed that SNP genotyping using next generation technologies, especially KASP™ chemistry, is highly accurate and facilitate the screening of the polymorphisms in one step if the sequence harboring the polymorphism is known. With the SSCP technique developed here to analyze genetic variants in the 3´-end of the *Ask-2* locus, was possible to find PCR-SSCP variants associated with a higher Trp and Lys content.

Finally, it was also evidenced that there was interaction between *opaque-2* and *Ask-2* genotypes for free amino acid accumulation in the endosperm. *Opaque-2* gene in a recessive form (*o2/o2*) remarkably increased the content of free amino acid, to drop drastically with the change to the heterozygous or dominant homozygous genotypes.

VI. REFERENCES

Arruda, P. L., E. L. Kemper, F. Papes F, and A. Leite. 2000. Regulation of lysine catabolism in higher plants. Trends in Plant Science 5: 324–330.

Atlin, G. N., N. Palacios, R. Babu, B. Das, S. Twumasi-Afriyie, D. K. Friesen, H. De Groote, B. Vivek, and K Pixley. 2012. Quality Protein Maize: progress and prospects. *In*: Plant Breeding Reviews 34. J. Janick (ed). Wiley Blackwell, USA. pp: 83-130.

Azevedo, R. A., and P. Arruda. 2010. High-lysine maize: the key discoveries that have made it possible. Amino Acids 39(4): 979-989**.**

Babu, R., S. K. Nair, A. Kumar, S. Venkatesh, J. C. Sekhar, N. N. Singh, G. Srinivasan, and H. S. Gupta. 2005. Two-generation marker-aided backcrossing for rapid conversion of normal maize lines to quality protein maize (QPM). Theoretical and Applied Genetics 111: 888-897.

Baker, D. H. 2007. Lysine, arginine, and related amino Acids: an introduction to the 6th amino acid assessment workshop. The Journal of Nutrition 1569S-1601S.

Balconi, C., H. Hartings, M. Lauria, R. Pirona, V. Rossi, M. Motto. 2007.Gene discovery to improve maize grain quality traits. Maydica 52: 357-373.

Berger, N.A. 1985. Poly(ADP-ribose) in the cellular response to DNA damage. Radiation Research 101: 4-15.

Bouis, H. E., and R. Welch. 2010. Biofortification -a sustainable agricultural strategy for reducing micronutrient malnutrition in the global south. Crop Science 50: 20–32.

Bressani, R. 1992. Nutritional value of high-lysine maize in humans. *In*: E.T. Mertz (ed.). Quality Protein Maize. American Association of Cereal Chemists, St. Paul, MN. USA.

Cairns, J. E., K. Sonder, P. H. Zaidi, N. Verhulst, , S. K. Nair, B. Das, B. Govaerts, M. T. Vinayan, Z. Rashid, J. J. Noor, P. Devi, F. San Vicente, and B. M. Prasanna. 2012. Maize production in a changing climate: impacts, adaptation, and mitigation strategies. Advances in Agronomy Journal 114: 1-58.

Castillo-Nonato, J., and C. Chávez-Mejía. 2013. Peasant characterization of management and use of maize diversity in San Felipe del Progreso, Estado de México. Agricultura, Sociedad y Desarrollo 10: 23-38.

Chin, E. C. L., M. L. Senior, H. Shu, and J. S. C. Smith. 1996. Maize simple repetitive DNA sequences: abundance and allele variation. Genome 39: 866-873.

Ching A., K. S. Caldwell, M. Jung, M. Dolan, O. S. (H.) Smith, S. Tingey, M. Morgante, and A. J. Rafalski. 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. BMC Genetics 3: 19.

CIMMYT. 2005. Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory. Third Edition. Mexico, D.F.: CIMMYT.

Cockerham, C. C. 1954. An extension of partitioning hereditary variance for analysis of covariance among relatives when epistasis is present. Genetics 39: 859-882.

Esen, A. 1987. A proposed nomenclature for the alcohol-soluble proteins (zeins) of maize (*Zea mays* L.). Journal of Cereal Science 5: 117–128.

FAO. 2001. Human Vitamin and Mineral Requirements. Food and Agriculture Organization (FAO), Rome, Italy. pp: 33-35

FAOSTAT. 2013. FAO. Statistical Database. Food and Agricultural Organization of United Nations. Available online: http://faostat.fao.org (verified on September 18, 2013).

Gao, S., C. Martínez, D. J. Skinner, A. F. Krivanek, J. H. Crouch, and Y Xu. 2008. Development of a seed DNA-based genotyping system for marker-assisted selection in maize. Molecular Breeding 22 (3): 477–494.

Gaut, B. S., M. Le Thierry d'Ennequin, A. S. Peek, and M. C. Sawkins. 2000. Maize as a model for the evolution of plant nuclear genomes. Proceedings of the National Academy of Sciences of United States of America 97: 7008-7015.

Gaziola, S. A., E. S. Alessi, P. E. O. Guimaraes, C. Damerval, and R. A. Azevedo. 1999. Quality protein maize: a biochemical study of enzymes involved in lysine metabolism. Journal of Agricultural and Food Chemistry 47: 1268-1275.

Gerdes, S., C. Lerma-Ortiz, O. Frelin, S. M. D. Seaver, C. S. Henry, V. de Crécy-Lagard, and A. D. Hanson. 2012. Plant B vitamin pathways and their compartmentation: a guide for the perplexed. Journal of Experimental Botany 63(15): 5379–5395.

Gibbon, B. C. and B. A. Larkins. 2005. Molecular approaches to developing quality protein maize. Trends in Genetics 21(4): 227-234.

Goodman, M. M. 1999. Broadening the genetic diversity in maize breeding by use of exotic germplasm. *In*: The Genetics and Exploitation of Heterosis in Crops. Coors J. C., and Pandey S. (eds). American Society of Agronomy, Inc. Crop Science Society of America. Madison, Wisconsin, USA. pp: 139–148.

Gupta P. K., J. K. Roy, and M. Prasad. 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.

Gupta, H.S., R. Babu, P. Agrawal, V. Mahajan, F. Hossain, and N. Thirunavukkarasu. 2012. Accelerated development of quality protein maize hybrid through marker-assisted introgression of *opaque-2* allele. Plant Breeding doi:10.1111/pbr.12009.

Habben, J. E., G. L. Moro, B. G. Hunter, B. R. Hamaker, and B. A. Larkins. 1995. Elongation factor 1 alpha concentration is highly correlated with the lysine content of maize endosperm. Proceedings of the National Academy of Sciences of the United States of America 92(19): 8640-8644.

Hamblin, M. T., Warburton M. L., and E. S. Buckler. 207. Empirical Comparison of Simple Sequence Repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. PlosOne 12:13367.

Hartings, H., M. Lauria, N. Lazzaroni, R. Pirona, and M. Motto. 2011. The *Zea mays* mutants *opaque-2* and *opaque-7* disclose extensive changes in endosperm metabolism as revealed by protein, amino acid, and transcriptome-wide analyses. BMC Genomics 12:41.

Hartings, H., M. Maddaloni, N. Lazzaroni, N. Di Fonzo, M. Motto, F. Salamini, and T. Thompson. 1989. The O2 gene which regulates zein deposition in maize

endosperm encodes a protein with structural homologies to transcriptional activators. EMBO Journal 8: 2795-2801.

Hayashi, K. 1991. PCR-SSCP: a simple and sensitive method for detections of mutations in the genomic DNA. PCR Methods Applications 1: 34-38.

Hayashi, K., and D. W. Yandell. 1993. How sensitive is PCR-SSCP? Human Mutation 2(5): 338-46.

Hodgkinson, V. H., J. Birungi, M. Haghpanah, S. Joshi, and L. E. Munstermann. 2002. Rapid identification of mitochondrial cytochrome B haplotypes by single strand conformation polymorphism in *Lutzomyia longipalpis* (Diptera: Psychodidae) Populations. Journal of Medical Entomology 39(4): 689-694.

Holding, D. R. and B. A. Larkins. 2009. Zein storage proteins. *In*: Molecular Genetics Approaches to Maize Improvement. Biotechnology in Agriculture and Forestry. A L. Kriz, B. A. Larkins (eds). Springer-Verlag. Heidelberg, Germany. pp: 269-286.

Hongyo, T., G. S. Buzardl, R. J.Calvert, and C. M. Weghorst. 1993. 'Cold SSCP': a simple, rapid and non-radioactive method for optimized single-strand conformation polymorphism analyses. Nucleic Acids Research 21(16):3637-3642.

Hossain, F., B. M. Prasanna, R. Kumar, and B. B. Singh. 2008. Genetic analysis of kernel modification in Quality Protein Maize (QPM) genotypes. Indian Journal of Genetics 68(1): 1-9.

Jones, E. S., H. Sullivan, D. Bhattramakki, and J. S. C. Smith. 2007. A comparison of simple sequence repeat and single nucleotide polymorphism marker

technologies for the genotypic analysis of maize (*Zea mays* L.). Theoretical and Applied Genetics 115: 361–371.

Kato, T. A., C. Mapes, L. M. Mera, J. A. Serratos, R. A. Bye. 2009. Origen y diversificación del maíz: una revisión analítica. Universidad Nacional Autónoma de México, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. 116 p. México, D.F.

Kourkine, I. V., C. N. Hestekin, and A. E. Barron. 2002. Technical challenges in applying capillary electrophoresis-single strand conformation polymorphism for routine genetic analysis. Electrophoresis 23: 1375–1385.

Krivanek, A. F., H. D. Groote, N. S. Gunaratna, A O Diallo, and D Friesen. 2007. Breeding and disseminating quality protein maize (QPM) for Africa. African Journal of Biotechnology 6(4): 312-324.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227 (5259): 680-685.

Lamkey K. R. and J. W Edwards. 1999. Quantitative genetics of heterosis. *In*: The genetics and exploitation of heterosis in crops. Coors J. G. and S. Pandey (eds). American Society of Agronomy, Inc. Crop Science Society of America. Madison, Wisconsin, USA. pp: 31-48.

Lia V. V., M. Bracco, A. M. Gottlieb , L. Poggio, and·V. A. Confalonieri. 2007. Complex mutational patterns and size homoplasy at maize microsatellite loci. Theoretical and Applied Genetics 115: 981–991.

Lohmer, S., M. Maddaloni, M. Motto, N. DiFonzo, H. Hartings, F. Salamini, and R. D. Thompson, 1991. The maize regulatory locus opaque-2 encodes a DNA binding

protein which activates the transcription of the B-32 gene. EMBO Journal 10: 617–624.

Maddaloni, M., N. Di Fonzo, H. Hartings, N. Lazzaroni, F. Salamini, R.D. Thompson and M. Motto. 1989. The sequence of the zein regulatory gene *opaque-2* (O2) of *Zea mays*. Nucleic Acids Research 17: 7532-7535.

MaizeGDB. 2010. Maize Genetics and Genomics Database. Available online: http://www.maizegdb.org. (Verified on 2010).

Manicacci, D., L. Camus-Kulandaivelu, M. Fourmann, Ch. Arar, S. Barrault, A. Rousselet, N. l. Feminias, L. Consoli, L. Francès, V. Méchin, A. Murigneux, J.-L. Prioul, A. Charcosset and C. Damerval. 2009. Epistatic interactions between *Opaque2* transcriptional activator and its target gene *CyPPDK1* control kernel trait variation in maize. Plant Physiology 150: 506–520.

Matsuoka, Y., Y. Vigouroux, M. M. Goodman, J. G. Sanchez, E. Buckler, and J. Doebley. 2002. A single domestication for maize shown by multilocus microsatellite genotyping. Proceedings of the National Academy of Sciences of the United States of America 99: 6080–6084.

Mertz, E. T., L. S. Bates, and O. E. Nelson. 1964. Mutant gene that changes protein composition and increases lysine content of maize endosperm. Science 145(3629): 279-280.

Motto, M., H. Hartings, M. Francassetti, and F. Consonni. 2011. Grain quality-related traits in maize: gene identification and exploitation. Maydica 56(1770): 291-314

Muehlbauer, G. J., D. A. Somers, B. F. Matthews, and B. C. Gengenbach. 1994. Molecular genetics of the maize (*Zea mays* L.) aspartate kinase-homoserine dehydrogenase gene family. Plant Physiology 106: 1303-1312.

NCBI. 2009. National Center for Biotechnology Information, U.S. National Library of Medicine. Available on: http://www.ncbi.nlm.nih.gov/. (Verified on 2009).

Nuss, E. T., and S. A. Tanumihardjo. 2011. Quality Protein Maize for Africa: closing the protein inadequacy gap in vulnerable populations. American Society of Nutrition. Advances in Nutrition 2: 217–224.

Orita M, H. Iwahana, H. Kanazawa, K. Hayashi, and T. Sekiya.1989. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proceedings of the National Academy of Sciences of the USA 86: 2766–2770.

Pérez-Massot, E., R. Banakar, S. Gómez-Galera, U. Zorrilla-López, G. Sanahuja, G. Arjó, B. Miralpeix, E. Vamvaka, G. Farré, S. M. Rivera, S. Dashevskaya, J. Berman, M. Sabalza, D. Yuan, Ch. Bai, L. Bassie, R. M. Twyman, T. Capell, P. Christou, and Changfu Zhu. 2013. The contribution of transgenic plants to better health through improved nutrition: opportunities and constraints. Genes Nutrition 8: 29–41.

Prasanna, B. M. S. K. Vasal, B. Kassahun, and N. N. Singh. 2001. Quality protein maize. Current Science 81(10): 1208-1319.

Reis-Filho, J. S. 2009. Next-generation sequencing. Breast Cancer Research doi:10.1186/bcr2431

Schmidt, R. J., F. A. Burr, M. J. Aukerman, and B. Burr. 1990. Maize regulatory gene *opaque-2* encodes a protein with a “leucine zipper” motif that binds to zein DNA. Proceedings of the National Academy of Sciences of United States of America 87:46 50.

Schnable, P., D. Ware, R. S. Fulton, J. C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton, T. A. Graves, P. Minx, A. D. Reily, L. Courtney, S. S. Kruchowski, C. Tomlinson, C. Strong, K. Delehaunty, C. Fronick, B. Courtney, S. M. Rock, E. Belter, F. Du, K. Kim, R. M. Abbott, M. Cotton, A. Levy, P. Marchetto, K. Ochoa, S. M. Jackson, B. Gillam, W. Chen, L. Yan, J. Higginbotham, M. Cardenas, J. Waligorski, E. Applebaum, L. Phelps, J. Falcone, K. Kanchi, T. Thane, A. Scimone, N. Thane, J. Henke, T. Wang, J. Ruppert, N. Shah, K. Rotter, J.R. Hodges, E. Ingenthron, M. Cordes, S. Kohlberg, J. Sgro, Brandon Delgado, K. Mead, A. Chinwalla, S. Leonard, K. Crouse, K. Collura, D. Kudrna, J. Currie, R. He, A. Angelova, S. Rajasekar, T. Mueller, R. Lomeli, G. Scara, A. Ko, K. Delaney, M. Wissotski, G. Lopez, D.Campos, M. Braidotti, E. Ashley, W. Golser, H. Kim, S. Lee, J. Lin, Z. Dujmic, W. Kim, J. Talag, A. Zuccolo, C. Fan, A. Sebastian, M. Kramer, L. Spiegel, L. Nascimento, T. Zutavern, B. Miller, C. Ambroise, S. Muller, W.Spooner, A. Narechania, L. Ren, S. Wei, S. Kumari, B. Faga, M. J. Levy, L.McMahan, P. van Buren, M. W. Vaughn, K. Ying, C.-T. Yeh, S. J. Emrich, Y.Jia, A. Kalyanaraman, A.-P. Hsia, W. B. Barbazuk, R. S. Baucom, T. P. Brutnell, N. C. Carpita, C. Chaparro, J.-M. Chia, J.-M. Deragon, J. C. Estill, Y. Fu, J. A. Jeddeloh, Y. Han, H. Lee, P. Li, D. R. Lisch, S. Liu, Z. Liu, D. H. Nagel, M. C. McCann, P. SanMiguel, A. M. Myers, D. Nettleton, J. Nguyen, B. W. Penning, L. Ponnala, K. L. Schneider, D. C. Schwartz, A. Sharma, C. l. Soderlund, N. M. Springer, Q. Sun, H. Wang, M. Waterman, R Westerman, T. K. Wolfgruber, L. Yang, Y. Yu, L. Zhang, S. Zhou, Q. Zhu, J. L. Bennetzen, R. K. Dawe, Ji. Jiang, N. Jiang, G. G. Presting, S. R. Wessler, S. Aluru, R. A. Martienssen, S. W. Clifton, W. R.

McCombie, R. A. Wing, and R. K. Wilson. The B73 Maize Genome: Complexity, Diversity, and Dynamics. Science 326 (5956): 1112.

Semang, K, R. Babu, S. Hearne, and M. Olsen. 2013. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Molecular Breeding doi 10.1007/s11032-013-9917-x.

Serratos, J. A. H. 2012. El origen y la diversidad del maíz en el continente americano. Greenpeace-México. Segunda Edición. 40 p.

Shai Ufaz, S. and G. Galili. 2008. Improving the content of essential amino acids in crop plants: goals and opportunities. Plant Physiology 147: 954–961.

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

Singleton, R. V. 1939. Recent linkage studies in maize. V. Opaque endosperm-2 (*o2*). Genetics, 24: 61–62.

Strable, J. and M. J. Scanlon. 2009. Maize (*Zea mays*): a model organism for basic and applied research in plant biology. Cold Spring Harbor Protocols doi:10.1101/pdb.emo132

Sun, Y., M. Thompson, G. Lin, H. Butler, Z. Gao, S. Thornburgh, K. Yau, D. A. Smith, and V. K. Shukla. 2009. Inositol 1,3,4,5,6-Pentakisphosphate 2-Kinase from Maize: Molecular and Biochemical Characterization. New Phytologist 182: 49–84.

Sunnucks, P., A. C. C. Wilson, L. B. Beheregaray, K. Zenger, J. French, and, A. C. Taylor. 2000. SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. Molecular ecology 9: 1699-1710.

Syvänen, A. C. 2001. Accessing genetic variation: genotyping of single nucleotide polymorphisms. Nature Review of Genetics 2: 930–942.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgin. 1997. The CLUSTAL X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25: 4876–4882.

Vasal, S.K. 2001. High quality protein corn. *In*: Specialty Corns. A.R. Hallauer (ed). CRC Press, Boca Raton, Florida, USA. pp: 85-129.

Vicente-Carbajosa, J., S. P. Moose, R. L. Parsons, and R. Schmidt. 1997. A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. Proceedings of the National Academy of Sciences of the United States of America 94: 7685-7690.

Vielle-Calzada, J-P., O. Martinez de la Vega, G. Hernández-Guzmán, E. Ibarra-Laclette, C. Alvarez-Mejía, J. C. Vega-Arreguín, B. Jiménez-Moraila, A. Fernández-Cortés, G. Corona-Armenta, L. Herrera-Estrella, and A. Herrera-Estrella. 2009. The Palomero genome suggests metal effects on domestication. Science 326 (5956): 1078.

Vivek, B. S., A.F. Krivanek, N. Palacios-Rojas, S. Twumasi-Afriyie, and A.O. Diallo. 2008. Breeding Quality Protein Maize (QPM): Protocols for Developing QPM Cultivars. Mexico, D.F.: CIMMYT.

Wang, X., D. K. Stumpf, and B. A. Larkins. 2001. Aspartate kinase 2. A candidate gene of a quantitative trait locus influencing free amino acid content in maize endosperm. Plant Physiology 125: 1778–1787.

Wang, X., J. A. López-Valenzuela, B. C. Gibbon, B. Gakiere, G. Galili, and B. A. Larkins. 2007. Characterization of monofunctional aspartate kinase genes in maize and their relationship with free amino acid content in the endosperm. Journal of Experimental Botany 58 (10): 2653–2660.

Wu, Y, D. R. Holding, and J. Messing. 2010. γ-Zeins are essential for endosperm modification in quality protein maize. Proceedings of the National Academy of Sciences of the United States of America 107(29): 12810-12815.

Ying, W. 2013. Roles of NAD+, PARP-1, and sirtuins in cell death, ischemic brain injury, and synchrotron radiation X-ray-induced tissue injury. Scientifica http://dx.doi.org/10.1155/2013/691251

Xu, Y., D. J. Skinner, H. Wu, N. Palacios-Rojas, J. L. Araus, J. Yan, S. Gao, M. L. Warburton, and J. H. Crouch. 2009. Advances in maize genomics and their value for enhancing gains from breeding. International Journal of Plant Genomics 2009:957602. doi:10.1155/2009/957602.

VII. APPENDIX

A.1. TWO-PHASE POLYACRYLAMIDE GELS FOR PCR-SSCP

Based on Laemmli´s system (1970)

Formulations to prepare gels using a MGV-216-33 Dual Triple Wide Mini-Vertical Electrophoresis System (C.B.S. Scientific, USA)

STOCK SOLUTIONS

Resolving Acrylamide:

	50 gels	1 gel
Acrylamide-bisacrylamide 40% 29:1 (Sigma A2792)	334 ml	6.7 ml
ddH_20	276 ml	5.5 ml
Tris 1.5 M, pH 8.8.	210 ml	4.2 ml

Stacking Acrylamide:

	50 gels	1 gel
Acrylamide-bisacrylamide 40% 29:1	23 ml	0.45 ml
ddH_20	109 ml	2.19 ml
Tris 0.5 M, pH 6.8	48 ml	0.96 ml

SOLUTIONS NEEDED TO POUR GELS (MIX IMMEDIATELY BEFORE USE)

Resolving (lower gel phase) for 1 gel:

Running acrylamide	31 ml
APS 10% (Sigma A3678)	120 μl
TEMED (1,2-Bis(dimethylamino-ethane)	25 μl

Stacking (upper gel phase) for 1 gel:

Stacking acrylamide	13 ml
APS 10%	125 μl
TEMED (1,2-Bis(dimethylamino)ethane)	16 μl

Running buffer

TG 10x	**2 L**	**8 L**	**10 L**
Trizma Base (Sigma T1503)	60 gr	240 gr	300 gr
Glicine (Sigma G8898)	288 gr	1152 gr	1440 gr